DOI: 10.1111/rmir.12224

FEATURE ARTICLE

Risk Management and Insurance Review WILEY

The multivariate Poisson-Generalized Inverse Gaussian claim count regression model with varying dispersion and shape parameters

George Tzougas[1] | Despoina Makariou[2,3]

[1]Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, UK

[2]Department of Statistics, London School of Economics and Political Science, London, UK

[3]Institute of Insurance Economics, School of Management, University of St. Gallen, St. Gallen, Switzerland

Correspondence
George Tzougas, Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh EH14 4AS, UK.
Email: G.Tzougas@hw.ac.uk

Abstract

We introduce a multivariate Poisson-Generalized Inverse Gaussian regression model with varying dispersion and shape for modeling different types of claims and their associated counts in nonlife insurance. The multivariate Poisson-Generalized Inverse Gaussian regression model is a general class of models which, under the approach adopted herein, allows us to account for overdispersion and positive correlation between the claim count responses in a flexible manner. For expository purposes, we consider the bivariate Poisson-Generalized Inverse Gaussian with regression structures on the mean, dispersion, and shape parameters. The model's implementation is demonstrated by using bodily injury and property damage claim count data from a European motor insurer. The parameters of the model are estimated via the Expectation-Maximization algorithm which is computationally tractable and is shown to have a satisfactory performance.

1 | INTRODUCTION

The regression analysis of multivariate count data for capturing the dependence structures between multiple count response variables based on explanatory variables is encountered across several disciplines such as biology, biometrics, genetics, medicine, marketing, ecology, sociology, econometrics, and insurance. In general, multivariate count data models can be classified into the following three classes: multivariate Poisson models, multivariate mixed Poisson (MVMP) models, and copula-based models. For more details, the interested reader can refer to the works of Aguero-Valverde and Jovanis (2009); Aitchison and Ho (1989); Cameron et al. (2004); Cameron and Trivedi (2013); Chen and Hanson (2017); Chib and Winkelmann (2001); El-Basyouny and Sayed (2009); Famoye (2010); Genest and Nešlehová (2007); Ghitany et al. (2012); Gurmu and Elder (2000); Ho and Singer (2001); Joe (1997); Johnson et al. (1997); Jung and Winkelmann (1993); Karlis and Meligkotsidou (2005); Kocherlakota (1988); Kocherlakota and Kocherlakota (2001); Krummenauer (1998); Lakshminarayana et al. (1999); Lee (1999); Ma et al. (2008); Marra and Wyszynski (2016); M'Kendrick (1925); Munkin and Trivedi (1999); Nikoloulopoulos (2013, 2016); Nikoloulopoulos and Karlis (2010); Park and Lord (2007); Rüschendorf (2013); Silva et al. (2019); Stein and Juritz (1987); Stein et al. (1987); Winkelmann (2008); Zhan et al. (2015); Zimmer and Trivedi (2006), and Chiquet et al. (2020).

In a nonlife insurance setting, the actuary may be concerned with modeling jointly different types of claims and their associated counts. In this market segment, there are several circumstances where the interest lies in developing models which can accommodate for positively correlated claims whilst accounting for overdispersion which is a direct consequence of unobserved heterogeneity due to systematic effects in the data. Furthermore, these dependence structures between different claim types may be observed within the same insurance policy, such as property damage and bodily injury claims in motor third party liability (MTPL) insurance, or in alternative types of coverage, such as home and auto insurance, bundled together under a single policy. Regarding the latter, some of the advantages for the policyholder are multi-product premium discounts, straightforward tracking of policy renewal dates, easy claims reporting, and a more "personal" relationship between the insured and their insurer where the latter closely identify their needs and mitigate possible insurance coverage gaps. From the insurer's perspective though, bundling multiple types of insurance for the same policyholder translates into a need to develop predictive models which can efficiently capture the joint dynamics of different claims types associated with various insurance business lines. With regard to the use of alternative multivariate count models in nonlife insurance, see for instance, Abdallah et al. (2016); Bermudez et al. (2018); Bermudez and Karlis (2011, 2012, 2017); Bolancé et al. (2020); Bolancé and Vernic (2019); Denuit et al. (2019); Fung et al. (2019); Gómez-Déniz and Calderín-Ojeda (2021); Jeong and Dey (2021); Pechon et al. (2019, 2021, 2018); Shi and Valdez (2014a,b) and Tzougas and di Cerchiara (2021).

In the current study, we develop a multivariate Poisson-Generalized Inverse Gaussian (MVPGIG) regression model with varying dispersion and shape for modeling positively correlated and overdispersed claim counts from different types of coverage in a flexible manner. In particular, within the adopted modeling framework, in addition to the marginal mean parameters, which are traditionally modeled using risk factors, regressors are allowed on the dispersion and shape parameters. The proposed approach allows us to model the skewness and kurtosis of the model explicitly as a function of the explanatory variables for the mean, dispersion and shape parameters. Instead, if only the mean parameter is modeled in terms of explanatory variables then this can lead to a misclassification of policyholders with a high number of claims due to the unobserved heterogeneity changes with covariates. Furthermore, the MVPGIG, is a broad family of models including many MVMP models considered in the aforementioned literature ones as special and/or limiting cases, such as, for example, the multivariate Negative Binomial (MVNB), or multivariate Poisson-Gamma, multivariate Poisson-Inverse Gaussian (MVPIG), multivariate Poisson-Inverse Exponential, multivariate Poisson-Inverse Chi Squared, and multivariate Poisson-Scaled Inverse Chi Squared distributions, depending on the estimated values of the dispersion and shape parameters which are modeled based on covariate information, hence enabling us to account for the tail behavior of observed data in versatile manner. The latter can be regarded as an important property for capturing overdispersion since this phenomenon is not necessarily attributed to an excess of zeros but it may be also caused by a heavy tail in the claim count data, see Shared (1980). For illustrative purposes, the bivariate Poisson-Generalized Inverse Gaussian (BPGIG) regression model with varying dispersion and shape is fitted on Motor Third Party Liability (MTPL) insurance bodily injury and property damage claim count data using a novel Expectation-Maximization (EM) type algorithm. The proposed maximum likelihood (ML) estimation scheme takes advantage of the stochastic mixture representation of the BPGIG model to reduce the problem of maximizing its cumbersome likelihood function which is expressed in terms of the modified Bessel function of the third kind to the simple problem of maximizing the likelihood function of its mixing density.

The remainder of this article is organized, as follows: Section 2 deals with the construction of the proposed MVPGIG regression model with varying dispersion and shape parameters. In Section 3, we describe the ML estimation procedure for the BPGIG model via the EM algorithm. A real data application based on the two-dimensional MTPL data set is presented in Section 4 and the fitting performance of the BPGIG regression model with varying dispersion and shape parameters is compared to that of the bivariate Negative Binomial (BNB) and Poisson-Inverse Gaussian (BPIG) regression models with varying dispersion of Tzougas and Pignatelli di Cerchiara (2021) that we use as a benchmark for comparison. In Section 5, the a posteriori, or Bonus-Malus, premiums determined by the from the BNB and BPIG models are compared with those resulting from the proposed BPGIG model using the expected value principle. Finally, concluding remarks can be found in Section 6.

2 | THE MULTIVARIATE POISSON-GENERALIZED INVERSE GAUSSIAN REGRESSION MODEL WITH VARYING DISPERSION AND SHAPE PARAMETERS

Consider that in a nonlife insurance policy of an insured j, where $j = 1, \ldots, n$, we observe multi-peril claim frequencies $K_{i,j}$, for $i = 1, \ldots, m$ types of coverage. Assume that given the random variables $Z_j > 0$, $K_{i,j}|Z_j$ per claim type i are distributed according to a Poisson distribution with probability mass function (pmf) given by

$$P(k_{i,j}|z_j) = \frac{\exp[-\mu_{i,j} z_j](\mu_{i,j} z_j)^{k_{i,j}}}{k_{i,j}!}, \tag{1}$$

for $k_{i,j} = 0, 1, 2, 3, \ldots$, where $\mu_{i,j} > 0$, with mean and variance $\mathbb{E}(k_{i,j}|z_j) = \mu_{i,j} z_j$ and $\mathrm{Var}(k_{i,j}|z_j) = \mu_{i,j} z_j$.

Also, suppose that Z_j are random variables from a Generalized-Inverse Gaussian (GIG) distribution with probability density function (pdf) given by

$$g(z_j;\sigma_j,\nu_j)=\frac{c_j^{\nu_j}}{2K_{\nu_j}\left(\frac{1}{\sigma_j}\right)}z_j^{\nu_j-1}\exp\left[-\frac{1}{2\sigma_j}\left(c_jz_j+\frac{1}{c_jz_j}\right)\right], \tag{2}$$

for $\sigma_j>0$ and $-\infty<\nu_j<\infty$, where $c_j=\frac{K_{\nu_j+1}\left(\sigma_j^{-1}\right)}{K_{\nu_j}\left(\sigma_j^{-1}\right)}$ and

$$K_{\nu_j}(\omega)=\int_0^\infty x^{\nu_j-1}\exp\left\{-\frac{1}{2}\omega\left(x+\frac{1}{x}\right)\right\}dx$$

is the modified Bessel function of the third kind of order ν_j and argument ω. This parameterization ensures that the model is identifiable since $\mathbb{E}(Z_j)=1$. Furthermore, note that

$$Var(Z_j)=\frac{K_{\nu_j+2}\left(\frac{1}{\sigma_j}\right)K_{\nu_j}\left(\frac{1}{\sigma_j}\right)}{K_{\nu_j+1}\left(\frac{1}{\sigma_j}\right)^2}-1.$$

Thus, considering the assumptions in Equations (1) and (2) it is easy to see that the unconditional distribution of $K_{i,j}$ will be a multivariate Poisson-Generalized Inverse Gaussian (MVPGIG) distribution with joint probability mass function (JPMF) given by

$$P(k_{1,j},k_{2,j},\ldots,k_{m,j})$$
$$=\frac{\prod_{i=1}^m\mu_{i,j}^{k_{i,j}}}{\prod_{i=1}^m k_{i,j}!}\frac{c_j^{\nu_j}}{2K_{\nu_j}\left(\frac{1}{\sigma_j}\right)}\left[\left(2\sum_{i=1}^m\mu_{i,j}+\frac{c_j}{\sigma_j}\right)c\sigma_j\right]^{-\frac{\sum_{i=1}^m k_{i,j}+\nu_j}{2}}2K_{\sum\limits_{i=1}^m k_{i,j}+\nu_j}\left[\sqrt{\frac{1}{c\sigma_j}\left(2\sum_{i=1}^m\mu_{i,j}+\frac{c}{\sigma_j}\right)}\right]. \tag{3}$$

Note that if we let $\nu_j=-0.5$ in Equation (3) the MVPGIG distribution reduces to a multivariate Poisson-Inverse Gaussian (MVPIG) distribution. Further, the multivariate Negative Binomial (MVNBB) distribution is a limiting case of Equation (3), obtained by letting $\sigma_j\to\infty$ for $\nu_j>0$ and $\nu_j<-1$ respectively.

Henceforth, for expository purposes, we will restrict attention to the bivariate case m = 2. We assume that the mean, dispersion and shape parameters of the bivariate Poisson-Generalized Inverse Gaussian (BPGIG) are modeled as functions of explanatory variables with parametric linear functional forms:

$$\mu_{1,j}=\exp\left(\boldsymbol{x}_{1,j}^T\boldsymbol{\beta}_1\right), \tag{4}$$

$$\mu_{2,j}=\exp\left(\boldsymbol{x}_{2,j}^T\boldsymbol{\beta}_2\right), \tag{5}$$

$$\sigma_j = \exp\left(\boldsymbol{x}_{3,j}^T \boldsymbol{\beta}_3\right) \tag{6}$$

and

$$\nu_j = \boldsymbol{x}_{4,j}^T \boldsymbol{\beta}_4, \tag{7}$$

where $\boldsymbol{x}_{1,j}, \boldsymbol{x}_{2,j}, \boldsymbol{x}_{3,j}$, and $\boldsymbol{x}_{4,j}$ are vectors of covariates with dimensions $p_1 \times 1$, $p_2 \times 1, p_3 \times 1$ and $p_4 \times 1$ respectively, with $(\beta_{1,1}, ..., \beta_{1,p_1})^T, (\beta_{2,1}, ..., \beta_{2,p_2})^T, (\beta_{3,1}, ..., \beta_{3,p_3})^T$ and $(\beta_{4,1}, ..., \beta_{4,p_4})^T$ the corresponding parameter vectors and where it is considered that the matrices $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$, and $\boldsymbol{X}_4$ with rows given by $x_{1,i}$, $x_{2,i}$, $x_{3,i}$, and $x_{4,i}$ respectively, are of full rank.

Finally, the following desirable properties ensure the flexibility of the proposed model for capturing overdispersion and accommodating for positive correlation structures[1] between the different claim count response variables.

1. The marginal distribution of $K_{i,j}$, for $i = 1, ..., m$ and $j = 1, ..., n$, is a Poisson-Generalized Inverse Gaussian, or Sichel, distribution. Also, the mean and the variance of $K_{i,j}$ are given by

$$\mathbb{E}(K_{i,j}) = \mu_{i,j} \tag{8}$$

and

$$\mathrm{Var}(K_{i,j}) = \mu_{i,j} + \mu_{i,j}^2\left(c_j^{-2} + \frac{2(\nu_j + 1)}{c_j}\sigma_j - 1\right). \tag{9}$$

2. Let $K_{i,j}|z_j$ and z_j, for $i = 1, ..., m$ and $j = 1, ..., n$, be distributed according to the Poisson and GIG distributions which are given Equations (1) and (2), respectively. Also, consider that the cumulative generating function of z_j is denoted by $C_{z_j}(t)$, then the cumulative generating function of the marginal distribution of $K_{i,j}$, which is denoted by $C_{K_{i,j}}(t)$, is given by

$$C_{K_{i,j}}(t) = C_{z_j}[\mu_{i,j}(e^t - 1)] \tag{10}$$

and hence, since z_j has a unit mean, the third and fourth cumulants of $K_{i,j}$ and z_j are related by

$$C_{3K_{i,j}} = \mu_{i,j} + 3\mu_{i,j}^2 Var(z_j) + \mu_{i,j}^3 C_{3z_j}, \tag{11}$$

and

$$C_{4K_{i,j}} = \mu_{i,j} + 7\mu_{i,j}^2 Var(z_j) + 6\mu_{i,j}^3 C_{3z_j} + \mu_{i,j}^4 C_{4z_j}, \tag{12}$$

where $C_{3K_{i,j}}$ and $C_{4K_{i,j}}$ are the third and fourth cumulants of $K_{i,j}$.

[1]A limitation of the proposed model is that it cannot allow for negative correlation between the claim count response variables. However, regarding MTPL data, such as those we use in this study, positive correlation between MTPL bodily injury and property damage claim counts is what we expect.

The skewness and kurtosis of $K_{i,j}$ are $\sqrt{\beta_1} = \kappa_{3K_{i,j}}/[Var(K_{i,j})]^{1.5}$ and $\beta_2 = 3 + \{\kappa_{4K_{i,j}}/[Var(K_{i,j})^2]\}$ respectively, where the cumulants of the mixing distribution are given by

$$C_{3z_j} = [g_2 - 3g_1] \tag{13}$$

and

$$C_{4\gamma} = \left(g_3 - 4g_2 + 6g_1 - 3g_1^2\right), \tag{14}$$

where $g_1 = \left[1/c_j^2 + 2\sigma_j(\nu_j + 1)/c_j - 1\right]$,
$g_2 = 2\sigma_j(\nu_j + 2)/c_j^3 + \left[4\sigma_j^2(\nu_j + 1)(\nu_j + 2) + 1\right]/c_j^2 - 1$
$g_3 = \left[1 + 4\sigma_j^2(\nu_j + 2)(\nu_j + 3)\right]/c_j^4 + \left[8\sigma_j^3(\nu_j + 1)(\nu_j + 2)(\nu_j + 3) + 4\sigma(\nu_j + 2)\right]/c_j^3 - 1.$

3. The covariance (Cov) between $K_{1,j}$ and $K_{2,j}$ is given by

$$\mathrm{Cov}(K_{1,j}, K_{2,j}) = \mu_{1,j}\mu_{2,j}\left(c_j^{-2} + \frac{2(\nu_j + 1)}{c_j}\sigma_j - 1\right). \tag{15}$$

3 | STATISTICAL INFERENCE: THE EM ALGORITHM

In this section, an EM type algorithm (Dempster et al., 1977; McLachlan & Krishnan, 2007) is developed to facilitate maximum likelihood (ML) estimation of the BPGIG regression model with varying dispersion and shape.

Furthermore, assume that $(k_{1,j}, k_{2,j}, \boldsymbol{x_{1,j}}, \boldsymbol{x_{2,j}}, \boldsymbol{x_{3,j}}, \boldsymbol{x_{4,j}}), j = 1, ..., n$, is a sample of independent observations, where $K_{1,j}$ and $K_{2,j}$ are the claim count variables and $\boldsymbol{x_{1,j}}, \boldsymbol{x_{2,j}}, \boldsymbol{x_{3,j}}$ and $\boldsymbol{x_{4,j}}$ are the vectors of covariates with dimensions $p_1 \times 1, p_2 \times 1, p_3 \times 1$ and $p_4 \times 1$ respectively. Also, assume that $Z_j, j = 1, ..., n$, are the random effects which are nonobservable and are considered to produce missing data. By augmentation of the unobserved Z_j one can write the complete log-likelihood as follows:

$$\begin{aligned}\ell_c(\theta) \propto & \textstyle\sum_{i=1}^{2}\sum_{j=1}^{n}[-\mu_{i,j}z_j + k_{i,j}\log(\mu_{i,j})] \\ & + \textstyle\sum_{j=1}^{n}\left[\nu_j\log(c_j) + (\nu_j - 1)\log(z_j) - \log\left(K_{\nu_j}\left(\sigma_j^{-1}\right)\right) - \frac{1}{2\sigma_j}\left(c_j z_j + \frac{1}{c_j z_j}\right)\right],\end{aligned} \tag{16}$$

where $\theta = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4)$ is the vector of the parameters.

We present below the E- and the M-Steps of our EM type algorithm. At the E-Step, we compute the Q-function, which is the conditional expectation of the complete log-likelihood function given by Equation (16), given θ^r, which is the estimated value of θ at the rth iteration. The M-Step consists in maximizing the Q-function. In particular, we want to find the updated parameters θ^{r+1} such that the Q-function is increased with respect to θ.

- **E-Step**: The Q-function at the r-th iteration can be written as

$$
\begin{aligned}
Q(\theta;\, \theta^{(r)}) \equiv\ & \mathbb{E}_z(\ell_c(\theta)|k_{1,j}, k_{2,j};\, \theta^{(r)}) \\
& \propto \textstyle\sum_{i=1}^{2} \sum_{j=1}^{n} \left[-\mu_{i,j}^{(r)} w_{1,j} + k_{i,j} \log\left(\mu_{i,j}^{(r)}\right)\right] \\
& \quad + \textstyle\sum_{j=1}^{n} \left[\nu_j^{(r)} \log\left(c_j^{(r)}\right) + \left(\nu_j^{(r)} - 1\right) w_{3,j}\right. \\
& \quad \left. - \log\left(K_{\nu_j^{(r)}}\left(1/\sigma_j^{(r)}\right)\right) - \frac{1}{2\sigma_j^{(r)}} \left(c_j^{(r)} w_{1,j} + \frac{w_{2,j}}{c_j^{(r)}}\right)\right],
\end{aligned}
\tag{17}
$$

where we have defined the pseudo-values $w_{1,j} = \mathbb{E}_{z_j}[z_j|k_{i,j};\, \theta^{(r)}]$, $w_{2,j} = \mathbb{E}_{z_j}\left[z_j^{-1}|k_{i,j};\, \theta^{(r)}\right]$ and $w_{3,j} = \mathbb{E}_{z_j}[\log(z_j)|k_{i,j};\, \theta^{r}]$.

- **M-Step**:
  - First, differentiate the $Q-$function with respect to $\boldsymbol{\beta}_1$:

$$
h_1(\boldsymbol{\beta}_1) = \frac{\partial Q(\theta;\, \theta^{(r)})}{\partial \beta_{1,l}} = \sum_{j=1}^{n} \left(k_{1,j} - \mu_{1,j}^{(r)} w_{1,j}\right) x_{1,j,l},
\tag{18}
$$

and

$$
H_1(\boldsymbol{\beta}_1) = \frac{\partial^2 Q(\theta;\, \theta^{(r)})}{\partial \beta_{1,l} \partial \beta_{1,l}^T} = \sum_{j=1}^{n} \left(-\mu_{1,j}^{(r)} w_{1,j}\right) x_{1,j,l} x_{1,j,l}^T,
\tag{19}
$$

for $j = 1, \ldots, n$ and $l = 1, \ldots, p_1$. Then, the Newton–Raphson iterative algorithm for $\boldsymbol{\beta}_1$ is as follows:

$$
\boldsymbol{\beta}_1^{(r+1)} \equiv \boldsymbol{\beta}_1^{(r)} - \left[H_1\left(\boldsymbol{\beta}_1^{(r)}\right)\right]^{-1} h_1\left(\boldsymbol{\beta}_1^{(r)}\right).
\tag{20}
$$

  - Second, differentiate the $Q-$function with respect to $\boldsymbol{\beta}_2$:

$$
h_2(\boldsymbol{\beta}_2) = \frac{\partial Q(\theta;\, \theta^{(r)})}{\partial \beta_{2,l}} = \sum_{j=1}^{n} \left(k_{2,j} - \mu_{2,j}^{(r)} w_{1,j}\right) x_{2,j,l},
\tag{21}
$$

and

$$
H_2(\boldsymbol{\beta}_2) = \frac{\partial^2 Q(\theta;\, \theta^{(r)})}{\partial \beta_{2,l} \partial \beta_{2,l}^T} = \sum_{j=1}^{n} \left(-\mu_{2,j}^{(r)} w_{1,j}\right) x_{2,j,l} x_{2,j,l}^T,
\tag{22}
$$

for $j = 1, \ldots, n$ and $l = 1, \ldots, p_2$. Then, the Newton–Raphson iterative algorithm for $\boldsymbol{\beta}_2$ is as follows:

$$
\boldsymbol{\beta}_2^{(r+1)} \equiv \boldsymbol{\beta}_2^{(r)} - \left[H_2\left(\boldsymbol{\beta}_2^{(r)}\right)\right]^{-1} h_2\left(\boldsymbol{\beta}_2^{(r)}\right).
\tag{23}
$$

– Third, differentiate the Q–function with respect to $\boldsymbol{\beta}_3$:

$$
\begin{aligned}
h_3(\boldsymbol{\beta}_3) = \frac{\partial Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(r)})}{\partial \beta_{3,l}} = \textstyle\sum_{i=1}^{n} \sigma_j^{(r)} \Bigg[& \frac{\nu_j^{(r)}}{\sigma_j^{(r)}} - \frac{c_j^{(r)}}{\left(\sigma_j^2\right)^{(r)}} \\
& + \frac{1}{2\left(\sigma_j^2\right)^{(r)}}\left(c_j^{(r)} w_{1,j} + \frac{w_{2,j}}{c_j^{(r)}}\right) + \frac{\nu_j^{(r)}}{c_j^{(r)}}\frac{\partial c_j^{(r)}}{\partial \sigma_j^{(r)}} \\
& - \frac{1}{2\sigma_j^{(r)}}\frac{\partial c_j^{(r)}}{\partial \sigma_j^{(r)}}\left(w_{1,j} - \frac{w_{2,j}}{\left(c_j^2\right)^{(r)}}\right)\Bigg] x_{3,j,l},
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
H_3(\boldsymbol{\beta}_3) &= \frac{\partial^2 Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(r)})}{\partial \beta_{3,l}\partial \beta_{3,j}^T} \\
&= \textstyle\sum_{i=1}^{n} \sigma_j^{(r)} \Bigg\{ \frac{\nu_j^{(r)}}{\sigma_j^{(r)}} - \frac{c_j^{(r)}}{\left(\sigma_j^2\right)^{(r)}} + \frac{1}{2\left(\sigma_j^2\right)^{(r)}}\left(c_j^{(r)} w_{1,j} + \frac{w_{2,j}}{c_j^{(r)}}\right) \\
&+ \frac{\nu_j^{(r)}}{c_j^{(r)}}\frac{\partial c_j^{(r)}}{\partial \sigma_j^{(r)}} - \frac{1}{2\sigma_j^{(r)}}\frac{\partial c_j^{(r)}}{\partial \sigma_j^{(r)}}\left(w_{1,j} - \frac{w_{2,j}}{\left(c_j^2\right)^{(r)}}\right) \\
&+ \sigma_j^{(r)}\Bigg[\left(\frac{\nu_j^{(r)}}{c_j^{(r)}} - \frac{1}{2\sigma_j^{(r)}}\left(w_{1,j} - \frac{w_{2,j}}{\left(c_j^2\right)^{(r)}}\right)\right)\frac{\partial^2 c_j^{(r)}}{\partial\left(\sigma_j^2\right)^{(r)}} \\
&+ \frac{1}{\left(\sigma_j^2\right)^{(r)}}\left(w_{1,j} - \frac{w_{2,j}}{\left(c_j^2\right)^{(r)}} - 1\right)\frac{\partial c_j^{(r)}}{\partial \sigma_j^{(r)}} \\
&- \left(\frac{\nu_j^{(r)}}{\left(c_j^2\right)^{(r)}} + \frac{w_{2,j}}{\left(c_j^3\right)^{(r)}\sigma_j^{(r)}}\right)\left(\frac{\partial c_j^{(r)}}{\partial \sigma_j^{(r)}}\right)^2 - \frac{\nu_j^{(r)}}{\left(\sigma_j^2\right)^{(r)}} \\
&+ \frac{2c_j^{(r)}}{\left(\sigma_j^3\right)^{(r)}} - \frac{1}{\left(\sigma_j^3\right)^{(r)}}\left(c_j^{(r)} w_{1,j} + \frac{w_{2,j}}{c_j^{(r)}}\right)\Bigg]\Bigg\} x_{3,j,l} x_{3,j,l}^T,
\end{aligned}
\tag{25}
$$

for $j = 1, \ldots, n$ and $l = 1, \ldots, p_3$, where

$$
\frac{\partial c_j^{(r)}}{\partial \phi_j^{(r)}} = \frac{c_j^{(r)}\left(2\nu_j^{(r)} + 1\right)}{\phi_j^{(r)}} + \frac{1 - \left(c_j^2\right)^{(r)}}{\left(\phi_j^2\right)^{(r)}}
\tag{26}
$$

and where

$$\frac{\partial^2 c_j^{(r)}}{\partial\left(\phi_j^2\right)^{(r)}} = \left(\frac{2\nu_j^{(r)}+1}{\phi_j^{(r)}} - \frac{2c_j^{(r)}}{\left(\phi_j^2\right)^{(r)}}\right)\frac{\partial c_j^{(r)}}{\partial \phi_j^{(r)}} - \frac{c_j^{(r)}\left(2\nu_j^{(r)}+1\right)}{\left(\phi_j^2\right)^{(r)}} + \frac{2\left(\left(c_j^2\right)^{(r)}-1\right)}{\left(\phi_j^3\right)^{(r)}}. \tag{27}$$

Then, the Newton–Raphson iterative algorithm for $\boldsymbol{\beta}_3$ is as follows:

$$\boldsymbol{\beta}_3^{(r+1)} \equiv \boldsymbol{\beta}_3^{(r)} - \left[H_3\left(\boldsymbol{\beta}_3^{(r)}\right)\right]^{-1} h_3\left(\boldsymbol{\beta}_3^{(r)}\right). \tag{28}$$

– Finally, differentiate the Q–function with respect to $\boldsymbol{\beta}_4$:

$$\begin{aligned} h_4(\boldsymbol{\beta}_4) = \frac{\partial Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(r)})}{\partial \beta_{4,l}} = \frac{\partial}{\partial \beta_{4,j}}\Bigg\{\textstyle\sum_{i=1}^{n} \Bigg[\nu_j^{(r)}\log\left(c_j^{(r)}\right) \\ + \left(\nu_j^{(r)}-1\right)w_{3,j} - \log\left(K_{\nu_j^{(r)}}\left(\frac{1}{\sigma_j^{(r)}}\right)\right) \\ -\frac{1}{2\sigma_j^{(r)}}\left(c_j^{(r)}w_{1,j} + \frac{w_{2,j}}{c_j^{(r)}}\right)\Bigg]\Bigg\}x_{4,j,l}, \end{aligned} \tag{29}$$

and

$$\begin{aligned} H_4(\boldsymbol{\beta}_4) = \frac{\partial^2 Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(r)})}{\partial \beta_{4,l}\partial \beta_{4,j}^T} = \frac{\partial^2}{\partial \beta_{4,j}\partial \beta_{4,j}^T}\Bigg\{\textstyle\sum_{i=1}^{n} \Bigg[\nu_j^{(r)}\log\left(c_j^{(r)}\right) \\ +\left(\nu_j^{(r)}-1\right)w_{3,j} - \log\left(K_{\nu_j^{(r)}}\left(\frac{1}{\sigma_j^{(r)}}\right)\right) \\ -\frac{1}{2\sigma_j^{(r)}}\left(c_j^{(r)}w_{1,j} + \frac{w_{2,j}}{c_j^{(r)}}\right)\Bigg]\Bigg\}x_{4,j,l}x_{4,j,l}^T, \end{aligned} \tag{30}$$

for $j = 1, ..., n$ and $l = 1, ..., p_4$. Thus, the Newton–Raphson iterative algorithm for $\boldsymbol{\beta}_4$ is as follows:

$$\boldsymbol{\beta}_4^{(r+1)} \equiv \boldsymbol{\beta}_4^{(r)} - \left[H_4\left(\boldsymbol{\beta}_4^{(r)}\right)\right]^{-1} h_4\left(\boldsymbol{\beta}_4^{(r)}\right). \tag{31}$$

4 | NUMERICAL ILLUSTRATION

We conducted an empirical analysis using a sample of claim frequency data which was randomly selected from a larger pool of MTPL insurance policies observed during the year 2017 from a major European insurance company. We are interested in modeling the MTPL bodily injury and property damage claims with their associated claim counts denoted by $K_{1,j}$ and $K_{2,j}$ respectively, for $j = 1, \ldots, n$. For each policy, the total number of claims for each type of claim were reported within this yearly period. The sample comprised insured parties with complete records; that is, with the availability of all a priori rating variables which affect both $K_{1,j}$ and $K_{2,j}$. Furthermore, an exploratory analysis was carried out to accurately select the subset of explanatory variables with the highest predictive power for both $K_{1,j}$ and $K_{2,j}$. There were $n = 5186$ observations and three explanatory variables that met our criteria. Table 1 summarizes the explanatory variables whilst Table 2 depicts some standard descriptive statistics for $K_{1,j}$ and $K_{2,j}$, along with the values of Kendall's τ and Spearman's ρ correlation coefficients. As it was expected, Table 2 shows the existence of positive correlation between $k_{1,j}$ and $k_{2,j}$ as well as their marginal overdispersion. Furthermore, we would like to call attention to the fact that, as is well known, the range of Kendall's τ and Spearman's ρ for discrete random variables is narrower than $[-1, 1]$, see Denuit and Lambert (2005); Mesfioui and Tajar (2005) and Mesfioui et al. (2022). Furthermore, Nikoloulopoulos and Karlis (2010) and Safari-Katesari et al. (2020) showed how to compute the population versions of Kendall's τ and Spearman's

TABLE 1 The explanatory variables and their description

	Categories		
Variables	**C1**	**C2**	**C3**
City population (v1)	≤1,000,000	1,000,001–2,000,000	≥2,000,001
Number of years that the policyholder has been registered with the insurance company (v2)	<5 years	>5 years	–
Horsepower of the vehicle (v3)	0–1400 cc	1400–1800 cc	≥1800cc

TABLE 2 Descriptive statistics for the two responses

K_1		K_2	
Statistic	**Value**	**Statistic**	**Value**
Minimum	0	Minimum	0
Median	0	Median	0
Mean	0.0954	Mean	0.0618
Variance	0.1375	Variance	0.0644
Maximum	4	Maximum	3
Kendall's τ: 0.1760			
Spearman's ρ: 0.1777			

TABLE 3 Parameter estimates of the BNB and BPIG regression models with varying dispersion and the BPGIG regression model with varying dispersion and shape

	BNB			BPIG			BPGIG			
Variable	**Coeff.** β_1	**Coeff.** β_2	**Coeff.** β_3	**Coeff.** β_1	**Coeff.** β_2	**Coeff.** β_3	**Coeff.** β_1	**Coeff.** β_2	**Coeff.** β_3	$\nu = -0.53$
Intercept	−2.3933	−2.9262	−1.1296	−2.3950	−2.9279	−0.5908	−2.3964	−2.9293	−0.5893	–
v1 C2	0.0524	0.1504	−0.1912	0.0535	0.1518	−0.1157	0.0543	0.1529	−0.1149	–
v1 C3	0.1556	0.1770	−0.2303	0.1587	0.1793	−0.1364	0.1615	0.1813	−0.1352	–
v2 C2	0.0452	0.1780	0.3627	0.0465	0.1790	0.1959	0.0475	0.1797	0.1968	–
v3 C2	−0.1216	−0.0203	−0.3144	−0.1203	−0.0190	−0.1769	−0.1187	−0.0174	−0.1756	–
v3 C3	0.1767	0.1712	−0.0883	0.1784	0.1731	−0.0716	0.1798	0.1747	−0.0711	–

Abbreviations: BNB, bivariate Negative Binomial; BPIG, bivariate Poisson-Inverse Gaussian; BPGIG, bivariate Poisson-Generalized Inverse Gaussian.

TABLE 4 BNB, BPIG and BPGIG models comparison based on global deviance, AIC and SBC

Model	*df*	Global deviance	AIC	SBC
BNB	18	4388	4424	4542
BPIG	18	4249	4285	4403
BPGIG	19	4223	4261	4386

Abbreviations: BNB, bivariate Negative Binomial; BPIG, bivariate Poisson-Inverse Gaussian; BPGIG, bivariate Poisson-Generalized Inverse Gaussian.

ρ by pairing copulas with discrete marginal distributions, respectively. Following their setup, we investigated the variability of the population versions of Kendall's τ and Spearman's ρ from lowest to highest attainable values for our data by pairing two marginal Poisson distributions with varying mean parameter μ from 1 up to 20 using the Normal copula. Also, we considered that the copula parameter θ can vary from -1 to 1. We observed that the values of Kendall's τ and Spearman's ρ stabilize close to 1 for the values of μ which are greater than 10. Therefore, the bivariate Negative Binomial (BNB) and Poisson-Inverse Gaussian (BPIG) regression models with varying dispersion and the bivariate Poisson-Generalized Inverse Gaussian (BPGIG) regression model with varying dispersion and shape which allow for positive correlation between the two types of claims are better assumptions than the bivariate Poisson model, as the latter is not equipped for handling overdispersion. Moreover, Table 3 presents the estimated regression coefficients for the BNB and BPIG regression models with varying dispersion and the BPGIG regression model with varying dispersion and shape.[2]

Furthermore, we compare the fit of the BNB and BPIG regression models with varying dispersion to that of the BPGIG regression model with varying dispersion and shape based on the standard specification tests DEV, AIC, and SBC. The DEV is given by

$$\mathrm{DEV} = -2\hat{l}\,(\hat{\theta}), \tag{32}$$

with $\hat{l}$ being the maximum of the log-likelihood and $\hat{\theta}$ the vector of estimated parameters of the model. Moreover, the AIC is defined as

$$\mathrm{AIC} = \mathrm{DEV} + 2 \times df \tag{33}$$

and the SBC is given by

$$\mathrm{SBC} = \mathrm{DEV} + \log(n) \times df, \tag{34}$$

where df are the degrees of freedom which correspond to the number of fitted parameters in the model and n is the number of observations in the sample. The values of the DEV, AIC, and SBC for the competing models are provided in Table 4. As is well known, two models can be considered to be significantly different if the difference in the log-likelihoods exceeds five, corresponding to a difference in their respective AIC and SBC values of greater than ten and

[2]All the parameters were statistically significant at a 5% threshold.

TABLE 5 Comparison of the A Posteriori, or Bonus-Malus, Premium Rates for $t = 1$

	$t=1$				$t=1$				$t=1$			
	$t = 1$ BNB regression model				**$t = 1$ BPIG regression model**				**$t = 1$ BPGIG regression model**			
	Best profile				**Best profile**				**Best profile**			
$k_{1,j}/k_{2,j}$	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
0	76.36	312.65	548.94	785.23	82.05	239.47	500.38	800.56	82.05	254.83	556.90	917.13
1	312.65	548.94	785.23	1021.52	239.47	500.38	800.56	1110.36	1 254.82	556.90	917.13	1297.67
2	548.94	785.23	1021.52	1257.81	500.38	800.56	1110.36	1422.85	556.91	917.13	1297.67	1688.07
3	785.23	1021.52	1257.81	1494.10	800.56	1110.36	1422.85	1736.37	917.13	1297.67	1688.07	2084.96
	$t = 1$				$t = 1$				$t=1$			
	BNB regression model				**BPIG regression model**				**BPGIG regression model**			
	Average profile				**Average profile**				**Average profile**			
$k_{1,j}/k_{2,j}$	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
0	73.18	334.41	595.64	856.88	80.00	247.48	528.29	849.50	80.00	264.20	590.75	978.13
1	334.41	595.64	856.88	1118.11	247.48	528.29	849.50	1179.87	264.20	590.75	978.13	1386.04
2	595.64	856.88	1118.11	1379.35	528.29	849.50	1179.87	1512.72	590.75	978.13	1386.04	1804.03
3	856.88	1118.11	1379.35	1640.58	849.50	1179.87	1512.72	1846.48	978.13	1386.04	1804.03	2228.75
	$t = 1$				$t = 1$				$t=1$			
	BNB regression mode				**BPIG regression model**				**BPLN regression model**			
	Worst profile				**Worst profile**				**Worst profile**			
$k_{1,j}/k_{2,j}$	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
0	71.49	283.19	494.88	706.58	79.05	210.99	425.43	674.37	79.05	223.39	470.01	766.39
1	283.19	494.88	706.58	918.27	210.99	425.43	674.37	932.83	223.39	470.01	766.39	1081.27
2	494.88	706.58	918.27	1129.97	425.43	674.37	932.83	1194.13	470.01	766.39	1081.27	1405.04
3	706.58	918.27	1129.97	1341.66	674.37	932.83	1194.13	1456.54	1734.49	766.39	1081.28	1405.04

five respectively. Thus, this case we see that the best fitting performances are provided by the BPGIG regression model with varying dispersion and shape.[3]

Finally, we compare the forecasting performance of the proposed model and the benchmark models using both in-sample estimation and out-of sample validation. For this purpose, we split the data into training and test data at the ratio of 9 : 1. Therefore, the training data for the re-estimation of the parameters of the models contains 4149 data points. The remaining 1037 data points are used for testing purposes. To measure the prediction performances the deviance statistic is used. The deviance value for the BNB, BPIG, and BPGIG models are 490.80, 475.25, and 472.35 respectively. Thus, the BPGIG model outperforms the two competing bivariate mixed Poisson models.

5 | CALCULATION OF THE A POSTERIORI PREMIUMS

In this subsection, the expected value premium principle is used to compute the a posteriori, or Bonus-Malus, premium rates determined by the BNB, BPIG, and BPGIG models for $t = 1$ for three risk class profiles that we classify as Best, Average, and Worst according to the values of the mean claim frequencies $\mu_{1,j}$ and $\mu_{2,j}$, which are calculated using the same set of explanatory variables per claim type $i = 1, 2$. The results are depicted in Table 5.

6 | CONCLUSIONS

In this article, we presented the MVPGIG claims count regression model with varying dispersion and shape for modeling different types of claims in nonlife insurance. The MVPGIG is a wide family of models which, under the proposed modeling framework, can provide sufficient flexibility for capturing overdispersion and positive correlation structures in highly-dimensional claim count data. For demonstration purposes, the bivariate version of the model, namely the BPGIG model, with regression specifications for the mean, dispersion, and shape parameters was fitted on MTPL property damage and bodily injury claim count data. The ML estimates of the parameters of the model were obtained via a novel EM type algorithm. However, it should be noted that a shortcoming of the proposed approach is that there is a strong discrepancy between the flexibility within the equations of the random effect distribution, and the rigidity between these equations. To relax this rigidity, the BPGIG model can be constructed either by using the so-called trivariate reduction method or by considering correlated GIG random effects (say $z_{1;j}$ and $z_{2;j}$) paired via a Gaussian copula following the approaches of Bermudez and Karlis (2017) and Pechon et al. (2018) in the former and latter case respectively. Both approaches are very efficient when modeling different types of claims from different types of coverage or household claim frequencies in MTPL insurance. Finally, in a forthcoming paper, time series components will be included to accommodate for both cross dependence between different types of claims and time dependence, proceeding along similar lines as in Bermudez et al. (2018) among others.

[3]Note that the stopping criterion for the EM algorithm was rather strict as the algorithm iterated between the E and the M-steps until the relative change in the log-likelihood between two successive iterations was smaller than 10^{-12}.

ACKNOWLEDGMENTS

The authors would like to thank the two anonymous referees for their very helpful comments and suggestions which have significantly improved this study paper.

ORCID

Despoina Makariou http://orcid.org/0000-0002-9001-2122

REFERENCES

Abdallah, A., Boucher, J.-P., & Cossette, H. (2016). Sarmanov family of multivariate distributions for bivariate dynamic claim counts model. *Insurance: Mathematics and Economics*, *68*, 120–133.

Aguero-Valverde, J., & Jovanis, P. P. (2009). Bayesian multivariate Poisson lognormal models for crash severity modeling and site ranking. *Transportation Research Record*, *2136*(1), 82–91.

Aitchison, J., & Ho, C. (1989). The multivariate Poisson-log normal distribution. *Biometrika*, *76*(4), 643–653.

Bermudez, L., Guillén, M., & Karlis, D. (2018). Allowing for time and cross dependence assumptions between claim counts in ratemaking models. *Insurance: Mathematics and Economics*, *83*, 161–169.

Bermudez, L., & Karlis, D. (2011). Bayesian multivariate Poisson models for insurance ratemaking. *Insurance: Mathematics and Economics*, *48*(2), 226–236.

Bermudez, L., & Karlis, D. (2012). A finite mixture of bivariate Poisson regression models with an application to insurance ratemaking. *Computational Statistics & Data Analysis*, *56*(12), 3988–3999.

Bermudez, L., & Karlis, D. (2017). A posteriori ratemaking using bivariate Poisson models. *Scandinavian Actuarial Journal*, *2017*(2), 148–158.

Bolancé, C., Guillen, M., & Pitarque, A. (2020). A Sarmanov distribution with beta marginals: An application to motor insurance pricing. *Mathematics*, *8*(11), 2020. https://doi.org/10.3390/math8112020

Bolancé, C., & Vernic, R. (2019). Multivariate count data generalized linear models: three approaches based on the Sarmanov distribution. *Insurance: Mathematics and Economics*, *85*, 89–103.

Cameron, A. C., Li, T., Trivedi, P. K., & Zimmer, D. M. (2004). Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. *The Econometrics Journal*, *7*(2), 566–584.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53). Cambridge university press.

Chen, Y., & Hanson, T. (2017). Copula regression models for discrete and mixed bivariate responses. *Journal of Statistical Theory and Practice*, *11*(4), 515–530.

Chib, S., & Winkelmann, R. (2001). Markov chain Monte Carlo analysis of correlated count data. *Journal of Business & Economic Statistics*, *19*(4), 428–435.

Chiquet, J., Mariadassou, M., & Robin, S. (2020). The Poisson-lognormal model as a versatile framework for the joint analysis of species abundances. *bioRxiv*.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Denuit, M., Guillen, M., & Trufin, J. (2019). Multivariate credibility modelling for usage-based motor insurance pricing with behavioural data. *Annals of Actuarial Science*, *13*(2), 378–399.

Denuit, M., & Lambert, P. (2005). Constraints on concordance measures in bivariate discrete data. *Journal of Multivariate Analysis*, *93*(1), 40–57.

El-Basyouny, K., & Sayed, T. (2009). Accident prediction models with random corridor parameters. *Accident Analysis & Prevention*, *41*(5), 1118–1123.

Famoye, F. (2010). On the bivariate negative binomial regression model. *Journal of Applied Statistics*, *37*(6), 969–981.

Fung, T. C., Badescu, A. L., & Lin, X. S. (2019). A class of mixture of experts models for general insurance: Application to correlated claim frequencies. *ASTIN Bulletin: The Journal of the IAA*, *49*(3), 647–688.

Genest, C., & Nešlehová, J. (2007). A primer on copulas for count data. *ASTIN Bulletin: The Journal of the IAA*, *37*(2), 475–515.

Ghitany, M., Karlis, D., Al-Mutairi, D., & Al-Awadhi, F. (2012). An EM algorithm for multivariate mixed Poisson regression models and its application. *Applied Mathematical Sciences*, *6*(137), 6843–6856.

Gómez-Déniz, E., & Calderín-Ojeda, E. (2021). A priori ratemaking selection using multivariate regression models allowing different coverages in auto insurance. *Risks*, *9*(7), 137.

Gurmu, S., & Elder, J. (2000). Generalized bivariate count data regression models. *Economics Letters*, *68*(1), 31–36.

Ho, L. L., & Singer, J. d.M. (2001). Generalized least squares methods for bivariate Poisson regression. *Communications in Statistics—Theory and Methods*, *30*(2), 263–277.

Jeong, H., & Dey, D. K. (2021). *Multi-peril frequency credibility premium via shared random effects*. Available at SSRN 3825435.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1997). *Discrete multivariate distributions* (Vol. 165). Wiley.

Jung, R. C., & Winkelmann, R. (1993). Two aspects of labor mobility: A bivariate poisson regression approach. *Empirical economics*, *18*(3), 543–556.

Karlis, D., & Meligkotsidou, L. (2005). Multivariate Poisson regression with covariance structure. *Statistics and Computing*, *15*(4), 255–265.

Kocherlakota, S. (1988). On the compounded bivariate Poisson distribution: A unified treatment. *Annals of the Institute of Statistical Mathematics*, *40*(1), 61–76.

Kocherlakota, S., & Kocherlakota, K. (2001). Regression in the bivariate Poisson distribution. *Communications in Statistics-Theory and Methods*, *30*(5), 815–827.

Krummenauer, F. (1998). Representation of multivariate discrete distributions by probability generating functions. *Statistics & Probability Letters*, *39*(4), 327–331.

Lakshminarayana, J., Pandit, S. N., & SrinivasaRao, K. (1999). On a bivariate Poisson distribution. *Communications in Statistics-Theory and Methods*, *28*(2), 267–276.

Lee, A. (1999). Applications: Modelling rugby league data via bivariate negative binomial regression. *Australian & New Zealand Journal of Statistics*, *41*(2), 141–152.

Ma, J., Kockelman, K. M., & Damien, P. (2008). A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis & Prevention*, *40*(3), 964–975.

Marra, G., & Wyszynski, K. (2016). Semi-parametric copula sample selection models for count responses. *Computational Statistics & Data Analysis*, *104*, 110–129.

McLachlan, G. J., & Krishnan, T. (2007). *The EM algorithm and extensions* (Vol. 382). John Wiley & Sons.

Mesfioui, M., & Tajar, A. (2005). On the properties of some nonparametric concordance measures in the discrete case. *Nonparametric Statistics*, *17*(5), 541–554.

Mesfioui, M., Trufin, J., & Zuyderhoff, P. (2022). Bounds on spearman's rho when at least one random variable is discrete. *European Actuarial Journal*, *12*, 321–348. https://doi.org/10.1007/s13385-021-00289-8

M'Kendrick, A. (1925). Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, *44*, 98–130.

Munkin, M. K., & Trivedi, P. K. (1999). Simulated maximum likelihood estimation of multivariate mixed-Poisson regression models, with application. *The Econometrics Journal*, *2*(1), 29–48.

Nikoloulopoulos, A. K. (2013). Copula-based models for multivariate discrete response data. In P. Jaworski, F. Durante, & W. K. Härdle (Eds.), *Copulae in mathematical and quantitative finance* (pp. 231–249). Springer.

Nikoloulopoulos, A. K. (2016). Efficient estimation of high-dimensional multivariate normal copula models with discrete spatial responses. *Stochastic Environmental Research and Risk Assessment*, *30*(2), 493–505.

Nikoloulopoulos, A. K., & Karlis, D. (2010). Regression in a copula model for bivariate count data. *Journal of Applied Statistics*, *37*(9), 1555–1568.

Park, E. S., & Lord, D. (2007). Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record*, *2019*(1), 1–6.

Pechon, F., Denuit, M., & Trufin, J. (2019). Multivariate modelling of multiple guarantees in motor insurance of a household. *European Actuarial Journal*, *9*(2), 575–602.

Pechon, F., Denuit, M., & Trufin, J. (2021). Home and motor insurance joined at a household level using multivariate credibility. *Annals of Actuarial Science*, *15*(1), 82–114.

Pechon, F., Trufin, J., & Denuit, M. (2018). Multivariate modelling of household claim frequencies in motor third-party liability insurance. *Astin Bulletin*, *48*(3), 969–993.

Rüschendorf, L. (2013). Copulas, Sklar's theorem, and distributional transform. In *Mathematical risk analysis*. Springer Series in Operations Research and Financial Engineering Springer series in operations research and financial engineering (pp. 3–34). Springer. https://doi.org/10.1007/978-3-642-33590-7_1

Safari-Katesari, H., Samadi, S. Y., & Zaroudi, S. (2020). Modelling count data via copulas. *Statistics*, *54*(6), 1329–1355.

Shared, M. (1980). On mixtures from exponential families. *Journal of the Royal Statistical Society: Series B (Methodological)*, *42*(2), 192–198.

Shi, P., & Valdez, E. A. (2014a). Longitudinal modeling of insurance claim counts using jitters. *Scandinavian Actuarial Journal*, *2014*(2), 159–179.

Shi, P., & Valdez, E. A. (2014b). Multivariate negative binomial models for insurance claim counts. *Insurance: Mathematics and Economics*, *55*, 18–29.

Silva, A., Rothstein, S. J., McNicholas, P. D., & Subedi, S. (2019). A multivariate Poisson-log normal mixture model for clustering transcriptome sequencing data. *BMC Bioinformatics*, *20*(1), 1–11.

Stein, G. Z., & Juritz, J. M. (1987). Bivariate compound Poisson distributions. *Communications in Statistics-Theory and Methods*, *16*(12), 3591–3607.

Stein, G. Z., Zucchini, W., & Juritz, J. M. (1987). Parameter estimation for the Sichel distribution and its multivariate extension. *Journal of the American Statistical Association*, *82*(399), 938–944.

Tzougas, G., & di Cerchiara, A. P. (2021). Bivariate mixed Poisson regression models with varying dispersion. *North American Actuarial Journal*, 1–31. https://doi.org/10.1080/10920277.2021.1978850

Winkelmann, R. (2008). *Econometric analysis of count data*. Springer Science & Business Media.

Zhan, X., Aziz, H. A., & Ukkusuri, S. V. (2015). An efficient parallel sampling technique for multivariate Poisson-lognormal model: Analysis with two crash count datasets. *Analytic methods in accident research*, *8*, 45–60.

Zimmer, D. M., & Trivedi, P. K. (2006). Using trivariate copulas to model sample selection and treatment effects: Application to family health care demand. *Journal of Business & Economic Statistics*, *24*(1), 63–76.

How to cite this article: Tzougas, G., & Makariou, D. (2022). The multivariate Poisson-Generalized Inverse Gaussian claim count regression model with varying dispersion and shape parameters. *Risk Management & Insurance Review*, 1–17. https://doi.org/10.1111/rmir.12224