"Science Stays True Here"
Journal of Mathematics and Statistical Science (ISSN 2411-2518, USA), Vol.12, Issue 11, 26-43 | Science Signpost Publishing

Three- or Four-parameter Over-dispersion Count Models in Disease and Treatment Management

Manasi Sheth[1] and N.Rao Chaganty[2]

1. *University of Wisconsin, Whitewater, WI, USA*
2. *Old Dominion University, Norfolk, VA, USA*

Received: August 16, 2026 / Accepted: September 11, 2026 / Published: November 25 2026.

Abstract: In order to model over-dispersed count data with covariates, negative binomial regression model (NBR) is a commonly used approach. There are several parameterizations for NBR, and the two well-known models, negative binomial – 1 regression model (NBR1) and negative binomial – 2 regression model (NBR2), have been applied by Cameron and Trivedi (1998). This paper introduces several forms of doubly-inflated negative binomial (DINB) regression models which can be fitted to inflated counts of zero and k that also lead to overdispersion of count data. Many important characteristics such as expectations, moments, and Fisher's Information Matrix are investigated. Relative efficiencies have also been calculated comparing the maximum likelihood estimation and method of moments. Applications of several forms of DINB have been illustrated using count data of dental caries along with presenting the generalized doubly inflated theory.

Key words: dental caries, zero inflated Poisson, zero inflated negative binomial, doubly inflated negative binomial

1 Introduction

Researchers in both natural and social sciences use regression models to explain the frequency of events such as the number of deaths, or daily cases due to flu,

number of epileptic seizures observed over a specific duration, etc. It is reasonable to model such situations using a suitable discrete distribution such as commonly used Poisson distribution. However, in real life, there is an underlying hazard to count data that it results in over-dispersion or underdispersion of counts. If the variance is greater than the mean, it indicates overdispersion and vice versa if underdispersion. Determining whether and how this hazard varies with covariates, including treatment effect, is typically the aim of data analysis. The general problem of overdispersion has been considered in the statistical literature. Cameron and Trivedi (1998) provides an excellent summary of approaches for verifying and analyzing overdispersed count data.

There has been a need for researchers to create mixture models of the discrete distributions because traditional discrete distributions such as binomial, Poisson, geometric and negative binomials distributions were unable to adequately characterize different types of real data by themselves. In the existing literature, many have developed discrete distributions that claim to model over- and under-dispersed datasets. In recent years, there has been considerable interest in models for count data that allow for excess zeros as well as another count value, say k, particularly in the economic and medical literature. These models complement more conventional models for overdispersion that focus on accurately modeling the variance–mean relationship. For counts with inflated zeros, application areas are diverse and have included manufacturing defects (Lambert (1992)), patent applications (Crepon and Duguet (1997)), road safety (Miaou (1994)), medical consultations (Gurmu (1997)), use of recreational facilities (Gurmu (1996)) and sexual behavior (Heilbron (1994)). Fernandez (2022) provides a great summary and comparison of zero-inflated count models for the application of hospital length of stay. For additional k inflation counts, there are instances in hospital stay and dental caries data (Sheth et al. (2019), Sheth-Chandra (2011)); ER and pap smear data (Arora and Chaganty (2021)); work week hours for couples (Mathews (2022)) as well as Australian health care and Malaysian motor insurance (Faroughi (2017)).

There are a variety of models that has been recommended for over-dispersed count data where the extra variability is caused by excess zeros and also unobserved heterogeneity. However, there are missing gaps in literature regarding the robust estimation of the doubly inflated negative binomial (DINB) regression models, especially since the DINB distribution not only models the inflated zero and the inflated k for a count outcome variable. In this paper, we will examine closely at a three parameter as well as a four parameter doubly inflated negative binomial (DINB) distribution models with an application to dental data. We will also introduce generalized inflated probability distributions, especially presenting the three parameter doubly inflated Poisson (DIP) distribution model. We will investigate characteristics of distribution function and compare the model using the maximum likelihood estimation and moment estimation techniques with its relative efficiencies.

2 Methodology Section

2.1 Review of Negative Binomial Distribution

The Poisson regression model can be generalized by introducing an unobserved heterogeneity term for observation i. The subjects are assumed to differ randomly in a manner that is not fully accounted for by the observed covariates. This is formulated as:

$$\begin{aligned} E(Y_i|X_i = x_i, \tau_i) &= \mu_i \tau_i \\ &= \exp x_i'\beta + \epsilon_i, \end{aligned}$$

where the unobserved heterogeneity term τ_i is independent of the vector of the predictor variables x_i. Then the conditional distribution of Y_i on $X_i = x_i$ is Poisson with conditional mean and conditional variance $\mu_i\tau_i$:

$$f(Y_i = y_i|X_i = x_i, \tau_i) = \frac{\exp -\mu_i\tau_i(\mu_i\tau_i)^{y_i}}{y_i!}, y_i = 0, 1, 2, ...$$

The negative binomial distribution is derived as a gamma mixture of Poisson random variables (insert reference). Most real-life count data is overdispersed. Negative Binomial Distribution is commonly used to account for overdispersed data. The most commonly used Negative Binomial distribution is the NB2 model with the following distribution:

$$g(y|\mu,\eta) = \frac{\Gamma(y+\eta^{-1})}{\Gamma(y+1)\Gamma(\eta^{-1})}\left(\frac{\eta^{-1}}{\mu+\eta^{-1}}\right)^{\eta^{-1}}\left(\frac{\mu}{\mu+\eta^{-1}}\right)^{y},$$

where $\eta > 0$ and $y = 0, 1, 2, \ldots$. The mean and variance of NB2 model is μ and $\mu(1+\eta\mu)$, respectively. If we replace η^{-1} with $\eta^{-1}\mu$, then we get the NB1 model which has the following distribution:

$$g(y) = \frac{\Gamma(y+\mu\eta^{-1})}{\Gamma(y+1)\Gamma(\mu\eta^{-1})}\left(\frac{\eta^{-1}}{1+\eta^{-1}}\right)^{\mu\eta^{-1}}\left(\frac{1}{1+\eta^{-1}}\right)^{y},$$

where $\eta > 0$ and $y = 0, 1, 2, \ldots$. The mean and variance of NB1 model is μ and $\mu(1+\eta)$, respectively.

Cameron and Trivedi (1998) discussed the NBr parameterization where $\tau = \mu^{2-r}\eta^{-1}$ and $\xi = \mu^{r-1}\eta$, where $1 \leq r \leq 2$. For $r = 1$ and $r = 2$, Negative Binomial Distribution (τ, ξ) or NB (τ, ξ) is

$$g(y|\tau,\xi) = \frac{\Gamma(\tau+y)}{\Gamma(\tau)\Gamma(y+1)}\frac{\xi^{y}}{(1+\xi)^{\tau+y}}, \text{ for } y = 0, 1, 2, \ldots$$

where, τ is a real positive number and $\xi > 0$. Note that $\xi = 1/(\pi - 1)$ where π is a Bernoulli parameter. The mean is $\mu = \tau\xi$ and variance $\sigma^2 = \tau\xi(1+\xi)$. One could also consider the generalized NBr model in which case r must be estimated.

2.2 Doubly Inflated Negative Binomial (p, τ, ξ) Distribution

The Doubly Inflated Negative Binomial (DINB) (p, τ, ξ) probability model is:

$$f(y|p,\tau,\xi) = \begin{cases} p^2 + q^2 g(0|\tau,\xi), & \text{for } y = 0; \\ 2pq + q^2 g(k|\tau,\xi), & \text{for } y = k; \\ q^2 g(y|\tau,\xi), & \text{for } y = 1, 2, \ldots \neq k. \end{cases}$$

where $g(y|\tau,\xi)$ is the negative binomial distribution. As $p \to 0$, this model reduces to a NB (τ,ξ). Further, since τ is a function of η, if $\eta \to 0$, it will further reduce to an ordinary Poisson distribution. The mean and variance of DINB (p, τ, ξ) are

$$E(Y) = 2pqk + q^2\tau\xi$$
$$Var(Y) = 2pqk^2 + q^2\tau\xi(2+\xi) - (2pqk + q^2\tau\xi)^2$$

The Fisher Information Matrix for the distribution is given by $I_1 =$

$$\begin{pmatrix} I(p) & I(p,\tau) & I(p,\xi) \\ I(\tau,p) & I(\tau) & I(\tau,\xi) \\ I(\xi,p) & I(\xi,\tau) & I(\xi) \end{pmatrix}$$

The Cramer-Rao lower bounds for the variances of the unbiased estimators of parameters can be obtained from the diagonal elements of I^{-1}.

2.2.1 Doubly Inflated Negative Binomial (p_1, p_2, τ, ξ) Distribution

The second DINB model, i.e. Doubly Inflated Negative Binomial (DINB) (p_1, p_2, τ, ξ) distribution is

$$f(y|p_1, p_2, \tau, \xi) = \begin{cases} p_1 + p_3 g(0|\tau, \xi), & \text{for } y = 0; \\ p_2 + p_3 g(k|\tau, \xi), & \text{for } y = k; \\ p_3 g(y|\tau, \xi), & \text{for } y = 1, 2, \ldots \neq k. \end{cases}$$

As $p_2 \to 0$, this model reduces to a Zero-Inflated Negative Binomial (ZINB) distribution. Further, since τ is a function of η, if $\eta \to 0$, this distribution reduces further to a Zero-Inflated Poisson (ZIP) distribution. The mean and variance of DINB (p_1, p_2, τ, ξ) are

$$\begin{aligned} E(Y) &= p_2 k + p_3 \tau \xi \\ Var(Y) &= p_2 k^2 + p_3 \tau \xi (2 + \xi) - (p_2 k + p_3 \tau \xi)^2 \end{aligned}$$

The Fisher Information Matrix (FIM) for the distribution is given by $I_2 =$

$$\begin{pmatrix} I(p_1) & I(p_1, p_2) & I(p_1, \tau) & I(p_1, \xi) \\ & I(p_2) & I(p_2, \tau) & I(p_2, \xi) \\ & & I(\tau) & I(\tau, \xi) \\ & & & I(\xi) \end{pmatrix}$$

The Cramer-Rao lower bounds for the variances of the unbiased estimators of parameters can be obtained from the diagonal elements of I^{-1}.

2.3 Maximum Likelihood Estimation

Let $\boldsymbol{y} = (y_1, y_2, ..., y_n)$ be a vector of n independent responses where there are inflated zeros and k's. Also, we assume that corresponding to each y_i, a vector $\boldsymbol{x_i^T} = (x_{i1}, x_{i2}, ..., x_{ip})$ of covariates has been observed. In this case, the likelihood function of the observed data for the DINB (p, τ, ξ) is given by

$$L(p, \tau, \xi|\boldsymbol{y}) = \sum_{y_i=0} p^2 + q^2 g(0|\tau, \xi)$$
$$\sum_{y_i=k} 2pq + q^2 g(k|\tau, \xi), \text{for } y = k;$$
$$\sum_{o.w.} q^2 g(y|\tau, \xi)$$

Similarly, the likelihood function of the observed data for the DINB (p_1, p_2, τ, ξ) is given by

$$L(p_1, p_2, \tau, \lambda|\boldsymbol{y}) = \sum_{y_i=0} p_1 + p_3 g(0|\tau, \xi)$$
$$\sum_{y_i=k} p_2 + p_3 g(k|\tau, \xi), \text{for } y = k;$$
$$\sum_{y_i o.w.} p_3 g(y|\tau, \xi)$$

2.3.1 DINB (p, τ, ξ) Regression Model

Suppose that data consisting of count responses $y_i, i = 1, \ldots, n$, for each subject i are independently distributed as doubly inflated negative binomial counts with parameters (p_i, μ_i). The DINB (p, τ, ξ) Regression can be modeled under the assumptions

$$\text{logit}(p_i) = \boldsymbol{B_i\beta}, \text{ and } \qquad \log(\mu_i) = \boldsymbol{G_i\gamma}$$

where $\boldsymbol{G_i}, \boldsymbol{B_i}$ are subject-specific covariates, $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are regression parameters, and $\boldsymbol{\mu_i} = \tau_i\xi_i$. Here, $\boldsymbol{\gamma}$'s will have interpretation in terms of the binomial probability, p_i; and, $\boldsymbol{\beta}$'s have interpretation in terms of the NB mean, μ_i.

The log-likelihood function is

$$\begin{aligned}\ell(\boldsymbol{\beta},\boldsymbol{\gamma}) &= n_0 \log(\exp(2\boldsymbol{B_i\beta}) + g(0|\tau,\xi)) \\ &\quad + n_k \log(\exp(2\boldsymbol{B_i\beta}) + \boldsymbol{g(k|\tau,\xi)}) \\ &\quad + \sum_{\substack{\{i:\, y_i=y \\ y\neq 0,k\}}} n_j \log(g(y|\tau,\xi)) \\ &\quad - 2\sum_{j=1}^{m} n_j \log(1+\exp(2\boldsymbol{B_i\beta}))\end{aligned}$$

To obtain the ML estimates $(\widehat{\boldsymbol{\gamma}},\widehat{\boldsymbol{\beta}})$, we differentiate the log-likelihood function with respect to the parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$, and set the score functions equal to zero. Thus, we solve

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} = 0 \qquad \frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \boldsymbol{\beta}} = 0$$

Taking the second-order derivatives with respect to the regression parameters, we obtain the Hessian Matrix. The standard errors for the parameter estimates are obtained by taking the inverse of the negative Hessian Matrix calculated at the ML estimates for the regression parameters.

2.4 DINB (p_1, p_2, τ, ξ) Regression Model

Suppose that data consisting of count responses $y_i, i = 1, \ldots, n$, for each subject i are independently distributed as doubly inflated negative binomial counts with parameters $(p_{1i}, p_{2i}, \tau_i, \xi_i)$. The DINB (p_1, p_2, τ, ξ) Regression can be modeled under the assumptions

$$\begin{aligned}\text{logit}(p_{1i}) &= \boldsymbol{A_i\alpha}, \\ \text{logit}(p_{2i}) &= \boldsymbol{B_i\beta}, \\ \log(\mu_i) &= \boldsymbol{G_i\gamma}\end{aligned}$$

where $\boldsymbol{G_i}, \boldsymbol{B_i}$ are subject-specific covariates, $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are regression parameters, and $\boldsymbol{\mu_i} = \tau_i \xi_i$ Here, $\boldsymbol{\alpha}$s and $\boldsymbol{\beta}$s will have interpretation in terms of the binomial probabilities and, $\boldsymbol{\gamma}$'s have interpretation in terms of the NB mean, μ_i.

2.4.1 Maximum Likelihood Estimation

The log-likelihood function is

$$\begin{aligned}
\ell(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = n_0 \log(&\exp(\boldsymbol{A_i\alpha}) + g(0|\tau, \xi) + \\
&(1 - g(0|\tau, \xi)) \exp(\boldsymbol{B_i\beta} + \boldsymbol{A_i\alpha})) \\
&+ n_k \log(\exp(\boldsymbol{A_i\alpha}) + \\
&g(k|\tau, \xi) + (1 - g(k|\tau, \xi)) \exp(\boldsymbol{B_i\beta} + \boldsymbol{A_i\alpha})) + \\
&\sum_{\substack{\{i: \\ y \neq 0,k\}}}^{y_i = y} n_j \log(g(y|\tau, \xi)(1 - \exp(\boldsymbol{B_i\beta} + \boldsymbol{A_i\alpha}))) \\
&+ \sum n_j \log((1 + \exp(\boldsymbol{A_i\alpha}))(1 + \exp(\boldsymbol{B_i\beta})))
\end{aligned}$$

To obtain the ML estimates $(\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\beta}})$, we differentiate the log-likelihood function with respect to the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ and set the score functions equal to zero. Thus, we solve the following first-order conditions

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}} &= 0 \\
\frac{\partial \ell(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta})}{\partial \boldsymbol{\gamma}} &= 0 \\
\frac{\partial \ell(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= 0
\end{aligned}$$

Taking the second-order derivatives with respect to the regression parameters, we obtain the Hessian Matrix. The standard errors for the parameter estimates are obtained by taking the inverse of the negative Hessian Matrix calculated at the ML estimates for the regression parameters.

3 Generalized Doubly Inflated Regression Theory

When there are situations where a major source of overdispersion is a preponderence of zero counts as well as counts of integer k. In such cases, one can use doubly inflated estimation methods. The doubly inflated estimation techniques permit the researcher to answer two questions that pertain to low base rate-dependent variables (a) what predicts whether or not the behavior occurs and (b) if the behavior occurs, what predicts the frequency of occurrence of the inflated counts ([**?**])? Additionally, doubly inflated models have a statistical advantage to standard Poisson and zero-inflated Poisson models in that they model the preponderence of zeros and k counts simultaneously utilizing binomial distribution or dummy variables for the first component and discrete distributions such as binomial, negative binomial, multinomial and so on for the second component. The probability mass function of Y, where k is known, is:

$$P_\omega(Y=y) = \begin{cases} \phi + (1-\phi-\psi)P(Y=0)\omega, & \text{for } y=0; \\ \psi + (1-\phi-\psi)P(Y=k)\omega, & \text{for } y=k; \\ (1-\phi-\psi)P(Y=y)\omega, & \text{otherwise.} \end{cases}$$

3.1 The DIP (p_1, p_2, λ) Model

The marginal distribution of Y distributed as the DIP (p_1, p_2, λ) model, where k is known, is

$$P(Y = y) = \begin{cases} p_1 + p_3 \exp(-\lambda), & \text{for } y = 0; \\ p_2 + p_3 \dfrac{\exp(-\lambda)\lambda^k}{k!}, & \text{for } y = k; \\ p_3 \dfrac{\exp(-\lambda)\lambda^y}{y!}, & \text{for } y = 1, 2, ... \neq k. \end{cases}$$

When $p_2 = 0$, DIP (p_1, p_2, λ) Model reduces to the Zero-Inflated Poisson Model. The first four moments of DIP (p_1, p_2, λ) Model are given by

$$\begin{aligned} E(Y) &= p_2 k + p_3 \lambda \\ E(Y^2) &= p_2 k^2 + p_3(\lambda^2 + \lambda) \\ E(Y^3) &= p_2 k^3 + p_3(\lambda + 3\lambda^2 + \lambda^3) \\ E(Y^4) &= p_2 k^4 + p_3(\lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4) \end{aligned}$$

The mean μ and the variance σ^2 are

$$\begin{aligned} \mu = E(Y) &= p_2 k + p_3 \lambda \\ \sigma^2 = Var(Y) &= p_2 k^2 + p_3(\lambda^2 + \lambda) - \mu^2 \\ &= \mu - \mu^2 + p_3 \lambda^2 - p_2(k^2 - k) \end{aligned}$$

Suppose data consists of $y_i, i = 1, \ldots, n$, independent distributed as doubly inflated poisson counts with parameters $(p_{1i}, p_{2i}, \lambda_i)$.Then, the DIP (p_1, p_2, λ) Regression can be modeled using maximum likelihood estimation method under

the assumptions

$$\text{logit}(p_{1i}) = \boldsymbol{A_i\alpha}, \qquad \text{logit}(p_{2i}) = \boldsymbol{G_i\gamma}, \text{ and}$$
$$\log(\lambda_i) = \boldsymbol{B_i\beta}$$

where $\boldsymbol{A_i}, \boldsymbol{G_i}$, and $\boldsymbol{B_i}$ are subject-specific covariates; and, $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\beta}$ are regression parameters. Also, $\boldsymbol{\alpha}$'s will have interpretation in terms of the additional proportions of zeros, p_{1i}; $\boldsymbol{\gamma}$'s will have interpretation in terms of the additional proportions of ks, p_{2i}; and, $\boldsymbol{\beta}$'s have interpretation in terms of the Poisson mean.

4 Simulations

The relative efficiency of moment estimator of p_1 with respect to the maximum likelihood estimator is

$$e(\widetilde{p_1}, \widehat{p_1}) = \frac{\sigma^2(\widehat{p_1})}{\sigma^2(\widetilde{p_1})}$$

The relative efficiency of moment estimator of p_2 with respect to the maximum likelihood estimator is

$$e(\widetilde{p_2}, \widehat{p_2}) = \frac{\sigma^2(\widehat{p_2})}{\sigma^2(\widetilde{p_2})}$$

The relative efficiency of moment estimator of λ with respect to the maximum likelihood estimator is

$$e(\widetilde{\lambda}, \widehat{\lambda}) = \frac{\sigma^2(\widehat{\lambda})}{\sigma^2(\widetilde{\lambda})}$$

ML estimators are more efficient than MOM estimators.

Table 1: Asymptotic Relative Efficiencies of p_1

		p_2				
λ	p_1	0.1	0.3	0.5	0.7	0.9
4	0.1	0.0369	0.0456	0.0593	0.0839	0.1320
	0.2	0.0738	0.0952	0.1310	0.2012	
	0.3	0.1186	0.1588	0.2303		
	0.4	0.1740	0.2429	0.3745		
	0.5	0.2445	0.3585			
	0.6	0.3371	0.5247			
	0.7	0.4637				
	0.8	0.6469				
	0.9	0.8625				

Table 2: Asymptotic Relative Efficiencies of p_2

		p_2				
λ	p_1	0.1	0.3	0.5	0.7	0.9
4	0.1	0.0015	0.0023	0.0030	0.0034	0.0033
	0.2	0.0017	0.0028	0.0039	0.0050	
	0.3	0.0020	0.0035	0.0055	0.0068	
	0.4	0.0024	0.0047	0.0089		
	0.5	0.0030	0.0070	0.0110		
	0.6	0.0041	0.0128			
	0.7	0.0062	0.0138			
	0.8	0.0124				
	0.9	0.0515				

Table 3: Asymptotic Relative Efficiencies of λ

		p_2				
λ	p_1	0.1	0.3	0.5	0.7	0.9
4	0.1	0.0235	0.0227	0.0207	0.0158	0.0105
	0.2	0.0254	0.0245	0.0217	0.0126	
	0.3	0.0285	0.0271	0.0221	0.0156	
	0.4	0.0325	0.0303	0.0197		
	0.5	0.0381	0.0337	0.0257		
	0.6	0.0460	0.0349			
	0.7	0.0579	0.0499			
	0.8	0.0767				
	0.9	0.0796				

Table 4: MLE Parameter Estimates for Raw Counts

	DINB Model 1		DINB Model 2	
Parameter	Est.	Std.Err.	Est.	Std.Err.
p	0.077	0.014	-	-
p_1	–	-	0.183	0.015
p_2	–	-	0.262	0.021
λ	–	-	–	–
τ	2.888	0.424	84.282	47.149
ξ	0.622	0.095	0.033	0.0166
NegLL	1701.135	-	1687.288	-

5 Results

In dental epidemiology, the DMFT-index is an important indicator and measures for the dental status of a person. It is a count number of decayed, missing, and filled teeth. Consider an example data for 1013 school children of age 7 . Only the eight deciduous molars are considered. Thus, the smallest possible value of the DMFT-index is 0 and the largest is 8. DMFT was measured at the beginning of the study and after a year. Thus, we obtain a count for the change in DMFT, $\delta DMFT$. The zero count corresponds to those children showing no improvement and/or consistent dental care. The one count corresponds to those children showing improvement in one cavity, and so on. The aim of the study is to compare methods of six schools of treatments: oral health education, enrichment of the school diet with rice bran, mouthwash with 0.2% of NaF solution, oral hygiene, all of the four treatments and control group. Race and gender were also obtained for each of the school children.

As seen in the table below, DINB distribution model 1 has almost similar log-likelihood value as the Doubly Inflated Regression Model ($NegLL = 1700.094$). Of course, adding a parameter in the second model would show a substantial decrease in the NegLL.

Table 5 represents the parameter estimates for the ZIP regression model on dental caries data. Exponentiating the parameter estimates for those variables gives us the odds ratio. The progression in NegLL from ZIP to DIP to DINB

Table 5: MLE Parameter Estimates for ZIP on Dental Data

Parameters	Estimates	Std Err	p-value
Logit Link for p			
`Intercept`	2.857	1.281	0.026
`Trtmt(Educ)`	0.461	1.684	0.784
`Trtmt(Enrich)`	−0.116	1.614	0.943
`Trtmt(Rinse)`	−1.001	0.792	0.207
`Trtmt(Hygiene)`	−0.237	1.275	0.853
`Trtmt(All)`	−0.827	0.959	0.388
`Gender(M)`	0.184	0.627	0.769
`Eth(White)`	−0.416	1.325	0.754
`Eth(Black)`	12.230	3.191	0.001
Log Link for λ			
`Intercept`	0.713	0.116	0.001
`Trtmt(Educ)`	−0.232	0.108	0.032
`Trtmt(Enrich)`	−0.082	0.116	0.518
`Trtmt(Rinse)`	−0.214	0.106	0.044
`Trtmt(Hygiene)`	−0.262	0.137	0.055
`Trtmt(All)`	−0.448	0.129	0.001
`Gender(M)`	0.107	0.074	0.152
`Eth(White)`	0.134	0.124	0.279
`Eth(Black)`	−0.273	0.111	0.014
`NegLL`	1713.094		

Model 1 to DINB Model 2, with a further drop to 1656.6 under the log link for tau, is exactly the head-to-head comparison that one needs to see with the progression of the models by addig a parameter, and it cleanly makes the case for the negative binomial extension along with the modeling for the inflated k.

4. Conclusions Usually, overdispersion and k-inflated values are seen in the real-life count data. In the presence of overdispersion, negative binomial is a commonly used distribution. Therefore, when there is an inflation of zero and k count values, a doubly inflated negative binomial model is more adequate for the modeling of the outcome variable DMFT and any other count values. This can be seen from the significant reduction of the negative log-likelihood values. Such models can account for increased variability due to overdispersion as well as inflated values. Furthermore, these models are highly reliable as they are more efficient using the maximum likelihood estimation (MLE) methods as compared to the method of moments (MOM) estimation methods.

Table 6: MLE Parameter Estimates for DIP on Dental Data

Parameters	Estimates	Std Err	p-value
Logit Link for p			
`Intercept`	1.484	0.677	0.029
`Trtmt(Educ)`	1.161	0.807	0.151
`Trtmt(Enrich)`	0.847	0.772	0.273
`Trtmt(Rinse)`	−0.928	0.544	0.088
`Trtmt(Hygiene)`	−0.760	0.572	0.185
`Trtmt(All)`	−1.195	0.543	0.028
`Gender(M)`	0.077	0.236	0.774
`Eth(White)`	−0.181	0.354	0.609
`Eth(Black)`	−0.766	0.035	0.035
Log Link for λ			
`Intercept`	0.878	0.190	0.001
`Trtmt(Educ)`	−0.451	0.169	0.008
`Trtmt(Enrich)`	−0.254	0.180	0.158
`Trtmt(Rinse)`	−0.003	0.159	0.983
`Trtmt(Hygiene)`	−0.041	0.200	0.838
`Trtmt(All)`	−0.097	0.167	0.561
`Gender(M)`	0.113	0.085	0.184
`Eth(White)`	0.175	0.109	0.111
`Eth(Black)`	0.060	0.137	0.659
`NegLL`	1700.094		

Table 7: MLE Parameter Estimates for DINB on Dental Data

Parameters	Estimates	Std Err	p-value
Logit Link for p			
`Intercept`	−3.088	0.808	0.0001
`Trtmt(Educ)`	0.502	0.706	0.238
`Trtmt(Enrich)`	0.206	0.738	0.390
`Trtmt(Rinse)`	1.130	0.689	0.051
`Trtmt(Hygiene)`	1.315	0.680	0.026
`Trtmt(All)`	1.557	0.643	0.007
`Gender(M)`	−0.411	0.396	0.149
`Eth(White)`	−0.975	0.638	0.063
`Eth(Black)`	1.825	0.421	0.001
`NegLL`	1682.254		

Acknowledgment Dr. Sheth acknowledges the support Departments of Mathematics and Computer Science, University of Wisconsin - Whitewater, WI, USA as well as Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA." Dr. Chaganty acknowledges the support of Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA.

Dr. Chaganty acknowledges the following individuals for their support: Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA.

References

1. Cameron, A. C., and Trivedi, P. K. (1998), *Regression Analysis of Count Data*, New York: Cambridge.

2. Lambert, D. (1992), "Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing", *Technometrics*, **34**, 1-14 .

3. Fernandez. G.A. and Vatcheva, K.P. (2022). A comparison of statistical methods for modeling count data with an application to hospital length of stay. BMC Medical Research Methodology, 22(1), 211.

4. Sheth-Chandra, M., Chaganty, N.R., Sabo, R.T. (2019). "A Doubly-Inflated Poisson Distribution and Regression Model." In: Diawara, N. (eds) Modern Statistical Methods for Spatial and Multivariate Data. STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health. Springer, Cham. https://doi.org/10.1007/978-3-030-11431-2

5. Sheth-Chandra, M. (2011). The doubly inflated Poisson and related regression models. Old Dominion University. Dissertation.

6. Crepon, B. and Duguet, E. (1997). "Estimating the Innovation Function from Patent Numbers: GMM on Count Panel Data." Journal of Applied Econometrics, 12(3), pp. 243 - 263.

7. Arora, M., Rao Chaganty, N., & Sellers, K. F. (2021). "A flexible regression model for zero- and k-inflated count data." Journal of Statistical Computation and Simulation, 91(9), 1815–1845.

8. Miaou, S. (1994). "The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regression." Accident Analysis and Prevention, 26(4), 471-482.

9. Gurmu, S. (1997). "Semi-parametric estimation of hurdle regression models with an application to medical utilization." Journal of Applied Econometrics, 225 - 242.

10. Gurmu, Shiferaw and Trivedi, Pravin, (1996), "Excess Zeros in Count Models for Recreational Trips", Journal of Business & Economic Statistics, 14, issue 4, p. 469-77.

11. Heilbron, D. "Zero-Altered and other Regression Models for Count Data with Added Zeros", Biometrical Journal, 531-547.

12. Mathews, J., Bhattacharya, S., Sen, S., and Das, I. "Multiple inflated negative binomial regression for correlated multivariate count data", Dependence Modelling, 290-307.

13. Faroughi, P., Karimi, M. S., Ismail, N., & Karimi, A. (2017). Estimation of Count Data using Bivariate Negative Binomial Regression Models. Quarterly Journal of Quantitative Economics, Summer 2017, 14(2): 143 -166.