Bayesian Models for Zero Truncated Count Data

.
ABSTRACT

It is important to fit count data with suitable model(s), models such as Poisson Regression, Quassi Poisson, Negative Binomial, to mention but a few have been adopted by researchers to fit zero truncated count data in the past. In recent times, dedicated models for fitting zero truncated count data have been developed, and they are considered sufficient. This study proposed Bayesian multi-level Poisson and Bayesian multi-level Geometric model, Bayesian Monte Carlo Markov Chain Generalized linear Mixed Models (MCMCglmms) of zero truncated Poisson and MCMCglmms Poisson regression model to fit health count data that is truncated at zero. Suitable model selection criteria were used to determine preferred models for fitting zero truncated data. Results obtained showed that Bayesian multi-level Poisson outperformed Bayesian multi-level Poisson Geometric model; also MCMCglmms of zero truncated Poisson outperformed MCMCglmms Poisson.

1. INTRODUCTION

Count data are type of data obtained by counting, and observations taken within a fixed period of time. Count data include zero and positive integers only. Example of studies where count data was modeled include the field of insurance [1], telecommunications [2], academics [3-4], medicine [5-9]. Other areas of study include but not limited to agriculture, sports, biology and transportation.

Count data can be equi-dispersed, over-dispersed or under-dispersed. Over-dispersion is a problem when conditional variance is larger than the conditional mean, while under-dispersion is when the conditional variance is less than the conditional mean. Poisson model is assumed to have equal mean and variance, making the model unsuitable to fit under-or over-dispersed count data [5]. Therefore, a more robust model relative to Poisson should be considered to fit data that is over-or under-dispersed. Apart from observing the conditional mean and variance, another way to check for over-dispersion or under-dispersion is to

carryout out a dispersion test, and this can be done by fitting quassi-Poisson regression model, [6] identified the robustness of quassi-poisson regression model in fitting count data.

Count data with many zeros can be modeled with zero-inflated or hurdle models, while count data with no zero count can be effectively modeled with zero truncated models. Studies such as [11-13] give details of such analysis. Linear Model (LM), sometimes called Ordinary Least Square (OLS) is considered inadequate in modelling count data because OLS cannot account for heteroscedasticity in count data [10], hence, the adoption of Generalized Linear Model (GLM) in fitting count data. GLM is implemented in such a way that a link function is assigned to the response variable, which links the response variable to the predictors, and GLM model for count data are members of exponential class of family given in equation (1).

A random variable Y has a distribution in the exponential family if its probability density function (pdf) has the form:

$$f(y|\theta) = c(y,\phi) \exp\{(y\theta - a(\theta))/\phi\}, \ g(\mu_i) = x'\beta \tag{1}$$

Where $'y'$ is the value of an observation Y, θ is a location parameter called the canonical parameter, ϕ is a dispersion parameter sometimes called the scale parameter and it determines the shape of the distribution, $c(y,\phi)$ is a normalizing factor producing unit total mass for the distribution. Equation (1) was defined by [14], and the equation for $f(y|\theta)$ indicates that the distribution of the response is in the exponential family. $g(\mu_i) = x'\beta$ identify that a transformation of the mean, $g(\mu_i)$, is linearly related to the explanatory variables contained in X. This study is centred on zero truncated data, so zero truncated models are itemized in equations (2) to (4).

Let $P_0(x; \ \theta)$ be the original Poisson distribution (pdf). Then the pdf of zero-truncated form of $P_0(x; \ \theta)$ is given as follows

$$P_0(x; \ \theta) = \frac{P_0(x; \ \theta)}{1 - P_0(0; \ \theta)} ; x = 1, 2, , ... n \tag{2}$$

And for zero-truncated binomial we have

$$P_N\left(x; \underline{\theta}\right) = \frac{1}{1-(1-p)^n} \binom{n}{k} p^k (1-p)^n \tag{3}$$

Where $\underline{\theta} = (n; p)$, and $n \in \mathbb{N}, 0 < p < 1,$

While that of Zero truncated geometric distribution is given as

$$p(x; \theta) = \theta(1 - \theta)^{x-1}, \; x = 1, 2, 3 \ldots \;\; 0 \le \theta < 1 \tag{4}$$

The link functions for fitting regression resulting from equation (2) and (4) is the log link, while that (3) is logit link.

<mark>Study by [3] showed sufficiently the superiority of MCMCglmms in fitting count data among other models used. However, [4] demonstrated the superiority of Dirichlet Process Mixture Prior of GLMMs (DPMglmms) over MCMCglmms and Bayesian Discrete Weibull to fit zero inflated data.</mark> This study proposed Bayesian multi-level model and MCMCglmms models to fit zero truncated count data. Next section is the Materials and Methods, section 3 is the Results and Discussion, and section 4 is the conclusion.

2. MATERIALS AND METHODS

2.1 Multi-Level Modelling

Multi-Level Modelling has to do with predicting the response variable (y) using the linear combination η of predictors transformed by the inverse link function (f_l) . For a given distribution $'d'$, it can be written as

$$y_i \sim d\left(f_l(\eta_i), \theta\right) \tag{5}$$

where d is called family of the distribution as usually represented in statistical software, then the parameter θ represents additional family parameters which naturally do not change as data points increases. We consider a general linear predictor that can be written as

$$\eta = \mathbf{A}\beta + \mathbf{B}\varepsilon \tag{6}$$

Where β and ε (fixed and random effect) are the coefficients at population-level and group-level respectively, $\mathbf{A}$, $\mathbf{B}$ are the corresponding design matrices. The response y as well as $\mathbf{A}$ and $\mathbf{B}$ makes up the data, while β , ε and θ are the model parameters to be estimated. Bayesian MCMC methods treat ε as a parameter unlike maximum likelihood which treats ε as error term [15]. Prior selection for Bayesian Multi-level modelling corresponds to the use of No-U-Turn Sampler (NUTS) instead of Inverse-Wishart prior distribution as discussed by [15].

2.2 Generalized Linear Mixed Models (GLMMs)

Generalized linear mixed models extension of Generalized Linear Models because as it incorporates fixed and random effects. The model can be simply put as follows:

$$y = X_i\boldsymbol{\beta} + Z_i\boldsymbol{\gamma} + e \qquad (7)$$

Where y is a $N \times 1$ column vector, $\boldsymbol{X}$ and $\boldsymbol{Z}$ are design matrices for the fixed and random predictors of the data respectively. These predictors have connected parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and e is a vector of residuals. For random effects, $\boldsymbol{\gamma_i} = (\gamma_{i1},.......\gamma_{iq})'$ explained the inclusion of covariates Z_i, the link can be written in extended form as

$$g\left(\mu_i\right) = \eta_i = X_i\boldsymbol{\beta} + Z_i\boldsymbol{\gamma} + e, \; i = 1, 2, \ldots, n \qquad (8)$$

Equation (8) was implemented using Bayesian technique of MCMCglmm, according to [3] MCMCglmms can be modelled using the R -and G –structure; G-structure was adopted in this study for the random part of the model; and the latent variables were assumed to have the multivariate normal distribution and the conjugate prior of the variance structure was inverse-Wishart prior distribution which were Gibbs sampled. MCMCglmm allows variance structures of the form:

$$\mathbf{G} = (\boldsymbol{\Lambda}_1 \otimes \mathbf{K}_1) \oplus (\boldsymbol{\Lambda}_2 \otimes \mathbf{K}_2) \oplus ....(\boldsymbol{\Lambda}_n \otimes \mathbf{K}_n) \qquad (9)$$

and the inverse structure written as

$$\mathbf{G} = (\boldsymbol{\Lambda}_1^{-1} \otimes \mathbf{K}_1^{-1}) \oplus (\boldsymbol{\Lambda}_2^{-1} \otimes \mathbf{K}_2^{-1}) \oplus ....(\boldsymbol{\Lambda}_n^{-1} \otimes \mathbf{K}_n^{-1}) \qquad (10)$$

Where $(\boldsymbol{\Lambda})$ and $(\mathbf{K})$ are matrices, $(\boldsymbol{\Lambda})$ is estimated, while $(\mathbf{K})$ are generally high dimensional and treated as known. Each component term is formed through the Kronecker product $\otimes$ which allows for possible dependence between random effects within a component term [16], while $\oplus$ is the direct sum. Expanded form of equation (9) gives

$$\mathbf{G} = \begin{bmatrix} \boldsymbol{\Lambda}_1 \otimes \mathbf{K}_1 & 0 \\ 0 & (\boldsymbol{\Lambda}_2 \otimes \mathbf{K}_2) \end{bmatrix} \qquad (11)$$

The zero off-diagonals signify independence between component term, which can be written in simplest form of Identity matrices as

$$(\boldsymbol{\Lambda}_1 \otimes \mathbf{K}_1) = \sigma_1^2 I \qquad (12)$$

According to [16], equation (12) is on the premise that random effects within a component term are independent but have a common variance. The G component may have the form:

$$\boldsymbol{\Lambda}_1 \otimes \mathbf{K}_1 = \begin{bmatrix} \sigma_{u_1}^2 & \sigma_{u_1,u_2} \\ \sigma_{u_2,u_1} & \sigma_{u_2}^2 \end{bmatrix} \otimes I \qquad (13)$$

2.3 Parameter Estimation and Model Selection

Quassi-Poisson regression analysis was carried out to identify the dispersion type of the data following the procedure by [17], further analysis was carried on the data with Bayesian multi-level analysis and MCMCglmms. Bayesian MCMC is such that for Monte Carlo, the first (MC) part has to do with generating pseudo-random numbers and the Markov Chain, the second (MC) is a sequence of number where each number is dependent on the previous number on the sequence.

Considering a Gaussian proposal $\theta_t = N(\mu, \sigma)$ for instance, the proposal shifts to the right each time a sample is drawn from the distribution and plot generated from θ_t is called trace plot, the density plot represents the sample generated, and can be expressed as

$$\theta_{t-1} \sim N(\theta_{t-2}, \sigma), \qquad \theta_t \sim N(\theta_{t-1}, \sigma) \tag{14}$$

As a default with Bayesian multi-level analysis, normal prior was adopted, using No-U-Turn Sampler (NUTS) to sample from the posterior distribution fitted with '**brms'** package in R by [18], the **'brms'** package does not work in isolation but with stan processor. In order to determine the model with a better fit between Poisson and Geometric models in the context of Bayesian multi-level modeling, the Watanabe-Akaike Information Criteria (WAIC), [19] and Leave-one-out cross validation LOO-CV proposed by [20-21] were used. Model with lower WAIC and LOO indicates a better fit for the data. On the other hand, model selection criteria such as Aikaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Deviance Information Criteria (DIC) were used, for MCMCglmms.

The trace and density plots of the predictors can be found under results and discussion. For Bayesian Multi-level models, Pareto k analysis was carried out to determine if any observation was left out in the process of the analysis. Any observation with $k > 0.7$ indicated a bad observation; consequently, the observation would be left out during analysis. Software package by [22] was used for the analysis, and "package AER" by [17] was used to carry out the dispersion test.

2.4 Data Description
The dataset used for this study was obtained from health facility in Ota, Ogun State, Nigeria, comprising of National Health Insurance Scheme (NHIS) data with no zero count.. A sample of 1647 patients under National Health Insurance Scheme was obtained from July 2016 to July 2017. Response variable (Nencounter) that is, number of encounter (visit to the doctor). The class (Eclass) indicated whether a patient was ever on admission for the period, that is, (in-patient=1, out-patient= 0). Another predictor is (follow-up), indicating whether a patient was on regular check-up or not, (follow-up=1, no follow-up=0). Gender (sex) of patients;

(male=1, female=0). Another predictor was Ndiagnosis, which represented the number of diagnosis a patient had for the period of encounter. The last predictor included is biological age of patient. Following the dispersion test, the data is under-dispersed with dispersion parameter $\phi =$0.7806. Table 1, Table 2, Figure 1 to Figure 4 was further used to describe the data.

The classification of gender according to the number of times patients had Encounter(s) at the health facility is presented in Table 1.

Table 1: Classification of gender according to Number of Encounter

Number of Encounter	Sex		Total
	Female	Male	
1-5	745	630	1375
6-10	105	91	196
11-15	31	21	52
16-21	6	9	15
21-27	7	2	9
Total	894	753	1647

Source: Authors' Computation

From Table 1, there is an indication that females had more encounter as compared to their male counterpart.
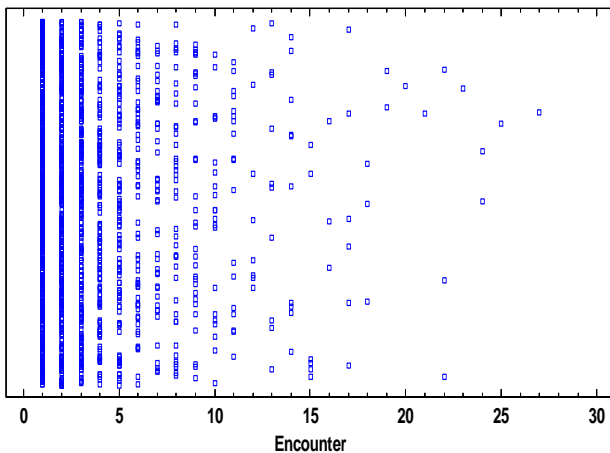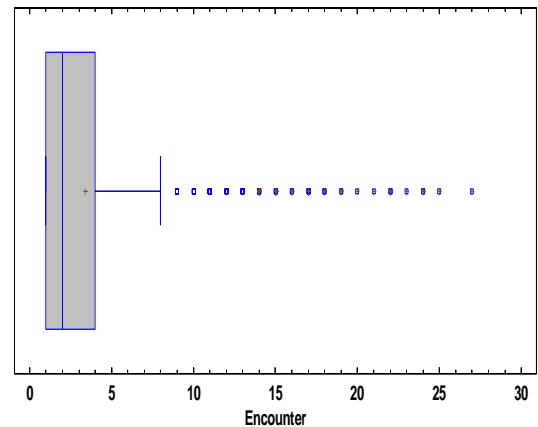

Figure 1: Scatter plot of Nencounter

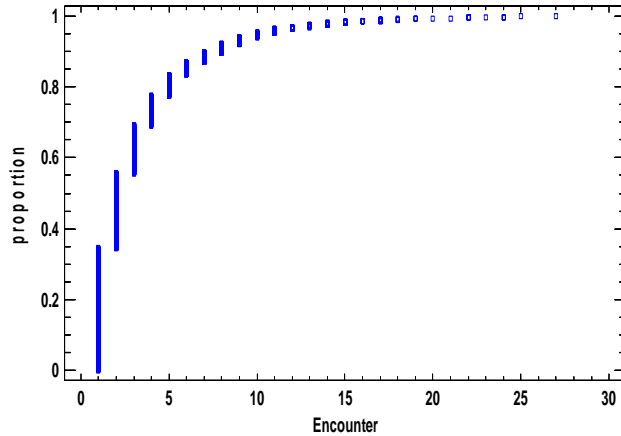**Figure 2:** Box and Whisker plot of Encounter
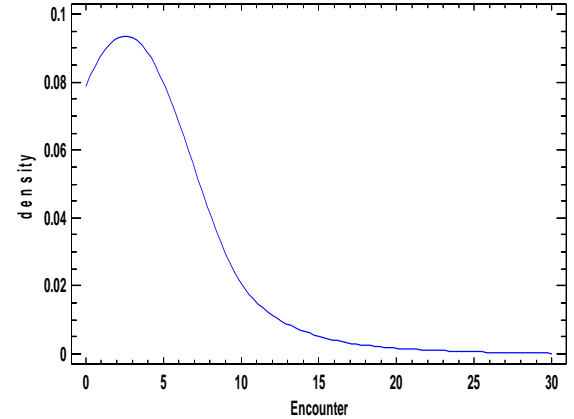
Figure 3: Quantile plot of encounter

Figure 4: Density trace plot of Nencounter

Information about patients on follow-up by gender is represented in Table 2.

Table 2: Information on follow-up status

NEncounter	Follow-up		Total
	Female	Male	
No Follow-up	737	633	1370
Follow-up	157	120	277
Total	894	753	1647

Source: Authors' Computation

Figure 1 represents scatter plot, the plot shows that patients that had regular encounters were fewer as compared to those who had encounters occasionally. This was further explained with Box-plot in Figure 2, Quantile plot in Figure 3, and Density trace plot in Figure 4. The Plots were made with use of STATGRAPHICS software.

3. RESULTS AND DISCUSSION

The result for the Quassi-Poisson regression analysis is presented in Table 3.

Table 3: Quassi-Poisson Regression Analysis

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.2941138	0.0295171	9.964	<2e-16 ***
Sex	0.0111539	0.0239250	0.466	0.64113
Age	0.0018506	0.0006451	2.869	0.00417 **
followup	-0.1534049	0.0324182	-4.732	2.41e-06 ***
Eclass	0.1700727	0.0955086	1.781	0.07515 .
Ndiagnosis	0.2677050	0.0037693	71.023	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Source: Authors' Computation

From the dispersion test carried out using quasi Poisson regression, dispersion parameter, $= \phi$ 0.7806, indicating that the data was under-dispersed since $\phi < 1$, also z = -3.0491, (P= 0.9989).

Table 4 shows that only follow-up have significantly negative effect on Encounter, so whether a patient <mark>was</mark> on follow-up or not <mark>did</mark> not necessarily mean that encounter <mark>would</mark> increase. With quasi-Poisson, [6] identified that the inadequacy inherent in Poisson model is being taken care of. Model selection for Bayesian multi-level shown in Table 4, and * is indicates model with lower value between Bayesian multi-level Geometric and Bayesian multi-level Poisson.

Table 4: Bayesian Multi-level Model Selection

Model	elpd_waic	p_waic	Waic	elpd_loo	p_loo	Looic	Waic=LOO
Geometric							
Est.	-3534	0.9	7069.1	-3534	0.9	7069.1	Yes
SE	30	0.1	60.1	30	0.1	60.1	
Poisson							
Est.	-2862.0	18.2	5724.4*	-2862.0	18.6	5724.8*	No
SE	56.3	7.2	112.5	56.4	7.4	112.8	

Source: Authors' Computation

All Pareto k estimates are good (k < 0.5) in the case of Geometric, but for Poisson model, two observations are bad with estimate of k>0.7. From Table 4, it can be seen that Bayesian Multi-level model with Poisson outperformed Geometric based on the WAIC and LOO criteria, contrary to maximum likelihood estimates (classical model) for Poisson. The results for population-level effects model for Geometric and Poisson models are presented in Tables 5 and 6.

Table 5: Population-Level Effects model for Bayesian Multi-level Geometric:

	Est.	Error	l-95% CI	u-95% CI	Eff.Sample	Rhat
Int.	-.03	.08	-.18	.22	4244	1.00
Eclass	.05	.25	-.41	.54	3908	1.00
Followup	.01	.08	-.14	.16	4260	1.00
Sex	.00	.06	-.11	.11	3868	1.00
Age	.00	.00	-.00	.00	5059	1.00
Ndiagnosis	.37	.02	.34	.43	4109	1.00

Source: Authors' Computation

Table 6: Population-Level Effects model for Bayesian Multi-level Poisson

	Est.	Error	l-95% CI	u-95% CI	Eff.Samp	Rhat
Intercept	0.29	0.03	-0.22	0.36	3624	1.00
Eclass	0.16	0.11	-0.05	0.37	2606	1.00
followup	-0.15	0.04	-0.23	-0.08	3283	1.00
Sex	0.01	0.03	-0.04	0.07	3159	1.00
Age	0.00	0.00	0.00	0.00	4203	1.00
Ndiagnosis	0.27	0.00	0.26	0.28	3586	1.00

Source: Authors' Computation

Table 7: Bayesian MCMCglmm Zero truncated Poisson and Ordinary Poisson regression Model

Model selection	Bayesian ZT Poisson	Bayesian Poisson
AIC	4709.586	5463.467
BIC	4747.433	5501.314
DIC	5027.495	5630.394

Source: Authors' Computation

For multi-level models, samples were drawn using sampling (NUTS). Scale reduction factor on split chains (at convergence, Rhat = 1). Log link was used for both Bayesian multi-level Geometric and Bayesian multi-level Poisson models, the 'Est.' in column 2 of Table 5 and Table 6 represents the posterior mean, while 'Error' stands for standard deviation of the posterior mean, 'CI' stands for the Confidence Interval, while 'Eff.Samp' stands for Efficient Sampling. All the parameters have positive relationship with "NEncounter" showing that number of Encounter can be determined by each of the predictor. By implication, patients on admission (inpatients) have more encounter than outpatients, since it was coded (0, 1), patients on follow-up had more encounters than patients that were not on follow-up, since it was coded (0, 1). Also, female patients had more encounters than male, coded as (male=0, Female=1). Biological age does not significantly have impact on encounter rate. From Table 5, Ndiagnosis (0.37) account for number of Encounter than Eclass, Follow-up, Sex, and Age. Table 5 relates to Figure 5, From Figure 5 the estimates falls within the two tails of the density plot, same with l-95%, and u-95%.

From Table 6, for Bayesian multi-level Poisson model, All the parameters have positive relationship with Encounter, except 'follow-up', showing that number of encounters can be determined by each of the predictor but not in the case of follow-up; "follow-up" had significantly negative effect on "encounter" as identified in Table 3 in the case of classical Quassi Poisson regression analysis, therefore, whether a patient is on follow-up or not did not necessarily mean that Encounter would increase. The trace and density plots for Bayesian multi-level Poisson for each of the predictors is represented in Figure 5.
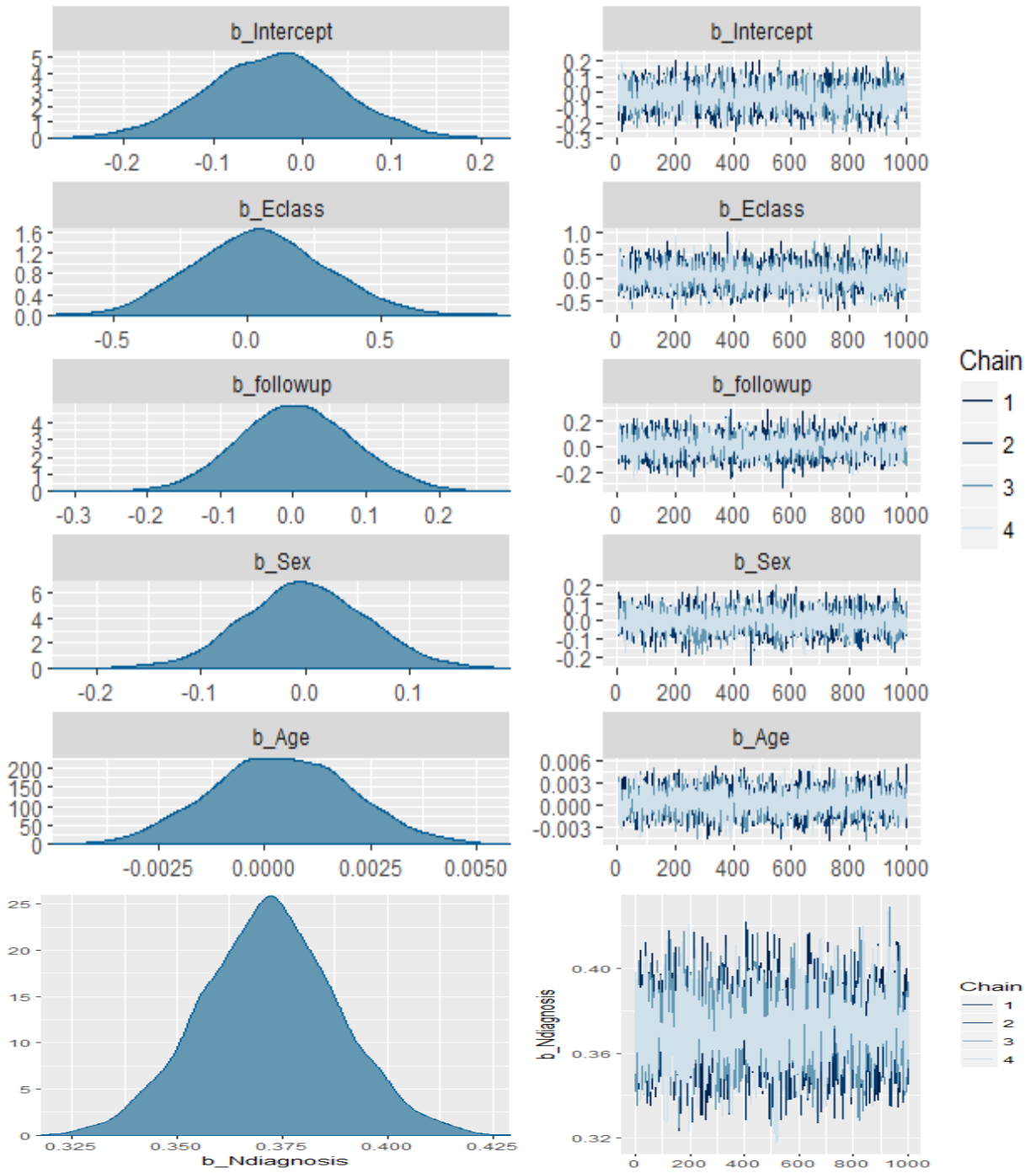
Figure 5: Trace and Density plots of all relevant parameters for Bayesian Multi-level with Geometric model

Table 8: Posterior Mean for Bayesian MCMC glmm for Zero Truncated Poisson Model

	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-0.2365188	-0.3428029	-0.1292316	678.7	<3e-04 ***
Sex	0.0042738	-0.0675917	0.0813161	1020.5	0.9137
Age	0.0014956	-0.0004543	0.0037423	1281.2	0.1545
Ndiagnosis	0.3638062	0.3455116	0.3798459	638.7	<3e-04 ***
followup	-0.0947547	-0.1988362	0.0025938	1159.1	0.0669 .
Eclass	0.1145752	-0.1929133	0.4219337	1283.6	0.4341

Source: Authors' Computation

In this study Bayesian multi-level and Bayesian MCMCglmms have been employed to fit zero truncated count data from health and insurance domain, basic descriptive analysis was carried out on the data and quassi-Poisson regression analysis was equally performed, first to determine the dispersion type and second, to determine the relationship the predictors had with the response variable. The quassi-Poisson analysis showed that it was only follow-up that did not necessarily have relationship with 'Nencounter', among all the variables. Bayesian Multi-level regression analysis was implemented and the result showed that Bayesian Multi-level Poisson regression outperformed Geometric model using "waic" and "looic" as presented in Table 4. In the class of Bayesian MCMCglmms, Bayesian Zero truncated Poisson model outperformed Ordinary Poisson regression Model using AIC, BIC and DIC as presented in Table 7.

For the Bayesian multi-level model based Geometric distribution, all the predictors have positive relationships with "Encounter", while Bayesian multi-level model based Poisson distribution showed the same relationship with that of quassi-Poisson model, indicating that follow-up was negative. That further showed how reliable quassi-Poisson is in fitting count data. The posterior mean (estimate) for MCMCglmm Zero truncated Poisson model (-0.0947547) and that of ordinary Poisson regression Model (-0.0903413) also have all estimates to be positive except that of follow-up, as shown in Table 8. The result of MCMCglmms Poisson was not tabulated but reported. The result agreed with that of the Multi-level Poisson model. Comparing the results obtained in this study with previous studies, the result obtained agreed with that of [3] and [23], although [4] demonstrated the superiority of Diritchlet Prior Mixture Model over MCMCglmms.

Other deductions from the study showed that more females were on hospital admission than their male counterpart, also patients on follow-up were significantly more than patients that

<mark>were</mark> not on follow-up, and this might indicate that more people <mark>were</mark> aware of the need to follow-up on their treatments and health status.

4. CONCLUSION

From the results obtained, it can be deduced that when comparing Bayesian Multi-level Poisson with Geometric model, Poisson outperformed the Geometric. On the other hand, when comparing Bayesian zero truncated family of MCMCglmms with MCMCglmms Poisson, the zero truncated family outperformed the Poisson family. The four models cannot be compared altogether because WAIC and LOO ([19-21]) was adopted for Bayesian Multi-level models which <mark>are</mark> considered an improvement on DIC, while AIC, BIC and DIC <mark>were</mark> used as model selection criteria for MCMCglmms Models. Therefore, when a researcher is considering a suitable multi-level model to fit under-dispersed count data that is truncated at zero, Poisson family should be considered, the problem associated with classical Poisson model has being taken care of when using Bayesian multi-level models because Bayesian multi-level models have unique sampler and unique prior selection procedure ([24-25]) and this strength makes the model suitable for any type of count data. <mark>Zero truncated Poisson family of MCMCglmms model should be considered relative to MCMCglmms Poisson.</mark>

References

[1] Jong P, Heller GZ. Generalized Linear Models For Insurance Data, 1st ed., Cambridge University Press, New York, ISBN-13 978-0-511-38677-0, 2008.

[2] Paleti R. Generalized Extreme Value models for count data: Application to worker telecommuting frequency choices. Transportation Research Part B: Methodological, 2016, 83: 104-120.

[3] Adesina OS, Agunbiade DA, Osundina SA. Bayesian Regression Model for Counts in Scholarship. *Journal of Mathematical Theory and Modelling*, 2017;7(9): 46-57

[4] Adesina OS, Olatayo TO, Agboola OO, Oguntunde PE. Bayesian Dirichlet Process Mixture Prior For Count Data . *International Journal of Mechanical Engineering and Technology,* 2018; 9(12): 630-646

[5] Cameron AC, Johansson P. *Count data regression using series expansion:With applications*, J. Appl. Econom, 1997; 12:203–223.

[6] Cameron AC, Trivedi PK. Count Panel Data. *Oxford Handbook of Panel Data Econometrics, Oxford University Press,* 2013*.*

[7] Kokonendji CC. Over- and under-dispersion models, in Methods and Applications of Statistics in Clinical Trials, Volume 2: Planning, Analysis, and Inferential Methods, N. Balakrishnan, ed., Wiley, Hoboken, NJ, 2014: 506–526.

[8] Haselimashhadi H, Vinciotti V, Yu K. A new Bayesian regression model for counts in medicine. *arXiv:1601.02820 [stat.ME];* 2016:1-21

[9] Workie MS, Lakew AM. Bayesian count regression analysis for determinants of antenatal care service visits among pregnant women in Amhara regional state, Ethiopia. *Journal of Big Data*, 2018; 5: 7. https://doi.org/10.1186/s40537-018-0117-8 5:7.

[10] Cameron AC, Trivedi PK. Regression Analysis for Count Data, *Econometric Society Monograph, 2nd ed.*, Cambridge, UK, Cambridge University Press, 1998.

[11] Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics*, 1992; 34: 1–14.

[12] Famoye F, Wang W. Censored generalized Poisson regression model*, J. Comput. Stat. Data. Anal*., 2004; 46:547–560.

[13] Famoye F, Singh KP. Zero-inflated generalized Poisson regression model with applications to domestic violence data, *J. Data Sci.*, 2006; 4:117–130.

[14] Nelder JA, Wedderburn RWA. Generalized Linear Models, *Journal of the Royal Statistical Society, Series A (General)*, 1972; 135(3):370–384.

[15] Fox J, Weisberg S. An R companion to Applied Regression, Second Edition. Sage, 2011.

[16] Hadfield JD. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software, 2010; 33(2), 1-22. URL http://www.jstatsoft.org/v33/i02/.

[17] Kleiber C, Zeileis A. Applied Econometrics with R. New York: Springer-Verlag, ISBN 978-0-387-77316-2, 2008 URL https://CRAN.R-project.org/package=AER

[18] Burkner PC. brms: An R Package for Bayesian Multilevel Models using Stan." Journal of Statistical Software, 2017.

[19] Watanabe S. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *The Journal of Machine Learning Research*, 2010; 11:3571-3594.

[20] Gelfand AE, Dey DK, Chang H. Model Determination Using Predictive Distributions with Implementation via Sampling-Based Methods." Technical report, DTIC Document, 1992.

[21] Vehtari A, Gelman A, Gabry J. Efficient Implementation of Leave-One-Out Cross-Validation and WAIC for Evaluating Fitted Bayesian Models." Unpublished manuscript, 2015; 1-22. URL http://www.stat.columbia.edu/~gelman/research/unpublished/loo_ stan.pdf.

[22] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2018, https://www.R-project.org

[23] Brooks ME, Kristensen K, Benthem KJ, A Magnusson, CW Berg, A Nielsen, HJ Skaug, M Maechle, M Bolker. Modelling Zero-Inflated Count Data With glmmTMB. bioRxiv preprint bioRxiv:132753; doi:https://doi.org/10.1101/132753; 2017.

[24] Hoffman MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo." The Journal of Machine Learning Research, 2014; 15(1):1593-1623.

[25] Lewandowski D, Kurowicka D, Joe H. Generating Random Correlation Matrices Based on Vines and Extended Onion Method." *Journal of Multivariate Analysis*, 2009; 100(9):1989-2001.