

Innovations

Negative Binomial Regression Model for Over-dispersed Fertility Count Data

¹Srinu Setti*, ²B. Muniswamy, ³B. Punyavathi

¹Research Scholar, Andhra University, Visakhapatnam, India

²Professor, Andhra University, Visakhapatnam, India

³Guest Faculty, Andhra University, Visakhapatnam, India

Corresponding Author: [Srinu Setti](#)

Abstract

Background: Researchers have long been concerned about the steady rise in the number of caesarean section deliveries in several nations, including India. Many medical professionals and the general population now choose to have caesarean section deliveries due to the growing medicalization of human health. In light of this, the proposed study looks at the main factors that influence caesarean deliveries in the Indian state of Andhra Pradesh. The study's primary goal is to use Poisson regression model (PRM) and negative binomial regression model (NBRM) to estimate parameters and to identify the significant factors influencing the number of caesarean-section deliveries amidst women in Andhra Pradesh, India, who are between ages of 15 and 49. **Methods:** The analysis makes use of the fertility count data set, real-world input of 2019–2021 Demographic and Health Surveys (DHS) phase VII input, National Family Health Survey (NFHS–5) input. The Generalized Linear Models (GLM) used in this study to fit the number of caesarean-section deliveries model in PRM and NBRM. The R package Modern Applied Statistics with S (MASS) is used for the analysis. **Results:** The parameters are estimated using both PRM and NBRM; that are Intercept (-1.0742; -1.0776), "Breech Presentation" ("Yes and Don't know") (-0.3199 and 0.5194; -0.3227 and 0.5149), "Currently has heart disease" ("Yes") (-0.8527; -0.8644), "High blood pressure" ("Yes and Don't know") (0.0305 and 0.0263; 0.0328 and 0.0230), "Prolonged labour" ("Yes and Don't know") (-0.0699 and -0.7673; -0.0684 and -0.7589), "Child is twin" ("1st multiple, 2nd multiple, 3rd multiple") (0.2146, 0.1041, -11.4900; 0.2182, 0.0997, -18.4904), "Age" (0.0222; 0.0223), "Educational level" ("Primary, Secondary, Higher") (0.1070, 0.4918, 0.7000; 0.1039, 0.4910, 0.7003) of PRM and NBRM respectively. According to the outcomes of applying the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), the AIC (7052.4) and BIC (7147.556); AIC (7062.1) and BIC (7151.32) are of NBRM and PRM, respectively. At the 0.1 level of significance, the variable "Currently has heart disease" (Yes) is significant. At the 0.001 level of

significance, the variables "Age," "Educational level" ("Secondary and Higher"), "Prolonged labour" ("don't know"), and "Breech Presentation" ("Yes and don't know") are significant. **Conclusion:** The parameters impacting the number of caesarean section deliveries include "Currently has heart disease" ("Yes"), "Breech presentation" ("Yes and don't know"), "Prolonged labour" ("Don't know"), "Age," and "Educational level" (Secondary and Higher) in the model fitted using PRM and NBRM. The health of women (those in the age range of 15 to 49) must be a priority for the governmental authorities that make policies for women's welfare. The NBRM exhibits a relatively better match with "NCSD" than the PRM. It is advised to compare PRM with various models that predict over-dispersion in count data in order to do additional study.

Keywords: fertility, count data, caesarean section, PRM, NBRM, GLM, AIC, BIC

I. Background

Due to resource scarcity and the medicalization of health issues, emerging countries have delicate health circumstances that have caused severe concerns and made achieving universal health, particularly for women and children, challenging. Consequently, we may enhance and protect the maternal health care system by creating a sensible plan and implementing the necessary treatments as needed. Giving birth is not an illness, but a completely normal process. Worldwide, the use of obstetrical procedures has expanded recently [3]. According to a study conducted in West Bengal, the number of caesarean deliveries has skyrocketed globally over the years, especially in high-income and emerging nations. On the other hand, numerous negative short- and long-term effects of caesarean delivery on reproductive health and delivery have been documented in research investigations [28]. A study in Uttar Pradesh shows that women who live in urban areas and give birth in private sector hospitals with different clinico-social profiles have an extremely high chance of having a caesarean section[21].

Count data shows how many times an event occurs over a given time frame [36, 9, 2, 35]. For instance, the quantity of caesarean sections [10] a woman has ever had in her lifetime. Nearly all fields of study, including management, industrial organizations, economics, health, and so on, deal with count data [15]. Count data is widely used in many domains, including marketing, public health, and biological sciences. Poisson models are frequently employed as the foundation for count data analysis as well as in regression analysis of count data [14, 8, 31, 20, 12, 30]. One of the most often used methods for counting data analysis is Poisson regression [12, 26, 13, 7]. PRM considering generously assumptions of variance is extended by negative binomial regression, which has the potential to converge into PRM when the dispersal parameter is becomes 0 [34]. Count data frequently shows excess zeroes and over-dispersion in real-world applications.

NBRM is applied for over-dispersed count data; however, excess zeroes in the data are addressed by both Hurdle [25, 17] and Zero-inflated [5, 22, 6, 11, 27] regressions.

The “NCSD” is modelled using NFHS-5 fertility count data. Estimates on fertility, mortality, adult, child, and maternal health, as well as women's health and children's nutrition, are provided by the NFHS-5. The majority of these metrics draw attention to significant facets of Indian families' well-being. NFHS-5 provides information on a number of indicators included in the Sustainable Development Goals (SDGs), to which India is committed. In order to achieve the SDGs by 2030 (NFHS-5, 2019–2021), SDG-3, which states, “Ensure healthy lives and promote well-being for all at all ages,” must be met [30, 18, 29].

II. Methods

The fifth survey in the NFHS series, the NFHS-5, offers data on India's population, health, and nutrition. In India, NFHS-5 fieldwork was carried out in two phases: Phase-I, which covered 17 states and 5 union territories (UT), ran from June 17, 2019 to January 30, 2020, and Phase-II, which covered 11 states and 3 UTs, ran from January 2, 2020 to April 30, 2021. 17 Field Agencies conducted enquiry with 7,24,115 women, gathering data that resulted in a 97 percent response rate [30, 18, 29]. The fieldwork for NFHS-5 in Andhra Pradesh was carried out by Sigma Research and Consulting Pvt. Ltd. from July 2, 2019 to November 14, 2019. 10,975 women provided information [30, 29, 19]. The study used the purposive sampling technique. In the initial phase, 7, 24,115 women are taken into account. 18,538 women from Andhra Pradesh were taken into consideration in the second step of the purposive sampling procedure, of whom 7,563 women were enquired in Phase II. At the end, 2,833 females between the ages of 15 and 49 are taken into account using the purposive sample method. The analysis makes use of NFHS-5 secondary data. 50.5%, 39.3% is the percentage of caesarean section deliveries in urban, in rural areas respectively, and total for births that occurred five years before to the study is 42.4%. The percentage of caesarean section deliveries in urban and rural areas that take place in private health facilities are 66.1% and 61.4%, respectively, while the total for births that occurred five years before to the survey is 63.0%. Caesarean section rates for new-borns in public health institutions in urban, in rural areas are 30.9%, 25.2%, respectively, and 26.6% for births that occurred five years prior to the survey [30, 29, 19].

The variables that are principally and secondary to the research of “NCSD” in Andhra Pradesh, India, are summarized in Table 1.

Table 1: A list of the Variables taken into account for the research

Variable	Type	Value Description
The number of caesarean section deliveries	Categorical	0 = "No caesarean section delivery", 1 = "One caesarean section delivery", 2 = "Two caesarean section deliveries"
Breech Presentation	Categorical	1 = "No", 2 = "Yes", 3 = "Don't know"
Currently has heart disease	Categorical	1 = "No", 2 = "Yes", 3 = "Don't know"
High blood pressure	Categorical	1 = "No", 2 = "Yes", 3 = "Don't know"
Prolonged labour	Categorical	1 = "No", 2 = "Yes", 3 = "Don't know"
Child is twin.	Categorical	1 = "Single birth", 2 = "1st of multiple", 3 = "2nd of multiple", 4 = "3rd of multiple", 5 = "4th of multiple", 6 = "5th of multiple"
Age	Interval	15,16,.....,49

Educational Level	Categorical	1 = "No education", 2 = "Primary", 3 = "Secondary", 4 = "Higher"
Type of place of residence	Categorical	1 = "Urban", 2 = "Rural"

The model is: "Number of caesarean section deliveries" = "Breech presentation" + "Currently has heart disease" + "High blood pressure" + "Prolonged labour" + "Child is twin" + "Educational level" This is written as follows $y = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7$. The explained variable "NCS D" is defined as "0", "1" and "2" or more "NCS D"[12]. While the variables, last birth, caesarean section, and delivery by caesarean section transform "NCS D", the variables "High blood pressure", and being told by a doctor or other health professional that they had high blood pressure on two or more occasions, combines to form x_3 and the variables "Prolonged labour", and was asked if they had experienced prolonged labour during delivery, combines to form x_4 . The other variables are explained below. Additionally, "3 = Do not know" is used to replace the missing numbers.

PRM is fitted to the count data outcome variable, which is the "NCS D" [8, 30]. First, PRM is fitted to the model, followed by NBRM. Although more usually referred to as simply over-dispersed, data with a higher variance than the mean are called Poisson over-dispersed [16]. The response variable is considered over-dispersed if its variance exceeds the mean [12, 4], and under-dispersed if the variance is less than the mean. "NCS D" model is fitted in PRM and NBRM using GLM [23]. MASS [33] package is used in the analysis.

$\log(y) = \alpha + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p$ is the general mathematical form of PRM, where y is the predictand variable, α and β are numeric coefficients, α is the intercept (occasionally α is sometimes denoted by β_0 , which is the same), and x is the predictor variable.

Examine the following equation, which has one predictand variable and seven predictor variables;

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

$$\log(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \beta_7x_7 \text{ ----- 2}$$

Where

- y = "Number of caesarean section deliveries"
- x_1 = "Breech presentation"
- x_2 = "Currently has heart disease"

x_3 = "High blood pressure"

x_4 = "Prolonged labour"

x_5 = "Child is twin"

x_6 = "Age"

x_7 = "Educational level"

Which is similar to: $y = e^{(\alpha + \beta(x))} = e^{\beta_0} + e^{\beta_p * x_p}$ ----- 3

The distribution known as negative binomial is determined by the mean (μ) and alpha (α). When α tends to zero, the distribution is known as the Poisson distribution [30, 34, 32]. The PRM is the format of the NBRM model equation. Thus, for NBRM, equations 1, 2, and 3 are obtained. Next, based on equation 3, the subsequent equations are

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

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

III. Results

The study takes into account the demographic data of the respondents, including their "Type of Residence," "Age," and "Educational Level." The research study conducted in Andhra Pradesh, India, on women aged 15 to 49 of the NFHS-5 data shows a range of outcomes. The following are the respondents' descriptive data.

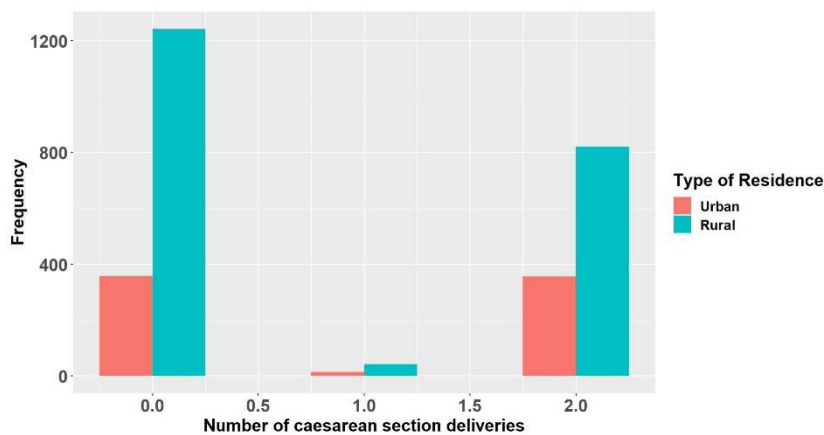


Figure 1: Plot of "NCS" and Type of Residence

According to Figure 1, the majority of rural women have two "NCS"s compared to the majority of urban women, and rural women also have a higher percentage of one "NCS" than urban women. The plot makes it very evident that most rural women give birth by caesarean section.

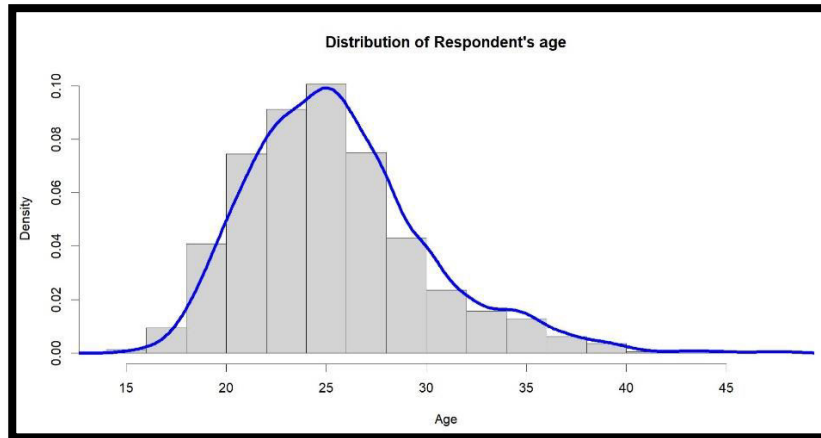


Figure 2: Distribution of the Respondent's Age

The age distribution of the respondents is shown in Figure 2. It demonstrates that women with two “NCSD” are most common when they are 25 years old, and 26 years old is the next age group.

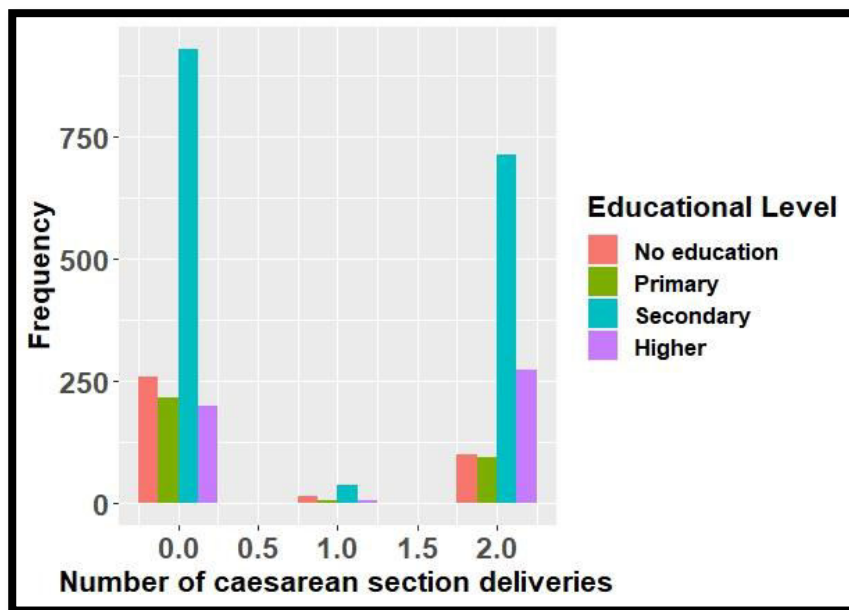


Figure 3: Plot of “NCSD” and Educational Level

According to Figure 3, women with two “NCSD” are mainly qualified for secondary education, which is followed by women qualified for higher education, women without any education, and women qualified for primary education. Second, women with one “NCSD” are mostly qualified secondary school graduates, followed by uneducated women, elementary school graduates, and women with higher education. Ultimately, the majority of women who do not give birth by caesarean section have completed secondary education, followed by women with no formal education, women with elementary education, and women with higher education.

The Fitted “NCSD” model in PRM using MASS. The result is given below:

Table 2: PRM Values Synopsis - Coefficients:				
Variables	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	-1.0742	0.1612	-6.665	0.0000 ***
Breech Presentation				
Yes	-0.3199	0.0700	-4.569	0.0000 ***
Don't know	0.5194	0.1148	4.525	0.0000 ***
Currently has heart disease				
Yes	-0.8527	0.5011	-1.702	0.0888 .
High blood pressure				
Yes	0.0305	0.0608	0.502	0.6159 @
Don't know	0.0263	0.1456	0.180	0.8568 @
Prolonged labour				
Yes	-0.0699	0.0631	-1.108	0.2677 @
Don't know	-0.7673	0.1868	-4.107	0.0000 ***
Child is twin				
1 st of multiple	0.2146	0.2131	1.007	0.3137 @
2 nd of multiple	0.1041	0.2097	0.497	0.6195 @
3 rd of multiple	-11.4900	284.6592	-0.040	0.9678 @
Age	0.0222	0.0047	4.701	0.0000 ***
Educational level				
Primary	0.1070	0.1006	1.064	0.2875 @
Secondary	0.4918	0.0764	6.437	0.0000 ***
Higher	0.7000	0.0814	8.600	0.0000 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 and InSignif. Code: @				

The PRM value summary is shown in Table 2. The elements “Currently has heart disease” (“Yes)is significant at the 0.1 level of significance (LOS). The “Breech presentation” (“Yes and Don’t know”), “Prolonged labour” (“Don't know”), “Age” and the “Educational level” (Secondary and Higher) are significant at the 0.001 LOS in the fitted PRM

Table 3: An overview of the -2L, AIC, and BIC PRM values

	-2L	AIC	BIC
PRM	7032.084	7062.1	7151.32

The values of the BIC [24, 1] and AIC [24, 1] are summarized in Table 3. PRM has a 2*Log-likelihood of -7032.084.

The Fitted “NCSD” model in NBRM using MASS. The result is shown below:

Table 4: NBRM Values Synopsis - Coefficients:

Variables	Estimate	Std.Error	z-value		Pr(> z)
(Intercept)	-1.0776	0.1739	-6.196		0.0000***
Breech presentation					
Yes	-0.3227	0.0747	-4.320		0.0000***
Don't know	0.5149	0.1300	3.960		0.0000***
Currently has heart disease					
Yes	-0.8644	0.5236	-1.651		0.0988 .
High blood pressure					
Yes	0.0328	0.0660	0.497		0.6189
Don't know	0.0230	0.1591	0.145		0.8851
Prolonged labour					
Yes	-0.0684	0.0688	-0.994		0.3202
Don't know	-0.7589	0.2065	-3.675		0.0000***
Child is twin					
multiple 1st of	0.2182	0.2313	0.943		0.3455
multiple 2nd of	0.0997	0.2287	0.436		0.6630
multiple 3rd of	-18.4904	9426.6169	-0.002		0.9984
Age	0.0223	0.0051	4.342		0.0000***
Educational Level					
Primary	0.1039	0.1065	0.976		0.3291
Secondary	0.4910	0.0810	6.061		0.0000***
Higher	0.7003	0.0871	8.042		0.0000***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 and InSignif. Code: @					

The NBRM value summary is explained in Table 4. In the fitted NBRM, the elements "Currently has heart disease" ("Yes") is significant at the 0.1 LOS. The "Breech presentation" ("Yes and Don't know"), "Prolonged labour" ("Don't know"), "Age," and "Educational level" (Secondary and Higher) are significant at the 0.001 LOS.

Table 5: An overview of the -2L, AIC, and BIC NBRM values

	-2L	AIC	BIC
PRM	7020.371	7052.4	7147.556

The 2*Log-likelihood of NBRM is -7020.371, given in Table 5.

Table 6: An overview of the -2L, AIC, and BIC values

Model	Model selection criteria		
	-2L	AIC	BIC
PRM	7032.084	7062.1	7151.32
NBRM	7020.371	7052.4	7147.556

Table 6 provides a clear explanation of the -2L, AIC, and BIC summary values. NBRM, 7052.4 has a lower AIC than PRM, 7062.1, while NBRM, 7147.556 has a lower BIC than PRM, 7151.32. As a result, NBRM fits the "NCSD" model better. Compared to the NBRM, the PRM has lower AIC and BIC values. Thus, there is more proof that the NBRM model fits the data better.

IV. Discussion

The parameters are estimated using PRM and NBRM for the "NCSD" model, NFHS-5, 2019–2021 data, that is available at DHS program. The intercept coefficient of PRM is (-1.0742) and is significant at 0.1 LOS. The coefficient of "Breech presentation" ("Yes and Don't know"), is(-0.3199 and 0.5194) and is significant at 0.001 LOS. The coefficient of "Currently has heart disease" ("Yes") is (-0.8527) and is significant at 0.1 LOS. The coefficient of "Prolonged labour" ("Don't know) is (-0.7673) and is significant at 0.001 LOS. The coefficient of "Age" is (0.0222) and is significant at 0.001 LOS. The coefficient of "Educational level "(Secondary and Higher) is (0.4918 and 0.7000) and is significant at 0.001 LOS. Therefore, it is concluded that these elements "Breech presentation" ("Yes and Don't know"), "Currently has heart disease" ("Yes"), "Prolonged labour" ("Don't know"), "Age" and "Educational level "(Secondary and Higher) are affecting "NCSD". MASS is used to model "NCSD" in Andhra Pradesh, India, between 2019 and 2021. It is possible to compare the "NCSD" model with different regression models. The coefficient of "High blood pressure" ("Yes and Don't know") is (0.0305 and 0.0263) and is not significant. The coefficient of "Prolonged labour" ("Yes") is (-0.0699) and is not significant. The coefficient of

“Child is twin” (“1st multiple, 2nd multiple, 3rd multiple”) is (0.2146, 0.1041, -11.4900) and is not significant. The coefficient of “Educational level” (“Primary”) is (0.1070) and is not significant.

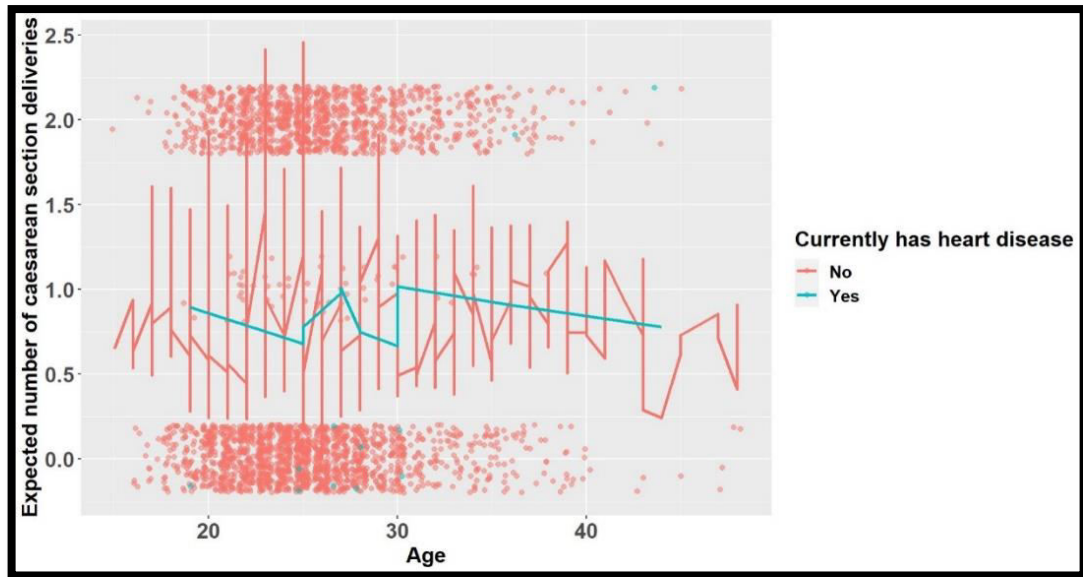


Figure 4: PRM's predicted “NCSD”

Figure 1 indicates the predicted “NCSD” for those respondents “currently has heart disease” with the respondents current “age” using PRM.

The intercept coefficient of NBRM is (-1.0776) and is significant at 0.1 LOS. The coefficient of “Breech presentation” (“Yes and Don’t know”), is (-0.3227 and 0.5149) and is significant at 0.001 LOS. The coefficient of “Currently has heart disease” (Yes) is (-0.8644) and is significant at 0.1 LOS. The coefficient of “Prolonged labour” (“Don’t know”) is (-0.7589) and is significant at 0.001 LOS. The coefficient of “Age” is (0.0223) and is significant at 0.001 LOS. The coefficient of “Educational level” (Secondary and Higher) is (0.4910 and 0.7003) and is significant at 0.001 LOS. Therefore, it is concluded that these elements “Breech presentation” (“Yes and Don’t know”), “Currently has heart disease” (“Yes”), “Prolonged labour” (“Don’t know”), “Age” and “Educational level” (Secondary and Higher) are affecting “NCSD”. The coefficient of “High blood pressure” (“Yes and Don’t know”) is (0.0328 and 0.0230) and is not significant. The coefficient of “Prolonged labour” (“Yes”) is (-0.0684) and is not significant. The coefficient of “Child is twin” (“1st multiple, 2nd multiple, 3rd multiple”) is (0.2182, 0.0997, -18.4904) and is not significant. The coefficient of “Educational level” (“Primary”) is (0.1039) and is not significant.

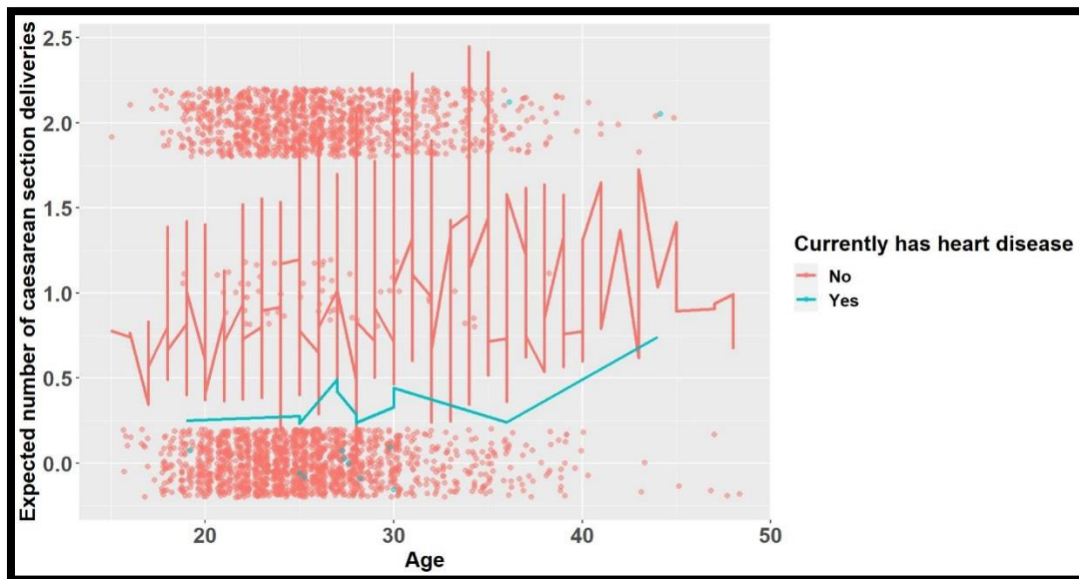


Figure 5: NBRM's predicted "NCSD"

Using NBRM, Figure 2 predicts the "NCSD": for respondents who "currently has heart disease" based on their present "age".

The following are the research limitations: the study does not take into account the respondents' socioeconomic status; their demographics are not proportionate; the respondents' caste and religion are not taken into account in the model; and a comparative analysis with other Indian states is not conducted. This would have required a very large study that took into account numerous variables under various circumstances. These factors can be taken into account for the study in the future to lower the "NCSD."

V. Conclusion

Our research indicates that 25-year-old respondents who reside in rural areas are more likely to give birth via caesarean section. In order to reduce the number of caesarean section births, local medical officers and government health policymakers should raise public awareness, establish well-equipped public healthcare facilities, and, most importantly, assign medical staff to counsel expectant mothers about normal vaginal deliveries. In order to prevent future complications, the government agencies responsible for formulating health policies pertaining to women's welfare must prioritize the health of women, defined as those between the ages of 15 and 49 to reduce the "NCSD" by raising maternal health literacy and awareness among women and the community. The variables "Currently has heart disease" ("Yes"), "Breech presentation" ("Yes and Don't know"), "Prolonged labour" ("Don't know"), "Age," and "Educational level" ("Secondary and Higher") have an effect on the "NCSD" in the model fitted using PRM and NBRM. Compared to the PRM, the NBRM shows a comparatively better match with "NCSD". It is discovered that the NBRM fits the best for "NCSD".

Further research is suggested by comparing PRM with different models that predict over-dispersion in count data.

List of abbreviations

NCSD =	Number of caesarean section deliveries
NFHS-5	= National Family Health Survey series 5
DHS =	Demographic and Health Surveys
PRM =	Poisson regression model
NBRM =	Negative Binomial Regression Model
GLM =	Generalized Linear Models
MASS =	Modern Applied Statistics with S
AIC =	Akaike Information Criterion
BIC =	Bayesian Information Criterion
LOS =	Level of significance
2L =	2*Log-likelihood
SDGs =	Sustainable Development Goals

Acknowledgement: The DHS Program provided the study's data, for which the authors are appreciative.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Conflicting Interests: The Authors declare that there is no conflict of interest.

References

1. Acquah, H. D. (2010). Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *Journal of Development and Agricultural Economics*, 2(1), pp. 001–006. www.academicjournals.org
2. Agresti, A. *An Introduction to Categorical Data Analysis*. NJ 07030, USA: John Wiley & Sons, Inc., 111 River Street, Hoboken. 1996.
3. Akyıldız, D., Çoban, A., Uslu, F. G., & Taşpınar, A. (2021). Effects of obstetric interventions during labor on birth process and newborn health. *Florence Nightingale Journal of Nursing* :, 29(1), 9–21. (www.fnjn.org)
4. AragawEshetieAguaade 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)
5. 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. (www.sciencedirect.com)
6. B.Muniswamy, DejenTesfawMolla 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)
 7. BilelBousselmi, J.-F. D. (2021). Censored count data regression with missing censoring information. *Electronic Journal of Statistics*, (www.imstat.org)
 8. Br`ann`as, K. (1992). Limited dependent Poisson regression. *Statistician* 41: pp. 413–423. (www.jstor.org)
 9. Cameron, A. C. and P. K. Trivedi. *Regression Analysis of Count Data*, Cambridge University Press, Cambridge. 1998.
 10. Cavallaro, F. L., Cresswell, J. A., Frana, G. V., Victora, C. G., Barros, A. J., & Ronsmans, C. (2013). Trends in caesarean delivery by country and wealth quintile: cross-sectional surveys in southern Asia and sub-Saharan Africa. *Bulletin of the World Health Organization*, 91(12), 914–922D. (www.who.int)
 11. David Giles. *Notes on the Zero-Inflated Poisson Regression Model*, Department of Economics, University of Victoria, March. 2010. (www.web.uvic.ca)
 12. DejenTesfawMolla & 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)
 13. DejenTesfawMolla & 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)
 14. E. L. Frome. (1983). The analysis of rates using Poisson regression models, *Biometrics*, 39, pp. 665-674. (www.jstor.org)
 15. Greene, W. *Functional Form and Heterogeneity in Models for Count Data*. *Foundations and Trends in Econometrics*. 2007. 1(2): 113–218. (www.nowpublishers.com)
 16. Hilbe, J. M. (2nd ed.). *Negative Binomial Regression*. Cambridge University Press. 2011.
 17. HILBE, J. M. *Negative Binomial Regression*. New York: Published in the United States of America by Cambridge University Press. 2007.
 18. 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)
 19. J. F. Lawless. (1987). Negative binomial and mixed Poisson regression. *Canadian Journal of Statistics*, 15, pp. 209-225. (www.jstor.org)
 20. Katara, S., Kumar, R., & Mishra, S. (2023). Analysis of determinants responsible for Cesarean delivery in Uttar Pradesh: Evidence from NFHS-5 (2019-21). In *Bharatiya Shiksha Shodh Patrika*: Vol. Vol. 42 (Issue No.

- 1(ii)).(www.researchgate.net)
21. Lambert, D. (1992). *Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing*, *Technometrics*, Vol. 34, No. 1, pp. 1 – 14. (www.jstor.org)
 22. McCullagh, P. and Nelder, J. A. *Generalized linear models, 2nd Ed.* Chapman and Hall, New York, USA. 1989.
 23. McElduff, F. (2012). *Models for discrete epidemiological and clinical data*. Doctoral Thesis, UCL (University College London). discovery.ucl.ac.uk(www.discovery.ucl.ac.uk)
 24. Mullahy, J. (1986). *Specification and Testing of Some Modified Count Data Models*, *Journal of Econometrics*, 33, 341-365. (www.sciencedirect.com)
 25. 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, (www.tandfonline.com)
 26. 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). (www.journals.utm.my)
 27. Sarkar, S. (2020). *Prevalence and determinants of the use of caesarean section (CS) in the dichotomy of 'public' and 'private' health facilities in West Bengal, India*. *Clinical Epidemiology and Global Health*, 8(4), 1377–1383. (www.cegh.net)
 28. Srinu Setti, B. Muniswamy, and B. Punyavathi. (2023). *Poisson Regression Model for fertility count data and its applications*. *Journal of chemical health risks*, Vol. 13. No. 4, pp. 1318-1327. (www.jchr.org)
 29. Srinu Setti, B. Muniswamy, and B. Punyavathi. (2023). *Right-Censored Poisson Regression Model for Fertility Count Data*. *Journal of chemical health risks*, Vol. 13. No. 6, pp. 2068-2078. (www.jchr.org)
 30. StefanyCoxe, 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, pp. 121-136, (www.tandfonline.com)
 31. Van Niekerk, J., Krainski, E. T., Rustand, D., & Rue, H. (2022). *A new avenue for Bayesian inference with INLA*. *arXiv (Cornell University)*. (www.arxiv.org)
 32. Venables WN, Ripley BD. *Modern Applied Statistics with S*. Fourth Edition. Springer. 2002. ISBN 0-387-98825-4, URL.
 33. WenSui Liu, Jimmy Cela. (2008). *Statistics and Data Analysis, SAS Global Forum 2008, Paper 371*. (www.support.sas.com)
 34. 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. (www.elsevier.com)
 35. Yu X, C. W. *Censored Count Data Analysis – Statistical Techniques And Applications*. digital commons. library.tmc.edu (www.digitalcommons.library.tmc.edu)