

KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0332

CLAIMS RESERVING USING GENERALIZED LINEAR MODELS

by

T. HOEDEMAKERS
J. BEIRLANT
M. GOOVAERTS
J. DHAENE

Claims Reserving Using Generalized Linear Models

Tom Hoedemakers*† Jan Beirlant‡ Marc J. Goovaerts*§ Jan Dhaene*§

Abstract

Renshaw and Verrall (1994) specified the generalized linear model (GLM) underlying the chain-ladder technique and suggested some other GLMs which might be useful in claims reserving. The purpose of this paper is to construct bounds for the discounted loss reserve within the framework of GLMs. Exact calculation of the distribution of the total reserve is not feasible, and hence the determination of lower and upper bounds with a simpler structure is a possible way out. The paper ends with numerical examples illustrating the usefulness of the presented approximations.

Keywords: IBNR, comonotonicity, simulation, generalized linear model, prediction.

*Department of Applied Economics, K.U.Leuven, Naamsestraat 69, 3000 Leuven, Belgium
†Corresponding author. E-mail address: Tom.Hoedemakers@econ.kuleuven.ac.be
‡University Centre for Statistics, K.U.Leuven, W. de Croylaan 54, 3001 Heverlee, Belgium
§Institute of Actuarial Science, University of Amsterdam, Roeterstraat 11, 1018 WB Amsterdam, Netherlands

1 Introduction

The correct estimation of the amount of money a company should set aside now to meet claims arising in the future on the written policies represents an important task for insurance companies (to get the correct picture of its liabilities). The past data used to construct estimates for the future payments consist of a triangle of incremental claims Y_{ij}. We use the standard notation, with the random variables Y_{ij} for $i = 1, 2, \ldots, t; j = 1, 2, \ldots, s$ denoting the claim figures for year of origin (or accident year) i and development year j, meaning that the claims were paid in calendar year $i + j - 1$. Year of origin, year of development and calendar year act as possible explanatory variables for the observation Y_{ij}. Most claims reserving methods usually assume that $t = s$. For (i, j) combinations with $i + j \leq t + 1$, Y_{ij} has already been observed, otherwise it is a future observation. To a large extent, it is irrelevant whether incremental or cumulative data are used when considering claims reserving in a stochastic context. The known values are presented in the form of a run-off triangle, as depicted in Figure 1.

We consider annual development in this paper (the methods can be extended easily to semi-annual, quarterly or monthly development) and we assume that the time it takes for the claims to be completely paid is fixed and known. The triangle is augmented each year by the addition of a new diagonal. The purpose is to complete this run-off triangle to a square, and even to a rectangle if estimates are required pertaining to development years of which no data are recorded in the run-off triangle at hand. To aid in the setting of reserves, the actuary can make use of a variety of techniques. The inherent uncertainty is described by the distribution of possible outcomes, and one needs to arrive at the best estimate of the reserve.

In this paper our aim is to model claim payments using Generalized Linear Models (GLMs) and to incorporate a stochastic discounting factor at the same time when estimating loss reserves. Distributions used to describe the claim size should have a thick right tail. Furthermore, the phenomena to be modelled are rarely additive in the collateral data. A multiplicative model is much more plausible. These problems cannot be solved by working with ordinary linear models, but with generalized linear models. The generalization is twofold. First, it is allowed that the random deviations from the mean obey another distribution than the normal. In fact, one can take any distribution from the exponential dispersion family, including for instance the Poisson, the binomial, the gamma and the inverse Gaussian distributions. Second, it is no longer necessary that the mean of the random variable is a linear function of the explanatory variables, but it only has to be linear on a certain scale. If this scale for instance is logarithmic, we have in fact a multiplicative model instead of an additive model.

Loss reserving deals with the determination of the uncertain present value of an unknown amount of future payments. One of the sub-problems in this respect consists of the discounting of the future estimates in the run-off triangle, where interest rates (and inflation)

Year of	*Development year*						
origin	1	2	$\cdots$	j	$\cdots$	$t-1$	t
1	Y_{11}	Y_{12}	$\cdots$	Y_{1j}	$\cdots$	$Y_{1,t-1}$	Y_{1t}
2	Y_{21}	Y_{22}	$\cdots$	Y_{2j}	$\cdots$	$Y_{2,t-1}$	
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		
i	Y_{i1}	$\cdots$	$\cdots$	Y_{ij}			
$\vdots$	$\cdots$	$\cdots$	$\cdots$				
t	Y_{t1}						

Figure 1: Random variables in a run-off triangle

are not known for certain. We will model the stochastic discount factor using a Brownian motion with drift. When determining the discounted loss reserve, we impose an explicit margin based on a risk measure (for example Value at Risk) from the total distribution of the discounted reserve. In general, it is hard or even impossible to determine the quantiles of S analytically, because in any realistic model for the return process the random variable S will be a sum of strongly dependent random variables. In the present setting we suggest to solve this problem by calculating upper and lower bounds for this sum of dependent random variables making efficient use of the available information. These bounds are based on a general technique for deriving lower and upper bounds for stop-loss premiums of sums of dependent random variables, as explained in Kaas et al. (2000). The first approximation we will consider for the distribution function of the discounted IBNR reserve is derived by approximating the dependence structure between the random variables involved by a comonotonic dependence structure. The second approximation, which is derived by considering conditional expectations, takes part of the dependence structure into account. We will include a numerical comparison of our approximations with a simulation study. The second approximation turns out to perform quite well. For details of this technique we refer to Dhaene et al. (2002a,b) and the references therein.

This paper is set out as follows. In section 2 we present a brief review of generalized linear models and their applications to claims reserving. In section 3 we explain the methodology for obtaining the bounds and we recall the main result concerning stochastic bounds for the scalar product of two independent random vectors, where the marginal distribution functions of each vector are given, but the dependence structures are unknown. To use these results for discounted IBNR evaluations we need some asymptotic results for model parameter estimates in GLMs. Some numerical illustrations for a simulated data set are provided in section 4. We also graphically illustrate the obtained bounds.

2 Generalized Linear Models and Claims Reserving

For a general introduction to generalized linear models we refer to McCullagh and Nelder (1992). This family encompasses normal error linear regression models and the nonlinear exponential, logistic and Poisson regression models, as well as many other models, such as loglinear models for categorical data. In this subsection we recall the structure of GLMs in the framework of claims reserving.

The first component of a GLM, the random component, assumes that the response variables Y_{ij} are independent and that the density function of Y_{ij} belongs to the exponential family with densities of the form

$$f(y_{ij};\theta_{ij},\phi) = \exp\left\{\left[y_{ij}\theta_{ij} - b(\theta_{ij})\right]/a(\phi) + c(y_{ij},\phi)\right\}, \tag{1}$$

where $a(.)$, $b(.)$ en $c(.,.)$ are known functions. The function $a(\phi)$ often has the form $a(\phi) = \phi$, where ϕ is called the dispersion parameter.

When ϕ is a known constant, (1) simplifies to the natural exponential family

$$f(y_{ij};\theta_{ij}) = \tilde{a}(\theta_{ij})\tilde{b}(y_{ij})\exp\left\{y_{ij}Q(\theta_{ij})\right\}. \tag{2}$$

We identify $Q(\theta)$ with $\theta/a(\phi)$, $\tilde{a}(\theta)$ with $\exp\{-b(\theta)/a(\phi)\}$, and $\tilde{b}(y)$ with $\exp\{c(y,\phi)\}$. The more general formula (1) is useful for two-parameter families, such as the normal or gamma, in which ϕ is a nuisance parameter. Let us denote the mean of Y_{ij} by μ_{ij}, then it is known that

$$\mu_{ij} = E[Y_{ij}] = b'(\theta_{ij}) \text{ and } \mathrm{Var}[Y_{ij}] = b''(\theta_{ij})a(\phi), \tag{3}$$

where the primes denote derivatives with respect to θ. The variance expressed as a function of the mean is given by

$$\mathrm{Var}[Y_{ij}] = a(\phi)V(\mu_{ij}) = \phi V(\mu_{ij}),$$

where $V(.)$ is called the variance function. The variance function captures the relationship, if any, between the mean and variance of Y_{ij}.

The possible distributions to work with in claims reserving include the normal, the Poisson, the gamma and the inverse Gaussian. Table 1 shows some of their characteristics. For a given distribution, link functions other than the natural link function can also be used. For example, the log-link is often used with the gamma distribution.

The systematic component of a GLM is based on a linear predictor

$$\eta_{ij} = (\mathbf{R}\vec{\beta})_{ij} = \beta_1 R_{ij,1} + \cdots + \beta_p R_{ij,p}, \qquad i,j = 1,\cdots,t, \tag{4}$$

Distribution	Density	ϕ	Canonical link $\theta(\mu)$	Mean function $\mu(\theta)$	Variance function $V(\mu)$
N(μ, σ^2)	$\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$	σ^2	μ	θ	1
Poisson(μ)	$e^{-\mu}\frac{\mu^y}{y!}$	1	$\log(\mu)$	e^θ	μ
Gamma(μ, ν)	$\frac{1}{\Gamma(\nu)}\left(\frac{\nu y}{\mu}\right)^\nu \exp\left(-\frac{\nu y}{\mu}\right)\frac{1}{y}$	$\frac{1}{\nu}$	$1/\mu$	$-1/\theta$	μ^2
IG(μ, σ^2)	$\frac{y^{-3/2}}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(y-\mu)^2}{2y\sigma^2\mu^2}\right)$	σ^2	$1/\mu^2$	$(-2\theta)^{-1/2}$	μ^3

Table 1: Characteristics of some frequently used distributions in loss reserving

where $\vec{\beta} = (\beta_1, \cdots, \beta_p)'$ are model parameters, and $\mathbf{R}$ is the regression (model) matrix of dimension $t^2 \times p$. Various choices are possible for this linear predictor. We give here a short overview of frequently used parametric structures.

A well-known and widely used linear predictor is the chain-ladder type, given by

$$\eta_{ij} = \alpha_i + \beta_j, \tag{5}$$

with α_i the parameter for each year of origin i and β_j for each development year j. Note that a parameter, for example β_1, must be set equal to zero, in order to have a non-singular regression matrix. Fitting a parametric curve, such as the Hoerl curve (Zehnwirth, 1985), to the run-off triangle reduces the number of parameters used to describe trends in the development years. For the lognormal model or models with a log-link function, the Hoerl or gamma curve is provided by replacing β_j in (5) with $\beta_i \log(j) + \gamma_i j$. Wright (1990) extends this Hoerl curve with a term to model possible claim inflation.

The separation predictor takes into account the calendar years and replaces in (5) α_i with γ_k $(k = i+j-1)$. It combines the effects of monetary inflation and changing jurisprudence.

For a general model with parameters in the three directions, we refer to De Vylder and Goovaerts (1979). Barnett and Zehnwirth (1998) publish a paper in which they describe the probabilistic trend family (PTF) of models with the following linear predictor

$$\alpha_i + \sum_{k=1}^{j-1} \beta_k + \sum_{t=1}^{i+j-2} \gamma_t.$$

The link function, the third component of a GLM, connects the expectation μ_{ij} of Y_{ij} to the linear predictor by

$$\eta_{ij} = g(\mu_{ij}), \tag{6}$$

where g is a monotone, differentiable function. Thus, a GLM links the expected value of the response to the explanatory variables through the equation

$$g(\mu_{ij}) = (\mathbf{R}\vec{\beta})_{ij} \qquad i, j = 1, \cdots, t. \tag{7}$$

For the canonical link g for which $g(\mu_{ij}) = \theta_{ij}$ in (1), there is the direct relationship between the natural parameter and the linear predictor. Since $\mu_{ij} = b'(\theta_{ij})$, the canonical link is the inverse function of b'.

Generalized linear models may have nonconstant variances σ_{ij}^2 for the responses Y_{ij}, but the variance σ_{ij}^2 must be a function of the predictor variables through the mean response μ_{ij}. Any regression model that belongs to the family of generalized linear models can be analyzed in a unified fashion. The maximum likelihood estimates of the regression parameters can be obtained by iteratively reweighted least squares (naturally extending ordinary least squares for normal error linear regression models).

Suppose that the claim amounts follow a lognormal distribution, then taking the logarithm of all Y_{ij}'s implies that they have a normal distribution. So, the link function is given by $\eta_{ij} = \mu_{ij}$ and the scale parameter is the variance of the normal distribution, i.e. $\phi = \sigma^2$. Remark that each incremental claim must be greater than zero and predictions from this model can yield unusable results.

The need for more general GLM models for modelling claims reserves becomes clear in the column of variance functions in Table 1. If the variance of the claims is proportional to the square of the mean, the gamma family of distributions can accommodate this characteristic. The Poisson and inverse Gaussian provide alternative variance functions. However, it may be that the relationship between the mean and the variance falls somewhere between the inverse Gaussian and the gamma models. Quasi-likelihood is designed to handle this broader class of mean-variance relationships. This is a very simple and robust alternative, introduced in Wedderburn (1974), which uses only the most elementary information about the response variable, namely the variance-mean relationship. It is extraordinary that this information alone is often sufficient to stay close to the full efficiency of maximum likelihood estimators. Suppose that we know that the response is always positive, the data are invariably skew to the right, and the variance increases with the mean. This does not enable to specify a particular distribution (for example it does not discriminate between Poisson or negative binomial errors), hence one cannot use techniques like maximum likelihood or likelihood ratio tests. However, quasi-likelihood estimation allows one to model the response variable in a regression context without specifying its distribution.

To avoid consistent over-estimation of ultimate claims in early development years, we assume that the sum of the incremental claims in each column is greater than zero (see Renshaw and Verrall (1994)). When this assumption is not satisfied, the authors suggest to add a small constant to the past data and to substract this constant from the forecasts at the end of the reserving process.

We recall that the only distributional assumptions used in GLMs are the functional relationship between variance and mean and the fact that the distribution belongs to the exponential family. When we consider the Poisson case, this relationship can be expressed as

$$\text{Var}[Y_{ij}] = E[Y_{ij}]. \tag{8}$$

One can allow for more dispersion in the data by generalizing (8) to $\text{Var}[Y_{ij}]=\phi E[Y_{ij}]$ without any change in the form and solution of the likelihood equations. For example, it is well known that an over-dispersed Poisson model with the chain-ladder type linear predictor (5) gives the same predictions as those obtained by the deterministic chain-ladder method (see Renshaw and Verrall, 1994).

Modelling the incremental claim amounts as independent gamma response variables, with a logarithmic link function and the chain-ladder type linear predictor (5) produces exactly the same results as obtained by Mack (1991). The relationship between this generalized linear model and the model proposed by Mack was first pointed out by Renshaw and Verrall, 1994). The mean-variance relationship for the gamma model is given by

$$\text{Var}[Y_{ij}] = \phi \left(E[Y_{ij}]\right)^2. \tag{9}$$

Using this model gives predictions close to those from the deterministic chain-ladder technique, but not exactly the same. Remark that we need to impose that each incremental value should be positive (non-negative) if we work with gamma (Poisson) models. This restriction can be overcome using a quasi-likelihood approach.

As in normal regression, the search for a suitable model may encompass a wide range of possibilities. The Bayesian information criterion (BIC) and the Akaike Information Criterion (AIC) are model selection devices that emphasize parsimony by penalizing models for having large numbers of parameters. Tests for model development to determine whether some predictor variables may be dropped from the model can be conducted using partial deviances. Two measures for the goodness-of-fit of a given generalized linear model are the scaled deviance and Pearson's chi-square statistic.

In cases where the dispersion parameter is not known, an estimate can be used to obtain an approximation to the scaled deviance and Pearson's chi-square statistic. One strategy is to fit a model that contains a sufficient number of parameters so that all systematic variation is removed, estimate ϕ from this model, and then use this estimate in computing the scaled deviance of sub-models. The deviance or Pearson's chi-square divided by its degrees of freedom is sometimes used as an estimate of the dispersion parameter ϕ.

3 Application to estimation of discounted IBNR reserves

In claims reserving, we are interested in the aggregated value $\sum_{i=2}^{t}\sum_{j=t+2-i}^{t} Y_{ij}$. The predicted value will be given by

$$\text{IBNR reserve} = \sum_{i=2}^{t}\sum_{j=t+2-i}^{t} \hat{\mu}_{ij}, \tag{10}$$

with $\hat{\vec{\mu}}_{ij} = g^{-1}\left((\mathbf{R}\hat{\vec{\beta}})_{ij}\right)$ for a given link function g.

In the case that the type of business allows for discounting, or in the case that the value of the reserve itself is seen as a risk in the framework of financial reinsurance, we add a discounting process. Of course, the level of the required reserve will strongly depend on how we will invest this reserve. Let us assume that the reserve will be invested such that it generates a stochastic return Y_j in year j, $j = 1, 2, \cdots, t-1$, i.e. an amount of 1 at time $j-1$ will become e^{Y_j} at time j. The discount factor for a payment of 1 at time i is then given by $e^{-(Y_1+Y_2+\cdots+Y_i)}$, because this stochastic amount will exactly grow to an amount 1 at time i. We will assume that the return vector $(Y_1, Y_2, \cdots, Y_{t-1})$ has a multivariate normal distribution, which is independent of the Y_{ij}'s. The present value of the payments is then a linear combination of dependent lognormal random variables. We introduce the random variable $Y(i)$ defined by

$$Y(i) = Y_1 + Y_2 + \cdots + Y_i \tag{11}$$

and assume that

$$Y(i) = (\delta + \frac{\sigma^2}{2})i + \sigma B(i), \tag{12}$$

where $B(i)$ is the standard Brownian motion and where δ is a constant force of interest. In order to obtain a net present value, that is consistent with pricing in the financial environment, we transform the total estimated IBNR-reserve as follows

$$S \overset{def}{=} \sum_{i=2}^{t}\sum_{j=t+2-i}^{t} \hat{\mu}_{ij} e^{-Y(i+j-t-1)} \tag{13}$$

$$= \sum_{i=2}^{t}\sum_{j=t+2-i}^{t} \hat{\mu}_{ij} \exp\left(-(\delta + \sigma^2/2)(i+j-t-1) - \sigma B(i+j-t-1)\right). \tag{14}$$

With this adaptation, we have that

$$E[e^{-Y(i)}] \,.\, e^{\delta i} = 1. \tag{15}$$

In order to study the distribution of the discounted IBNR reserve (13), we will use recent results concerning bounds for sums of stochastic variables. In the following section, we will explain the methodology we used for finding the desired answers. We will briefly repeat the most important results.

4 Methodology and Asymptotic Results

Because the discounted IBNR reserve is a sum of dependent random variables, its distribution function cannot be determined analytically. Therefore, instead of calculating the exact distribution, we will look for bounds, in the sense of "more favourable/less dangerous" and "less favourable/more dangerous", with a simpler structure. This technique is common practice in the actuarial literature. When lower and upper bounds are close to each other, together they can provide reliable information about the original and more complex variable. The notion "less favourable" or "more dangerous" variable will be defined by means of the convex order.

4.1 Convex order and comonotonicity

Definition 1 *A random variable V is smaller than a random variable W in convex order if*

$$E[u(V)] \leq E[u(W)], \tag{16}$$

for all convex functions u: $\mathbb{R} \to \mathbb{R} : x \mapsto u(x)$, provided the expectations exist. This is denoted as

$$V \leq_{cx} W. \tag{17}$$

Roughly speaking, convex functions are functions that take on their largest values in the tails. Therefore, $V \leq_{cx} W$ means that W is more likely to take on extreme values than V. In terms of utility theory, $V \leq_{cx} W$ means that the loss V is preferred to the loss W by all risk averse decision makers, i.e. $E[u(-V)] \geq E[u(-W)]$ for all concave utility functions u. This means that replacing the (unknown) distribution function of V by the distribution function of W, can be considered as a prudent strategy with respect to setting reserves.

It follows that $V \leq_{cx} W$ implies $E[V] = E[W]$ and $\text{Var}[V] \leq \text{Var}[W]$, see for example Dhaene et al. (2002a). The next theorem, given in Hoedemakers et al. (2003), extends the results of Dhaene et al. (2002a) and Kaas et al. (2002) for ordinary sums of variables to sums of scalar products of independent random variables.

Theorem 1 *Assume that the vectors $\mathbf{X} = (X_1, \cdots, X_n)$ and $\mathbf{Y} = (Y_1, \cdots, Y_n)$, given the random variable Z, are mutually independent and that Z is independent of $\mathbf{Y}$. Consider two mutually independent uniform(0,1) random variables U and V. If the X_i and Y_i are non-negative random variables, then we find that the following relations hold:*

$$S_l \leq_{cx} S \leq_{cx} S'_u \leq_{cx} S_u, \tag{18}$$

with

$$\begin{aligned}
S &= X_1Y_1 + X_2Y_2 + \ldots + X_nY_n && (19)\\
S_l &= E[X_1|Z]E[Y_1] + E[X_2|Z]E[Y_2] + \ldots + E[X_n|Z]E[Y_n] && (20)\\
S'_u &= F^{-1}_{X_1|Z}(U)F^{-1}_{Y_1}(V) + F^{-1}_{X_2|Z}(U)F^{-1}_{Y_2}(V) + \ldots + F^{-1}_{X_n|Z}(U)F^{-1}_{Y_n}(V) && (21)\\
S_u &= F^{-1}_{X_1}(U)F^{-1}_{Y_1}(V) + F^{-1}_{X_2}(U)F^{-1}_{Y_2}(V) + \ldots + F^{-1}_{X_n}(U)F^{-1}_{Y_n}(V), && (22)
\end{aligned}$$

and where U, V and Z are mutually independent.

We will apply the results of previous theorem to the discounted IBNR reserve as formulated in (13). Before starting with this, we have to say something about the distribution of $\hat{\vec{\mu}}$. This is done in what follows.

4.2 The distribution of $\hat{\vec{\mu}}$

Let $\hat{\phi}, \hat{\vec{\beta}}$, $\hat{\vec{\eta}} = \mathbf{R}\hat{\vec{\beta}}$ and $\hat{\vec{\mu}} = g^{-1}(\hat{\vec{\eta}})$ be the maximum likelihood estimates of $\phi, \vec{\beta}, \vec{\eta}$ and $\vec{\mu}$ respectively. Denote the regression matrix corresponding to the upper triangle by $\mathbf{U}$. The estimation equations for $\hat{\vec{\beta}}$ are given by

$$\mathbf{U}'\hat{\mathbf{W}}\mathbf{U}\hat{\vec{\beta}} = \mathbf{U}'\hat{\mathbf{W}}\hat{y}^{\star}, \tag{23}$$

where $\mathbf{W} = \text{diag}\{w_{11}, \cdots, w_{t1}\}$, with $w_{ij} = \text{Var}[Y_{ij}]^{-1}(d\mu_{ij}/d\eta_{ij})^2$, and $y^{\star} = (y^{\star}_{11}, \cdots, y^{\star}_{t1})'$, with $y^{\star}_{ij} = \eta_{ij} + (y_{ij} - \mu_{ij})d\eta_{ij}/d\mu_{ij}$ and y_{ij} denote the sample values. Note that $\hat{\mathbf{W}}$ is $\mathbf{W}$ evaluated at $\hat{\vec{\beta}}$.

It is well-known that for a statistic with a large-sample normal distribution, many functions of that statistic are also asymptotically normal. Because $\mathbf{R}\hat{\vec{\beta}} = \left((\mathbf{R}\hat{\vec{\beta}})_{11}, \cdots, (\mathbf{R}\hat{\vec{\beta}})_{tt}\right)'$ is asymptotically multivariate normal with mean $\mathbf{R}\vec{\beta} = \left((\mathbf{R}\vec{\beta})_{11}, \cdots, (\mathbf{R}\hat{\vec{\beta}})_{tt}\right)'$ and variance-covariance matrix $\mathbf{Z} = \{z_{ij}\} = \mathbf{R}(\mathbf{U}'\mathbf{W}\mathbf{U})^{-1}\mathbf{R}'$ and the function $g^{-1}(\eta_{11}, \cdots, \eta_{tt})$ has a nonzero differential $\vec{\psi} = (\psi_{11}, \cdots, \psi_{tt})'$ at $(\mathbf{R}\vec{\beta})$, where $\psi_{ij} = d\mu_{ij}/d\eta_{ij}$, it follows from the delta method that

$$\left[\hat{\vec{\mu}} - \vec{\mu}\right] \xrightarrow{d} N\left(0, \mathbf{\Sigma}(\hat{\vec{\mu}})\right), \tag{24}$$

where $\mathbf{\Sigma}(\hat{\vec{\mu}}) = \vec{\psi}'\mathbf{Z}\vec{\psi}$. Hence, for large samples the distribution of $\hat{\vec{\mu}} = g^{-1}(\mathbf{R}\hat{\vec{\beta}})$ can be approximated by a normal distribution with mean $\vec{\mu}$ and variance-covariance matrix $\mathbf{\Sigma}(\hat{\vec{\mu}})$.

Maximum likelihood estimates may be biased when the sample size or the total Fisher information is small. The bias is usually ignored in practice, because it is negligible compared with the standard errors. In small or moderate-sized samples, however, a bias correction can be appreciable, and it is helpful to have a rough estimate of its size.

In deriving the convex bounds, one need the expected values. Since there is no exact expression for the expectation of $\hat{\vec{\mu}}$, we approximate it using a general formula for the first-order bias of the estimate of $\vec{\mu}$, derived by Cordeiro and McCullagh (1991):

$$\mathbf{B}(\hat{\tilde{\mu}}) = \frac{1}{2}\left(\mathbf{G}_2 - \mathbf{G}_1\mathbf{Z}\mathbf{F}\right)\mathbf{Z}_d\mathbf{1}, \tag{25}$$

with $\mathbf{G}_1 = \text{diag}\{\psi_{11}, \cdots, \psi_{tt}\}$, $\mathbf{G}_2 = \text{diag}\{\varphi_{11}, \cdots, \varphi_{tt}\}$ where $\psi_{ij} = d\mu_{ij}/d\eta_{ij}$ and $\varphi_{ij} = d^2\mu_{ij}/d\eta_{ij}^2$, $\mathbf{Z}_d = \text{diag}\{z_{11}, \cdots, z_{tt}\}$, $\mathbf{1}$ is a t^2x1 vector of ones, and $\mathbf{F} = \text{diag}\{f_{11}, \cdots, f_{tt}\}$ with $f_{ij} = \text{Var}[Y_{ij}]^{-1}(d\mu_{ij}/d\eta_{ij})\ (d^2\mu_{ij}/d\eta_{ij}^2)$.

So, we can define adjusted values as $\hat{\tilde{\mu}}_c = \hat{\tilde{\mu}} - \hat{\mathbf{B}}(\hat{\tilde{\mu}})$, which should have smaller biases than the corresponding $\hat{\tilde{\mu}}$. Note that $\hat{\mathbf{B}}(.)$ means the value of $\mathbf{B}(.)$ at the point $(\hat{\phi}, \hat{\tilde{\mu}})$.

4.3 Upper and lower bounds for the discounted IBNR reserve

We will compute the lower and upper bound for the following conditioning normal random variable Z

$$Z = \sum_{i=2}^{t}\sum_{j=t+2-i}^{t} \nu_{ij} Y_{i+j-t-1}, \tag{26}$$

with

$$\begin{aligned}\nu_{ij} \;=\; & \sum_{k=i+1}^{t}\sum_{l=t+2-k}^{t} \exp\left(-(k+l-t-1)(\delta + \frac{1}{2}\sigma^2)\right) + \\ & \sum_{l=j}^{t} \exp\left(-(i+l-t-1)(\delta + \frac{1}{2}\sigma^2)\right).\end{aligned} \tag{27}$$

We have chosen this random variable by following the same strategy as explained in Kaas, Dhaene and Goovaerts (2000). By this choice, the lower bound will perform well in these cases. This is due to the fact that this choice makes Z a linear transformation of a first-order approximation to

$$\sum_{i=2}^{t}\sum_{j=t+2-i}^{t} e^{-Y(i+j-t-1)}. \tag{28}$$

We introduce the random variable W_{ij} defined by

$$W_{ij} = -Y(i+j-t-1). \tag{29}$$

(W_{ij}, Z) has a bivariate normal distribution. Conditionally given $Z = z$, W_{ij} has a univariate normal distribution with mean and variance given by

$$E[W_{ij}|Z=z] = E[W_{ij}] + \rho_{ij}\frac{\sigma_{W_{ij}}}{\sigma_Z}\left(z - E[Z]\right) \tag{30}$$

and

$$\text{Var}[W_{ij}|Z=z] = \sigma^2_{W_{ij}}\left(1-\rho^2_{ij}\right) \tag{31}$$

where ρ_{ij} denotes the correlation between Z and W_{ij}.

Using the same notation as in Theorem 1, we can calculate the lower, upper and improved upper bound for the discounted reserve (13). If the vector $\mathbf{X}$ describes in this setting the statistical part and the vector $\mathbf{Y}$ the financial part (the discounting process), we obtain the following expressions for S_l, S_u and S'_u

$$S_l \cong \sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}\right)\exp\left(E[W_{ij}]+\rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U)+\frac{1}{2}(1-\rho^2_{ij})\sigma^2_{W_{ij}}\right) \tag{32}$$

$$\begin{aligned} S'_u \cong & \sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}+\sqrt{\mathbf{\Sigma}(\hat{\vec{\mu}})_{ij}}\Phi^{-1}(V)\right) \\ & .\exp\left(E[W_{ij}]+\rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U)+\sqrt{1-\rho^2_{ij}}\sigma_{W_{ij}}\Phi^{-1}(W)\right), \end{aligned} \tag{33}$$

$$\begin{aligned} S_u \cong & \sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}+\sqrt{\mathbf{\Sigma}(\hat{\vec{\mu}})_{ij}}\Phi^{-1}(V)\right) \\ & .\exp\left(E[W_{ij}]+\sigma_{W_{ij}}\Phi^{-1}(U)\right), \end{aligned} \tag{34}$$

where U, V and W are mutually independent uniform$(0,1)$ random variables and Φ is the cumulative distribution function of the $N(0,1)$ distribution.

Proof.

1. If a random variable X is lognormal(μ,σ^2) distributed, then $E[X]=\exp(\mu+\frac{1}{2}\sigma^2)$. Hence for $Z=\sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\nu_{ij}Y_{i+j-t-1}$, we find, taking $U=\Phi\left(\frac{Z-E[Z]}{\sigma_Z}\right)\sim$ uniform$(0,1)$, that

$$\begin{aligned} E[\hat{\mu}_{ij}]E[V_{ij}|Z] \cong & \left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}\right) \\ & .\exp\left(E[W_{ij}]+\rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U)+\frac{1}{2}(1-\rho^2_{ij})\sigma^2_{W_{ij}}\right) \end{aligned} \tag{35}$$

2. If a random variable X is lognormal(μ,σ^2) distributed, then $F^{-1}_X(p)=\exp(\mu+\sigma\Phi^{-1}(p))$. So, we find that

$$\begin{aligned} F^{-1}_{\hat{\mu}_{ij}}(p)F^{-1}_{V_{ij}|Z}(q) \cong & \sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}+\sqrt{\mathbf{\Sigma}(\hat{\vec{\mu}})_{ij}}\Phi^{-1}(p)\right) \\ & .\exp\left(E[W_{ij}]+\rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U)+\sqrt{1-\rho^2_{ij}}\sigma_{W_{ij}}\Phi^{-1}(q)\right) \end{aligned} \tag{36}$$

3. Equation (34) follows from (36).

Since we have no equality of the first moments, the convex order relationship (18) between the three approximations and S is not valid here. This doens not impose any restrictions on the use of the approximations. In fact, we can say that the convex order only holds asymptotically in this application.

In Hoedemakers et al. (2003) the reader can find some more details concerning the (calculation of the) distributions of the different bounds.

5 Numerical illustrations

In this section we illustrate the effectiveness of the bounds derived for the discounted IBNR reserve S. We investigate the accuracy of the proposed bounds, by comparing their cumulative distribution function (cdf) to the empirical cdf obtained with Monte Carlo simulation, which serves as a close approximation to the exact distribution of S.

In order to illustrate the power of the bounds, namely inspecting the deviation of the cdf of the convex bounds S_l, S_u and S'_u from the true distribution of the total IBNR reserve S, we simulate a triangle from a particular model. In these illustrations we model the incremental claims Y_{ij} with a logarithmic link function to obtain a multiplicative parametric structure and we link the expected value of the response to the chain-ladder type linear predictor. Formally, this means with the notation introduced in section 2

$$\begin{aligned} E[Y_{ij}] &= \mu_{ij}, \\ \mathrm{Var}[Y_{ij}] &= \phi\mu_{ij}^{\kappa}, \\ \log(\mu_{ij}) &= \eta_{ij}, \\ \eta_{ij} &= \alpha_i + \beta_j. \end{aligned} \tag{37}$$

The choice of the error distribution is determined by κ. Note that a parameter, for example β_1, must be set equal to zero, in order to have a non-singular regression matrix.

We also specify the multivariate distribution function of the random vector $(Y_1, Y_2, \ldots, Y_{t-1})$. In particular, we will assume that the random variables Y_i are i.i.d. and normal$(\delta+\frac{1}{2}\sigma^2, \sigma^2)$ distributed with $\delta = 0.08$ and $\sigma = 0.11$. This enables now to simulate the cdf's. The conditioning random variable Z is defined as in (26) and (27).

In a first example we consider model (37) with the Poisson error distribution (κ=1). The simulated triangle for this model is depicted in Table 2. Parameter estimates and standard errors for this fit are shown in Table 3.

Since this model is a generalized linear model, standard statistical software can be used to obtain maximum (quasi) likelihood parameter estimates, fitted and predicted values.

	1	2	3	4	5	6	7	8	9	10	11
1	362,505	493,876	323,065	237,574	249,850	152,221	139,293	95,961	70,812	53,395	35,902
2	399,642	545,274	357,788	263,414	276,500	168,064	153,603	105,760	78,736	58,612	
3	805,843	1,100,020	722,110	531,220	557,195	337,606	309,306	213,416	158,611		
4	728,762	994,975	653,231	478,728	502,797	306,071	278,436	193,201			
5	661,713	899,778	591,647	434,626	456,763	276,588	253,297				
6	539,789	737,394	484,415	355,175	372,800	226,865					
7	983,897	1,341,585	881,786	647,431	679,264						
8	889,268	1,217,248	798,387	585,099							
9	487,823	666,590	437,987								
10	442,982	601,706									
11	1,087,672										

Table 2: Ex. 1 (κ=1): Run-off triangle with non-cumulative claim figures.

Parameter	Model parameter	Estimate	Standard error
α_1	12.8	12.7990566	0.0007918770
α_2	12.9	12.8989406	0.0007631003
α_3	13.6	13.6001742	0.0006060520
α_4	13.5	13.4989356	0.0006283423
α_5	13.4	13.4007436	0.0006556928
α_6	13.2	13.1997559	0.0007180990
α_7	13.8	13.7991616	0.0005991796
α_8	13.7	13.6998329	0.0006464691
α_9	13.1	13.0989431	0.0008707837
α_{10}	13.0	12.9987252	0.0010370987
α_{11}	13.9	13.8995502	0.0009710197
β_2	0.31	0.3106789	0.0005310346
β_3	−0.11	−0.1099061	0.0006026958
β_4	−0.42	−0.4189677	0.0006804776
β_5	−0.37	−0.3700452	0.0007168115
β_6	−0.87	−0.8685181	0.0009462170
β_7	−0.96	−0.9585385	0.0010542829
β_8	−1.33	−1.3284870	0.0013825136
β_9	−1.63	−1.6269622	0.0018947413
β_{10}	−1.92	−1.9170757	0.0030880359
β_{11}	−2.31	−2.3105083	0.0054029754
ϕ	1	1.025663	

Table 3: Ex. 1 (κ=1): Model specification, maximum likelihood estimates and standard errors.

Standard statistical theory also suggests goodness-of-fit measures and appropriate residual definitions for diagnostic checks of the fitted model.

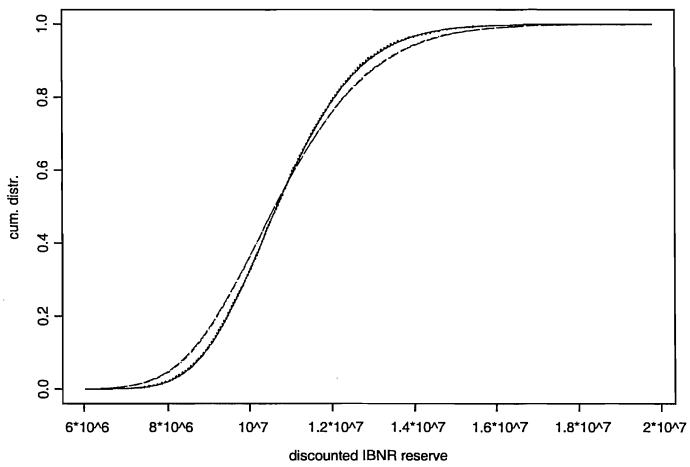


Figure 2: Ex. 1 (κ=1): The cdf's of the lower bound S_l (dotted line), the upper bound S_u (dot-dashed line) and of the improved upper bound S'_u (dashed line) vs. the distribution of the discounted IBNR reserve S approximated by extensive simulation (solid line) for the run-off triangle in Table 2.

Figure 2 shows the distribution functions of the different bounds compared to the empirical distribution obtained by simulation. The distribution functions are remarkably close to each other and enclose the simulated cdf nicely. This is confirmed by the QQ-plot in Figure 3 where we also see that the comonotonic improved upper bound has somewhat heavier tails. Numerical values of some high quantiles of S, S_l and S'_u are given in Table 5. We find that the improved upper bound S'_u is very close to the comonotonic upper bound S_u. This could be expected because in this example ρ_{ij} is close to ρ_{kl} for any pair (ij, kl) with ij and kl sufficient close. This implies that for any such pair (ij, kl) $\left(F^{-1}_{e^{-Y(i+j-t-1)}|Z}(U), F^{-1}_{e^{-Y(k+l-t-1)}|Z}(U)\right)$ is close to $\left(F^{-1}_{e^{-Y(i+j-t-1)}}(U), F^{-1}_{e^{-Y(k+l-t-1)}}(U)\right)$.

Table 4 summarizes the numerical values of the 95th percentiles of the two bounds S_l and S'_u vs. S, together with their means and standard deviations. This is also provided for the row totals

$$S_i = \sum_{j=t+2-i}^{t} \hat{\mu}_{ij} e^{-Y(i+j-t-1)}, \qquad i = 2, \cdots, t. \tag{38}$$

	F_{S_l}			F_S			$F_{S'_u}$		
year	95%	mean	st. dev.	95%	mean	st. dev.	95%	mean	st. dev.
2	41,197	36,621	2,679	43,523	36,620	3,972	43,623	36,621	4,032
3	206,740	177,582	16,843	213,586	177,580	20,745	216,333	177,582	22,686
4	335,050	280,309	31,314	341,367	280,310	35,061	340,344	280,311	39,689
5	482,333	396,021	49,038	488,207	396,025	52,194	499,841	396,027	60,213
6	603,492	490,201	64,100	606,876	490,200	66,411	618,662	490,203	77,768
7	1,506,886	1,204,980	170,088	1,507,845	1,204,970	172,583	1,557,930	1,204,982	202,984
8	1,971,284	1,574,980	222,932	1,977,962	1,574,977	225,404	2,024,461	1,574,983	267,694
9	1,388,485	1,093,661	165,352	1,390,932	1,093,943	165,734	1,426,650	1,093,666	196,364
10	1,621,765	1,278,628	196,488	1,625,542	1,278,619	196,551	1,690,363	1,278,629	235,170
11	5,430,624	4,275,045	650,349	5,437,003	4,275,009	650,354	5,665,896	4,275,047	782,517
total	13,604,855	10,808,003	1,570,742	13,639,223	10,808,012	1,579,730	14,153,682	10,808,049	1,889,011

Table 4: Ex. 1 (κ=1): 95th percentiles, means and standard deviations of the distributions of S_l and S'_u vs. S. ($\delta = 0.08, \sigma = 0.11$)

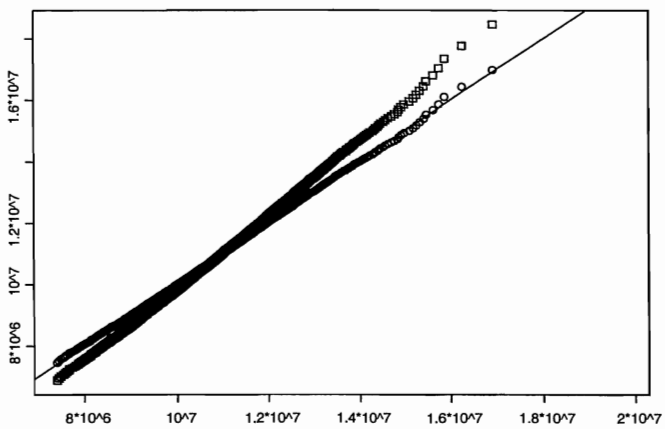


Figure 3: Ex. 1 (κ=1): QQ-plot of the quantiles of S_l (○) and S'_u (□) versus those of S.

We can conclude that the lower bound approximates the "real discounted reserve" very well. Remark that the means of S, S_l and S'_u are almost equal. The small differences come from the fact that these random variables are only "asymptotically" convex ordered, but by using a bias correction this problem is almost neutralized.

In a second example, we illustrate the method using a gamma regression model instead of a Possion regression model. The results are very similar. The simulated run-off triangle

p	$F_{S_l}^{-1}(p)$	$F_S^{-1}(p)$	$F_{S'_u}^{-1}(p)$
0.95	13,604,855	13,639,223	14,153,682
0.975	14,268,088	14,325,593	14,968,214
0.99	15,087,300	15,090,853	15,982,232
0.995	15,524,557	15,525,160	16,716,836
0.999	16,593,144	16,579,201	18,354,296

Table 5: Ex. 1 (κ=1): Quantiles of S_l and S'_u versus those of S.

	1	2	3	4	5	6	7	8	9	10
1	292,686	683,476	701,376	747,034	504,265	312,468	284,954	170,814	249,348	69,752
2	423,113	991,584	1,032,142	945,156	500,205	413,863	434,622	206,319	342,383	
3	344,386	936,335	971,651	1,104,206	575,666	416,179	359,195	246,463		
4	308,603	830,615	864,751	981,609	504,837	372,329	353,145			
5	338,073	884,174	895,252	927,435	647,289	391,208				
6	322,270	927,791	980,275	952,298	577,483					
7	387,598	1,084,439	1,126,376	1,035,701						
8	385,603	1,143,038	1,209,301							
9	388,795	951,100								
10	308,586									

Table 6: Ex. 2 (κ=2): Run-off triangle with non-cumulative claim figures.

and the table of the parameter values are shown in Table 6 and Table 7 respectively.

Since the upper and lower bounds appear to be rather close to each other in Figure 4, they prove to be quite good approximations for the unknown distribution of S. We can see that in contrast to example 1, here the improved upper bound S'_u indeed improves the upper bound S_u. From the QQ-plot in Figure 5, we can conclude that the improved upper bound (slightly) overestimates the tails of S, whereas the accuracy of the lower bond is extremely high for the chosen set of parameter values. Table 9 confirms these observations.

We can conclude that in each of the two examples the lower bound approximates the "real discounted reserve" very well. Using bootstrap methodology it is possible to provide statistical confidence intervals for the given bounds incorporating the estimation error. This was performed in Hoedemakers et al. (2003) for the case of the simple loglinear model.

The chain-ladder technique and the associated GLM are widely used methods of claims reserving. The range of the models can be expanded significantly by using other assumptions for the distribution of the data, and for the link function. Other parametric models can easily be incorporated using different linear predictors.

Parameter	Model parameter	Estimate	Standard error
α_1	12.56	12.51790600	0.03610258
α_2	12.88	12.80591922	0.03610258
α_3	12.84	12.79630916	0.03663606
α_4	12.72	12.67925064	0.03753296
α_5	12.79	12.74885712	0.03883950
α_6	12.83	12.74961540	0.04071262
α_7	12.91	12.88770262	0.04347811
α_8	13.02	12.94869876	0.04784851
α_9	12.87	12.83535778	0.05571696
α_{10}	12.75	12.63975585	0.07475215
β_2	0.91	0.96676725	0.03523850
β_3	0.93	1.00976556	0.03685331
β_4	0.99	1.02578624	0.03861620
β_5	0.41	0.50519662	0.04071262
β_6	0.11	0.13617431	0.04337087
β_7	−0.05	0.07957371	0.04697864
β_8	−0.45	−0.47029820	0.05233710
β_9	−0.06	−0.07666105	0.06153780
β_{10}	−1.43	−1.36520463	0.08301374
ϕ	0.005	0.0055879	

Table 7: Ex. 1 (κ=1): Model specification, maximum likelihood estimates and standard errors.

	F_{S_l}			F_S			$F_{S'_u}$		
year	95%	mean	st. dev.	95%	mean	st. dev.	95%	mean	st. dev.
2	98,490	86,120	7,206	107,422	86,118	7,210	109,588	86,122	9,483
3	451,438	387,695	36,861	474,399	387,661	36,873	477,521	387,696	46,308
4	612,987	503,661	62,363	628,801	503,659	62,514	629,408	503,663	72,386
5	1,038,878	842,693	111,288	1,052,407	842,693	111,988	1,069,225	842,699	128,800
6	1,432,101	1,143,062	163,211	1,448,157	1,143,043	173,024	1,476,022	1,143,069	188,317
7	2,287,505	1,816,883	265,204	2,318,270	1,816,860	266,371	2,367,733	1,816,886	309,432
8	3,592,631	2,866,107	409,581	3,653,365	2,866,110	411,061	3,730,824	2,866,112	485,496
9	4,199,902	3,315,398	497,675	4,289,402	3,315,399	499,049	4,369,312	3,315,401	589,304
10	4,198,309	3,271,581	519,961	4,311,722	3,271,577	522,461	4,377,799	3,271,584	617,083
total	17,912,242	14,233,213	2,073,272	18,023,261	14,233,155	2,077,677	18,582,249	14,233,232	2,446,443

Table 8: Ex. 2 (κ=2): 95th percentiles, means and standard deviations of the distributions of S_l and S'_u vs. S. ($\delta = 0.08, \sigma = 0.11$)

6 Conclusions and possibilities for future research

In this paper, we considered the problem of deriving the distribution function of the discounted loss reserve using a generalized linear model together with some stochastic return

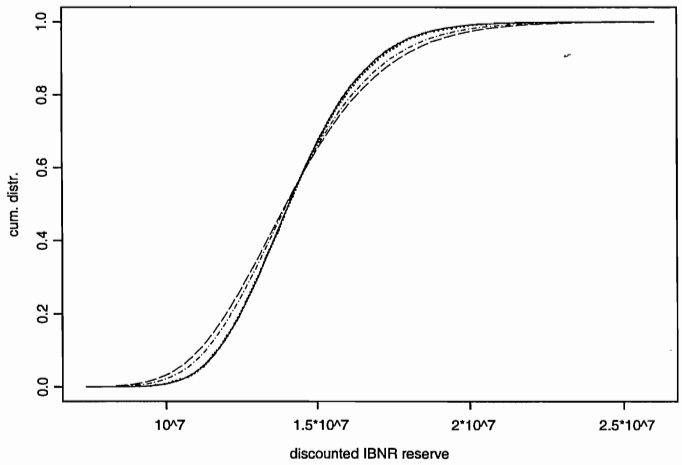


Figure 4: Ex. 2 (κ=2): The cdf's of the lower bound S_l (dotted line), the upper bound S_u (dot-dashed line) and of the improved upper bound S'_u (dashed line) vs. the distribution of the discounted IBNR reserve S approximated by extensive simulation (solid line) for the run-off triangle in Table 6.

p	$F^{-1}_{S_l}(p)$	$F^{-1}_{S}(p)$	$F^{-1}_{S'_u}(p)$
0.95	17,912,242	18,023,261	18,582,249
0.975	18,974,353	19,014,554	19,634,348
0.99	19,935,094	20,205,096	20,941,335
0.995	20,765,334	20,994,297	21,886,306
0.999	22,239,709	22,243,174	23,987,206

Table 9: Ex. 2 (κ=2): Quantiles of S_l and S'_u versus those of S.

process. An advantage of working with GLMs is the fact that it is not necessary anymore that the Y_{ij}'s are non-negative or integers. The incremental claim amounts can for instance be modelled as independent normal, Poisson, gamma or inverse Gaussian response variables together with a logarithmic link function and a specified linear predictor. When using the logarithmic link function, which provides a multiplicative parametric structure and produces positive fitted values, the technique is not applicable for incurred data with a large number of negative incremental claims in the later stages of development.

Because an explicit expression for the distribution function is hard to obtain, we presented

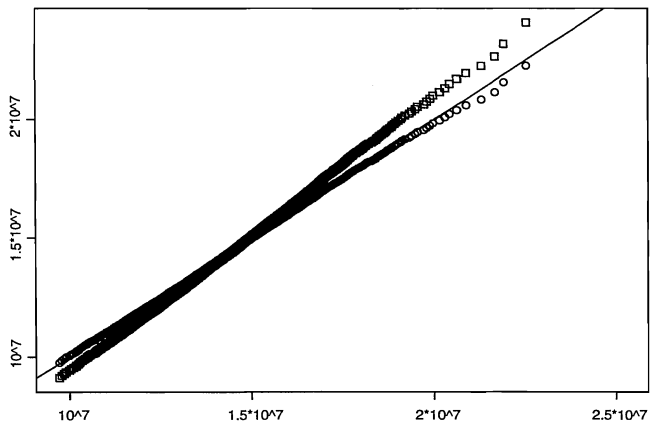


Figure 5: Ex. 2 (κ=2): QQ-plot of the quantiles of S_l (○) and S'_u (□) versus those of S.

three approximations for this distribution function, in the sense that these approximations are larger or smaller in convex order sense than the exact distribution. This technique is common practice in the actuarial literature. When lower and upper bounds are close to each other, together they can provide reliable information about the original and more complex variable.

An essential point in the derivation of the presented approximations is the choice of the conditioning random variable Z. The improved upper bound becomes closer to the original variable S, the more the variables Z and S are alike. When dealing with very large variances in the statistical part of our model, an adaptation of the random variable Z will be necessary. This could be a research object for a next paper.

Acknowledgements

The authors acknowledge financial support by the Onderzoeksfonds K.U.Leuven (GOA/02: Actuariële, financiële en statistische aspecten van afhankelijkheden in verzekerings- en financiële portefeuilles).

References

[1] Barnett, G. and Zehnwirth, B., 1998. Best estimates for reserves. *Casualty Actuarial Society Forum*, 1-54.

[2] Cordeiro, G.M. and McCullagh P., 1991. Bias correction in generalized linear models. *Journal of the Royal Statistical Society B*, **53, No. 3**, 629-643.

[3] De Vylder, F. and Goovaerts, M.J., 1979. Proceedings of the first meeting of the contact group "Actuarial Sciences", K.U.Leuven, nr 7904B, wettelijk Depot: D/1979/23761/5.

[4] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D., 2002. (a) The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, **31(1)**, 3-33.

[5] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D., 2002. (b) The concept of comonotonicity in actuarial science and finance: Applications. *Insurance: Mathematics and Economics*, **31(2)**, 133-161.

[6] England, P.D. and Verrall, R.J., 1999. Analytic and bootstrap estimates of prediction errors in claim reserving. *Insurance: Mathematics and Economics*, **25**, 281-293.

[7] England, P.D. and Verrall, R.J., 2001. A flexible framework for stochastic claims reserving. *Proceedings of the Casualty Actuarial Society*, **88**.

[8] England, P.D. and Verrall, R.J., 2002. Stochastic claims reserving in general insurance. *British Actuarial Journal*, **8(3)**, 443-518.

[9] Goovaerts, M.J., Kaas, R., Dhaene, J. and Denuit, M., 2001. *Modern Actuarial Risk Theory*. Kluwer Academic, The Netherlands.

[10] Hoedemakers T., Beirlant J., Goovaerts M.J. and Dhaene J., 2003. Confidence bounds for discounted loss reserves. *Insurance: Mathematics and Economics*, To be published.

[11] Kaas, R., Dhaene, J. and Goovaerts, M.J., 2000. Upper and lower bounds for sums of random variables. *Insurance: Mathematics and Economics*, **27**, 151-168.

[12] Kremer, E., 1982. IBNR-Claims and the two-way Model of ANOVA. *Scandinavian Actuarial Journal*, 47-55.

[13] Mack, T., 1991. A simple parametric model for rating automobile insurance or estimating IBNR claims reserves. *ASTIN Bulletin*, **22(1)**, 93-109.

[14] McCullagh P. and Nelder J.A., 1992. *Generalized Linear Models*. Chapman and Hall, 2nd edition.

[15] Renshaw, A.E., 1989. Chain ladder and interactive modelling (claims reserving and GLIM). *Journal of the Institute of Actuaries*, **116(III)**, 559-587.

[16] Renshaw, A.E., 1994. On the second moment properties and the implementation of certain GLIM based stochastic claims reserving models. Actuarial Research Paper No. 65, Department of Actuarial Science and Statistics, City University, London, EC1V 0HB.

[17] Renshaw, A.E. and Verrall, R.J., 1994. A stochastic model underlying the chain-ladder technique. *Proceedings XXV Astin Colloquium*, Cannes.

[18] Verrall, R.J., 1989. A state space representation of the chain-ladder linear model. *J.I.A.*, **116**, 589-610.

[19] Verrall, R.J., 1991. On the unbiased estimation of reserves from loglinear models. *Insurance: Mathematics and Economics*, **10**, 75-80.

[20] Wedderburn, R.W.M., 1974. Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss-Newton Method. *Biometrika*, **61**, 439-447.

[21] Wright, T.S., 1990. A stochastic method for claims reserving in general insurance. *J.I.A.*, **117**, 677-731.

[22] Zehnwirth, B., 1985. *ICRFS version 4 manual and users guide.* Benhar Nominees PTy Ltd, Turramurra, NSW, Australia

[23] Zehnwirth, B., 1989. The chain ladder technique - A stochastic model. *Claims Reserving Manual*, **Vol 2**, Institute of Actuaries, London.