Lab 10 – MATH 240 – Computational Statistics

Henry Sun
Colgate University
Department of Mathematics
hlsun@colgate.edu

4/8/25

Abstract

In this week's lab, we conducted statistical analysis on Gallup poll data. Gallup suggests that doubling the sampling size will halve the margin of error (MOE). We show that this is actually not the case, as the margin of error depends on both the sample size (n) and the population proportion (p).

Keywords: Simulations; resampling; Wilson Estimate; margin of error

1 Introduction

Gallup polls published a document called "*How Are Polls Conducted?*" that describes how Gallup selects people to include in its poll and other details. Near the end of the document, Gallup stated that:

"For example, with a sample size of 1,000 national adults (derived using careful random selection procedures), the results are highly likely to be accurate within a margin of error of ±4 *percentage points."*

"If Gallup increases a poll sample size to 2,000, the results would then be accurate within $\pm2\%$ *of the underlying population value, a gain of two percentage points in terms of accuracy, but with a* 100% *increase in the cost of conducting the survey."*

Gallup provided the results of a February 3-16, 2025 poll of 1,004 adults aged 18+ living in all 50 U.S. states and the District of Columbia, revealing that 39% of of respondents were satisfied with the position of the United States in the world today, compared to 59% that were dissatisfied and 2% with no opinion. For this poll, Gallup reported a $\pm4\%$ margin of error (MOE).

When providing the MOE, we are aiming to describe the variation in our sample using the sample proportion $\hat{p}$ as an estimate of p. We are basically asking how much we expect sample proportion $\hat{p}$ to differ from the population proportion p. In order to maintain a small MOE, we have to make sure that the sample size is large enough, and that the sample is representative of the population, which we will explore in this lab.

It is convention in statistics to report a margin of error that provides 95% confidence, meaning that if we were to conduct the poll repeatedly, the interval would contain the actual proportion 95% of the time.

In this lab, we first began by calculating the MOE by conducting a basic simulation using 10k polls, assuming that we know p, the population proportion.

We then used resampling to obtain a sampling distribution for $\hat{p}$ and calculated the MOE.

Finally, we varied n and p, conducting a simulation using 10k polls for each combination of n and p, seeing how the estimated MOE and Wilson MOE varied with n and p.

2 Methods

For this lab, we are working with data given by Gallup polls. The results data from the polls were generated by creating a data frame with the exact proportions given in the poll. In order to plot our graphs and create tables, we used `ggplot2` (Wickham, 2016) and `xtable` (Dahl, 2022). `patchwork` (Pedersen, 2024) was used to combined plots together.

2.1 Basic Simulation

For our basic simulation, we assumed that the true probability/population proportion, p, of someone being satisfied with the position of the United States today is 0.39. We could then simulate 10k polls using `rbinom()` using these conditions, also using the sample size given by the Gallup poll, $n = 1004$. We then repeated another simulation using the same conditions, except we doubled the sample size.

2.2 Resampling

In the previous section, we had made the assumption that the population proportion p was exactly 0.39. However, in nearly any other situations, we are never given the true population proportion, nor are we able to just magically double the sample the sample size. Instead, we can use resampling of the original Gallup survey and plot the sampling distribution of $\hat{p}$, the sample proportion, for 1000 resamples.

2.3 Simulation over n and p

In order to show how both sample size (n) and the population proportion (p) affect MOE, we conducted as similar procedure in the "Basic Simulation" Section, except we varied n from 100 to 3000 in increments of 10 and p from 0.01 to

0.99 in increments of 0.01, and calculated the estimated MOE for each case. We then summarized the results by creating a `geom_raster()` plot, showing MOE as function of n and p.

2.4 Actual Margin of Error

When finding the actual margin of error, we used the Wilson margin of error formula, given as:

$$\text{MOE}_\text{w} = z_{1-\alpha/2} \times \frac{\sqrt{n\hat{p}(1-\hat{p}) + \frac{z_{1-\alpha/2}{}^2}{4}}}{n + z_{1-\alpha/2}{}^2}.$$

Where $z_{\alpha/2}$ is the z-score at the $\alpha/2$th percentile.

For our lab, $\alpha = 0.05$. We repeated the same procedure highlighted above in "Simulation over n and p" and summarized the results by creating another `geom_raster()` plot.

3 Results

3.1 Basic Simulation

Figure 1 shows the histogram and superimposed density of the sampling distribution for p and sample sizes. The red lines denote the population proportion p. Both plots show an approximately Gaussian distribution, but the histogram with the larger sample size has considerably less variation, showing that a larger sample size will lead to convergence towards the true p.

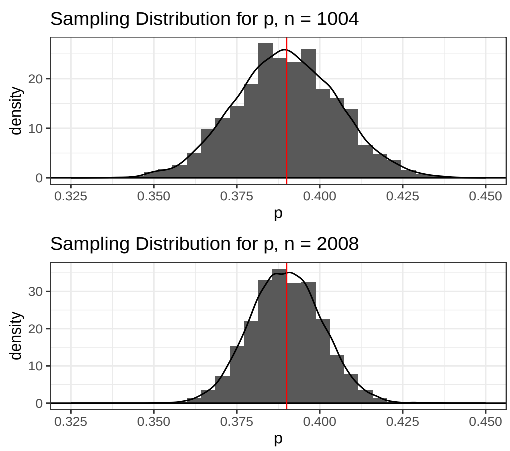

Figure 1: Plots for the sampling distribution of p. The distribution is approximately Gaussian for both, but there is variation with $n = 2008$.

n	Middle 95	Estimated MOE	Gallup MOE
1004	0.0608	0.0303	0.04
2008	0.0423	0.0211	0.02

Table 1: Table comparing the estimated MOE compared to Gallup's reported MOE.

As it turns out, Gallup's reported MOE for $n = 1004$ is significantly off from the estimated MOE. For our lab, we calculated the estimated MOE by halving the range of the middle 95%, which is calculated by subtracting the 97.5th percentile by the 2.5th percentile. As we can see in the table (Table 1), Gallup actually overestimated its MOE for the smaller sample size, although it's MOE for the larger sample size is roughly the same as the estimated MOE. In effect, we've showed that the original MOE of 4% is not very precise, and that doubling the sample size will not halve the MOE.

3.2 Resampling

The figure (Figure 3) in the appendix shows this sampling distribution of $\hat{p}$, which is once again approximately Gaussian due to its large sample size, and centered around 0.39, The estimated MOE calculating for this sampling distribution is 0.0304, yielding similar results to the basic simulation MOE for the sample size. Once again, Gallup's reported MOE of 4% is overestimating the estimated MOE from resampling.

3.3 Simulation of n and p

The figure below (Figure 2) show the `geom_raster()` plots for both the estimated MOE (calculated by halving the middle 95%) and the Wilson MOE (calculated by the formula given before). In the top plot (the plot for the estimated MOE), we can see that the sample size is only one variable affecting the MOE. While sample size does reduce the MOE, MOE also changes with p. That is, when p is extreme, the MOE is smaller. The intersection between the black line and the red line shows the estimated MOE at $n = 1004$ and $p = 0.39$, and the intersection between the black line and the red line shows the estimated MOE at $n = 2008$ and $p = 0.39$. We can clearly see that MOE decreases with a larger sample size, but not linearly.

3.4 Actual Margin of Error

The bottom plot in figure 2 shows pretty much the same thing highlighted in the `geom_raster()` plot for the estimated MOE, except that we are calculating the actual Wilson MOE plotting it as a function of n and p. This paints a nearly identical picture and yields similar results to the estimated MOE. The only noticeable difference is that the gradient between MOE boundaries are slightly smoother.

4 Discussion

As we've shown in this lab, the story isn't as simple as Gallup made it out to be. First of all, Gallup's estimates for the MOE were inaccurate to begin with. The estimated MOE for $n = 1004$ is actually more like 3% compared to Gallup's 4%. Gallup is actually *overreporting* the MOE. However, Gallup's MOE and the estimated MOE are nearly identical when doubling the sample size.

Although we did show in the basic simulations that doubling sampling size decreased the MOE, Gallup implies that doubling the sample size will correspond to halving the MOE. As shown above this is not true. Why? because MOE also depends on p, as shown by the `geom_raster()` plots of the estimated MOE and the Wilson MOE.

References

Dahl, D. B. (2022). *xtable: Export Tables to LaTeX or HTML*. R package version 1.8-4.
Pedersen, T. L. (2024). *patchwork: The Composer of Plots*. R package version 1.3.0.
Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

5 Appendix

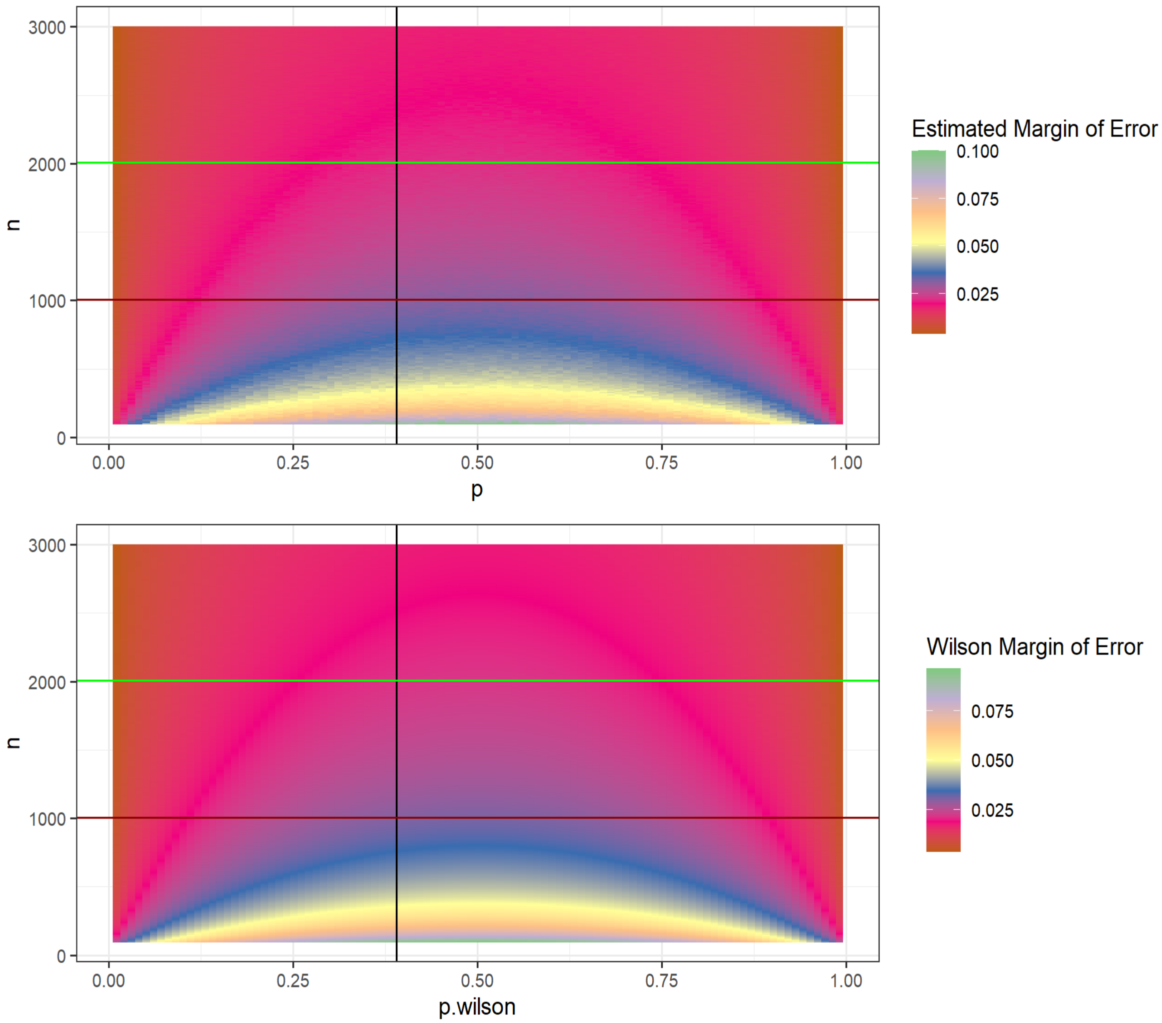

Figure 2: Plots showing how the estimated MOE varies with n and p (top) and how the Wilson MOE varies with n and p.

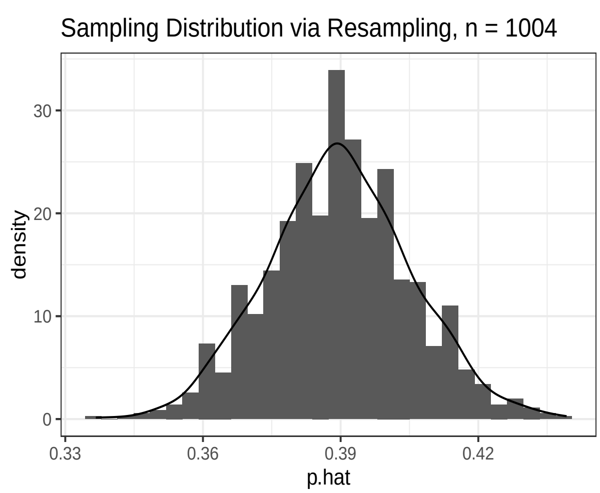

Figure 3: Plot showing the sampling distribution for $\hat{p}$ using resampling. For this plot, we conducted 1000 resamples.