STATISTICAL MODELLING OF COVID-19 DEATHS IN NIGERIA

[1]Okereke, Emmanuel Wilfred and [2]George, Uchechukwu

[1,2]Department of Statistics Michael Okpara University of Agriculture, Umudike, Nigeria

[2]Email: george.uchechukwu@mouau.edu.ng

ABSTRACT

In this study, we used Poisson hidden Markov model to model an over dispersed daily data on Covid-19 deaths in Nigeria. The performance of the 2-state and 3-state Poisson hidden Markov models fitted to the data was compared with the help of the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Both criteria favoured the selection of the 3-state model. We therefore recommend the model for predicting Covid-19 deaths in Nigeria until further studies prove otherwise.

Keywords: Covid-19 pandemic, over dispersed data, multimodal data, serially correlated observations, Poisson hidden Markov model

INTRODUCTION

The COVID-19 pandemic which originated from Wuhan, China has posed a serious challenge to mankind. On 30 January 2020, World Health Organization (WHO) officially declared the COVID-19 epidemic as a public health emergency of international concern (Xiang *et al.* (2020) as cited in Ayinde *et al.*, 2020) and later on 11th March 2020 WHO reported it as a pandemic based on its alarming level of spread and severity over the world Bedford *et al.* (2020). According to World Health Organization (WHO), twenty percent of persons who contact COVID-19 becomes seriously ill and develop difficulty in breathing. Older people, and those with underlying medical problems like high blood pressure, heart and lung problems, diabetes, or cancer, are at higher risk of developing serious illness.

COVID-19 is a major global health event of the twenty-first century as the pandemic leads to lock down in so many places. This affected economic and social activities adversely. For instance in Nigeria, markets, schools, religious and social gatherings, international and national movements and non-essential services were closed down completely. However, the essential services were exempted from the ban. This led to job losses and slowed down economic activities with Nigeria slipping into a recession in November 2020 after its gross domestic product contracted for the second consecutive quarter (Aljazeera online News ,2020 ; CBN, 2020). The health sector was not spared as the health facilities were stretched this is because those who were being treated for COVID-19 were quarantined from those being treated of other ailments and this led to so many deaths. Many of the health workers who treated COVID-19 patients contracted COVID-19. Some of these workers died in the process.

It is over one year since Nigeria recorded the first death from COVID-19. According to figures from https://www.worldometers.info/coronavirus/#countries, as at March, 18, 2021 COVID-19 is reported in 221 countries with reported total cases of COVID-19 as 121,943,902 and total deaths globally as 2,694,960, total cases of COVID-19 in USA was 30,295,501, total death in USA was 550,671, total cases of COVID-19 in Europe 36,876,356, total death in Europe 863,280, total cases of COVID-19 in Asia 26,510,809, total death due to COVID-19 in Asia 413,140, total cases in Africa 4,096,309 death due to COVID-19 in Africa 108,903, total cases of COVID-19 in Nigeria 161,261 and deaths due to COVID-19 in Nigeria 2027. Despite the significant number of people who died of COVID-19 disease in Nigeria, not much has been done to model deaths due to the disease. Therefore, this study models COVID-19 deaths in Nigeria using daily data from Nigeria Centre for Disease Control (NCDC). Modelling of COVID-19 deaths provides rich insight into changes in mortality trends that are hidden in the population. The model can also assist in decision-making by making projections regarding important issues such as intervention-induced changes in the spread of a disease.

2. Overview of Hidden Markov Model

In hidden Markov models (HMMs), we assume that the parameter process is serially dependent so as to allow for serial dependence in the observations. Specifically, the parameter process is considered to be a Markov chain (George and Thomas, 2018). The state of an underlying system at any time t is unobservable in HMMs while the observations are essentially the outputs of another stochastic process influenced by a hidden process (George and Thomas, 2018).

Suppose $\{C_t, t=1,2,...\}$ is the unobserved parameter process with Markov property and $\{Z_t, t=1,2,...\}$ is the state-dependent process such that if C_t is known, the distribution of Z_t depends only on current state C_t. Let $Z^{(t)}$ and $C^{(t)}$ represent histories from time 1 to t. Then the hidden Markov model is a dependent mixture such that

$$P\left(C_t \middle| C^{(t-1)}\right) = P\left(C_t \middle| C_{(t-1)}\right), t = 2,3,4,...$$

$$P\left(Z_t \middle| \mathbf{Z}^{(t-1)}, C^{(t)}\right) = P\left(Z_t \middle| C_t\right), t \in \mathbb{N}.$$

Given discrete observations, the probability mass function of Z_t when the Markov chain is in state j becomes

$$P_i(z) = P\left(Z_t = z \middle| C_t = i\right).$$

With $w_i(t) = P\left(C_t = i\right), t = 1,2,...,$

$$P(Z_t = z) = \sum_{i=1}^{m} P\left(C_t = i\right) P\left(Z_t \middle| C_t = i\right) = \sum_{i=1}^{m} w_i(t) P_i\left(z\right)$$

$$= \left(w_1(t), w_2(t), \ldots, w_m(t)\right) \begin{pmatrix} p_1(z) & 0 & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & p_m(z) \end{pmatrix} \begin{pmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix} = \mathbf{w}(t)\mathbf{P}(z)\mathbf{1}',$$

where $\mathbf{P}(z)$ is the diagonal matrix with jth diagonal element $p_i(z)$, $\mathbf{w}(t) = \left(w_1(t), w_2(t), \ldots, w_m(t)\right)$ represents the initial distribution of the Markov chain. Let $\boldsymbol{\Gamma} = (\gamma_{ij})$ be the transition matrix associated with the Markov chain. If the Markov chain is stationary with stationary distribution $\boldsymbol{\delta}$, $\boldsymbol{\delta}\boldsymbol{\Gamma}^{t-1} = \boldsymbol{\delta}$ for all $t \in \mathbb{N}$. This implies that

$$\mathbf{P}(Z_t = z) = \boldsymbol{\delta}\mathbf{P}(z)\mathbf{1}'.$$

For a stationary m-state Poisson-HMM $\{Z_t, t = 1, 2, \ldots\}$, let $\boldsymbol{\Gamma}$, $\boldsymbol{\lambda} = \left(\lambda_1, \lambda_2, \ldots, \lambda_m\right)$, $\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{\lambda})$ and $\boldsymbol{\delta} = \left(\delta_1, \delta_2, \ldots, \delta_m\right)$. Here, $\boldsymbol{\Gamma}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\delta}$ are the transition probability matrix, state-dependent means and stationary distribution of the Markov chain respectively. First and second moments of Z_t are available in George and Thomas (2018).

Let $z_1, z_2, \ldots, z_n$ denote the sequence of observations generated by a hidden Markov model and $\boldsymbol{\delta}$ be the initial distribution assumed to be the same as the stationary distribution. In order to estimate the parameters of the HMM, we consider the numerical optimization of the likelihood function (Zucchini and MacDonald, 2009).

$$L = \boldsymbol{\delta}\boldsymbol{\Gamma}\mathbf{P}(z_1)\boldsymbol{\delta}\boldsymbol{\Gamma}\mathbf{P}(z_2)\ldots\boldsymbol{\delta}\boldsymbol{\Gamma}\mathbf{P}(z_n)\mathbf{1}'.$$

3. RESULTS

We begin this section by establishing the need to fit a dependent mixture model to the data on daily reported deaths due to Covid-19 infection in Nigeria. First, the Poisson dispersion and Hartigan's Dip tests performed on the data yielded the p values 0.0000 and 1.184×10^{-6} respectively. Thus, the series is over dispersed and multimodal. Evidence of serial correlation in the series can obtained via the inspection of the sample correlogram.

Next, we fit stationary Poisson hidden Markov models to the data. The estimates $\boldsymbol{\Gamma}_2$ and $\boldsymbol{\Gamma}_3$ of the transition probability matrices corresponding to the two-state and three-state Markov chains are respectively, given as

$$\boldsymbol{\Gamma}_2 = \begin{pmatrix} 0.9122 & 0.0878 \\ 0.0965 & 0.9035 \end{pmatrix} \text{ and } \boldsymbol{\Gamma}_3 = \begin{pmatrix} 0.9618 & 0.0358 & 6.0771 \times 10^{-15} \\ 0.0027 & 0.7102 & 0.2627 \\ 1.5135 \times 10^{-96} & 0.5602 & 0.4398 \end{pmatrix}.$$

The corresponding estimates of $\boldsymbol{\delta}$ and $\boldsymbol{\lambda}$ as well as the AIC and BIC values are presented in Table 1.

Table 1: Poisson-hidden Markov model fitted to the reported deaths due to Covid-19 in Nigeria.

Model	i	δ_i	λ_i	-log L	AIC	BIC
2-state P-HMM	1	0.5237	1.8719	-962.4071	1932.8142	1948.1884
	2	0.4763	10.1592			
3-state P-HMM	1	0.3405	1.3054	-882.2275	1776.4550	1799.5163
	2	0.4490	5.6898			
	3	0.2105	14.6901			

Based on minimum AIC and BIC values, the 3-state Poisson hidden model outperforms the other model.

CONCLUSION

We have studied the properties of the time series data on daily deaths due to COVID-19 in Nigeria with a view to providing an appropriate model for the data. In particular, the series is overdispersed and multimodal. The observations are also serially correlated, as can be observed from the sample correlogram. The 2-state and 3-state Poisson hidden Markov models were fitted to the data and their performance was compared using the selection criteria. On the basis of the minimum AIC and BIC values, the 3-state model is a better model for the series than the 2-state model.

REFERENCES

Ayinde, K, Lukman, A.F., Rauf, I.R., Alabi, O.O, Okon, E.C, Ayinde, O.E. (2020) "Modeling NigerianCovid-19 cases: A comparative analysis of models and estimators" Chaos, Solitons and Fractals138(2020)109911.https://doi.org/10.1016/j.chaos.2020.109911

Aljazeera News(2020) https://www.aljazeera.com/news/2020/11/21/nigeria-slips-into-recession-blamedon-covid-19-and-oil-prices Acessed 18th March,2021.

Bedford J, Enria D, Giesecke J, Heymann DL, Ihekweazu C, Kobinger G. (2020) COVID-19: towardscontrolling of a pandemic. Lancet North Am Ed 2020;395(10229):1015–18.

Central Bank of Nigeria (CBN) (2020) Communiqué No. 133 of the Monetary Policy Committee MeetingHeld On Monday 23rd and Tuesday 24th November 2020

Gecili E, Ziady, A, Szczesniak RD (2021) Forecasting COVID-19 confirmed cases, deaths and recoveries: Revisiting established time series modeling through novel applications for the USA and Italy. PLoS ONE 16(1): e0244173. https://doi.org/10.1371/journal.pone.0244173.

George, J and Thomas, S (2018) Modelling of Dengue Death Counts using Hidden Markov Model. Journalof Environmental Statistics, 8(9), 1-11.

NCDC (2020): http://covid19.ncdc.gov.ng. Accessed 10th March, 2021.

Lazaros K. Gallos, Nina H. Fefferman (2015) The Effect of Disease-Induced Mortality on Structural Network Properties PLoS One. ; 10(8): e0136704. Published online 2015 Aug 27. doi: 10.1371/journal.pone.0136704 PMCID: PMC4552173.

Worldometers: https://www.worldometers.info/coronavirus/#countries Access 18th March,2021

Xiang Y, Li W, Zhang Q, Jin Y, Rao W, Zeng L. Timely research papers about COVID-19 in China. Lancet North Am Ed 2020. doi:10.1016/S0140-6736(20)30375-5.

Zucchini, W and MacDonald, I. L (2009) Hidden Markov Models for Time Series: An introduction Using R, Monographs on Statistics and Applied Probability, 110.