# Lab 10 – MATH 240 – Computational Statistics

Anya Suko

April 10, 2024

## 1    Introduction

In this lab, we examine Gallup's study on polling. The study explains that a poll of 1,000 people can have a margin of error of 4 percent, while a poll of 2,000 people can reduce that margin to 2. Our objective in this lab is to run simulations to better understand how sample size impacts the margin of error and compare our results to Gallup's findings. We also look at re sampling and other methods to see how different factors, like sample size and population proportion, influence the margin of error, helping us understand how Gallup improves its polling accuracy.

## 2    Basic Simulation

First, we conducted a basic simulation of 10,000 polls using a sample size of 1,004. For the population, we assume 0.39 is the proportion of satisfied respondents.We use a binomial distribution to model the responses.
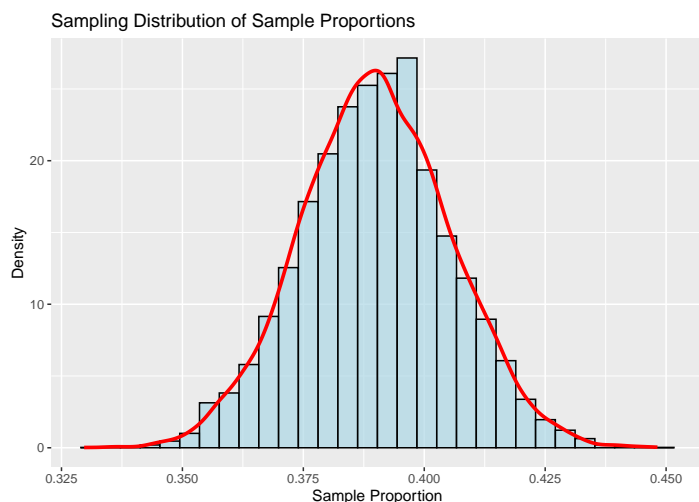


Figure 1: Basic Simulation with size 1,004

The distribution is bell shaped and symmetric. The middle 95 percent of the data ranges from approximately [0.3613, 0.4181], which has a margin of error of 2.84 percent, which is a little bit smaller than the margin of error Gallup describes in the report (4 percent).

When the sample size is doubled to become 2,008, the margin of error becomes smaller, becoming 1.85 percent. This is close to what Gallup describes ( two percent).
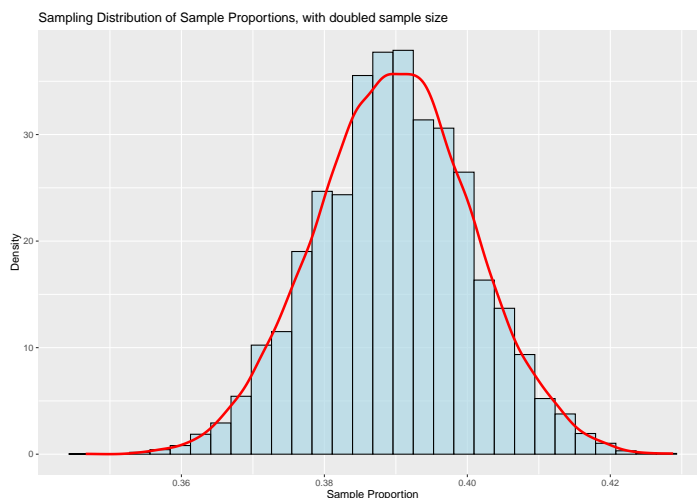


Figure 2: Basic Simulation with size 2,008

## 3    Resampling

We generated 1,000 re-samples from the observed Gallup responses (with 39 percent satisfaction and 59 percent unsatisfaction, and one percent not answering).
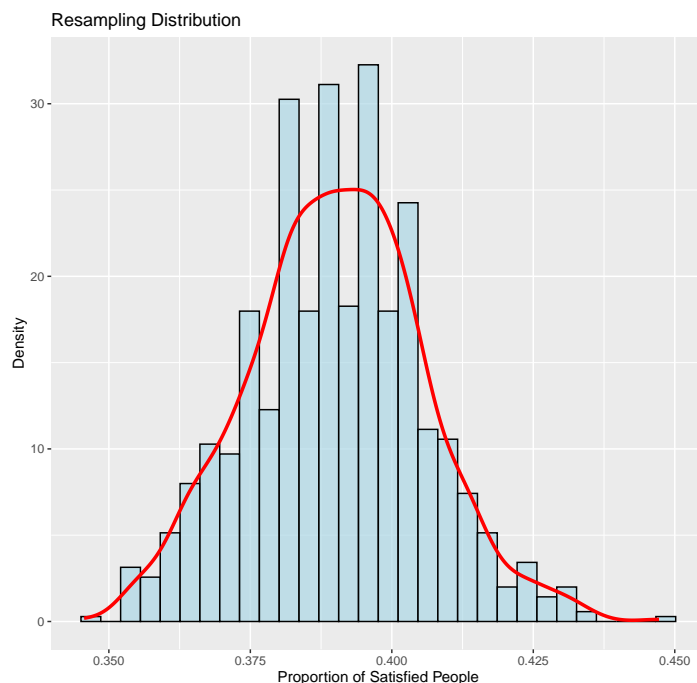


Figure 3: Resampling Plot

The resampled margin of error is approximately 2.67 per-

cent, which aligns with the simulaiton results and Gallup's 4 percent margin of error.

# 4 Simulation over n and p

We simulated 10,000 polls over various values of n and p to estimate margin of error.
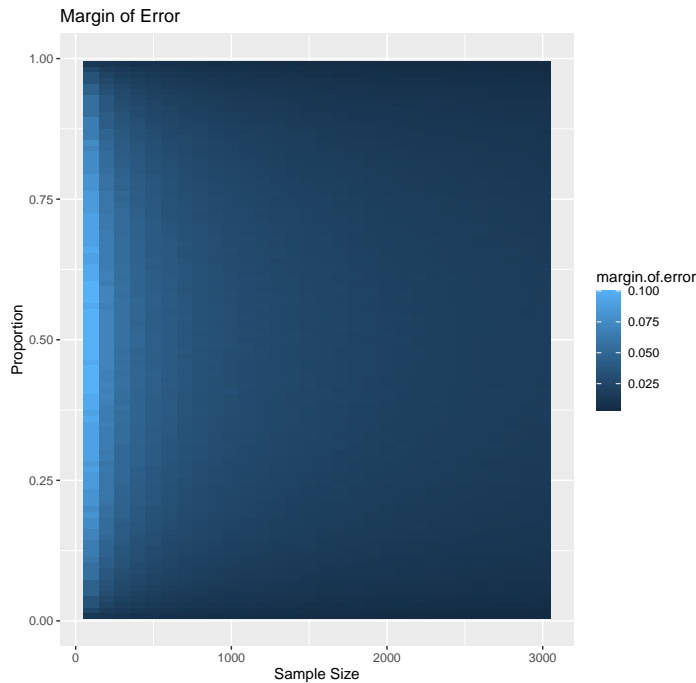


Figure 4: Simluation with n and p plot

The margin of error is highest near p equals 0.5, and decreases with bigger samples. This plot could help Gallup understand tradeoffs between cost and accuracy of his experiment.

# 5 Wilson Margin of Error Calculation

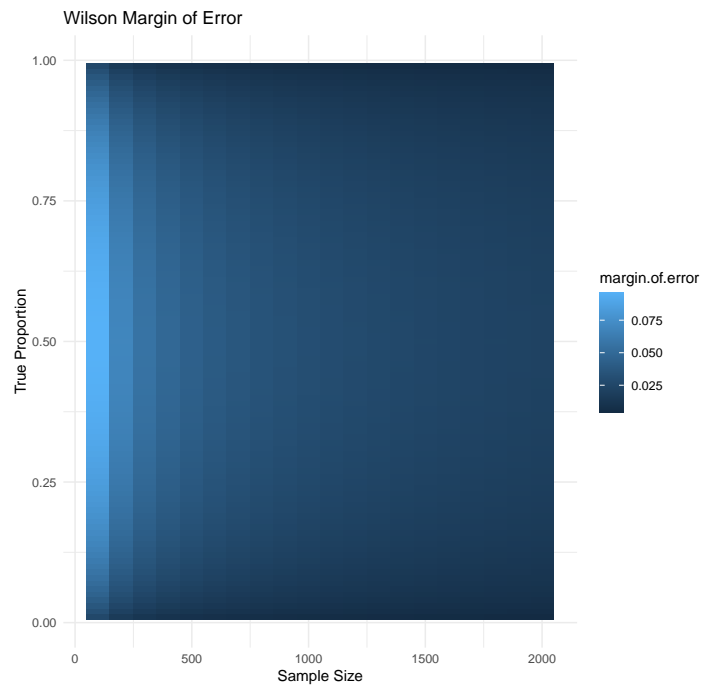We computed wilson confidence intervals for the same values of n and p.



Figure 5: Margin of Error

The Wilson margins are similar to simulated ones. This shows that this can be useful when data is really sparse, or when p is extreme.

# 6 Discussion

Overall, these simulations align with Gallup's claim about how sample size affects margin of error. Larger samples can lead to smaller margins, and this relationship held true across both basic and resampling methods. The Wilson interval also gave similar results, showing it can be a good alternative, especially when data is limited or when the proportion is very low or high. This lab helped us understand why those who study polls, like Gallup, make choices about sample size and how they try to keep their error low while still being efficient.

**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.

Wickham, H. (2023). *Tidyverse: Easily Install and Load the 'Tidyverse'* [R package version 2.0.0]. Retrieved from https://www.tidyverse.org
Gallup, Inc. (2023). *Margin of Error in Polling.* Retrieved from https://news.gallup.com/polling-methodology.aspx