Lab 10 – MATH 240 – Computational Statistics

Jack Schaeffer
Professor Cipolli
Math 240
jschaeffer@colgate.edu

April 8, 2025

Abstract

Focused on analyzing effects on the margin of error, we calculated differing approximate margins of errors based on Gallup polls. We first determined how this value changes based on sample size before comparing the effect within a binomial distribution in comparison to sample size and p.

Keywords: Binomial distribution, margin of error, resampling

1 Introduction

Questions concerning margin of error stemmed from the Gallup polls article "*How Are Polls Conducted?*" that described the formation of polls and how statistical calculations are made. The article asserted that doubling their 1000 person sample size would serve to halve the margin of error, and my work was testing this with my own calculations. My initial work was testing margin of error after doubling sample size of a known p value before moving on to compare these values with resampling of the Gallup polls data. The final test on margin of error was comparing the margin of error of varying sample size and p values through both simulations of the binomial distribution and calculation of the Wilson Estimate.

2 Methods

My work was mainly focused on simulating or estimating data that could then be summarized with `tidyverse` and `patchwork` to obtain more useful information (Wickham et al., 2019; Pedersen, 2024). To first answer the question of doubling sample size, I ran two separate simulations of the binomial distribution under the assumption that Gallup polls produced an accurate p value. Although I could not use resampling to test a larger sample size than the given 1004, resampling was an effective way to compare the simulated margin of error to ensure that the given p value is a reasonable assumption.

I tested the margin of error over numerous combinations of sample size and p value in the same way as my initial simulations, but scaled to include far more combinations of p and n values. I used a similar combination fo p and n values to test the margin of error calculated through the Wilson Estimate, so that I could compare the effect of sample size and p value on margin of error through both methods.

3 Results

To compare the margin of error of different sample sizes, I constructed Figure 1, which features a smaller range for the plot with a larger sample size. Figure 1 contains data made with the assumption that $p = .39$ and simulating this data for different sample sizes. Calculation on the data reveals that a sample size of 1004 leads to a 3.0% margin of error while a sample size of 2008 has a 2.1% margin of error.

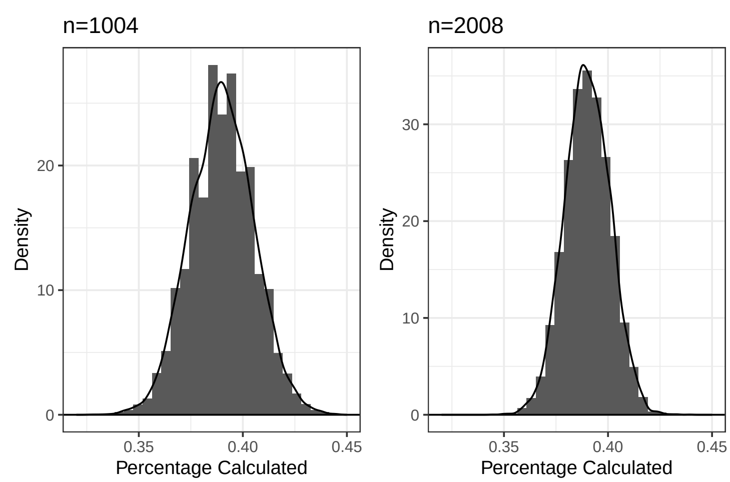

Figure 1: A comparison of simulations for different sample sizes

Calculating the same values using resampling of a sample size $n = 1004$ yielded similar results in Figure 2. When resampling, I calculated a 2.9% margin of error, nearly the same as the 3.0% under the assumption that $p = .39$ is correct.

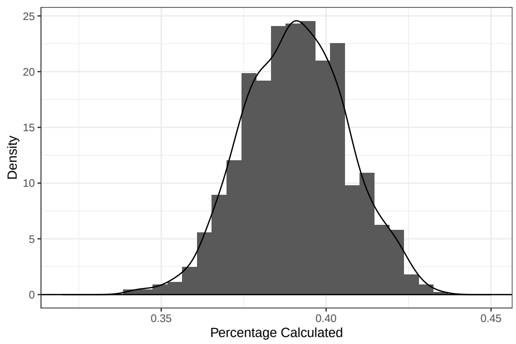

Figure 2: Calculated p values with resampling

Figures 3 and 4 showcase how margin of error changes dependent on samples size and p.

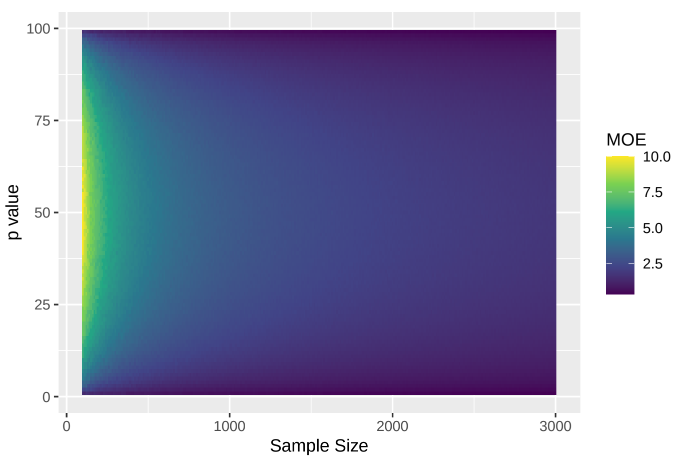

Figure 3: Simulated margin of error for binomial distribution

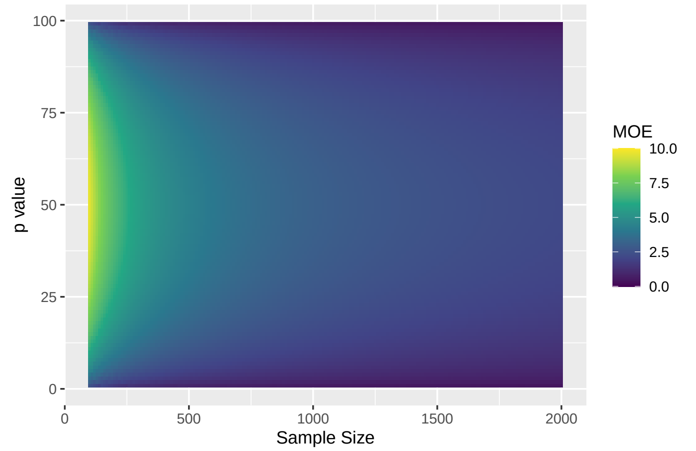

Figure 4: Margin of error calculated through Wilson Estimate

4 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?

References

Pedersen, T. L. (2024). *patchwork: The Composer of Plots*. R package version 1.3.0.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686.