

NEGATIVE BINOMIAL-CRACK (NB-CR) DISTRIBUTION

Pornpop Saengthong¹, Winai Bodhisuwan² §

^{1,2}Department of Statistics

Kasetsart University

Chatuchak, Bangkok, 10900, THAILAND

Abstract: The objective of this paper is to provide an alternative distribution for modeling overdispersed count data. We propose a negative binomial-Crack (NB-CR) distribution which is obtained by mixing the NB distribution with a CR distribution. This new formulation distribution contains as special cases three parameter distribution, namely, negative binomial-inverse Gaussian (NB-IG), negative binomial-Birnbaum-Saunders (NB-BS) and negative binomial-length biased inverse Gaussian (NB-LBIG). In addition, we present some properties of the new distribution such as the factorial moments, the first four moments, variance, skewness and kurtosis. Parameters estimation are also implemented using maximum likelihood method and the application of NB-CR distribution is carried out on a sample of count data. The results show that the NB-CR provides a better fit compared to the Poisson and the NB distribution.

AMS Subject Classification: 60E05

Key Words: mixed negative binomial distribution, negative binomial-crack distribution, count data, overdispersion

1. Introduction

Count data arise in many contexts and can be modeled by a Poisson distribu-

Received: December 23, 2012

© 2013 Academic Publications, Ltd.
url: www.acadpubl.eu

§Correspondence author

tion, which is a proper model for counting the number of occurrences over a time interval at random when not many occurrences are observed within a short period of time, they occur at a constant rate through time, and one occurrence of the phenomenon does not alter the probability of any future occurrence. Equality of mean and variance, called equal dispersion, is a quintessential characteristic of the Poisson distribution [12]. However, many count data often exhibit overdispersion, with a variance larger than the mean and an extension to the Poisson model is more appropriate. The negative binomial (NB) distribution is a popular alternative distribution for modeling overdispersed count data because it is more flexible in accommodating over-dispersion in comparison with the Poisson model.

The NB distribution is a mixture of Poisson distribution by mixing the Poisson and gamma distribution. Applications using the NB distribution can be found in many areas, for instance, economics [6], accident statistics [10], biostatistics [1], [7] and actuarial science [3], [16]. Although, the NB distribution allows for over-dispersion, it does not take care of excess zeros in the data. [9] studied on a tool for analyzing crash data characterized by a large amount of zeros. They pointed out that traditional statistical distributions or model, such as the Poisson and the NB distributions, cannot be used efficiently in models for count data with many zeros. The Poisson distribution tends to under-estimate the number of zeros given the mean of the data, while the NB may over-estimate zeros, but under-estimate observations with a count.

The problem of overdispersion and excess zeros is usually solved by introducing mixed Poisson or mixed NB distribution. In several studies, it is shown that mixed Poisson and mixed NB distribution provided better fit on count data compared to the Poisson and the NB distribution. These include the Poisson-inverse Gaussian [8], [13], [15], negative binomial-inverse Gaussian [5], negative binomial-Lindley [16], beta-negative binomial [14] and negative binomial-Beta Exponential [11]. Therefore, in order to provide another competitive alternative to the models above, a new mixed model is considered. We propose the negative binomial-Crack (NB-CR) distribution which is a new mixed NB distribution obtained by mixing the distribution of $NB(r,p)$ where, $p = \exp(-a)$ with distribution of $CR(\lambda, \theta, \gamma)$. This new distribution has a heavy tail, seems to be positively skewed and may be considered as a competitive alternative for modeling overdispersed count data.

The purpose of this paper is to investigate the properties of the NB-CR distribution and its application. In Section 2 we present the distributions. Section 3 is concerned with mixture of the NB distribution with the CR distribution. Properties of the NB-CR distribution including the factorial moments, the first

four moments, variance, skewness, kurtosis and parameter estimation are given in Section 4 and 5, respectively. Section 6 provides an application to illustrate our methodology by a real data set and some conclusions are presented in the last section.

2. Preliminaries

As mentioned earlier, the NB-CR distribution is a mixture of the NB and Crack distributions. First we present the NB distribution and some of its properties. The probability mass function (pmf) of NB distribution can be given as

$$f(x) = \binom{r+x-1}{x} p^r (1-p)^x, \quad x = 0, 1, 2, \dots \tag{1}$$

where $r > 0$ and $0 < p < 1$.

The first two moments about zero and the factorial moment of a NB distribution (see [2]) are respectively given by

$$\begin{aligned} E(X) &= \frac{r(1-p)}{p}, \\ E(X^2) &= \frac{r(1-p)[1+r(1-p)]}{p^2}, \\ \mu_{[k]}(X) &= E[X(X-1)\cdots(X-k+1)] \\ &= \frac{\Gamma(r+k)}{\Gamma(r)} \frac{(1-p)^k}{p^k}, \quad k = 1, 2, \dots \end{aligned} \tag{2}$$

where $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(t) = \int_0^\infty x^{t-1} \exp(-x) dx, \quad t > 0.$$

In order to maintain a self-contained analysis, we present the CR distribution with parameter λ, θ, γ , with its probability density function (pdf)

$$\begin{aligned} g(x) &= \frac{1}{\theta\sqrt{2\pi}} \left[\gamma\lambda\left(\frac{\theta}{x}\right)^{3/2} + (1-\gamma)\left(\frac{\theta}{x}\right)^{1/2} \right] \\ &\times \exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{\theta}} - \lambda\sqrt{\frac{\theta}{x}}\right)^2 \right], \quad x > 0 \end{aligned} \tag{3}$$

where $\lambda > 0$, $\theta > 0$ and $0 \leq \gamma \leq 1$.

The CR distribution was studied by [4] and it was shown there that its moment generating function is given by

$$M_X(t) = \frac{\exp[\lambda(1 - \sqrt{1 - 2\theta t})]}{\sqrt{1 - 2\theta t}} \left[1 - \gamma(1 - \sqrt{1 - 2\theta t})\right], \quad (4)$$

and exists if $t < \frac{1}{2\theta}$.

Definition 1. Let X be a random variable which follows the negative binomial-Crack distribution with parameter r , λ , θ and γ , $X \sim \text{NB-CR}(r, \lambda, \theta, \gamma)$, when X has a NB distribution with parameter $r > 0$ and $p = \exp(-a)$ where a is distributed as CR with positive parameters λ , θ and γ , i.e., $X|a \sim \text{NB}(r, p = \exp(-a))$ and $a \sim \text{CR}(\lambda, \theta, \gamma)$.

3. The Negative Binomial-Crack (NB-CR) Distribution

Now, we prove that the distribution of NB-CR is distributed according to Eq. 1 and that the unknown parameter, a , follows a prior distribution as defined in Eq. 3.

Theorem 2. Let $X \sim \text{NB-CR}(r, \lambda, \theta, \gamma)$. The probability mass function of X is given by

$$h(x) = \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \frac{\exp\left[\lambda\left(1 - \sqrt{1 + 2\theta(r+j)}\right)\right]}{\sqrt{1 + 2\theta(r+j)}} \times \left[1 - \gamma\left(1 - \sqrt{1 + 2\theta(r+j)}\right)\right] \quad x = 0, 1, 2, \dots \quad (5)$$

where $r, \lambda, \theta > 0$ and $0 \leq \gamma \leq 1$.

Proof. If $X|a \sim \text{NB}(r, p = \exp(-a))$ in Eq. 1 and $a \sim \text{CR}(\lambda, \theta, \gamma)$ in Eq. 3, then the pmf of X can be obtained by

$$h(x) = \int_0^\infty f(x|a)g(a; \lambda, \theta, \gamma) da, \quad (6)$$

where

$$f(x|a) = \binom{r+x-1}{x} \exp(-ar)(1 - \exp(-a))^x$$

$$= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \exp(-a(r+j)). \tag{7}$$

By substituting Eq. 7 into Eq. 6. we obtain

$$\begin{aligned} h(x) &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \int_0^{\infty} \exp(-a(r+j)) g(a; \lambda, \theta, \gamma) da \\ &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j M_a(-(r+j)). \end{aligned} \tag{8}$$

Substituting the moment generating function of CR distribution in Eq. 4 into Eq. 8, the pmf of NB-CR($r, \lambda, \theta, \gamma$) is finally given as

$$\begin{aligned} h(x) &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta(r+j)} \right) \right]}{\sqrt{1 + 2\theta(r+j)}} \\ &\quad \times \left[1 - \gamma \left(1 - \sqrt{1 + 2\theta(r+j)} \right) \right]. \end{aligned}$$

□

In order to study the behavior of the distribution for different value of $r, \lambda, \theta, \gamma$, the pmf is calculated. We display some of these in the following corollaries and their graphs are shown in Fig. 1. From the graphs it can be shown that all of them seem unimodal, positively skewed.

Corollary 3. *If $\gamma = 1, \lambda = \frac{\psi}{\mu}$ and $\theta = \frac{\mu^2}{\psi}$ then the NB-CR distribution reduces to the negative binomial-inverse Gaussian (NB-IG) distribution [5] with pmf given by*

$$\begin{aligned} h(x) &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \\ &\quad \times \exp \left[\frac{\psi}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2(r+j)}{\psi}} \right) \right], \quad x = 0, 1, 2, \dots \end{aligned} \tag{9}$$

where r, μ and $\psi > 0$.

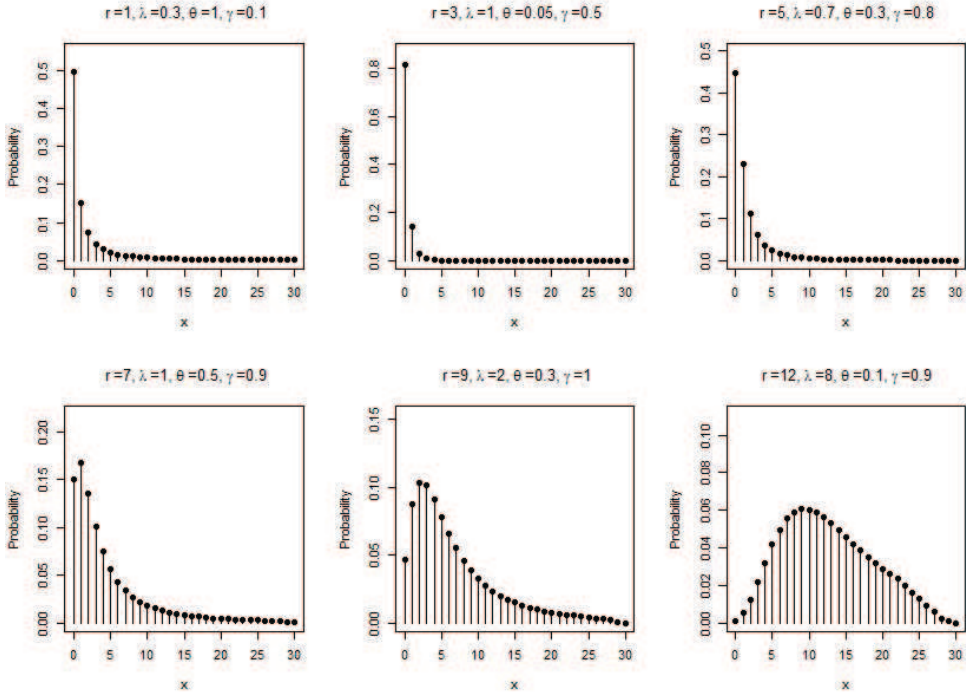


Figure 1: Some examples of probability mass functions of the NB-CR random variable with some specified values of r , λ , θ and γ .

Proof. Substituting $\gamma = 1$, $\lambda = \frac{\psi}{\mu}$ and $\theta = \frac{\mu^2}{\psi}$ into Eq. 5, then the pmf of X becomes

$$\begin{aligned}
 h(x) &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \frac{\exp \left[\frac{\psi}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2(r+j)}{\psi}} \right) \right]}{\sqrt{1 + \frac{2\mu^2(r+j)}{\psi}}} \\
 &\quad \times \left[1 - 1 \left(1 - \sqrt{1 + \frac{2\mu^2(r+j)}{\psi}} \right) \right] \\
 &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \exp \left[\frac{\psi}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2(r+j)}{\psi}} \right) \right].
 \end{aligned}$$

□

Corollary 4. *If $\gamma = \frac{1}{2}$ then the NB-CR distribution reduces to the negative binomial-Birnbaum-Saunders (NB-BS) distribution with pmf given by*

$$\begin{aligned}
 h(x) &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta(r+j)} \right) \right]}{\sqrt{1 + 2\theta(r+j)}} \\
 &\quad \times \frac{1}{2} \left[1 + \sqrt{1 + 2\theta(r+j)} \right], \quad x = 0, 1, 2, \dots
 \end{aligned}
 \tag{10}$$

where r, λ and $\theta > 0$.

Proof. Substituting $\gamma = \frac{1}{2}$ into Eq. 5, then we get the pmf of the NB-BS distribution of the random variable X as

$$\begin{aligned}
 h(x) &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta(r+j)} \right) \right]}{\sqrt{1 + 2\theta(r+j)}} \\
 &\quad \times \left[1 - \frac{1}{2} \left(1 - \sqrt{1 + 2\theta(r+j)} \right) \right] \\
 &= \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta(r+j)} \right) \right]}{\sqrt{1 + 2\theta(r+j)}} \\
 &\quad \times \frac{1}{2} \left[1 + \sqrt{1 + 2\theta(r+j)} \right].
 \end{aligned}$$

We can see that the NB-BS distribution is a new mixed NB distribution by mixing the distributions of NB and BS. The pmf of NB-BS(r, λ, θ) can be derived by substituting the moment generating function of BS distribution in Eq. 11 into Eq. 8.

$$M_X(t) = \exp \left[\lambda \left(1 - \sqrt{1 - 2\theta t} \right) \right] \frac{1}{2} \left[1 + \frac{1}{\sqrt{1 - 2\theta t}} \right].
 \tag{11}$$

□

Corollary 5. *If $\gamma = 0$ then the NB-CR distribution reduces to the negative binomial-length biased inverse Gaussian (NB-LBIG) distribution with pmf given by*

$$h(x) = \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j$$

$$\times \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta (r + j)} \right) \right]}{\sqrt{1 + 2\theta (r + j)}}, \quad x = 0, 1, 2, \dots \tag{12}$$

where r, λ and $\theta > 0$.

Proof. Substituting $\gamma = 0$ into Eq. 5, then we get the pmf of the NB-LBIG distribution of the random variable X as

$$h(x) = \binom{r+x-1}{x} \sum_{j=0}^x \binom{x}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta (r + j)} \right) \right]}{\sqrt{1 + 2\theta (r + j)}}.$$

The NB-LBIG is a new mixed NB distribution which can be obtained by mixing the NB distribution with a LBIG distribution. We can show the pmf of NB-LBIG(r, λ, θ) by substituting the moment generating function of LBIG distribution in Eq. 13 into Eq. 8.

$$M_X(t) = \frac{\exp \left[\lambda \left(1 - \sqrt{1 - 2\theta t} \right) \right]}{\sqrt{1 - 2\theta t}}. \tag{13}$$

□

From Corollary 3-5, we find that the NB-IG distribution and the new mixed NB; NB-BS distribution and NB-LBIG distribution are all special cases of the NB-CR distribution.

4. The Properties of the NB-CR Distribution

We now give an useful expansions for the pmf in Eq. 5. With the expansion, we can obtain mathematical properties of the NB-CR distribution and its special cases. The main result of this section is theorem and the subsequent corollaries which are concerned with the factorial moments. We hardly need to emphasize the necessity and importance of factorial moment in any statistical analysis especially in applied work. Some of the most important features and characteristics of a distribution can be studied through factorial moments (e.g., mean, variance, skewness and kurtosis).

Theorem 6. *If $X \sim \text{NB-CR}(r, \lambda, \theta, \gamma)$. Then the factorial moment of order k of X is given by*

$$\mu_{[k]}(X) = \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 - 2\theta (k - j)} \right) \right]}{\sqrt{1 - 2\theta (k - j)}}$$

$$\times \left[1 - \gamma \left(1 - \sqrt{1 - 2\theta(k - j)} \right) \right], \quad k = 1, 2, \dots \tag{14}$$

where $r, \lambda, \theta > 0$ and $0 \leq \gamma \leq 1$.

Proof. From [5], the factorial moment of order k of mixed NB distribution where $p = \exp(-a)$ can be expressed in the terms of elementary function by

$$\begin{aligned} \mu_{[k]}(X) &= E_a \left(\frac{\Gamma(r + k)}{\Gamma(r)} \frac{(1 - \exp(-a))^k}{\exp(-ak)} \right) \\ &= \frac{\Gamma(r + k)}{\Gamma(r)} E_a(\exp(a) - 1)^k. \end{aligned} \tag{15}$$

Using of a binomial expansion for the term $(\exp(a) - 1)^k$, we can write $\mu_{[k]}(x)$ as

$$\begin{aligned} \mu_{[k]}(X) &= \frac{\Gamma(r + k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j E(\exp(a(k - j))) \\ &= \frac{\Gamma(r + k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j M_a(k - j). \end{aligned} \tag{16}$$

From the moment generating function of CR distribution in Eq. 4 with $t = k - j$, we have finally that $\mu_{[k]}(X)$ can be written as

$$\begin{aligned} \mu_{[k]}(X) &= \frac{\Gamma(r + k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 - 2\theta(k - j)} \right) \right]}{\sqrt{1 - 2\theta(k - j)}} \\ &\quad \times \left[1 - \gamma \left(1 - \sqrt{1 - 2\theta(k - j)} \right) \right]. \end{aligned}$$

□

As a consequence of this theorem, we have the following results.

Corollary 7. *If $\gamma = 1$, $\lambda = \frac{\psi}{\mu}$ and $\theta = \frac{\mu^2}{\psi}$ then the factorial moments of negative binomial-crack reduces to*

$$\mu_{[k]}(X) = \frac{\Gamma(r + k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j$$

$$\times \exp \left[\frac{\psi}{\mu} \left(1 - \sqrt{1 - \frac{2(k-j)\mu^2}{\psi}} \right) \right], \quad k = 1, 2, \dots \quad (17)$$

where r, λ and $\theta > 0$. which is the same as the factorial moment of order k of NB-IG distribution (see [5]).

Proof. Substituting $\gamma = 1, \lambda = \frac{\psi}{\mu}$ and $\theta = \frac{\mu^2}{\psi}$ into Eq.14, we get

$$\begin{aligned} \mu_{[k]}(X) &= \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \frac{\exp \left[\frac{\psi}{\mu} \left(1 - \sqrt{1 - \frac{2(k-j)\mu^2}{\psi}} \right) \right]}{\sqrt{1 - \frac{2(k-j)\mu^2}{\psi}}} \\ &\quad \times \left[1 - \left(1 - \sqrt{1 - \frac{2(k-j)\mu^2}{\psi}} \right) \right] \\ &= \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \\ &\quad \times \exp \left[\frac{\psi}{\mu} \left(1 - \sqrt{1 - \frac{2(k-j)\mu^2}{\psi}} \right) \right]. \end{aligned} \quad (18)$$

□

Corollary 8. If $\gamma = \frac{1}{2}$ then the factorial moments of negative binomial-crack reduces to the factorial moment of order k of NB-BS distribution which is then given as

$$\begin{aligned} \mu_{[k]}(X) &= \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 - 2\theta(k-j)} \right) \right]}{\sqrt{1 - 2\theta(k-j)}} \\ &\quad \times \frac{1}{2} \left[1 + \sqrt{1 - 2\theta(k-j)} \right], \quad k = 1, 2, \dots \end{aligned} \quad (19)$$

where r, λ and $\theta > 0$.

Proof. Substituting $\gamma = \frac{1}{2}$ into Eq.14, we get

$$\mu_{[k]}(X) = \frac{\Gamma(r+k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 - 2\theta(k-j)} \right) \right]}{\sqrt{1 - 2\theta(k-j)}}$$

$$\times \frac{1}{2} \left[1 + \sqrt{1 - 2\theta(k - j)} \right].$$

which is the same as the factorial moment of order k of NB-BS distribution, can be derived by substituting the moment generating function of BS distribution in Eq. 11 into Eq. 16. □

Corollary 9. *If $\gamma = 0$ then the factorial moments of negative binomial-crack reduces to the factorial moment of order k of NB-LBIG distribution which can be written as*

$$\begin{aligned} \mu_{[k]}(X) &= \frac{\Gamma(r + k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \\ &\times \frac{\exp \left[\lambda \left(1 - \sqrt{1 - 2\theta(k - j)} \right) \right]}{\sqrt{1 - 2\theta(k - j)}}, \quad k = 1, 2, \dots \end{aligned} \tag{20}$$

where r, λ and $\theta > 0$.

Proof. Substituting $\gamma = 0$ into Eq.14, we get

$$\mu_{[k]}(X) = \frac{\Gamma(r + k)}{\Gamma(r)} \sum_{j=0}^k \binom{k}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 - 2\theta(k - j)} \right) \right]}{\sqrt{1 - 2\theta(k - j)}}.$$

We can see that the above equation is the same as the factorial moment of order k of NB-LBIG distribution which is obtained by substituting the moment generating function of LBIG distribution in Eq. 13 into Eq. 16. □

From the factorial moments of NB-CR distribution, for convenience we let $\delta = \sqrt{1 - 2\theta}$, $\zeta = \sqrt{1 - 4\theta}$, $\vartheta = \sqrt{1 - 6\theta}$ and $\kappa = \sqrt{1 - 8\theta}$, the first four order moments about zero, variance, skewness and kurtosis are respectively given by

$$\begin{aligned} E(X) &= r \left(\frac{(1 - \gamma(1 - \delta)) \exp(\lambda(1 - \delta))}{\delta} \right) - r, \\ E(X^2) &= \exp(\lambda) \left(\frac{(r^2 + r)(1 - \gamma(1 - \zeta)) \exp(-\lambda\zeta)}{\zeta} \right. \\ &\quad \left. - \frac{(2r^2 + r)(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right) + r^2, \\ E(X^3) &= \exp(\lambda) \left(\frac{(r^3 + 3r^2 + 2r)(1 - \gamma(1 - \vartheta)) \exp(-\lambda\vartheta)}{\vartheta} \right) \end{aligned}$$

$$- \frac{(3r^3 + 6r^2 + 3r)(1 - \gamma(1 - \zeta)) \exp(-\lambda\zeta)}{\zeta} + \frac{(3r^3 + 3r^2 + r)(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \Big) - r^3,$$

$$\begin{aligned} E(X^4) &= \exp(\lambda) \left(\frac{(r^4 + 6r^3 + 11r^2 + 6r)(1 - \gamma(1 - \kappa)) \exp(-\lambda\kappa)}{\kappa} \right. \\ &\quad - \frac{(4r^4 + 18r^3 + 26r^2 + 12r)(1 - \gamma(1 - \vartheta)) \exp(-\lambda\vartheta)}{\vartheta} \\ &\quad + \frac{(6r^4 + 18r^3 + 19r^2 + 7r)(1 - \gamma(1 - \zeta)) \exp(-\lambda\zeta)}{\zeta} \\ &\quad \left. - \frac{(4r^4 + 6r^3 + 4r^2 + r)(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right) + r^4, \end{aligned}$$

$$\begin{aligned} Var(X) &= \exp(\lambda) \left(\frac{(r^2 + r)(1 - \gamma(1 - \zeta)) \exp(-\lambda\zeta)}{\zeta} \right. \\ &\quad - \frac{r(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \\ &\quad \left. - r^2 \left(\frac{(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right)^2 \exp(\lambda) \right), \end{aligned}$$

$$\begin{aligned} \text{Skewness}(X) &= \exp(\lambda) \left(\frac{(r^3 + 3r^2 + 2r)(1 - \gamma(1 - \vartheta)) \exp(-\lambda\vartheta)}{\vartheta} \right. \\ &\quad - \frac{(3r^2 + 3r)(1 - \gamma(1 - \zeta)) \exp(-\lambda\zeta)}{\zeta} \\ &\quad + \frac{r(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \\ &\quad \left. + 3r^2 \left(\frac{(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right)^2 \exp(\lambda) \right), \end{aligned}$$

$$\frac{(3r^3 + 3r^2)(1 - \gamma(1 - \delta))(1 - \gamma(1 - \zeta)) \exp(\lambda(1 - \delta - \zeta))}{\delta\zeta} + 2r^3 \left(\frac{(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right)^3 \exp(2\lambda) \Big) / (\text{var}(X))^{3/2},$$

Kurtosis (X)

$$\begin{aligned} = & \exp(\lambda) \left(\frac{(r^4 + 6r^3 + 11r^2 + 6r)(1 - \gamma(1 - \kappa)) \exp(-\lambda\kappa)}{\kappa} \right. \\ & - \frac{(6r^3 + 18r^2 + 12r)(1 - \gamma(1 - \vartheta)) \exp(-\lambda\vartheta)}{\vartheta} \\ & + \frac{(7r^2 + 7r)(1 - \gamma(1 - \zeta)) \exp(-\lambda\zeta)}{\zeta} \\ & - \frac{r(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \\ & - 4r^2 \left(\frac{(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right)^2 \exp(\lambda) \\ & + \frac{(12r^3 + 12r^2)(1 - \gamma(1 - \delta))(1 - \gamma(1 - \zeta)) \exp(\lambda(1 - \delta - \zeta))}{\delta\zeta} \\ & - \frac{(4r^4 + 12r^3 + 8r^2)(1 - \gamma(1 - \delta))(1 - \gamma(1 - \vartheta)) \exp(\lambda(1 - \delta - \vartheta))}{\delta\vartheta} \\ & + (6r^4 + 6r^3) \left(\frac{(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right)^2 \frac{(1 - \gamma(1 - \zeta)) \exp(\lambda(2 - \zeta))}{\zeta} \\ & - 6r^3 \left(\frac{(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right)^3 \exp(2\lambda) \\ & \left. - 3r^4 \left(\frac{(1 - \gamma(1 - \delta)) \exp(-\lambda\delta)}{\delta} \right)^4 \exp(3\lambda) \right) / (\text{var}(X))^2. \end{aligned}$$

5. Parameters Estimation

The parameters estimation of NB-CR distribution will be done via the Maximum Likelihood Estimation (MLE) procedure. The likelihood function of the NB-CR($r, \lambda, \theta, \gamma$) is given by

$$\mathcal{L}(r, \lambda, \theta, \gamma) = \prod_{i=1}^n \binom{r + x_i - 1}{x_i} \sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j$$

$$\begin{aligned} & \times \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta (r + j)} \right) \right]}{\sqrt{1 + 2\theta (r + j)}} \\ & \times \left[1 - \gamma \left(1 - \sqrt{1 + 2\theta (r + j)} \right) \right], \end{aligned}$$

from which we calculate the log-likelihood function

$$\begin{aligned} \ell(r, \lambda, \theta, \gamma) &= \log \mathcal{L}(r, \lambda, \theta, \gamma) \\ &= \sum_{i=1}^n (\log (r + x_i - 1)! - \log x_i! - \log (r - 1)!) \\ & \quad + \sum_{i=1}^n \left(\log \sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp \left[\lambda \left(1 - \sqrt{1 + 2\theta (r + j)} \right) \right]}{\sqrt{1 + 2\theta (r + j)}} \right. \\ & \quad \left. \times \left[1 - \gamma \left(1 - \sqrt{1 + 2\theta (r + j)} \right) \right] \right). \end{aligned}$$

It can be verified that the first partial derivatives $\ell(r, \lambda, \theta, \gamma)$ with respect to r, λ, θ and γ , we then obtain the following differential equations;

$$\begin{aligned} & \frac{\partial}{\partial r} \ell(r, \lambda, \theta, \gamma) \\ &= \sum_{i=1}^n \left(\frac{\frac{\partial}{\partial r} \Gamma(r + x_i)}{\Gamma(r + x_i)} - \frac{\frac{\partial}{\partial r} \Gamma(r)}{\Gamma(r)} \right) \\ & \quad + \sum_{i=1}^n \frac{\partial}{\partial r} \log \sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1 - Z_j)]}{Z_j} [1 - \gamma(1 - Z_j)] \\ &= \sum_{i=1}^n \psi(r + x_i) - n\psi(r) \\ & \quad + \sum_{i=1}^n \left(\frac{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\partial}{\partial r} \left(\frac{\exp[\lambda(1 - Z_j)]}{Z_j} [1 - \gamma(1 - Z_j)] \right)}{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1 - Z_j)]}{Z_j} [1 - \gamma(1 - Z_j)]} \right), \end{aligned}$$

where $\psi(k) = \frac{\Gamma'(k)}{\Gamma(k)}$ is the digamma function.

$$\frac{\partial}{\partial \lambda} \ell(r, \lambda, \theta, \gamma)$$

$$\begin{aligned}
 &= \sum_{i=1}^n \frac{\partial}{\partial \lambda} \log \sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1-Z_j)]}{Z_j} [1 - \gamma(1-Z_j)] \\
 &= \sum_{i=1}^n \left(\frac{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{[1-\gamma(1-Z_j)]}{Z_j} \exp[\lambda(1-Z_j)] (1-Z_j)}{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1-Z_j)]}{Z_j} [1 - \gamma(1-Z_j)]} \right),
 \end{aligned}$$

$$\begin{aligned}
 &\frac{\partial}{\partial \theta} \ell(r, \lambda, \theta, \gamma) \\
 &= \sum_{i=1}^n \frac{\partial}{\partial \theta} \log \sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1-Z_j)]}{Z_j} [1 - \gamma(1-Z_j)] \\
 &= \sum_{i=1}^n \left(\frac{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\partial}{\partial \theta} \left(\frac{\exp[\lambda(1-Z_j)]}{Z_j} [1 - \gamma(1-Z_j)] \right)}{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1-Z_j)]}{Z_j} [1 - \gamma(1-Z_j)]} \right),
 \end{aligned}$$

$$\begin{aligned}
 &\frac{\partial}{\partial \gamma} \ell(r, \lambda, \theta, \gamma) \\
 &= \sum_{i=1}^n \frac{\partial}{\partial \gamma} \log \sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1-Z_j)]}{Z_j} [1 - \gamma(1-Z_j)] \\
 &= \sum_{i=1}^n \left(\frac{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^{j+1} \frac{\exp[\lambda(1-Z_j)]}{Z_j} (1-Z_j)}{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j \frac{\exp[\lambda(1-Z_j)]}{Z_j} [1 - \gamma(1-Z_j)]} \right),
 \end{aligned}$$

where $Z_j = \sqrt{1 + 2\theta(r + j)}$.

These four derivative equations cannot be solved analytically, as they need to rely on Newton-Raphson: a simple iterative numerical method to approximate MLE. In practice \hat{r} , $\hat{\lambda}$, $\hat{\theta}$ and $\hat{\gamma}$ are the solution of the estimating equations obtained by differentiating the likelihood in terms of r , λ , θ and γ and solving them to zero. Therefore, \hat{r} , $\hat{\lambda}$, $\hat{\theta}$ and $\hat{\gamma}$ will be obtained by maximizing the log-likelihood function using a numerical iterative method.

6. Results and Discussion

The results of the proposed distribution in this study show that the NB-CR distribution is a new mixed distribution to extend the NB distribution and CR distribution studied by [4]. The NB-CR distribution represents generalization of distribution for count data, including the NB-IG [5], NB-BS and NB-LBIG distribution.

For one application of NB-CR distribution is applied on a sample of count data. We used a data set which was obtained from [8], provides information on 9,461 automobile insurance policies whereby the number of accidents of each policy has been recorded. The Poisson, NB and NB-CR distributions are fitted to the data. We summarize the observed and expected frequencies, grouped in classes of expected frequency greater than five for the chi-square goodness of fit test (Table 1). Based on the p-value, the NB-CR distribution provides a better fit compared to the Poisson and the NB for count data where the probability at zero has a large value.

Table 1: Observed and expected frequencies of accident data

No.of claims	No.of drivers	Fitting distribution		
		Poisson	NB	NB-CR
0	7840	7638.3	7841.2	7842.2
1	1317	1634.6	1291.7	1313.2
2	239	174.9	258.1	234.2
3	42	12.5	54.6	51.2
4	14	0.7	11.9	13.6
5	4	0	2.6	4.2
6	4	0	0.6	1.5
7	1	0	0.2	0.6
8+	0	0	0.1	0.3
Estimated parameters		$\hat{\lambda} = 0.214$	$\hat{r} = 0.701$ $\hat{p} = 0.765$	$\hat{r} = 3.036$ $\hat{\lambda} = 1.269$ $\hat{\theta} = 0.052$ $\hat{\gamma} = 0.909$
Chi-squares		293.803	8.657	2.623
d.f.		2	2	1
p-value		< 0.001	0.014	0.105

7. Conclusion

This paper offers NB-CR distribution which is obtained by mixing the NB distribution with a CR distribution. We find that the NB-IG distribution, NB-BS distribution and NB-LBIG distribution are all special cases of this distribution. We have derived several properties of the NB-CR distribution which includes the factorial moments, mean, variance, skewness and kurtosis. In addition, the parameters estimation of the NB-CR distribution via maximum likelihood method are also implemented. Finally, we have compared efficiencies of the NB-CR distribution with the Poisson and the NB distribution, fitting distribution by using real data. Based on the results, it is shown that the NB-CR distribution provides a better fit compared to the Poisson and the NB distribution.

Acknowledgments

This work was supported by research funding from Assumption University, and their contribution is gratefully acknowledged. The authors also wish to thank the anonymous reviewers for their valuable suggestions.

References

- [1] N. Alexander, R. Moyeed, J. Stander, Spatial modelling of individual-level parasite counts using the negative binomial distribution, *Biostatistics*, **1** (2000), 453-463, **doi:** 10.1093/biostatistics/1.4.453.
- [2] N. Balakrishnan, V. Nevzorov, *A Primer on Statistical Distributions*, John Wiley and Sons, USA (2003).
- [3] J. P. Boucher, M. Denuit, M. Guillen, Models of insurance claim counts with time dependence based on generalization of poisson and negative binomial distributions, *Variance*, **2**, No. 1 (2008), 135-162.
- [4] P. Bowonrattanasat, K. Budsaba, Some properties of the three-parameter Crack distribution, *Thailand Statistician*, **9**, No. 2 (2011), 195-203.
- [5] E. Gómez and J. María and E. Calderín, Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications, *Insurance: Mathematics and Economics*, **42** (2008), 39-49, **doi:** 10.1016/j.insmatheco.2006.12.001.

- [6] J. Hausman, B.H. Hall, Z. Griliches, Econometric models for count data with an application to the patents R-D relationship, *Econometrica*, **52** (1984), 909-938, **doi:** 10.2307/1911191.
- [7] D. Hoffman, Negative binomial control limits for count data with extra-Poisson variation, *Pharmaceut. Statist*, **2** (2003), 127-132, **doi:** 10.1002/pst.51.
- [8] S.A. Klugman, H.H. Panjer, G.E. Willmot, *Loss Models: From Data to Decision*, John Wiley and Sons, USA (2008).
- [9] Z. lord, S.R. Geedipally, The negative binomial-Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros, *Accident Analysis and Prevention*, **43** (2011), 1738-1742, **doi:** 10.1016/j.aap.2011.04.004.
- [10] M. Poch, F. Mannering, Negative binomial analysis of intersection accident frequency, *Journal of Transportation Engineering*, **122** (1996), 105-113, **doi:** 10.1061/(ASCE)0733-947X(1996)122:2(105).
- [11] C. Pudprommarat, W. Bodhisuwan, A new mixed negative binomial distribution, *Journal of Applied Sciences*, **12**, No. 17 (2012), 1853-1858, **doi:** 10.3923/jas.2012.1853.1858.
- [12] W. Rainer, *Econometric Analysis of Count Data*, Springer-Verlag, Berlin (2000).
- [13] L. Tremblay, Using the Poisson inverse Gaussian in bonus-malus systems, *Astin Bulletin*, **22**, No. 1 (1992), 97-106, **doi:** 10.2143/AST.22.1.2005129.
- [14] Z. Wang, One mixed negative binomial distribution with application, *J. Stat. Plann. Inference*, **141** (2011), 1153-1160, **doi:** 10.1016/j.jspi.2010.09.020.
- [15] G.E. Willmot, The Poisson-inverse Gaussian distribution as an alternative to the negative binomial, *Scandinavian Actuarial Journal*, (1987), 113-127.
- [16] H. Zamani, N. Ismail, Negative binomial-Lindley distribution and its application, *Journal of Mathematics and Statistics*, **6**, No. 1 (2010), 4-9.