Contents

List of Tables

List of Figures

EM Estimation for the Bivariate Mixed Gamma Regression Model

Abstract

In this paper, we present a new family of bivariate mixed Gamma regression models for taking into account the positive correlation between the cost of claims from motor third-party liability bodily injury and property damage in a versatile manner. Furthermore, we demonstrate how maximum likelihood estimation of the model parameters can be achieved via a novel Expectation- Maximization algorithm. The implementation of two members of this family, namely the Bivariate Gamma-Inverse Gaussian regression and Bivariate Gamma-Weibull models is illustrated by a real data application which involves fitting motor insurance data from a European motor insurance company.

1 Introduction

Over the last few decades, there has been a vast increase in actuarial research works focusing on modelling costs of a particular claim type based on various claim severity modelling approaches such as heavy-tailed families of models, log phase-type distributions, mixture models, composite or splicing models and combinations of composite models with mixture models, see, for example, Frees and Valdez (2008), Ramirez-Cobo et al. (2010),Lee and Lin (2010), Pigeon and Denuit (2011), Scollnik and Sun (2012), Frees et al. (2014), Hürlimann (2014), Tzougas et al. (2014), Nadarajah and Bakar (2014), Bakar et al. (2015), Miljkovic and Grün (2016), Calderín-Ojeda and Kwok (2016), Reynkens et al. (2017), Tzougas et al. (2018), Tzougas et al. (2019), Laudagé et al. (2019), Grün and Miljkovic (2019), Jeong (2020), Parodi (2020), Wang et al. (2020), Tzougas and Karlis (2020) and Fung et al. (2021) among many more. However, even if the literature in the univariate setting contains large number of articles, the bivariate, and/or multivariate, extensions of such models have not been explored in depth even if in non-life insurance, the actuary may often be concerned with modelling jointly different types of claims and their associated costs. In this paper, motivated by a European Motor Third Party Liability (MTPL) insurance data set, we introduce a family of bivariate mixed Gamma regression models for joint modelling the costs from positively correlated bodily injury and property damage claims in terms of covariates.

The proposed class of bivariate claim severity regression models is based on a mixing between two marginal Gamma distributions and a unit mean continuous and at least twice differentiable mixing density. The regression structure is employed to those parameters which are involved in the mean and dispersion of the bivariate mixed Gamma regression models. The modelling framework we consider can account for the positive dependency between the two claim types in a flexible manner since it allows for a variety of alternative distributional assumptions for the mixing density. Furthermore, depending on the choice of the mixing density the bivariate mixed Gamma model can be used to model both moderate and large bodily injury and property damage claim sizes. Finally, we develop an Expectation-Maximization (EM) type algorithm1 which takes advantage of the stochastic mixture representation of the bivariate mixed Gamma regression model for maximizing its log-likelihood in a computationally efficient and parsimonious manner. For expository purposes,bivariate Gamma-Inverse Gaussian (BGIG) and bivariate Gamma Weibull (BGW) regression models are fitted on the MTPL bodily injury and property damage data set. The rest of the paper proceeds as follows. Section 2 discusses how the bivariate mixed Gamma regression model can be constructed and the joint probability density functions (jpdfs) of the and BGIG and BGW regression models, which are used for demonstration purposes, are derived. Section 3 deals with parameter estimation for the proposed model based on the EM algorithm. In Section 4, the models presented in Section 2 are fitted to the MTPL bodily injury and property damage claims data set. Finally, concluding remarks are provided in Section 5.

2 The Bivariate mixed Gamma regression model

Consider a non-life MTPL insurance which contains bodily injury and property damage claims and their are associated costs. Since it is possible that there exists a positive correlation between the two types of claims we propose the following family of models. The claim amounts of both types are denoted as $Y_i, i = 1, 2$, which are well-defined when there is at least one claim for each type of claim. Furthermore, we consider that conditional on a random effect $Z > 0$, the random variables $Y_i, i = 1, 2$ are independent Gamma random variables with rates $\alpha_i z$. The random effect Z is a continuous random variable with density $g_{\phi_j}(z_j)$ where, $j = 1, 2, \cdot, n$ which takes positive values only and it mainly controls the variation and correlation of the whole bivariate sequence. To avoid the identifiability problem, we have to restrict the expectation $\exp[z]$ to be a fixed constant and one usually lets $\mathbf{E}(Z) = 1$. On the other hand, to account for the impact of heterogeneity between different policyholders, the rates α_i ,$i = 1, 2$ are modelled as functions of explanatory variables $xi \in \mathbb{R}^{d_i \times 1}$ such that $\alpha_i = \exp\{x_i^\top \beta_i\}$, where $d_i \in N_+ and \beta_i \in \mathbb{R}^{d_i \times 1}$ are the corresponding coefficients. Then, the unconditional joint density function, $f_Y(y)$, of this bivariate sequence $Y = (Y_1, Y_2)$ is given by

$$f_Y(y) = \int_0^\infty (\prod_{i=1}^{2} f_{Y_i|z}(y, z)) g_{\phi_j}(z_j) dZ_j \tag{1}$$

$$= \int_0^\infty \frac{1}{\Gamma\lambda^2 \left(\alpha_1\alpha_2\right)^{\lambda} z_j^{2\lambda}} e^{-\frac{\frac{y_1}{\alpha_1}+\frac{y_2}{\alpha_2}}{z_j}} \left(y_1 y_2\right)^{\lambda - 1} g_{\phi_j}(z_j) dZ_j.$$

For $j = 1, 2, \cdots, n$. In the following,for demonstration purposes, we specialize with two different mixing densities, the Weibull and Inverse Gaussian (IG) distributions, which lead to the Bivariate Gamma-Inverse Gaussian (BGIG) and Bivariate Gamma Weibull (BGW) regression models respectively.

2.1 Bivariate Gamma-Inverse Gaussian regression model

The unit mean Inverse Gaussian (IG) density function is given by

$$g_{\phi_j}(z_j) = \frac{\phi_j}{\sqrt{2\pi}} e^{\phi_j^2} z_i^{-3/2} e^{\left\{\frac{1}{2}\left(z+\frac{1}{z}\right)\left(-\phi_j^2\right)\right\}} \tag{2}$$

The random effect Z now has a unit mean and variance $\frac{1}{\phi_j^2}$, To link available covariates with parameter ϕ_j we use log link function as $\phi_j = \exp(x_3^\top \beta_3)$. The unconditional joint density function of the bivariate Gamma-Inverse Gaussian (BGIG) can be derived as follows

$$f_Y(y) = \int_0^\infty (\prod_{i=1}^{2} f_{Y_i|z}(y, z)) g_{\phi_i}(z_i) dz_i$$

$$= \int_0^\infty \frac{1}{\Gamma\lambda^2 \left(\alpha_1\alpha_2\right)^{\lambda} z_i^{2\lambda}} e^{-\frac{\frac{y_1}{\alpha_1}+\frac{y_2}{\alpha_2}}{z_i}} \left(y_1 y_2\right)^{\lambda-1} \times \frac{\phi_j}{\sqrt{2\pi}} e^{\phi_j^2} z_j^{-3/2} e^{\left\{\frac{1}{2}\left(z_j+\frac{1}{z_j}\right)\left(-\phi_j^2\right)\right\}} dz_j$$

where $\alpha_i > 0$ and $\phi_j > 0$. Unfortunately, the last integral cannot be simplified but it can be computed via numerical integration. The mean, variance, covariance and correlation in the case of the BGIG model are given by

$$\mathbb{E}[Y_i] = \mathbb{E}_z[\mathbb{E}[Y_i|Z_j]] = \mathbb{E}[\alpha_i\lambda] = \alpha_i\lambda \tag{3}$$

$$\begin{aligned}\mathbb{V}(Y_i) &= \mathbb{E}[\mathbb{V}_y(Y_i|Z_j)] + \mathbb{V}(\mathbb{E}_z[Y_i|Z_j]) \\ &= \mathbb{E}[\lambda(\alpha_i Z_j)^2] + \mathbb{V}(\alpha_i Z_j \lambda) \qquad (4) \\ &= \alpha^2\lambda\left[\frac{\phi_j^2 + 1 + \lambda}{\phi_j^2}\right] \qquad (5)\end{aligned}$$

$$
\begin{aligned}
Cov(Y_1, Y_2) &= \mathbb{E}[Cov(Y_1, Y_2|Z_j) + Cov(\mathbb{E}[Y_1|Z_j], \mathbb{E}[Y_2|Z_j])] \\
&= 0 + Cov(\alpha_1 Z\lambda, \alpha_2 Z_j \lambda) = \frac{\alpha_1\alpha_2\lambda^2}{\phi_j^2}
\end{aligned} \tag{6}
$$

$$
Corr(Y_1, Y_2) = \frac{Cov(Y_1, Y_2)}{\sqrt{\mathbb{V}(Y_1)\mathbb{V}(Y_2)}} = \frac{\lambda}{\phi_j^2 + 1 + \lambda} \tag{7}
$$

2.2 Bivariate Gamma - Weibull regression model

The unit mean Weibull distribution having parameter ϕ_i with pdf given by

$$
g(z_j; \phi_j) = \phi_j. \left[\Gamma\left(1 + \frac{1}{\phi_j}\right)\right] e^{-\left(z_j\left[\Gamma(1+\frac{1}{\phi_j})\right]\right)^{\phi_j}} \left(z_j\left[\Gamma\left(1 + \frac{1}{\phi_j}\right)\right]\right)^{\phi_j - 1} \tag{8}
$$

$z_j > 0$ and $\mathbb{E}(Z_j) = 1$ and $\mathbb{V}ar(Z_j) = \left(\frac{1}{\left[\Gamma\left(1+\frac{1}{\phi_j}\right)\right]}\right)^2 \left[\Gamma\left(1 + \frac{2}{\phi_j}\right) - \Gamma^2\left(1 + \frac{1}{\phi_j}\right)\right]$, where $j = 1, 2, ..., n$.To link available covariates with parameter ϕ_j we use log link function as $\phi_j = \exp(x_3^\top \beta_3)$

The unconditional joint density function of the bivariate Gamma Weibull (BGW) can be derived as follows

$$
f_Y(y) = \int_0^\infty (\prod_{i=1}^{2} f_{Y_i|z}(y, z)) g_{(z_j);\phi_j} dz_j
$$

$$
= \int_0^\infty \frac{1}{\Gamma\lambda^2 (\alpha_1\alpha_2)^{\lambda} z_i^{2\lambda}} e^{-\frac{\frac{y_1}{\alpha_1} + \frac{y_2}{\alpha_2}}{z_i}} (y_1 y_2)^{\lambda - 1} \times \phi_j. \left[\Gamma\left(1 + \frac{1}{\phi_j}\right)\right] e^{-\left(z_j\left[\Gamma(1+\frac{1}{\phi_j})\right]\right)^{\phi_j}} \left(z_j\left[\Gamma\left(1 + \frac{1}{\phi_j}\right)\right]\right)^{\phi_j - 1} dz_j
$$

where $\alpha_i > 0$ and $\phi_j > 0$. Unfortunately ,the last integral cannot be simplifed but it can be computed via numerical integration. The mean, variance, covariance and correlation in the case of the BGW model are given by

$$\mathbb{E}[Y_i] = \mathbb{E}_z[\mathbb{E}[Y_i|Z_j]] = \mathbb{E}[\alpha_i\lambda] = \alpha_i\lambda \tag{9}$$

$$\begin{aligned}\mathbb{V}(Y_i) &= \mathbb{E}[\mathbb{V}_y(Y_i|Z_j)] + \mathbb{V}(\mathbb{E}_z[Y_i|Z_j]) \\ &= \mathbb{E}[\lambda(\alpha_i Z_j)^2] + \mathbb{V}(\alpha_i Z_j\lambda) \end{aligned}\tag{10}$$

$$= \frac{\alpha^2\lambda\left(-(\lambda+1)\Gamma\left(\frac{1}{\phi_j}\right)^2 + 2(\lambda+1)\phi_j\Gamma\left(\frac{2}{\phi_j}\right) + (\phi_j+1)^2\right)}{(\phi_j+1)^2} \tag{11}$$

$$\begin{aligned}Cov(Y_1, Y_2) &= \mathbb{E}[Cov(Y_1, Y_2|Z_j) + Cov(\mathbb{E}[Y_1|Z_j], \mathbb{E}[Y_2|Z_j])] \\ = 0 + Cov(\alpha_1 Z_j\lambda, \alpha_2 Z_j\lambda) &= \frac{\alpha_1\alpha_2\lambda^2\left(2\phi_j\Gamma\left(\frac{2}{\phi_j}\right) - \Gamma\left(\frac{1}{\phi_j}\right)^2\right)}{(\phi_j+1)^2}\end{aligned} \tag{12}$$

$$Corr(Y_1, Y_2) = \frac{Cov(Y_1, Y_2)}{\sqrt{\mathbb{V}(Y_1)\mathbb{V}(Y_2)}} = \frac{\lambda\left(2\phi_j\Gamma\left(\frac{2}{\phi_j}\right) - \Gamma\left(\frac{1}{\phi_j}\right)^2\right)}{\sqrt{\left(-(\lambda+1)\Gamma\left(\frac{1}{\phi_j}\right)^2 + 2(\lambda+1)\phi_j\Gamma\left(\frac{2}{\phi_j}\right) + (\phi_j+1)^2\right)^2}} \tag{13}$$

3 The EM algorithm for the bivariate mixed Gamma regression model

In this Section, an Expectation-Maximization (EM) algorithm is applied to facilitate the maximization likelihood estimation of the bivariate mixed Gamma regression model. Consider the observed bivariate response sequence $Yjj = 1, ..., n$ and the corresponding covariates $\{x_1, j\}_{j=1,...,n}$ and $\{x_2, j\}_{j=1,...,n}$. Also, let $\Theta = \{\beta_1, \beta_2, \beta_3\}$ be the parameter space for this model. Then, the log-likelihood function can be written as

$$l(\Theta) = \sum_{j=1}^{2} log\bigg(\int_0^\infty \frac{1}{\Gamma\lambda^2\, (\alpha_1\alpha_2)^{\,\lambda} z_j^{2\lambda}} e^{-\frac{\frac{y_1}{\alpha_1}+\frac{y_2}{\alpha_2}}{z_j}} \,(y_1y_2)^{\,\lambda-1} g_{\phi_j}(z_j) dZ_j \bigg)$$

The direct maximization of the above Equation with respect to parameter space Θ is complicated. Fortunately, in such cases, the EM algorithm can be used to simplified the maximization problem for the above Equation . In particular, if we augment the unobserved variable $\{Z_j\}_{j=1,...,n}$, then the complete log-likelihood function is given by

$$l_c(\Theta) = \sum_{i=1}^{2}\sum_{j=1}^{n} \bigg(-\lambda \log(\alpha_{i,j}) - 2\lambda \log(z_j) - 2\log\Gamma\lambda - \frac{y_{ij}}{\alpha_{ij} z_j} + (\lambda-1)\log(y_{ij}) \bigg) + \sum_{j=1}^{n} \log(g_{\phi_j}(z_j))$$

The two-steps of EM algorithm are described in what follows.

E − step : The Q-function, $Q(\Theta; \Theta^{(r)})$, which is the conditional posterior expectation of Equation , is given by

$$\begin{aligned} Q(\Theta; \Theta^{(r)}) = \sum_{i=1}^{2}\sum_{j=1}^{n} \bigg(&-\lambda \log(\alpha_{i,j}) - 2\lambda \mathbb{E}^{(r)}_{z_{i,j}}[\log(z_j)] - 2\log\Gamma\lambda - \frac{y_{ij}}{\alpha_{ij}} \mathbb{E}^{(r)}_{z_{i,j}}[z_j^{-1}] \\ &+ (\lambda-1)\log(y_{ij}) \bigg) + \sum_{j=1}^{n} \mathbb{E}^{(r)}_{z_{i,j}}[\log(g_{\phi_j}(z_j))] \end{aligned} \qquad (14)$$

Where $\alpha^{(r)}_{i,j}$=$\exp\{x^\top_{i,j}\beta^{(r)}_i\}$ and where the conditional expectation $\mathbb{E}^{(r)}_{z_j}$ $[h(z)]$ for any real function ,$h(.)$,is defined as follows

$$\mathbb{E}_{z_j}^{(r)} = \mathbb{E}[h(z_j)|\Theta^{(r)}, y_j, x_{1,j}, x_{2,j}] = \int_0^\infty h(z_j)\pi(z|\Theta^{(r)}, y_j, x_{1,j}, x_{2,j})dz_j$$

$\mathbf{M-step}$: After calculating the Q- function, we find its maximum global point,$\Theta^{(j+1)}$, i.e. we update the parameters by computing the gradient function, g(.), and the Hessian matrix, H(.), of the Q-function. In particular, the Newton-Raphson algorithm is used for maximizing the Q-function and the parameters β_1,β_2,β_3 for the Gamma part and the parameter λ for the randnom effect part are updated separately as shown below.

- For the Gamma part,

$$\beta_i^{(r+1)} = \beta_i^{(r)} - H^{-1}(\beta_i^{(r)})g(\beta_i^{(r)}), \quad i = 1, 2$$

$$g(\beta_i^{(r)}) = X_i^\top V_i \qquad H(\beta_i^{(r)}) = X_i^\top D_i X_i$$

$$V_i = \left(\left\{\frac{y_{i,j}}{\alpha_{i,j}^{(r)}}\mathbb{E}_{z_{i,j}}^{(r)}[z^{(-1)}] - \lambda\right\}_{j=1,\ldots,n}\right)$$

$$D_i = diag\left(\left\{-\frac{y_{i,j}}{\alpha_{i,j}^{(r)}}\mathbb{E}_{z_{i,j}}^{(r)}[z^{(-1)}]\right\}_{j=1,\ldots,n}\right),$$

where $X_i = (x_i, 1, ..., x_{i,n})$ is the design matrix for $\alpha_{i,j}^{(r)}$. For λ,we derive the first and second order deratives of λ and then we take the gradient function and the Hessian Matrix.Finally we update λ by using the one-step Newton-Raphson.

$$\lambda^{(r+1)} = \lambda^{(r)} - \frac{h(\lambda^{(r)})}{H(\lambda^{(r)})} \tag{15}$$

where,

$$h(\lambda^{(r)}) = -\sum_{i=1}^{2}\sum_{j=1}^{n}\lambda_{ij}^{(r)} - 4n\psi^{(0)}(\lambda^{(r)}) - 4\sum_{j=1}^{n}\mathbb{E}^{(r)}(\log Z_i) + \sum_{i=1}^{2}\sum_{j=1}^{n}\log y_{ij}$$

$$H(\lambda 6(r)) = -4n\psi^{(1)}(\lambda^{(r)})$$

We can improve the estimates of λ using (5).

.

3.1 The EM algorithm for ML estimation of the Gamma-Inverse Gaussian regression model

The task of obtaining maximum likelihood estimates of the parameters of the model Pareto-Inverse Gaussian is carried out using EM algorithm. The joint log-likelihood function of the complete data can be written as follows,

$$\begin{aligned} l_c(\mathbf{\Theta}) = \sum_{j=1}^{n} (\log(\alpha_j) + \alpha_j . \log(z_j) - (\alpha_j + 1) . \log(y_j)) \\ + \sum_{j=1}^{n} \left(\log(\phi_j) + \phi_j^2 - \frac{3}{2} \log(z_j) - \frac{\phi_j^2 . \left(\frac{1}{z_j} + z_j \right)}{2} \right) \end{aligned} \quad (16)$$

where $\phi_j^{(r)} = \exp(\mathbf{x}_{2,j}^{\top} \boldsymbol{\beta}_{\mathbf{3}}^{(\boldsymbol{r})})$.

The M-steps required conditional expectation for the terms $\log(z_j)$ and $\left(1 - \frac{1}{2}\left(\frac{1}{z_j} + z_j\right)\right)$. Hence the algorithm can be written as

- **E-step:** The conditional expectation can be computed as

$$t_j = \mathbb{E}_{Z_j}\left(\log(Z_j) | y_j ; \mathbf{\Theta}^{(r)}\right)$$

$$t_j = \frac{\int_0^y \log(z_j) \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} \frac{\phi_j^{(r)} \exp\left(\left(1 - \frac{1}{2}\left(z_j + \frac{1}{z_j}\right)\right)(\phi_j^{(r)})^2\right)}{\sqrt{2\pi} z_j^{3/2}} dz_j}{\int_0^y \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} \frac{\phi_{ij}^{(r)} \exp\left(\left(1 - \frac{1}{2}\left(z_j + \frac{1}{z_j}\right)\right)(\phi_j^{(r)})^2\right)}{\sqrt{2\pi} z_j^{3/2}} dz_j} \quad (17)$$

and

$$d_j = \mathbb{E}_{Z_j}\left(1 - \frac{1}{2}\left(\frac{1}{z_j} + z_j\right) | y_j ; \mathbf{\Theta}^{(r)}\right)$$

$$d_j = \frac{\int_0^y \left(1 - \frac{1}{2}\left(\frac{1}{z_j} + z_j\right)\right) \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} \frac{\phi_j^{(r)} \exp\left(\left(1-\frac{1}{2}\left(z_j+\frac{1}{z_j}\right)\right)(\phi_j^{(r)})^2\right)}{\sqrt{2\pi} z_j^{3/2}} dz_j}{\int_0^y \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} \frac{\phi_j^{(r)} \exp\left(\left(1-\frac{1}{2}\left(z_j+\frac{1}{z_j}\right)\right)(\phi_j^{(r)})^2\right)}{\sqrt{2\pi} z_j^{3/2}} dz_j} \tag{18}$$

- **M-step:**

$$h_2(\boldsymbol{\beta_3}) = \left(1 + 2\phi_j^{(r)} d_j\right) x_{2,jk} \tag{19}$$

and

$$H_2(\boldsymbol{\beta_3}) = \left(4\phi_j^{(r)} d_j\right) x_{2,jk} x_{2,jk}^\top = \mathbf{X}_3^\top \mathbf{D}_3 \mathbf{X}_3 \tag{20}$$

where $j = 1, 2, \cdots, n$, $k = 1, 2, \cdots, p_2$ and $\mathbf{D}_3 = diag\{4\phi_j^{(r)} d_j\}$. Equation (**??**) can be used to improve the estimates of $\boldsymbol{\beta_3^{(r)}}$.

- Finally, iterate between the E-step and the M-step until some convergence criterion is satisfied, for example the relative change in log-likelihood between two successive iterations is smaller than 10^{-12}.

3.2 EM algorithm for ML estimation of Gamma-Weibull regression model

Due to the complex structure of density function of the Weibull model, direct maximization of the log-likelihood through usual way will not result into efficient estimates of the parameters of the model. An EM algorithm can be used to efficiently compute the estimates of the various parameter of the Weibull model as follows

The complete data log-likelihood takes the form

$$l_c(\theta) = \sum_{j=1}^{n} \left[\log(\alpha_j) + \alpha_j . \log(z_j) - (\alpha_j + 1). \log(y_j)\right]$$
$$+\sum_{j=1}^{n} \left[\log(\phi_j) + \log\left(\Gamma\left(1+\frac{1}{\phi_j}\right)\right) + (\phi_j - 1).\left(\log(z_j) + \log\left(\Gamma\left(1+\frac{1}{\phi_j}\right)\right)\right) - z_j.\Gamma\left(1+\frac{1}{\phi_j}\right)\right] \tag{21}$$

where $\phi_j^{(r)} = \exp(\mathbf{x}_{2,j}^{\top}\boldsymbol{\beta_3^{(r)}})$.

The expectation of $\log(z_j)$ and z_j are needed for the process of M-step. The EM algorithm can be written as

- E-step: The required expectations for $j = 1, 2...n$ can be computed as

$$t_j = \mathbb{E}_{Z_j}\left[\log(Z_j)|y_j; \theta^{(r)}\right]$$

$$t_j = \frac{\int_0^y \log(z_j) \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} . \phi_j^{(r)} . \left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right] e^{-\left(z_j\left[\Gamma(1+\frac{1}{\phi_j^{(r)}})\right]\right)^{\phi_j^{(r)}}} \left(z_j\left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right]\right)^{\phi_j^{(r)}-1} dz_j}{\int_0^y \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} . \phi_j^{(r)} . \left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right] e^{-\left(z_j\left[\Gamma(1+\frac{1}{\phi_j^{(r)}})\right]\right)^{\phi_j^{(r)}}} \left(z_j\left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right]\right)^{\phi_j^{(r)}-1} dz_j} \tag{22}$$

and

$$d_j = \mathbb{E}_{Z_j}\left[Z_j|y_j; \theta^{(r)}\right]$$

$$d_j = \frac{\int_0^y (z_j) \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} . \phi_j^{(r)} . \left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right] e^{-\left(z_j\left[\Gamma(1+\frac{1}{\phi_j^{(r)}})\right]\right)^{\phi_j^{(r)}}} \left(z_j\left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right]\right)^{\phi_j^{(r)}-1} dz_j}{\int_0^y \frac{\alpha_j^{(r)} z_j^{\alpha_j^{(r)}}}{y_j^{\alpha_j^{(r)}+1}} . \phi_j^{(r)} . \left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right] e^{-\left(z_j\left[\Gamma(1+\frac{1}{\phi_j^{(r)}})\right]\right)^{\phi_j^{(r)}}} \left(z_j\left[\Gamma\left(1+\frac{1}{\phi_j^{(r)}}\right)\right]\right)^{\phi_j^{(r)}-1} dz_j} \tag{23}$$

where $\phi_j^{(r)} = \exp(\mathrm{x}_2^{\top}\beta_2^{(r)})$.

The closed form expressions for the above expectations are not easily available hence numerical approximations are required to compute above mentioned quantities.

- M-step: Using the numerical approximate value of t_j, update the regression parameters β_1. Update the regression parameters β_2 using the approximate values of t_j and d_j, the Newton-Raphson algorithm to improve the estimates of regression parameters given as

$$h_2(\beta_3) = \left[1 + \left(\left(1 - \Gamma\left(1 + \frac{1}{\phi_j^{(r)}}\right)^{\phi_j^{(r)}} .d_i^{\phi_j^{(r)}}\right)\left(\phi_j^{(r)}.\log\left(\Gamma\left(1 + \frac{1}{\phi_j^{(r)}}\right)\right) - F\left(1 + \frac{1}{\phi_j^{(r)}}\right)\right.\right.\right.$$
$$\left.\left.\left. + \phi_j^{(r)}.t_j\right)\right)\right]\mathbf{x}_{2,jk} \tag{24}$$

$$H_2(\beta_3) = [A_1 + A_2 + A_3 + A_4 + A_5]\,\mathbf{x}_{2,jk}\mathbf{x}_{3,jk}^\top = X_3^\top D_3 X_3 \tag{25}$$

where $j = 1, 2, 3, ...n$ and $k = 1, 2, 3....p_2$. The matrix D_3 can be written as $D_3 = diag\{A_1 + A_2 + A_3 + A_4 + A_5\}$.

Where $A_1 = \phi_j^{(r)}\log(\Gamma(1+\frac{1}{\phi_j^{(r)}}))-\Psi^{(0)}(1+\frac{1}{\phi_j^{(r)}})$, $A_2 = \frac{\Psi^{(1)}\left(1+\frac{1}{\phi_j^{(r)}}\right)}{\phi_j^{(r)}}$, $A_3 = \phi_j^{(r)}\log(z_j)$, $A_4 = \log(z_j)\left[1 + \phi_j^{(r)}.\log(z_j) + \phi_j^{(r)}.\log(\Gamma(1 + \frac{1}{\phi_j^{(r)}})) - \Psi^{(0)}(1 + \frac{1}{\phi_j^{(r)}})\right].z_j^{\phi_j^{(r)}}\left(\Gamma(1 + \frac{1}{\phi_j^{(r)}})\right).\phi_j^{(r)}$ and

$A_5 = z_j^{\phi_j^{(r)}}.\left(\Gamma(1 + \frac{1}{\phi_j^{(r)}})\right)\left[\left(\phi_j^{(r)}\right)^2.\log\left(\Gamma\left(1 + \frac{1}{\phi_j^{(r)}}\right)\right)\left(1 + \phi_j^{(r)}.\log(z_j) + \phi_j^{(r)}.\log\left(\Gamma\left(1 + \frac{1}{\phi_j^{(r)}}\right)\right)\right) - \phi_j^{(r)}\right.$

The improved estimates of $\beta_3^{(r)}$ can be obtained using Newton-Raphson mathod.

- Finally, iterate between the E-step and the M-step until some convergence criterion is satisfied, for example the relative change in log-likelihood between two successive iterations is smaller than 10^{-12}.

4 Empirical Analysis

The study is based on data from automobile policies from a major insurance European company for the underwriting years 2012 - 2019. This data set contains bodily injury (BI) and property damage (PD) claims and their associated claim costs, denoted by Y_1 and Y_2 respectively, and risk factors that affect both Y_1 and Y_2. There were 7263 observations in total which met our criteria of having at least one claim made The summary statistics for Y_1 and Y_2 are shown in Table 1 and Figure 1. As it was expected, both Y_1 and Y_2 are positively skewed. Also, the Pearson correlation test indicates that it is appropriate to model both types of claim costs based on a single bivariate model rather than two independent univariate models.

Table 1: Summary statisitcs of two types of d claim amount. The correlation test is an one-sided test, where the alternative hypothesis is 'true correlation is greater than 0'.

Aggregated Claim	min	median	mean	max	standard deviation	correlation	p-value
Y_1	0.9	2413.4	11017.3	251958.2	27128.85	0.1095	0.0000
Y_2	6.2	1012.4	1871.2	14818.2	2217.138		

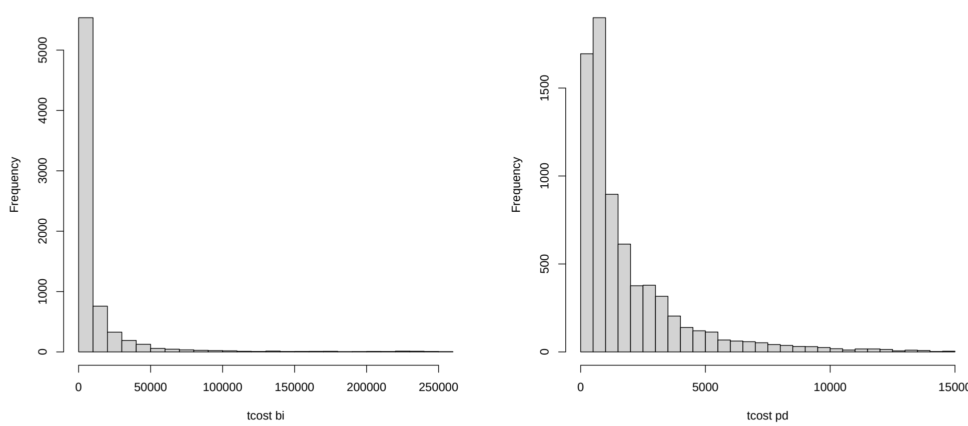

Figure 1: Empirical distribution of two types of d claim amount

Furthermore, a description of the explanatory variables which we included in the 144 regression analysis for Y1 and Y2 is provided below.

• The variable Driver's age. Policyholders aged 18 to 90 years old.

• The variable Vehicle's age. Vehicles aged 0 to 60 years old.

• The variable Car cubism, 'CC', consists of four categories. Vehicles with horse power '0-1299 cc' (C1), '1300 -1399 cc' (C2), '1400 - 1599 cc' and ' greater or equal 1600 cc' (C3).

• The variable 'PT' consisted of three types of policy, 'Economic type which includes only MTPL coverage' (C1) , 'Middle type which includes apart from MTPL coverage other types of coverage' (C2), and 'Expensive type – Own coverage' (C3)

• The variable 'Region' consisted of three board regions, 'Captial city'(C1), 'province cities of the mainland'(C2), and 'province cities of the island area' (C3)

Additionally, the empirical distributions of the categorical and continuous explanatory variables are shown in Table 2 and Figure 2 respectively. The BPA and BEIG regression models were fitted to the claim costs $Y = (Y_1, Y_2)$. All computing was done using the **R** software. The vector of parameters$\Theta = \{\beta_1, \beta_2, \beta_3\}$ was estimated using the EM algorithm which was presented in Section 3 and their standard deviations weree obtained through expressions that were directly produced by the EM algorithm for the observed information matrix of each model. Finally, the fit of the competing models was compared by employing

Table 2: Empirical distributions of categorical variables

	Horse power(CC)	**Policy Type(PT)**	**Region**
C1	2036	1144	4220
C2	2417	1940	2333
C3	1833	4179	710
C4	977	-	-

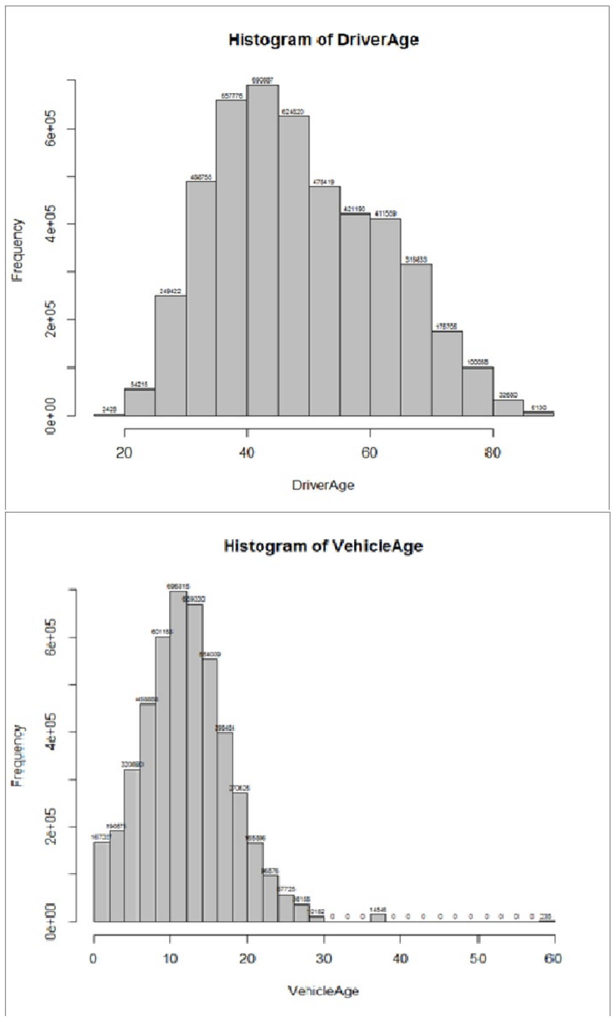

Figure 2: Empirical distributions of continuous explanatory variables

Table 3: Estimated parameters and standard errors in parentheses for the BGIG and BGW regressions model. AIC: Akaike information criterion; BIC: Bayesian information criterion

	BGIG			**BGW**		
Response	β_1	β_2	β_3	β_1	β_2	β_3
Intercept	8.0395	7.7080	-0.9201	8.1177	7.5724	0.4663
Driver's Age	-0.0002	0.0012	0.0022	0.0013	0.0003	0.0076
CC:C2	-0.0009	0.1945	-0.0892	-0.0505	0.1706	-0.2686
CC:C3	0.0489	0.0487	0.0098	0.0285	0.0708	-0.1824
CC:C4	0.0489	0.0487	0.0098	0.0022	0.1727	-0.2971
PT:C2	0.6450	0.1864	-0.0763	0.3705	-0.4283	1.3525
PT:C3	0.1801	-0.4467	0.3701	0.2879	-0.2617	0.3226
Vehicle's Age	0.2309	-0.3615	0.2449	0.0124	-0.0110	0.0323
Region:C2	0.0088	-0.0115	0.0105	-0.1029	0.1771	-0.3498
Region:C3	0.1976	-0.1137	0.1175	0.2428	-0.1237	0.4724
AIC	**275811.22**			**273941.22**		
BIC	**275868.88**			**273998.88**		

5 Concluding Remarks

In this paper, we developed a class of bivariate mixed Gamma regression models which can approximate moderate and large claim costs in an efficient manner based on the choice of mixing density. We illustrated our approach by fitting the BGW and BGIG regression models on MTPL data which were provided by a European insurance company. The proposed family of models can accommodate the positive correlation between MTPL bodily injury and property damage claims and their associated costs, when explanatory variables for each type of claims are taken into account through regression structure for their mean as well as dispersion parameters. The main achievement is that we developed an EM-type algorithm which computa- tionally efficient. This was demonstrated by obtaining reliable estimates when applying the models to the real data. Here we put the covariates in α_i which is a part of mean and also in ϕ_j which is a part of dispersion for getting better results. The BGW regression model provides good fit as compare to BGIG regression model for the MTPL dataset

Abbreviations

The following abbreviations are used in this work

BGW	Bivariate Gamma Weibull
BGIG	Bivariate Gamma- Inverse Gaussian
EM	Expectation-Maximization
IG	Inverse Gaussian

References

[1] Bakar, SA Abu, Nor A Hamzah, Mastoureh Maghsoudi, and Saralees Nadarajah. (2015). Modeling loss data using composite models. *Insurance: Mathematics and Economics* 61, 146–154.

[2] Calderín-Ojeda, Enrique and Chun Fung Kwok. (2016). Modeling claims data with composite Stoppa models. *Scandinavian Actuarial Journal* (2016)(9), 817–836.

[3] Chen, Z., Dassios, A., & Tzougas, G. (2022). EM Estimation for the Bivariate Mixed Exponential Regression Model. *Risks*, 10(5), 105.

[4] Frees, Edward W, Richard A Derrig, and Glenn Meyers. (2014). Predictive modeling applications in actuarial science, Volume 1. *Cambridge University Press.*

[5] Frees, E. W. and E. A. Valdez. (2008). Hierarchical insurance claims modelling. *Journal of the American Statistical Association.*

[6] Fung, Tsz Chai, Andrei L. Badescu, and X. Sheldon Lin. (2021). A new class of severity regression models with an application to ibnr prediction. *North American Actuarial Journal* 25(2), 206–231.

[7] Grün, Bettina and Tatjana Miljkovic. (2019). Extending composite loss models using a general framework 201 of advanced computational tools. *Scandinavian Actuarial Journal* (2019)(8), 642–660.

[8] Hürlimann, Werner. (2014). Pareto type distributions and excess-of-loss reinsurance. *International Journal of Research and Reviews in Applied Sciences* 18(3), 1.

[9] Jeong, Himchan. (2020). Testing for random effects in compound risk models via Bregman divergence. 205 *ASTIN Bulletin: The Journal of the IAA 50(3)*, 777–798.

[10] Laudagé, Christian, Sascha Desmettre, and Jörg Wenzel. (2019). Severity modeling of extreme insurance claims for tariffication. *Insurance: Mathematics and Economics* 88, 77–92.

[11] Version March 29, (2022) submitted to Risks 10 of 10 Lee, Simon CK and X. Shelson Lin. 2010. Modeling and evaluating insurance losses via mixtures of *Erlang distributions. North American Actuarial Journal* 14(1), 107–130.

[12] Miljkovic, Tatjana and Bettina Grün. (2016). Modeling loss data using mixtures of distributions. *Insurance: Mathematics and Economics* 70, 387 – 396. doi:https://doi.org/10.1016/j.insmatheco.2016.06.019.

[13] Nadarajah, Saralees and SA Abu Bakar. (2014). New composite models for the Danish fire insurance data. *Scandinavian Actuarial Journal* 2014(2), 180–187.

[14] Parodi, Pietro. (2020). A generalised property exposure rating framework that incorporates scale- independent losses and maximum possible loss uncertainty. *Astin Bulletin* 50(2), 513–553.

[15] Pigeon, Mathieu and Michel Denuit. (2011). Composite lognormal-Pareto model with random threshold. *Scandinavian Actuarial Journal* (2011)(3), 177–192.

[16] Ramirez-Cobo, Pepa, Rosa E Lillo, Simon Wilson, Michael P Wiper, et al. (2010). Bayesian inference for double pareto lognormal queues. *The Annals of Applied Statistics* 4(3), 1533–1557.

[17] Reynkens, Tom, Roel Verbelen, Jan Beirlant, and Katrien Antonio. (2017). Modelling censored losses using splicing: A global fit strategy with mixed Erlang and extreme value distributions. *Insurance: Mathematics and Economics* 77, 65–77.

[18] Scollnik, David PM and Chenchen Sun. (2012). Modeling with Weibull-Pareto models. *North American Actuarial Journal* 16(2), 260–272.

[19] Tzougas, George and Dimitris Karlis. (2020). An em algorithm for fitting a new class of mixed exponential 226 regression models with varying dispersion. *Astin Bulletin.*

[20] Tzougas, George, Spyridon Vrontos, and Nicholas Frangos. (2014). Optimal bonus-malus systems using finite mixture models. *Astin Bulletin* 44(2), 417–444.

[21] Tzougas, George, Spyridon Vrontos, and Nicholas Frangos. (2018). Bonus-malus systems with two- component mixture models arising from different parametric families. *North American Actuarial Journal* 22(1), 55–91.

[22] Tzougas, George, Woo Hee Yik, and Muhammad Waqar Mustaqeem. (2019). Insurance ratemaking using the exponential-lognormal regression model. |textitAnnals of Actuarial Science, 1–30.

[23] Wang, Yinzhi, Ingrid Hobæk Haff, and Arne Huseby. (2020). Modelling extreme claims via composite models and threshold selection methods. *Insurance: Mathematics and Economics* 91, 257–268.