Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 1318-1327 | ISSN:2251-6727

Poisson Regression Model for Fertility Count Data and Its Applications

[1]Srinu Setti, [1]B.Muniswamy, [1]B.Punyavathi

[1]Department of Statistics, Andhra University, Visakhapatnam, Andhra Pradesh, India

(Received: 02 September 2023 Revised: 14 October Accepted: 07 November)

KEYWORDS

fertility, caesarean-section delivery, Poisson, negative binomial, INLA

ABSTRACT:

Introduction: Count data represents the number of occurrences of an event within a fixed period. For example the number of caesarean-section delivery in the lifetime of women. Count data is encountered in almost all research areas including economics, medicine, management, industrial organizations, and many more. Count data is very common in various fields such as biomedical science, public health, and marketing.

Objectives: The main aim of this study is to estimate the parameters of interest and compare the number of caesarean-section deliveries (NCSD) among women aged 15-49, in the state of Andhra Pradesh, India, using the Poisson regression model (PRM) and negative binomial regression model (NBRM). The fertility counts data set, the real-world data of the National Family Health Survey (NFHS-5), 2019-2021, from the Demography and Health Survey (DHS), 2019-2021 phase VII data is used for the analysis.

Methods: Investigating the delivery patterns among pregnant women. This study develops an algorithm based on Integrated Nested Laplace Approximation (INLA) for fitting the model NCSD in PRM and NBRM. The analysis is carried out using the INLA package in R.

Results: By use of the Deviance Information Criterion (DIC) and Watanabe-Akaike information criterion (WAIC), the result shows that the NBRM; DIC (7079.61) and WAIC (7079.61) present a comparatively better fit in modelling the NCSD than the PRM; DIC (7096.79) and WAIC (7097.89).

Conclusions: The INLA provides an efficient algorithm to model in PRM and NBRM. For further research, comparing the PRM with other models that estimate over-dispersion in count data is recommended.

1. Introduction

Count data represents the number of occurrences of an event within a fixed period [1, 2, 3, 4]. For example the number of caesarean-section delivery in the lifetime of women. Count data is encountered in almost all research areas including economics, medicine, management, industrial organizations, and many more [5]. Count data is very common in various fields such as biomedical science, public health, and marketing. Poisson models are widely used in the regression analysis of count data and as a basis for count data analysis [6, 7, 8, 9, 10]. Poisson regression is one of the most popular techniques for the analysis of count data [10, 11, 12, 13]. Negative binomial regression is the extension of Poisson with a more liberal variance assumption and could collapse into Poisson regression with the dispersion parameter equal to 0 [14]. In real-life applications, count data often exhibits over-dispersion and excess zeroes. While Negative binomial regression can model count data with over-dispersion, both Hurdle [15, 16] and Zero-inflated [17, 18, 19, 20, 21] regressions address the issue of excess zeroes in their rights.

The fertility count data of NFHS-5 is used for modelling the NCSD. NFHS-5 can provide information on important indicators, which help track the progression of Sustainable Development Goals at various levels for

SDG-1 "No Poverty", SDG-2 "Zero Hunger", SDG-3 "Good Health and Well-being" and SDG-5 "Gender Equality" [22]. The NFHS-5 provides much-needed estimates on fertility, mortality, maternal, child, and adult health, women and child nutrition, etc. Most of these indicators highlight important aspects of family well-being in India. The NFHS-5 also provides information on several indicators covered in the Sustainable Development Goals (SDGs), which India is committed to. SDG-3, which says "Ensure healthy lives and promote well-being for all at all ages" in achieving the SDGs by 2030 (NFHS-5, 2019-2021) [22].

2. Objectives

The main aim of this study is to estimate the parameters of interest and compare the number of caesarean-section deliveries (NCSD) among women aged 15-49, in the state of Andhra Pradesh, India, using the Poisson regression model (PRM) and negative binomial regression model (NBRM). The fertility counts data set, the real-world data of the National Family Health Survey (NFHS-5), 2019-2021, from the Demography and Health Survey (DHS), 2019-2021 phase VII data is used for the analysis.

3. Methods

3.1. Population and sample design: The National Family Health Survey 2019-21 (NFHS-5), the fifth in the NFHS series, provides information on population, health, and nutrition for India. NFHS-5 fieldwork for India was conducted in two phases— Phase-I from 17 June 2019 to 30 January 2020 covering 17 states and 5 union territories (UT), and Phase-II from 2 January 2020 to 30 April 2021 covering 11 states and 3 UTs—by 17 Field Agencies and interviews were completed with 724,115 women, gathered the information, for a response rate of 97 percent [22].

3.2. Sample in the study: NFHS-5 fieldwork for Andhra Pradesh was conducted from 2nd July 2019 to 14 November 2019 by Sigma Research and Consulting Pvt. Ltd. Information was gathered from 10,975 women [23]. The purposive sampling technique is used for the study. In the first stage, 724,115 women are considered. In the second stage of the purposive sampling method, 18,538 women from Andhra Pradesh were considered of which 7,563 women of Andhra Pradesh were interviewed in Phase-II. In the final stage, 2,833 women aged 15-49 are considered by the purposive sampling technique. Births delivered by caesarean section in urban and rural is 50.5% and 39.3% respectively and the total is 42.4% for births in 5 years before the survey. Births in private health facilities that are delivered by caesarean section in urban and rural areas are 66.1% and 61.4% respectively and the total is 63.0% for births in 5 years before the survey. Births in public health facilities that are delivered by caesarean section in urban and rural are 30.9% and 25.2% respectively and the total is 26.6% for births in 5 years before the survey [23].

3.3. Variables in the study: The following Table 1, briefs about the variables that are primarily and secondarily involved in the study of the number of caesarean section delivery in Andhra Pradesh, India.

Table 1: Summary of variables considered for the study

Variable	Type	Value Description
The number of caesarean section delivery	Numeric	0 = "No caesarean section delivery", 1 = "One caesarean section delivery", 2 = "Two caesarean section deliveries"
During delivery, did you experience a breech presentation?	Factor	0 = "No", 1 = "Yes", 3 = "Don't know"
Currently has heart disease	Factor	0 = "No", 1 = "Yes", 3 = "Don't know"
High blood pressure	Factor	0 = "No", 1 = "Yes", 3 = "Don't know"
Prolonged labour	Factor	0 = "No", 1 = "Yes", 3 = "Don't know"
Child is twin.	Numeric	0 = "Single birth", 1 = "1st of multiple", 2 = "2nd of multiple", 3 = "3rd of multiple", 4 = "4th of multiple", 5 = "5th of multiple"

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(4), 1318-1327 | ISSN:2251-6727

Respondent's current age	Numeric	15 to 49
Highest educational level	Factor	0 = "No education", 1 = "Primary", 2 = "Secondary", 3 = "Higher"

Where NCSD is transformed by combining variables of last birth a caesarean section and delivery by caesarean section, high blood pressure is transformed by combining variables told about pregnancy complication: high blood pressure and told had high BP on two or more occasions by a doctor or other health professional and prolonged labour is transformed by combining variables told about pregnancy complication: prolonged labour and during delivery, did you experience prolonged labour?. And the missing values are replaced with "3 = Do not know". The model is: $\mathrm{NCSD} = \mathrm{S441} + \mathrm{S728E} + \mathrm{HBP} + \mathrm{PL} + \mathrm{B0} + \mathrm{V012} + \mathrm{V106}$

The number of caesarean section delivery = During delivery, did you experience a breech presentation? + Currently has heart disease + High blood pressure + Prolonged labour + Child is twin + Respondent's current age + Highest educational level. The explained variable the number of caesarean section delivery is defined as "0", "1" and "2" or more NCSD. The explanatory variables are during delivery, did you experience a breech presentation?, currently has heart disease, high blood pressure, prolonged labour, the child is twin, respondent's current age, and highest educational level.

3.4. Statistical Model and Parameter Estimation: The count data outcome variable number of caesarean section delivery is modelled to fit the Poisson regression model. The model is fitted with PRM [11] and checked for the fulfilment of the assumption of equidispersion [10, 21]. In practicality this assumption may not be fulfilled then the model is tested for either under-dispersion or over-dispersion. If the mean is greater than the variance then it is said to be over-dispersed [24] otherwise it is said to be under-dispersed. The NBRM [25] is popularly used to address the issue of over-dispersion [26]. The integrated nested Laplace approximation (INLA) [1, 27, 28, 29, 30, 31, 32, 33] was used for estimating unknown parameters in the PRM. The INLA methodology ensures computational efficiency by using sparse representations of high dimensional matrices used in latent Gaussian models (LGMs) [27, 28, 34, 35].

One of the most important characteristics of Poisson distribution and PRM is equidispersion [21], which means that the mean and variance of the distribution are equal. A PRM is also called the log-linear model. The general mathematical form of PRM is $log(y) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$ where y is the explained variable, α and β are numeric coefficients; α being the intercept, sometimes α also is represented by β_0, it's the same and x is the explanatory variable. Consider an equation with seven predictor variables and one predictand variable:

$$\log(y) = \alpha + \beta_p x_p \text{ , where } p = 1,2,\dots,7 \quad \text{------------ } 1$$

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 \text{ -------------------------------------- } 2$$

where y = number of caesarean section delivery, x_1 = during delivery, did you experience a breech presentation? x_2 = currently has heart disease, x_3 = high blood pressure, x_4= prolonged labour, x_5= child is twin, x_6= respondent's current age and x_7 = highest educational level. This is equivalent to: $y = e^{(\alpha + \beta(x))} = e^{\beta_0} + e^{\beta_{p^*} x_p}$ ------------------------------ 3

The negative binomial distribution is a function of both mean (μ) and alpha (α); the dispersion parameter, as α →0; the distribution becomes the Poisson distribution [14, 36]. The form of the model equation for NBRM is the PRM. The log of the outcome is predicted with a linear combination of predictors [36]: Hence equations 1, 2, and 3 are obtained for NBRM. Then from equation 3, the following equations are

$$y = e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7)} \text{ ---- } 4$$

$$y = e^{\beta_0} * e^{\beta_1 x_1} * e^{\beta_2 x_2} * e^{\beta_3 x_3} * e^{\beta_4 x_4} * e^{\beta_5 x_5} * e^{\beta_6 x_6} * e^{\beta_7 x_7} \text{ -- } 5$$

4. Results

4.1. Descriptive Statistics: The following are a few tables that explain the descriptive Statistics of the respondents:

Table 2: Summary of the highest education level of the respondent

Highest educational level	Frequency	Percent
No education	369	13
Primary	313	11
Secondary	1676	59

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 1318-1327 | ISSN:2251-6727

Higher	475	16.8
Total	2833	100

Table 2, the highest educational level of the respondents describes that 1676 respondents (59%) have secondary education as their highest educational level, followed by 475 respondents (17%) have higher education, then 313 respondents (11%) have primary education and 369 respondents (13%) have no education.

Table 3: Summary of the respondent's current age

Respondent's current age	Frequency	Percent
15	1	0
16	6	0.2
17	10	0.4
18	43	1.5
19	86	3
20	146	5.2
21	184	6.5
22	238	8.4
23	256	9
24	261	9.2
25	300	10.6
26	270	9.5
27	218	7.7
28	207	7.3
29	112	4
30	132	4.7
31	68	2.4
32	65	2.3
33	41	1.4
34	48	1.7
35	48	1.7
36	24	0.8
37	23	0.8
38	12	0.4
39	14	0.5
40	6	0.2
41	2	0.1
42	1	0
43	3	0.1
44	2	0.1
45	2	0.1
47	2	0.1
48	2	0.1
Total	2833	100

The respondent's current age Table 3, explains that 300 (11%) women are 25 years old, followed by 270 (10%) women who are 26 years old. Then 1(0%) respondent each is 15 and 42 years old.

Table 4: Summary of cross tabulation of the respondent's current age and the number of caesarean section delivery

	The number of caesarean section delivery		
Respondent's current age	0	1	2
15	0	0	1
16	5	0	1
17	8	0	2
18	33	0	10
19	49	2	35
20	94	0	52
21	116	4	64
22	124	9	105
23	155	5	96
24	154	4	103
25	174	2	124
26	148	7	115
27	124	7	87
28	114	3	90
29	60	2	50
30	74	4	54
31	30	0	38
32	23	2	40
33	24	1	16
34	17	3	28
35	25	0	23
36	13	0	11
37	9	0	14
38	6	1	5
39	10	0	4
40	4	0	2
41	0	0	2
42	0	0	1
43	2	0	1
44	0	0	2
45	1	0	1
47	2	0	0
48	2	0	0

The cross-tabulation of respondent's current age and the NCSD Table 4 in the appendix, shows that the age of 25, with the highest number of 124 women have 2 NCSD, 2 women have 1 NCSD, and 174 women have zero NCSD,

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 1318-1327 | ISSN:2251-6727

followed by age 26, with the number of 115 women have 2 NCSD, 7 women have 1 NCSD and 148 women have zero NCSD. The ages of 47 and 48, with the least of no women having 2 NCSD, no women having 1 NCSD, and 2 women having zero NCSD respectively.

4.2. Women's socio-demography findings in the fitted model: Below are a few tables that show women’s socio-demography findings:

Table 5: Summary of the number of caesarean section delivery

Number of caesarean section delivery	Frequency	Percent
No caesarean section delivery	600	56.5
One caesarean section delivery	56	2.0
Two caesarean section deliveries	1177	41.5
Total	2833	100.0

Table 5 gives the details that 1177 respondents (41.5%) have 2 caesarean section delivery, followed by 1600 respondents (56.5%) who have no caesarean section delivery and 56 respondents (2%) have 1 caesarean section delivery.

Table 6: Summary of the type of place of residence

Type of place of residence	Frequency	Percent
Urban	728	25.7
Rural	2105	74.3
Total	2833	100

Table 6 describes that the majority of 2105 respondents (74.3%) are from rural areas. A few 728 respondents (25.7%) are from urban areas.

Table 7: Summary of cross-tabulation of type of place of residence and the number of caesarean section delivery

	Number of caesarean section delivery		
Type of place of residence	0	1	2
Urban	358	14	356
Rural	1242	42	821
Total	1600	56	1177

Table 7 shows the cross-tabulation of the type of place of residence and the number of caesarean section deliveries. For most women in rural areas, 821 have 2 caesarean section delivery, 42 women have 1 caesarean section delivery, and 1242 women have no caesarean section delivery. Fewer women in urban areas 356 have 2 caesarean section delivery, 14 women have 1 caesarean section delivery and 358 women have no caesarean section delivery.

4.3. Models Comparisons Criteria: The model number of caesarean section delivery is fitted in PRM using INLA, fixed effects. The result is as follows:

Table 8: Summary of values of fixed effects and model hyperparameters of PRM

Fixed effects:	Mean	Standard deviation	0.025quant
(Intercept)	-1.1	0.134	-1.37
During delivery, did you experience a breech presentation?	0.03	0.03	-0.03
Currently has heart disease	-1.1	0.5	-2.03
High blood pressure	0	0.041	-0.08
Prolonged labour	-0.1	0.05	-0.2
Child is twin	0.05	0.092	-0.13
Respondent's current age	0.02	0.004	0.01
Highest educational level	0.26	0.025	0.21

Fixed effects:	0.5quant	0.975quant	Mode	KLD
(Intercept)	-1.1	-0.84	-1.1	0
During delivery, did you experience a breech presentation?	0.03	0.088	0.03	0
Currently has heart disease	-1.1	-0.07	-1.05	0
High blood pressure	0	0.084	0	0
Prolonged labour	-0.1	-0.01	-0.11	0
Child is twin	0.05	0.231	0.05	0
Respondent's current age	0.02	0.031	0.02	0

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(4), 1318-1327 | ISSN:2251-6727

Highest educational level	0.26	0.304	0.26	0

Table 8 exhibits estimates of PRM. The mean and mode of the posterior distribution for each model parameter are determined, which are Bayesian parameter point estimates of the model [16].

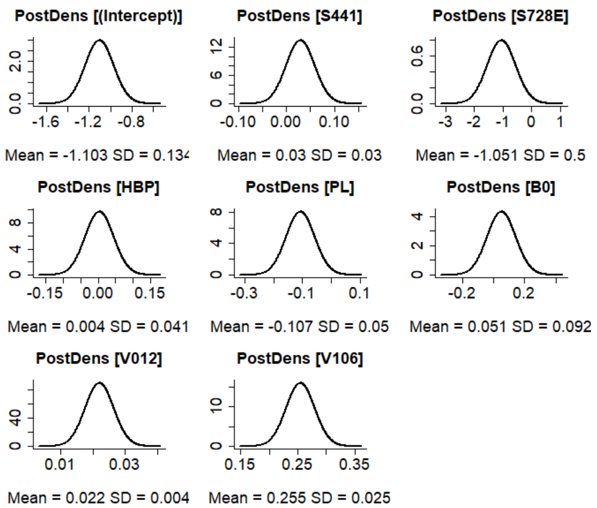

Figure 1: Fixed effects of PRM

Figure 1 shows the 'Fixed' effects of PRM represent a set of summaries from the posterior distribution. The values 0's of the Kullback-Leibler Divergence (KLD) indicate that the posterior distribution is well approximated by a Gaussian distribution. Through an estimate of the model's predictive accuracy, the best fit of a model is assessed. Cross-validation includes Akaike's Information Criterion (AIC) and Deviance Information Criterion (DIC) [37] when associated with Bayesian analyses. The DIC is a measure of the "goodness of fit" of a model penalizing for "complexity", and similar to AIC. The smaller the DIC better the model is. It can be used for comparing and ranking competing models.

Table 9: Summary of values of marginal log-likelihood, DIC and WAIC

	MLIK	DIC	WAIC
PRM	-3590.41	7096.79	7097.89

Table 9 briefs the values of marginal log-likelihood, DIC and WAIC. The DIC of PRM is 7096.79. Wantanabe-Akaike Information Criterion or Widely Applicable Information Criterion (WAIC) [38], by contrast to AIC (and DIC) WAIC is a more fully Bayesian approach for estimating the out-of-sample expectation based on the log point-wise posterior predictive density. The WAIC of PRM is 7097.89. The marginal log-likelihood of PRM is -3590.41.

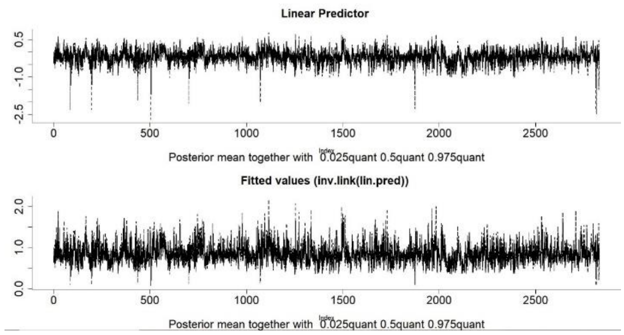

Figure 2: Predictor of PRM

Figure 2 explains the predictor of PRM. Conditional Predictive Ordinate (CPO) is computed. The sum of the CPO values is a measure of fit. In PRM, there are no non-zero CPO values. Therefore none of the observations are unusual concerning the model. Probability Integral Transforms (PIT) provides a version of CPO that reveals whether or not any of the values are 'small' (all values must be between 0 and 1).

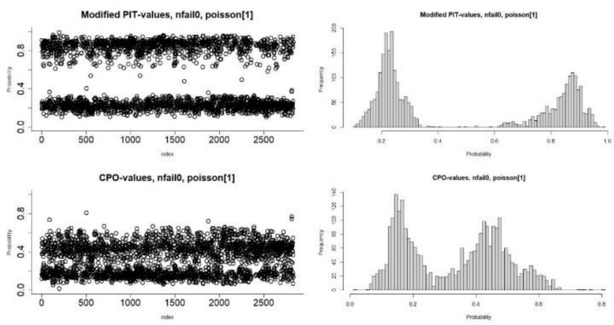

Figure 3: CPO & PIT of PRM

Figure 3 briefs about the CPO and PIT of PRM. In PRM the PIT does not indicate a lack of fit for the values do not appear to deviate from a uniform distribution.

The model number of caesarean section delivery is fitted in NBRM using INLA, fixed effects. The result is as follows: The parameters of interest are estimated [16].

Table 10: Summary of values of fixed effects and model hyperparameters of NBRM

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 1318-1327 | ISSN:2251-6727

Fixed effects:	Mean	Standard deviation	0.025 quant
(Intercept)	-1.1	0.147	-1.39
During delivery, did you experience a breech presentation?	0.03	0.032	-0.04
Currently has heart disease	-1.1	0.524	-2.08
High blood pressure	0	0.045	-0.09
Prolonged labour	-0.1	0.055	-0.21
Child is twin	0.05	0.102	-0.15
Respondent's current age	0.02	0.005	0.01
Highest educational level	0.26	0.027	0.2
Model hyperparameters	4.93	1.79	2.85

Fixed effects:	0.5quant	0.975quant	Mode	KLD
(Intercept)	-1.1	-0.82	-1.1	0
During delivery, did you experience a breech presentation?	0.03	0.092	0.03	0
Currently has heart disease	-1.1	-0.03	-1.05	0
High blood pressure	0	0.092	0	0
Prolonged labour	-0.1	0.001	-0.11	0
Child is twin	0.05	0.254	0.05	0
Respondent's current age	0.02	0.032	0.02	0
Highest educational level	0.26	0.309	0.26	0
Model hyperparameters	4.51	9.4	3.97	

Table 10 portrays the estimation of the parameters of NBRM.

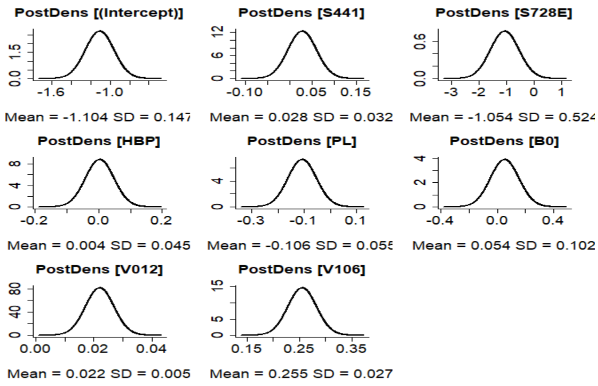

Figure 4: Fixed effects of NBRM

Figure 4 in the appendix showcases the 'Fixed' effects of NBRM [39], representing a set of summaries from the posterior distribution. The Kullback-Leibler Divergence (KLD) values “0” of NBRM indicate that the posterior distribution is well approximated by a Gaussian distribution.

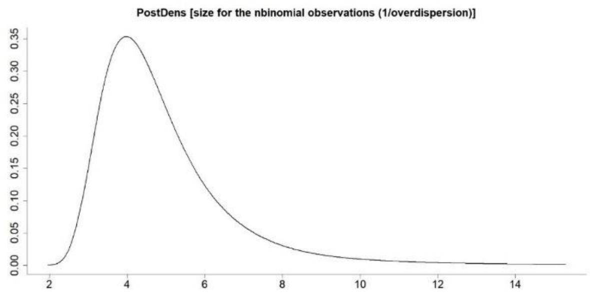

Figure 5: Hyper-parameters of NBRM

Figure 5 indicates the model hyper-parameters of NBRM, representing a set of summaries of the hyper-parameters.

Table 11: Summary of NBRM values of marginal log-likelihood, DIC and WAIC

	MLIK	DIC	WAIC
NBRM	-3583.31	7079.61	7079.61

Table 11 shows clearly the values of marginal log-likelihood, DIC, and WAIC of NBRM. The DIC of NBRM is 7079.61and WAIC is 7079.61. The marginal log-likelihood of NBRM is -3583.31.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 1318-1327 | ISSN:2251-6727

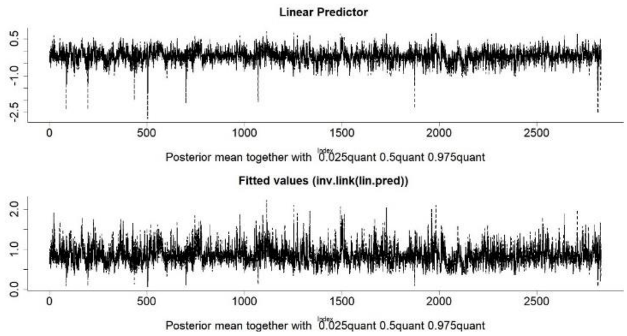

Figure 6: Predictor of NBRM

Figure 6 explains the predictor of NBRM. In NBRM, there are no non-zero CPO values. Therefore none of the observations are surprising concerning the NBRM.

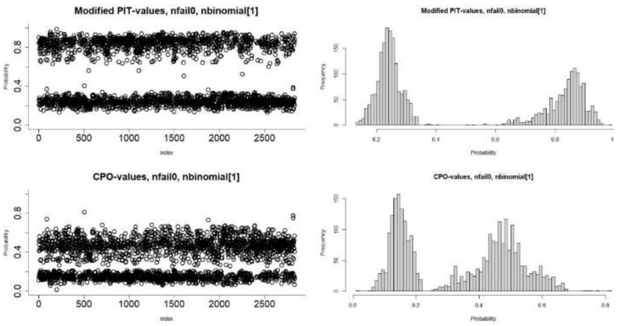

Figure 7: CPO & PIT of NBRM

Figure 7 tells about the CPO and PIT of NBRM. In NBRM the PIT values do not indicate a lack of fit for the values do not appear to deviate from a uniform distribution.

Table 12: Summary of values of marginal log-likelihood, DIC, and WAIC

	Model selection criteria		
Model	MLIK	DIC	WAIC
P	-3590.41	7096.79	7097.89
NB	-3583.31	7079.61	7079.61

Table 12 vividly explains the marginal log-likelihood of NBRM, -3583.31 is less than PRM, -3590.41. The DIC of NBRM, 7079.61 is less than PRM, 7096.79, and the WAIC of NBRM, 7079.61 is less than PRM, 7097.89. Hence NBRM is a better fit to the model NCSD. The NBRM has lower DIC, WAIC, and MLIK values compared to PRM. Hence it is more evidence that NBRM is the correct model and better fit.

5. Discussion

This paper briefly describes the INLA algorithm to estimate marginal posterior mean and mode for parameters and hyperparameters for Bayes spatial to spatio-temporal models. The NCSD data set that fits PRM and NBRM is used to illustrate the INLA estimation results. The INLA produces great computational benefits rather than the other methods in solving problems that cover random and fixed effects to every specific region and time on its spatio-temporal analysis. In this study, the INLA additive model of fixed effects is computed. The NBRM is the best-fit model compared to PRM. The model NCSD can also be computed on other models. The number of caesarean section deliveries in Andhra Pradesh, India, 2019-2021 is modelled using Bayes spatial with INLA specification. The model NCSD can be compared with other regression models.

This study aimed to fit the model number of caesarean section delivery using a count data regression model for the real-world data National Family Health Survey (NFHS-5), 2019-2021. The parameters of interest are estimated and compared to the number of caesarean-section deliveries (NCSD) among women aged 15-49, in the state of Andhra Pradesh, India. The NBRM is found to be the best and concludes that during delivery, did you experience a breech presentation?, currently has heart disease, high blood pressure, prolonged labour, the child is twin, respondent's current age and highest educational level are important factors that determine the number of caesarean section delivery.

Acknowledgement

The authors are grateful to the DHS Program for providing the data for the study.

References

1. Yu X, C. W. *Censored Count Data Analysis – Statistical Techniques And Applications.* https://digitalcommons.library.tmc.edu/uthsph_dissertsopen/1: UT School of Public Health Dissertations (Open Access).1. 2018. (*www.digitalcommons.library.tmc.edu*)

2. Cameron, A. C. and P. K. Trivedi. Regression Analysis of Count Data, Cambridge University Press, Cambridge. 1998.
3. Agresti, A. An Introduction to Categorical Data Analysis. NJ 07030, USA: John Wiley & Sons, Inc., 111 River Street, Hoboken. 1996.
4. Yadav, B., Jeyaseelan, L., Jeyaseelan, V., Durairaj, J., George, S., Selvaraj, K., & Bangdiwala, S. I. 2021. Can Generalized Poisson model replace any other count data models? An evaluation. Clinical Epidemiology and Global Health, 11, 100774. https://doi.org/10.1016/j.cegh.2021.100774. *(www.elsevier.com)*
5. Greene, W. Functional Form and Heterogeneity in Models for Count Data. Foundations and Trends in Econometrics. 2007. 1(2): 113–218. *(www.nowpublishers.com)*
6. E. L. Frome, 1983. The analysis of rates using Poisson regression models, Biometrics, 39, 665-674. *(www.jstor.org)*
7. Br¨ann¨as, K.1992. Limited dependent Poisson regression. Statistician 41: 413–423. *(www.jstor.org)*
8. Stefany Coxe, Stephen G. West & Leona S. Aiken. 2009. The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives, Journal of Personality Assessment, 91:2, 121-136, DOI: 10.1080/00223890802634175. *(www.tandfonline.com)*
9. J. F. Lawless 1987. Negative binomial and mixed Poisson regression. Canadian Journal of Statistics, 15, 209-225. https://doi.org/10.2307/3314912. *(www.jstor.org)*
10. Dejen Tesfaw Molla & B. Muniswamy, 2012. "Power of Tests for Negative Binomial Regression Coefficients in Count Data", International Journal of Mathematical Archive, Vol. 3(8), PP. 3150-3156. ISSN 2229 – 5046. *(www.ijma.info)*
11. Peter G. Moffatt. 1995. Grouped Poisson regression models: theory and an application to public house visit frequency, Communications in Statistics - Theory and Methods, 24:11, 2779-2796, DOI: 10.1080/03610929508831649. *(www.tandfonline.com)*
12. Dejen Tesfaw Molla & B. Muniswamy, 2012. "Power of Tests for Overdispersion Parameter in Negative Binomial Regression Model", IOSR Journal of Mathematics Vol. 1, Issue4, PP. 29-36. ISSN: 2278-5728. *(www.iosrjournals.org)*
13. Bilel Bousselmi, J.-F. D. 2021. Censored count data regression with missing censoring information. *Electronic Journal of Statistics, https://doi.org/10.1214/21-EJS1897*, 4343–4383. *(www.imstat.org)*
14. WenSui Liu, Jimmy Cela. 2008. Statistics and Data Analysis, SAS Global Forum 2008, Paper 371. *(www.support.sas.com)*
15. Mullahy, J. 1986. Specification and Testing of Some Modified Count Data Models, Journal of Econometrics, 33, 341-365. *(www.sciencedirect.com)*
16. Hilbe, J. M. Negative Binomial Regression. New York: Published in the United States of America by Cambridge University Press. 2007.
17. Argawu, A. S., & Mekebo, G. G. 2022. Risk factors of under-five mortality in Ethiopia using count data regression models, 2021. *Annals of Medicine and Surgery*, *82*. https://doi.org/10.1016/j.amsu.2022.104764 *(www.sciencedirect.com)*
18. Lambert, D. 1992. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing, Technometrics, Vol. 34, No. 1, 1 – 14. https://doi.org/10.2307/1269547 *(www.jstor.org)*
19. B.Muniswamy, Dejen Tesfaw Molla and Konda Reddy. 2015. "Comparison of Test Statistic for Zero-Inflated Negative Binomial against Zero Inflated Poisson Model", Indian Journal of Science and Technology, Vol 8(4), pp 349-357, ISSN, 0974-5645. *(www.indjst.org)*
20. David Giles. Notes on the Zero-Inflated Poisson Regression Model, Department of Economics, University of Victoria, March. 2010. *(www.web.uvic.ca)*
21. Saffari, S. E., Adnan, R., Greene, W. H., & Ahmad, M. H. 2013. A Poisson Regression Model For Analysis of Censored Count Data with Excess Zeroes. Jurnal Teknologi, 63(2). https://doi.org/10.11113/jt.v63.1915 *(www.journals.utm.my)*
22. http://rchiips.org/nfhs/NFHS-5Reports/NFHS-5_INDIA_REPORT.pdf National Family Health

Survey. NFHS-5: India Report. (Accessed August 24, 2023). *(www.rchiips.org)*

23. https://rchiips.org/nfhs/NFHS-5_FCTS/COMPENDIUM/Andhra_Pradesh.pdf National Family Health Survey. NFHS-5: Compendium of Fact Sheet, KEY INDICATORS State and Districts of Andhra Pradesh. (Accessed August 24, 2023). *(www.rchiips.org)*
24. Aragaw Eshetie Aguade and B. Muniswamy. 2018. “Proposed Score Test for Over-dispersion Parameter in the Multilevel Negative Binomial Regression Model”, Journal of Emerging Technologies and Innovative Research, Volume 5, Issue 12, pp. 709-720, ISSN-2349-5162. *(www.jetir.org)*
25. Hashim, L. H., Dreeb, N. K., Hashim, K. H., & Shiker, M. A. K. 2021. An application comparison of two negative binomial models on rainfall count data. Journal of Physics, 1818(1), 012100. https://doi.org/10.1088/1742-6596/1818/1/012100 *(www.opscience.iop.org)*
26. Greene, W. H. Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models, Leonard N. Stern School of Business, New York University. 1994.
27. Van Niekerk, J., Bakka, H., Rue, H., & Schenk, O. 2021. New Frontiers in Bayesian Modeling Using the INLA Package in *R. Journal of Statistical Software*, *100*(2). https://doi.org/10.18637/jss.v100.i02 *(www.jstatsoft.org)*
28. Rue, H., S. Martino, and N. Chopin. 2009. “Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with Discussion).” *Journal of the Royal Statistical Society: Series B (Methodological)* 71 (2): 319–92. *(www.rss.onlinelibrary.wiley.com)*
29. Martins, Thiago G, Daniel Simpson, Finn Lindgren, and Håvard Rue. 2013. “Bayesian Computing with INLA: New Features.” *Computational Statistics and Data Analysis* 67: 68–83. *(www.econpapers.repec.org)*
30. Rue, Håvard, Andrea Riebler, Sigrunn H. Sørbye, Janine B. Illian, Daniel P. Simpson, and FinnLindgren. 2017. “Bayesian Computing with INLA: A Review.” *Annual Review of Statistics and Its Application* 4: 395–421. https://doi.org/10.1146/annurev-statistics-060116-054045 *(www.annualreviews.org)*
31. Martino, S., & Riebler, A. 2019. Integrated Nested Laplace Approximations (INLA). arXiv (Cornell University). https://arxiv.org/pdf/1907.01248.pdf *(www.arxiv.org)*
32. Berg, J. 2022. Multistate models in survival analysis using INLA applied on data for resuscitation after cardiac arrest [Master's thesis in Mathematical Sciences]. Norwegian University of Science and Technology, Faculty of Information Technology and Electrical Engineering, Department of Mathematical Sciences.
33. Van Niekerk, J., Krainski, E. T., Rustand, D., & Rue, H. 2022. A new avenue for Bayesian inference with INLA. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2204.06797 *(www.arxiv.org)*
34. Fahrmeir, L., and G. Tutz. *Multivariate Statistical Modeling Based on Generalized Linear Models*. Berlin: Springer. 2001.
35. Blangiardo, Marta, and Michela Cameletti. *Spatial and Spatio-Temporal Bayesian Models with R-INLA*. Chichester: John Wiley & Sons. 2015.
36. Antony Mutungi. 2018. Estimating Adolescents Fertility in Kenya Using Poisson and Negative Binomial Regression Models With 2014 DHS Data, Research Report in Mathematics, Number 32. *(www.erepository.uonbi.ac.ke)*
37. Spiegelhalter, David J, Nicola G Best, Bradley P Carlin, and Angelika Van Der Linde. 2002. “Bayesian Measures of Model Complexity and Fit.” Journal of the Royal Statistical Society: Series B (Methodological) 64 (4): 583–639. *(www.rss.onlinelibrary.wiley.com)*
38. Gelman, Andrew, Jessica Hwang, and Aki Vehtari. 2014. “Understanding Predictive Information Criteria for Bayesian Models.” Statistics and Computing 24 (6): 997–1016. *(www.springer.com)*
39. B. Muniswamy and Aragaw Eshetie Aguade. 2019. “Proposed Score Test of Homogeneity of Groups in the Multilevel Poisson Regression Model for Clustered Discrete Data”, International Journal of Scientific Research and Reviews, 8(1), 1036-1052, ISSN: 2279–0543. *(www.ijsrr.org)*