BAYESIAN INFERENCE OF MULTINOMIAL DATA USING SEQUENTIAL UPDATING: A CASE STUDY OF THE NIGERIAN PRESIDENTIAL ELECTION

Consul, Juliana Iworikumo*, Okrinya, Aniayam

Department of Mathematics and Computer Science, Niger Delta University, Bayelsa State, Nigeria.

IJASR 2019
VOLUME 2
ISSUE 3 MAY – JUNE **ISSN: 2581-7876**

Abstract – This research is focused on the Bayesian inference of multinomial data using sequential updating of the Nigerian presidential polls. The research has supposed that there are three distinct proportions which are the probabilities that the sample of the politicians would be allocated into each of the categories of the political parties. The prior beliefs about the proportions were taken as the choices of the parties of the three hundred and sixty elected House of Representatives. This is given a Dirichlet distribution. Based on sequential updating, the posterior distribution after observing the 2015 Nigerian election data is the prior distribution before observing the 2019 Nigerian election data. The posterior distribution and credible interval of the parameters were also summarised. It was seen that the data had little effect on the prior distribution and obviously the Nigerian electoral system makes it almost impossible for new political parties to become successful. It was also seen that the credible intervals for all parameters were very narrow which implied a smaller chance of obtaining an observation outside the interval, high accuracy and more precise estimates.

Keywords: Prior distribution, posterior distribution, sequential updating, Dirichlet distribution, multinomial distribution..

1.0 Introduction

Election is the formal process of selecting a person for public office by voting. It is often thought of that election is the very core of democracy which involves electing the decision makers and hence getting rid of those that are disliked. Democracy is the best form of government in which the expressions of the people are shown through legitimacy and leadership succession. In the democratic system, people vote for candidates of their choice and these votes are counted. The major aim of democracy should be ensuring political stability and promoting human rights.

Elections in Nigeria are administered by the Independent National Electoral Commission (INEC). The main functions of the INEC include conducting and registration of voters' card, registering and regulating political parties and organising and supervising elections. The electoral system in Nigeria consists of a set of rules that govern all aspects of the voting process. This process involves who is allowed to vote, stand as a candidate, how ballots are marked and cast and the factors that affect the outcome of the election. In Nigeria, election depends heavily on financial contributions from candidates for the elections. It is well known that money plays a critical role in politics since it is used for lobbying for both legislators and administrative agencies. It is also almost always that the candidate that raises the most money wins.

Nigeria has an active registration system that requires all citizens to vote in person at the polling unit where they have registered. The electronic voter registration software is being used for the elections. The INEC registers all voters using a biometric data system which include a photograph and fingerprints. The voters' registration card must be presented on the day of election. Nigeria adopted the open ballot system in the third Nigeria republic where voters vote openly by queuing to indicate the candidate of their choice. The majority system is one where the candidate with majority of the votes is elected as the winner.

It was recorded that the 2011 Nigeria Presidential election was the most smoothly ran elections that have held since the restoration of democracy 12 years before. One of the problems encountered during the 2015 election was that the website of the INEC was hacked on the day of election. There were also technical problems with electronic card readers in the 2011, 2015 and 2019 elections (Wikipedia, 2011, 2015 and 2019 Nigerian election). Nwokeke et al. (2011) examined the effects of election rigging on the democratic consolidation in Nigeria. Election rigging has claimed lives of citizens of Nigeria because of the politicians are desperate of power. This has made politics to be seen as the game that is played with seriousness.

In this research, we will look at Bayesian Inference in probability models particularly involving more than one unknown parameters by applying it to the analysis of a multinomial data such as the Nigerian election data using a conjugate Dirichlet prior distribution and sequential updating. The 2011, 2015 and 2019 Nigerian election data will be used to study the parameters which represent the proportions of the voters that voted for the candidate of the political parties using Bayesian inference and sequential updating. In this case, the beliefs are represented by a model containing several unknown parameters.

The basics of Bayesian inference requires that the prior experience with observed data (in the form of likelihood)is used to interpret these data (in the form of a posterior distribution) (Catanach et al, 2018). The prior distribution should reflect information about the model parameters (Garthwaite et al, 2005). It is very important in Bayesian inference because it is often used to bring the analyst closer to what is being modelled. The prior distribution could be an opinion of an expert within the field of investigation from whom information is being elicited (Berger et al, 2009). In the case of this research, the purpose is simply to communicate the results of an investigation and hence, reasonable prior specifications are used. It is possible that we have multivariate prior elicitation and we belief that our unknown parameters are independent and hence information about some of the parameters would not affect the beliefs about the other parameters (Revie et al., 2010). The prior distribution can also be informative. For instance, the posterior distribution of the previous model which is similar to the form of a present model is used as the prior distribution of the present model. The present model might not start from scratch and not based only on the present data but on the cumulative effects of past and present data which are taken into account.

We refer to Bayes theorem (Ronquist et al.,2003 and Gelman et al., 2004) for some basis of Bayesian inference. When applying Bayesian inference, the prior is multiplied by the likelihood function and hence the prior affects the posterior distribution. This requires calculating the normalising term which being an integral can sometimes be problematic to evaluate. The normalising constant is usually found by integrating over all parameters. In some cases, the (unnormalised) posterior is the product of the prior and likelihood. Sometimes, the posterior distribution cannot be represented analytically in realistically complex problems because of the intractability of the normalising constant (Berger, 2005). In practice, it is very common that such integrations are carried out numerically using computer soft wares or suitable numerical methods such as Markov chain Monte Carlo methods (MCMC) like Metropolis and Metropolis Hastings algorithm, Gibbs sampler and Metropolis within Gibbs algorithm etc. The MCMC methods are used to draw samples from the posterior distribution to approximate the posterior. Once the posterior distribution is calculated, summaries of the parameters can also be calculated without any complications.

The purpose of Bayesian inference from the election data is mostly to study the parameters used to represent the proportions of voters that voted for the candidate of the political parties. Andrew Gelman (2017) discussed various choices of designing and analysing opinion polls and election data using a Bayesian approach by differentiating between the election studies in Spain and the United States. Gabriel et al. (2015) performed a Bayesian inference on proportional elections using Monte Carlo simulation technique by considering the Brazilian system of seats distribution. Rigdon et al., (2009) presented a Bayesian approach to forecasting by incorporating some estimators into a dynamic programming algorithm to determine the probability that a candidate for election will win. Some researchers like Bailer – Jones (2013) also treated data and parameter probabilistically by inferring the probability distribution over the model parameters of interest. The efficient sampling of multidimensional distribution by applying Bayesian inference using Monte Carlo methods was also discussed in Bailer – Jones (2013).

2.0 Methodology

2.1 The Multinomial and Dirichlet Distribution

We refer to a binomial distribution as a distribution with probability associated with two outcomes in which the experiment consists of a set of N independent trials with probability of success, θ. We suppose that the distribution of the two possible outcomes in a set of N trials is known as the number of "success", Y_1 and the number of "failures", $Y_2 = N - Y_1$. The binomial distribution can be generalised to a multinomial distribution by allowing for more than two possible outcomes in each of the set of N trials (Avetisyan et al., 2012). Suppose that we have c distinct number of categories, $1,2, \ldots, c$ and we observe $Y_1, Y_2, \ldots . Y_c$ with $\sum_{i=1}^{c} Y_i = N$. The probabilities $\theta_1, \ldots, \theta_c$ for the categories are such that $\sum_{i=1}^{c} \theta_i = 1$. Given $\theta = (\theta_1, \ldots, \theta_c)^T$, the multinomial distribution of $Y_1, Y_2, \ldots . Y_c$ is given by

$$\Pr(Y_1 = y_1, \ldots, Y_c = y_c) = \frac{N\,!}{\prod_{i=c}^{c} y_i!} \prod_{i=1}^{c} \theta_i^{\,y_i}$$

The probability density function of the Dirichlet distribution of a set of probabilities $\theta = \theta_1, \theta_{2,} \ldots .. \ \theta_c$ where $\theta_i \geq 0$ and $\sum_{i=1}^{c} \theta_i = 1$ is given by

$$f\left(\theta_1, \theta_{2,} \ldots .. \ \theta_c\right) = \frac{\Gamma(A)}{\prod_{i=1}^{c} \Gamma(a_i)} \prod_{i=1}^{c} \theta_i^{\,a_i - 1}$$

where we have that $A = \sum_{i=1}^{c} a_i$ and $a_1, a_2, \ldots \ldots a_c$ are parameters for $i = 1,2, \ldots c$. The mean of θ_k is given as

$$E(\theta_k) = \int \ldots . \int \theta_k \frac{\Gamma(A)}{\prod_{i=1}^{m} \Gamma(a_i)} \prod_{i=1}^{c} \theta_i^{\,a_i - 1} \ d\theta_1, d\theta_2 \ldots \ldots . . d\theta_c$$

$$= \frac{a_k}{A}$$

The variance of $\theta_k$can be expressed as

$$Var(\theta_k) = E\left({\theta_k}^2\right) - \left(E(\theta_k)\right)^2$$

$$= \frac{A(a_k+1)a_k - A + 1\,(a_k)^2}{A^2(A+1)}$$

2.2 Bayesian Inference

The prior belief about the parameters, $\pi(\theta)$ expresses the probability which is a means of quantifying uncertainty before taking the data into account. The posterior distribution, $\pi(\theta \mid y)$ combines the likelihood, $L(\theta, y)$ and the prior $\pi(\theta)$, by capturing all that is known about the parameters using the Bayes' formula.

In this case, the likelihood from a multinomial $(N, \theta_1, \ldots, \theta_c)$ distribution is given by

$$L(\theta, y) \propto \prod_{i=1}^{c} \theta_i^{\,y_i}$$

We choose the conjugate form of prior for the likelihood since the posterior distribution is in the same family as the prior distribution (Raiffa et al., 1961). In this case, the prior and posterior are called conjugate distribution. The conjugate prior for the categories of probability of a multinomial distribution, $\theta = (\theta_1, \theta_{2,} \ldots .. \ \theta_c)^T$ is a Dirichlet distribution since the calculation work out neatly.

We wish to determine the posterior distribution of the unknown parameters. The posterior distribution is

$$\pi(\theta \mid y) \propto prior \ \times likelihood$$

$$\propto \pi(\theta)\, L(\theta, y)$$

$$= \prod_{i=1}^{c} \theta_i^{\,a_i - 1} \times \prod_{i=1}^{c} \theta_i^{\,y_i}$$

$$= \prod_{i=1}^{c} \theta_i^{\,a_i + y_i - 1}$$

The posterior probability density function (pdf) is proportional to that of a Dirichlet distribution with parameters $a_{1+}y_1, a_2 + y_2, \ldots \ldots, a_c + y_c$.

We note that the calculation is simple using a conjugate prior distribution since it matches the likelihood in the sense that the posterior distribution belongs to the same family as the prior distribution. The pressure of using a convenient conjugate prior can be removed by solving the integration numerically using computer. In cases where the prior beliefs are not represented by a conjugate distribution, we use a different distribution and resort to numerical evaluation of the posterior distribution (Gelfand et al., 1990).

Some researchers like Kern (2006) have conducted Bayesian inference on multinomial probabilities using data collected from a pig game where prior information on the parameter of interest were available and was easily incorporated in a Dirichlet prior distribution. Avetisyan et al., (2012) extended the Dirichlet – Multinomial model for categorical randomised response data and sampling algorithms produced samples from the posterior distribution using Markov chain Monte Carlo to produce samples.

2.3 Sequential Updating

Our beliefs about our unknown parameter θ changes from a prior distribution, $f^{(0)}(\theta)$ to a posterior distribution with density $f^{(1)}(\theta)$ when we observe data $y^{(1)}$. So,
$f^{(1)}(\theta) \propto f^{(0)}(\theta)\, L\big(\theta; y^{(1)}\big)$

where the likelihood is $L(\theta; y)$.

We will suppose that the process is repeated and we start with a prior distribution with density $f^{(0)}(\theta)$. If we observe data $y^{(1)}$ and we represent our beliefs about the parameter by $f^{(1)}(\theta)$. Again, we suppose that we observe further data $y^{(2)}$ so that the density becomes $f^{(2)}(\theta)$. The joint probability density of θ is

$$f^{(0)}(\theta) f_1\big(y^{(1)} \mid \theta\big)$$

where $f_1(y^{(1)} \mid \theta)$ is the conditional probability density of $y^{(1)}$ given θ and $f_2(y^{(2)} \mid \theta, y^{(1)})$ is the conditional probability density of $y^{(2)}$ given θ and $y^{(1)}$. We have that

$$f^{(1)}(\theta) \propto f^{(0)}(\theta) f_1\big(y^{(1)} \mid \theta\big)$$

$$= f^{(0)}(\theta) L_1(\theta, y^1)$$

and

$$f^{(2)}(\theta) \propto f^{(1)}(\theta)\, f_2(y^{(2)} \mid \theta, y^{(1)})$$

$$= f^{(1)}(\theta)\, L_2(\theta, y^{(2)} \mid y^{(1)})$$

We will assume that $y^{(1)}$ and $y^{(2)}$ are independent of θ. In such case, we have that

$$f_2\big(y^{(2)} \mid \theta, y^{(1)}\big) = f_2\big(y^{(2)} \mid \theta\big)$$

Therefore,

$$f^{(2)}(\theta) \propto f^{(0)}(\theta) L_1\big(\theta, y^{(1)}\big) L_2(\theta,\ y^{(2)})$$

We will note that the posterior distribution after observing data $y^{(1)}$ is the prior distribution before observing data $y^{(2)}$.

In sequential updating, the current state of knowledge regarding parameters in terms of the posterior distribution is used as prior distribution when new data becomes available. The posterior distribution is then constructed in such a way that we avoid repeating the computation of the likelihood of the old data as the new data becomes available. Jones (2016) illustrated the advantages of sequential Bayesian updating with a set of data to study Alzheimer's Dementia. Friedman et al., (1996) described the sequential updating of Bayesian network by including necessary modifications for learning Bayesian network by evaluating the effectiveness and also extending to the case of missing data. Catanach et al.,(2018) discussed some MCMC methods for Bayesian updating and system reliability assessment called sequential tempered algorithms.

2.4 Bayesian Credible Interval for the parameters

It is very natural that we may want to make a probability statement regarding the parameters after observing the data. The Bayesian credible interval of the parameter comes to play in this case. The Bayesian credible interval of size $1-\alpha$ for proportion P_i is an interval (L_{P_i}, U_{P_i}) such that

$$Pr\left(L_{P_i} \le P_i \le U_{P_i} \mid y\right) = 1-\alpha$$

We calculate a credible interval for all parameters by calculating the credible intervals for each parameter marginally since the marginal of a Dirichlet distribution is a Beta distribution. The marginal distribution of P_i is $Beta\,(a_i, \sum_{i=1}^{k} a_i - a_i)$ for $i = 1, \dots k$ and $0 < p_i < 1$ (Balakrishnan et al, 2004). We can therefore find L_{P_i} and U_{P_i} by calculating the $\alpha/2$ and $1-\alpha/2$ quantiles of the marginal beta distribution using R software (R Core Team, 2014).

3.0 Application to the Nigerian Election Data

3.1 *Data Collection*

Nigeria is a multi-party system in which two or more strong parties can be electorally successful. Political parties are formal organisations that recruit citizens into various offices through an electoral process. These political parties strive to produce leaders such as Presidents, Senate President, Speakers, Governors, House of Representative members e.t.c. The highest elected office at the Federal level in Nigeria is the presidency. The Federal legislative is the Nigerian National Assembly with 109 Senators and 360 member House of Representative.

The 2015 and 2019 Nigerian Presidential elections data will be used in this research. The 2015 and 2019 Nigerian Presidential elections were held on the 29th of March, 2015 and 23rd of February, 2019 respectively. In 2015, some political parties were registered in Nigeria but only 14 had candidates for the Presidential election (Wikipedia, 2015 Nigerian election) while 72 parties registered in 2019.

In this research, we will use all results from election because during election campaigns people were encouraged to vote by claims that ``every vote matters". The electoral system allows as many as possible parties to register. For instance, in the 2011 election, the votes for 20 political parties were recorded but only the two leading candidates (Peoples Democratic Party (PDP) and Congress for Progressive Change (CPC)) had very high votes (22,495,187 and 12,214,853 respectively) (Wikipedia, 2011 Nigerian election). Prior to the 2015 Nigerian Presidential elections, the All Progressives Congress (APC) was formed as an alliance of four opposition parties. These four parties were the Action Congress of Nigeria (ACN), the Congress for Progressive Change (CPC), the All Nigeria Peoples Party (ANPP) and the All Progressive Grand Alliance (APGA). In the 2015 election, the votes for 14 political parties were recorded and the two leading candidates (APC and PDP) had very high votes (15,424,921 and 12,853,162 respectively). In the 2019 election, the votes for 72 political parties were recorded and the two leading candidates (APC and PDP) had very high votes (15,191,847 and 11,262,978 respectively).

3.2 Bayesian inference of electoral data using sequential updating

We wish to apply our discussion on the Bayesian inference of probability models involving more than two unknown parameters with application to the Nigerian election data as a Multinomial distribution and a conjugate Dirichlet prior distribution using sequential updating. We assume there are three categories of parties, `APC', `PDP' and `others'. In the case of `others', we sum up the values of all other parties. The votes of `ACN', `CPC', `ANPP' and `APGA' will added up for the votes for `APC' since it was formed as an alliance of these four opposition parties. The summary of the categories of the parties for the 2015 and 2019 elections are given in Table 1.

Table 1: Summary of the categories of the Political parties for the 2015 and 2019 elections.

Categories	2015 election votes	2019 election votes
APC	15,424,921	15,191,847
PDP	12,853,162	11,262,978
Others	309,481	894,057
Total	28,587,564	27,348,882

We suppose that the Nigerian election is an experiment in one of c$=3$ (number of political parties) distinct outcomes and that each outcome i occurs independently of the other outcomes with a probability $\theta_i > 0$. We will also suppose that Y_i is the frequency of outcome i in $``N"$ runs of the experiment. A possible distribution for the outcomes of the election data is a multinomial distribution $M(N, \theta)$. In this research, we will assume that the population is large so that the information gathered may be considered independent given the true underlying proportions. In this case, we will let $_1, \dots\dots,$ be the probability that the sample of politicians would be allocated into each of the categories of parties. We have multivariate prior elicitation and we belief that our unknown parameters are independent and hence information about some of the parameters would not affect the beliefs about the other parameters. In this case, we will suppose that the 360 elected House of representatives were asked their choice of which party that they would want to vote for the 2015 Presidential election. Again, we suppose that their choices are their respective parties and hence we use the number of seats of the 2015 House of Representative election as the prior distribution for $_1, _2$ and $_3$ is a Dirichlet distribution. Table 2 summarises the categories of the House of Representative and the number of seats for the 2015 election.

Table 2: Summary of the categories of the House of Representative and the number of Seats.

Categories	Number of Seats
APC	225
PDP	125
Other parties	10
Total	360

Table 3 gives the summary of the prior distribution of the parameters.

Table 3: Summary of the prior distribution of the parameters.

Categories		Prior mean	Prior variance	Prior standard deviation
APC	225	0.625	0.00065	0.025
PDP	125	0.35	0.00063	0.025
Others	10	0.03	7.49×10^{-5}	0.009
Total	360	1		

The prior means, variance and standard deviation for the parameters are also calculated and given in Table 3. Then we can summarise the posterior distribution in Table 4.

We want to summarise the posterior distribution after observing the data $y^{(1)}$. We suppose that the observed data is the 2015 election votes, $y^{(1)}$ as given in Table 1.

At this point, the belief about the unknown parameters for θ_1, θ_2 and θ_3 have changed from the prior distribution, $f^{(0)}(\theta)$ (the number of seats by categories of the 360 elected House of Representatives) to a posterior distribution with density $f^{(1)}(\theta)$ after observing the data (case the 2015 election votes) with the likelihood contribution $L_1(\theta; y^{(1)})$.

We might want to find the credible interval (CI) such that there is a posterior probability of 0.95 that a parameter θ_i lies in the interval $(L_{\theta_i}, U_{\theta_i})$. Such an interval is sometimes called a 95% credible interval for the parameter. We establish a beta distribution for the Dirichlet distribution with three parameters by treating θ_1, θ_2 and θ_3 each as an independent variable. For instance, the probability density function of θ_1 is $Beta(a_1, a_2 + a_3)$. The marginal distribution of all three parameters can also be given as $\theta_i \sim Beta(a_i + y_i^{(1)}, A + n - a_i - y_i^{(1)})$ for $i = 1,2,3$.

Table 4 summarises the posterior means, variance, standard deviation and credible interval of the proportions of voters that voted for the candidate in each category.

Table 4: Summary of the posterior distribution after observing the 2015 election data

Categories	$a_i + y_i^{(1)}$	Posterior mean	Posterior variance	Posterior standard deviation	Lower limit CI	Upper limit CI
APC PDP Others	15,425,146 12,853,287 309,491	0.5396 0.4496 0.0108	8.69×10^{-9} 8.656×10^{-9} 3.746×10^{-10}	9.322×10^{-5} 9.304×10^{-5} 1.935×10^{-5}	0.5394 0.4494 0.0108	0.5398 0.4498 0.0109
Total	28,587,924	1				

In this research, we use sequential updating by supposing that we observe a further data $y^{(2)}$ which is the 2019 election data votes with density $f^{(2)}(\theta)$. We assume that $y^{(1)}$ and $y^{(2)}$ are independent of θ. The posterior distribution after observing data $y^{(1)}$ is the prior distribution before observing data $y^{(2)}$ with the likelihood contribution of $L_2(\theta; y^{(2)})$. We want to summarise the posterior distribution after observing the data $y^{(2)}$. The observed data is the 2019 election votes $y^{(2)}$ as given in Table 1. We will also find an interval such that there is a posterior probability of 0.95 that a parameter θ_i lies in the interval $(L_{\theta_i} U_{\theta_i})$ using sequential updating. The marginal distribution of all three parameters can also be given as $\theta_i \sim Beta(a_i + y_i^{(1)} + y_i^{(2)}, A + n - a_i - y_i^{(1)} - y_i^{(2)})$ for $i = 1,2,3$. Table 5 summarises the posterior distribution after using sequential updating.

Table 5: Summary of the posterior distribution after observing the 2019 election data

Categories	$a_i + y_i^{(1)} + y_i^{(2)}$	Posterior mean	Posterior variance	Posterior standard deviation	Lower limit CI	Upper limit CI
APC PDP Others	30,616,993 24,116,265 1,203,548	0.5473 0.4311 0.0215	4.43×10^{-9} 4.38×10^{-9} 3.76×10^{-10}	6.66×10^{-5} 6.62×10^{-5} 1.94×10^{-5}	0.5472 0.4310 0.02147	0.5475 0.4313 0.0216
Total	55,936,806	1				

DISCUSSIONS

In this research, we have supposed that there are three parameters. These are the probability that the `APC', `PDP' or `other party' candidate wins (represented by θ_1, θ_2 and θ_3 respectively). We needed to specify prior distributions for the three parameters and we have supposed that the 360 elected House of representatives were asked their choice of party that they would want to vote for the 2015 Presidential election. It is most likely that they will want to vote for candidates in their respective parties. Their responses were used as the prior information. We have used the Dirichlet distribution with parameters 225,125,10 or $D_3(225,125,10)$. The posterior distribution was constructed using sequential updating where the posterior distribution after observing

the data from the 2015 election votes are the prior distribution before observing the data from the 2019 election votes.
The proportion of the voters that will vote for the `APC', `PDP' candidate and `others' are 0.5473, 0.4311 and 0.0215 respectively using sequential updating. We notice that the posterior means for APC parameter, θ_1 is smaller than the expected in the prior. For θ_1, we have a posterior mean of 0.5473 and a posterior standard deviation of 6.66×10^{-5} compared to the prior mean of 0.625 and the prior standard deviation of 0.025. The reverse is the case for the PDP parameter. The results from the posterior distribution using the likelihood from the 2015 votes and 2019 votes did not vary much. The results also show that after seeing the data, it appears that a greater proportion of the voters still voted for the APC candidate than the PDP candidate. It is also seen that the data had little effect on the prior distribution. It is clear that it does not matter whether we update our beliefs using the 2015 election votes or the 2019 election votes first. Using sequential updating or not gave the same posterior means but slightly different posterior variances. It is seen that the credible intervals for the parameters were very narrow.

5.0 CONCLUSION

Bayesian methods capture our uncertainty about parameters using probability distribution and update this understanding as new information are available. We have used the conjugate Bayesian analysis where we have the multinomial likelihood and Dirichlet prior. We note that we must not be compelled to use conjugate prior. It is often that a conjugate prior will be able to represent the beliefs about our parameters closely enough and the calculations become much simpler.

We conclude that since there is no difference in the posterior mean using the likelihood from the 2015 and 2019 election votes but the standard deviation varied. The proportion of the voters that voted for the APC candidate was the highest and so the candidate wins. We will also conclude sequential updating or not gave the same posterior means but slightly different posterior variances. The posterior mean for the category `others' is 0.0215. This is too small.

It is very obvious and expected that voters vote for parties and not candidate. The reverse should be the case. Voters are supposed to vote for specific candidates and not parties. It is also obvious that the electoral system makes it almost impossible for new political parties to become successful. The system makes it hard for a new political party to be taken seriously and for it to gradually increase its standing in the political system. It is typical that new parties grow by starting small and then gaining credibility first in the local level before moving to the higher levels of the system.

The credible intervals for all parameters were very narrow. This implies that there is a smaller chance of obtaining an observation outside the interval and hence the accuracy is very high. It is seen that a large sample of data (the Nigerian election data) tend to give more precise estimates and hence narrower credible intervals than the smaller samples.

Reference

Andrew Gelman (2007), Bayesian Analysis of Election Surveys and forecasts: learning from the State of Valencia Experience.https://pdfs.semanticscholar.org/eff2/a9a9eec812386a06f 2c6bca25d4ac2bb5cff.pdf?_ga=2.143606070.2044591365.15544398822055290400.1554439882.

Avetisyan, M and Fox, J.P (2012). The Dirichlet – Multinomial model for multivariate randomised response data and small samples.

Bailer – Jones C.A.L (2013). Bayesian parameter estimation and model comparison, Astrostats 2013 Lecture, Max Planck Institute for Astronomy, Heidelberg, http://www.mpia.de/~calj/.

Balakrishnan .N and Nevzorov .V.B (2004). A primer on statistical distributions, chapter 27, John Wiley & Sons.

Berger, J. O., Bernardo, J. M. and Sun, D (2009). The formal definition of reference priors. The Annals of Statistics 37 (2), 905-938.

Berger, J. O. (1985). Statistical Decision Theory and Bayesian Analysis (second Edition).

New York: Springer-Verlag.

Catanach .T.A and Beck .J.L (2018). Bayesian updating and uncertainty quantification using sequential tempered MCMC with the Rank- one modified metropolis algorithm.

Final Report (2012), National Democratic Institute for International Affairs,. https://www.ndi.org/sites/default/files/NDI%20Final%20Report%20on%20the%20Nigeria%202011%20Electionsnew_Part1.pdf

Friedman, N. & Goldszmidt .M (1996). Learning Bayesian networks with local structure. In E. Horvits & F. Jensen (Eds.), *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence,* 252–262. San Francisco, CA: Morgan Kaufmann.

Brunello, .G.H.V, Nakano, E.Y (2015), Bayesian Inference on Proportional Elections. PLOS ONE 10(3): e0116924. Doi:10:1371/journal.pone.0116924.

Garthwaite, P. H., Kadane, J. B. & O'Hagan, A (2005). Statistical methods for eliciting probability distributions. Journal of the American Statistical Association 100 (470), 680 -701.

Gelfand A.E and Smith A.F.M (1990), Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85, 398-409.

Gelman .A, Carlin .J, Stern .H and Rubin .D (2004). Bayesian Data Analysis, Second Edition, Chapman & Hall, Boca Raton, FL

Hammersley J.M, Handscomb D.C (1964) Monte Carlo methods, London, Methuen.

Jones .M.N (2016). Big Data in Cognitive Science (Frontiers of Cognitive Psychology). CRS Press, Taylor & Francis Group.

Kern J.C (2006), Pig data and Bayesian inference on multinomial probabilities. Journal of Statistics Education 14(3). DOI: 10.1080110691898.2006.11910593.

Nwokeke P.O, Jayuw J.A (2011). The Electoral process and democratic consolidation in Nigeria. Journal of Politics and Law. Volume 4(2).

R Core Team (2014), R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria.

Raiffa .H, Schaifer .R (1961). Applied Statistical decision theory. Division of Research Graduate school of Business Administration, Havard University.

Revie, M., Bedford, T. & Walls, L. (2010). Evaluation of elicitation methods to quantify Bayes linear models. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 224 (4), 322-332.

Rigdon S.E, Jacobson S.H, Cho, Wendy K.T, Sawell E.C and Rigdon C.J (2009), A Bayesian prediction model for the United States presidential election: American politics Research 37(4), 700-724.

Ronquist .F and Huelsenbeck .J.P (2003). Bayesian Phylogenetic inference under mixed models. Bioinformatics Volume 19 No. 12, 1572 – 1574.

2011 Nigerian Election: https://en.wikipedia.org/wiki/2011_Nigerian_presidential_election

2015 Nigerian Election: https://en.wikipedia.org/wiki/2015_Nigerian_general_election

2019 Nigerian Election: https://en.wikipedia.org/wiki/2019_Nigerian_general_election