

Bayesian Trivariate Analysis of an Opinion Poll: With Application to the Kenyan Polls

Jeremiah Kiingati*, Samuel Mwalili, Anthony Waititu

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University Agriculture and Technology, Nairobi, Kenya

Email address:

jkiingati@gmail.com (J. Kiingati), samuel.mwalili@gmail.com (S. Mwalili), agwaititu@gmail.com (A. Waititu)

*Corresponding author

To cite this article:

Jeremiah Kiingati, Samuel Mwalili, Anthony Waititu. Bayesian Trivariate Analysis of an Opinion Poll: With Application to the Kenyan Polls. *International Journal of Data Science and Analysis*. Vol. 6, No. 1, 2020, pp. 58-63. doi: 10.11648/j.ijdsa.20200601.17

Received: January 15, 2020; **Accepted:** February 4, 2020; **Published:** March 24, 2020

Abstract: There has been a growing interest by political pundits and scholars alike to predict the winner of the presidential elections. Although forecasting has now quite a history, we argue that the closeness of recent Kenyan presidential opinion polls and the wide accessibility of data should change how presidential election forecasting is conducted. We present a Bayesian forecasting model that concentrates on the national wide pre-election polls prior to 2013 general elections and considers finer details such as third-party candidates and self-proclaimed undecided voters. We incorporate our estimators into WinBUGS to determine the probability that a candidate will win an election. The model predicted the outright winner for the 2013 Kenyan election.

Keywords: Presidential Elections, Election Forecasting, Operations Research, Bayesian Prediction Models

1. Introduction

1.1. Background

Several researchers in the field of political science focus their attention on the historic problem of understanding and predicting election outcomes. Understanding voter preferences and forecasting the final outcome of elections is of critical importance to politicians, as they can use the insights gained from the exercise to fine-tune their campaign strategies. Consequently a substantial literature on the prediction of election is available. Some authors like [7, 16] have used regression models for nationwide polling data to forecast the outcome of the popular vote in the US elections while others use state-level polls to make quantitative estimate of the proportion of votes for the two major party candidates in each state [3, 9, 15, 10, 18, 4]. However, prediction of who will actually win the presidency, the issue of principal interest, is not addressed by the analyses of popular opinion trends.

Multilevel logistic regression model to generate estimates of state-level vote shares was employed by [17]. Their model employs national opinion data and state level demographic covariates to obtain estimates in a manner that is related to

the small-area estimation problem.

Predicting election day voting outcomes based on early pre-election polling is a very complicated problem because such a prediction would require a consideration of opinion trends, future campaign spending, and historical voter behavior. Further, the actual election day results will be affected by many unpredictable factors arising in the final days of the campaign, including world events and candidate mistakes. Effectively the majority will is difficult to measure and fundamentally ambiguous [14, 12, 13]. Consequently, we focus on the simpler problem of estimating the probability that the incumbent president would win the election if it were held on the day of the recent poll.

1.2. Kenya's Presidential Elections

Recent general elections indicate that the Kenya's voting population is highly polarized along ethnic lines. Election campaigns in Kenya have often focused on internal issues, often driven by ethnic agenda and lacking in regional or geopolitical focus. Indeed a recent discussion with a varied group of African enthusiasts revealed that Kenyan politics ignores what should be fundamental to the country's development – that is geopolitics. Luckily, there are emerging internal sobering political dynamics with a new generation of

young urban citizens that consider themselves as distinctly different from their older counterparts, least bothered about ethnic affiliations and therefore difficult to manipulate using tribal messages.

All presidential candidates enter the contest with a full knowledge of this and therefore tailor their campaign strategies accordingly. Candidates spend far more of their campaign funds and time in regions that are close and pay scant attention to voters in regions where the outcome is basically foregone [11, 5, 1, 2].

2. Methodology

2.1. Model Description

More formally, define θ_i to be the true proportion of voters in a county who intend to vote for candidate i in the election (for simplicity, let $i = 1$ correspond to the incumbent candidate, $i = 2$ correspond to the main opposition candidate, $i = 3$ collectively correspond to all the other minor candidates (also called third force), and $i = 4$ correspond to no candidate or voters who have declared that they are still undecided). These proportions are assumed to be continuous (between 0 and 1) and sum to 1. The joint prior distribution for $\theta = (\theta_1, \theta_2, \theta_3)$ is assumed to be a conjugate prior distribution (i.e., the resulting posterior belongs to the same distributional family as the prior distribution). To satisfy this requirement, assume that θ follows a Dirichlet distribution, $\theta \sim \text{DIRICHLET}(\alpha_1, \alpha_2, \alpha_3)$, which is a multivariate generalization of the beta distribution and is often used as a prior for the probability of a success in Bernoulli trials. Therefore, the joint probability density function of θ can be written as

$$f(\theta_1, \theta_2, \theta_3) \propto \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} \theta_3^{\alpha_3-1}, \theta_i > 0, i = 1, 2, 3 \text{ and } \sum \theta_i = 1$$

The probability that a candidate wins a given county can be computed using the marginal probability densities. To obtain these marginals, we sequentially integrated the remaining variables out of the joint Dirichlet probability density function. We now illustrate this process by first rewriting the joint Dirichlet probability density function as

$$f(\theta_1, \theta_2) \propto \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} (1 - \theta_1 - \theta_2)^{\alpha_3-1}, \theta_1, \theta_2 \geq 0, i = 1, 2 \text{ and } \sum \theta_i \leq 1$$

Integrating over θ_3 leads to an expression for the joint probability density of θ_1 and θ_2

$$\begin{aligned} f(\theta_1, \theta_2) &= \int_0^{1-\theta_1-\theta_2} c \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} (1 - \theta_1 - \theta_2)^{\alpha_3-1} d\theta_3 \\ &= c \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} \int_0^{1-\theta_1-\theta_2} (1 - \theta_1 - \theta_2)^{\alpha_3-1} d\theta_3 \end{aligned}$$

These results leads to the expression

$$f(\theta_1, \theta_2) \propto \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} (1 - \theta_1 - \theta_2)^{\alpha_3-1}, \theta_1, \theta_2 \geq 0 \text{ and } \theta_1 + \theta_2 \leq 1$$

Integrating over all possible values of θ_2 gives the marginal density of θ_1 ,

$$\begin{aligned} f(\theta_1) &= \int_0^{1-\theta_1} K \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} (1 - \theta_1 - \theta_2)^{\alpha_3-1} d\theta_2 \\ &= K \theta_1^{\alpha_1-1} \int_0^{1-\theta_1} \theta_2^{\alpha_2-1} (1 - \theta_1 - \theta_2)^{\alpha_3-1} d\theta_2 \end{aligned}$$

The results leads to the expression

$$f(\theta_1) \propto \theta_1^{\alpha_1-1} (1 - \theta_1)^{\alpha_2+\alpha_3-1}, 0 \leq \theta_1 \leq 1$$

Therefore, by the form of $f(\theta_1)$, θ_1 is distributed as a beta random variable with parameters α_1 and $\alpha_2 + \alpha_3$. Using the identical argument, θ_2 and θ_3 are also distributed as beta random variables; hence, $\theta_i \sim \text{Beta}(\alpha_i, \sum \alpha_j - \alpha_i)$ $i = 1, 2, 3$

2.2. Choice of Prior Parameters

The beta prior distribution is characterized by 2 shape parameters, here, α_1 and $\alpha_2 + \alpha_3$, which must be chosen. Different choices can be incorporated, which would result in quite varied substantive implications. For example, one possibility is to set these values so that the expected value and the variance, or the first and second moments, for $\theta_i, i = 1, 2, 3$, closely match observed elections. However, there are an infinite number of combinations of α_1, α_2 and α_3 that result in the same values for the expectations and the variances for θ_1, θ_2 and θ_3 . That is, if the expectation and variance of a beta random variable are given by

$$E(\theta_i) = \frac{\alpha_i}{\sum_{i=1}^3 \alpha_i} \text{ and } \text{var}(\theta_i) = \frac{\alpha_i (\sum_{i=1}^3 \alpha_i - \alpha_i)}{(\sum_{i=1}^3 \alpha_i)^2 (\sum_{i=1}^3 \alpha_i + 1)}$$

The choice of the shape parameters is essentially arbitrary if we do not issue any constraints or do not use any substantive guidance. Fortunately, in presidential forecasting, we have a great deal of substantive knowledge that can be integrated. One way to constrain our choices is to choose the $\alpha_{i,s}$ so that

- $\alpha_i / \sum \alpha_i$ equals what θ_i is expected to be, prior to observing the polling data, and
- the spread of the prior distribution for θ_i (determined by $\sum \alpha_i$, with larger values indicating less uncertainty) reflects the perceived uncertainty in θ_i .

3. Results from the Bayesian Model Fit

A Bayesian was fitted to the Kenyan 2013 opinion data comparing three candidates; the two leading candidates and the 'other', which comprises of all the other remaining candidates. The parameters of interest here are $(\alpha_1, \alpha_2, \alpha_3)$ and $(\theta_1, \theta_2, \theta_3)$ which predict the probability of each of the three candidates. The model was fitted in WinBUGS with post-analysis in R.

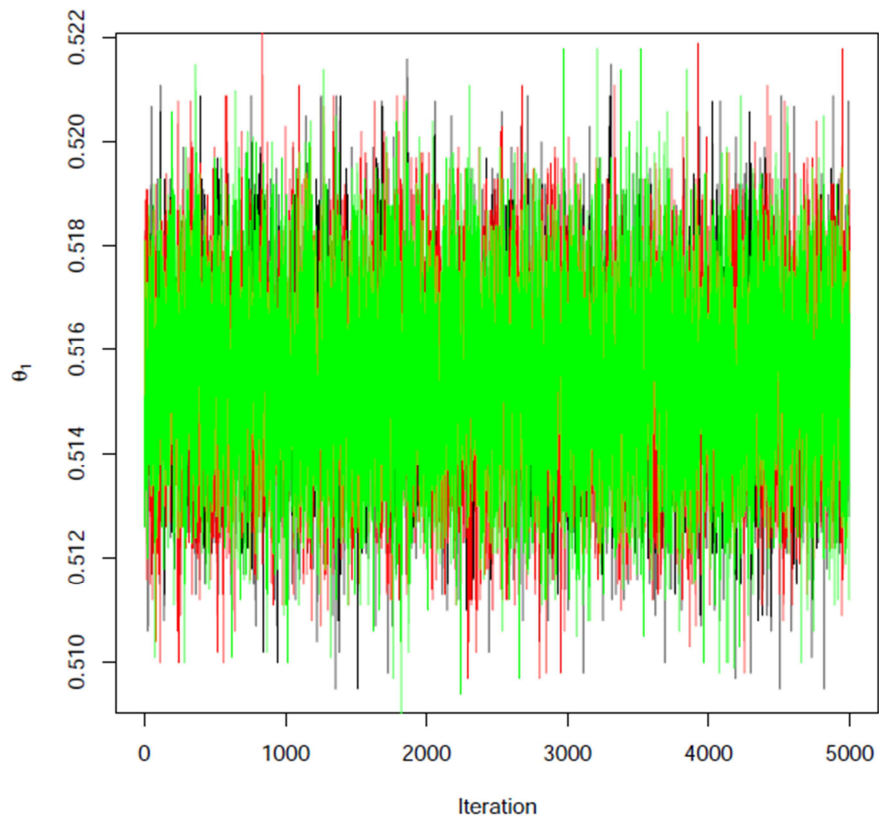


Figure 1. Trace plot of the θ_1 , the parameter for the leading candidate, from the trivariate model of the 2013 Kenyan poll.

From plots of the three chains, for each parameter, as can be seen in Figures 1-3, it is immediately clear that the posterior estimates of the parameters are converging because the chain mix like spaghetti.

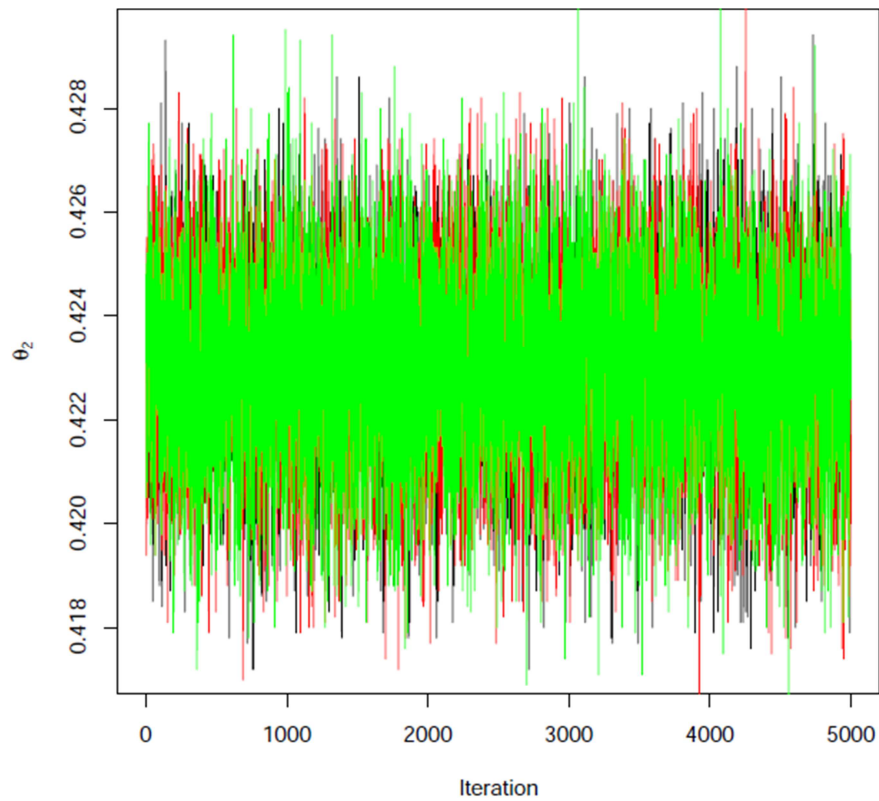


Figure 2. Trace plot of the θ_2 , the parameter for the second candidate, from the trivariate model of the 2013 Kenyan poll.

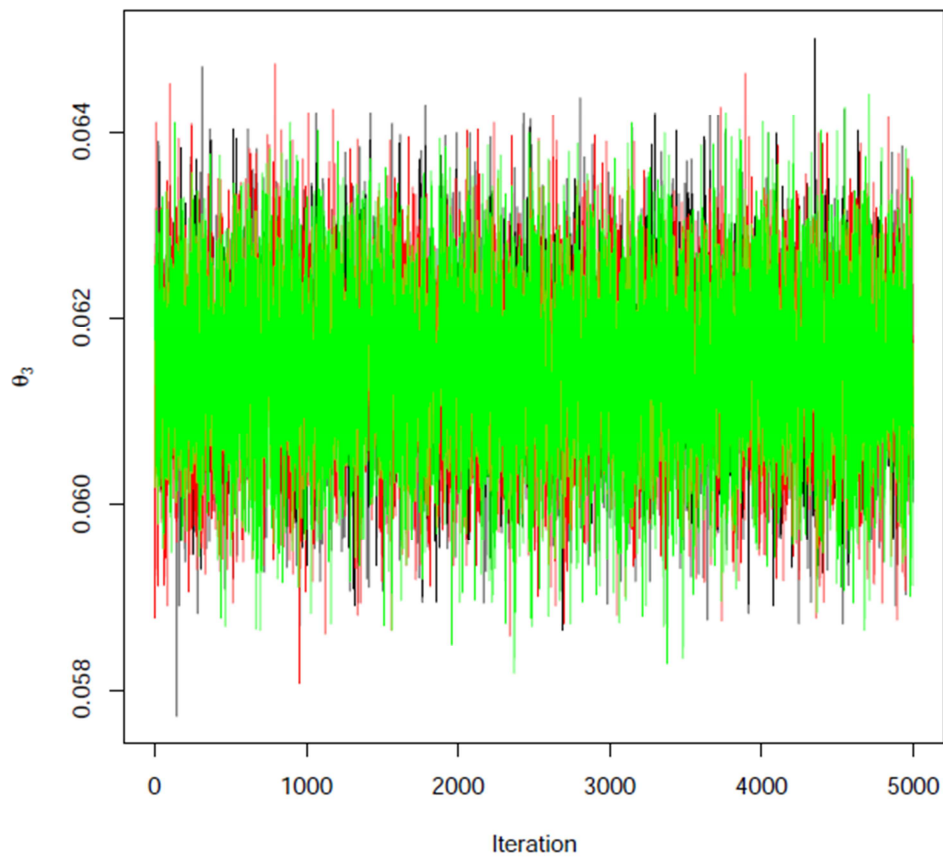


Figure 3. Trace plot of the θ_3 , the parameter for the other candidates, from the trivariate model of the 2013 Kenyan poll.

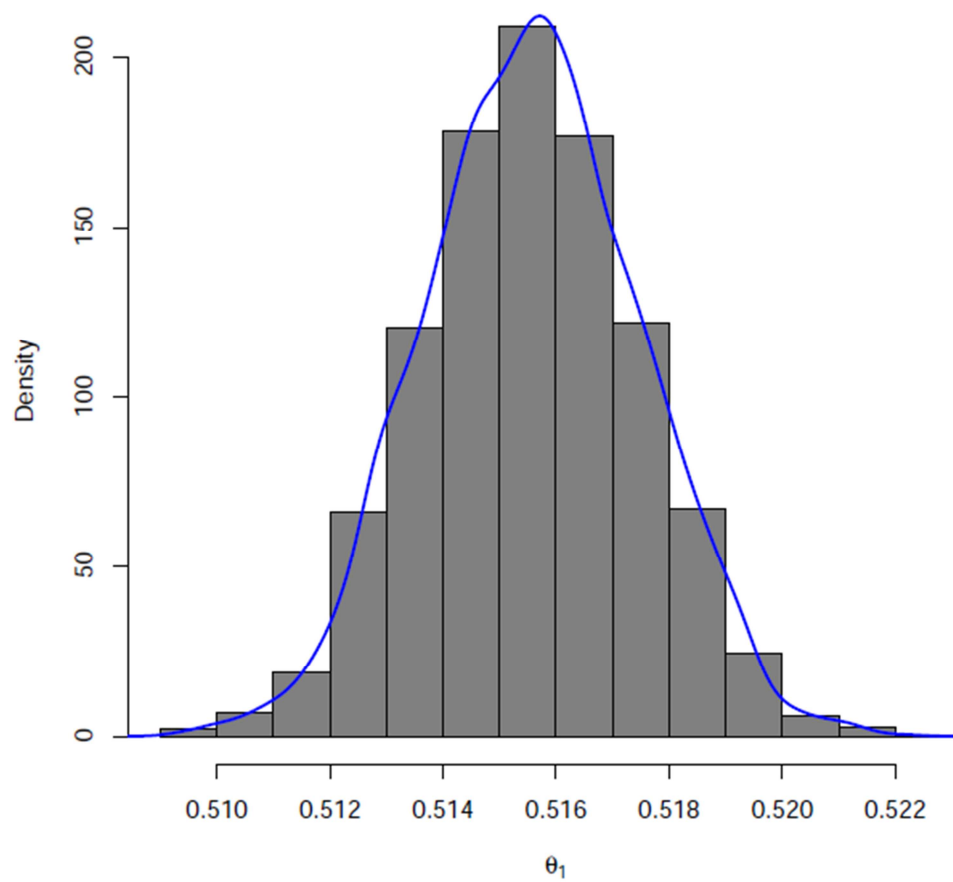


Figure 4. Density plot of the θ_1 , the parameter for the leading candidate, from the trivariate model of the 2013 Kenyan poll.

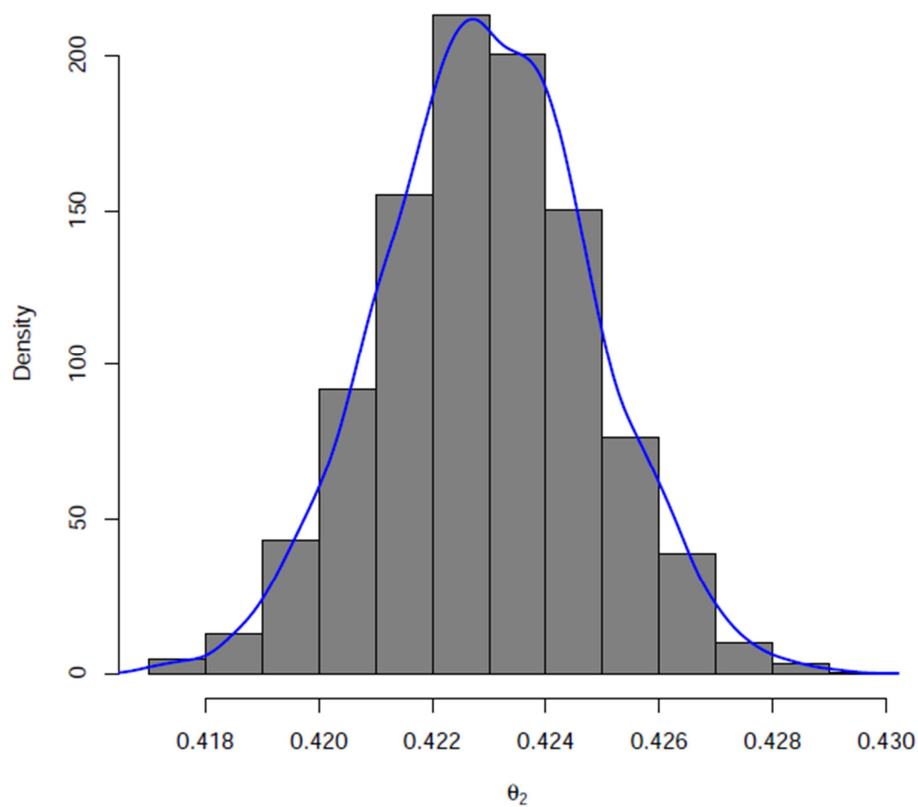


Figure 5. Density plot of the θ_2 , the parameter for the second candidate, from the trivariate model of the 2013 Kenyan poll.

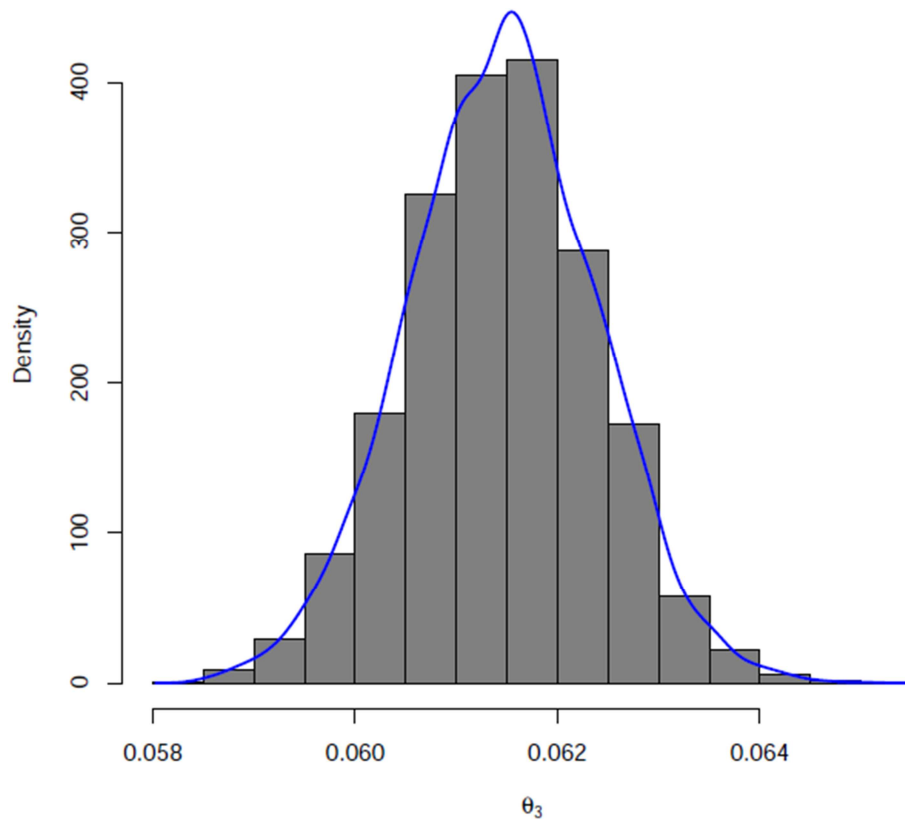


Figure 6. Density plot of the θ_3 , the parameter for the other candidates, from the trivariate model of the 2013 Kenyan poll.

The density plots of the parameters of interested, shown in Figures 4-6, indicate that the posterior distributions of $(\theta_1, \theta_2, \theta_3)$ are symmetrical and take nearly a normal

distribution shape. This implies that the posterior point estimates can estimated by posterior mean or median.

Table 1. Parameter point and standard errors from the trivariate model of the 2013 Kenyan poll.

Parameter	Estimate	SE	95% Credibility Interval (CI)	
			Lower	Upper
α_1 *	0.000		—	—
α_2	-0.198	0.0078	-0.213	-0.183
α_3	-2.127	0.0166	-2.160	-2.095
θ_1	0.515	0.0019	0.512	0.519
θ_2	0.423	0.0019	0.419	0.427
θ_3	0.0615	9.3E-04	6.0E-02	0.063

* α_1 is set to 0 as a baseline parameter.

From these results the leading Candidate is projected to win with 51.5% (95% CI 51.2-51.9) with the second candidate at with 42.3% (95% CI 41.9-42.7) and the other remaining candidate at 6.2% (95% CI 6.0-6.3). Since the 95% credibility interval do not overlap, the model seem to predict incumbent as the outright winner.

4. Conclusions

Bayesian analysis provided a very powerful approach to modeling opinion poll data. The posterior mode gives unbiased estimate of the parameter of interest. Furthermore, it averaged over many data points, thus taking care of possible inconsistencies in some data points. Even though in this case we considered three candidate the model to be easily extend to more than three candidates.

Acknowledgements

We acknowledge the Kenyan media houses and opinion pollsters who were the making sources of these Kenyan opinion polls data.

References

- [1] Abaramowitz, Alan I. 2012. "Forecasting in a Polarized Era: The Time for Change Model and the 2012 Presidential Election." P. S. Political Science and Politics 45, 4, 618-619.
- [2] Abramowitz, Alan I. 2008. "Forecasting the 2008 Presidential Election with the Time-for-Change Model." P. S. Political Science and Politics 41, 4, 691-695.
- [3] Abrams, Burton A. and James L. Butkiewicz. 1995. "The Influence of State-Level Economic Conditions on the 1992 U.S. Presidential Election." Public Choice 85, 1, 1-10.
- [4] Ahammed, M. S., Newaz, M. N., & Dey, A. (2019). Analyzing Political Opinions and Prediction of Voting Patterns in the US Election with Data Mining Approaches. Global Journal of Computer Science and Technology.
- [5] Althaus, S. L., Nardulli, P. F., & Shaw, D. R. (2002). Candidate appearances in presidential elections, 1972-2000. Political Communication, 19 (1), 49-72.
- [6] Campbell, J. E. (2014). Issues in presidential election forecasting: election margins, incumbency, and model credibility. PS: Political Science & Politics, 47 (2), 301-303.
- [7] Campbell, J. E. (2016). The Trial-Heat and Seats-in-Trouble forecasts of the 2016 Presidential.
- [8] Congressional elections. PS: Political Science & Politics, 49 (4), 664-668.
- [9] Cohen, J. E. (2006). Public opinion in state politics. Stanford University Press.
- [10] Cohen, J. E. (2015). Presidential Leadership in Public Opinion. Cambridge University Press.
- [11] Dorsey, Margaret E., and Miguel Diaz-Barriga. "Senator Barack Obama and immigration reform." Journal of Black Studies 38, no. 1 (2007): 90-104.
- [12] Feld, S. L., & Grofman, B. (2010). Puzzles and paradoxes involving averages: An intuitive approach. In Collective Decision Making (pp. 137-150). Springer, Berlin, Heidelberg.
- [13] Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. Social Science Computer Review, 31 (6), 649-679.
- [14] Gelman, A., Katz, J. N., & Tuerlinckx, F. (2002). The mathematics and statistics of voting power. Statistical Science, 420-435.
- [15] Holbrook, T. M., & DeSart, J. A. (1999). Using state polls to forecast presidential election outcomes in the American states. International Journal of Forecasting, 15 (2), 137-142.
- [16] Kennedy, R., Wojcik, S., & Lazer, D. (2017). Improving election prediction internationally. Science, 355 (6324), 515-520.
- [17] Park, D. K., Gelman, A., & Bafumi, J. (2004). Bayesian multilevel estimation with poststratification: State-level estimates from national polls. Political Analysis, 12 (4), 375-385.
- [18] Wang, W., Rothschild, D., Goel, S., & Gelman, A. (2015). Forecasting elections with non-representative polls. International Journal of Forecasting, 31 (3), 980-991.