Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Characterisation of the multivariate negative binomial-generalised exponential distribution

Winai Bodhisuwan [1,*] and Sirinapa Aryuyuen [2]

[1] Department of Statistics, Faculty of Science, Kasetsart University, Chatuchak, Bangkok, 10900, Thailand
[2] Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Rangsit-Nakhonnayok Road, Thanyaburi, Pathum Thani 12110, Thailand
* Corresponding author, e-mail: fsciwnb@ku.ac.th

Received: 16 March 2015 / Accepted: 24 July 2016 / Published: 29 July 2016

Abstract: The negative binomial-generalised exponential distribution was recently developed. A multivariate negative binomial-generalised exponential (MNB-GE) distribution is consequently introduced and applied in a multivariate count data analysis. Some probabilistic properties of the proposed distribution are studied. A bivariate negative binomial-generalised exponential distribution is also shown as a special case of the MNB-GE distribution. Joint probability functions and characteristics of the proposed distributions are derived. We also consider both dependent and independent bivariate random variables. The maximum likelihood estimation technique is used to estimate the parameters of the proposed distributions. Furthermore, the application of accidents data is illustrated for both the univariate and bivariate versions.

Keywords: bivariate count data, multivariate count data, negative binomial-generalised exponential, over-dispersion, mixed negative binomial

---

INTRODUCTION

The distribution mixture defines one of the most important ways to obtain new probability distributions in applied probability and operational research. For a distribution set, a negative binomial-generalised exponential (NB-GE) distribution has been used to model the number of rare events that occur at one time and in one area, region, volume or space, and for which the NB-GE distribution is a mixed negative binomial (NB) distribution. Examples include the number of automobile liability policies for private cars, the number of telephone calls within a business, and the number of accidents at an intersection. For the fitting of count data, the NB-GE distribution is an

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

alternative to the Poisson distribution, especially when the data present problems of over-dispersion [1].

Kocherlakota and Kocherlakota [2] presented a bivariate negative binomial distribution under various forms of 'chance mechanism'. Kotz et al. [3] reviewed the bivariate negative multinomial or binomial distribution systematically in connection with historical remarks and applications. One of the most prominent applications is presented by Lundberg [4].

A new bivariate mixed negative binomial distribution, i.e. a bivariate negative binomial-generalised exponential distribution, is introduced in this work.

When there are multiple random variables associated with an experiment or process, we usually denote them as $X_1, X_2, ..., X_k$, which is a multivariate (*k*-variate) distribution. G´omez-D´eniz et al. [5] proposed a new compound negative binomial distribution by mixing the *p* negative binomial parameter with an inverse Gaussian distribution. It provides a tractable model with attractive properties, which makes it suitable for application not only in the insurance setting, but also in other fields where overdispersion is observed. A multivariate version of the negative binomial-inverse Gaussian distribution was also introduced and some examples of application for both univariate and bivariate cases were given [5]. Liu and Tian [6] proposed a multivariate zero-inflated Poisson (ZIP) distribution, called Type I multivariate ZIP distribution, to model correlated multivariate count data with extra zeros. Two real data sets, the defect data of Nortel's telecommunications products and the Lacistema aggregatum and Protium guianense data, were used to illustrate the Type I multivariate ZIP distribution.

In this paper we propose a new multivariate version of the NB-GE distribution. Joint probability functions and some characteristics of the distribution are presented. The maximum likelihood estimation (MLE) is used to estimate the parameters of the distribution. Finally, an application of the univariate and bivariate versions of the NB-GE distribution is illustrated.

METHODS

Basic Results of Univariate Version of NB-GE Distribution

As discussed above, let X be a random variable of the NB-GE distribution, denoted as X ~NB-GE(r,α,β), which is a mixture of the NB distribution with parameters r and $p = e^{-\lambda}$ and a generalised exponential (GE) distribution with positive parameters α and β [7], i.e.

$$X \mid \lambda \sim \text{NB}\left(r, p = e^{-\lambda}\right) \text{ and } \lambda \sim \text{GE}\left(\alpha, \beta\right). \tag{1}$$

The probability mass function (pmf) of X is given by

$$f(x) = \binom{r+x-1}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \frac{\Gamma(\alpha+1)\Gamma\left(1+\dfrac{r+j}{\beta}\right)}{\Gamma\left(\alpha+\dfrac{r+j}{\beta}+1\right)}, \; x = 0,1,2,\ldots, \tag{2}$$

where $r,\alpha,\beta > 0$ and $\Gamma(\cdot)$ is a gamma function denoted as $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} \text{dx}$, for $t > 0$.

The factorial moment of order k is given by

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

$$\mu'_k(\mathrm{X}) = \frac{\Gamma(r+x)}{\Gamma(r)} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \frac{\Gamma(\alpha+1)\Gamma\left(1-\frac{k-j}{\beta}\right)}{\Gamma\left(\alpha-\frac{k-j}{\beta}+1\right)}, \; k = 1,2,\ldots. \tag{3}$$

The mean and variance are

$$\mathrm{E}(X) = r(\delta_{(1)} - 1) \text{ and } \mathrm{Var}(X) = \sigma_X^2 = r\left(r+1\right)\delta_{(2)} - r\left(r\delta_{(1)}+1\right)\delta_{(1)}, \tag{4}$$

where

$$\delta_{(u)} = \frac{\Gamma(\alpha+1)\Gamma(1-u/\beta)}{\Gamma(\alpha-u/\beta+1)}. \tag{5}$$

Theorem 1. Let X ~NB-GE(r,α,β), λ~GE(α,β) and $\breve{X}$~NB$\left(r, p=\left[\mathrm{E}(e^{\lambda})\right]^{-1}\right)$. Then

(a) $\mathrm{E}\left(\breve{X}\right) = \mathrm{E}(X)$ and $\mathrm{Var}\left(X\right) > \mathrm{Var}\left(\breve{X}\right)$,

(b) $\mathrm{Var}\left(X\right) > \mathrm{E}\left(X\right)$.

Proof. Explicitly, $\mathrm{E}(e^{\lambda}) > 1$; then $p = \left[\mathrm{E}(e^{\lambda})\right]^{-1} = 1/\mathrm{E}(e^{\lambda})$. Now, the same mean of X and $\breve{X}$ is $\mathrm{E}(X) = \mathrm{E}\left[\mathrm{E}(X \mid \lambda)\right] = \mathrm{E}(\breve{X}) = r\left[\mathrm{E}(e^{\lambda})-1\right]$ and the variance of X and $\breve{X}$ are, respectively,

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}\left[\mathrm{Var}(X \mid \lambda)\right] + \mathrm{Var}\left[\mathrm{E}(X \mid \lambda)\right] \\ &= r\left[\mathrm{E}(e^{2\lambda}) - \mathrm{E}(e^{\lambda})\right] + r^2\mathrm{Var}(e^{\lambda}), \\ \mathrm{Var}(\breve{X}) &= r\left[\mathrm{E}(e^{\lambda})-1\right]\mathrm{E}(e^{\lambda}).\end{aligned}$$

We then obtain that $\mathrm{Var}\left(X\right) > \mathrm{E}\left(X\right)$ and consequently,

$$\begin{aligned}\mathrm{Var}(X) - \mathrm{Var}(\breve{X}) &= r\left[\mathrm{E}(e^{2\lambda}) - \mathrm{E}(e^{\lambda})\right] + r^2\mathrm{Var}\left(e^{\lambda}\right) - r\left[\mathrm{E}\left(e^{\lambda}\right)-1\right]\mathrm{E}\left(e^{\lambda}\right) \\ &= r\left[\mathrm{E}(e^{2\lambda}) - \left(\mathrm{E}(e^{\lambda})\right)^2\right] + r^2\mathrm{Var}\left(e^{\lambda}\right) \\ &= \left(r^2+r\right)\mathrm{Var}(e^{\lambda}) > 0.\end{aligned}$$

Bivariate Negative Binomial-Generalised Exponential (BNB-GE) Distribution

In this section we present the joint and conditional probability mass functions of dependent and independent BNB-GE distribution. The joint pmf can be used to infer the conditional pmf, so consequently we can use the conditional pmf to find the joint probabilities of events that both occur. Moreover, in probability theory a conditional probability measures an event's probability given that another event has occurred. Therefore, the conditional probability mass functions are useful for updating information of the event based upon the knowledge of other related events.

Let X_1 and X_2 be discrete random variables defined in the same probability space $\left(\Omega, \mathcal{F}, P\right)$, consisting of a sample space Ω, a σ-field $\mathcal{F}$ of subsets of Ω, and a probability measure on $\mathcal{F}$. If we have $X_1, X_2 \in \Omega$, then we will collect data for the random variables X_1 and X_2. We wish to analyse the dependencies between these random variables, where X_1 and X_2 are independent if and only if

$$f\left(x_1,x_2\right) = f\left(x_1\right)\cdot f\left(x_2\right); \quad \forall x_1 \in \Omega_{X_1}, \forall x_2 \in \Omega_{X_2}.$$

If X_1 and X_2 are independent random variables, then the covariance of the two variables is

$$\sigma_{X_1,X_2} = Cov\left(X_1,X_2\right) = 0 \text{ or } \mathrm{E}\left(X_1X_2\right) = \mathrm{E}\left(X_1\right)\mathrm{E}\left(X_2\right).$$

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

Dependent BNB-GE distribution

Let X_1 and X_2 be dependent NB-GE random variables. The joint probability function of the dependent BNB-GE random variables is defined in Definition 1, and some graphs of the joint pmf of the dependent BNB-GE are shown in Figure 1.

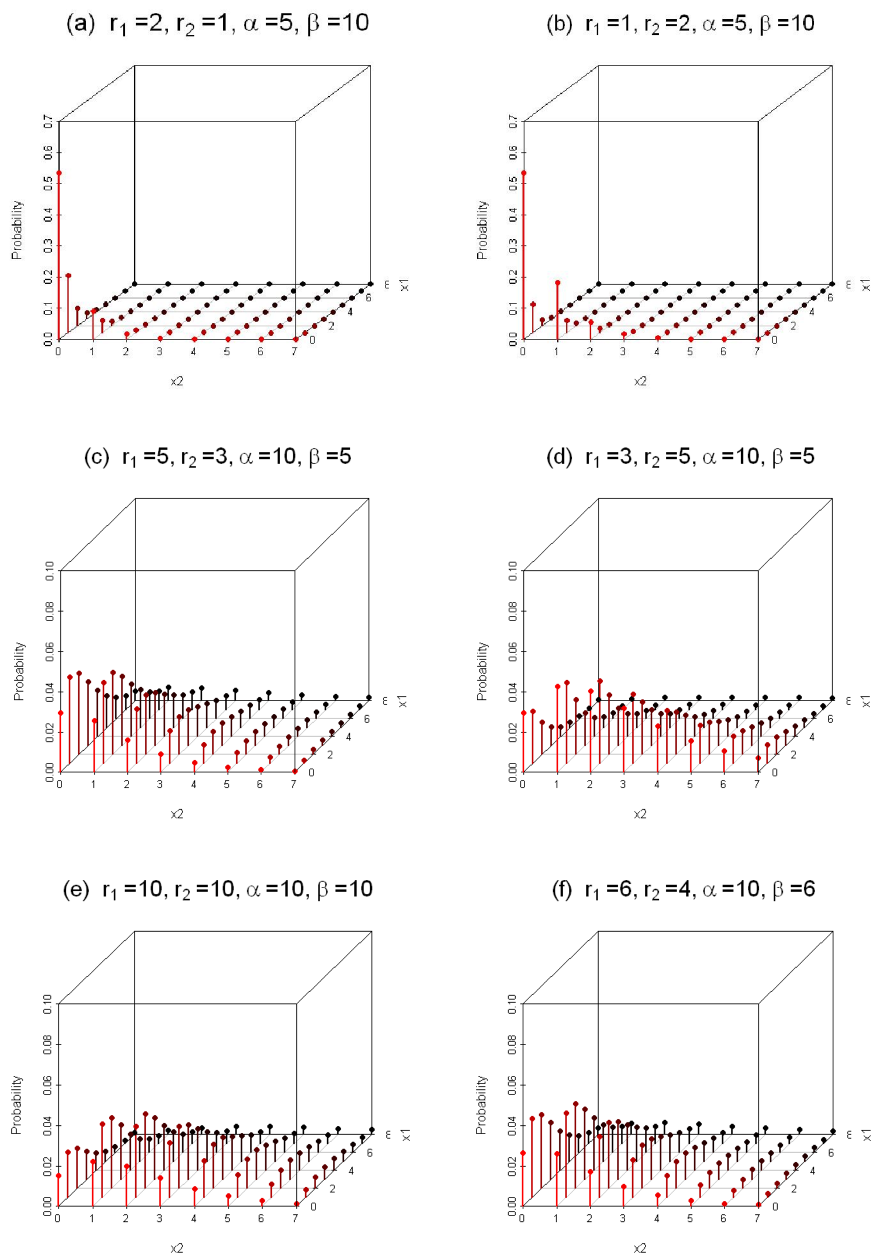

Figure 1. The pmf of dependent BNB-GE random variables X_1 and X_2 with specified parameters

Definition 1. Let X_1 and X_2 be dependent random variables. A bivariate negative binomial-distribution (X_1, X_2) is defined by the stochastic representation:

$$X_i \mid \lambda \sim \mathrm{NB}(r_i, e^{-\lambda}), \; i = 1,2 \text{ and } \lambda \sim \mathrm{GE}(\alpha, \beta).$$

Using this definition and arguments similar to those used in the basic results of the univariate version of NB-GE distribution, we obtain the joint pmf of X_1 and X_2 given by

$$f\left(x_1,x_2\right)=\binom{r_1+x_1-1}{x_1}\binom{r_2+x_2-1}{x_2}\sum_{j=0}^{\tilde{x}}\binom{\tilde{x}}{j}(-1)^j\frac{\Gamma\left(\alpha+1\right)\Gamma\left(1+\frac{\tilde{r}+j}{\beta}\right)}{\Gamma\left(\alpha+\frac{\tilde{r}+j}{\beta}+1\right)}, \tag{6}$$

where $x_i=0,1,...,\ i=1,2,\ \tilde{x}=x_1+x_2,\ \tilde{r}=r_1+r_2,$ and $r_1,r_2,\alpha,\beta>0.$

Theorem 2. Let X_1 and X_2 be dependent random variables for the joint pmf as in (6). Then the factorial moment of order k is

$$\mu_k'\left(X_1X_2\right)=\frac{\Gamma\left(r_1+k\right)\Gamma\left(r_2+k\right)}{\Gamma\left(r_1\right)\Gamma\left(r_2\right)}\sum_{j=0}^{2k}\binom{2k}{j}(-1)^j\frac{\Gamma\left(\alpha+1\right)\Gamma\left(1-\frac{2k-j}{\beta}\right)}{\Gamma\left(\alpha-\frac{2k-j}{\beta}+1\right)}, \tag{7}$$

where $k=1,2,...$ and $r_1,r_2,\alpha,\beta>0.$

Proof. If X_i~NB$(r_i,p=e^{-\lambda})$, $i=1,2$ and λ~GE(α,β), then the factorial moment of order k of X can be obtained by

$$\begin{aligned}\mu_k'\left(X_1X_2\right)&=\mathrm{E}_\lambda\left[\mu_k'\left(X_i\mid\lambda\right)\right]\\&=\mathrm{E}_\lambda\left[\frac{\Gamma\left(r_1+k\right)\left(1-e^{-\lambda}\right)^k}{\Gamma\left(r_1\right)e^{-\lambda k}}\times\frac{\Gamma\left(r_2+k\right)\left(1-e^{-\lambda}\right)^k}{\Gamma\left(r_2\right)e^{-\lambda k}}\right]\\&=\mathrm{E}_\lambda\left[\frac{\Gamma\left(r_1+k\right)\Gamma\left(r_2+k\right)}{\Gamma\left(r_1\right)\Gamma\left(r_2\right)}\left(\frac{1-e^{-\lambda}}{e^{-\lambda}}\right)^{2k}\right]\\&=\frac{\Gamma\left(r_1+k\right)\Gamma\left(r_2+k\right)}{\Gamma\left(r_1\right)\Gamma\left(r_2\right)}\mathrm{E}_\lambda\left[\left(e^{\lambda}-1\right)^{2k}\right].\end{aligned}$$

Using a binomial expansion $\left(e^{\lambda}-1\right)^{2k}$, $\mu_k'\left(X_1X_2\right)$ can be written as

$$\begin{aligned}\mu_k'\left(X_1X_2\right)&=\frac{\Gamma\left(r_1+k\right)\Gamma\left(r_2+k\right)}{\Gamma\left(r_1\right)\Gamma\left(r_2\right)}\sum_{j=0}^{2k}\binom{2k}{j}(-1)^j\mathrm{E}_\lambda\left(e^{\lambda(2k-j)}\right)\\&=\frac{\Gamma\left(r_1+k\right)\Gamma\left(r_2+k\right)}{\Gamma\left(r_1\right)\Gamma\left(r_2\right)}\sum_{j=0}^{2k}\binom{2k}{j}(-1)^j\mathrm{M}_\lambda\left(2k-j\right)\end{aligned}$$

From the moment generating function of GE distribution, $\mathrm{M}_\lambda(t)=\frac{\Gamma\left(\alpha+1\right)\Gamma\left(1-t/\beta\right)}{\Gamma\left(\alpha-t/\beta+1\right)}=\delta_{(t)}$, where t is replaced with $2k-j$. Then we obtain $\mu_k'\left(X_1X_2\right)$, which can be written as shown in (7).

From the factorial moment in (7), we also obtain the mean and covariance of X_1 and X_2 respectively as:

$$\mathrm{E}(X_1X_2)=r_1r_2\left[\delta_{(2)}-2\delta_{(1)}+1\right]\text{ and }\sigma_{X_1,X_2}=r_1r_2\left(\delta_{(2)}-\delta_{(1)}^2\right), \tag{8}$$

where $\delta_{(u)}$ is defined in (5). Note that $\delta_{(2)}-\delta_{(1)}^2=\mathrm{Var}\left(e^{\lambda}\right)$ and the correlation coefficient of X_1 and X_2 is $\rho_{X_1,X_2}=\frac{\sigma_{X_1,X_2}}{\sigma_{X_1}\sigma_{X_2}};\ 0<\rho_{X_1,X_2}<1$, where $\sigma_{X_i}=\sqrt{Var\left(X_i\right)},\ i=1,2$, which is the same as the expression in (4). The covariance and correlation matrices of X_1 and X_2 are equal to

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

$$\mathbf{\Sigma}=\begin{pmatrix}\sigma_{X_1}^2 & \sigma_{X_1,X_2}\\ \sigma_{X_1,X_2} & \sigma_{X_2}^2\end{pmatrix} \text{ and } \boldsymbol{\rho}=\begin{pmatrix}1 & \rho_{X_1,X_2}\\ \rho_{X_1,X_2} & 1\end{pmatrix}.$$

Next, we analyse the situation where a related event occurs and consider the situation where X_1 is used to explain X_2, which is called the condition distribution of X_2 given by $X_1 = x_1$. The conditional pmf of the dependent BNB-GE distribution is

$$\mathrm{P}\left(X_2 \mid X_1 = x_1\right)=\frac{f(x_1,x_2)}{f(x_1)},\ f(x_1)>0. \tag{9}$$

Independent BNB-GE distribution

Let X_1 and X_2 be independent random variables. With the stochastic representation of $X_i \mid \lambda \sim \mathrm{NB}(r_i, e^{-\lambda_i})$, $i=1,2$; $\lambda_i \sim \mathrm{GE}(\alpha_i,\beta_i)$ and arguments similar to those used in the basic results of the univariate version of NB-GE distribution , we obtain the joint pmf of X_1 and X_2 as given by Definition 2. Moreover, some graphs of the joint pmf of independent X_1 and X_2 are shown in Figure 2.

Definition 2. Let X_1 and X_2 be independent random variables. Assuming that $X_i \sim$NB-GE $(r_i, \mathrm{e}^{-\lambda_i})$, $i=1,2$; $\lambda_i \sim \mathrm{GE}(\alpha_i,\beta_i)$, then the joint pmf of X_1 and X_2 is given by

$$f(x_1,x_2)=\binom{r_1+x_1-1}{x_1}\binom{r_2+x_2-1}{x_2}\prod_{i=1}^{2}\left(\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j\frac{\Gamma(\alpha_i+1)\Gamma\left(1+\frac{r_i+j}{\beta_i}\right)}{\Gamma\left(\alpha_i+\frac{r_i+j}{\beta_i}+1\right)}\right), \tag{10}$$

where $x_i = 0,1,...,\ i=1,2,$ and $r_i,\alpha_i,\beta_i > 0.$

When X_1 and X_2 are independent, we have the mean and variance of X_i as defined in (4), and $\sigma_{X_1,X_2}=0$. The covariance and correlation matrices of X_1 and X_2 are thus equal to

$$\mathbf{\Sigma}=\begin{pmatrix}\sigma_{X_1}^2 & 0\\ 0 & \sigma_{X_2}^2\end{pmatrix} \text{ and } \boldsymbol{\rho}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

and the conditional pmf of the dependent BNB-GE distribution is given by $\mathrm{P}\left(X_2 \mid X_1 = x_1\right)= f(x_2)$.

Multivariate Negative Binomial-Generalised Exponential (MNB-GE) Distribution

We propose an MNB-GE distribution, which is a natural extension of the BNB-GE distribution. The MNB-GE distribution can be considered as a mixture of independent random variables $X_i \sim$NB-GE$(r_i, e^{-\lambda})$, $i=1,2,\ldots,m$ combined with $\lambda \sim \mathrm{GE}(\alpha,\beta)$. This proposed distribution is obtained by using a method in accordance with G´omez-D´eniz et al. [5].

Definition 3. Let $X_i \sim$NB-GE(r_i,α,β), $i=1,2,\ldots,m$ be independent and identically distributed random variables. A multivariate negative binomial-distribution $(X_1, X_2,..., X_m)$ is defined by the stochastic representation

$X_i \mid \lambda \sim \mathrm{NB}(r_i, e^{-\lambda})$, $i=1,2,...,m$,

$\lambda \sim \mathrm{GE}(\alpha,\beta)$.

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

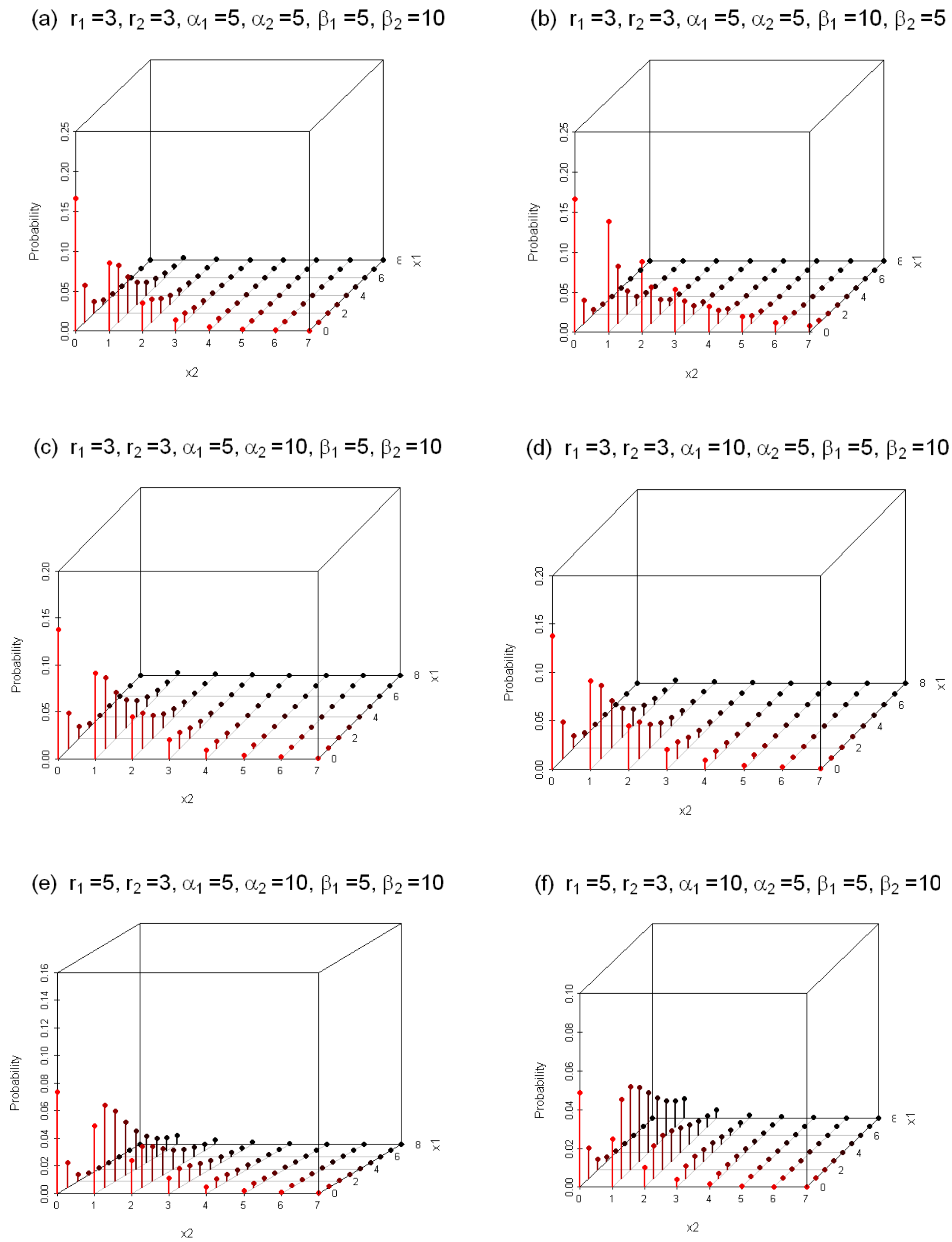

Figure 2. The pmf of independent BNB-GE random variables X_1 and X_2 with specified parameters

Using this definition and arguments similar to those used in the basic results of the univariate version of NB-GE distribution, we obtain the joint pmf of the MNB-GE distribution defined by:

$$f\left(x_1, x_2, \ldots, x_m\right) = \prod_{i=1}^{m} \binom{r_i + x_i - 1}{x_i} \left(\sum_{j=0}^{\tilde{x}} \binom{\tilde{x}}{j} (-1)^j \frac{\Gamma\left(\alpha+1\right)\Gamma\left(1+\dfrac{\tilde{r}+j}{\beta}\right)}{\Gamma\left(\alpha+\dfrac{\tilde{r}+j}{\beta}+1\right)} \right), \tag{11}$$

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

where $x_i = 0,1,...,\ i = 1,2,\ldots,m,\ \tilde{x} = x_1 + x_2 + \cdots + x_m,\ \tilde{r} = r_1 + r_2 + \cdots + r_m,$ and $r_i, \alpha, \beta > 0.$

A simpler expression of the joint pmf of the MNB-GE distribution in (11) can be written as

$$f(x_1, x_2, \ldots, x_m) = \frac{\prod_{i=1}^{m} \binom{r_i + x_i - 1}{x_i}}{\binom{\tilde{r} + \tilde{x} - 1}{\tilde{x}}} \cdot f(y; \tilde{r}, \tilde{x}), \text{ where } Y \sim \text{NB-GE}(\tilde{r}, \alpha, \beta).$$

Next, some characteristics of the MNB-GE distribution, such as the mean vector, covariance matrix and correlation coefficient matrix, are proposed respectively as follows:

$$\boldsymbol{\mu} = \begin{pmatrix} \mathrm{E}(X_1) \\ \mathrm{E}(X_2) \\ \vdots \\ \mathrm{E}(X_m) \end{pmatrix},\ \boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2_{X_1} & \sigma_{X_1,X_2} & \cdots & \sigma_{X_1,X_m} \\ \sigma_{X_1,X_2} & \sigma^2_{X_2} & \cdots & \sigma_{X_2,X_m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_1,X_m} & \sigma_{X_2,X_m} & \cdots & \sigma^2_{X_m} \end{pmatrix},\ \text{and } \boldsymbol{\rho} = \begin{pmatrix} 1 & \rho_{X_1,X_2} & \cdots & \rho_{X_1,X_m} \\ \rho_{X_1,X_2} & 1 & \cdots & \rho_{X_2,X_m} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{X_1,X_m} & \rho_{X_2,X_m} & \cdots & 1 \end{pmatrix},$$

where

$$\mathrm{E}(X_i) = r_i(\delta_{(1)} - 1),\quad i = 1,2,\ldots,m,$$

$$\sigma^2_{X_i} = r_i(r_i + 1)\delta_{(2)} - r_i(r_i\delta_{(1)} + 1)\delta_{(1)},$$

$$\sigma_{X_i,X_j} = r_i r_j(\delta_{(2)} - \delta^2_{(1)}) > 0 \text{ for } i \neq j,$$

$$\rho_{X_i,X_j} = \frac{\sigma_{X_i,X_j}}{\sigma_{X_i}\sigma_{X_j}} \text{ for } 0 < \rho_{X_i,X_j} < 1.$$

Parameter Estimation

In this section the method of parameter estimation is discussed based on the MLE technique. There are procedures for three different cases of the NB-GE distribution, viz. univariate case, dependent bivariate case and independent bivariate case.

Univariate case

The likelihood function for a random sample of size n from univariate NB-GE distribution with the pmf in (2) is provided as

$$L_u = \prod_{i=1}^{n} \binom{r + x - 1}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \frac{\Gamma(\alpha+1)\Gamma\left(1 + \frac{r+j}{\beta}\right)}{\Gamma\left(\alpha + \frac{r+j}{\beta} + 1\right)} \tag{12}$$

Thus, by taking the logarithm of (12), the log-likelihood function is given by

$$\ell_u = \sum_{i=1}^{n} \left[\log\Gamma(r + x_i) - \log\Gamma(r) - \log\Gamma(x_i + 1) + \log\Gamma(\alpha + 1) + \log\xi(r + j, \alpha, \beta)\right],$$

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

where $\xi(r+j,\alpha,\beta)=\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j\dfrac{\Gamma(\alpha+1)\Gamma\left(1+\frac{r+j}{\beta}\right)}{\Gamma\left(\alpha+\frac{r+j}{\beta}+1\right)}$.

In order to obtain the unit score vector $\mathbf{U}_u=\left(\frac{\partial\ell_u}{\partial r},\frac{\partial\ell_u}{\partial\alpha},\frac{\partial\ell_u}{\partial\beta}\right)^T$, we take the derivative of the log-likelihood ℓ_u with respect to r, α and β. Then equating the unit score vector to zero results in a non-linear equation that can be solved numerically through many well-known statistical softwares. Therefore, the non-linear model function in the R language [8] is employed to find the estimates of $(r,\ \alpha,\ \lambda)^T$.

Bivariate cases

Next, two kinds of bivariate case are considered. Suppose X_1 and X_2 are distributed as the BNB-GE distribution. The likelihood functions can be computed from the joint pmf in (6) and (10). Consequently, the log-likelihood functions can be written as follows.

For the dependent bivariate case, we obtain from the pmf of the dependent BNB-GE in (6) that the log-likelihood function is

$$\ell_{db}=\sum_{i=1}^{n}\Big[\log\Gamma(r_1+x_{1i})-\log\Gamma(r_1)-\log\Gamma(x_{1i}+1)+\log\Gamma(r_1+x_{1i})-\log\Gamma(r_2) -\log\Gamma(x_{2i}+1)+\log\Gamma(\alpha+1)+\log\xi(\tilde{r}+j,\alpha,\beta)\Big], \tag{13}$$

where $\xi(\tilde{r}+j,\alpha,\beta)=\sum_{j=0}^{\tilde{x}}\binom{\tilde{x}}{j}(-1)^j\dfrac{\Gamma(\alpha+1)\Gamma\left(1+\frac{\tilde{r}+j}{\beta}\right)}{\Gamma\left(\alpha+\frac{\tilde{r}+j}{\beta}+1\right)}$, $\tilde{r}=r_1+r_2$ and $\tilde{x}=x_1+x_2$.

For the independent bivariate case, the log-likelihood function of the independent BNB-GE in (10) is given by

$$\ell_{idb}=\sum_{i=1}^{n}\Big[\log\Gamma(r_1+x_{1i})-\log\Gamma(r_1)-\log\Gamma(x_{1i}+1)+\log\Gamma(\alpha_1+1)+\log\xi(r_1+j,\alpha_1,\beta_1)\Big] +\sum_{i=1}^{n}\Big[\log\Gamma(r_2+x_{2i})-\log\Gamma(r_2)-\log\Gamma(x_{2i}+1)+\log\Gamma(\alpha_2+1)+\log\xi(r_2+j,\alpha_2,\beta_2)\Big]. \tag{14}$$

In the same manner as mentioned above, the unit score vector associated with the log-likelihood functions in (13) and (14) are equated to zero and solved numerically to obtain the estimates of parameters.

RESULTS AND DISCUSSION

For the application of this study, one example of dataset is used to fit count data to the proposed distribution based on a bivariate case. The number of accident proneness of 122 experienced railroad hunters is used, as appearing in Dunn [9]. Let X_1 refer to the number of accidents suffered by an individual in the 6-year period between 1937-1942 and X_2 refer to the number of accidents suffered by the individual in the following 5 years from 1943 to 1947.

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

Table 1 shows the observed and expected values of the accidents and provides the estimates for parameters in (6) for this data set, where $\hat{r}_1$=13.1444, $\hat{r}_2$=20.1991, $\hat{\alpha}$=3.3105 and $\hat{\beta}$=29.9360. The estimated parameters were obtained using the MLE procedure, along with the non-linear model function in R language [8]. We found that the BNB-GE distribution appropriately fits the dependent bivariate count data. In addition, plots of observed and expected frequencies of accident among 122 experienced railroad hunters by fitting the data with the dependent BNB-GE distribution as shown in Figure 3.

Table 1. Observed and expected (in brackets) values of accidents among 122 experienced railroad hunters

In years 1943-47	**In years 1937-42 (X_1)**				Total of X_2
(X_2)	0	1	2	3-7	
0	21 (25.6)	18 (18.2)	8 (9.0)	3 (6.4)	50 (59.2)
1	13 (11.8)	14 (11.2)	10 (6.9)	6 (6.4)	43 (36.3)
2	4 (3.9)	5 (4.6)	4 (3.4)	4 (4.1)	17 (16.0)
3-7	-	4 (4.0)	4 (2.2)	4 (4.3)	12 (10.5)
Total of X_1	38 (41.3)	41 (38.0)	26 (21.5)	17 (21.2)	
Parameter estimates	$\hat{r}_1$ = 13.1444, $\hat{r}_2$ = 20.1991, $\hat{\alpha}$ = 3.3105 and $\hat{\beta}$ = 29.9360				
Chi-squared test	χ^2 = 6.6252, degree of freedom = 10, p-value = 0.7603				

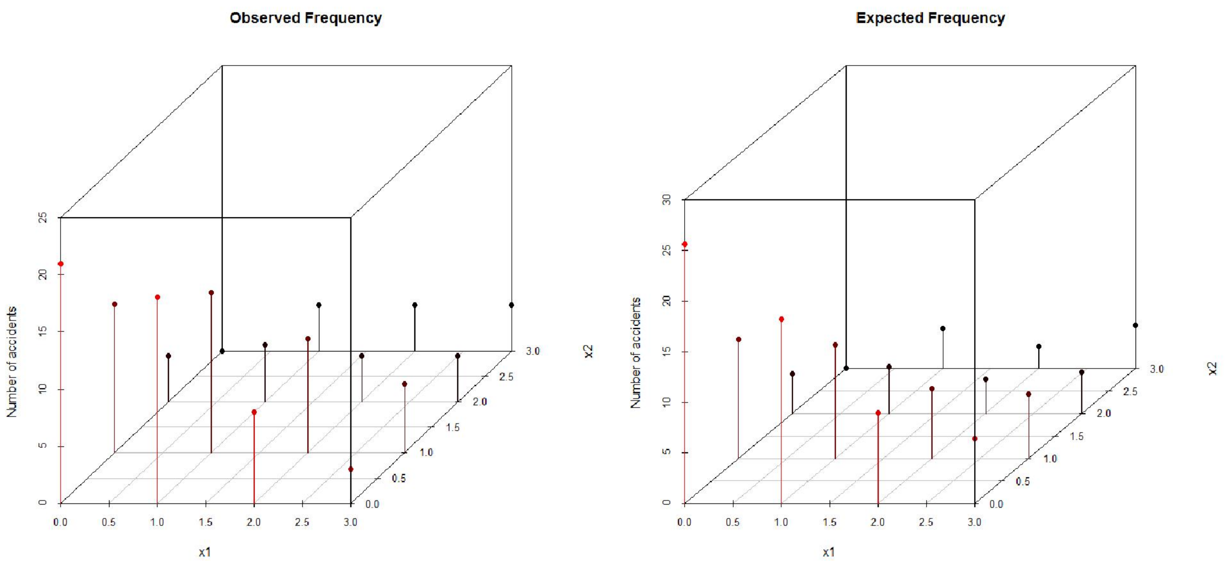

Figure 3. Plots of observed and expected frequencies of accidents among 122 experienced railroad hunters by fitting the data with dependent BNB-GE distribution

Table 2 shows the expected values of the number of accidents suffered by an individual in the 5-year period (1943-47) (X_2) when using the marginal probability function in (2) and the conditional probability function in (9). Based upon the Chi-squared statistic value, the results show that the distribution of X_2 with the conditional probability is better fitted than using the marginal probability. Figure 4 illustrates a bar chart of frequency of accidents suffered by an individual in the 5-year period, from 1943 to 1947 (X_2).

Table 2. Observed and expected values of accidents suffered by an individual in 5 years, from 1943 to 1947 (X_2)

In years 1943-47 (X_2)	**Observed value**	**Expected value by fitting distribution**	
		Marginal	**Conditional**
0	50	39.3	59.2
1	43	39.9	36.3
2	17	24.0	16.0
3-7	12	18.8	10.5
Parameter estimates		$\hat{r}_2$ =11.1158, $\hat{\alpha}$ =23.5887 $\hat{\beta}$ =34.7328	$\hat{r}_1$ =13.1444 $\hat{r}_2$ =20.1991 $\hat{\alpha}$ =3.3105 $\hat{\beta}$ =29.9360
χ^2		7.7584	2.9432

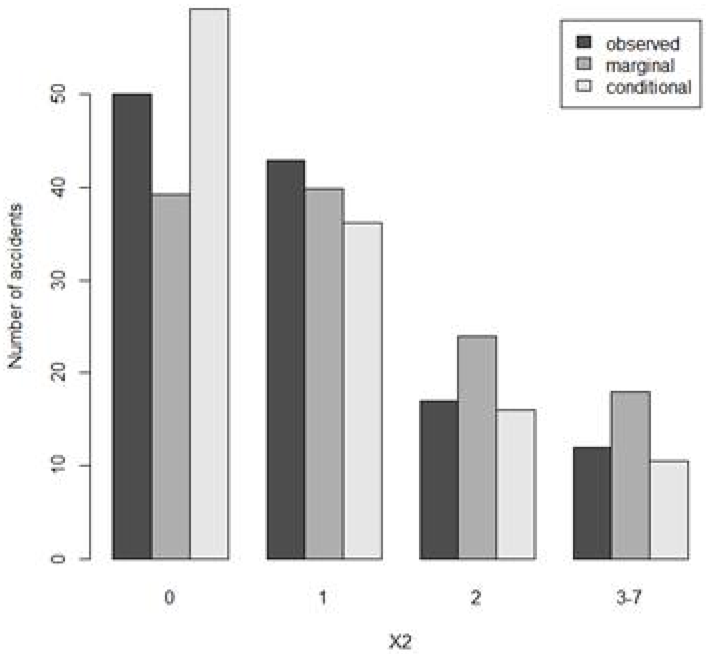

Figure 4. Frequency of accidents suffered by an individual in 5 years, from 1943 to 1947 (X_2). Expected values are obtained by fitting data with marginal and conditional NB-GE distributions**.**

CONCLUSIONS

This work proposes a new multivariate mixed negative binomial distribution which is called a multivariate negative binomial-generalised exponential distribution including the closed form. Some characteristics of the proposed distribution have been introduced. A bivariate version of the NB-GE distribution has been shown as a special case of the MNB-GE distribution, where the dependent and independent random variables are included. Parameters of the distribution are

Maejo Int. J. Sci. Technol. **2016**, *10*(02), 197-208

estimated by using the maximum likelihood estimation technique, and are computed using numerical optimisation under the non-linear model function in R language. In addition, the expected frequencies show a satisfactory goodness of fit; thus, the new distribution may be used to model the count data. Finally, the application of the univariate and bivariate random variables is illustrated.

ACKNOWLEDGEMENTS

The authors are grateful to the referees of this paper, who helped to clarify and improve the presentation. We acknowledge support from the Faculty of Science, Kasetsart University and the Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi.

REFERENCES

1. S. Aryuyuen and W. Bodhisuwan, “The negative binomial-generalised exponential (NB-GE) distribution”, *Appl. Math. Sci.*, **2013**, *7*, 1093-1105.
2. S. Kocherlakota and K. Kocherlakota, “Bivariate Discrete Distributions”, Marcel Dekker, New York, **1992**.
3. L. N. Johnson, S. Kotz, and N. Balakrishnan, “Discrete Multivariate Distributions”, Wiley, New York, **1997**.
4. O. Lundberg, “On Random Processes and Their Application to Sickness and Accident Statistics”, 2nd Edn., Almquist and Wiksell, Uppsala, **1964.**
5. E. G´omez-D´eniz, J. M. Sarabia and E. Calder´ın-Ojeda, “Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications”, *Insur. Math. Econ.*, **2008**, *42*, 39-49.
6. Y. Liu and G. L. Tian, “Type I multivariate zero-inflated Poisson distribution with applications”, *Comput. Stat. Data Anal.*, **2015**, *83*, 200-222.
7. R. D. Gupta and D. Kundu, “Generalised exponential distributions”. *Aust. NZ J. Stat.*, **1999**, *41*, 173-188.
8. The R Core Team, “R: A language and environment for statistical computing”, R Foundation for Statistical Computing, Vienna, **2013.**
9. J. E. Dunn, “Characterization of the bivariate negative binomial distribution”, *Arkansas Acad. Sci. Proc.*, **1967**, *21*, 77-86.