Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://journal.nsps.org.ng/index.php/jnsps

J. Nig. Soc. Phys. Sci. 3 (2021) 224–233

Journal of the Nigerian Society of Physical Sciences

Bayesian Multilevel Models for Count Data

Olumide Sunday Adesina[a,*]

[a] *Department of Mathematical Sciences, Redeemer's University, Nigeria*

Abstract

The traditional Poisson regression model for fitting count data is considered inadequate to fit over- or under-dispersed count data and new models have been developed to make up for such inadequacies inherent in the model. In this study, Bayesian Multi-level model was proposed using the No-U-Turn Sampler (NUTS) sampler to sample from the posterior distribution. A simulation was carried out for both over-and under-dispersed data from discrete Weibull distribution. Pareto k diagnostics was implemented, and the result showed that under-dispersed and over-dispersed simulated data has all its k value to be less than 0.5, which indicate that all the observations are good. Also all WAIC were the same as LOO-IC except for Poisson in the over-dispersed simulated data. Real-life data set from National Health Insurance Scheme (NHIS) was used for further analysis. Seven multi-level models were fitted and the Geometric model outperformed other model.

DOI:10.46481/jnsps.2021.168

Keywords: Count data, Health, Insurance, Dispersion, Multilevel models

Article History :
Received: 17 February 2021
Received in revised form: 09 July 2021
Accepted for publication: 30 July 2021
Published: 29 August 2021

Communicated by: T. Latunde

1. Introduction

Count data contains non-negative integers and zero obtained within a fixed period. Various studies have been carried out on count data and modelling, [1] applied it to medical data [2] applied it to model microbiome data and many others. Frequentist and Bayesian estimation have equally been used to model count data, and the widely used is the Bayesian estimation technique. There are three major types of count data, the over-dispersed, under-dispersed and over-dispersed. More about types of count data can be found in the study by [3,4]. There are models dedicated to modelling under-dispersion due to their suitability while a model such as Negative Binomial is suitable for fitting over-dispersion. Models such as Dirichlet Process Prior, Negative Weibull can be used to fit both under-dispersion and over-dispersion. Models such as zero inflated and hurdle models can effectively handle over and under-dispersed count data with many zeros.

The zero-truncated regression models are specifically designed to fit count data with no zero count. The categorized regression model is designed to fit count data that its response variable is categorized. Some of the improved techniques relative to Poisson regression model can be found in [5], [6], [3], amongst others. [7] carried out a on hidden markov model in multiple testing on dependent count data, [8] showed that the exponentiated-exponential Geometric distribution can be applied to fit under-dispersed or over-dispersed count data, in the same manner [4] demonstrated that Dirichlet Process Mixture Prior of Generalized Linear Mixed Models (DPMglmm) can fit either over-dispersed or under-dispersed count data well.

[9] sufficiently showed that multi-level zero-inflated Poisson (ZIP) regression model can adequately fit both over-and under-dispersed count data that have zero counts. The authors adopted

*Corresponding author tel. no:
Email address: olumidestats@gmail.com (Olumide Sunday Adesina)

EM algorithm along with the penalized likelihood and restricted maximum likelihood (REML). [10] adopted multilevel Zero-inflated Poisson (ZIP) regression and zero-inflated negative binomial (ZINB) and applied the models to fit count data relating to decay, missing and filled teeth of children aged 12 years old. A related and recent study was carried out by [11], the authors proposed multilevel zero-inflated generalized Poisson (ZIGP) that is suitable in fitting both over- and under-dispersed count data and compared with multilevel models of zero-inflated Poisson, zero-inflated negative binomial. The result showed that the multilevel ZIGP produced more accurate parameter estimates, particularly for under-dispersed data.
In this study, Bayesian multilevel modelling was proposed and implemented for some basic distributions used in fitting count data (Poisson, Negative Binomial and Geometric), zero inflated and hurdle models, and identify a most suitable model for fitting under-and over-dispersed respectively. The remaining part of this paper is sectionalized as follows; multi-level modelling is described in section 2, parameter estimation and model selection can be found in 3. Section 4 is the results of simulation study and that of real-life data. Lastly, summary and conclusion in section 5.

2. Model Description

2.1. Multilevel Modelling

The multilevel modelling technique follows a similar process involved when fitting the Generalized Linear Model. In GLM, a link function links the response y variable to the predictor(s), same with multilevel modelling. Let (f) be an inverse link function that links the response variable y to the predictor, then Υ is the linear combination of the predictors transformed by the inverse link function (f), and d be a parametric model (distribution), the model can be simply be written as

$$y_i \sim d\left(f\left(\Upsilon_i\right), \theta\right) \tag{1}$$

And linear predictor:

$$\Upsilon = \boldsymbol{X}\beta + \mathbf{K}\epsilon \tag{2}$$

The data is made up of the response variable y, $\boldsymbol{X}$ and $\mathbf{K}$, while β, ϵ and θ are the model parameters to be estimated. while β is the fixed effect coefficient at population level, while ϵ is the coefficient at group-level. The Bayesian estimation technique by Monte Carlo Markov Chain (MCMC) procedure considers ϵ as a parameter relative to maximum likelihood that considers ϵ as error term [12].

2.2. Zero-Inflated distribution

If Yfollows zero-inflated Poisson (ZIP) distributions, given by

$$P(Y=y) = \left\{ \begin{array}{ll} \omega + (1-\omega)\exp(-\lambda), & y=0 \\ (1-\omega)\exp(-\lambda)\lambda^y/y!, & y>0 \end{array} \right\} \tag{3}$$

Where is ω in the range$0 < \omega < 1$, in order to accommodate more zeros than those allowed under the Poisson assumption($\omega =$ 0), and the case of $\omega < 0$imply zero inflated. [9] estimated multi-level parameters of ZIP regression in the generalized linear mixed models (GLMMs) context. The authors generalized the ZIP model so that the model will be able to withstand more complex correlation structure.
Zero-inflated negative binomial for counts is formed from ZIP, the mean and variance defined as:

$$E(Y) = (1-\omega)\lambda = \mu, \tag{4}$$

The use of regression models based on ZIP was established by [13-15], [5]. Following [5] we have:
$\log(\lambda) = X\beta$ and

$$\log\left(\omega/1-\omega\right) = Z\Upsilon \tag{5}$$

where X and Z are matrices of covariates, β and $\Upsilon$are vector of parameters. Assuming two linear predictors are related in some ways, [16] provided a simplest form of (3) which is refers to the $ZIP(\tau)$model as follows:

$$\log(\lambda) = X\beta, \log\left(\omega/1-\omega\right) = \tau X\beta \tag{6}$$

Where τ is a scalar parameter, which implies that$\omega = (1 + \lambda^{-r})^{-1}$.
Following equation (3) in multi-level case, [9] identified the extension of ZIP model to include random components w_i and u_i within logistic and Poisson linear predictors to take care of dependence of observations in given clusters. The random effects w_i and $u_i$are specific to the i^{th} cluster. In a three-level hierarchical situation of Y_{ijk}, the k^{th} observation of the j^{th} individual within the i^{th} clusters is measured through random effects associated with the linear predictors as follows:

$$\begin{array}{l} \log\left[\frac{\phi_{ijk}}{(1-\phi_{ijk})}\right] = \xi_{ijk} = a_{ijk}^T\alpha + w_i + s_{ij} \\ \log(\lambda_{ijk}) = \gamma_{ijk} = x_{ijk}^T\beta + u_i + v_{ij} \end{array} \tag{7}$$

The covariates $a_{ijk}^T$and x_{ijk}^T are not always the same α and β are the corresponding vectors of regression coefficients. $s_{ij}$and $v_{ij}$are variations at subject level.

2.3. Hurdle Models

If the distribution of Yfollows zero-truncated Poisson distribution it follows that:

$$\pi_0 y > 0$$

$$P(Y=y) = \frac{(1-\pi_0)e^{-\lambda}\lambda^y}{(1-e^{-\lambda})y!} y = 0 \tag{8}$$

Reparametizing the zero-inflated Poisson model in equation (3) with$\pi_0 = \omega + (1-\omega)e^{-\lambda}$, [16], gave Poisson hurdle regression model is given as;

$$\log\left(\lambda\right) = X\beta,$$

$$\log[-\log(1-\pi_+) = \tau X\beta \tag{9}$$

Where $\pi_+ = 1 - \pi_0$ is the probability of clearing the "hurdle" and generating a non-zero count.

2.4. Prior Distributions

Prior distribution is specified at population and group-level. At population-level parameters have an improper prior [17]. At group level it is assumed that parameters ε comes from a multivariate normal distribution having zero mean and unknown covariance matrix$\mathbf{\Sigma}$.

$$epsilon \sim N(0, \mathbf{\Sigma}) \tag{10}$$

Covariances are between group-level parameters are generally of different groupings factors and assumed to be zero. By implication, $\mathbf{K}$and ε can be divided to form several matrices $\mathbf{K}_i$ and parameter vectors ε_i , where i indexes grouping factors, thus, the model can be simplified to

$$\epsilon_i \sim N(0, \mathbf{\Sigma_i}) \tag{11}$$

Sometimes, it can be assumed that group-level parameters for different levels of the same grouping factors are not dependent. If the other level is indexed by j, (11) leads to:

$$\epsilon_{ij} \sim N\left(0, \mathbf{M_j}\right) \tag{12}$$

The covariance matrices $\mathbf{M}_j$ will become the model parameters. No-U-Turn Sampler (NUTS) by (2014) [18] is used for $\mathbf{M}_j$ as instead of the Inverse-Wishart prior distribution used in most studies and packages. Inverse-Wishart distribution is used because it has good conjugacy characteristics for Gibbs-Sampler. The choice of Inverse-Wishart prior distribution was criticized in the studies by [19], and [20]. The parameters of $\mathbf{M}_j$ is selected in terms of correlation matrix $\mathbf{\Omega_j}$ and a vector of standard deviations σ_j through,

$$\mathbf{M}_j = \mathbf{D}\left(\sigma_j\right)\mathbf{\Omega_j}\mathbf{D}\left(\sigma_j\right) \tag{13}$$

and $\mathbf{D}\left(\sigma_j\right)$imply the diagonal matrix with diagonal elements σ_j. Then, prior would be specified for $\mathbf{D}\left(\sigma_j\right)\mathbf{\Omega_j}\mathbf{D}\left(\sigma_j\right)$. In the case of $\mathbf{\Omega_j}$, LKJ-Correlation prior by [21] is used, with $\zeta > 0$. That is, $\Omega_j \sim LKJ(\zeta)$

sectionParameter Estimation and Model Selection

Sampling from the posterior require appropriate sampling procedure, two basic sampling procedures are discussed here. First is the Hamiltonian Monte-Carlo (HMC) Sampler, also known as Hybrid Monte-Carlo [22-23]. [18] extended HMC to No-U-Turn Sampler (NUTS), because HMC has some drawbacks as discussed by [17]. The NUTS Sampler allows setting parameters, and eliminates the need for hand-tuning, [18] stated that setting the parameters automatically makes it least efficient as compared to a well-tuned Hamiltonian Monte-Carlo. Software package by R core team (2020) was used to fit the model with **brms** package by [17] along with **Stan** processor without which the analysis cannot be run, it can be assessed on http://cran.r-project.org/bin/windows/Rtools/.

The Watanabe-Akaike Information Criteria proposed (WAIC) by [24] and Leave-one-out cross validation LOO-CV by [25,26] were used for model selection in this study. The WAIC was used for estimating the out-of-sample expectation and considered an improvement upon the DIC, with WAIC, correction for effective number of parameters to adjust over-fitting is added. According to [27], WAIC can be computed in two possible ways, first is calculated using simulation $\theta^s, s = 1, ...S$ and given as

$$p_{WAIC_1} = 2\sum_{i=1}^{n}\left(\log\left(\frac{1}{S}\sum_{s=1}^{S}\log p(y_i|\theta^s)\right) - \frac{1}{S}\sum_{s=1}^{S}\log p(y_i|\theta^s)\right) \tag{14}$$

For the second WAIC computation approach, the variance of individual terms in the log predictive density is added up over the $n$data points and express as follows:

$$p_{WAIC_2} = \sum_{i=1}^{n} var_{post}\left(\log p(y_i|\theta)\right) \tag{15}$$

The advantages of WAIC over AIC and DIC was adequately discussed by [27] In the case of Leave-one-out cross-validation (LOO-CV) in Bayesian analysis, the data are repeatedly subdivided into a training set y_{train} and a holdout set $y_{holdout}$ with the objective of fitting y_{train} yielding a posterior distribution $p_{train}(\theta) = p_{train}(\theta|y_{train})$.

The Bayesian LOO-CV estimate of out-of-sample predictive fit is

$$lppd_{loo-cv} = \sum_{i=1}^{n}\log p_{post(-i)}(y_i) \tag{16}$$

and estimated as

$$\sum_{i=1}^{n}\log\left(\frac{1}{S}\sum_{s=1}^{S}\log p(y_i|\theta^{is})\right) \tag{17}$$

Lower WAICs and LOOs suggest better model fit.

2.5. Pareto-k-diagnostics

The shape parameter k of the generalized Pareto distribution can be used to assess the reliability and approximate convergence rate of the Pareto smoothed importance sampling (PSIS). It follows that if, $k < 0.5$(that is, 'good') then the central limit theorem holds. Similarly, If$0.5 \leq k < 1$, (that is, 'ok') then the variance of the raw importance ratios is infinite, but the mean exists. In the same manner, If $k > 0.7$(that is, 'bad'), unreasonable convergence rates is observed and unreliable Monte Carlo error estimates, and finally, if $k \geq 1$ (that is, 'very bad'), then neither the variance nor the mean of the raw importance ratios exists.

3. Result

3.1. Simulation Study

Simulation of over and under-dispersed count data was carried out and the response count variable was obtained from Discrete Weibull distribution. On simulating count data from Discrete Weibull (DW) distribution, [28] identified that the parameter β of DW should contain the range$0 \leq \beta \leq 1$; irrespective the value of parameter$q$. For under-dispersed count data, β should be specified such that$\beta \geq 2$, irrespective of the value of q. Analysis for simulation study was carried using software package by

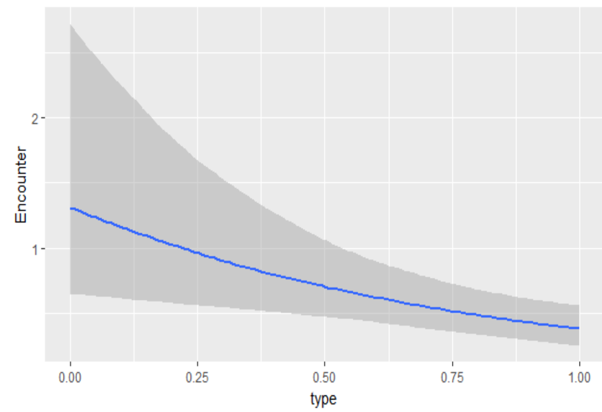

Figure 1. Marginal plot of relationship between Encounter and type

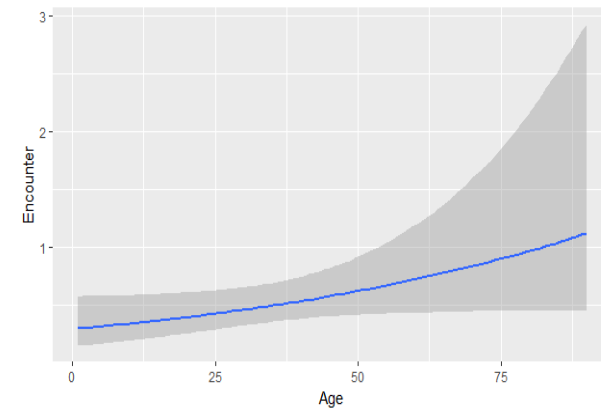

Figure 2. Marginal plot of relationship between Encounter and Age

[29], and R package "DWreg" by [30]. Random numbers consisting of 500 observations were generated and two predictors were uniformly generated in interval (0, 1) and (0, 2) for over- and under-dispersion respectively. For over dispersed, beta=0.8 and for under-dispersed beta=2.1.
The parameter of discrete Weibull follows that $\theta_0 = 0.25$, $\theta_1 = 0.35$, $\theta_2 = 0.5$ and the corresponding equation for logit link is as follows

$$\log\left(q\left(1-q\right)\right) = 0.25 + 0.35x_1 + 0.5x_2 \tag{18}$$

All Pareto k estimates are good (k < 0.5)

If p_waic estimates greater than 0.4. We recommend trying loo Instead.
Table 1 shows results for Under-dispersed simulated count data. The best two performed model are Geometric and Hurdle Poisson distribution indicated with ** and * respectively with lowest WAIC and LOO, following implementation using brms, a package for Bayesian multilevel modelling in R. from Table 1, it shows that WAIC=LOO for all the models.
Table 2 shows results for Over-dispersed simulated count data. The best two performed model are negative Binomial and zero inflated negative binomial distribution indicated with ** and * respectively with lowest WAIC and LOO, following implementation using **brms**, a package for Bayesian multilevel modelling in R. All the model shows that WAIC=LOO expect for Poisson model.

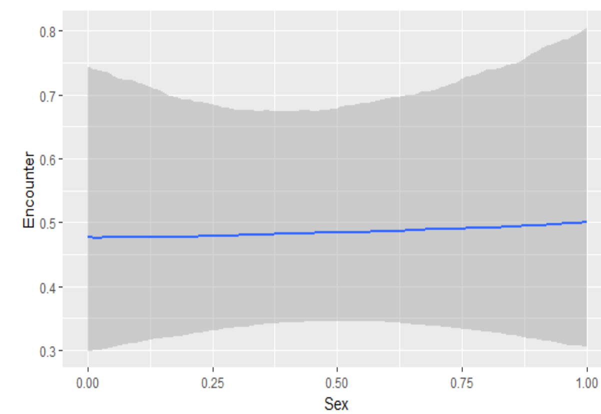

Figure 3. Marginal plot of relationship between Encounter and Sex

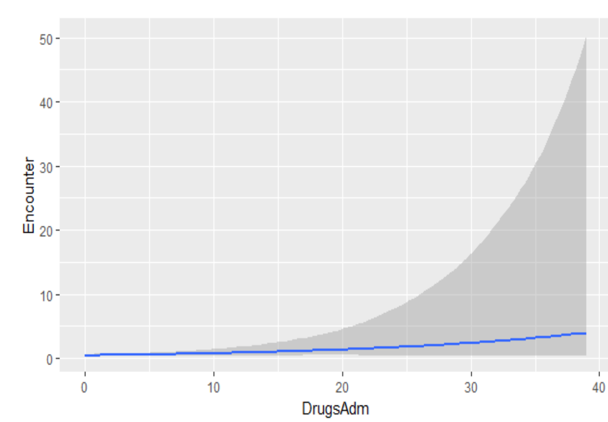

Figure 4. Marginal plot of relationship between Encounter and DrugsAdm

3.2. 4.2 Application to Health Insurance Data

3.2.1. Data Description

The data set was obtained from National Health Insurance Scheme (NHIS), and it contains excess zero count. Sample of 116 users of NHIS users was obtained from September 2016 to July 2017. Response variable is number of encounter (Encounter), out of 116 observed, eighty two (82) persons made claims. Encounter is the time a user of National Health Insurance Scheme (NHIS) visits the health facility, and possibly makes claims. The predictors include type of Encounter (type), which is either primary or (secondary= 0, primary= 1) primary are users who registered primarily to use the health facility, while secondary are users who were referred from another health facility to that of State Hospital Ota, because of availability of specialists. Another predictor Sex (male=1 and female=0), Biological age of patients (Age). Number of drugs administered (DrugsAdm) that is, both oral and injection.

Drugs-out-of-stock (DrugsOS) is another predictor. The

Table 1. Simulated Under-Dispersed count data from Discrete Weibull Model

Model	elpd_waic	p_waic	waic	elpd_loo	p_loo	looic	Waic=looic
Poisson							
Est.	-628.1	1.9	1256.2	-628.1	1.9	1256.2	Yes
SE	10.2	0.1	20.3	10.2	0.1	20.3	
Negbin							
Est.	-629.3	1.9	1258.5	-629.3	1.9	1258.5	Yes
SE	10.1	0.1	20.2	10.1	0.1	20.2	
Geometric							
Est.	-717.3	0.9	1434.6**	-717.3	0.9	1434.6**	Yes
SE	12.9	0.1	25.8	12.9	0.1	25.8	
hurdle_poisson							
Est.	-620.5	3.3	1241.0*	-620.5	3.3	1241.0*	Yes
SE	13.1	0.2	26.1	13.1	0.2	26.1	
hurdle_negbin							
Est.	621.2	3.3	1242.5	621.2	3.3	1242.5	Yes
SE	13.1	0.2	26.2	13.1	0.2	26.2	
zero_inflated_poisson							
Est.	-629.1	2.0	1258.3	-629.1	2.0	1258.3	Yes
SE	10.1	0.1	20.2	10.1	0.1	20.2	
zero_inflated_negbin							
Est.	630.2	1.9	1260.5	630.2	1.9	1260.5	Yes
SE	10.1	0.1	20.2	10.1	0.1	20.2	

Table 2. Simulated Over-Dispersed count data from Discrete Weibull Model

Model	elpd_waic	p_waic	waic	elpd_loo	p_loo	looic	Waic=LOO
Poisson							
Est.	-1984.6	21.6	3969.2	-1984.7	21.7	3969.4	No
SE	108.6	3.1	217.2	108.6	3.1	217.2	
Negbin							
Est.	-1217.4	3.9	2434.8**	-1217.4	3.9	2434.8**	Yes
SE	27.9	0.5	55.8	27.9	0.5	55.8	
Geometric							
Est.	-1228.7	4.2	2457.4	-1228.7	4.2	2457.4	Yes
SE	27.9	0.6	55.8	27.9	0.6	55.8	
hurdle_poisson							
Est.	-1682.3	19.0	3364.6	-1682.3	19.0	3364.6	Yes
SE	87.6	2.9	175.1	87.6	2.9	175.1	
hurdle_negbin							
Est.	-1229.7	4.5	2459.4	-1229.7	4.5	2459.4	Yes
SE	27.7	0.4	55.5	27.7	0.4	55.5	
zero_inflated_poisson							
Est.	-1680.7	18.7	3361.3	-1680.7	18.7	3361.4	Yes
SE	87.6	2.9	175.2	87.6	2.9	175.2	
zero_inflated_negbin							
Est.	-1218.2	4.0	2436.4*	-1218.2	4.1	2436.4*	Yes
SE	27.9	0.5	55.9	27.9	0.5	55.9	

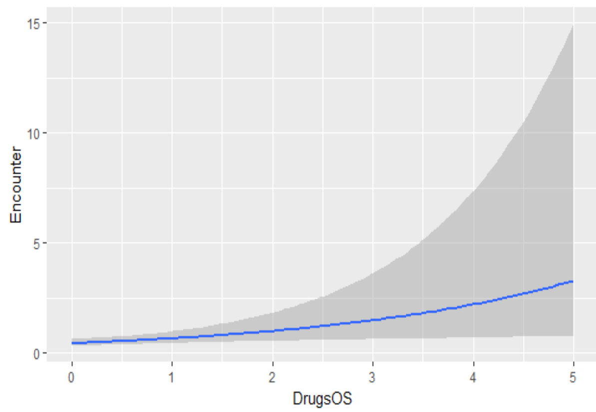

Figure 5. Marginal plot of relationship between Encounter and DrugOS

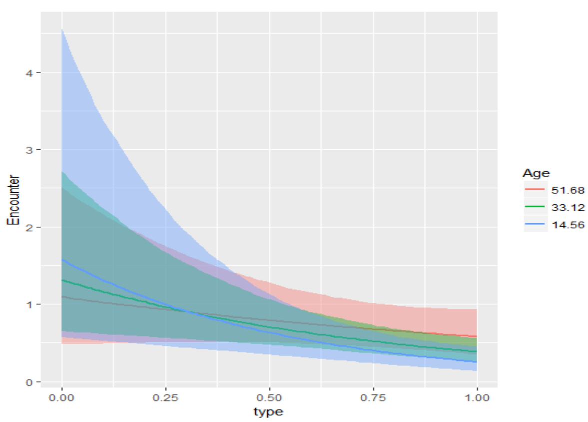

Figure 6. Marginal plot of relationship between Encounter and Type along with Age

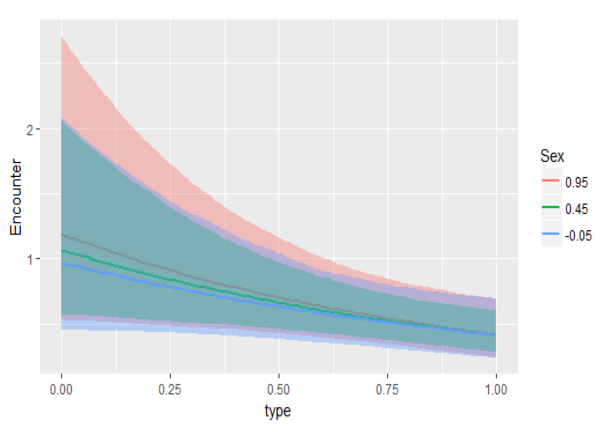

Figure 7. Marginal plot of relationship between Encounter and Type along Sex

idea of National Health Insurance Scheme (NHIS) is that treatments and drugs administered attract 10 percent of the total cost. It becomes disadvantage to patients having cases of DrugsOS. Out of 116 users, 97 are primary, while 19 are secondary users, 70 out of 97 (72.16%) primary users did not make claims within the period observed. 4 out of 19 secondary users did not make claims (21.05%). Out of 116 NHIS users of the health facility, 64 are females and 52 are males. 41 females had zero claims (64.06%), while 33 males had zero claims (63.46%). 11 females and 8 males are secondary users, while 53 females and 44 males are primary users. The data can be obtained https://data.mendeley.com/drafts/6hcf5mf7fy.

Mean of the response variable (Encounter) is 0.7068, while variance is 1.7916 which potentially suggests over-dispersion. Further test on the data shows that the dispersion parameter is δ=1.49799, indicating that the data set is over-dispersed with δ>1.The dispersion test was performed using AER package in R by [31].

From Table 3 shows results for the real-life Over-dispersed count data. The best two performed model are Geometric and Negative binomial distribution indicated with ** and * respectively with lowest WAIC and LOO, following implementation using brms, a package for Bayesian multilevel modelling in R, In all the models Waic LOO as shown in Table 6. As earlier stated; for any observation for $k > 0.7$indicate unreasonable convergence rates is observed and unreliable Monte Carlo error estimates. Table 4 shows that hurdle negative binomial (4) has the highest of such observations, Poisson, negative binomial, and hurdle Poisson model has 3 of such observations while Geometric, zero_inflated_poisson and zero_inflated_negbin has two (2) each.

Each parameter is presented in Table 5 with the posterior mean as the 'Estimate' and the 'Est.Error' as the standard deviation of the posterior distribution, the Table also contain a two-sided 95% Credible intervals (l-95% CI and u-95% CI) established on quantiles. From Table 5, the 'intercept', 'type' and 'type:Sex"-interaction has a negative posterior mean. For "type", the model predicts longer periods for encounter for secondary users than primary users; 'Sex' (0.19) accounts for more Encounter than 'Age' (0.01). "drugsOS (0.41) tells us that there significant cases of drugs-out-stock which suggest ineffectiveness of Nigeria Health Insurance Scheme.

Drawing samples from (NUTS) follows that for each parameter, Efficient Sample is a real measure of effective sample size, while Rhat is the would-be scale reduction factor on split chains (at convergence, Rhat = 1).

Figure 1 shows that Encounter has positive relation with type; the figure suggests that the effect of Encounter on type is higher for secondary user of the facility since it is higher on zero end than 1.

Figure 2 shows that Encounter has positive relation with Age; the figure suggests that as age increases Encounter increases

Figure 3 shows that Encounter has positive relation with Sex; the figure suggests that male account for more Encounter than female.

Figure 4 shows that Encounter has positive relation with type; Number of DrugsAdm increases with number of Encounter

Figure 5 shows that Encounter has positive relation with type; Number of DrugsOS increases significantly with number of Encounter

Figure 6 shows that as primary users of the facility increases, the number of Encounter increases across ages

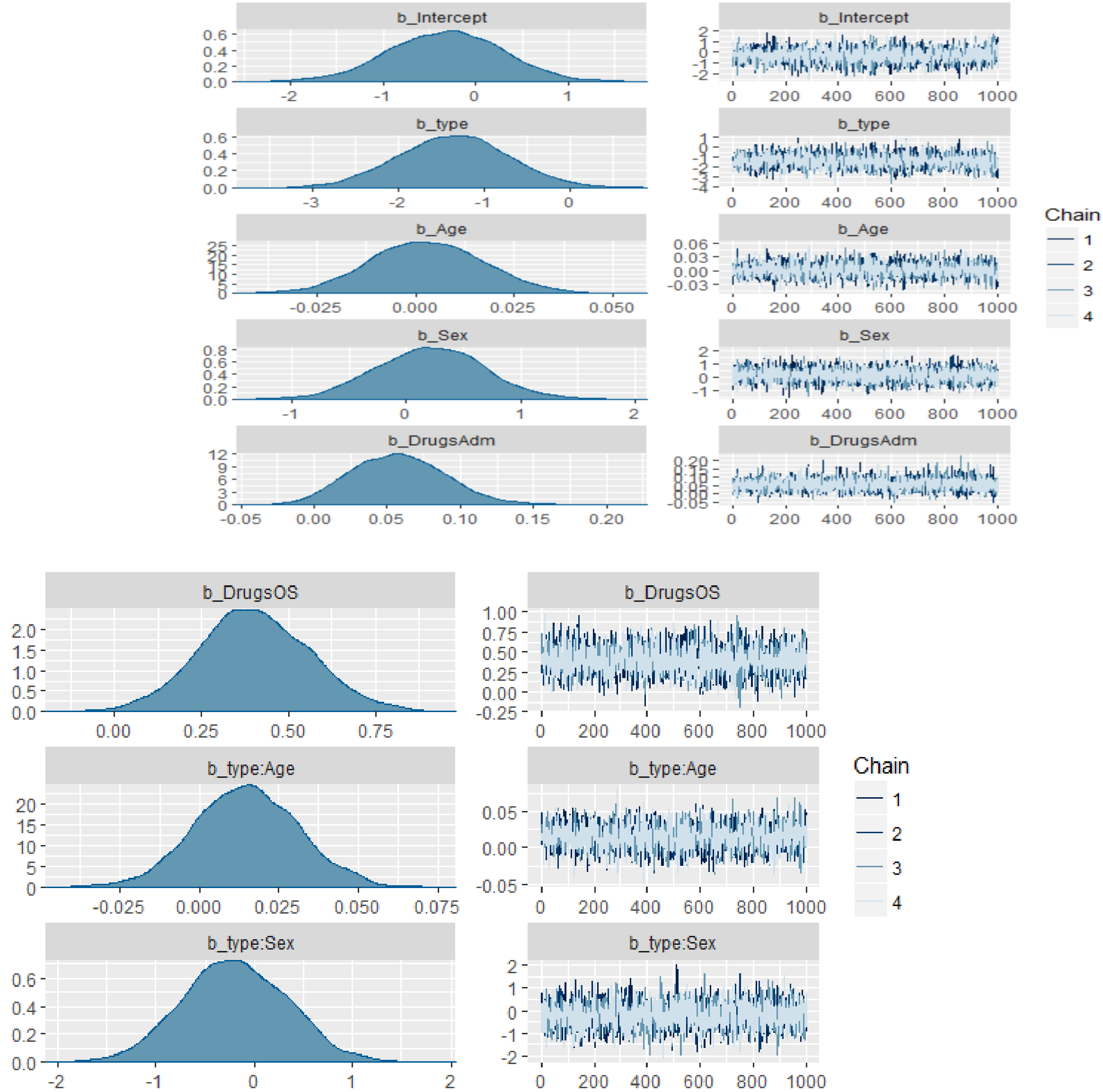

Figure 8. Density and Trace plots for all covariates

Figure 7 shows that as primary users of the facility increases, the number of Encounter increases across Sexes

Figure 8 shows that the density plots for the tail area of the distribution, which corresponds to l-95% CI and u-95% CI in Table 5 and trace plot, the Trace plot shows that the chains are stable.

4. Summary and Conclusion

In this study Bayesian multi-level model was proposed and implemented. The No-U-Turn Sampler (NUTS) sampler was used to sample from the posterior distribution, and implementations were done using the 'package brms' in R. Simulation study was carried out for both over-and under-dispersed and response variables were generated through discrete Weibull distribution while the predictors generated from Uniform distribution. Results from under-dispersed revealed that Geometric distribution is the most appropriate model to fit count data using multi-level approach. While for over-dispersed simulated data, negative binomial shows to outperform the Poisson, Geometric, hurdle-Poisson, hurdle-negbin, zero-inflated-Poisson, zero-inflated-negbin. Pareto K diagnostics shows that for under-dispersed and over-dispersed simulated data all k are less than 0.5, which makes all the observations to be good, also all WAIC were the same as LOO-IC except for Poisson in the over-dispersed simulated data.

Real-life data set of count of Encounter of patients from National Health Insurance Scheme was used for further analysis. The model that performs best was the Geometric distribution followed by negative binomial model. Contrary to the simulated data not all WAIC were the same as LOO-IC, except for Poisson model in the over-dispersed simulated data. The need to carry out LOO-IC was informed by having observa-

Table 3. Real Life Data

Model	elpd_waic	p_waic	waic	elpd_loo	p_loo	looic	Waic=loo
Poisson							
Est.	-127.4	15.9	254.8	-129.7	18.2	259.3	No
SE	15.3	6.3	30.6	15.9	7.0	31.8	
Negbin							
Est.	-122.3	10.6	244.6*	-124.0	12.3	248.0*	No
SE	12.8	3.5	25.6	13.5	4.3	27.0	
Geometric							
Est	-120.8	7.4	241.6**	-121.5	8.1	242.9**	No
SE.	12.3	2.3	24.6	12.5	2.7	25.0	
hurdle_poisson							
Est.	-132.6	11.4	265.1	-135.0	13.8	269.9	No
SE	13.3	4.2	26.6	14.2	5.4	28.5	
hurdle_negbin							
Est.	-131.6	5.0	263.3	-133.1	6.4	266.2	No
SE	12.7	1.3	25.5	13.3	2.3	26.5	
zero_inflated_poisson							
Est.	-126.7	12.8	253.3	-127.8	13.9	255.6	No
SE	14.8	5.1	29.6	15.1	5.4	30.1	
zero_inflated_negbin							
Est.	-122.7	9.4	245.3	-123.4	10.2	246.8	No
SE	12.9	3.3	25.9	13.2	3.7	26.4	

Table 4. Pareto *k* diagnostics

Model	Pareto k diag.	remark	count	Pct	Min_neff	obs. k>0.7
Poisson	(-Inf, 0.5]	Good	112	96.6%	1367	3
	(0.5, 0.7]	Ok	1	0.9	236	
	(0.7, 1]	Bad	2	1.7%	17	
	(1, Inf)	Very bad	1	0.9%	4	
Negbin	(-Inf, 0.5]	Good	112	96.6%	1068	3
	(0.5, 0.7]	Ok	1	0.9%	446	
	(0.7, 1]	Bad	3	2.6%	19	
	(1, Inf)	Very bad	0	0.0%	-	
Geometric	(-Inf, 0.5]	Good	114	98.3%	1036	2
	(0.5, 0.7]	Ok	0	0.0%	-	
	(0.7, 1]	Bad	2	1.7%	88	
	(1, Inf)	Very bad	0	0.0%	-	
hurdle_poisson	(-Inf, 0.5]	Good	108	93.1%	859	3
	(0.5, 0.7]	Ok	5	4.3%	413	
	(0.7, 1]	Bad	1	0.9%	16	
	(1, Inf)	Very bad	2	1.7%	7	
hurdle_negbin	(-Inf, 0.5]	Good	109	94.0%	2909	4
	(0.5, 0.7]	Ok	3	2.6%	594	
	(0.7, 1]	Bad	3	2.6%	135	
	(1, Inf)	Very bad	1	0.9%	9	
zero_inflated_poisson	(-Inf, 0.5]	Good	111	95.7%	1885	2
	(0.5, 0.7]	Ok	3	2.6%	165	
	(0.7, 1]	Bad	1	0.9%	26	
	(1, Inf)	Very bad	1	0.9%	10	
zero_inflated_negbin	(-Inf, 0.5]	Good	113	97.4%	1509	2
	(0.5, 0.7]	Ok	1	0.9%	213	
	(0.7, 1]	Bad	2	1.7%	50	
	(1, Inf)	Very bad	0	0.0%	-	

Table 5. Population-Level Effects

	Estimate	Est. Error	l-95% CI	u-95% CI	Eff.Sample
Intercept	-0.36	0.63	-1.63	0.83	2290
type	-1.33	0.65	-2.57	-0.06	1972
Age	0.01	0.01	-0.02	0.03	2024
Sex	0.19	0.48	-0.73	1.13	2839
DrugsAdm	0.06	0.03	-0.00	0.13	3444
DrugsOS	0.41	0.17	0.10	0.74	2609
type:Age	0.01	0.02	-0.02	0.05	1818
type:Sex	-0.18	0.56	-1.26	0.90	2897

tions for all the cases. Figures 1 to Figure 7 contains marginal plots to identify the relationship between the response variable (Encounter) and covariates; type, Sex, DrugsAdm, Age, and DrugsOS.
As identified by [8] that Geometric family have the ability to model count data effectively, the same has been demonstrated in this study using Bayesian multi-level regression approach. Future direction can consider fitting multi-level regression model to fit count data using distribution such as the Weibull-exponential distribution and Exponentiated Generalized Weibull proposed by [32] and [33] respectively.

Acknowledgements

My appreciation goes to the Medical Director of the State Hospital, Ota, Ogun state where the data set was obtained. I also appreciate the inputs of reviewers and editors on the manuscript.

References

[1] M. S. Workie & A. M. Lakew, "Bayesian count regression analysis for determinants of antenatal care service visits among pregnant women in Amhara regional state, Ethiopia" Journal of Big Data **5** (2018) https://doi.org/10.1186/s40537-018-0117-8.
[2] K. H. Lee, B. A Coull, A. B. Moscicki, B.J. Paster & J. R. Starr, "Bayesian Variable Selection for Multivariate Zero-Inflated Models: Application to Microbiome Count Data", Biostatistics **21** (2020) 499. **1**.
[3] F. Famoye & K.P. Singh, "Zero-inflated generalized Poisson regression model with applications to domestic violence data", Journal of Data Science **4** (2006) 117.
[4] O. S. Adesina, T. O. Olatayo, O. O. Agboola, & P. E. Oguntunde, "Bayesian Dirichlet process mixture prior for count data", International Journal of Mechanical Engineering and Technology **9** (2018) 630.
[5] D. Lambert, "Zero-inflated Poisson regression, with an application to defects in manufacturing" Technometrics **34** (1992) .
[6] F. Famoye & W. Wang, "Censored generalized Poisson regression model" Journal of Computational Statistics and Data Analysis **46** (2004) 547.
[7] W. Su & X. Wang, "Hidden Markov model in multiple testing on dependent count data", Journal of Statistical Computation and Simulation **5** (2020) 889.
[8] F. Famoye & L. Carl "Exponentiated-exponential geometric regression model", Journal of Applied Statistics **44** (2017) 2963.
[9] A. H. Lee, K. Wang, J. A Scott, K. W Yau & G. J. McLachlan, "Multilevel zero-inflated Poisson regression modelling of correlated count data with excess zeros", Statistical Methods in Medical Research **15** (2006) 47.
[10] A. Moghimbeigi, M.R. Eshraghian, K. Mohammad & B. Mcardle "Multilevel zero-inflated negative binomial regression modeling for over-dispersed count data with extra zeros", Journal of Applied Statistics **35** (2008) 1193.
[11] A. Almasi, M.R. Eshraghian, A. Moghimbeigic, A. Rahimib, K. Mohammad & S. Fallahigilan, "Multilevel zero-inflated generalized Poisson regression modelling for dispersed correlated count data", Statistical Methodology **30** (2016) 1.
[12] J. Fox & S. Weisberg, *An R companion to applied regression*, Second Edition, Sage (2010), ISBN-13: 978-1412975148
[13] J. Mullahy "Specification and testing of some modified count data models", Journal of Econometrics **33** (1986) 341.
[14] D. Heilbron, "Generalized linear models for altered zero probabilities and overdispersion in count data" SIMS Technical Report 9, Department of Epidemiology and Biostatistics, University of California, San Francisco **9** (1989).
[15] D. Heilbron "Zero-altered and other regression models for count data with added zeros", Biometrical Journal **36** (1994) 531.
[16] M. Ridout, C. G. B. Demetrio & J. Hinde "Models for count data with many zeros", International Biometric Conference Cape Town (1998) **1**.
[17] P. C. Burkner. "brms: An R package for Bayesian multilevel models using Sta", Journal of Statistical Software (2017).
[18] M. D. Hoffman & A. Gelman "The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo", The Journal of Machine Learning Research **15** (2014) 1593.
[19] R. Natarajan & R. E. Kass, "Reference Bayesian methods for generalized linear mixed models", Journal of the American Statistical Association **95** (2000) 227.
[20] R. E. Kass, R. Natarajan, "A default conjugate prior for variance components in generalized linear mixed models (comment on article by Browne and Draper)", Bayesian Analysis **1** (2006) 535.
[21] D. Lewandowski, D. Kurowicka & H. Joe, "Generating random correlation matrices based on vines and extended onion method", Journal of Multivariate Analysis **100** (2009) 1989.
[22] S. Duane, A. D. Kennedy, B. J. Pendleton & D. Roweth, "Hybrid Monte Carlo", Physics Letters B, **195** (1987) 216.
[23] R. M. Neal, *Handbook of Markov Chain Monte Carlo*, volume 2, chapter MCMC Using Hamiltonian Dynamics. CRC Press **2** (2011) **1**.
[24] S. Watanabe, "Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory", The Journal of Machine Learning Research **11** (2011) 3571.
[25] A. E. Gelfand, D. K. Dey, & C. H. Hang, "Model determination using predictive distributions with implementation via sampling-based methods." Technical report, DTIC Document (1992)
[26] A. Vehtari', A. Gelman & J. Gabry Efficient Implementation of Leave-One-Out Cross-Validation and WAIC for Evaluating Fitted Bayesian Models", Unpublished manuscript, (2015).
[27] A. Gelman, J. B. Carlin, H. S. Stern & D. B. Rubin, *Bayesian Data Analysis*, Taylor & Francis **2** (2014).
[28] H. S. Klakattawi, V. Vinciotti & K. Yu, "A simple and adaptive dispersion regression model for count data", Entropy **20** (2016) 1.
[29] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org (2020).
[30] R. Vinciotti, "DWreg: Parametric Regression for Discrete Response", R package version 2.0. https://CRAN.R-project.org/package=DWreg (2016)
[31] C. Kleiber & A. Zeileis, *Applied Econometrics with R*, New

York: Springer-Verlag. ISBN 978-0-387-77316-2. URL https://CRAN.R-project.org/package=AER (2008)
[32] P. E. Oguntunde, A. O. Adejumo & E. A. Owoloko, "Exponential Inverse Exponential (EIE) distribution", Asian Journal of Scientific Research, **10** (2017) 169.
[33] P. E. Oguntunde, O. A. Odetunmibi & A. O. Adejumo, " On the exponentiated generalized Weibull distribution: a generalization of the Weibull distribution", Indian Journal of Science and Technology **8** (2015) **1**.