Trump expected to win 54% popular vote in 2020 US Election!

An evidence found from survey data using GLMM model with 95% CI from 52% to 56%

Yangming Hu

2020-11-02

1 Model

The study uses a generalized linear mixed logistic model with random effects across states as below:

$$Y_{ij} \sim \text{Bernoulli}(\lambda_{ij})$$

$$log(\lambda_{ij}) = \mu + X\beta + state_i + Z_{ij}$$

Let we assume that the covariance matrix Σ_{state} is among voters for state level and Σ_Z is the covariance matrix among individuals within state level.

$$state_i \sim N(0, \sigma^2_{state}\Sigma_{state})$$
$$Z_{ij} \sim N(0, \sigma^2_Z\Sigma_Z)$$

And

1. Y_{ij} is whether vote for Trump or not for ith state, j-th voter.

2. λ_{ij} is the chance of vote for Trump.

3. μ is the intercept, X are covariates,

4. β are coefficients.

5. State is the random effect for state level.

6. Z is random effect for invidual level.

As binary outcome is appropriate for logistic model, and random effect across states require random effect model, so the GLMM model is appropriate instead of linear regression model. Also, model diagnostics are used to verify the model is valid.

2 Results

Table 1 Forecasting Trump's winning probability with 95% confidence interval

Variable	Forecast
Probability	0.54
S.E.	0.01
95% CI	(0.52,0.56)

Using the model, the main results are shown in the table 1 which is the forecasting Trump's winning probability with 95% confidence interval, so the results show that Trump expected to win about 54% vote in 2020 US Election, that we are 95% confident that this predicted proportion of vote fall into the range from (52%, 56%). As 52% is still higher than 50%, so this study forecasts that Trump would win the 2020 US Election.

3 Discussion

This sutdy mainly uses a generalized linear mixed logistic model with random effects across states to forecast the US 2020 election, it was found that Trump expected to win about 54% vote in 2020 US Election, that we are 95% confident that this predicted proportion of vote fall into the range from (52%, 56%). And because the 52% is still higher than 50%, so this study forecasts that Trump would win the 2020 US Election. However, besides the findings, there are still some weaknesses in this study. For examples, the non-response biases in the survey data, the data cleaning processes and etc, in future work, these issues could be invistigated to improve the performance of the model and fndings in this study.

4 Appendix

The r codes and soure files are hold in the Github Repo: https://github.com/Jaydenhu123/PS3.

5 References

1. Hadley Wickham and Evan Miller (2020). haven: Import and Export 'SPSS','Stata' and 'SAS' Files. R package version 2.3.1. https://CRAN.R-project.org/package=haven

2. Mollie E. Brooks, Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Maechler and Benjamin M. Bolker (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal, 9(2), 378-400.

3. Patrick Brown (2020). Pmisc: Various Utilities for knitr and inla. R package version 0.3.2/r2380. https://R-Forge.R-project.org/projects/diseasemapping/

4. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

5. Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020.https://doi.org/10.18128/D010.V10.0

6. Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape, October 10-17, 2019 (version 20200814). Retrieved from https://www.voterstudygroup.org/downloads?key=196fe68b-363c-46f1-880b-75b48cd5dc4d

7. Wickham,Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

8. Yihui Xie (2020). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.27.