Available online at www.sciencedirect.com

ScienceDirect

Insurance: Mathematics and Economics 42 (2008) 39–49

INSURANCE MATHEMATICS & ECONOMICS

www.elsevier.com/locate/ime

Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications

Emilio Gómez-Déniz[a,*], José María Sarabia[b], Enrique Calderín-Ojeda[a]

[a] *Department of Quantitative Methods, University of Las Palmas de Gran Canaria, Gran Canaria, Spain*
[b] *Department of Economics, University of Cantabria, Santander, Spain*

Received September 2005; received in revised form December 2006; accepted 4 December 2006

Abstract

In this paper we propose a new compound negative binomial distribution by mixing the p negative binomial parameter with an inverse Gaussian distribution and where we consider the reparameterization $p = \exp(-\lambda)$. This new formulation provides a tractable model with attractive properties which make it suitable for application not only in the insurance setting but also in other fields where overdispersion is observed. Basic properties of the new distribution are studied. A recurrence for the probabilities of the new distribution and an integral equation for the probability density function of the compound version, when the claim severities are absolutely continuous, are derived. A multivariate version of the new distribution is proposed. For this multivariate version, we provide marginal distributions, the means vector, the covariance matrix and a simple formula for computing multivariate probabilities. Estimation methods are discussed. Finally, examples of application for both univariate and bivariate cases are given.

Classification: C10; C11; IM01; IM31

Keywords: Compound negative binomial distribution; Inverse Gaussian distribution; Overdispersion; Recursion; Premium

1. Introduction

Distribution mixtures define one of the most important ways to obtain new probability distributions in applied probability and operational research. In this sense, and looking for a more flexible alternative to the Poisson distribution, especially under the overdispersion phenomena, the negative binomial (Klugman et al., 1998; Lemaire, 1979; Simon, 1961) obtained as a mixture of Poisson and gamma distributions, negative binomial-Pareto (Klugman et al., 1998; Meng et al., 1999; Gómez and Vázquez, 2003) and Poisson-inverse Gaussian distribution (Gómez and Vázquez, 2003; Klugman et al., 1998; Tremblay, 1992; Willmot, 1987, among others) have been proposed in actuarial contexts, particularly in the automobile insurance setting. In order to provide another competitive alternative to the models above, a new mixture model is considered. We propose a new compound negative binomial distribution by mixing the p negative binomial parameter. Variations in individual claim propensity are assumed taking $p = \exp(-\lambda)$ and assuming that λ is distributed according to an inverse Gaussian distribution, obtaining the negative binomial-inverse Gaussian distribution, which we will call $\mathcal{NBIG}$. The distribution obtained can be viewed as a competitive alternative to the negative binomial and Poisson-inverse Gaussian distributions. This last distribution has been studied extensively; see Hougaard et al. (1997) and Willmot (1987, 1988).

The new distribution has thick tails, seems unimodal, positively or negatively skewed and the parameter estimators (in both univariate and multivariate models) present simple and closed expressions. Recursive expressions for computing probabilities are easily obtained. This expression is also used to obtain a recursive formula for the $\mathcal{NBIG}$ compound distribution. This recursive formula plays an important role in many insurance problems, particularly in the collective risk

* Corresponding address: Quantitative Methods in Economics, University of Las Palmas de Gran Canaria, Campus Universitario de Tafira, 35017 Las Palmas de Gran Canaria, Spain. Tel.: +34 928451803; fax: +34 928451829.
E-mail address: egomez@dmc.ulpgc.es (E. Gómez-Déniz).

0167-6687/$ - see front matter
doi:10.1016/j.insmatheco.2006.12.001

model. A multivariate version with positive correlations is presented and several properties are derived. Since modelling the claim frequency data is one of the most important areas in actuarial theory, we will apply the moment and maximum likelihood methods to fit observed claims distributions in Klugman et al. (1998) and Simon (1961). In the bivariate case we have fitted absenteeism data in Arbous and Sichel (1954) and Stein et al. (1987). Expected frequencies show a satisfactory goodness of fit. Although the inverse Gaussian distribution is not conjugated with respect to the negative binomial, posterior distribution can be easily computed. Therefore this new mixture model may be used to explain claim experience.

The contents of the paper are as follows. In Section 2 we study the basic properties of the model including the probability mass function (pmf), factorial and ordinary moments and overdispersion property. Section 3 studies the compound negative binomial-inverse Gaussian distribution. We begin by obtaining a recurrence for the probabilities of an $\mathcal{NBIG}$ distribution and then an integral equation is derived for the probability density function (pdf) of the compound version, when the claim severities are absolutely continuous, from the basic principles. Section 4 extends the univariate case to the multivariate one. For the bivariate case, conditional expectation is given in order to make a prediction from models. In Section 5 some methods of estimation of parameters are given for both the univariate and the multivariate case. Numerical examples are provided in Section 6. Afterwards, by using the Bayesian methodology, the posterior distribution is computed, as an application of this distribution in a particular insurance problem. Finally, some conclusions are presented in Section 7.

2. Basic results

In this section we introduce the definition and some basic properties of the $\mathcal{NBIG}$ distribution. We begin with two previous definitions. A classical negative binomial distribution with probability mass function:

$$\Pr(X=x)=\binom{r+x-1}{x}p^r(1-p)^x,\quad x=0,1,\ldots$$

will be denoted as $X\sim\mathcal{NB}(r,p)$, where $r>0$ and $0<p<1$. Because it will be used later, we present some characteristics of this distribution. The first three moments about zero and the factorial moment of a negative binomial distribution (see Balakrishnan and Nevzorov, 2003) are respectively given by:

$$E(X)=\frac{r(1-p)}{p},$$

$$E(X^2)=\frac{r(1-p)[1+r(1-p)]}{p^2},$$

$$E(X^3)=\frac{r(1-p)}{p^3}[1+(3r+1)(1-p)+r^2(1-p)^2],$$

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

where ($s>0$)

$$\Gamma(s)=\int_0^\infty x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$$

denotes the complete gamma function. On the other hand, a random variable Z has an inverse Gaussian distribution if its pdf is given by:

$$f(z;\mu,\psi)=\left(\frac{\psi}{2\pi z^3}\right)^{1/2}\exp\left\{-\frac{\psi(z-\mu)^2}{2\mu^2 z}\right\},\quad z>0 \tag{2}$$

where $\psi,\mu>0$. We will represent $Z\sim\mathcal{IG}(\mu,\psi)$. If $Z\sim\mathcal{IG}(\mu,\psi)$, the moment generating function is given (Tweedie, 1957) by,

$$M_Z(t)=E(\mathrm{e}^{tZ})=\exp\left[\frac{\psi}{\mu}\left(1-\sqrt{1-2\mu^2 t/\psi}\right)\right]. \tag{3}$$

Definition 1. We say that a random variable X has a negative binomial-inverse Gaussian distribution if it admits the stochastic representation:

$$X|\lambda\sim\mathcal{NB}(r,p=\mathrm{e}^{-\lambda}), \tag{4}$$

$$\lambda\sim\mathcal{IG}(\mu,\psi), \tag{5}$$

with $r,\mu,\psi>0$. We will denote this distribution by $X\sim\mathcal{NBIG}(r,\mu,\psi)$.

The next theorem provides closed formulas for the probability mass function and for the factorial moments:

Theorem 1. *Let $X\sim\mathcal{NBIG}(r,\mu,\psi)$ be a negative binomial-inverse Gaussian distribution defined in (4) and (5). Some basic properties are:*

(a) *The probability mass function is given by*

$$\Pr(X=x)=\binom{r+x-1}{x}\sum_{j=0}^{x}(-1)^j\binom{x}{j}\times\exp\left\{\frac{\psi}{\mu}\left[1-\sqrt{1+\frac{2(r+j)\mu^2}{\psi}}\right]\right\}, \tag{6}$$

with $x=0,1,2,\ldots$ and $r,\mu,\psi>0$.

(b) *The factorial moment of order k is given by*

$$\mu_{[k]}(X)=\frac{\Gamma(r+k)}{\Gamma(r)}\sum_{j=0}^{k}(-1)^j\binom{k}{j}\times\exp\left\{\frac{\psi}{\mu}\left[1-\sqrt{1-\frac{2(k-j)\mu^2}{\psi}}\right]\right\} \tag{7}$$

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

(c) *The mean, second order moment and variance are given by:*

$$E(X)=r[M_\lambda(1)-1], \tag{8}$$

$$E(X^2)=(r+r^2)M_\lambda(2)-(r+2r^2)M_\lambda(1)+r^2, \tag{9}$$

$$var(X)=(r+r^2)M_\lambda(2)-rM_\lambda(1)-r^2M_\lambda^2(1), \tag{10}$$

where $M_\lambda(u)$ is defined in (3).