Week 6, Day 2

Sarah Moore

2022-10-27

Continuing on with Uncertainty

- First, let's recap why we deal with uncertainty.

Potential measures and ways to visualize uncertainty:

Continuing on with Uncertainty

- First, let's recap why we deal with uncertainty.

Potential measures and ways to visualize uncertainty:

- Margin of error, confidence intervals, or standard errors are the most likely to be published in text. We can easily translate these from one to another

Continuing on with Uncertainty

- First, let's recap why we deal with uncertainty.

Potential measures and ways to visualize uncertainty:

- Margin of error, confidence intervals, or standard errors are the most likely to be published in text. We can easily translate these from one to another

- Visually we are likely to encounter error bars, highlighted confidence intervals, or some other sort of visual differentiation between a given estimate and the surrounding potential range of values.

Confidence Intervals

- We will encounter confidence intervals here (CIs for short). I'd like for us to know how to interpret this correctly, especially related to our confidence levels.
- Confidence intervals specify a potential range of values for a potential measurement of interest, given an associated confidence level.

$$CI = \text{Measurement} \pm MoE = [\text{Upper Bound, Lower Bound}]$$

- Where the **Margin of Error (MoE)** is the standard error ($\frac{s}{\sqrt{n}}$) multiplied by the Z-or t-score relative to the chosen confidence level (typically 95%, but also 90% and 99% possible).

Quick note of difference: standard deviation and standard error

- Does anyone know right off the bat what the difference is?

Quick note of difference: standard deviation and standard error

- Does anyone know right off the bat what the difference is?
- *Standard deviation* is an estimated statistical parameter about the spread of our data. Basically the expected amount of variation from observation to observation in the sample, and consequentially inferred about the population.

Quick note of difference: standard deviation and standard error

- Does anyone know right off the bat what the difference is?
- *Standard deviation* is an estimated statistical parameter about the spread of our data. Basically the expected amount of variation from observation to observation in the sample, and consequentially inferred about the population.
- *Standard error* is a measure of precision relative to the sample and its associated standard deviation.

Formulae

Sample Standard Deviation of Mean s:

$$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{n-1}}$$

where the numerator is the squared sum of differences between each observation x_i of the variable of interest and the variable's mean ($\bar{x}$). The denominator is the number of observations n minus 1. Easily done in R by `sd(var, na.rm = T)`.

Sample Standard Deviation of Proportion s:

$$s = \sqrt{\hat{\pi}(1 - \hat{\pi})}$$

Where $\hat{\pi}$ is the sample proportion.

Formulae

Standard Error (SE):

If based on a mean:

$$SE = \frac{s}{\sqrt{n}}$$

Where s is the sample standard deviation.

If based on proportion:

$$SE = \frac{s}{\sqrt{n}} = \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}}$$

Margin of Error $\frac{s}{\sqrt{n}} \times z$

$$MOE = \frac{s}{\sqrt{n}} \times z$$

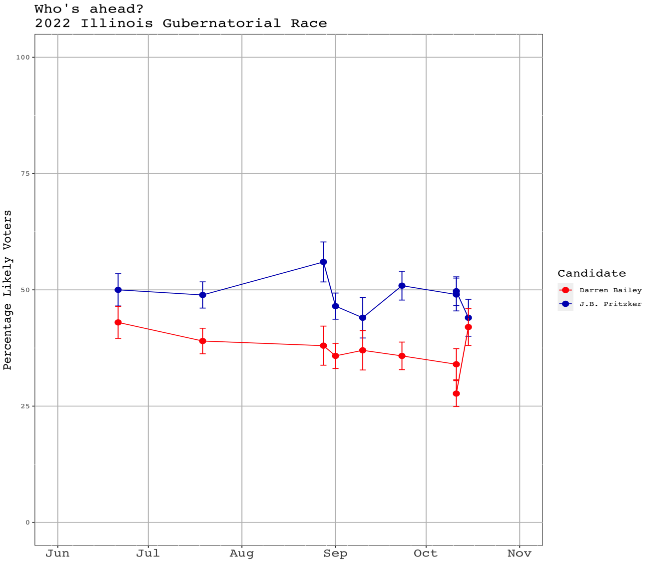
Where z is one of the following values:

Confidence Level	z
0.70	1.04
0.75	1.15
0.80	1.28
0.85	1.44
0.90	1.645
0.92	1.75
0.95	1.96
0.96	2.05
0.98	2.33
0.99	2.58

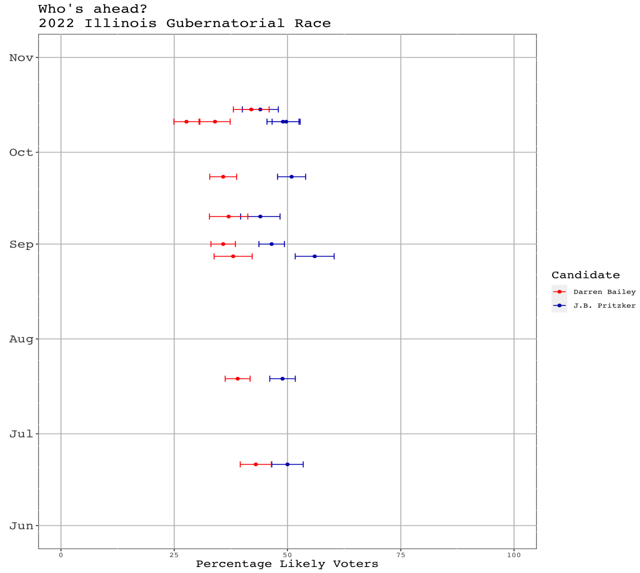
Confidence levels

- Typically we choose 95% as our baseline. However if we visualize confidence intervals, can easily account for 90% and 99% visually also.
- The higher the confidence level, the more wider the CI and thus less precise.

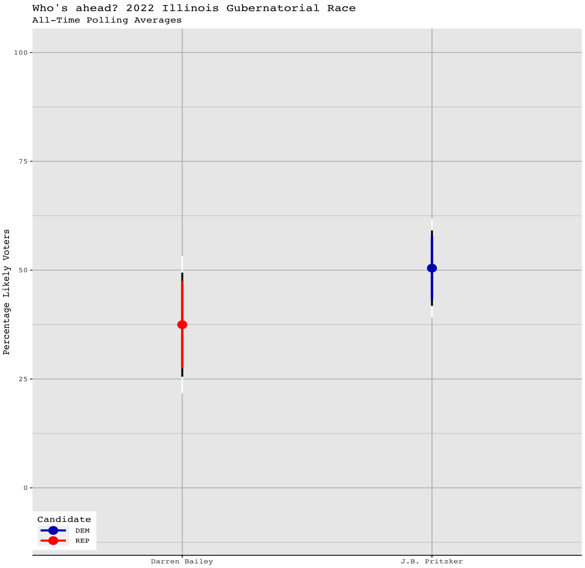
Point Estimates and Uncertainty



Point Estimates and Uncertainty



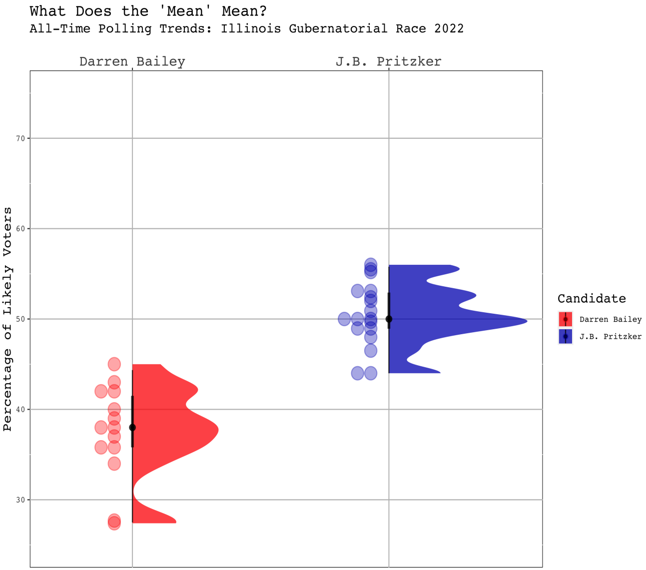
All Time?


Who's ahead? 2022 Illinois Gubernatorial Race
All-Time Polling Averages
Percentage Likely Voters
100
75
50
25
0
Candidate
DEM
REP
Darren Bailey
J.B. Pritzker

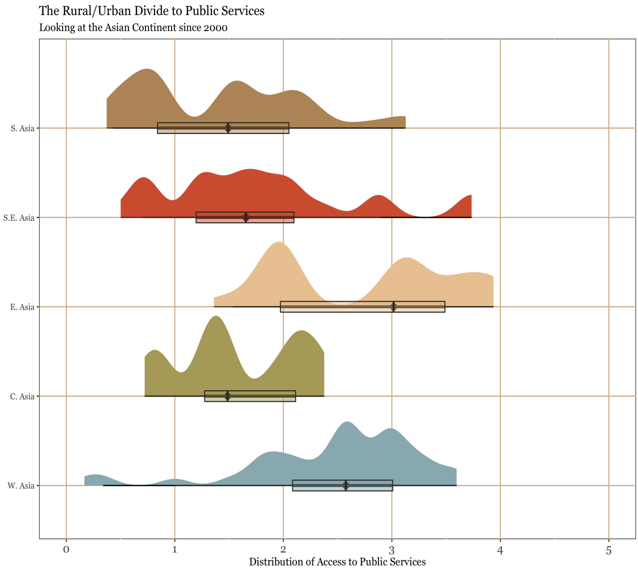
Another option for uncertainty + distributions

https://www.cedricscherer.com/2021/06/06/visualizing-distributions-with-raincloud-plots-and-how-to-create-them-with-ggplot2/

With the IL Data



Switching Data



Switching Region