Ratemaking of dependent risks*

Andrade e Silva, J.M. and Centeno, M.L.

CEMAPRE, ISEG, Universidade de Lisboa

Abstract

We start by describing how, in some cases, we can use variance related premium principles in ratemaking, when the claim numbers and individual claim amounts are independent. We use quasi-likelihood generalized linear models, under the assumption that the variance function is a power function of the mean of the underlying random variable.

We extend this approach to the cases where the claim numbers are correlated. Some alternatives to deal with dependent risks are presented, taking explicitly into account overdispersion. We present regression models covering the bivariate Poisson, the generalized bivariate negative binomial and the bivariate Poisson-Laguerre polynomial, which nest the bivariate negative binomial. We apply these models to a portfolio of the Portuguese insurance company Tranquilidade and compare the results obtained.

Keywords: ratemaking, variance related premium principles, generalized linear models, quasi-likelihood, bivariate distributions.

*The authors are grateful to the insurance company Tranquilidade, that provided the data used in this work. The authors were partially supported by the project CEMAPRE – MULTI/00491 financed by FCT/MEC through national funds and when applicable co-financed by FEDER, under the Partnership Agreement PT2020.

1 Introduction

The aim of an actuary when defining a tariff, is that the premiums match the risks as closely as possible. This is achieved by differentiating the risks on the basis of observable risk factors.

In the last decades, actuaries have used Generalized Linear Models (GLM) - see Nelder and Wedderburn (1972) or Denuit et al (2006) - to construct motor insurance tariff structures. Brockman et al (1992) provides a survey of application of Generalized Linear Models to Motor Insurance ratemaking. Implicitly, when doing so, the premium calculation principle used is the expected value principle. The premium is calculated proportionally to the conditional expected value of the aggregate claims, given a set of tariff variables. It is assumed, in most cases, independence between claim numbers and individual claim amounts. The conditional expected value of these two variables is estimated separately and the two estimates are multiplied. When the tariff structure is multiplicative the tariff can be interpreted in an aggregated way. In the GLM framework the estimators are calculated by maximizing the log-likelihood of the underlying distributions, which are assumed to belong to the exponential family. As it is well known, this assumption can be relaxed and the parameters estimated consistently using quasi-likelihood, in which case we only have to model the expected value and the variance function.

First, we show that, although in the GLM framework, most actuaries use the expected value principle for ratemaking purposes, there is no reason for not using variance related premium principles as explained in Section 2. Second, we discuss ratemaking when risks are correlated through claim numbers, which is a situation that occurs very often in practice. There is some work already done on this topic, namely Bermudez (2009). We introduce, in Section 3, other options to model bivariate counting data taking overdispersion explicitly into account. We present four models, which include the bivariate Poisson and the bivariate Negative Binomial.

In Section 4 we provide an example with data from the Portuguese insurance company Tranquilidade and compare the results obtained with the four models, as well as with the corresponding independent cases.

Finally Section 5 concludes.

2 Variance related premium principles

Let S_i be the aggregate claim amount of policy i, for a given period of time. We assume that S_i is a compound random variable, i.e. that

$$S_i = \sum_{j=0}^{N_i} Y_{ij},$$

where N_i is the number of claims for the same period of time, $Y_{i0} \equiv 0$ and $\{Y_{ij}\}_{j=1,2,...}$ are i.i.d. random variable and independent of N_i, representing the individual claim amounts. Under these assumptions

$$E[S_i] = E[N_i]E[Y_i] = \mu_{N_i}\mu_{Y_i}$$

and

$$Var[S_i] = \mu_{N_i} Var[Y_i] + Var[N_i]\mu_{Y_i}^2,$$

where Y_i is identically distributed to Y_{ij}.

In most regression models it is assumed that $Var[N_i]$ and $Var[Y_i]$ are functions of μ_{N_i} and of μ_{Y_i} respectively, which implies that $Var[S_i]$ can be expressed as a function of μ_{N_i} and μ_{Y_i}. For example if

$$Var[N_i] = \psi\mu_{N_i} \tag{1}$$

and

$$Var[Y_i] = \phi\mu_{Y_i}^2, \tag{2}$$

we get the quasi-Poisson-Gamma model, for which

$$Var[S_i] = (\psi + \phi)\mu_{N_i}\mu_{Y_i}^2. \tag{3}$$

Given the row vectors of covariates $\mathbf{X}_i = (X_1, X_2, ..., X_{p_1})$ and $\mathbf{Z}_i = (Z_1, Z_2, ..., Z_{p_2})$ to explain the conditional expected values of N_i and Y_i respectively (all or some of the covariates can be the same), and using a log-link function, the conditional expected values are given by

$$E[N_i|\mathbf{X}_i] = \exp(\mathbf{X}_i\boldsymbol{\beta}) \tag{4}$$

and

$$E[Y_i|\mathbf{Z}_i] = \exp(\mathbf{Z}_i\boldsymbol{\gamma}) \tag{5}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are two vectors with dimensions p_1 and p_2 respectively.

When using the expected value principle, the premium associated to policy i, given $\mathbf{X}_i$ and $\mathbf{Z}_i$, is of the form

$$P_i^{EV} = (1+\alpha)\exp(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\gamma}),\ \alpha > 0, \tag{6}$$

where α is the loading coefficient, and when using a related variance premium principle for the quasi-Poisson-Gamma model is

$$P_i^{RV} = \exp(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\gamma}) + g\left((\psi+\phi)\exp(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{2Z_i}\boldsymbol{\gamma})\right), \tag{7}$$

where $g(.)$ is an increasing and non-negative function. For the variance principle $g(x) = \delta x$, with $\delta > 0$, and for the standard deviation principle $g(x) = \delta\sqrt{x}$, with $\delta > 0$. Note that, for a given α, for the variance principle, and for δ calculated in such a way that the global premiums are the same, the main difference between (6) and (7) is that the variance principle penalizes more the policies with higher expected claim amount.

The parameter estimation is straightforward: $\boldsymbol{\beta}$ is obtained by quasi-likelihood of the claim number observations and $\boldsymbol{\gamma}$ is obtained by quasi-likelihood of the individual claim amounts. The parameters ψ and ϕ, are estimated in each model using, for instance, the moment estimator based on the chi-square statistic (see McCullagh and Nelder (1989)).

3 Bivariate Regression Models

In this section we consider, for each policy, a model involving two risks, dependent through the number of claims. For policy i, let $\left\{Y_{ij}^{(k)}\right\}_{j=1,2,\ldots}$ be the claim size random variables for risk k, $k = 1, 2$. We assume that, for $k = 1, 2$, $\left\{Y_{ij}^{(k)}\right\}_{j=1,2,\ldots}$ are i.i.d. random variables. Let $S_i^{(1)}$ and $S_i^{(2)}$ be the aggregate claim amounts for the first and second risks respectively, with

$$S_i^{(k)} = \sum_{j=0}^{N_i^{(k)}} Y_{ij}^{(k)}, \tag{8}$$

where $N_i^{(k)}$ is the number of claims of policy i and risk k in a given period of time. We consider that $Y_{i0}^{(1)} \equiv Y_{i0}^{(2)} \equiv 0$, that $\left\{Y_{ij}^{(1)}\right\}_{j=1,2,\ldots}$ are independent of $\left\{Y_{ij}^{(2)}\right\}_{j=1,2,\ldots}$ and that both are independent of

$N_i^{(1)}$ and $N_i^{(2)}$. The aggregate claim amount for policy i is

$$S_i = S_i^{(1)} + S_i^{(2)}, \tag{9}$$

with expected value

$$E[S_i] = \mu_{N_i^{(1)}}\mu_{Y_i^{(1)}} + \mu_{N_i^{(2)}}\mu_{Y_i^{(2)}}, \tag{10}$$

and variance

$$\begin{aligned} Var[S_i] \quad = \quad & \mu_{N_i^{(1)}} Var[Y_i^{(1)}] + Var[N_i^{(1)}]\left(\mu_{Y_i^{(1)}}\right)^2 \\ & + \mu_{N_i^{(2)}} Var[Y_i^{(2)}] + Var[N_i^{(2)}]\left(\mu_{Y_i^{(2)}}\right)^2 \\ & + 2\mu_{Y_i^{(1)}}\mu_{Y_i^{(2)}} Cov(N_i^{(1)}, N_i^{(2)}). \end{aligned} \tag{11}$$

As the main purpose of the paper is to analyse dependency on claim frequencies our focus will be on the claim numbers behaviour. We begin by presenting two models based on the bivariate Poisson distribution and then we consider two generalizations of the bivariate negative binomial distribution to deal with overdispersion, which is a phenomenon that is present in many insurance data sets. For sake of simplicity we present the results using the gamma distribution for the claim severity but these results are easily extended to all distributions where the variance is given by $Var[Y_i] = \phi\mu_{Y_i}^k$, for a given constant k.

3.1 Bivariate Poisson Models

Let

$$N_i^{(1)} = K_i^{(1)} + K_i \qquad \text{and} \qquad N_i^{(2)} = K_i^{(2)} + K_i \tag{12}$$

where $K_i^{(1)}$, $K_i^{(2)}$ and K_i are independent Poisson random variables with parameters $\lambda_i^{(1)}$, $\lambda_i^{(2)}$ and λ_i respectively, i.e. $\left(N_i^{(1)}, N_i^{(2)}\right)$ is the bivariate Poisson distribution, studied by Holgate (1964) (see also Johnson et al (1997)). Hence $S_i^{(1)}$ and $S_i^{(2)}$ are correlated through K_i. This model was, among others, considered in the context of actuarial science by Wang (1998), Walhin (2001), Cossete and Marceau (2000), and Centeno (2005). In this model

$$Cov(N_i^{(1)}, N_i^{(2)}) = \lambda_i.$$

In the Poisson-Gamma framework (11) is equivalent to

$$Var[S_i] = \left(1+\phi^{(1)}\right)\mu_{N_i^{(1)}}\left(\mu_{Y_i^{(1)}}\right)^2 + \left(1+\phi^{(2)}\right)\mu_{N_i^{(2)}}\left(\mu_{Y_i^{(2)}}\right)^2 + 2\mu_{Y_i^{(1)}}\mu_{Y_i^{(2)}}\lambda_i, \qquad (13)$$

where $\phi^{(1)}$ and $\phi^{(2)}$ are the dispersion parameters of risks 1 and 2 respectively (see (2)).

Note that under these models the covariance between the two risks is always assumed positive, since it is equal to the mean of the common Poisson random variable, which is not a real problem for most insurance applications. Note also that in the Poisson regression models the heterogeneity of the portfolio is only considered through the means of the endogenous random variables which can be quite different.

The main difference between the next two models lies on the way the covariates are introduced. While the Kocherlakota and Kocherlakota (2001) model explains the means of the observed variables $N_i^{(1)}$ and $N_i^{(2)}$, the Karlis and Ntzoufras (2005) model explains the means of the latent variables $K_i^{(1)}$, $K_i^{(2)}$ and K_i. These different approaches lead to different result interpretations and as it will be seen in our example to quite different estimates for the premiums.

3.1.1 K-K Bivariate Poisson Regression Model (KKBIP)

Kocherlakota and Kocherlakota (2001) considered that the common random variable K_i has mean λ, independent of i, i.e. that $\lambda_i = \lambda$, $i = 1, 2, ...$ The bivariate Poisson distribution is

$$f(n_i^{(1)}, n_i^{(2)}) = \exp(-\lambda_i^{(1)} - \lambda_i^{(2)} - \lambda)h(n_i^{(1)}, n_i^{(2)}), \qquad (14)$$

where

$$h(n_i^{(1)}, n_i^{(2)}) = \sum_{j=0}^{\min(n_i^{(1)}, n_i^{(2)})} \frac{(\lambda_i^{(1)})^{n_i^{(1)}-j}(\lambda_i^{(2)})^{n_i^{(2)}-j}\lambda^j}{(n_i^{(1)}-j)!(n_i^{(2)}-j)!j!}. \qquad (15)$$

For this distribution $\mu_{N_i^{(1)}} = \lambda_i^{(1)} + \lambda$, $\mu_{N_i^{(2)}} = \lambda_i^{(2)} + \lambda$ and $Cov\left(N_i^{(1)}, N_i^{(2)}\right) = \lambda$. The response variables $\left(N_i^{(1)}, N_i^{(2)}\right)$ are related with the covariates $\mathbf{X}_i^{(1)}$ and $\mathbf{X}_i^{(2)}$ through

$$E[N_i^{(k)}|\mathbf{X}_i^{(k)}] = \exp(\mathbf{X}_i^{(k)}\boldsymbol{\beta}^{(k)}),\ k = 1, 2. \qquad (16)$$

The maximum likelihood estimators of $\boldsymbol{\beta}^{(1)} = \left(\boldsymbol{\beta}_1^{(1)}, ..., \boldsymbol{\beta}_{p_1}^{(1)}\right)$, $\boldsymbol{\beta}^{(2)} = \left(\boldsymbol{\beta}_1^{(2)}, ..., \boldsymbol{\beta}_{p_1}^{(2)}\right)$ and λ were derived in Kocherlakota and Kocherlakota (2001), as well as the second order derivatives of the likelihood function. In this model the parameter λ is a nuisance parameter, used only for estimation purposes.

The related variance premium principle for a policy i is

$$P_i^{RV} = \sum_{k=1,2} \exp(\mathbf{X}_i^{(k)}\boldsymbol{\beta}^{(k)} + \mathbf{Z}_i^{(k)}\boldsymbol{\gamma}^{(k)}) + g\left(Var[S_i]\right), \tag{17}$$

where for the Bivariate Poisson-Gamma model

$$\begin{aligned} Var[S_i|\mathbf{X}_i^{(1)}, \mathbf{X}_i^{(2)}, \mathbf{Z}_i^{(1)}, \mathbf{Z}_i^{(2)}] &= \left(1+\phi^{(1)}\right)\exp(\mathbf{X}_i^{(1)}\boldsymbol{\beta}^{(1)} + 2\mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}) + \\ &+ \left(1+\phi^{(2)}\right)\exp(\mathbf{X}_i^{(2)}\boldsymbol{\beta}^{(2)} + 2\mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}) \\ &+ 2\lambda\exp(\mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)} + \mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}). \end{aligned} \tag{18}$$

The parameters $\boldsymbol{\beta}^{(k)}$, $\boldsymbol{\gamma}^{(k)}$, $k = 1, 2$, and λ are estimated by the maximum likelihood function of the bivariate random variable.

The bivariate Poisson model has a limitation on its applicability. As pointed out by Holgate (1964) the correlation coefficient between $N_i^{(1)}$ and $N_i^{(2)}$, equals to $\lambda/\sqrt{\mu_{N_i^{(1)}}\mu_{N_i^{(2)}}}$, cannot exceed the square root of the ratio of the smaller to the larger of the means of the two marginal distributions, i.e.

$$\text{corr}(N_i^{(1)}, N_i^{(2)}) = \frac{\lambda}{\sqrt{\mu_{N_i^{(1)}}\mu_{N_i^{(2)}}}} < \min\left(\sqrt{\frac{\mu_{N_i^{(1)}}}{\mu_{N_i^{(2)}}}}, \sqrt{\frac{\mu_{N_i^{(2)}}}{\mu_{N_i^{(1)}}}}\right), \tag{19}$$

which is equivalent to

$$\lambda < \min(\mu_{N_i^{(1)}}, \mu_{N_i^{(2)}}). \tag{20}$$

This condition is obviously implicit in the model, due to (12). In Kocherlakota and Kocherlakota (2001), as $\mu_{N_i^{(1)}}$ and $\mu_{N_i^{(2)}}$ are explained by the covariates $\mathbf{X}_i^{(1)}$ and $\mathbf{X}_i^{(2)}$ and λ needs to be estimated (depending indirectly on the covariates) the test to the appropriateness of the model can not be done *a priori*. This may be a reason why Kocherlakota and Kocherlakota (2001) model, although highly cited, is rarely used.

3.1.2 K-N Bivariate Poisson Regression Model (KNBIP)

In the model considered by Karlis and Ntzoufras (2005) and applied by Bermudez (2009) in the context of motor insurance, the random variable K_i has mean λ_i, not necessarily constant for all i. The regression model considered is

$$\begin{aligned} \lambda_i^{(1)} &= \exp\left(\mathbf{X}_i^{(1)}\mathbf{w}^{(1)}\right), \\ \lambda_i^{(2)} &= \exp(\mathbf{X}_i^{(2)}\mathbf{w}^{(2)}), \\ \lambda_i &= \exp(\mathbf{X}_i\mathbf{w}). \end{aligned} \tag{21}$$

where $\mathbf{X}_i^{(1)}$, $\mathbf{X}_i^{(2)}$ and $\mathbf{X}_i$ are now the set of covariates used to model the parameters $\lambda_i^{(1)}$, $\lambda_i^{(2)}$ and λ_i respectively. In this model

$$E[N_i^{(k)}|\mathbf{X}_i^{(k)}] = \exp\left(\mathbf{X}_i^{(k)}\mathbf{w}^{(k)}\right) + \exp(\mathbf{X}_i\mathbf{w}),\ k = 1, 2.$$

So, if an effect is present in $\mathbf{w}^{(1)}$ and $\mathbf{w}$, its effect is no longer multiplicative. Even in the case where λ_i is assumed constant (no covariates) and the covariates used to estimate $\lambda_i^{(1)}$ and $\lambda_i^{(2)}$ are the same used in regression (16) this model is different from the model described earlier. Karlis and Ntzoufras (2005) have also considered a generalization of the model by inflating the diagonal. In their article they implemented the model in R using an EM algorithm to maximize the log-likelihood function of the claim numbers.

In this model we have

$$\begin{aligned} E[S_i|\mathbf{X}_i^{(1)}, \mathbf{X}_i^{(2)}, \mathbf{X}_i, \mathbf{Z}_i^{(1)}, \mathbf{Z}_i^{(2)}] &= \exp\left(\mathbf{X}_i^{(1)}\mathbf{w}^{(1)} + \mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}\right) + \exp(\mathbf{X}_i\mathbf{w} + \mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}) + \\ &+ \exp\left(\mathbf{X}_i^{(2)}\mathbf{w}^{(2)} + \mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}\right) + \exp(\mathbf{X}_i\mathbf{w} + \mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}), \end{aligned} \tag{22}$$

and in the Bivariate Poisson - Gamma framework we have

$$\begin{aligned} Var[S_i|\mathbf{X}_i^{(1)}, \mathbf{X}_i^{(2)}, \mathbf{X}_i, \mathbf{Z}_i^{(1)}, \mathbf{Z}_i^{(2)}] &= \left(1 + \phi^{(1)}\right)\left(\exp\left(\mathbf{X}_i^{(1)}\mathbf{w}^{(1)}\right) + \exp(\mathbf{X}_i\mathbf{w})\right)\exp\left(2\mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}\right) + \\ &+ \left(1 + \phi^{(2)}\right)\left(\exp\left(\mathbf{X}_i^{(2)}\mathbf{w}^{(2)}\right) + \exp(\mathbf{X}_i\mathbf{w})\right)\exp\left(2\mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}\right) + \\ &+ 2\exp(\mathbf{X}_i\mathbf{w} + \mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)} + \mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}). \end{aligned} \tag{23}$$

Karlis and Ntzoufras (2005) model is forcing condition (19) to be satisfied, so one should be careful about the appropriateness of the model in a specific situation.

3.2 Generalized bivariate negative binomial models

For many insurance data sets the Poisson regression model does not capture all the heterogeneity of the portfolio. A mixed Poisson regression model appears as a natural alternative. Among these models the negative binomial model (considering the gamma as the mixing distribution) is the most popular. In the first model Gurmu and Elder propose a generalization of the gamma for the mixing distribution, which is the same for both risks. In the second model each risk is influenced by its own mixing distribution, which is a generalization of the gamma distribution, and the two mixing distributions may be correlated.

3.2.1 Generalized bivariate negative binomial regression model (GBINB)

In this subsection we consider the model proposed by Gurmu and Elder (2000) to describe $\left(N_i^{(1)}, N_i^{(2)}\right)$ to allow for overdispersion. In this model V_i is an unobserved heterogeneity component with density $g(v_i)$ given by

$$g(v_i) = \frac{v_i^{\alpha-1}\beta^\alpha}{\Gamma(\alpha)(1+c^2)} e^{-\beta v_i} \left[1 + c\frac{\alpha - \beta v_i}{\sqrt{\alpha}}\right]^2, \; v_i > 0. \tag{24}$$

It is also assumed that $(V_1, V_2, ...)$ are i.i.d. random variables and that given $V_i = v_i$ the variables $N_i^{(k)}$, $k = 1, 2$ are independent Poisson random variables with mean $\mu_{N_i^{(k)}} v_i$, $k = 1, 2$. This model, referred as generalized bivariate negative binomial, nests when $c = 0$ a bivariate negative binomial.

As it is usual, the mean of the unobserved heterogeneity is set equal to unity, which is to say,

$$\beta = \frac{1}{1+c^2}\left(\alpha - 2c\sqrt{\alpha} + c^2(\alpha+2)\right). \tag{25}$$

The conditional expected values are

$$E\left[N_i^{(k)} | \mathbf{X}_i^{(k)}\right] = \exp(\mathbf{X}_i^{(k)} \boldsymbol{\beta}^{(k)}), \; k = 1, 2 \tag{26}$$

and Gurmu and Elder (2000) show that the probability function of $\left(N_i^{(1)}, N_i^{(2)}\right)$ can be written as

$$f(n_i^{(1)}, n_i^{(2)}) = \left[\prod_{k=1}^{2} \frac{\left(\mu_{N_i^{(k)}}\right)^{n_i^{(k)}}}{n_i^{(k)}!}\right] \frac{\Gamma(n_i^{(1)} + n_i^{(2)} + \alpha)}{\Gamma(\alpha)} \beta^{-n_i^{(1)} - n_i^{(2)}} \left(1 + \frac{\mu_{N_i^{(1)}} + \mu_{N_i^{(2)}}}{\beta}\right)^{-(\alpha + n_i^{(1)} + n_i^{(2)})} \Psi_i, \tag{27}$$

where

$$\Psi_i = \frac{1}{1+c^2}\left[1 + 2c\sqrt{\alpha}(1-\eta_i) + c^2\alpha(1 - 2\eta_i + \eta_i\zeta_i)\right], \tag{28}$$

with $\eta_i = \frac{\alpha + n_i^{(1)} + n_i^{(2)}}{\alpha}\left(1 + \frac{\mu_{N_i^{(1)}} + \mu_{N_i^{(2)}}}{\beta}\right)^{-1}$ and $\zeta_i = \frac{\alpha + 1 + n_i^{(1)} + n_i^{(2)}}{\alpha}\left(1 + \frac{\mu_{N_i^{(1)}} + \mu_{N_i^{(2)}}}{\beta}\right)^{-1}$.

As, given $V_i = v_i$, $S_i^{(1)}$ and $S_i^{(2)}$ are independent compound Poisson random variables, we have that

$$\begin{aligned} Var[S_i] &= E[Var[S_i|V_i]] + Var[E[S_i|V_i]] = \\ &= \mu_{N_i^{(1)}}\left(Var\left[Y_i^{(1)}\right] + \left(\mu_{Y_i^{(1)}}\right)^2\right) + \mu_{N_i^{(2)}}\left(Var\left[Y_i^{(2)}\right] + \left(\mu_{Y_i^{(2)}}\right)^2\right) + \\ &\quad + \left(\mu_{N_i^{(1)}}\mu_{Y_i^{(1)}} + \mu_{N_i^{(2)}}\mu_{Y_i^{(2)}}\right)^2 Var[V_i], \end{aligned} \tag{29}$$

with

$$Var[V_i] = \frac{1}{1+c^2}\frac{(\alpha+1)}{\beta^2}\left[\alpha - 4c\sqrt{\alpha} + c^2(\alpha+6)\right] - 1, \tag{30}$$

which is easily expressed in terms of the parameters to be estimated, when the individual claim amounts are Gamma distributed by

$$\begin{aligned} Var[S_i|\mathbf{X}_i^{(1)}, \mathbf{X}_i^{(2)}, \mathbf{Z}_i^{(1)}, \mathbf{Z}_i^{(2)}] \quad = \quad & \left(1+\phi^{(1)}\right)\exp\left(\mathbf{X}_i^{(1)}\boldsymbol{\beta}^{(1)} + 2\mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}\right) + \\ & + \left(1+\phi^{(2)}\right)\exp\left(\mathbf{X}_i^{(2)}\boldsymbol{\beta}^{(2)} + 2\mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}\right) + \\ & + \left[\exp\left(\mathbf{X}_i^{(1)}\boldsymbol{\beta}^{(1)} + \mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}\right) + \exp\left(\mathbf{X}_i^{(2)}\boldsymbol{\beta}^{(2)} + \mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}\right)\right]^2 Var[V_i]. \end{aligned} \tag{31}$$

3.2.2 Bivariate Poisson-Laguerre polynomial regression model (BIPL)

This model, proposed by Gurmu and Elder (2007), differs from the previous one, in the sense that the dependence between, $N_i^{(1)}$ and $N_i^{(2)}$ is modelled by means of correlated unobserved heterogeneity components $V_i^{(1)}$ and $V_i^{(2)}$. Each component affects only the respective event count, but S_i will be affected by both. Let the mixing distribution be $g(v_i^{(1)}, v_i^{(2)})$, so that

$$f(n_i^{(1)}, n_i^{(2)}) = \int_0^\infty \int_0^\infty \prod_{k=1}^{2} \frac{\exp(-\mu_{N_i^{(k)}} v_i^{(k)})\,(\mu_{N_i^{(k)}} v_i^{(k)})^{n_i^{(k)}}}{n_i^{(k)}!} g(v_i^{(1)}, v_i^{(2)}) dv_i^{(1)} dv_i^{(2)}. \tag{32}$$

Denoting by $M(-\mu_{N_i^{(1)}}, -\mu_{N_i^{(2)}}) = E[\exp(-\mu_{N_i^{(1)}} v_i^{(1)} - \mu_{N_i^{(2)}} v_i^{(2)})]$ the bivariate moment generating function of $(v_i^{(1)}, v_i^{(2)})$ evaluated at $(-\mu_{N_i^{(1)}} - \mu_{N_i^{(2)}})$, (32) can be written as

$$f(n_i^{(1)}, n_i^{(2)}) = \left[\prod_{k=1}^{2} \frac{\left(\mu_{N_i^{(k)}}\right)^{n_i^{(k)}}}{n_i^{(k)}!}\right] M^{(n_i^{(1)}, n_i^{(2)})}(-\mu_{N_i^{(1)}}, -\mu_{N_i^{(2)}}), \tag{33}$$

where $M^{(n_i^{(1)}, n_i^{(2)})}(-\mu_{N_i^{(1)}}, -\mu_{N_i^{(2)}})$ denotes the derivative of $M(-\mu_{N_i^{(1)}}, -\mu_{N_i^{(2)}})$ of order $r = n_i^{(1)} + n_i^{(2)}$, i.e. $M^{(n_i^{(1)}, n_i^{(2)})}(-\mu_{N_i^{(1)}}, -\mu_{N_i^{(2)}}) = \partial^r M(-\mu_{N_i^{(1)}}, -\mu_{N_i^{(2)}})/\partial(-\mu_{N_i^{(1)}})^{n_i^{(1)}} \partial(-\mu_{N_i^{(2)}})^{n_i^{(2)}}$. The authors propose the mixing density

$$g(v_i^{(1)}, v_i^{(2)}) = \frac{w(v_i^{(1)}) w(v_i^{(2)})}{(1+\rho^2)}\left[1 + \rho P_1^{(1)}(v_i^{(1)}) P_1^{(2)}(v_i^{(2)})\right]^2 \tag{34}$$

where $w(v_i^{(k)})$ is a gamma density with parameters $(\alpha^{(k)}, \beta^{(k)})$ and

$$P_1^{(k)}(v_i^{(k)}) = \sqrt{\alpha^{(k)}} - \frac{\beta^{(k)}}{\sqrt{\alpha^{(k)}}} v_i^{(k)} \tag{35}$$

is the first order polynomial with unit variance and $\rho = \text{corr}(P_1^{(1)}(V^{(1)}), P_1^{(2)}(V^{(2)}))$ is an unknown correlation parameter. $g(v_i^{(1)}, v_i^{(2)})$ can be regarded as a variant of a bivariate gamma distribution ($\rho = 0$ leads to two independent negative binomials). The bivariate probability density function of the claim numbers are in this case

$$f(n_i^{(1)}, n_i^{(2)}) = \left[\prod_{k=1}^{2} \frac{\Gamma(n_i^{(k)} + \alpha^{(k)})}{\Gamma(\alpha^{(k)}) n_i^{(k)}!} \left(\frac{\mu_{N_i^{(k)}}}{\beta^{(k)}}\right)^{n_i^{(k)}} \left(1 + \frac{\mu_{N_i^{(k)}}}{\beta^{(k)}}\right)^{-(\alpha^{(k)} + n_i^{(k)})}\right] \Psi_i^* \tag{36}$$

with

$$\beta^{(k)} = \frac{1}{1+\rho^2}\left[\alpha^{(k)} + \rho^2(\alpha^{(k)} + 2)\right], \; k = 1, 2, \tag{37}$$

and

$$\Psi_i^* = \frac{1}{1+\rho^2}\left[1 + 2\rho\sqrt{\alpha^{(1)}\alpha^{(2)}}(1 - \eta_i^{(1)})(1 - \eta_i^{(2)}) + \rho^2\alpha^{(1)}\alpha^{(2)}(1 - 2\eta_i^{(1)} + \eta_i^{(1)}\zeta_i^{(1)})(1 - 2\eta_i^{(2)} + \eta_i^{(2)}\zeta_i^{(2)})\right] \tag{38}$$

where

$$\eta_i^{(k)} = \frac{n_i^{(k)} + \alpha^{(k)}}{\alpha^{(k)}}\left(1 + \frac{\mu_{N_i^{(k)}}}{\beta^{(k)}}\right)^{-1} \text{ and } \zeta_i^{(k)} = \frac{n_i^{(k)} + 1 + \alpha^{(k)}}{\alpha^{(k)}}\left(1 + \frac{\mu_{N_i^{(k)}}}{\beta^{(k)}}\right)^{-1}, \; k = 1, 2.$$

The p.d.f. (36) in the form (33) with

$$M^{(n_i^{(1)}, n_i^{(2)})}(-\mu_{N_i^{(1)}}, -\mu_{N_i^{(2)}}) = \left[\prod_{k=1}^{2} \frac{\Gamma(n_i^{(k)} + \alpha^{(k)})}{\Gamma(\alpha^{(k)})} \left(\beta^{(k)}\right)^{\alpha^{(k)}} \left(\beta^{(k)} + \mu_{N_i^{(k)}}\right)^{-(\alpha^{(k)} + n_i^{(k)})}\right] \Psi_i^*. \tag{39}$$

We can derive $Var[S_i]$ obtaining

$$\begin{aligned} Var[S_i] \;=\; & \mu_{N_i^{(1)}}\left(Var\left[Y_i^{(1)}\right] + \left(\mu_{Y_i^{(1)}}\right)^2\right) + \mu_{N_i^{(2)}}\left(Var\left[Y_i^{(2)}\right] + \left(\mu_{Y_i^{(2)}}\right)^2\right) + \\ & + \left(\mu_{N_i^{(1)}}\right)^2 \left(\mu_{Y_i^{(1)}}\right)^2 Var[V_i^{(1)}] + \left(\mu_{N_i^{(2)}}\right)^2 \left(\mu_{Y_i^{(2)}}\right)^2 Var[V_i^{(2)}] + \\ & + 2\mu_{N_i^{(1)}}\mu_{N_i^{(2)}}\mu_{Y_i^{(1)}}\mu_{Y_i^{(2)}}\text{Cov}(V_i^{(1)}, V_i^{(2)}) \end{aligned} \tag{40}$$

with

$$Var[V_i^{(1)}] = M^{(2,0)}(0,0) - 1 = \frac{\left(\alpha^{(1)} + 1\right)\left[\alpha^{(1)} + \rho^2(\alpha^{(1)} + 6)\right]}{\left(\beta^{(1)}\right)^2(1+\rho^2)} - 1, \tag{41}$$

$$Var[V_i^{(2)}] = M^{(0,2)}(0,0) - 1 = \frac{\left(\alpha^{(2)} + 1\right)\left[\alpha^{(2)} + \rho^2(\alpha^{(2)} + 6)\right]}{\left(\beta^{(2)}\right)^2(1+\rho^2)} - 1, \tag{42}$$

and

$$\mathrm{Cov}(V_i^{(1)}, V_i^{(2)}) = M^{(1,1)}(0,0) - 1 = \frac{\alpha^{(1)}\alpha^{(2)} + 2\rho\sqrt{\alpha^{(1)}\alpha^{(2)}} + \rho^2(\alpha^{(1)}+2)(\alpha^{(2)}+2)}{\beta^{(1)}\beta^{(2)}} - 1 \tag{43}$$

We can express (40) in terms of the regressors, when the individual claim amounts are Gamma distributed, obtaining

$$\begin{aligned} Var[S_i|\mathbf{X}_i^{(1)}, \mathbf{X}_i^{(2)}, \mathbf{Z}_i^{(1)}, \mathbf{Z}_i^{(2)}] &= \left(1+\phi^{(1)}\right)\exp\left(\mathbf{X}_i^{(1)}\boldsymbol{\beta}^{(1)} + 2\mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}\right) \\ &+ \left(1+\phi^{(2)}\right)\exp\left(\mathbf{X}_i^{(2)}\boldsymbol{\beta}^{(2)} + 2\mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}\right) \\ &+ \exp\left(2\mathbf{X}_i^{(1)}\boldsymbol{\beta}^{(1)} + 2\mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)}\right) Var[V_i^{(1)}] + \exp\left(2\mathbf{X}_i^{(2)}\boldsymbol{\beta}^{(2)} + 2\mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}\right)^2 Var[V_i^{(2)}] \\ &+ 2\exp\left(\mathbf{X}_i^{(1)}\boldsymbol{\beta}^{(1)} + \mathbf{Z}_i^{(1)}\boldsymbol{\gamma}^{(1)} + \mathbf{X}_i^{(2)}\boldsymbol{\beta}^{(2)} + \mathbf{Z}_i^{(2)}\boldsymbol{\gamma}^{(2)}\right) \mathrm{Cov}(V_i^{(1)}, V_i^{(2)}). \end{aligned} \tag{44}$$

4 Application of the models

Our database consists of the full sample of the Ambulatory Health Insurance portfolio of the Portuguese insurance company Tranquilidade. We have the data of all policies of the year 2007 and each policy can have several persons insured. The total number of persons insured (for the full year) is 19457. Our unit risk is the person insured and our database includes for each person: age (at the date of the policy renewal), policy age (at the date of the policy renewal), region, gender and information on the number of doctor visits, $N^{(1)}$, and its costs, $Y^{(1)}$, as well as the number and severity of other treatments, $N^{(2)}$ and $Y^{(2)}$, respectively. From Table 1, which gives the conditional frequency of $N^{(2)}$ given $N^{(1)}$ (for example 0.0471 is the frequency that $N^{(2)} = 1$ given that $N^{(1)} = 0$), we can conclude that these variables are strongly dependent, since the values are completely different from one row to another. The observed correlation coefficient is 0.5761. As it is expected we have a large dispersion for each of the counting variables ($\overline{x}_{N^{(1)}} = 1.900$ while $s^2_{N^{(1)}} = 6.771$ and $\overline{x}_{N^{(2)}} = 1.186$ while $s^2_{N^{(2)}} = 4.639$) which is expected to be, at least partially, kept by the Poisson regression model.

Table 1: Conditional frequency of $N^{(2)}$ given $N^{(1)}$

$N^{(1)} \backslash N^{(2)}$	0	1	2	3	4	5	6	7	>7
0	0.9131	0.0471	0.0195	0.0098	0.0055	0.0021	0.0015	0.0007	0.0007
1	0.5653	0.2653	0.0830	0.0456	0.0219	0.0102	0.0038	0.0012	0.0038
2	0.3712	0.2696	0.1608	0.0994	0.0529	0.0252	0.0098	0.0047	0.0064
3	0.2871	0.2245	0.1739	0.1316	0.0893	0.0381	0.0292	0.0131	0.0131
4	0.2379	0.1990	0.1651	0.1499	0.1050	0.0584	0.0364	0.0220	0.0262
5	0.2202	0.1619	0.1476	0.1345	0.1202	0.0774	0.0512	0.0357	0.0512
6	0.1594	0.1449	0.1576	0.1087	0.1250	0.1033	0.0851	0.0471	0.0688
7	0.1733	0.1570	0.1270	0.1109	0.0878	0.0785	0.0808	0.0647	0.1201

Table:2: Explanatory variables

Variable	Definition
X_{i1}	Equals 1 for women
X_{i2}	Equals 1 when age $\in [2,6)$
X_{i3}	Equals 1 when age $\in [6,11)$
X_{i4}	Equals 1 when age $\in [11,16)$
X_{i5}	Equals 1 when age $\in [16,21)$
X_{i6}	Equals 1 when age $\in [21,31)$
X_{i7}	Equals 1 when age $\in [31,41)$
X_{i8}	Equals 1 when age $\in [41,51)$
X_{i9}	Equals 1 when age $\in [51,61)$
X_{i10}	Equals 1 when age ≥ 61
X_{i11}	Equals 1 when policy age $\in [1,4)$
X_{i12}	Equals 1 when policy age ≥ 4
X_{i13}	Equals 1 when the region $\in$ Interior districts

We classified the data, using only dummy variables in the models and, for the sake of simplicity of presentation, we used the same explanatory variables for all the models ($\mathbf{X}_i^{(1)} = \mathbf{X}_i^{(2)} = \mathbf{Z}_i^{(1)} = \mathbf{Z}_i^{(2)} = \mathbf{X}_i$).

The explanatory variables are summarized in Table 2. It is worth mentioning that, in our study, a change on the region variables, can induce problems on the application of Kocherlakota and Kocherlakota (2001) model. The estimate for λ would be too high, violating for a couple of groups the constraint $\lambda < \min(\mu_{N_i^{(1)}}, \mu_{N_i^{(2)}})$.

The reference group (intercept) is considered to be a male, less than 2 years old, policy age less than 1 year and from one of the coastal districts of Aveiro, Braga, Coimbra, Faro, Leiria, Lisboa, Porto, Santarém and Setúbal.

All the results are based on the maximum likelihood estimates and the standard errors are calculated using the asymptotic distribution of the maximum likelihood estimators.

Table 3, shows for each model and the respective independent cases (Double Poisson -DP- for the Poisson model and PBIPL with $\rho = 0$ for the negative binomial model) the estimated parameters, the log-likelihood and the AIC. The results obtained with the KNBIP model are not presented, since they are not comparable with the others. The log-likelihood and the AIC of the KNBIP is -71085.11 and 142248.22 respectively which are similar to the values obtained by the KKBIP model.

From the Akaike's Information Criterion (AIC) or the Likelihood Criterion the worst fit is given by the Double Poisson. This model does not capture neither the heterogeneity nor the dependence of the data. Even, introducing dependence using the bivariate Poisson as in the KKBIP model the fit is poor. The independent Negative Binomial model (given by BIPL with $\rho = 0$) improves the fit, but the Generalized Bivariate Negative Binomial regression model GBINB is the best of the models, followed by the same model with $c = 0$ (bivariate negative binomial).Note that, even when $c = 0$ there is some kind of dependence in the GBINB.

We can conclude that the models assuming independence are strongly rejected using any statistical criteria as expected. Moreover the Poisson regression structure was unable to capture the heterogeneity of the data and consequently the negative binomial models and its generalizations present a much better fit. Finally it is interesting to observe that although the BIPL model has more parameters than the GBINB the latter presents a better fit.

In Table 4 we detailed the output obtained for the GBINB model, presenting the estimated coefficients as well as their standard error to illustrate the statistical significance of the results.

Table 5 shows the parameter estimates of the severity for both variables, considered independent and Gamma distributed. The significance of some of the parameters indicates that we could merge some of the policy age groups and not to discriminate according to geographical zone the severity of the Doctor's Visits. With respect to the severity of the Other Treatments a merger of the youngest age groups could be considered. As the main purpose of the paper is to discuss dependency through the claim numbers we do not pursue this analysis and keep the covariates unchanged.

Table 3: Parameter estimates for claim numbers

	DP		KKBIP		GBINB $c=0$		GBINB		BIPL $\rho=0$		BIPL	
	$k=1$	$k=2$	$k=1$	$k=2$	$k=1$	$k=2$	$k=1$	$k=2$	$k=1$	$k=2$	$k=1$	$k=2$
Intercept	1.3330	-0.7991	1.4286	-0.2797	1.3390	-0.7299	1.3153	-0.7535	1.3339	-0.7294	1.3073	-0.8030
$\beta_1^{(k)}$	0.3176	0.5397	0.2189	0.2991	0.3738	0.4988	0.3693	0.4949	0.3654	0.5036	0.2740	0.3854
$\beta_2^{(k)}$	-0.4412	0.0709	-0.4387	-0.0478	-0.4372	0.0720	-0.4122	0.0973	-0.4405	0.0756	-0.3876	0.1751
$\beta_3^{(k)}$	-1.0206	-0.2144	-0.9136	-0.0457	-1.0233	-0.2192	-0.9821	-0.1779	-1.0288	-0.2163	-0.9123	-0.0354
$\beta_4^{(k)}$	-1.3575	-0.1963	-1.1798	-0.0393	-1.3637	-0.2013	-1.3488	-0.1866	-1.3658	-0.2053	-1.2350	0.0020
$\beta_5^{(k)}$	-1.2509	0.3990	-1.2099	0.2114	-1.2961	0.3436	-1.3047	0.3358	-1.2964	0.3651	-1.1426	0.5881
$\beta_6^{(k)}$	-1.1136	0.5626	-1.1753	0.1672	-1.1607	0.4943	-1.1224	0.5327	-1.1527	0.4947	-1.0987	0.6311
$\beta_7^{(k)}$	-1.0567	0.7156	-1.1444	0.3014	-1.0983	0.6535	-1.0630	0.6891	-1.0914	0.6548	-1.0541	0.7764
$\beta_8^{(k)}$	-1.0036	0.8919	-1.1257	0.4493	-1.0366	0.8478	-1.0014	0.8831	-1.0282	0.8417	-1.0151	0.9470
$\beta_9^{(k)}$	-0.7854	1.0811	-0.9307	0.6207	-0.7935	1.0614	-0.7620	1.0931	-0.7932	1.0574	-0.8206	1.1181
$\beta_{10}^{(k)}$	-0.4065	1.3411	-0.4852	0.9663	-0.3964	1.3380	-0.3835	1.3511	-0.3992	1.3398	-0.5698	1.2151
$\beta_{11}^{(k)}$	0.1351	0.1546	0.1023	0.0438	0.1254	0.1562	0.1171	0.1470	0.1324	0.1503	0.1109	0.1328
$\beta_{12}^{(k)}$	0.2089	0.2437	0.1613	0.1018	0.1961	0.2393	0.1833	0.2259	0.2033	0.2284	0.1633	0.1894
$\beta_{13}^{(k)}$	-0.5447	-0.5563	-0.2958	-0.0728	-0.5480	-0.5937	-0.5291	-0.5746	-0.5437	-0.5706	-0.3986	-0.4165
Other Parameters			$\widehat{\lambda_3}$ =0.6639		$\widehat{\alpha}$ =0.5908		$\widehat{\alpha}$ =0.5673 $\widehat{c}$ =0.2157		$\widehat{\alpha}_1$ =0.7872 $\widehat{\alpha}_2$ =0.4634		$\widehat{\alpha}_1 = 0.8385$ $\widehat{\alpha}_2 = 0.3944$ $\widehat{\rho}$ =0.697422413	
Log-Likelihood	-78202.92		-71381.13		-57726.35		-57697.79		-62531.53		-58441.28	
AIC	156348.84		142819.26		115509.70		115453.58		125120.06		116939.56	

Table 4: Parameter estimates and standard errors for GBINB

	$N^{(1)}$		$N^{(2)}$	
	Coef.	St. Errors	Coef.	St. Errors
Intercept	1.3153	0.0543	-0.7535	0.0713
$\beta_1^{(k)}$	0.3693	0.0215	0.4949	0.0231
$\beta_2^{(k)}$	-0.4122	0.0629	0.0973	0.0819
$\beta_3^{(k)}$	-0.9821	0.0650	-0.1779	0.0845
$\beta_4^{(k)}$	-1.3488	0.0745	-0.1866	0.0934
$\beta_5^{(k)}$	-1.3047	0.0736	0.3358	0.0886
$\beta_6^{(k)}$	-1.1224	0.0577	0.5327	0.0740
$\beta_7^{(k)}$	-1.0630	0.0578	0.6891	0.0740
$\beta_8^{(k)}$	-1.0014	0.0606	0.8831	0.0761
$\beta_9^{(k)}$	-0.7620	0.0653	1.0931	0.0801
$\beta_{10}^{(k)}$	-0.3835	0.0974	1.3511	0.1091
$\beta_{11}^{(k)}$	0.1171	0.0279	0.1470	0.0303
$\beta_{12}^{(k)}$	0.1833	0.0324	0.2259	0.0347
$\beta_{13}^{(k)}$	-0.5291	0.0320	-0.5746	0.0349

We have calculated the premium associated to S using both the expected value and the standard deviation principle for all the models. To compare the results obtained by the different models we started by setting the total premium of the portfolio to a value, such that the loading is 25% of the expected aggregate claim amount, for the Double Poisson. Then the loading coefficient for each of the premiums/models was calculated, to get the set premium.

Table 5: Parameter estimates for claim severity

	$Y^{(1)}$		$Y^{(2)}$	
	Coef.	St. Errors	Coef.	St. Errors
Intercept	3.8103	0.0072	3.8841	0.0522
$\gamma_1^{(k)}$	0.0174	0.0038	-0.0347	0.0141
$\gamma_2^{(k)}$	-0.0286	0.0085	-0.0374	0.0594
$\gamma_3^{(k)}$	-0.0485	0.0097	-0.0365	0.0622
$\gamma_4^{(k)}$	-0.0932	0.0126	0.1413	0.0682
$\gamma_5^{(k)}$	-0.1374	0.0125	0.2572	0.0622
$\gamma_6^{(k)}$	-0.1260	0.0081	0.3752	0.0533
$\gamma_7^{(k)}$	-0.1052	0.0080	0.4106	0.0530
$\gamma_8^{(k)}$	-0.1175	0.0086	0.4802	0.0536
$\gamma_9^{(k)}$	-0.1321	0.0093	0.5523	0.0548
$\gamma_{10}^{(k)}$	-0.0998	0.0137	0.5212	0.0640
$\gamma_{11}^{(k)}$	-0.0040	0.0050	-0.0313	0.0187
$\gamma_{12}^{(k)}$	0.0347	0.0058	-0.0538	0.0207
$\gamma_{13}^{(k)}$	-0.0058	0.0067	0.0480	0.0237
Other Parameters	$\phi^{(1)} = 0.1331$		$\phi^{(2)} = 1.0820$	
Log-Likelihood	-148303		-120877	
AIC	296633		241781	

Although, apparently similar, the two dependent Poisson models, KKBIP and KNBIP, can lead to different results, as can be seen on Figure 1, where the box-plot of the ratio between the premiums is presented, for both the expected value principle and the standard deviation principle.

The premiums calculated according to the standard deviation principle tend to originate a narrower band of premiums than the ones obtained with the expected value principle for our data set. The exception is the results obtained for the BIPL model. Figure 2 shows for each model the ratio between the standard deviation and the expected value premium principles as a function of the latter premium.

Figure 1: KNBIB versus KKBIP models

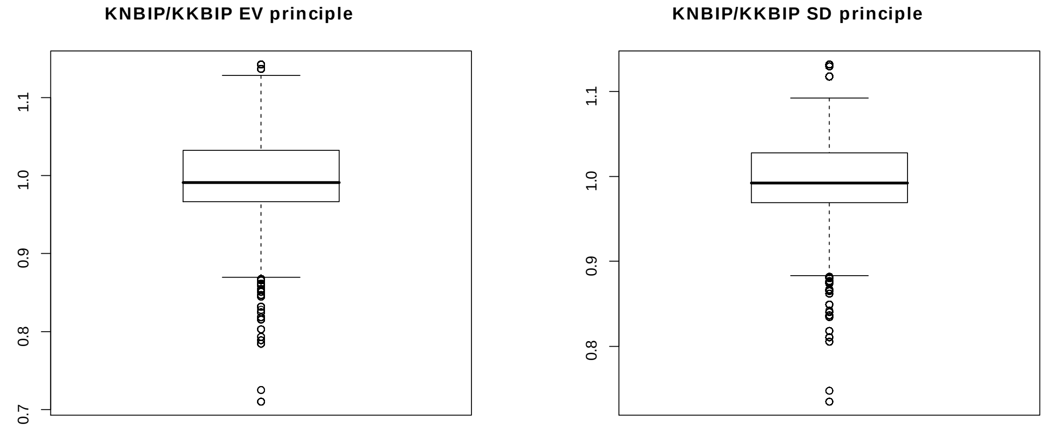
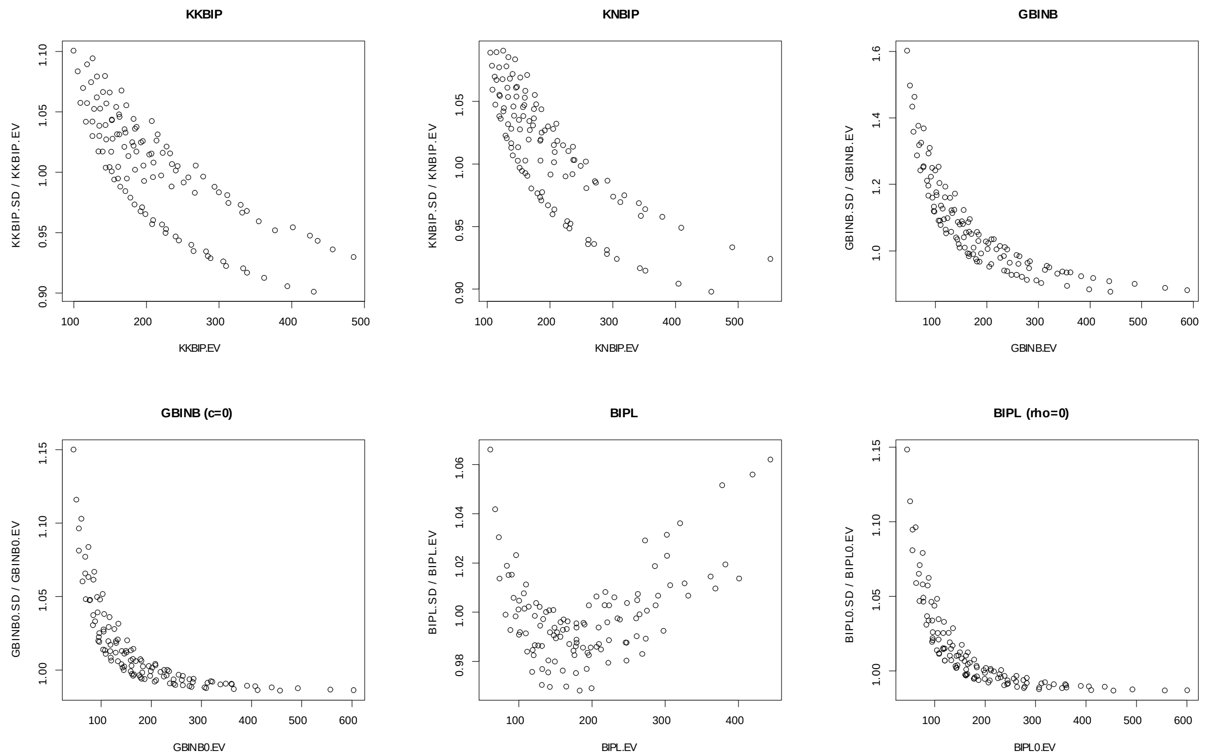


Figure 2: Standard Deviation Principle *versus* Expected Value Principle



To have a closer look at the impact of the different models, at the individual level, we defined five profiles corresponding to the 5, 25, 50, 75 and 95 percentiles of the rank of the premiums obtained for the double Poisson using the expected value principle. Profile 1 is a male, between 31 and 40 years old, new policy and from the coast. Profile 2 is a male, between 21 and 30 years old, policy age between 1 and 3 years and from the interior. Profile 3 differs from profile 2 only on the age, which is between 2 and 5 years old. The same happens for Profile 4 which age is between 51 and 60 years old. Finally profile 5 is a Female, between 51 and 60 years old, new policy and from the interior. Tables 6 and 7 show the premiums for these profiles, when the different models are used, and when using the expected value and the standard deviation principles, respectively. Information about the highest and the lowest of the premiums for each model is also given.

Table 6: Premiums for the different models, according to the expected value principle

Prof.	N.risks	DP	KKBIP	KNBIP	GBINB $c=0$	GBINB	BIPL $\rho=0$	BIPL
1	82	89.34	142.76	133.73	86.24	89.13	87.82	109.88
2	1100	149.65	152.33	157.25	146.16	147.36	146.63	157.04
3	564	184.57	207.76	204.10	186.77	185.87	186.73	195.70
4	372	255.87	242.67	229.30	262.72	263.06	261.71	260.29
5	90	351.54	299.62	291.45	359.76	361.94	358.87	326.45
Highest Premium		582.73	485.13	551.45	603.52	588.19	600.74	443.64
Lowest Premium		44.5	99.45	104.76	44.77	45.32	45.06	61.12
Ratio		13.09	4.88	5.26	13.48	12.98	13.33	7.26

Table 7: Premiums for the different models according to the standard deviation principle

Prof.	N.risks	DP	KKBIP	KNBIP	GBINB $c=0$	GBINB	BIPL $\rho=0$	BIPL
1	82	101.12	154.14	145.15	92.01	116.74	93.30	111.12
2	1100	155.56	158.88	163.27	147.79	159.09	148.17	156.41
3	564	178.32	198.84	195.89	185.63	179.79	185.59	192.29
4	372	257.48	243.91	231.66	261.68	258.82	260.86	259.64
5	90	344.38	294.63	287.55	356.39	338.46	355.67	330.27
Highest Premium		542.91	451.10	509.59	596.95	548.13	594.69	473.14
Lowest Premium		52.59	109.46	114.08	46.58	53.37	46.82	59.44
Ratio		10.32	4.12	4.47	12.81	10.27	12.70	7.96

Although the total amount of premiums is the same whatever the model and the principle used, we can see that, at individual level, things are different. When the expected value principle is used we obtain two main groups: the first one is composed by the double Poisson and the approaches based on the negative binomial distribution except the BIPL. The second one includes the approaches based on the (correlated) bivariate Poisson (KKBIP and KNBIP). The BIPL behaves somewhere in the middle.

The ratios between the highest premium over the lowest one confirms that the first group leads to a much larger variability among risks (ratio around 13 versus ratio around 5) when compared to the second one. As expected, the BIPL presents a ratio of 7.3 between the ratios of the two groups.

When the standard deviation principle is used results are quite similar: the same groups appear and individual premiums are in line with those obtained using the expected value principle. For our data the standard deviation principle originates lower ratios between the extreme premiums which can be a good point for the definition of the tariff.

5 Conclusions

For this data set, we can conclude that taking dependence into account matters. Moreover the Poisson regression structure was unable to capture the heterogeneity of the data and consequently the negative binomial models, and their generalizations, present a much better fit. The standard deviation principle leads to narrower scales than the expected value principle.

Based on the comments of the previous section and on the fitness of the models, if we had to choose a model, we would decide by the BIPL model.

Acknowledgement: The authors gratefully acknowledge the constructive comments made by two anonymous referees, which helped to improve the quality of the paper.

References

Bermudez, L. (2009) A Priori Ratemaking Using Bivariate Poisson Regression Models, *Insurance: Mathematics & Economics*, **44,** pp. 135-141**.** doi:10.1016/j.insmatheco.2008.11.005

Brockman, M.J. and Wright, M.A., (1992), Statistical Motor Rating: Making Effective Use of Your Data Set, *Journal of the Institute of Actuaries*, **119**/III, pp. 457-543.

Centeno, M.L., (2005), Dependent Risks and Excess of Loss Reinsurance, *Insurance: Mathematics & Economics*, **37**, pp. 229-238.

doi:10.1016/j.insmatheco.2004.12.001

Cossete, H. and Marceau, E., (2000), The Discrete-time Risk Model with Correlated Classes of Businesses, *Insurance: Mathematics & Economics* , **26**, pp. 133-149.

Denuit, M., Maréchal, X., Pitrebois, S. and Walhin, J.-F., (2006), *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Scales,* Wiley, New York.

Gurmu, S. and Elder, J., (2000), Generalised bivariate count data regression models. *Economics Letters,* pp. 31-36.

Gurmu, S. and Elder, J., (2007), A simple bivariate count data regression model. *Economics Bulletin,* **3***,* pp. 1-10.

Holgate, P., (1964), Estimation for Biavariate Poisson Distributions, *Biometrika*, **51**, pp. 241-245.

Johnson, L., Kotz, S. and Balakrishnan, N., (1997), *Discrete Multivariate Distributions*, Wiley.

Karlis, D. and Ntzoufras, I., (2005), Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in R. *Journal of Statistical Software,* **14**/10, pp. 1-36.

Kocherlakota, S and Kocherlakota, K. (2001) Regression in the Bivariate Poisson Distribution. *Communications in Statistics - Theory and Methods,* **30**/5, pp. 815-825**.**

McCullagh, P. and Nelder, J., (1989) Generalized Linear Models, Second Edition, Chapman & Hall.

Nelder, J.A. and Wedderburn, R.W, (1972), Generalized Linear Models, *Journal of the Royal Statistical Society,* **A**, 135, pp.370-384.

Whalin, J.F. (2001) Bivariate ZIP Models. *Biometrical Journal,* **43**/2**,** pp. 147-160.

Wang, S., (1998), *Aggregation of Correlated Risk Portfolios: Mothels and Algorithms*, pp. 848-939, Proceedings of the Casualty Actuarial Society, Vol. LXXXV.