# Unconditional Distributions Obtained from Conditional Specification Models with Applications in Risk Theory

**E. Gómez-Déniz[a] and E. Calderín–Ojeda[b]**

**Abstract**

Bivariate distributions, specified in terms of their conditional distributions, provide a powerful tool to obtain flexible distributions. These distributions play an important role in specifying the conjugate prior in certain multiparameter Bayesian settings. In this paper, the conditional specification technique is applied to look for more flexible distributions than the traditional ones used in the actuarial literature, as the Poisson, negative binomial and others. The new specification draws inferences about parameters of interest in problems appearing in actuarial statistics. Two unconditional (discrete) distributions obtained are studied and used in the collective risk model to compute the right-tail probability of the aggregate claim size distribution. Comparisons with the compound Poisson and compound negative binomial are made.

**Keywords**: Actuarial, Collective Risk Model, Conditional Distributions; Bivariate Distributions

# 1 Introduction

Many statisticians applying their studies in actuarial framework have examined the distribution of annual claim numbers (Willmot (1987), Kokonendji and Khoudar (2004), Meng et al. (1999), Sarabia et al. (2004), Denuit et al. (2007), Klugman et al. (2008) and Gómez-Déniz et al. (2008a, 2008b); among others). One of the most popular models in this context is the Poisson-gamma, which is based on the assumptions that the portfolio is heterogeneous and that all policyholders have constant but unequal underlying risks of having an accident. Under the above hypotheses, it is straightforward to prove that the unconditional distribution of the number of claims follows a negative binomial distribution. In the context of count regression models, the negative binomial distribution can be thought of as a Poisson distribution with unobserved heterogeneity. This in turn can be considered a mixture of two probability distributions, Poisson and gamma. However, since the Poisson model for count data imposes the restriction that the conditional mean should equal the conditional variance, the negative binomial distribution derived is open to criticism, because the random count process should exhibit overdispersion (i.e. distributions of counts often have a variance greater than the mean). Thus, alternative models to the standard Poisson distribution have been proposed. The binomial-beta and the Poisson-inverse Gaussian have the advantage of using conjugate distributions, which is an attractive feature in the scenario of credibility theory. Indeed, the use of the bivariate distribution by conditional specification is not a novelty in this field (see Sarabia et al. (2004, 2005), Sarabia and Guillén (2008) and Gómez-Déniz et al. (2009)). In the latter paper, a Bayesian robustness methodology was also developed showing that the premiums obtained under a conditional specification model are much more robust than those ones computed under classical models.

In the present paper, a new distribution is derived to predict the count data process, on the basis of the Conway-Maxwell distribution (see Ahmad (2007) and Johnson et al. (2005)) and its natural conjugate, the extended bivariate gamma distribution (see Kadane et al. (2006)). As a generalization of the former model, we seek the most general distribution of a bivariate random variable $(X, \Theta)$ such that their conditional distributions are Conway-Maxwell and gamma. This new family of distributions is very flexible and it contains, as particular cases, many other distributions proposed in the literature. It is well–known that a bivariate distribution can be specified through its conditional distributions. If we assume that conditional distributions belong to certain parametric families of distributions, it is possible to obtain the joint distribution using the methodology described in Arnold et al. (1999). See also Arnold et al. (2001) as an introduction to this topic. In order to obtain the joint distribution, it is necessary to determine the resolution of some functional equations, and by means of this methodology, highly flexible multiparametric distributions can be obtained.

The applicability of the model is shown by fitting empirical automobile frequency of claims within the framework of the collective risk model, where the pair $(X, N)$ represents the two random variables of interest. Here, $N$ is a count variable that refers to the random number of claims and $X = \sum_{i=1}^{N} X_i$ is the total claim amount (aggregate claim size). See Hernández et al. (2009), Klugman et al. (2008), Rolski et al. (1999), Sarabia and Guillén (2008) and Willmot (1987) for details on the collective risk model. As an important result, a new expression for the aggregate claim size distribution is obtained. This new expression can be considered an alternative to the classical compound Poisson model, where the individual claim size follows an exponential distribution and the number of claims is distributed according to a classical Poisson distribution.

The remainder of this paper is arranged as follows. In order to provide the necessary background, the Conway-Maxwell-extended bivariate gamma Bayesian model proposed in Kadane et al. (2006) is presented in Section 2. The new bivariate distribution, marginal and conditionals are presented in Section 3. Some interesting sub-models are shown in Section 4. Some new results in the collective risk model are given in Section 5. Section 6 includes the applications proposed and Section 7 concludes.

# 2 Background

The Conway-Maxwell-Poisson distribution (CMP) was first considered in 1962 from a need to model queuing systems with dependent service times; nevertheless, it has rarely been explored and its statistical and probabilistic properties have not been published. The CMP belongs to the family of two-parameter power series distributions (Johnson et al. (2005)) and forms an exponential family of distributions which includes as particular cases the Poisson and geometric discrete distributions and it has proved to be useful to model both under and over-dispersed data. Kadane et al. (2006) explored a Bayesian analysis of this distribution.

The CMP distribution is a generalization of the Poisson distribution, its probability mass function is

$$\Pr(X = x|\theta, \nu) = \frac{1}{Z(\theta, \nu)} \frac{\theta^x}{(x!)^\nu}, \quad \nu > 0,\, \theta > 0,\, x = 0, 1, \ldots \tag{1}$$

where

$$Z(\theta, \nu) = \sum_{j=0}^{\infty} \frac{\theta^j}{(j!)^\nu}.$$

When $\nu = 1$ a Poisson distribution is obtained. On the other hand, if $\nu = 0$, $\theta < 1$, the pmf (1) itself is geometric with parameter $0 < \theta < 1$.

Since the CMP forms an exponential family, there exists a conjugate family of prior distributions such that, whatever the data, the posterior is of the same form. For this distribution, the conjugate prior density (assuming that the two parameters are unknown and random variables themselves) is of the form (see Kadane et al. (2006)):

$$\pi(\theta,\nu) = \theta^{a-1} e^{-\nu\, b}\, Z^{-c}(\theta,\nu)\, \kappa(a,b,c), \tag{2}$$

which can be thought of as an extended bivariate gamma distribution.

If we assume $\nu$ to be known and $\lambda = e^{-\theta}$ unknown, a conjugate prior in the single-parameter exponential family of distributions also exists given in the following form (see Jewell (1974)):

$$\pi(\theta) = [c(\theta)]^{-n_0}\, \frac{\theta^{x_0}}{d(n_0, x_0)}.$$

being $c(\theta)$ and $d(n_0, x_0)$ normalizing factors.

Bayesian methodologies play an important role in many fields of research, in particular in actuarial statistics and more specifically in credibility theory. Under this approach, the premium charged to a policyholder is computed on the basis of past claims made and the accumulated past claims of the corresponding portfolio of policyholders. In order to obtain an appropriate formula for this technique, several procedures, mostly in the field of Bayesian decision-making methodologies, have been proposed in actuarial literature.

It is well–known (Jewell (1974)) that when a conjugate family of distributions is used, exact credibility expressions are obtained under the net premium calculation principle, based only on the mean value. Then, the Bayes net premium (the posterior mean of the parameter of interest) can be written as an exact credibility expression. That is,

$$\text{Bayes Net Premium } = Z(t)\, \bar{X} + (1 - Z(t))\, E(\theta),$$

where $Z(t)$ is the credibility factor, a number between 0 and 1, depending on the size of the sample, $t$ and $\bar{X} = (1/t) \sum_{i=1}^{t} X_i$. For this reason, if the CMP and its natural conjugate prior are used a credibility premium is obtained.

Detailed information about premium computation and credibility expression are provided in Gómez-Déniz et al. (2009), Hernández et al. (2009), Jewell (1974), Klugman et al. (2008), Sarabia et al. (2004, 2005) and Sarabia and Guillén (2008); among others.

Since the analytical expression of (1) includes a series, parameter estimation problems arise when the distribution is used to fit a data set. Model extensions have been proposed in recent years; for example, the Conway-Maxwell-hyper Poisson distribution proposed by Ahmad (2007). A generalization of the CMP distribution can also be obtained by using a conditional specification model and searching for the most general bivariate distribution with CMP and gamma conditional distribution. This is one of the main purposes of the remainder of the paper.

# 3 The new conditional specification model

A bivariate distribution can be specified through its conditional distributions. By assuming that both conditional distributions belong to certain parametric classes of distributions, it is possible to obtain the joint distribution using the methodology proposed by Arnold et al. (1999, 2001) (see also Sarabia and Gómez-Déniz (2008) as introduction to the topic).

The use of this kind of distribution in risk analysis and in economics is relatively new. Some applications have been provided by Sarabia et al. (2004, 2005). In addition to this, the class of bivariate distributions with lognormal conditionals has been fully characterized by Sarabia et al. (2005).

We seek the more general bivariate distribution $(X, \Theta)$ whose conditional distributions satisfy:

$$X|\Theta = \theta \sim \mathrm{CMP}(\lambda(\theta), \mu(\theta)), \tag{3}$$

$$\Theta|X = x \sim \mathcal{G}(\alpha(x), \beta(x)), \tag{4}$$

for given functions

$$\begin{aligned} \alpha(x) : \mathbb{N} &\longrightarrow \mathbb{R}^+, \\ \beta(x) : \mathbb{N} &\longrightarrow \mathbb{R}^+, \\ \lambda(\theta) : \mathbb{R}^+ &\longrightarrow \mathbb{R}^+, \\ \mu(\theta) : \mathbb{R}^+ &\longrightarrow \mathbb{R}^+. \end{aligned}$$

This question lead us to provide the following result:

**Theorem 1** *The most general bivariate distribution with conditional distributions* $X \mid \Theta = \theta \sim CMP(\lambda(\theta), \mu(\theta))$ *and* $\Theta \mid X = x \sim \mathcal{G}(\alpha(x), \beta(x))$*, is* *given by:*

$$f(x,\theta) = \frac{1}{\theta} \exp\left\{ (1 \quad x \quad -\log x!)\, \mathcal{M} \begin{pmatrix} 1 \\ -\theta \\ \log\theta \end{pmatrix} \right\}, \tag{5}$$

*where* $x = 0, 1, \ldots,$ $\theta > 0$ *and* $\mathcal{M} = (m_{ij})_{\substack{i=0,1,2 \\ j=0,1,2}}$ *is a three–dimensional matrix whose elements are constant parameters.*

**Proof:** Since both conditional distributions belong to the exponential family, this is a simple particular case of theorem 4.1 in Arnold et al. (1999). ∎

The parameter $m_{00}$ is the normalizing constant and is a function of the rest of the parameters. Therefore, the parameters $\{m_{ij}\}$ must be selected to satisfy $\int_0^\infty f_\Theta(\theta)d\theta < \infty$ or $\sum_{x=0}^{\infty} \Pr(X = x) < \infty$.

The bivariate distribution given in (5) may be rewritten as

$$f(x,\theta) = \frac{\theta^{m_{02}+m_{12}\,x-1}}{(x!)^{m_{20}-m_{21}\,\theta+m_{22}\,\log\theta}} \exp\left\{m_{00} + m_{10}\,x - (m_{01} + m_{11}\,x)\,\theta\right\}. \tag{6}$$

Henceforward, we will denote this model as the bivariate Conway-Maxwell gamma conditional distribution (BCMG); to the best of our knowledge it has not been considered previously in the literature.

Observe that the bivariate distribution (6) includes both a continuous and a discrete random variable, thus providing the possibility of modelling different types of phenomena. This kind of bivariate distribution is uncommon in the statistical literature. For instance, it is not easy to find bivariate distributions where one marginal is continuous and the other one is discrete (see for example Kotz et al. (2000) or Spanos (1999)).

The most relevant properties of the BCMG model are displayed below. The marginal distributions of $X$ and $\Theta$ are given in the next result.

**Proposition 1** *The marginal distribution of* $X$ *is given by:*

$$\Pr(X = x) = \frac{\Gamma(\alpha(x))}{\beta(x)^{\alpha(x)}} \exp\left\{m_{00} + m_{10}\,x - m_{20}\,\log x!\right\}, \tag{7}$$

*for* $x = 0, 1, \ldots,$ *and the marginal distribution of* $\Theta$ *is:*

$$f(\theta) = \theta^{m_{02}-1}\, Z(\lambda(\theta), \mu(\theta)) \exp\left\{m_{00} - m_{01}\,\theta\right\}, \tag{8}$$

*for $\theta > 0$, where*

$$\begin{aligned}
\alpha(x) &= m_{02} + m_{12}\,x - m_{22}\,\log x!, && (9)\\
\beta(x) &= m_{01} + m_{11}\,x - m_{21}\,\log x!, && (10)\\
\lambda(\theta) &= \exp\left\{m_{10} - m_{11}\,\theta + m_{12}\,\log\theta\right\}, && (11)\\
\mu(\theta) &= m_{20} - m_{21}\,\theta + m_{22}\,\log\theta. && (12)
\end{aligned}$$

**Proof:** It is straightforward to show that the marginal distribution of $X$ is obtained by integrating (6) with respect to $\theta$ and the marginal distribution of $\Theta$ is calculated by summing (6) with respect to $x$. ∎

The conditional distribution of $X$ given $\Theta = \theta$ is CMP with parameters $(\lambda(\theta), \mu(\theta))$. On the other hand, the conditional distribution of $\Theta$ given $X = x$ is gamma with parameters $(\alpha(x), \beta(x))$, where $\alpha(x)$, $\beta(x)$, $\lambda(\theta)$ and $\mu(\theta)$ are given in (9), (10), (11) and (12), respectively.

Note that the normalizing constant is given by:

$$m_{00} = -\log\left\{\int_0^\infty \theta^{m_{02}-1}\exp\{-m_{01}\theta\}\,Z(\lambda(\theta),\mu(\theta))\,d\theta\right\}. \qquad (13)$$

Now, any univariate integration rule could be used to obtain (13). For this situation, Gauss–Hermite rules (Davis and Rabinowitz, 1984; among others) could be used to approximate the normalizing constant. The conditions to ensure that (6) is a genuine probability density function (pdf henceforward) are $m_{01} > 0$, $m_{02} > 0$, $m_{11} \geq 0$, $m_{12} \geq 0$.

Finally, Bayesian methodology can be used to calculate premiums within the framework of credibility theory since a conjugate pair of likelihood and prior distribution can be built. The model obtained by assuming (3) as the likelihood and (8) as the prior distribution of $\Theta$ is not conjugate. However, by considering this new prior pdf:

$$\pi(\theta) = \mathcal{K}_3\,\left\{Z\left[\lambda(\theta),\mu(\theta)\right]\right\}^{-c}\theta^{m_{02}-1}\,e^{m_{01}\,\theta}, \qquad (14)$$

where $\mathcal{K}_3$ is a function of the parameters $m_{ij}$ and the new parameter $c$, thus extending the prior distribution considered by Kadane et al. (2006). This is shown in the next result.

**Proposition 2** *Let $f(x|\theta)$ the pdf given in (3) and the prior distribution of the parameter $\theta$ with pdf given in (14). The posterior distribution of $\theta$ given* *the sample information $\underline{X} = (X_1, X_2, \ldots, X_t)$ is as in (14), i.e. conjugate, but with the updated parameters*

$$
\begin{aligned}
m_{01}^* &= m_{01} - t\,\bar{X}\,m_{11} + m_{21}\sum_{i=1}^{t}\log X_i!, \\
m_{02}^* &= m_{02} + t\,\bar{X}\,m_{12} - m_{22}\sum_{i=1}^{t}\log X_i!, \\
m_{10}^* &= m_{10}, \quad m_{11}^* = m_{11}, \quad m_{12}^* = m_{12}, \\
m_{20}^* &= m_{20}, \quad m_{21}^* = m_{21}, \quad m_{22}^* = m_{22}, \\
c^* &= c + t,
\end{aligned}
$$

*where $\bar{X}$ is the sample mean.*

**Proof:** To see this, consider the prior distribution (14) and the sample $\underline{X} = (X_1, X_2, \ldots, X_t)$. Then the likelihood

$$
f(X_1, \ldots, X_t|\theta) = \frac{1}{\{Z[\lambda(\theta), \mu(\theta)]\}^t} \frac{\lambda(\theta)^{t\bar{X}}}{\left(\prod_{i=1}^{t} x_i!\right)^{\mu(\theta)}}
$$

and the posterior pdf is proportional to

$$
\begin{aligned}
\pi(\theta|\underline{X}) &\propto f(X_1, \ldots, X_t|\theta)\,\pi(\theta) \\
&\propto \{Z[\lambda(\theta), \mu(\theta)]\}^{-c-t}\,\theta^{m_{02}+m_{12}\,t\bar{X}-m_{22}\sum_{i=1}^{t}\log X_i!-1} \\
&\times \exp\left\{(m_{01} - m_{11}t\bar{X} + m_{21})\theta\right\}.
\end{aligned}
$$

Since $\lambda(\theta)$ and $\mu(\theta)$ do not depend on the parameters $m_{02}$ and $m_{01}$, the posterior pdf results of the same form as the pdf (14) and therefore is conjugate with respect to the pdf (3) with the updated parameters given above. ∎

# 4 Some interesting sub-models

The bivariate distribution given in (6) includes as a particular case some interesting sub-models which have been considered in the literature. For example, the choice $m_{20} = 1, \;\; m_{21} = m_{22} = 0$ provides the Poisson-gamma conditionals model, which appears in Arnold et al. (1999, p.97) and Sarabia et al. (2004).

Now, by assuming this first choice, the marginal distribution of $X$ given by (7) could be considered as a broad-based model that contains as particular cases several well-known probability distributions proposed in the literature depending on the values of its parameters:

The Poisson distribution corresponds to the choice $m_{11} = m_{12} = 0$. In this case, we have that $X \sim \mathcal{P}(e^{m_{10}})$ and the random variables $X$ and $\Theta$ in (6) are independent.

The negative binomial distribution corresponds to the choice $m_{10} = m_{11} = 0, \;\; m_{12} = 1$. In this case, $X \sim NB(m_{02}, 1 - m_{01}^{-1})$, with $m_{01} > 1$.

Furthermore, a nested negative binomial distribution is obtained by assuming $m_{10} = 0, \;\; m_{12} = 1$. This new model depends on three parameters and presents the advantage of including the negative binomial distribution as a particular case (nested). Under this nested model, the probability mass function is given by:

$$\Pr(X = x) = \frac{e^{m_{00}}}{x!} \frac{\Gamma(m_{02} + x)}{(m_{01} + m_{11}x)^{m_{02}+x}}, \quad x = 0, 1, 2, \dots \tag{15}$$

Obviously, the negative binomial distribution corresponds to the choice $m_{11} = 0$. Other models can be derived from (6) by choosing different parameter values. Since they will be used in the remainder of the paper, we have devoted two separate sections to each one of these models. They have been denoted as model S1 and model S2.

## 4.1 The sub-model S1

This sub-model corresponds to $m_{12} = m_{20} = m_{21} = m_{22} = 0$, for which the bivariate joint distribution can be written as

$$f(x, \theta) = \frac{\theta^{m_{02}-1}}{x!} \exp\left\{m_{00} + m_{10}\, x - (m_{01} + m_{11}\, x)\, \theta\right\}. \tag{16}$$

In this case, the normalization constant can be expressed in a closed form and it is given by

$$e^{m_{00}} = \mathcal{K}_1 = \frac{m_{11}^{m_{02}}}{\phi(e^{m_{10}}, m_{02}, \frac{m_{01}}{m_{11}})},$$

where $\phi(z,s,a) = \sum_{k=0}^{\infty} z^k\,(k+a)^{-s}$, is the Lerch transcendent function, which is available in Mathematica or a similar software.

The marginal distributions of $X$ and $\Theta$ are now given by the following expressions:

$$\Pr(X=x) = \frac{\mathcal{K}_1\, e^{m_{10}\,x}}{(m_{01}+m_{11}\,x)^{m_{02}}}, \quad x = 0,1,\ldots, \tag{17}$$

$$f(\theta) = \mathcal{K}_1\,\theta^{m_{02}-1}\,e^{-m_{01}\,\theta}\left(1-e^{m_{10}-m_{11}\,\theta}\right)^{-1}$$

for $\theta > 0$.

The probability generating function of (17) is

$$G_X(s) = \frac{\phi(s\,e^{m_{10}}, m_{02}, \frac{m_{01}}{m_{11}})}{\phi(e^{m_{10}}, m_{02}, \frac{m_{01}}{m_{11}})},$$

from which we can obtain its derivative of any order to obtain the factorial moments. In particular we have that the mean is given by

$$E(X) = \frac{1}{\phi(e^{m_{10}}, m_{02}, \frac{m_{01}}{m_{11}})}\left[\phi(e^{m_{10}}, m_{02}-1, \frac{m_{01}}{m_{11}}) - \frac{m_{01}}{m_{11}}\right].$$

Also, from (17) we get the ratio between successive probabilities given by

$$\frac{\Pr(X=x)}{\Pr(X=x-1)} = \left[\frac{m_{01}+m_{11}(x-1)}{m_{01}+m_{11}x}\right]^{m_{02}} e^{m_{10}}, \quad x = 1,2,\ldots \tag{18}$$

where $\Pr(X=0) = \mathcal{K}_1\,m_{01}^{-m_{02}}$.

Furthermore, from (18) it is simple to see that the distribution is unimodal with modal value at zero.

Let us prove that the pmf given in (17) is infinitely divisible.

**Proposition 3** *The discrete distribution with pmf given in (17) is infinitely divisible.*

**Proof:** Firstly, we have that $\Pr(X=0) \neq 0$, $\Pr(X=1) \neq 0$. Then, we must prove that $\{\Pr(X=j)/\Pr(X=j-1)\}$, $j = 1,2,\ldots$ forms a monotone increasing sequence. If we define $\Pr(X=x)$ also for non–integer values of $x$, we have that for $x \geq 1$

$$\frac{d}{dx}\left(\frac{\Pr(X=x)}{\Pr(X=x-1)}\right) = \frac{\Pr(X=x)}{\Pr(X=x-1)}\frac{m_{02}\,m_{11}^2}{[m_{01}+m_{11}(x-1)]\,[m_{01}+m_{11}x]} > 0.$$

Now, the result follows by applying Theorem 2.1 in Warde and Katti (1971). ■

The fact that $\{p_j/p_{j-1}\}$, $j=1,2,\ldots$, forms a monotone increasing sequence requires that $\{p_n\}$ be a decreasing sequence (see Johnson and Kotz, 1982, p.75), which is congruent with the zero vertex of the new distribution. Moreover, as any infinitely divisible distribution defined on nonnegative integers is a compound Poisson distribution (see Proposition 9 in Karlis and Xekalaki, 2005), therefore the distribution (17) presented in this subsection is a compound Poisson distribution.

Furthermore, the infinitely divisible distribution plays an important role in many areas of statistics, for example, in stochastic processes and in actuarial statistics. When a distribution $G$ is infinitely divisible then for any integer $n \geq 2$, there exists a distribution $G_n$ such that $G$ is the $n$–fold convolution of $G_n$, namely, $G = G_n^{*n}$.

Since the discrete distribution is infinitely divisible, an upper bound for the variance can be obtained for $m_{01} > 0$, $m_{02} > 0$, $m_{11} \geq 0$ and any $m_{10} \in \mathbb{R}$ (see Johnson and Kotz, 1982, p.75), which is given by

$$\mathrm{var}(X) \geq \frac{\Pr(X=1)}{\Pr(X=0)} = \left(\frac{m_{01}}{m_{01}+m_{11}}\right)^{m_{02}} e^{m_{10}}.$$

## 4.2 The sub-model S2

This another sub-model is obtained by taking

$$m_{02}=1, \quad m_{12}=m_{21}=m_{22}=0, \quad m_{20}=2.$$

whose bivariate joint distribution can be expressed as

$$f(x,\theta) = \frac{1}{(x!)^2} \exp\left\{m_{00} + m_{10}\,x - (m_{01} + m_{11}\,x)\,\theta\right\}. \tag{19}$$

In this sub–model, the normalization constant, obtained from (13), can be written in a closed form as a function of the generalized hypergeometric function in the following way, after some algebra

$$e^{m_{00}} = \mathcal{K}_2 = \frac{m_{01}}{{}_1F_2(m_{01}/m_{11}; 1, 1 + m_{01}/m_{11}; e^{m_{10}})},$$

where ${}_1F_2\,(a;b,c;x) = \sum_{j=0}^{\infty} \frac{(a)_j x^j}{(b)_j (c)_j j!}$ represents the generalized hypergeometric function and $(p)_j = \Gamma(p+j)/\Gamma(p)$ is the Pochhammer symbol.

The marginal distributions of $X$ and $\Theta$ are given by:

$$\Pr(X=x)=\frac{\mathcal{K}_2}{(x!)^2}\frac{e^{m_{10}\,x}}{m_{01}+m_{11}\,x},\quad x=0,1,\dots, \tag{20}$$

$$f(\theta) \;=\; \mathcal{K}_2\; e^{-m_{01}\,\theta}\, I_0\left(2\exp\left[\frac{1}{2}(m_{10}-m_{11}\,\theta)\right]\right),\quad \theta>0, \tag{21}$$

where $I_0(x)=\sum_{j=0}^{\infty}\frac{\left(\frac{1}{4}x^2\right)^j}{(j!)^2}$ represents the modified Bessel function of the first kind.

We are now interested in analyzing some properties of the discrete random variables given in (20):

Hence the probability generating function can be obtained as

$$G_X(s)=\frac{\mathcal{K}_2}{m_{01}}\,{}_1F_2\left(\frac{m_{01}}{m_{11}};1,1+\frac{m_{01}}{m_{11}};s\,e^{m_{10}}\right),$$

the mean is given by

$$E(X)=\frac{\mathcal{K}_2\,e^{m_{10}}}{m_{01}+m_{11}}\,{}_1F_2\left(1+\frac{m_{01}}{m_{11}};2,2+\frac{m_{01}}{m_{11}};e^{m_{10}}\right),$$

and the cumulative distribution function is

$$\begin{aligned}\Pr(X\le x) &= 1-\frac{\mathcal{K}_2\,e^{((1+x)m_{10})}}{(\Gamma(x+2))^2\,(m_{01}+(1+x)\,m_{11})}\\ &\times\;{}_2F_3\left(1,1+x+\frac{m_{01}}{m_{11}};2+x,2+x,2+x+\frac{m_{01}}{m_{11}};e^{m_{10}}\right)\end{aligned}$$

where ${}_2F_3(a,b;c,d,e;x)=\sum_{j=0}^{\infty}\frac{(a)_j(b)_jx^j}{(c)_j(d)_j(e)_jj!}$ represents the generalized hypergeometric function.

From (20) we have

$$\Pr(X=0)=\frac{\mathcal{K}_2}{m_{01}}, \tag{22}$$

and the remaining probabilities can be calculated recursively using (22) together with

$$\Pr(X=x)=\frac{e^{m_{10}}}{x^2}\left(1-\frac{m_{11}}{m_{01}+m_{11}\,x}\right)\Pr(X=x-1),\quad x=1,2,\dots$$

## 5 Results in the collective risk model

Within the framework of risk theory, the classical collective model (see Klugman et al. (2008) and Rolski et al. (1999)) we will assume that the amount of interest is the pair $(Y, X)$, where $X$ is a count variable that refers to the random number of claims and $Y = \sum_{i=1}^{X} Y_i$ is the total claim amount (aggregate claim size). The two components are usually termed the claim frequency, and the severity component and the respective densities are known as the primary and secondary distributions, respectively. In a given time period there can be several claims under a single policy. On the other hand the amount of each claim can vary in a significant way.

The natural way of modelling the pair $(Y, X)$ is in terms of the simple hierarchical model $Y|X \sim f_{Y_i}^{*x}(y)$ and $X \sim f_X(x)$, where $f_{Y_i}^{*x}(x)$ represents the $x$-fold convolution of the random variable $Y_i$.

Numerical evaluation of compound distributions is a useful tool in insurance mathematics to solve the problem of large claims made by policyholders, in particular regarding to reinsurance within the collective risk model; in this respect, two widely-used techniques are the Panjer recursion algorithm and transform methods such as Fast Fourier Transform. Although this method is very simple to implement, in practice, it presents the drawback that some improvements are required, especially for heavy-tailed distributions since aliasing errors imply the need to consider the whole distribution.

The use of the extrapolated tail, after fitting a distribution to claim size data, to estimate an extreme large claim value could sometimes lead us into possible errors since the estimated value is larger than the actual one. For that reason, it is crucial not to use a probability distribution whose tail fades away to zero rapidly. Two notorious examples that have been used in the risk theory literature are Pareto and log-normal distributions, in particular in reinsurance premium computation. For the sake of its simplicity, if the size of a single claim is described by an exponential distribution, the compound Poisson model has been commonly used. In this paper, several different models have been developed after considering that the number of claims can be explained by the pmf's introduced in the previous section.

The results presented in the following show that closed forms are obtained when the pmf (17), with $m_{02} = 1$ and (20), are assumed as primary distribution and exponential distributions as secondary distributions.

**Theorem 2** *If we assume the pmf (17) with $m_{02} = 1$ as the primary distri-*

*bution and an exponential distribution with parameter* $\gamma > 0$ *as the secondary distribution, then the pdf of the random variable* $Y = \sum_{i=1}^{X} Y_i$ *is given by*

$$
\begin{aligned}
f_S(y) &= \frac{\gamma\,(\gamma\,y)^{-m_{01}/m_{11}-1} e^{-(\gamma\,y + m_{10} m_{01}/m_{11})}}{\phi(e^{m_{10}}, 1, m_{01}/m_{11})} \\
&\times \left[\Gamma\left(1+\frac{m_1}{m_{11}}\right) - \Gamma\left(1+\frac{m_{01}}{m_{11}}, \gamma\,y\,e^{-m_{10}}\right)\right], \quad y > 0, \qquad (23)
\end{aligned}
$$

*while* $f_S(0) = \frac{m_{11}}{m_{01}\,\phi(e^{m_{10}},1,m_{01}/m_{11})}$. *Here* $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,dt$ *is the gamma function and* $\Gamma(a,z) = \int_z^\infty t^{a-1}e^{-t}\,dt$ *is the incomplete gamma function.*

**Proof:** Since the $x$-th fold convolution of the exponential distribution follows a gamma distribution with parameters $x$ and $\gamma$ (see Klugman et al. (2008) and Rolski et al. (1999))

$$
f^{*x}(y) = \frac{\gamma^x}{(x-1)!}\, y^{x-1}\, e^{-\gamma\,y},\ x = 1, 2, \ldots \qquad (24)
$$

the result follows after after simple but tedious algebra. ∎

**Theorem 3** *If we assume the pmf (20) as the primary distribution and an exponential distribution with parameter* $\gamma > 0$ *as the secondary distribution, then the pdf of the random variable* $Y = \sum_{i=1}^{X} Y_i$ *is given by*

$$
f_S(y) = \frac{\mathcal{K}_1\,\gamma\,e^{-(\gamma\,y - m_{10})}}{(m_{01}+m_{11})}\,{}_1F_3\left(1+\frac{m_{01}}{m_{11}}; 2, 2, 2+\frac{m_{01}}{m_{11}}; \gamma\,y\,e^{m_{10}}\right), \quad y > 0, \qquad (25)
$$

*while* $f_S(0) = \frac{\mathcal{K}_1}{m_{01}}$.

**Proof:** Again, by assuming the same argument the theorem holds after some computation. ∎

# 6 Applications

The most traditionally used aggregate claim model is the one obtained assuming the Poisson as primary and the exponential as secondary distribution. In this particular case (see Rolski et al. (1999) and Hernández et al. (2009);

among others), the distribution of the random variable aggregate claim size is given by

$$f_s(y) = \sqrt{\frac{\gamma\alpha}{y}} e^{-(\alpha+\gamma y)}\, I_1\left(2\sqrt{\alpha\gamma y}\right), \qquad y > 0, \tag{26}$$

with $f_S(0) = e^{-\alpha}$. In this case, $\alpha = e^{m_{10}} > 0$ and $\gamma > 0$ are the parameters of the Poisson and exponential distributions, respectively and $I_\nu(z) = \sum_{k=0}^{\infty}(z/2)^{2k+\nu}/(\Gamma(k+1)\,\Gamma(\nu+k+1)), \quad z \in \mathbb{R},\ \nu \in \mathbb{R}$, represents the modified Bessel function of the first kind. In addition to this, as a negative binomial distribution with parameters $r > 0$ and $0 < p < 1$ is assumed as primary distribution, the pdf of the random variable aggregate claim size is now given by the expression

$$f_S(y) = \gamma_1\, r\, p^r\,(1-p)e^{-\gamma\, y}\, {}_1F_1(1+r, 2, \gamma(1-p)r), \quad y > 0, \tag{27}$$

being $f_S(0) = p^r$. Recall that $r = m_{02}$ and $p = 1 - 1/m_{01}$.

As an application to real data, let us consider the claim counts of the third party liability vehicle insurance from an insurance company in Zaire (see Willmot (1987)) which correspond to claims from 4000 vehicle policies. We only have reproduced the data set with frequencies $\{0, 1, 2, \ldots\}$ and observed values $\{3719, 232, 38, 7, 3, 1\}$.

For comparison purposes, the Poisson distribution, the negative binomial distribution and the pmf given in (17) and (20) and (17) with $m_{20} = 1$ have been used to fit the data. The differences between the observed and expected frequencies were plotted and appear in Figure 1. As it can be observed from that pictorial, errors of estimation are only significant for the Poisson model. Parameters have been estimated, for all models, by the maximum likelihood method. Maximum of the log-likelihood ($\ell_{\max}$) and parameter estimates are shown in Table 1.

Those estimates have been used to compute to compute the right-tail cumulative probability of the aggregate claim distribution for the expressions (23), (25), (26) and (27). The results for different values of $\gamma$ are displayed in Tables 2 and 3. From those charts, it can be observed that expressions (23) and (25) fades away to zero slower than expressions (26) and (27), showing the efficiency of the distributions introduced in this paper to model extreme data when the severity component follows an exponential distribution. Certainly, the effectiveness of these distributions will improve by considering more heavy-tailed distributions to model the severity component.

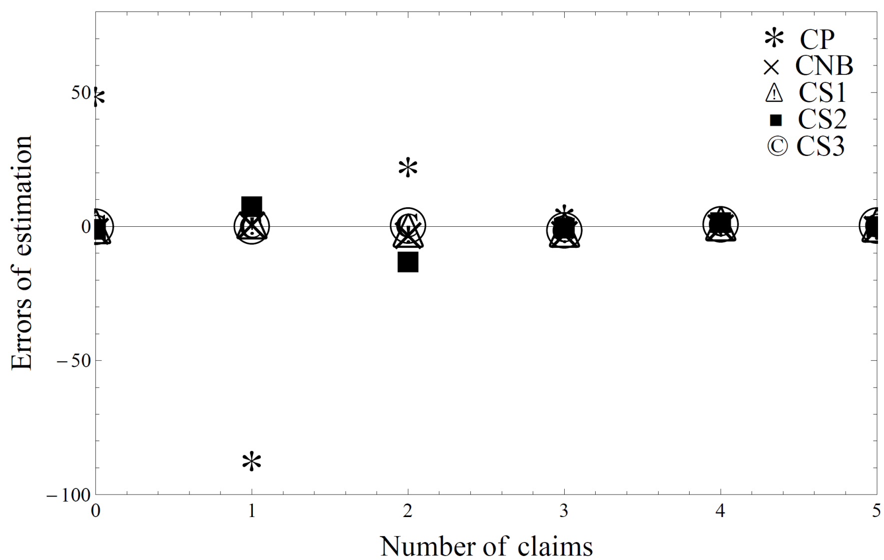


Figure 1: Errors of estimation, observed minus expected, Zaire (1974)

# 7 Conclusions

It is convenient in statistical modelling to find tractable multivariate distributions with given marginals in order to be able to quantify the dependence effect of the variables in the model. In this article, we have approached that problem by using given conditional distributions. As a result of this methodology, a bivariate distribution including both a continuous random variable and a discrete random variable has been obtained, providing the possibility of modelling different phenomena. In spite of this kind of multivariate distribution is unusual in statistical literature, it has been traditionally used in the actuarial framework, in particular within the collective risk model.

Additionally, this paper also provides a significant extension to the CMP-gamma Bayesian model used in Kadane et al. (2006).

The appropriateness of this model is determined by the simplicity and effectiveness to estimate parameters, thus providing a class of distributions

Table 1: Parameters estimates and the maximum of the log-likelihood

| | Parameter Estimates | | | | $\ell_{\max}$ |
|---|---|---|---|---|---|
| | $\widehat{m}_{01}$ | $\widehat{m}_{02}$ | $\widehat{m}_{10}$ | $\widehat{m}_{11}$ | |
| Poisson | – | – | –2.447 | | –1246.08 |
| NB | 3.496 | 0.216 | – | – | –1183.55 |
| S1 | 0.551 | 2.077 | –0.835 | 0.850 | –1183.36 |
| S1 ($m_{02} = 1$) | 0.470 | 1 | –1.177 | 1.868 | –1183.48 |
| S2 | 0.067 | – | 0.567 | 1.907 | –1189.67 |

to actuarial community that can be used in the collective risk model.

# References


Ahmad, M. (2007). A short note on Conway-Maxwell-hyper Poisson distribution. *Pakistan Journal of Statistics*, 23, 2, 135–137.

Arnold, B.C, Castillo, E., Sarabia, J.M. (1999). *Conditional Specification of Statistical Models.* Springer Series in Statistics, Springer Verlag, New York.

Arnold, B.C., Castillo, E., Sarabia, J.M. (2001). Conditionally Specified Distributions: An Introduction (with discussion). *Statistical Science*, 16, 3, 249–274.

Davis, P. and Rabinowitz, P. (1984). *Methods of numerical integration.* San Diego: Academic Press.

Denuit, M., Marechal, X., Pitrebois, S. and Walhin, J-F. (2007). *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems.* Wiley

Gómez–Déniz, E., Sarabia, J.M., Pérez–Sánchez, J.M. and Vázquez, F.J. (2008a). Using a Bayesian hierarchical model for fitting automobile claim frequency data. *Communications in Statistics: Theory and Methods*, 37, 3, 1425–1435.

Table 2: Right-tail cumulative probability of the aggregate claim distribution

Compound Poisson model

| $y$ | $\gamma = 0.10$ | $\gamma = 0.25$ | $\gamma = 0.50$ | $\gamma = 0.75$ | $\gamma = 1.00$ |
|---|---|---|---|---|---|
| 1 | 0.075298 | 0.065225 | 0.051338 | 0.040407 | 0.031802 |
| 2 | 0.068423 | 0.051338 | 0.031802 | 0.019697 | 0.012198 |
| 3 | 0.062175 | 0.040407 | 0.019697 | 0.009598 | 0.004676 |
| 4 | 0.056498 | 0.031802 | 0.012198 | 0.004676 | 0.001791 |
| 5 | 0.051338 | 0.025028 | 0.007553 | 0.002277 | 0.000686 |
| 6 | 0.046650 | 0.019697 | 0.004676 | 0.001108 | 0.000262 |
| 7 | 0.042389 | 0.015501 | 0.002894 | 0.000539 | 0.000100 |
| 8 | 0.038517 | 0.012198 | 0.001791 | 0.000262 | 0.000038 |
| 9 | 0.034999 | 0.009598 | 0.001108 | 0.000127 | 0.000014 |
| 10 | 0.031802 | 0.007553 | 0.000686 | 0.000062 | 5.612E–6 |

Compound Negative Binomial model

| $y$ | $\gamma = 0.10$ | $\gamma = 0.25$ | $\gamma = 0.50$ | $\gamma = 0.75$ | $\gamma = 1.00$ |
|---|---|---|---|---|---|
| 1 | 0.064653 | 0.057190 | 0.046630 | 0.038035 | 0.031036 |
| 2 | 0.059575 | 0.046630 | 0.031036 | 0.020687 | 0.013809 |
| 3 | 0.054900 | 0.038035 | 0.020687 | 0.011289 | 0.006179 |
| 4 | 0.050595 | 0.031036 | 0.013809 | 0.006179 | 0.002780 |
| 5 | 0.046630 | 0.025334 | 0.009231 | 0.003393 | 0.001257 |
| 6 | 0.042979 | 0.020687 | 0.006179 | 0.001868 | 0.000571 |
| 7 | 0.039616 | 0.016899 | 0.004142 | 0.001031 | 0.000260 |
| 8 | 0.036518 | 0.013809 | 0.002780 | 0.000571 | 0.000119 |
| 9 | 0.033664 | 0.011289 | 0.001868 | 0.000316 | 0.000054 |
| 10 | 0.031036 | 0.009231 | 0.001257 | 0.000176 | 0.000025 |

Table 3: Right-tail cumulative probability of the aggregate claim distribution

Compound S2 model

| $y$ | $\gamma = 0.10$ | $\gamma = 0.25$ | $\gamma = 0.50$ | $\gamma = 0.75$ | $\gamma = 1.00$ |
|---|---|---|---|---|---|
| 1 | 0.179465 | 0.161281 | 0.135051 | 0.113165 | 0.0948906 |
| 2 | 0.167123 | 0.135051 | 0.094890 | 0.066850 | 0.0472166 |
| 3 | 0.155647 | 0.113165 | 0.066850 | 0.039718 | 0.0237271 |
| 4 | 0.144976 | 0.094890 | 0.047216 | 0.023727 | 0.0120316 |
| 5 | 0.135051 | 0.079619 | 0.033431 | 0.014246 | 0.0061516 |
| 6 | 0.125820 | 0.066850 | 0.023727 | 0.008594 | 0.0031690 |
| 7 | 0.117232 | 0.056164 | 0.016877 | 0.005208 | 0.0016436 |
| 8 | 0.109243 | 0.047216 | 0.012031 | 0.003169 | 0.0008577 |
| 9 | 0.101809 | 0.039718 | 0.008594 | 0.001935 | 0.0004500 |
| 10 | 0.094890 | 0.033431 | 0.006151 | 0.001186 | 0.0002373 |

Compound S3 model

| $y$ | $\gamma = 0.10$ | $\gamma = 0.25$ | $\gamma = 0.50$ | $\gamma = 0.75$ | $\gamma = 1.00$ |
|---|---|---|---|---|---|
| 1 | 0.140517 | 0.122129 | 0.096661 | 0.076496 | 0.060530 |
| 2 | 0.127974 | 0.096661 | 0.060530 | 0.037886 | 0.023703 |
| 3 | 0.116549 | 0.076496 | 0.037886 | 0.018745 | 0.009266 |
| 4 | 0.106142 | 0.060530 | 0.023703 | 0.009266 | 0.003616 |
| 5 | 0.096661 | 0.047891 | 0.014823 | 0.004576 | 0.001409 |
| 6 | 0.088026 | 0.037886 | 0.009266 | 0.002258 | 0.000548 |
| 7 | 0.080161 | 0.029969 | 0.005790 | 0.001113 | 0.000213 |
| 8 | 0.072997 | 0.023703 | 0.003616 | 0.000548 | 0.000082 |
| 9 | 0.066473 | 0.018745 | 0.002258 | 0.000270 | 0.000032 |
| 10 | 0.060530 | 0.014823 | 0.001409 | 0.000132 | 0.000012 |

Gómez–Déniz, E., Sarabia, J.M. and Calderín, E. (2008b). Univariate and multivariate versions of the negative binomial inverse Gaussian distribution with applications. *Insurance: Mathematics and Economics*, 42, 39–49.

Gómez-Déniz, E., Sarabia, J.M. and Vázquez, F.J. (2009). Robust Bayesian bonus-malus premiums under the conditional specification model. *Statistical Papers*, 50, 465–480.

Hernández-Bastida, A., Gómez-Déniz, E., Pérez-Sánchez, J.M. (2009). Bayesian Robustness of the Compound Poisson Distribution under Bidimensional Prior: an Application to the Collective Risk Model. *Journal of Applied Statistics*, 36, 8, 853–869.

Jewell, W. (1974). Credible means are exact Bayesian for exponential families. *Astin Bulletin*, 8, 1, 77–90.

Johnson, N.L. and Kotz, S. (1982). Developments in discrete distribution, 1969–1980. *International Statistical Review*, 50, 71–101.

Johnson, N.L., Kemp, A.W., and Kotz, S. (2005). *Univariate Discrete Distributions*. Third Edition. John Wiley, New York.

Kadane, J.B., Shmueliy, G., Minkaz, T.P., Borlex, S. and Boatwright, P. (2006). Conjugate Analysis of the Conway-Maxwell-Poisson Distribution. *Bayesian Analysis*, 1, 2, 363–374.

Karlis, D., Xekalaki, E. (2005). Mixed Poisson distributions. *International Statistical Review*, 73, 35–58.

Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2008). *Loss Models: From Data to Decisions*. Third Edition. Wiley.

Kokonendji, C. and Khoudar, M. (2004). On strict arcsine distribution. *Communications in Statistics. Theory and Methods*, 33, 5, 993–1006.

Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000). *Continuous Multivariate Distributions*. Vol. 1: Models and Applications. John Wiley, New York.

Meng, S., Yuan, W., Whitmore, G.A. (1999). Accounting for individual over-dispersion in a bonus-malus automobile insurance system. *Astin Bulletin*, 29, 2, 327–337.

Rolski, T., Schmidli, H. Schmidt, V. and Teugel, J. (1999). *Stochastic processes for insurance and finance.* John Wiley & Sons.

Sarabia, J.M. and Gómez–Déniz, E. (2008). Construction of multivariate distributions: a review of some recent results. *Statistics and Operational Research Transactions (SORT)*, 32, 1, 3–36. (invited paper with discussion).

Sarabia, J.M., Gómez–Déniz, E. and Vázquez, F.J. (2004). On the use of conditional specification models in claim count distribution: an application to Bonus–Malus Systems. *Astin Bulletin*, 6, 435–463.

Sarabia, J.M., Castillo, E., Gómez–D éniz, E. and Vázquez, F.J.. (2005). A class of conjugate priors for log–normal claims based on conditional specification. *Journal of Risk and Insurance*, 72, 3, 479–495.

Sarabia, J.M. and Guillén, M. (2008). Joint modelling of the total amount and the number of claims by conditionals. *Insurance: Mathematics and Economics*, 43, 466–473.

Spanos, A. (1999). *Probability Theory and Statistical Inference.* Cambridge University Press, Cambridge.

Warde, W.D. and Katti, S.K. (1971). Infinite divisibility of discrete distributions II. *The Annals of Mathematical Statistics*, 42, 3, 1088–1090.

Willmot, G.E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial Journal*, 113–127.