Pak.j.stat.oper.res. Vol.21 No.4 2025 pp 619-630 DOI: http://dx.doi.org/10.18187/pjsor.v21i4.4566

Pakistan Journal of Statistics and Operation Research

The Negative Binomial-Bilal Distribution: Regression Model and Applications to Health Care Data

Yupapin Atikankul[1], Chanakarn Jornsatian[2*]

*Corresponding author

1. Department of Mathematics and Statistics, Rajamangala University of Technology Phra Nakhon, Bangkok, Thailand, yupapin.a@rmutp.ac.th
2. Department of Mathematics, Faculty of Science, Srinakharinwirot University, Bangkok, Thailand, chanakarnj@g.swu.ac.th

Abstract

In health care research, overdispersion often arises in count data. The Poisson distribution is a traditional distribution for modeling count data. However, it cannot handle overdispersed count data. This article introduces a new count distribution for overdispersed data. Statistical properties and a multivariate version of the proposed distribution are derived. Two parameter estimation methods are discussed by the maximum likelihood method and Bayesian approach. A simulation study is conducted to assess the performance of the estimators. A regression model based on the proposed distribution is constructed. Finally, two health care applications are analyzed to show the potential of the proposed distribution and its associated regression model.

Key Words: Count Data; Regression Model; Bayesian Approach; Maximum Likelihood Estimation.

Mathematical Subject Classification: 60E05, 62JO5.

1. Introduction

Counting data frequently appear in health care studies. The well-known model for analyzing count data is the Poisson distribution with the variance equal to the mean, equidispersion. Unfortunately, a practical problem of count data is the variance larger than the mean, overdispersion. Mixtures distributions can be applied to deal with overdispersion in count data. Recently, various mixed negative binomial (NB) distributions have been proposed to model count data, such as the NB-Pareto distribution (Meng et al., 1999), the NB-inverse Gaussian distribution (Gómez-Déniz et al., 2008), the NB-Lindley distribution (Zamani and Ismail, 2010), the NB-beta exponential distribution (Pudprommarat et al., 2012), the NB-generalized exponential distribution (Aryuyuen and Bodhisuwan, 2013), the NB-Sushila distribution (Yamrubboon et al., 2017), the NB-weighted Garima distribution (Bodhisuwan and Saengthong, 2020), the NB-reciprocal inverse Gaussian distribution (Hassan et al., 2021), the four parameters NB-Lindley distribution (Tajuddin et al., 2022), and the NB-generalized Lindley distribution (Atikankul et al., 2022). Although these mixed NB distributions are more flexible than the Poisson and NB distributions, they are not developed for modeling count data with covariates.

A new lifetime distribution was proposed by Abd-Elrahman (2013), called the Bilal distribution. Its probability density function (pdf) with scale parameter θ is

$$g(x) = \frac{6}{\theta} e^{-\frac{2x}{\theta}} \left(1 - e^{-\frac{x}{\theta}}\right),$$

for $x > 0$ and $\theta > 0$.

The Bilal distribution has a closed-form expression for the pdf. The distribution is unimodel. Moreover, two real lifetime data sets were fitted with the Bilal distribution. The distribution provided good fits for real-life data.

In this paper, we propose a new count distribution for modeling count data. The distibution is a mixed NB distribution by using the Bilal distribution to a mixing distribution. We call the negative binomial-Bilal (NBB) distribution. The two parameters of the NB distribution are r, number of successes; and p, probability of success. We assume $p = e^{-\Lambda}$ and $\Lambda \sim$Bilal(θ). The proposed distribution with parameters r and θ is unimodal. Parameter estimation is addressed by the maximum likelihood (ML) and Bayesian approaches. The Bayesian approach is used to construct the NBB regression model. We apply the NBB distribution and its regression model to overdispersed health care data sets. The results indicate that the proposed model provides better fits than classical models.

The rest of this article is organized as follows. The proposed distribution and its properties are presented in Section 2. In Section 3, parameter estimation is derived by the ML and Bayesian approaches. A simulation study is presented in order to evaluate estimators in Section 4. In Section 5, a new regression model based on the proposed distribution is developed. In Section 6, health care applications are analyzed to show the utility of the new distribution and its regression model. Lastly, the article is concluded in Section 7.

2. The Negative Binomial-Bilal Distribution

In this section, we define the definition of the NBB distribution. Moreover, shape and basic properties of the distribution are presented.

The probability mass function (pmf) of the NB distribution is

$$f(y) = \binom{r+y-1}{y} p^r (1-p)^y, \tag{1}$$

for $y = 0, 1, \ldots, r > 0$ and $0 < p < 1$.

Suppose Λ be a Bilal random variable, then the pdf of Λ is

$$g(\lambda) = \frac{6}{\theta} e^{-\frac{2\lambda}{\theta}} \left(1 - e^{-\frac{\lambda}{\theta}}\right),$$

for $\lambda > 0$ and $\theta > 0$.

The moment generating function (mgf) of the Bilal(θ) distribution is defined by

$$M_\Lambda(t) = \frac{6}{(3-\theta t)(2-\theta t)}. \tag{2}$$

The NBB distribution is the mixture of the NB and Bilal distributions. Its definition can be define as follows

Definition 2.1. *A random variable Y follows the NBB distribution if it admits the stochastic representation*

$$\begin{aligned} Y|\Lambda &\sim \mathit{NB}(r, p = e^{-\Lambda}), \\ \Lambda &\sim \mathit{Bilal}(\theta), \end{aligned}$$

for $r > 0$ and $\theta > 0$. The unconditional distribution of Y is denoted by NBB(r, θ) and its pmf is given by Theorem 2.1.

Theorem 2.1. *If $Y \sim NBB(r, \theta)$, then the pmf of Y is*

$$f(y) = \binom{r+y-1}{y} \sum_{j=0}^{y} \binom{y}{j} (-1)^j \frac{6}{(3+\theta(r+j))\,(2+\theta(r+j))},$$

for $y = 0, 1, \ldots$, $r > 0$ and $\theta > 0$.

Proof. If $Y|\Lambda \sim \mathrm{NB}(r, e^{-\Lambda})$ and $\Lambda \sim \mathrm{Bilal}(\theta)$, then

$$f(y) = \int_0^\infty f(y|\lambda) g(\lambda) d\lambda. \tag{3}$$

Since $f(y|\lambda) = \binom{r+y-1}{y} e^{-\lambda r}(1-e^{-\lambda})^y$ and $(1-e^{-\lambda})^y = \sum_{j=0}^{y} \binom{y}{j} (-1)^j e^{-\lambda j}$,

$$f(y|\lambda) = \binom{r+y-1}{y} \sum_{j=0}^{y} \binom{y}{j} (-1)^j e^{-\lambda(r+j)}. \tag{4}$$

Substituting Equation (4) into Equation (3), we get

$$\begin{aligned} f(y) &= \binom{r+y-1}{y} \sum_{j=0}^{y} \binom{y}{j} (-1)^j \int_0^\infty e^{-\lambda(r+j)} g(\lambda) d\lambda \\ &= \binom{r+y-1}{y} \sum_{j=0}^{y} \binom{y}{j} (-1)^j M_\Lambda(-(r+j)). \end{aligned} \tag{5}$$

To obtain the pmf of the NBB distribution, we substitute the mgf of the Bilal distriubtion in Equation (2) with $t = -(r+j)$ into Equation (5). Hence, the pmf of the NBB distribution is given by

$$f(y) = \binom{r+y-1}{y} \sum_{j=0}^{y} \binom{y}{j} (-1)^j \frac{6}{(3+\theta(r+j))\,(2+\theta(r+j))}.$$

□

Figure 1 presents some pmf plots of the NBB distribution for different parameter values of r and θ. We can see that the pmf of the NBB distribution is unimodal.

Proposition 2.1. *Let Y denote a NBB random vaiable, then the kth factorial moment of Y is*

$$\mu_{[k]}(Y) = \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \frac{6}{(3-\theta(k-j))(2-\theta(k-j))},$$

for $k = 1, 2, \ldots$.

Proof. If $Y|\Lambda \sim \mathrm{NB}(r, e^{-\Lambda})$ and $\Lambda \sim \mathrm{Bilal}(\theta)$, then the *k*th factorial moment of the NBB distribution is given by

$$\mu_{[k]}(Y) = \mathbb{E}_\Lambda[\mu_{[k]}(Y|\Lambda)].$$

The *k*th factorial moment of the $\mathrm{NB}(r, e^{-\lambda})$ is

$$\mu_{[k]}(Y|\Lambda) = \frac{\Gamma(r+k)}{\Gamma(r)} \frac{(1-e^{-\lambda})^k}{e^{-\lambda k}}. \tag{6}$$

Thus

$$\mu_{[k]}(Y) = \mathbb{E}_\Lambda \left[\frac{\Gamma(r+k)}{\Gamma(r)} \frac{(1-e^{-\lambda})^k}{e^{-\lambda k}} \right] = \frac{\Gamma(r+k)}{\Gamma(r)} \mathbb{E}_\Lambda \left(e^{\lambda} - 1 \right)^k.$$

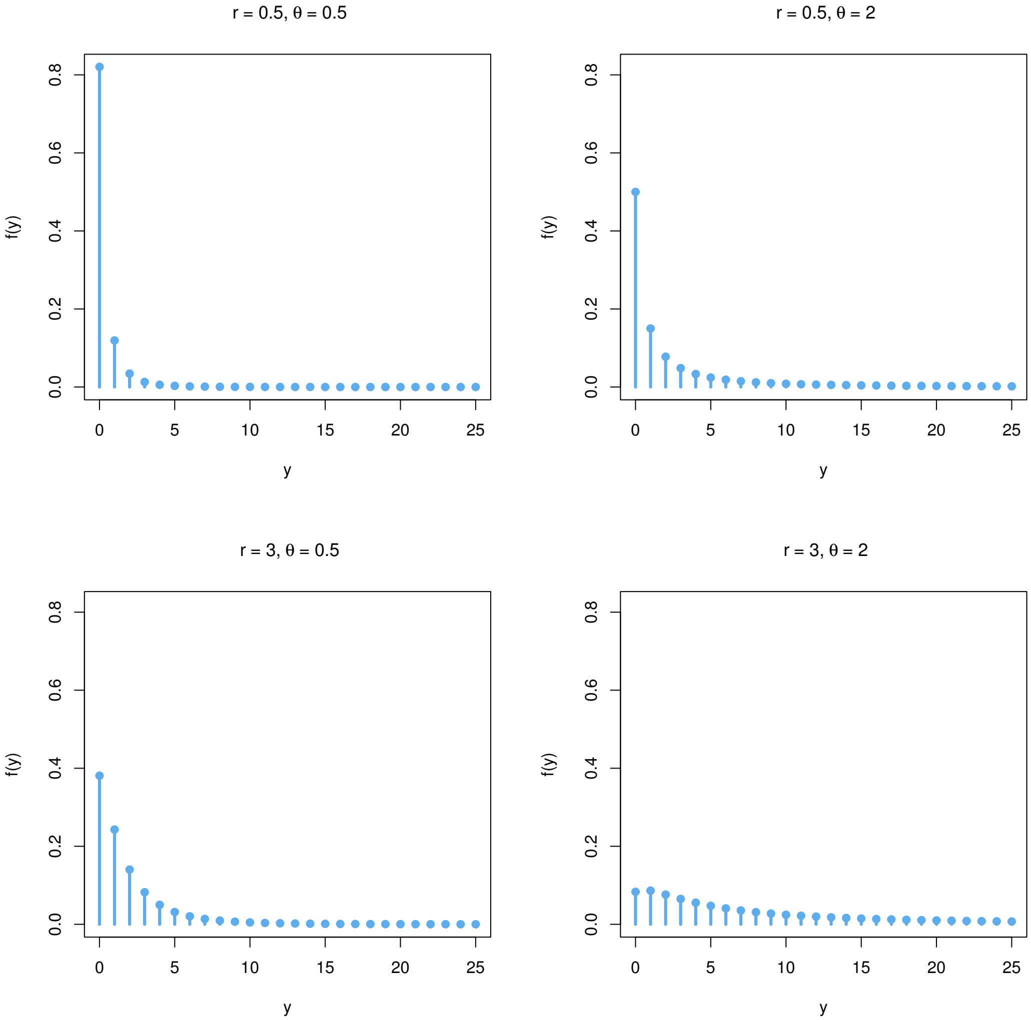

Figure 1: Pmf plots of the NBB distribution.

By the bimomial expansion of $(e^{-\lambda}-1)^k$, $\mu_{[k]}(Y)$ is defined by

$$\mu_{[k]}(Y) = \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \mathbb{E}_\Lambda(\mathrm{e}^{\lambda(k-j)})$$

$$= \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^{k} \binom{k}{j} (-1)^j M_\Lambda(k-j). \tag{7}$$

The mgf of the Bilal distribution in Equation (2) with $t = k - j$ is substituted for Equation (7). Thus, $\mu_{[k]}(Y)$ is

$$\mu_{[k]}(Y) = \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \frac{6}{(3-\theta(k-j))(2-\theta(k-j))}.$$

□

If Y be a random variable with the NBB distribution, the first and second moments about the origin of Y are

$$\begin{aligned}\mathbb{E}(Y) &= r(M_\Lambda(1)-1),\\ \mathbb{E}(Y^2) &= (r+r^2)M_\Lambda(2)-(r+2r^2)M_\Lambda(1)+r^2.\end{aligned}$$

The variance and the index of dispersion (ID) of Y are

$$\begin{aligned}\mathrm{V}(Y) &= (r+r^2)M_\Lambda(2)-\left[r+r^2M_\Lambda(1)\right]M_\Lambda(1),\\ \mathrm{ID}(Y) &= \frac{(r+1)M_\Lambda(2)-M_\Lambda(1)-rM_\Lambda^2(1)}{M_\Lambda(1)-1},\end{aligned}$$

where $M_\Lambda(t)$ is defined by Equation (2).

Theorem 2.2. *The pmf of the NBB distribution can be calculated by the recursive formula*

$$f_r(y)=\frac{r+y-1}{y}\left[f_r(y-1)-\frac{r}{r+y-1}f_{r+1}(y-1)\right],$$

for $y=1,2,\ldots$.

Proof. The pmf of the NB distribution can be expressed as

$$f(y|\lambda)=\binom{r+y-1}{y}e^{-\lambda r}(1-e^{-\lambda})^y,$$

for $y=0,1,\ldots$.

The simple recursion is given by

$$\frac{f(y|\lambda)}{f(y-1|\lambda)}=\frac{r+y-1}{y}\left(1-e^{-\lambda}\right).$$

Thus

$$f(y|\lambda)=f(y-1|\lambda)\frac{r+y-1}{y}\left(1-e^{-\lambda}\right), \tag{8}$$

for $y=1,2,\ldots$.

By the definition of the NBB distribution and Equation (8), then

$$\begin{aligned}f_r(y) &= \int_0^\infty f(y|\lambda)g(\lambda)d\lambda\\ &= \frac{r+y-1}{y}\int_0^\infty\left(1-e^{-\lambda}\right)f(y-1|\lambda)g(\lambda)d\lambda\\ &= \frac{r+y-1}{y}\left[f_r(y-1)-\int_0^\infty e^{-\lambda}f(y-1|\lambda)g(\lambda)d\lambda\right].\end{aligned}$$

Since

$$\begin{aligned}\int_0^\infty e^{-\lambda}f(y-1|\lambda)g(\lambda)d\lambda &= \int_0^\infty e^{-\lambda}\binom{r+y-2}{y-1}e^{-\lambda r}(1-e^{-\lambda})^{y-1}g(\lambda)d\lambda\\ &= \int_0^\infty\frac{r}{r+y-1}\binom{r+1+y-2}{y-1}e^{-\lambda(r+1)}(1-e^{-\lambda})^{y-1}g(\lambda)d\lambda\\ &= \frac{r}{r+y-1}f_{r+1}(y-1),\end{aligned}$$

the recursive equation of the NBB is given by

$$f_r(y) = \frac{r+y-1}{y}\left[f_r(y-1) - \frac{r}{r+y-1} f_{r+1}(y-1)\right].$$

□

Definition 2.2. *A multivariate NBB distribution, **Y** = $(Y_1, \ldots, Y_d)^T$ can be defined by*

$$Y_i|\Lambda \sim NB(r_i, e^{-\Lambda}),$$

for $i = 1, 2, \ldots, d$ are independent, and

$$\Lambda \sim \textit{Bilal}(\theta).$$

The same arguments in Theorem 2.1 is mentioned, thus the joint pmf of the NBB distribution is given by Theorem 2.3.

Theorem 2.3. *The pmf of the multivariate NBB distribution is*

$$f(y_1, \ldots, y_d) = \prod_{i=1}^{d}\binom{r_i+y_i-1}{y_i}\sum_{j=0}^{\tilde{y}}\binom{\tilde{y}}{j}(-1)^j\frac{6}{(3+\theta(\tilde{r}+j))\,(2+\theta(\tilde{r}+j))},$$

for $y_1, \ldots, y_d = 0, 1, \ldots; \theta > 0; r_1, \ldots, r_d > 0$; $\tilde{r} = \sum_{i=1}^{d} r_i$ and $\tilde{y} = \sum_{i=1}^{d} y_i$.

3. Parameter Estimation

In this section, the parameter estimation methods of the NBB distribution are discussed by the ML method and the Bayesian approach.

3.1. Maximum Likelihood Estimation

Suppose $\boldsymbol{y} = (y_1, \ldots, y_n)^T$ be a random variable of n observations from the NBB distribution. The likelihood function for the vector of parameters $\Theta = (r, \theta)^T$ is given by

$$L(\Theta|\boldsymbol{y}) = \prod_{i=1}^{n}\binom{r+y_i-1}{y_i}\sum_{j=0}^{y_i}\binom{y_i}{j}(-1)^j\frac{6}{(3+\theta(r+j))\,(2+\theta(r+j))}.$$

The corresponding log-likelihood function is

$$\begin{aligned}\log L(\Theta|\boldsymbol{y}) &= \sum_{i=1}^{n}[\log\Gamma(r+y_i) - \log\Gamma(y_i+1) - \log\Gamma(r)] \\ &\quad + \sum_{i=1}^{n}\log\left[\sum_{j=0}^{y_i}\binom{y_i}{j}(-1)^j\frac{6}{(3+\theta(r+j))\,(2+\theta(r+j))}\right].\end{aligned}$$

Taking the first partial derivative of the log-likelihood with respect to each parameter, we obtain

$$\begin{aligned}\frac{\partial \log L(\Theta|\boldsymbol{y})}{\partial r} &= \sum_{i=1}^{n}\psi(r+y_i)-n\psi(r)\\ &\quad -\sum_{i=1}^{n}\left[\frac{\sum_{j=0}^{y_i}\binom{y_i}{j}(-1)^j\left(\frac{6\theta(2r\theta+2j\theta+5)}{(3+\theta(r+j))^2\,(2+\theta(r+j))^2}\right)}{\sum_{j=0}^{y_i}\binom{y_i}{j}(-1)^j\frac{6}{(3+\theta(r+j))\,(2+\theta(r+j))}}\right],\\ \frac{\partial \log L(\Theta|\boldsymbol{y})}{\partial \theta} &= -\sum_{i=1}^{n}\left[\frac{\sum_{j=0}^{y_i}\binom{y_i}{j}(-1)^j\left(\frac{6(r+j)(2r\theta+2j\theta+5)}{(3+\theta(r+j))^2\,(2+\theta(r+j))^2}\right)}{\sum_{j=0}^{y_i}\binom{y_i}{j}(-1)^j\frac{6}{(3+\theta(r+j))\,(2+\theta(r+j))}}\right],\end{aligned}$$

where $\psi(\cdot)$ is the digamma function.

These equations are nonlinear. Since they cannot be explicitly solved, the numerical methods can be solved them. In this article, the optimx function of the optimx package (Nash and Varadhan, 2011) in the R programming language (R Core Team, 2023) is applied.

3.2. Bayesian Approach

Let $\pi(\Theta)$ be the set of prior distribution and $L(\boldsymbol{y}|\Theta)$ is the likelihood function. Then the joint distribution of Θ and $\boldsymbol{y}$ is

$$\pi(\boldsymbol{y},\Theta)=L(\boldsymbol{y}|\Theta)\pi(\Theta).$$

The marginal distribution of $\boldsymbol{y}$ is

$$\pi(\boldsymbol{y})=\int_{\Theta}L(\boldsymbol{y}|\Theta)\pi(\Theta)d(\Theta).$$

The joint posterior distribution is expressed by

$$\pi(\Theta|\boldsymbol{y})=\frac{L(\boldsymbol{y}|\Theta)\pi(\Theta)}{\int_{\Theta}L(\boldsymbol{y}|\Theta)\pi(\Theta)d(\Theta)}.$$

Since the denominator is normalization constant, it is not used in determining the posterior distribution. Then

$$\pi(\Theta|\boldsymbol{y})\propto L(\boldsymbol{y}|\Theta)\pi(\Theta).$$

In this paper, the Gamma(0.001, 0.001) distribution is applied for noninformative prior distribution, given by

$$\begin{aligned} r &\sim \text{Gamma}(0.001, 0.001),\\ p &\sim \text{Gamma}(0.001, 0.001).\end{aligned}$$

Due to the complexity of the joint posterior distribution. The Markov Chain Monte Carlo techniques can be used to calculated the posterior distribution. Here, the Metropolis-Hastings-within Gibbs with 10,000 iterations in the LaplaceDemon function of the LaplaceDemon package (Statisticat, LLC., 2021) of the R programming language (R Core Team, 2023) is employed.

4. Simulation Study

In this section, the performance of the ML estimators and Bayesian estimators are present. The following steps can be used to generate NBB(r,θ) random variates.

1. Generate λ_i from the Bilal distribution by the inversion method.
2. Generate y_i from $\text{NB}(r, e^{-\lambda_i})$.

The simulation study is carried out 1,000 times for sample sizes $n = 50, 100, 150, 200$ with a specified parameters vector $r = 0.5$ and $\theta = 0.5$. The root mean square error (RMSE) and bias are calculated by

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{1,000} (\hat{\Theta}_i - \Theta)^2}{1,000}}, \qquad \text{and} \qquad \text{Bias} = \frac{\sum_{i=1}^{1,000} (\hat{\Theta}_i - \Theta)}{1,000}.$$

Table 1 displays the ML and Bayesian estimates for $r = 0.5$ and $\theta = 0.5$, and the RMSEs and biases of the ML estimates and posterior means for the NBB distribution are shown in Table 2. The RMSEs of r and θ for ML estimators and Bayesian estimators decrease with increasing sample size. While the biases of r and θ for ML estimators and Bayesian estimators approach zero with increasing sample size. Moreover, the MSEs and biases of the ML estimators are slightly lower than the Bayesian estimators because the ML estimators can be biased for small sample sizes, which small sample sizes are not used in this paper.

Table 1: ML estimates and posterior means for $r = 0.5$ and $\theta = 0.5$.

n	ML		Bayesian	
	r	θ	r	θ
50	3.022	0.442	3.627	0.464
100	1.356	0.473	2.195	0.469
150	0.738	0.487	1.240	0.478
200	0.652	0.485	0.992	0.476

Table 2: Average RMSEs (average biases) of the simulated estimates.

n	ML		Bayesian	
	r	θ	r	θ
50	6.467 (**2.522**)	**0.346** (−0.058)	**5.557** (3.127)	0.366 (−**0.036**)
100	**3.336** (**0.856**)	**0.265** (−**0.027**)	4.191 (1.695)	0.292 (−0.031)
150	**1.250** (**0.238**)	**0.200** (−**0.013**)	1.987 (0.740)	0.228 (−0.022)
200	**0.775** (**0.152**)	**0.176** (−**0.015**)	1.433 (0.492)	0.202 (−0.024)

5. The Negative-Binomial Bilal Regression Model

A new count regression model based on the NBB response variable is presented in this section. Let Λ follows the Bilal distribution, the pdf and the mean of the Bilal distribution are, respectively

$$\begin{aligned} g(\lambda) &= \frac{6}{\theta} e^{-\frac{2\lambda}{\theta}} \left(1 - e^{-\frac{\lambda}{\theta}}\right), \\ \mathbb{E}(\Lambda) &= \frac{5\theta}{6}. \end{aligned}$$

We parameterize $\theta = \frac{6\mathbb{E}(\Lambda)}{5}$ to the pdf of the Bilal distribution. Thus

$$g(\lambda) = \frac{5}{\mathbb{E}(\Lambda)} e^{-\frac{5\lambda}{3\mathbb{E}(\Lambda)}} \left(1 - e^{-\frac{5\lambda}{6\mathbb{E}(\Lambda)}}\right).$$

The Bayesian approach is employed to develop the NBB regression model. Let $Y_1, \ldots, Y_n$ be independent random variables follows the NBB(r, μ_i) distribution. The NBB regression model can be constructed by

$$\begin{aligned} Y_i|\Lambda_i &\sim \text{NB}(r, p = e^{-\Lambda_i}) \\ \Lambda_i &\sim \text{Bilal}(\mu_i), \\ \log(\mu_i) &= \boldsymbol{x_i}^T \boldsymbol{\beta}, \end{aligned}$$

where $\boldsymbol{x_i} = (1, x_{i1}, x_{i2}, \ldots, x_{ip})^T$ is the vector of the explanatory variables, and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^T$ is the parameter vector of each covariate.

In this paper, the prior distributions for all unknown parameters considered are

$$\begin{aligned} r &\sim \text{Gamma}(0.001, 0.001), \\ \boldsymbol{\beta} &\sim \text{N}(0, 10000). \end{aligned}$$

The Metropolis-Hastings-within Gibbs with 10,000 iterations in the LaplaceDemon function of the LaplaceDemon package (Statisticat, LLC., 2021) in the R programming language (R Core Team, 2023) is applied for the posterior densities of parameters.

6. Applications

In this section, two health care data sets are considered to show the performance of the NBB model. The proposed model is compared with two traditional count models including the Poisson, and NB models. The criteria of classical parameter estimation are the Akaike information criterion (AIC) (Akaike, 1974) the Bayesian information criterion (BIC) (Schwarz, 1978), and the discrete Anderson-Darling (AD) goodness of fit test for discrete distributions (Choulakian et al., 1994). The model with the lowest the AIC, the lowest BIC, and the highest of p-value based on the discrete AD test is recommended that better models. Based on the Bayesian approach, we apply the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) . The lowest of DIC indicates better models.

Application 1

The first data set concerns the number of doctor visits in the past two weeks from the Australian Health Survey between 1977 and 1978 (Cameron and Trivedi, 2013). The data set is overdispersed because the sample variance, $s^2 = 0.637$, is larger than the sample mean, $\bar{y} = 0.302$.

The results based on the method ML and Bayesian approach are shown in Table 3 and Table 4, respectively. They are evident that the NBB distribution provides a better fit than the competing distributions, since it provides the smallest AIC, smallest BIC, smallest DIC and highest p-values based on the discrete AD test. In addition, the p-values based on the discrete AD test of the Poisson distribution for the data set is less than the 5% significance level because the data set is overdispersed. Thus, the Poisson distribution cannot be applied for it.

Application 2

The second health care data set with covariates is considered to show the performance of the NBB regression model. The data set appears in Hilbe and Greene (2007), and Hilbe (2011). It contains 1,127 observations from German health survey data for 1998. Table 5 displays variables and their summary statistics. We consider the number of visits to a doctor during 1998 as a response variable. The response variable is overdispersed because the variance is larger than the mean.

Table 3: ML estimates, expected frequencies, AIC values, BIC values, and p-values based on the discrete AD test of fitted distributions for the doctor visit data set.

Number of doctor visits	Observed frequencies	Expected frequencies		
		Poisson	NB	NBB
0	4141	3838.185	4158.052	4155.221
1	782	1158.111	697.124	728.218
2	174	174.721	213.274	193.287
3	30	17.573	75.078	64.704
4	24	1.326	28.160	25.351
5	9	0.080	10.951	11.153
6	12	0.004	4.360	5.366
7	12	0	1.765	2.773
8	5	0	0.723	1.520
9	1	0	0.299	0.874
Estimated parameters		$\hat{\lambda} = 0.302$	$\hat{r} = 0.377$	$\hat{r} = 0.817$
			$\hat{p} = 0.556$	$\hat{\theta} = 0.346$
log-likelihood		-3983.194	-3585.992	-3565.924
AIC		7968.389	7175.983	7135.847
BIC		7974.943	7189.092	7148.956
AD		63.106	1.942	0.822
p-value		< 0.001	0.055	0.233

Table 6 shows the posterior means of parameters, the lower bound (LB) and the upper bound (UB) of the 95% credible intervals for parameters, and the DIC values. The table indicates that the NBB regression model is the best model among the compared regression models because it provides the smallest DIC values. Furthermore, all parameters are significant at the 0.05 when the 95% credible intervals for parameters are considered.

7. Conclusions

This paper proposes a new mixed NB distribution. The distribution is the mixture of the NB and Bilal distributions, called the NBB distribution. Several statistical properties were presented. Parameter estimators of the proposed distribution were derived by the ML method and Bayesian approach. Moreover, the simulation study was conduced to assess the performance of the ML estimators and the Bayesian estimators. Both ML and Bayesian approaches provided good performance with respect to the RMSE and bias of the estimators. In addition, we developed a count regression model based on the NBB distribution. The usefulness of NBB distribution and its associated regression model were shown through applications to health care data. The results indicate that the proposed model provides better fits than compared models. Thus, the proposed model can be considered as an alternative to any of count models.

Acknowledgements

The authors are grateful to the editor and anonymous reviewers for their valuable comments and suggestion. The first author would like to thank Rajamangala University of Technology Phra Nakhon.

Table 4: Posterior means, expected frequencies, DIC values, and *p*-values based on the discrete AD test of fitted distributions for the doctor visits data set.

Number of doctor visits	Observed frequencies	Expected frequencies		
		Poisson	NB	NBB
0	4141	3844.745	4157.819	4155.572
1	782	1153.525	694.808	728.275
2	174	173.044	213.878	193.151
3	30	17.306	75.868	64.599
4	24	1.298	28.692	25.287
5	9	0.078	11.254	11.115
6	12	0.004	4.520	5.343
7	12	0	1.846	2.759
8	5	0	0.763	1.511
9	1	0	0.318	0.869
Posterior means		$\hat{\lambda} = 0.306$	$\hat{r} = 0.373$	$\hat{r} = 0.818$
			$\hat{p} = 0.551$	$\hat{\theta} = 0.345$
DIC		7967.228	7173.907	7134.073
AD		61.72	2.028	0.818
p-value		< 0.001	0.050	0.235

Table 5: Variables and summary statistics.

Variable	Measurement	Mean	Variance
numvisit	The number of visits to a doctor	2.353	11.982
badh	Patients claims		
	1=bad health		
	0=not bad health		
age	Age of patient	37.229	117.265

References

1. Abd-Elrahman, A. M. (2013). Utilizing Ordered Statistics in Lifetime Distributions Production: A New Lifetime Distribution and Applications. *Journal of Probability and Statistical Science*, 11:153–164.
2. Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.
3. Aryuyuen, S. and Bodhisuwan, W. (2013). The Negative Binomial-Generalized Exponential (NB-GE) Distribution. *Applied Mathematical Sciences*, 7(22):1093–1105.
4. Atikankul, Y., Wattanavisut, A., and Liu, S. (2022). The Negative Binomial-Generalized Lindley Distribution for Overdispersed Data. *Lobachevskii Journal of Mathematics*, 43(9):2378–2386.
5. Bodhisuwan, W. and Saengthong, P. (2020). The Negative Binomial-Weighted Garima Distribution: Model, Properties and Applications. *Pakistan Journal of Statistics and Operation Research*, 16(1):1–10.
6. Cameron, A. C. and Trivedi, P. K. (2013). *Regression Analysis of Count Data*, volume 53. Cambridge Univer-

Table 6: Posterior means, credible intervals, and DIC values of fitted distributions for the doctor visits data set.

Parameter	Posterior mean (LB, UB)		
	Poisson	NB	NBB
β_0	0.157	0.185	0.362
	(0.118, 0.201)	(0.118, 0.268)	(0.311, 0.442)
β_1	1.049	1.068	1.105
	(0.953, 1.148)	(0.889, 1.297)	(0.889, 1.328)
β_2	0.013	0.013	0.008
	(0.012,0.013)	(0.012, 0.013)	(0.007, 0.009)
r	-	0.997	1.000
	-	(0.883, 1.133)	(0.879, 1.132)
DIC	5655.007	4476.518	4472.723

sity Press.

7. Choulakian, V., Lockhart, R. A., and Stephens, M. A. (1994). Cramér-von Mises Statistics for Discrete Distributions. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 22:125–137.
8. Gómez-Déniz, E., Sarabia, J. M., and Calderín-Ojeda, E. (2008). Univariate and Multivariate Versions of the Negative Binomial-Inverse Gaussian Distributions with Applications. *Insurance Mathematics and Economics*, 42(1):39–49.
9. Hassan, A., Shah, I., and Bilal, P. (2021). Negative Binomial-Reciprocal Inverse Gaussian Distribution: Statistical Properties with Applications. *Thailand Statistician*, 19(3):437–449.
10. Hilbe, J. M. (2011). *Negative Binomial Regression*. Cambridge University Press.
11. Hilbe, J. M. and Greene, W. H. (2007). 7 Count Response Regression Models. *Handbook of Statistics*, 27:210–252.
12. Meng, S., Yuan, W., and Whitmore, G. (1999). Accounting for Individual Over-Dispersion in a Bonus-Malus Automobile Insurance System. *ASTIN Bulletin*, 29:327–338.
13. Nash, J. C. and Varadhan, R. (2011). Unifying Optimization Algorithms to Aid Software System Users: Optimx for R. *Journal of Statistical Software*, 43(9):1–14.
14. Pudprommarat, C., Bodhisuwan, W., and Zeephongsekul, P. (2012). A New Mixed Negative Binomial Distribution. *Journal of Applied Sciences*, 12(17):1853–1858.
15. R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
16. Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464.
17. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.
18. Statisticat, LLC. (2021). *LaplacesDemon: Complete Environment for Bayesian Inference*. Package Version 16.1.6.
19. Tajuddin, R. R. M., Ismail, N., Ibrahim, K., and Bakar, S. A. A. (2022). A Four-Parameter Negative Binomial-Lindley Distribution for Modeling Over and Underdispersed Count Data with Excess Zeros. *Communications in Statistics-Theory and Methods*, 51(2):414–426.
20. Yamrubboon, D., Bodhisuwan, W., Pudprommarat, C., and Saothayanun, L. (2017). The Negative Binomial-Sushila Distribution with Application in Count Data Analysis. *Thailand Statistician*, 15(1):69–77.
21. Zamani, H. and Ismail, N. (2010). Negative Binomial-Lindley Distribution and Its Application. *Journal of Mathematics and Statistics*, 6(1):4–9.