On the Distribution of Discounted Loss Reserves 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
†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 to meet claims arising in the future on the written policies represents an important task for insurance companies in order 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 claim amounts 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 subexponential 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) are not known for certain. We will model the stochastic discount factor using a Brownian motion with drift. When determining the discounted loss reserve S, 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

Year of origin	*Development year* 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

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 illustrate the obtained bounds graphically.

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

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

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

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

where $a(.)$, $b(.)$ 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

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

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. Denoting the mean of Y_{ij} by μ_{ij}, it is known that

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

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

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

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

The possible distributions to work with in claims reserving include for instance 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 canonical 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

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

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 of 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 further 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) described 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}. Then the variance σ_{ij}^2 can be taken as a function of the predictor variables through the mean response μ_{ij}, or the variance can be modelled using a parameterised structure (see Renshaw (1994)). 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).

Supposing 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. $\phi = \sigma^2$. We remark that each incremental claim must be greater than zero and predictions from this model can yield unusable results.

We end this section with some extra comments concerning GLMs.

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 mean-variance relationship. 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.

When using a logarithmic link function, the quasi-likelihood equations are given by

$$\begin{aligned} \sum_{j=1}^{t+1-i} e^{\eta_{ij}} &= \sum_{j=1}^{t+1-i} Y_{ij} \quad 1 \leq i \leq t; \\ \sum_{i=1}^{t+1-j} e^{\eta_{ij}} &= \sum_{i=1}^{t+1-j} Y_{ij} \quad 1 \leq j \leq t. \end{aligned} \tag{8}$$

As can easily be seen from these equations, it is necessary to impose the constraint that the sum of the incremental claims in every row and column has to be non-negative. This implies that the described technique is not applicable for modelling incurred data with a large number of negative incremental claims in the later stages of development, which is the result of overestimates of case reserves in the first development years.

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

$$\text{Var}[Y_{ij}] = E[Y_{ij}]. \tag{9}$$

One can allow for more or less dispersion in the data by generalizing (9) to $\text{Var}[Y_{ij}]{=}\phi E[Y_{ij}]$ ($\phi \in (0, \infty)$) 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

$$\mathrm{Var}[Y_{ij}] = \phi \left(E[Y_{ij}]\right)^2 . \tag{10}$$

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

$$\text{IBNR reserve} = \sum_{i=2}^{t}\sum_{j=t+2-i}^{t} \hat{\mu}_{ij}, \tag{11}$$

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

$$Y(i) = Y_1 + Y_2 + \cdots + Y_i \tag{12}$$

and assume that

$$Y(i) = (\delta + \sigma^2/2)i + \sigma B(i), \tag{13}$$

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

$$S \stackrel{def}{=} \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \hat{\mu}_{ij} e^{-Y(i+j-t-1)} \tag{14}$$

$$= \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \hat{\mu}_{ij} \exp\left(-(\delta + \sigma^2/2)(i+j-t-1) - \sigma B(i+j-t-1)\right). \tag{15}$$

With this adaptation, we have that

$$E[e^{-Y(i)}] \,.\, e^{\delta i} = 1. \tag{16}$$

In order to study the distribution of the discounted IBNR reserve (14), 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*

$$E[u(V)] \leq E[u(W)], \tag{17}$$

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

$$V \leq_{cx} W. \tag{18}$$

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. (2000) 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:*

$$S_l \leq_{cx} S \leq_{cx} S'_u \leq_{cx} S_u, \tag{19}$$

with

$$\begin{aligned} S &= X_1Y_1 + X_2Y_2 + \ldots + X_nY_n & (20)\\ S_l &= E[X_1|Z]E[Y_1] + E[X_2|Z]E[Y_2] + \ldots + E[X_n|Z]E[Y_n] & (21)\\ S'_u &= F^{-1}_{X_1|Z}(U)F^{-1}_{Y_1}(V) + F^{-1}_{X_2|Z}(U)F^{-1}_{Y_2}(V) + \ldots + F^{-1}_{X_n|Z}(U)F^{-1}_{Y_n}(V) & (22)\\ S_u &= F^{-1}_{X_1}(U)F^{-1}_{Y_1}(V) + F^{-1}_{X_2}(U)F^{-1}_{Y_2}(V) + \ldots + F^{-1}_{X_n}(U)F^{-1}_{Y_n}(V), & (23) \end{aligned}$$

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 (14). Before starting with this, we have to specify further 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 equation for $\hat{\vec{\beta}}$ is then given by

$$\mathbf{U}'\hat{\mathbf{W}}\mathbf{U}\hat{\vec{\beta}} = \mathbf{U}'\hat{\mathbf{W}}\hat{\vec{y}}^*, \tag{24}$$

where $\mathbf{W} = \text{diag}\{w_{11}, \cdots, w_{t1}\}$, with $w_{ij} = \text{Var}[Y_{ij}]^{-1}(d\mu_{ij}/d\eta_{ij})^2$, $\vec{y}^* = (y^*_{11}, \cdots, y^*_{t1})'$, and denoting $y^*_{ij} = \eta_{ij} + (y_{ij} - \mu_{ij})d\eta_{ij}/d\mu_{ij}$ where 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 asymptotically normal statistics, many functions of such statistics 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 $\mathbf{R}\vec{\beta} = \left((\mathbf{R}\vec{\beta})_{11}, \cdots, (\mathbf{R}\vec{\beta})_{tt}\right)'$ and variance-covariance matrix $\mathbf{\Sigma}(\mathbf{R}\hat{\vec{\beta}}) = \mathbf{\Sigma}^a = \{\sigma^a_{ij}\} = \mathbf{R}(\mathbf{U}'\mathbf{W}\mathbf{U})^{-1}\mathbf{R}'$ 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

$$\left[\hat{\vec{\mu}} - \vec{\mu}\right] \xrightarrow{d} N\left(0, \mathbf{\Sigma}(\hat{\vec{\mu}})\right), \tag{25}$$

where $\xrightarrow{d}$ means convergence in distribution and $\mathbf{\Sigma}(\hat{\vec{\mu}}) = \vec{\psi}'\mathbf{\Sigma}^a\vec{\psi}$. 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 necessary, 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):

$$\mathbf{B}(\hat{\vec{\mu}}) = \frac{1}{2}\left\{\mathbf{G}_2\mathbf{\Sigma}^a_d\tilde{\mathbf{1}} - \mathbf{G}_1\mathbf{R}\mathbf{\Sigma}^b\mathbf{U}'\mathbf{\Sigma}^c_d\mathbf{F}_d\bar{\mathbf{1}}\right\}, \tag{26}$$

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 $\varphi_{ij} = d^2\mu_{ij}/d\eta^2_{ij}$, $\mathbf{\Sigma}^b = \mathbf{\Sigma}(\hat{\vec{\beta}}) = \{\sigma^b_{ij}\} = (\mathbf{U}'\mathbf{W}\mathbf{U})^{-1}$, $\mathbf{\Sigma}^c = \mathbf{\Sigma}(\mathbf{U}\hat{\vec{\beta}}) = \{\sigma^c_{ij}\} = \mathbf{U}\mathbf{\Sigma}^b\mathbf{U}'$, $\mathbf{\Sigma}^a_d = \text{diag}\{\sigma^a_{11}, \cdots, \sigma^a_{tt}\}$, $\mathbf{\Sigma}^c_d = \text{diag}\{\sigma^c_{11}, \cdots, \sigma^c_{t1}\}$, $\tilde{\mathbf{1}}$ is a $t^2 \times 1$ vector of ones, $\bar{\mathbf{1}}$ is a $t(t+1)/2 \times 1$ vector of ones, and $\mathbf{F}_d = \text{diag}\{f_{11}, \cdots, f_{t1}\}$ with $f_{ij} = \text{Var}[Y_{ij}]^{-1}(d\mu_{ij}/d\eta_{ij})\ (d^2\mu_{ij}/d\eta^2_{ij})$.

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

4.3 Upper and lower bounds for the discounted IBNR reserve

We will compute the lower and upper bound using the conditioning normal random variable

$$Z = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \nu_{ij} Y(i+j-t-1), \tag{27}$$

with

$$\nu_{ij} = \left(\mu_{ij} + \mathbf{B}(\hat{\vec{\mu}})_{ij}\right) \exp\left(-(i+j-t-1)\delta\right). \tag{28}$$

We have chosen this random variable by following the same strategy as explained in Kaas et al. (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

$$\sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}\right)e^{-Y(i+j-t-1)}. \tag{29}$$

When introducing the random variable $W_{ij}=-Y(i+j-t-1)$, we have that (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

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

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 (14). If the vector $\mathbf{Y}$ describes in this setting the statistical part and the vector $\mathbf{X}$ the financial part (the discounting process), we obtain the following approximate expressions for S_l, S_u and S'_u

$$\begin{aligned}
S_l &= \sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}\right)\exp\left(E[W_{ij}]+\rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U)+\frac{1}{2}(1-\rho^2_{ij})\sigma^2_{W_{ij}}\right), && (32)\\
S'_u &= \sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}+\sqrt{\mathbf{\Sigma}(\hat{\vec{\mu}})_{ij}}\Phi^{-1}(V)\right)\times\\
&\quad\times\exp\left(E[W_{ij}]+\rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U)+\sqrt{1-\rho^2_{ij}}\sigma_{W_{ij}}\Phi^{-1}(W)\right), && (33)\\
S_u &= \sum_{i=2}^{t}\sum_{j=t+2-i}^{t}\left(\mu_{ij}+\mathbf{B}(\hat{\vec{\mu}})_{ij}+\sqrt{\mathbf{\Sigma}(\hat{\vec{\mu}})_{ij}}\Phi^{-1}(V)\right)\times\\
&\quad\times\exp\left(E[W_{ij}]+\sigma_{W_{ij}}\Phi^{-1}(U)\right), && (34)
\end{aligned}$$

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

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

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

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

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

3. Equation (34) follows from (36). ■

Since we have no equality of the first moments, the convex order relationship (19) between the three approximations and S is not valid here. This does 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, under the model studied. 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. The simulation results are based on generating 100.000 random paths. The estimates obtained from this time-consuming simulation will serve as benchmark. The random paths are based on antithetic variables in order to reduce the variance of the Monte Carlo estimates.

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 that

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

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$(\delta + \frac{1}{2}\sigma^2, \sigma^2)$

	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.

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 (27) and (28).

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.

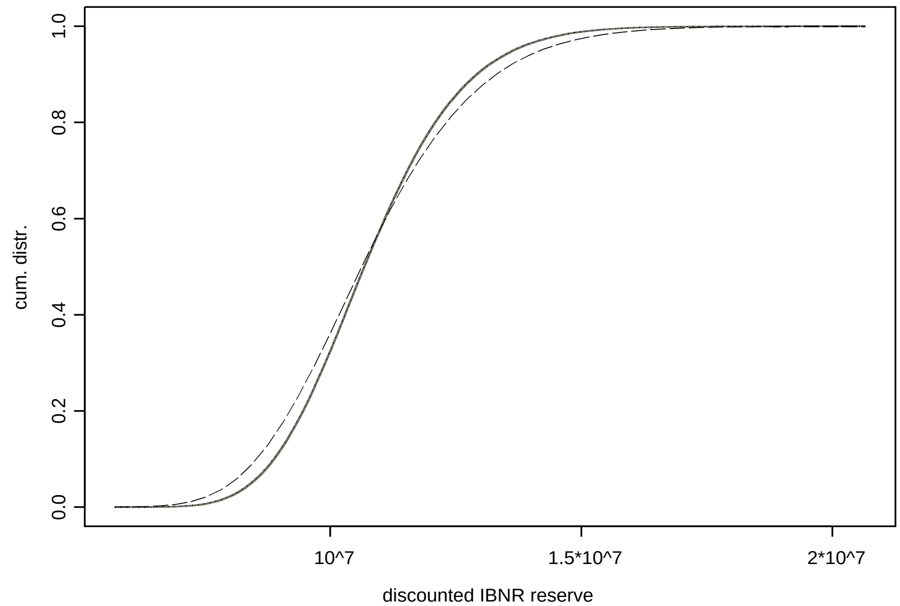

Figure 2: Ex. 1 (κ=1): The cdf's of the lower bound S_l (dotted line), the upper bound S_u (dashed line) vs. the distribution of the discounted IBNR reserve S approximated by extensive simulation (solid line) for the run-off triangle in Table 2.

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

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 upper bound has somewhat heavier tails. Numerical values of some high quantiles of S, S_l and S_u are given in Table 5.

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

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

We can conclude that the lower bound approximates the "real discounted reserve" very well.

year	F_{S_l}			F_S			F_{S_u}		
	95%	mean	st. dev.	95%	mean	st. dev.	95%	mean	st. dev.
2	43,622	36,623	4,041	43,624	36,623	4,042	43,631	36,623	4,046
3	214,142	177,600	21,002	214,428	177,600	21,040	217,352	177,600	22,751
4	342,589	280,318	35,595	343,011	280,318	35,691	350,360	280,318	39,805
5	489,087	396,089	52,976	489,689	396,089	53,194	502,853	396,089	60,398
6	608,891	490,289	67,401	609,535	490,289	67,565	628,672	490,289	78,021
7	1,514,480	1,205,224	175,099	1,516,799	1,205,224	175,658	1,567,945	1,205,224	203,692
8	1,977,737	1,575,313	227,703	1,980,868	1,575,313	228,343	2,054,475	1,575,313	268,661
9	1,390,601	1,093,992	167,320	1,392,957	1,093,992	167,862	1,444,660	1,093,992	197,121
10	1,632,675	1,278,947	199,110	1,634,653	1,278,947	199,693	1,702,375	1,278,947	236,121
11	5,439,986	4,276,121	655,280	5,446,107	4,276,121	656,472	5,685,932	4,276,121	785,741
total	13,631,905	10,810,476	1,594,152	13,648,695	10,810,476	1,597,507	14,200,226	10,810,476	1,896,219

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

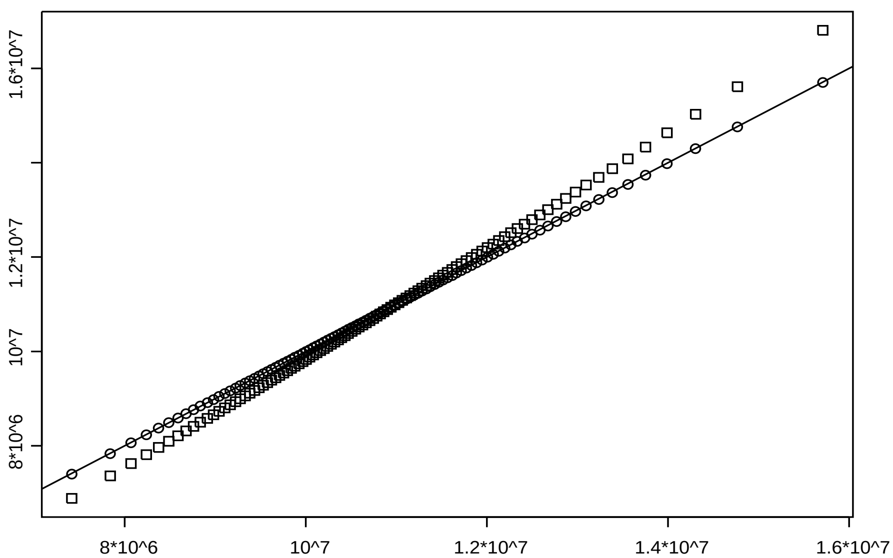

Figure 3: Ex. 1 (κ=1): QQ-plot of the quantiles of S_l ($\circ$) and S_u ($\square$) versus those of S.

p	$F_{S_l}^{-1}(p)$	$F_S^{-1}(p)$	$F_{S_u}^{-1}(p)$
0.95	13,631,905	13,648,695	14,200,226
0.975	14,296,448	14,305,657	15,027,414
0.99	15,115,189	15,122,840	16,057,613
0.995	15,702,702	15,709,497	16,804,206
0.999	16,996,374	17,018,860	18,469,110

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.

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. 2 (κ=2): Model specification, maximum likelihood estimates and standard errors.

In a second example, we illustrate the method using a gamma regression model instead of a Poisson regression model. The results are very similar. The simulated run-off triangle 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. From the QQ-plot in Figure 5, we can conclude that the 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. Some numerical values for the row totals are given in Table 8.

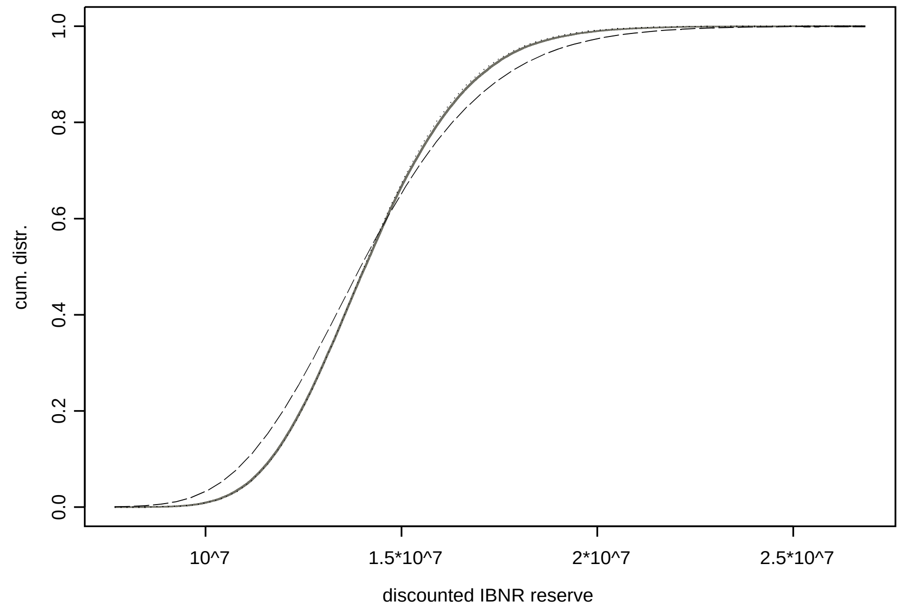

Figure 4: Ex. 2 (κ=2): The cdf's of the lower bound S_l (dotted line), the upper bound S_u (dashed line) vs. the distribution of the discounted IBNR reserve S approximated by extensive simulation (solid line) for the run-off triangle in Table 6.

	F_{S_l}			F_S			F_{S_u}		
year	95%	mean	st. dev.	95%	mean	st. dev.	95%	mean	st. dev.
2	102,356	85,934	9,481	103,187	85,934	9,747	106,553	85,934	11,857
3	462,847	387,251	43,602	466,609	387,251	44,775	479,913	387,251	53,038
4	619,090	503,187	66,173	624,112	503,187	68,014	642,819	503,187	79,110
5	1,042,181	842,092	113,871	1,050,345	842,092	117,188	1,087,242	842,092	138,274
6	1,432,744	1,142,369	164,543	1,444,486	1,142,369	169,224	1,498,433	1,142,369	199,885
7	2,286,615	1,815,836	266,221	2,305,985	1,815,836	273,721	2,400,469	1,815,836	327,286
8	3,590,200	2,864,235	410,836	3,619,252	2,864,235	422,643	3,785,691	2,864,235	515,535
9	4,197,088	3,312,169	499,465	4,231,171	3,312,169	513,473	4,442,318	3,312,169	630,417
10	4,197,710	3,264,577	524,580	4,231,798	3,264,577	539,321	4,487,925	3,264,577	679,607
total	17,888,702	14,217,631	2,076,583	18,033,971	14,217,631	2,135,185	18,926,155	14,217,631	2,631,780

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

We remark that the improved upper bound S'_u is very close to the comonotonic upper bound S_u. This could be expected because 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 $\left(F^{-1}_{e^{-Y(i+j-t-1)}}(U), F^{-1}_{e^{-Y(k+l-t-1)}}(U)\right)$. Since the improved upper bound requires more computational time, see Hoedemakers et al. (2003) for more details, the results for the improved upper bound are not displayed in this section.

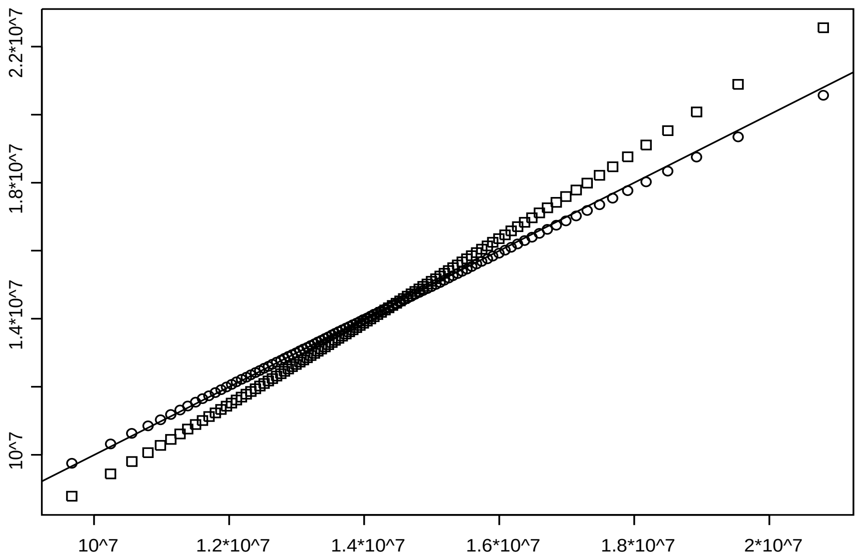

Figure 5: Ex. 2 (κ=2): QQ-plot of the quantiles of S_l ($\circ$) and S_u ($\square$) versus those of S.

p	$F_{S_l}^{-1}(p)$	$F_S^{-1}(p)$	$F_{S_u}^{-1}(p)$
0.95	17,888,702	18,033,971	18,926,155
0.975	18,749,885	18,923,975	20,077,389
0.99	19,809,569	19,986,346	21,511,663
0.995	20,569,107	20,799,492	22,551,353
0.999	22,239,104	22,410,022	24,870,374

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

We can conclude that in each of the two examples the lower bound approximates the "real discounted reserve" very well. The precision of the bounds only depends on the underlying variance of the statistical and financial part. As long as the yearly volatility does not exceed $\sigma = 35\%$, the financial part of the comonotonic approximation provides a very accurate fit. These parameters are consistent with historical capital market values as reported by Ibbotson Associates (2002). The underlying variance of the statistical part depends on the estimated dispersion parameter and error distribution or mean-variance relationship. For example, in case of the gamma distribution one obtains excellent results as long as the dispersion parameter is smaller than 1. This is again in line with the volatility structure in practical IBNR data sets. Since the parameters in the paper for the statistical part of the bounds, obtained through the quasi-likelihood approach, have small standard errors, it follows that results would be similar when simulating from a GLM with the same linear predictor, but for instance with another distribution type. In that sense our findings are robust.

	Distribution of bootstrapped 95th percentiles of S_l	Simulated distribution of $F_S^{-1}(0.95)$
1 st percentile	13,614,404	13,604,314
2.5 th percentile	13,617,028	13,609,425
5 th percentile	13,619,474	13,613,048
10 th percentile	13,622,664	13,618,053
25 th percentile	13,626,759	13,624,369
50 th percentile	13,631,651	13,631,622
75 th percentile	13,636,506	13,638,997
90 th percentile	13,641,168	13,645,812
95 th percentile	13,643,882	13,649,574
97.5 th percentile	13,646,720	13,652,995
99 th percentile	13,648,833	13,656,178

Table 10: Ex. 1 (κ=1): Percentiles of the bootstrapped 95th percentile of the distribution of the lower bound $S^B_{l(95)}$ vs. the simulation.

	Distribution of bootstrapped 95th percentiles of S_l	Simulated distribution of $F_S^{-1}(0.95)$
1 st percentile	16,661,827	16,333,152
2.5 th percentile	16,861,353	16,576,586
5 th percentile	17,048,933	16,759,301
10 th percentile	17,233,865	17,101,271
25 th percentile	17,551,891	17,450,048
50 th percentile	17,913,169	17,904,390
75 th percentile	18,284,619	18,380,651
90 th percentile	18,641,949	18,832,716
95 th percentile	18,850,593	19,117,307
97.5 th percentile	18,999,178	19,264,184
99 th percentile	19,187,288	19,481,477

Table 11: Ex. 2 (κ=2): Percentiles of the bootstrapped 95th percentile of the distribution of the lower bound $S^B_{l(95)}$ vs. the simulation.

Using bootstrap methodology it is possible to provide statistical confidence intervals for the given bounds incorporating the estimation error. The estimation error arises from the estimation of the vector parameters $\hat{\beta}$ from the data, and the statistical error stems from the stochastic nature of the underlying model. We bootstrap an upper triangle using the non-parametric bootstrap procedure. This involves resampling, with replacement, from the original residuals and then creating a new triangle of past claims payments using the resampled residuals together with the fitted values. For a description of the bootstrap technique to claims reserving we refer to Lowe (1994), Taylor (2000) and England and Verrall (2002). These authors used this procedure to obtain prediction errors for different claims reserving methods and also to obtain a predictive distribution of reserves.

For each bootstrap sample, we calculate the desired percentile of the distribution of S_l. This two-step procedure is repeated a large number of times. The first column of Table 10 and Table

11 shows the results, concerning the 95th percentile, for 5000 bootstrap samples applied to the run-off triangle in Table 2, respectively Table 6. When compared with the simulated distribution of $F_S^{-1}(0.95)$ (obtained through 5000 simulated triangles), we can conclude that the bootstrap distribution yields appropriate confidence bounds when applied to the lower bound procedure.

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.

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 process. The use of GLMs offers a great gain in modelling flexibility over the simple lognormal model. 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 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 the financial support of the Onderzoeksfonds K.U.Leuven (GOA/02: Actuariële, financiële en statistische aspecten van afhankelijkheden in verzekerings- en financiële portefeuilles).

Corresponding author

Tom Hoedemakers
Catholic University of Leuven
Department of Applied Economics
Naamsestraat 69, B-3000 Leuven, Belgium

tel: ++32 16 32.67.61, fax: ++32 16 32.67.32
email: tom.hoedemakers@econ.kuleuven.ac.be

References

[1] Barnett, G. and Zehnwirth, B. (1998). Best estimates for reserves. *CAS Forum Fall 1998*, 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. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, **8(3)**, 443-518.

[7] Hoedemakers T., Beirlant J., Goovaerts M.J. and Dhaene J. (2003). Confidence bounds for discounted loss reserves. *Insurance: Mathematics and Economics*, **33(2)**, 297-316.

[8] Ibbotson Associates (2002). Stocks, Bonds, Bills and Inflation: 1926-2001. Chicago, IL.

[9] Kaas, R., Dhaene, J. and Goovaerts, M.J. (2000). Upper and lower bounds for sums of random variables. *Insurance: Mathematics and Economics*, **27(2)**, 151-168.

[10] Lowe, J., 1994. A practical guide to measuring reserve variability using: Bootstrapping, operational time and a distribution free approach. *Proceedings of the 1994 General Insurance Convention*, Institute of Actuaries and Faculty of Actuaries.

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

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

[13] Renshaw, A.E. (1994). Claims reserving by joint modelling. *Actuarial Research Paper No. 72*, Department of Actuarial Science and Statistics, City University, London.

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

[15] Taylor, G.C., 2000. *Loss Reserving: An Actuarial Perspective*. Kluwer Academic Publishers.

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

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

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