KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0263

**CONFIDENCE BOUNDS FOR DISCOUNTED
LOSS RESERVES**

by

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

D/2002/2376/63

Confidence Bounds for Discounted Loss Reserves

Tom Hoedemakers^{*†} Jan Beirlant[‡] Marc J. Goovaerts^{*§} Jan Dhaene^{*§}

Abstract

In this paper we give some methods to set up confidence bounds for the *discounted* IBNR reserve. We start with a loglinear regression model and estimate the parameters by maximum likelihood methods such as e.g. given in Doray, 1996. We want to find an answer to the following question: "What is the amount we have to reserve at time 0 in order to be able to meet our future obligations?" The knowledge of the distribution function of the discounted IBNR reserve (S) will help us to determine the initial reserve V_0 . E.g. we could determine this reserve as $V_0 = F_S^{-1}(0.95)$, or $V_0 = E[S|S > F_S^{-1}(0.95)]$, to assume with 95% confidence that we can meet our future obligations. The results are based on convex order techniques, such that our approximations for the distribution function of S are larger or smaller, in convex order sense, than the "true distribution function of S ".

Keywords: IBNR, confidence bound, comonotonicity, bootstrap, simulation.

^{*}Department of Applied Economics, K.U.Leuven, Naamsestraat 69, 3000 Leuven, Belgium

[†]Corresponding author. E-mail adress: Tom.Hoedemakers@econ.kuleuven.ac.be

[‡]Department of Mathematics, K.U.Leuven, Celestijnenlaan 200B, 3001 Heverlee, Belgium

[§]Institute of Actuarial Science, University of Amsterdam, Roeterstraat 11, 1018 WB Amsterdam, Netherlands

1 Introduction

An important problem in insurance is to determine, at the end of an insurance period, the provision for claims already incurred but not yet reported (hence IBNR), or not fully paid. The past data used to construct estimates for the future payments consist of a triangle of incremental claims Y_{ij} , as depicted in Figure 1. This is the simplest shape of data that can be obtained and it avoids us having to introduce complicated notation to cope with all possible situations.

The random variables Y_{ij} for $i, j = 1, 2, \dots, t$ denote the claim figures for year of origin 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 explanatory variables for the observation Y_{ij} . For (i, j) combinations with $i + j \leq t + 1$, Y_{ij} has already been observed, otherwise it is a future observation. As well as claims actually paid, these figures can also be used to denote quantities such as loss ratios. To a large extent, it is irrelevant whether incremental or cumulative data are used when considering claims reserving in a stochastic context.

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.

The statistical model that will be used, in this paper, to describe the past and the future payments is a loglinear model. So, the total IBNR reserve will be a sum of lognormal random variables which implies that its exact distribution function (d.f.) cannot be determined analytically. If we consider now the *discounted* IBNR reserve (S), we have to incorporate a certain dependence structure. We will come back to this later on. 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.

The essential problem to be solved is the management of the risk associated with the future payments contained in the lower triangle. Most methods estimate the lower triangle cell-by-cell, and do not pay (enough) attention to the structure describing the dependencies between these cells. Each cell must be considered as a univariate random variable being part of the multivariate random variable describing the lower triangle. Hence, the IBNR reserve must be considered as a (univariate) random variable being the sum of the dependent components of the random vector describing the lower triangle. Estimating the correlations from the past data, and using them for multivariate simulations of the lower triangle is a dangerous technique because the insurer is especially interested in the tail of the distribution function by choosing the reserve as a percentile. In practice, the

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

Figure 1: Random variables in a run-off triangle

insurer will choose a very high percentile as basis for his reserve. Hence, only very high time-consuming multivariate simulations will lead to a sufficient number of simulated values in such an extreme tail. Of course, a multivariate simulation technique will only be possible if the whole dependence structure of the lower triangle is known. In practice, we encounter situations where only the distribution of each cell can be estimated with enough accuracy, but where only limited information of the dependence structure can be obtained (mostly due to the fact that not enough data are available). We hence conclude that a multivariate simulation technique is not the most appropriate way to determine IBNR reserves.

As mentioned above, the "true" multivariate distribution function of the lower triangle cannot be determined in most cases, because the mutual dependencies are not known, or difficult to cope with. The only conceivable solution is to find upper and lower bounds for this sum of dependent random variables making efficient use of the available information. Hence, within a certain class of random vectors (with given marginals, and eventually additionally information), we propose to look for upper and lower bounds for the sum of the cells of the lower triangle. In this paper, we develop approximations for the distribution function of the discounted IBNR reserve, by deriving accurate bounds. For these bounds we follow an approach that is 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, Dhaene and Goovaerts (2000). The first approximation we will consider for the d.f. 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 takes into account part of the dependence structure. It is derived by considering conditional expectations. We will include a numerical comparison of our approximations with a simulation study. The second approximation turns out to perform very well. For details of this technique we refer to Dhaene, Denuit, Goovaerts, Kaas and Vyncke (2002a,b) and the references therein.

As stated above, the goal is to obtain upper limits to the forecasts and to associate a

confidence level to those limits. It is necessary to be able to estimate the variability of claims reserves, and ideally to be able to estimate a full distribution of possible outcomes, from which percentiles (or other measures) of that distribution can be obtained. Stochastic claims reserving methods extend traditional techniques to allow those additional measures to be estimated. It is also important to emphasize that this paper proposes a method to calculate the confidence bounds of the IBNR distribution given a specific underlying statistical model, here the lognormal regression model. The choice of a good model is another separate problem. In fact, it is the starting point of the analysis. Therefore, the data set should be examined in detail in order to find an appropriate model, rather than using the same modelling approach in all circumstances. It is important to test different models. Different stochastic methods will give different results. Once within a stochastic framework, there is considerable flexibility in the choice of predictor structures, with predictors including continuous covariates offering the possibility of extrapolation beyond the range of data observed. Once within a statistical framework, diagnostic checks of the fitted models are possible, such as goodness-of-fit tests and analysis of residuals (which highlight systematic and isolated departures from the fitted model). In England and Verrall (2002) the reader finds an excellent review of the stochastic models which have already been suggested. Of course, there will also be cases where none of the approaches set out will be appropriate, but most of the models have a wide applicability. An appropriate model will enable the calculation of the distribution of the reserve that reflects both the process variability producing the future payments and the parameter estimation error (parameter or statistical uncertainty).

This paper is set out as follows. Section 2 gives a summary of the use of loglinear models in claims reserving. Doray (1996) studied these models extensively, taking into account the estimation error on the parameters and the statistical prediction error in the model. He has derived various estimators for the IBNR reserve under the independence assumption, among them the maximum likelihood estimator and the uniformly minimum variance unbiased estimator (UMVUE), as well as an expression for the variance of the latter estimator. The variance of the IBNR reserve is also calculated. The joint distribution of the amounts in each cell of the lower triangle is shown to follow a multivariate lognormal distribution. Section 3 derives 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. We will describe how these results can be used in the IBNR evaluations. Finally, we will calculate the cdf's of these bounds. In section 4 we explain how we can use the bootstrap technique, as presented in England and Verrall (1999) and in Pinheiro, Andrade e Silva and Centeno (2002), in order to obtain confidence bounds for the predictive distribution of the total reserve. Section 5 gives some numerical illustrations for a simulated data set. We also illustrate the obtained bounds graphically.

2 Loglinear Models

We consider the following loglinear regression model

$$\vec{Z}_i = \ln \vec{Y}_i = \mathbf{X}_i \vec{\beta} + \vec{\epsilon}_i, \quad \vec{Y}_i > 0 \quad (1)$$

where

- $\vec{Y}_i$ is the i th element of the datavector $\vec{Y}$, of dimension $\frac{t(t+1)}{2}$,
- $\mathbf{X}$ is the regression matrix of dimension $[\frac{t(t+1)}{2}] \times p$; the i th row is denoted by $\mathbf{X}_i$, and element (i, j) is denoted X_{ij} ,
- $\vec{\beta}$ is the vector (of dimension p) of unknown parameters,
- $\vec{\epsilon}_i$ are independent normal random errors with mean 0 and variance σ^2 .

In matrix notation this linear model can be represented as

$$\vec{Z} = \ln \vec{Y} = \mathbf{X} \vec{\beta} + \vec{\epsilon}, \quad \vec{\epsilon} \sim MN(0, \sigma^2 \mathbf{I}) \quad (2)$$

The use of the logarithmic transform immediately imposes a limitation on this class of models, in that incremental claim amounts must be positive. The normal responses Z_{ij} are assumed to decompose (additively) into a deterministic non-random component with mean $(\mathbf{X}\vec{\beta})_{ij}$ and a homoscedastic normally distributed random error component with zero mean.

For the regression parameters, various choices are possible. A well-known and widely used model is the stochastic chain-ladder model

•

$$Z_{ij} = \ln Y_{ij} = \alpha_i + \beta_j + \epsilon_{ij}, \quad i + j \leq t + 1 \quad (3)$$

(α is the parameter for each year of origin i and, β for each development year j). It should be noted that this representation implies the same development pattern for all years of origin, where that pattern is defined by the parameters β_j . The relationship between this loglinear model and the chain-ladder technique was first pointed out by Kremer (1982) and used by Renshaw (1989), Verall (1989) and Christofides (1990), amongst others. Use of this model produces predicted values close, but not identical, to those from the simple chain ladder technique.

For a general model with parameters in the three directions, we refer to De Vylder F. and Goovaerts M.J. (eds.) (1979). We give here some frequently used special cases

•

$$Z_{ij} = \ln Y_{ij} = \alpha_i + \sum_{k=1}^{j-1} \beta_k + \sum_{t=1}^{i+j-2} \gamma_t + \epsilon_{ij}, \quad i + j \leq t + 1 \quad (4)$$

for the probabilistic trend family (PTF) of models as suggested in Barnett and Zehnwirth (1998) (γ denotes the calendar year effect; it combines the effects of monetary inflation and changing jurisprudence).

•

$$Z_{ij} = \ln Y_{ij} = \alpha_i + \beta_i \log(j) + \gamma_i j + \epsilon_{ij}, \quad (j > 0) \quad i + j \leq t + 1 \quad (5)$$

for the Hoerl curve as in Zehnwirth (1985). A special case is created by setting $\beta_i = \beta$ and $\gamma_i = \gamma$ for all i , where the decay pattern is the same for all accident years represented by only two parameters. This imposes a strict parametric form on the shape of the run-off triangle. Although this sacrifices goodness-of-fit, it has the advantage that payments can be predicted by extrapolation beyond the range of the observed j .

•

$$Z_{ij} = \ln Y_{ij} = \begin{cases} \alpha_i + \beta_j + \epsilon_{ij} & \text{if } j \leq q; \\ \alpha_i + \beta_i \log(j) + \gamma_i j + \epsilon_{ij} & \text{if } j > q. \end{cases} \quad i + j \leq t + 1 \quad (6)$$

for some integer q specified by the modeller. This is a mixture of cases 1 and 3 as in England and Verrall (2001).

The parameters are estimated by maximum likelihood, which in the case of the normal error structure is equivalent to minimizing the residual sum of squares. The unknown variance σ^2 is estimated by the residual sum of squares divided by the degrees of freedom (the number of observations minus the numbers of parameters estimated)

$$\tilde{\sigma}^2 = \frac{1}{n-p} (\vec{Z} - \mathbf{X} \hat{\vec{\beta}})' (\vec{Z} - \mathbf{X} \hat{\vec{\beta}}), \quad (7)$$

where $\tilde{\sigma}^2$ is an unbiased estimator of σ^2 . The residual sum of squares divided by the number of observations is the maximum likelihood estimator of σ^2

$$\hat{\sigma}^2 = \frac{1}{n} (\vec{Z} - \mathbf{X} \hat{\vec{\beta}})' (\vec{Z} - \mathbf{X} \hat{\vec{\beta}}) \quad (8)$$

The maximum likelihood estimator of $\vec{\beta}$ is

$$\hat{\vec{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\vec{Z}. \quad (9)$$

Let $\mathbf{B}$ be the regression matrix corresponding to the lower triangle, of dimension $\lceil \frac{t(t-1)}{2} \rceil \times p$. (defined analogously to the regression matrix $\mathbf{X}$).

Now we can forecast the total IBNR reserve as follows:

$$\text{IBNR reserve} = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} + \epsilon_{ij}} \quad (10)$$

This definition of the IBNR reserve is taken from Doray (1996), among others. Remark that another definition could be

$$\text{IBNR reserve} = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} + \epsilon_{ij}} \quad (11)$$

The approach taken in (10) partly uses the information contained in the upper triangle (through $\tilde{\beta}$), and acknowledges the underlying stochastic structure (through ϵ_{ij}).

We have that

$$\epsilon_{ij} \sim \text{i.i.d } N(0, \sigma^2) \quad (12)$$

$$(\mathbf{B}\tilde{\beta})_{ij} \sim N((\mathbf{B}\tilde{\beta})_{ij}, \sigma^2 (\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij}) \quad (13)$$

Starting from model (1), we summarize now some properties of the IBNR reserve, taken from Doray (1996).

1. The mean of the IBNR reserve:

$$W = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} + \frac{1}{2}\sigma^2(1 + (\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij})} \quad (14)$$

2. The unique UMVUE of the mean of the IBNR reserve:

$$\hat{W}_U = {}_0F_1\left(\frac{n-p}{2}; \frac{SS_z}{4}\right) \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij}}, \quad (15)$$

where ${}_0F_1(\alpha; z)$ denotes the hypergeometric function.

3. The MLE of the mean of the IBNR reserve:

$$\hat{W}_M = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} + \frac{1}{2}\hat{\sigma}^2(1 + (\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij})} \quad (16)$$

4. Verall (1991) has considered an estimator similar to $\hat{W}_M$, but with $\hat{\sigma}^2$ replaced with $\hat{\sigma}^2$.

$$\hat{W}_V = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\hat{\beta})_{ij} + \frac{1}{2}\hat{\sigma}^2(1 + (\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij})} \quad (17)$$

5. Doray (1996) has derived an estimator of the mean of the IBNR reserve:

$$\hat{W}_D = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\hat{\beta})_{ij} + \frac{1}{2}\hat{\sigma}^2} \quad (18)$$

Now we have the following order relation:

$$\hat{W}_U < \hat{W}_D < \hat{W}_V, \quad (19)$$

which implies that

$$W = E[\hat{W}_U] < E[\hat{W}_D] < E[\hat{W}_V] \quad (20)$$

Hence both the estimators $\hat{W}_D$ and $\hat{W}_V$ exhibit a positive bias.

3 Methodology

3.1 Comonotonicity, stop-loss order and convex order

Definition 1 The Fréchet space $R_n(F_1, F_2, \dots, F_n)$ determined by the (univariate) distribution functions $F_1, F_2, \dots, F_n$ is the class of all n -variate distribution functions F (or the corresponding random variables) with marginals $F_1, F_2, \dots, F_n$.

In the Fréchet space $R_n(F_1, F_2, \dots, F_n)$ any random variable $\mathbf{X}$ is constrained from above by

$$F_{\mathbf{X}}(\mathbf{x}) \leq \min\{F_1(x_1), F_2(x_2), \dots, F_n(x_n)\} =: W_n(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R} \quad (21)$$

A comonotone risk is a random variable with cdf W_n , see e.g. Dhaene et al (1997).

Definition 2 A random vector $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is said to be comonotone (the random variables $X_1, X_2, \dots, X_n$ are said to be mutually comonotone) if any of the following conditions hold:

1. For the n -variate cdf we have

$$F_{\mathbf{X}}(\mathbf{x}) = \min\{F_1(x_1), F_2(x_2), \dots, F_n(x_n)\}, \quad \forall \mathbf{x} \in \mathbb{R}^n; \quad (22)$$

2. There exist a random variable Z and non-decreasing functions $g_1, g_2, \dots, g_n: \mathbb{R} \rightarrow \mathbb{R}$ such that

$$(X_1, X_2, \dots, X_n) \stackrel{d}{=} (g_1(Z), g_2(Z), \dots, g_n(Z)); \quad (23)$$

3. For any random variable U uniformly distributed on $(0, 1)$, we have:

$$(X_1, X_2, \dots, X_n) \stackrel{d}{=} (F_1^{-1}(U), F_2^{-1}(U), \dots, F_n^{-1}(U)). \quad (24)$$

As usual " $\stackrel{d}{=}$ " denotes equality in distribution and F^{-1} represents the inverse of the cdf F defined as

$$F_X^{-1}(p) = \inf\{x \in \mathbb{R} \mid F_X(x) \geq p\}, \quad p \in [0, 1]. \quad (25)$$

It can be seen from condition 2 that comonotonic random variables possess a very strong positive dependence: increasing one of the X_i will lead to an increase of all other random variables X_j involved. A comonotone risk is a random variable with cdf W_n , see e.g. Dhaene et al (1997).

Definition 3 Consider two random variables X and Y . Then X is said to precede Y in the stop-loss order sense, notation $X \leq_{sl} Y$, if and only if X has lower stop-loss premiums than Y :

$$E[(X - d)_+] \leq E[(Y - d)_+], \quad \forall d \quad (26)$$

Definition 4 Consider two random variables X and Y . Then X is said to precede Y in the convex order sense, notation $X \leq_{cx} Y$, if and only if

$$\begin{aligned} E[X] &= E[Y] \\ E[(X - d)_+] &\leq E[(Y - d)_+], \quad \forall d \end{aligned} \quad (27)$$

From Definition 2 we have that a random variable V precedes a random variable W in the $\leq_{cx}$ -sense, if and only if, the following condition holds:

$$E[v(V)] \leq E[v(W)] \quad \text{for all convex functions } v: \mathbb{R} \rightarrow \mathbb{R} \quad (28)$$

provided the expectations exist.

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.

In this paper, we will consider loss variables V for which the distribution function cannot be determined explicitly. We will construct a new random variable W which is larger in the convex order sense, meaning that $E[V] = E[W]$, and such that for each retention d , the stop-loss premium $E[(V - d)_+]$ is smaller than or equal to the corresponding stop-loss premium of W . Replacement of the loss V by the loss W is safe in the sense that all risk aversion decision makers will consider W as a less favorable loss. Practically, of course, applying the technique of replacing a loss by a less favorable loss will only make sense if the new loss variable has a simpler dependence structure, making it easier to determine its distribution function.

3.2 Convex bounds for sums of random variables

Proposition 1 *Let U be a uniform(0,1) random variable. For any random vector $(X_1, X_2, \dots, X_n)$ with marginal cdf's $F_{X_1}, F_{X_2}, \dots, F_{X_n}$, we have*

$$X_1 + X_2 + \dots + X_n \leq_{cx} F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U) + \dots + F_{X_n}^{-1}(U). \quad (29)$$

Proof. See e.g. [16].

Now we can generalize this proposition to a convex upper bound for $\sum_{i=1}^n \phi_i(X_i)$. Let $\phi_1, \dots, \phi_n$ be real valued functions. If we take the same assumptions into account as in Proposition 1, we find that

$$\phi_1(X_1) + \dots + \phi_n(X_n) \leq_{cx} F_{\phi_1(X_1)}^{-1}(U) + \dots + F_{\phi_n(X_n)}^{-1}(U) \quad (30)$$

This is because $(\phi_1(X_1), \dots, \phi_n(X_n))$ and $(F_{\phi_1(X_1)}^{-1}(U), \dots, F_{\phi_n(X_n)}^{-1}(U))$ belong to the same Fréchet space, while the latter random vector is comonotonic.

When all the functions ϕ_i are continuous and non-decreasing (or all are non-increasing), then we have that

$$F_{\phi_1(X_1)}^{-1}(U) + \dots + F_{\phi_n(X_n)}^{-1}(U) \stackrel{d}{=} \phi_1(F_{X_1}^{-1}(U)) + \dots + \phi_n(F_{X_n}^{-1}(U)) \quad (31)$$

so we find the following stochastic inequality:

$$\phi_1(X_1) + \dots + \phi_n(X_n) \leq_{cx} \phi_1(F_{X_1}^{-1}(U)) + \dots + \phi_n(F_{X_n}^{-1}(U)) \quad (32)$$

Consider now the random variable $S \stackrel{def}{=} X_1Y_1 + \dots + X_nY_n$, where X_i ($i = 1, \dots, n$) could e.g. represent a stochastic cash flow at time i and Y_i ($i = 1, \dots, n$) represents the stochastic discount factor for a payment made at time i . Hence, the random variable $X_1Y_1 + \dots + X_nY_n$ could e.g. represent the present value at time 0, of a sequence of random payments to be made at times 1, 2, $\dots$, n . In general, the random payments $(X_1, \dots, X_n)$ will not be mutually independent, but even if they were, the discount factors will in general not be mutually independent.

Application of the previous Proposition leads to

$$X_1Y_1 + \cdots + X_nY_n \leq_{cx} F_{X_1Y_1}^{-1}(U) + \cdots + F_{X_nY_n}^{-1}(U). \quad (33)$$

Several approximations for the distribution function of sums of dependent random variables based on the concept 'comonotonicity' have been described in the actuarial literature. The concept of "comonotonicity" describes a strong positive dependency between random variables. For an overview of applications of comonotonicity in the field of actuarial sciences, see Dhaene, Denuit, Goovaerts, Kaas and Vyncke (2002a,b) and the references therein.

3.3 Scalar products of independent random vectors

In this subsection we will generalize the main results from Kaas, Dhaene and Goovaerts (2000) for scalar products of independent random vectors.

Proposition 2 *Assume that the vectors $\mathbf{X}$ and $\mathbf{Y}$ are mutually independent and also the vectors $\mathbf{X}$ and $\mathbf{Z}$. If*

$$\sum_{i=1}^n x_i Y_i \leq_{cx} \sum_{i=1}^n x_i Z_i \quad (34)$$

for all outcomes $\mathbf{x}$ of $\mathbf{X}$, then

$$\sum_{i=1}^n X_i Y_i \leq_{cx} \sum_{i=1}^n X_i Z_i. \quad (35)$$

Proof. Let ϕ be a convex function. By conditioning on $\mathbf{X}$ and taking the assumptions into account, we find that

$$E \left[\phi \left(\sum_{i=1}^n X_i Y_i \right) \right] = E_{\mathbf{X}} \left\{ E \left[\phi \left(\sum_{i=1}^n X_i Y_i \right) | \mathbf{X} \right] \right\} \quad (36)$$

$$\leq E_{\mathbf{X}} \left\{ E \left[\phi \left(\sum_{i=1}^n X_i Z_i \right) | \mathbf{X} \right] \right\} \quad (37)$$

$$= E \left[\phi \left(\sum_{i=1}^n X_i Z_i \right) \right] \quad (38)$$

holds for any convex function ϕ . ■

3.3.1 A convex upper bound for $\sum_{i=1}^n X_i Y_i$

The next theorem states a convex order - upper bound that is sharper than (33) if we assume that the vectors $\mathbf{X}$ and $\mathbf{Y}$ are mutually independent and if the X_i and Y_i are non-negative random variables.

Theorem 1 *Let U and V be mutually independent uniform(0,1) random variables. Assume that the vectors $\mathbf{X}$ and $\mathbf{Y}$ are mutually independent and that the X_i and Y_i are non-negative random variables. Then*

$$\sum_{i=1}^n X_i Y_i \leq_{cx} \sum_{i=1}^n F_{X_i}^{-1}(U) F_{Y_i}^{-1}(V) \leq_{cx} \sum_{i=1}^n F_{X_i Y_i}^{-1}(U). \quad (39)$$

Proof. From Kaas, Dhaene and Goovaerts (2000)

$$\sum_{i=1}^n X_i Y_i \leq_{cx} \sum_{i=1}^n F_{X_i Y_i}^{-1}(U), \quad (40)$$

where U is a uniformly distributed random variable. Because $F_{\alpha X}^{-1}(p) = \alpha F_X^{-1}(p)$ for α positive, we have that

$$\sum_{i=1}^n F_{X_i Y_i}^{-1}(U) = \sum_{i=1}^n Y_i F_{X_i}^{-1}(U) \quad (41)$$

From the previous Proposition

$$\sum_{i=1}^n X_i Y_i \leq_{cx} \sum_{i=1}^n Y_i F_{X_i}^{-1}(U), \quad (42)$$

if U is independent of $\mathbf{Y}$ (*).

Proceeding the same for $\sum_{i=1}^n Y_i F_{X_i}^{-1}(U)$ proves the first inequality of the theorem. Note that the independence assumption (*) can be omitted by replacing e.g. U by U' . The second inequality follows from (33) and because $\sum_{i=1}^n F_{X_i Y_i}^{-1}(U)$ is an extremal element/distribution in the Fréchet space. ■

Generalizations

The results of the previous subsections can be generalized in order to derive upper bounds (in terms of convex order) for sums of products of three or more random variables.

Let us assume that the vectors $\mathbf{X}$ and $\mathbf{Z}$ are independent and also the vectors $\mathbf{Y}$ and $\mathbf{Z}$. If the X_i , Y_i and Z_i are non-negative random variables, then we find:

$$\sum_{i=1}^n X_i Y_i Z_i \leq_{cx} \sum_{i=1}^n F_{X_i Y_i}^{-1}(U) F_{Z_i}^{-1}(V) \quad (43)$$

where U and V are mutually independent uniformly distributed random variables.

3.3.2 An improved upper bound for $\sum_{i=1}^n X_i Y_i$

Let us now assume that we have complete (or partial) information concerning the dependence structure of the random vector $(X_1 Y_1, X_2 Y_2, \dots, X_n Y_n)$, but that exact computation of the cdf of the sum $X_1 Y_1 + X_2 Y_2 + \dots + X_n Y_n$ is very time-consuming or even impossible. In this case we can, derive improved upper bounds for S by using part of the information on the dependence structure, by conditioning on some random variable Z which is assumed to be some function of the random vector $\mathbf{X}$. We will assume that we know the conditional cdf's, given $Z = z$, of the random variables X_i .

Theorem 2 *Assume that the vectors $\mathbf{X}$ and $\mathbf{Y}$ are mutually independent and that the X_i and Y_i are non-negative random variables. Consider a random variable Z and two mutually independent uniform(0,1) random variables U and V . Then we have*

$$S \stackrel{\text{def}}{=} \sum_{i=1}^n X_i Y_i \leq_{cx} S'_u \stackrel{\text{def}}{=} \sum_{i=1}^n F_{X_i|Z}^{-1}(U) F_{Y_i}^{-1}(V) \leq_{cx} S_u \stackrel{\text{def}}{=} \sum_{i=1}^n F_{X_i}^{-1}(U) F_{Y_i}^{-1}(V) \quad (44)$$

where U , V and Z are mutually independent.

Proof. We have that

$$\sum_{i=1}^n X_i Y_i \leq_{cx} \sum_{i=1}^n F_{X_i Y_i|Z}^{-1}(U) = \sum_{i=1}^n Y_i F_{X_i|Z}^{-1}(U). \quad (45)$$

Proceeding now the same as in the previous theorem, proves the first inequality.

Now, we have that the random vector $(F_{X_1|Z}^{-1}(U), \dots, F_{X_n|Z}^{-1}(U))$ has marginals $F_{X_1}, \dots, F_{X_n}$, because

$$F_{X_i}(x) = P(X_i \leq x) \quad (46)$$

$$= \int_{-\infty}^{\infty} P(X_i \leq x | Z = z) dF_Z(z) \quad (47)$$

$$= \int_{-\infty}^{\infty} P(F_{X_i|Z=z}^{-1}(U) \leq x) dF_Z(z) \quad (48)$$

$$= F_{F_{X_i|Z}^{-1}(U)}(x). \quad (49)$$

In view of Proposition 1 this implies the second inequality. ■

3.3.3 A lower bound for $\sum_{i=1}^n X_i Y_i$

Let $\mathbf{X}$ be a random vector with marginals $F_{X_1}, \dots, F_{X_n}$, and assume that we want to find a lower bound, in the sense of convex order, for $S = X_1 Y_1 + \dots + X_n Y_n$. We can obtain such a bound by conditioning on some random variable Z , again assumed to be a function of the random vector $\mathbf{X}$.

Theorem 3 Assume that the vectors $\mathbf{X}$ and $\mathbf{Y}$, given the random variable Z , are mutually independent and that Z is independent of $\mathbf{Y}$. Then we have

$$S_l \stackrel{\text{def}}{=} E[X_1|Z]E[Y_1] + \cdots + E[X_n|Z]E[Y_n] \leq_{cx} X_1Y_1 + \cdots + X_nY_n \quad (50)$$

Proof. By Jensen's inequality, we find that for any convex function ϕ , the following inequality holds:

$$E[\phi(X_1Y_1 + \cdots + X_nY_n)] = E_Z E[\phi(X_1Y_1 + \cdots + X_nY_n)|Z] \quad (51)$$

$$\geq E_Z[\phi(E[X_1Y_1 + \cdots + X_nY_n|Z])] \quad (52)$$

$$= E_Z[\phi(E[X_1Y_1|Z] + \cdots + E[X_nY_n|Z])] \quad (53)$$

$$= E_Z[\phi(E[X_1|Z]E[Y_1] + \cdots + E[X_n|Z]E[Y_n])] \quad (54)$$

This proves the stated result. ■

3.4 Upper and lower bounds for the discounted IBNR reserve

Recall that

$$\text{IBNR reserve} = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\hat{\beta})_{ij} + \epsilon_{ij}} \quad (55)$$

In case the type of business allows for discounting, or in case 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 , $j = 1, 2, \dots, t-1$, 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 $(Y_1, Y_2, \dots, Y_{t-1})$ has a multivariate normal distribution. The present value of the payments is then a linear combination of dependent lognormal random variables. Setting

$$Y(i) = Y_1 + Y_2 + \cdots + Y_i, \quad (56)$$

and

$$Y(i) = (\mu + \frac{\delta^2}{2})i + \delta B(i), \quad (57)$$

where $B(i)$ is the standard Brownian motion and where μ 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

$$S \stackrel{def}{=} \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\hat{\vec{\beta}})_{ij} - Y(i+j-t-1) + \epsilon_{ij}} \quad (58)$$

$$= \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left((\mathbf{B}\hat{\vec{\beta}})_{ij} - (\mu + \delta^2/2)(i+j-t-1) - \delta B(i+j-t-1) + \epsilon_{ij} \right) \quad (59)$$

Now we have that

$$E[e^{-\mu i - (\delta^2/2)i - \delta B(i)}] = e^{-\mu i} \quad (60)$$

We introduce the random variables V_{ij} and W_{ij} defined by

$$W_{ij} = (\mathbf{B}\hat{\vec{\beta}})_{ij} - Y(i+j-t-1); \quad V_{ij} = e^{W_{ij}} \quad (61)$$

Consider now a conditioning normally distributed random variable Z defined as follows:

$$Z = \sum_{i=2}^t \sum_{j=t+2-i}^t \nu_{ij} Y_{i+j-t-1} \quad (62)$$

We will compute the lower and upper bound for the following choice of the parameters

$$\begin{aligned} \nu_{ij} = & \sum_{k=i+1}^t \sum_{l=t+2-k}^t \exp \left((\mathbf{B}\vec{\beta})_{kl} - (k+l-t-1)(\mu + \frac{1}{2}\delta^2) \right) \\ & + \sum_{l=j}^t \exp \left((\mathbf{B}\vec{\beta})_{il} - (i+l-t-1)(\mu + \frac{1}{2}\delta^2) \right) \end{aligned} \quad (63)$$

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

$$\tilde{S} = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\vec{\beta})_{ij} - Y(i+j-t-1)} \quad (64)$$

(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

$$E[W_{ij}|Z = z] = E[W_{ij}] + \rho_{ij} \frac{\sigma_{W_{ij}}}{\sigma_Z} (z - E[Z]) \quad (65)$$

and

$$\text{Var}[W_{ij}|Z = z] = \sigma_{W_{ij}}^2 (1 - \rho_{ij}^2) \quad (66)$$

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

Proposition 3 *Let S, S_l, S'_u and S_u be defined as follows:*

$$S = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp(W_{ij} + \epsilon_{ij}), \quad (67)$$

$$S_l = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp\left(E[W_{ij}] + \rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U) + \frac{1}{2}(1 - \rho_{ij}^2)\sigma_{W_{ij}}^2 + \frac{1}{2}\sigma_{\epsilon_{ij}}^2\right), \quad (68)$$

$$S'_u = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp\left(E[W_{ij}] + \rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U) + \sqrt{1 - \rho_{ij}^2}\sigma_{W_{ij}}\Phi^{-1}(V) + \sigma_{\epsilon_{ij}}\Phi^{-1}(W)\right), \quad (69)$$

$$S_u = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp\left(E[W_{ij}] + \sigma_{W_{ij}}\Phi^{-1}(U) + \sigma_{\epsilon_{ij}}\Phi^{-1}(V)\right), \quad (70)$$

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

$$S_l \leq_{cx} S \leq_{cx} S'_u \leq_{cx} S_u. \quad (71)$$

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 $U = \Phi\left(\frac{Z - E[Z]}{\sigma_Z}\right) \sim \text{uniform}(0, 1)$, that

$$E[V_{ij}|Z]E[e^{\epsilon_{ij}}] = \exp\left(E[W_{ij}] + \rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U) + \frac{1}{2}(1 - \rho_{ij}^2)\sigma_{W_{ij}}^2 + \frac{1}{2}\sigma_{\epsilon_{ij}}^2\right). \quad (72)$$

From Theorem 3, we find $S_l \leq_{cx} S$.

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

$$F_{V_{ij}|Z}^{-1}(p)F_{\epsilon_{ij}}^{-1}(q) = \exp\left(E[W_{ij}] + \rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U) + \sqrt{1 - \rho_{ij}^2}\sigma_{W_{ij}}\Phi^{-1}(p) + \sigma_{\epsilon_{ij}}\Phi^{-1}(q)\right). \quad (73)$$

From Theorem 2, we find $S \leq_{cx} S'_u$.

3. The stochastic inequality $S'_u \leq_{cx} S_u$ follows from Theorem 2. ■

It remains to derive expressions for the cdfs of S_l, S'_u and S_u .

3.5 Distributions

In this section we provide some more details concerning the (calculation of the) distributions of the different bounds derived in Proposition 2.

Consider a random variable W which is defined as the product of two non-negative independent variables X and Y :

$$W = XY \quad (74)$$

The cdf of W follows from

$$F_W(z) = \int_{-\infty}^{\infty} F_Y\left(\frac{z}{x}\right) dF_X(x) = \int_0^1 F_Y\left(\frac{z}{F_X^{-1}(u)}\right) du \quad (75)$$

3.5.1 The cdf of S_u

From Theorem 2, we can write the convex upper bound for the discounted IBNR reserve as follows

$$S_u = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp\left(F_{(\mathbf{B}\tilde{\beta})_{ij}-Y(i+j-t-1)}^{-1}(U)\right) \exp\left(F_{\epsilon_{ij}}^{-1}(V)\right) \quad (76)$$

$$= e^{\sigma\Phi^{-1}(V)} \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} - (\mu + \frac{1}{2}\delta^2)(i+j-t-1) + \sqrt{\sigma^2(\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij} + \delta^2(i+j-t-1)}\Phi^{-1}(U)}. \quad (77)$$

There are several possibilities to derive the cdf of S_u . From previous results

$$F_{S_u}(z) = \int_0^1 F_N\left(\ln(z) - \ln(F_{S_u''}^{-1}(u))\right) du \quad (78)$$

with $F_N(x)$ the cdf of $N(0, \sigma^2)$ and

$$S_u'' = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp\left(F_{(\mathbf{B}\tilde{\beta})_{ij}-Y(i+j-t-1)}^{-1}(U)\right) \quad (79)$$

$$= \sum_{i=2}^t \sum_{j=t+2-i}^t \exp\left(E[W_{ij}] + \sigma_{W_{ij}}\Phi^{-1}(U)\right). \quad (80)$$

So, we have that $F_{S_u}(z) =$

$$\int_0^1 F_N\left(\ln(z) - \ln\left(\sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} - (\mu + \frac{1}{2}\delta^2)(i+j-t-1) + \sqrt{\sigma^2(\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij} + \delta^2(i+j-t-1)}\Phi^{-1}(u)}\right)\right) du. \quad (81)$$

We can also derive an algorithm for the determination of the cdf of S_u . We have that

$S_u \mid V = v$ is the sum of $t(t-1)/2$ comonotonic risks. This implies

$$F_{S_u|V=v}^{-1}(p) = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(F_{(\mathbf{B}\hat{\beta})_{ij}-Y(i+j-t-1)}^{-1}(p) \right) \exp \left(F_{\epsilon_{ij}}^{-1}(v) \right) \quad (82)$$

and

$$F_{S_u|V=v}(x) = \left\{ p \in [0, 1] \mid \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(F_{(\mathbf{B}\hat{\beta})_{ij}-Y(i+j-t-1)}^{-1}(p) \right) \exp \left(F_{\epsilon_{ij}}^{-1}(v) \right) \leq x \right\}. \quad (83)$$

If we in addition assume that the cdf's $F_{(\mathbf{B}\hat{\beta})_{ij}-Y(i+j-t-1)}$ are strictly increasing and continuous, then $F_{S_u|V=v}(x)$ follows from

$$\sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(F_{(\mathbf{B}\hat{\beta})_{ij}-Y(i+j-t-1)}^{-1}(F_{S_u|V=v}(x)) \right) \exp \left(F_{\epsilon_{ij}}^{-1}(v) \right) = x. \quad (84)$$

In any case, the cdf of S_u follows from

$$F_{S_u}(x) = \int_0^1 F_{S_u|V=v}(x) dv. \quad (85)$$

3.5.2 The cdf of S_l

From Theorem 3, we obtained the convex lower bound for the discounted IBNR reserve from

$$S_l = \sum_{i=2}^t \sum_{j=t+2-i}^t E \left[\exp \left((\mathbf{B}\hat{\beta})_{ij} - Y(i+j-t-1) \right) \middle| \sum_{i=2}^t \sum_{j=t+2-i}^t \nu_{ij} Y_{i+j-t-1} \right] E[\exp(\epsilon_{ij})]. \quad (86)$$

Taking into account that $Z = \sum_{i=2}^t \sum_{j=t+2-i}^t \nu_{ij} Y_{i+j-t-1}$ is normally distributed, we find that

$$F_Z^{-1}(1-p) = E[Z] - \sigma_Z \Phi^{-1}(p), \quad (87)$$

and hence

$$F_{s_l}^{-1}(p) = F_{\sum_{i=2}^t \sum_{j=t+2-i}^t E[V_{ij}|Z]E[e^{\epsilon_{ij}}]}^{-1}(p), \quad p \in (0, 1) \quad (88)$$

$$= \sum_{i=2}^t \sum_{j=t+2-i}^t F_{E[V_{ij}|Z]E[e^{\epsilon_{ij}}]}^{-1}(p) \quad (89)$$

$$= \sum_{i=2}^t \sum_{j=t+2-i}^t E[V_{ij}|Z = F_Z^{-1}(1-p)] E[e^{\epsilon_{ij}}] \quad (90)$$

$$= \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(E[W_{ij}] - \rho_{ij} \sigma_{W_{ij}} \Phi^{-1}(p) + \frac{1}{2} (1 - \rho_{ij}^2) \sigma_{W_{ij}}^2 + \frac{1}{2} \sigma_{\epsilon_{ij}}^2 \right). \quad (91)$$

In order to derive the above result, we used the fact that for a non-increasing continuous function g , we have

$$F_{g(X)}^{-1}(p) = g(F_X^{-1}(1-p)), \quad p \in (0, 1). \quad (92)$$

Here, $g = E[e^{W_{ij}}|Z]$ is a non-increasing function of Z since ρ_{ij} is always negative. So, we have that

$$F_{S_t}^{-1}(p) = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} - (\mu + \frac{1}{2}\delta^2)(i+j-t-1) + \frac{\delta \sum_{k=2}^t \sum_{l=t+2-k}^t \nu_{kl}}{\sqrt{\sum_{k=2}^t \sum_{l=t+2-k}^t \nu_{kl}^2}}} \Phi^{-1}(p)} \quad (93)$$

$$e^{\frac{1}{2} \left(\sigma^2 (\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij} + \delta^2(i+j-t-1) - \frac{\delta^2 \left(\sum_{k=2}^t \sum_{l=t+2-k}^t \nu_{kl} \right)^2}{\sum_{k=2}^t \sum_{l=t+2-k}^t \nu_{kl}^2} \right) + \frac{1}{2}\sigma^2}, \quad p \in (0, 1). \quad (94)$$

$F_{S_t}(x)$ can be obtained from solving the equation

$$\sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(E[W_{ij}] - \rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(F_{S_t}(x)) + \frac{1}{2}(1 - \rho_{ij}^2)\sigma_{W_{ij}}^2 + \frac{1}{2}\sigma_{\epsilon_{ij}}^2 \right) = x. \quad (95)$$

3.5.3 The cdf of S'_u

From Theorem 2, we can obtained the improved convex upper bound for the discounted IBNR reserve from

$$S'_u = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(F_{(\mathbf{B}\tilde{\beta})_{ij} - Y(i+j-t-1)|Z}^{-1}(U) \right) \exp \left(F_{\epsilon_{ij}}^{-1}(V) \right), \quad (96)$$

with $Z = \sum_{i=2}^t \sum_{j=t+2-i}^t \nu_{ij} Y_{i+j-t-1}$.

From previous results

$$F_{S'_u}(z) = \int_0^1 F_N \left(\ln(z) - \ln(F_{S'_u}^{-1}(y)) \right) dy \quad (97)$$

with $F_N(x)$ the cdf of $N(0, \sigma^2)$ and $S''_u = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(F_{(\mathbf{B}\tilde{\beta})_{ij} - Y(i+j-t-1)|Z}^{-1}(U) \right)$.

Now, we will derive an algorithm for the determination of $F_{S''_u}(y)$. From Proposition 3 we know that

$$S''_u = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(E[W_{ij}] + \rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(U) + \sqrt{1 - \rho_{ij}^2}\sigma_{W_{ij}}\Phi^{-1}(V) \right). \quad (98)$$

Since $F_{S''_u|U=u}$ is a sum of $t(t-1)$ comonotonous random variables, we have

$$F_{S''_u|U=u}^{-1}(p) = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(E[W_{ij}] + \rho_{ij}\sigma_{W_{ij}}\Phi^{-1}(u) + \sqrt{1 - \rho_{ij}^2}\sigma_{W_{ij}}\Phi^{-1}(p) \right). \quad (99)$$

$F_{S_u''''|U=u}$ also follows implicitly from

$$\sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(E[W_{ij}] + \rho_{ij} \sigma_{W_{ij}} \Phi^{-1}(u) + \sqrt{1 - \rho_{ij}^2} \sigma_{W_{ij}} \Phi^{-1}(F_{S_u''''|U=u}(y)) \right) = y. \quad (100)$$

The cdf of S_u'''' then follows from

$$F_{S_u''''}(y) = \int_0^1 F_{S_u''''|U=u}(y) du. \quad (101)$$

3.6 Bounds constructed on the basis of $\hat{W}_D$

The estimator $\hat{W}_D$ for the mean of the IBNR reserve,

$$\hat{W}_D = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} + \frac{1}{2}\tilde{\sigma}^2} \quad (102)$$

given in Doray (1996), constitutes a close upper bound for the UMVUE of the mean of the IBNR reserve if $\frac{t(t+1)}{2} - p$ is large and the residual sum of squares is small. It should be noted that $e^{(\mathbf{B}\tilde{\beta})_{ij} + \tilde{\sigma}^2/2}$ is the estimator of the mean of a lognormal distribution $LN((\mathbf{B}\tilde{\beta})_{ij}, \sigma^2)$ obtained by replacing the parameters $\tilde{\beta}$ and σ^2 by their unbiased estimates.

Adding now a discount process to (102) like in the previous subsection, gives

$$\hat{W}_{DD} = \sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} - Y(i+j-t-1) + \frac{1}{2}\tilde{\sigma}^2} \quad (103)$$

Now, we can apply the same methodology as explained before. Proposition 3 is still applicable. The only difference is that ϵ_{ij} is changed by $\frac{1}{2}\tilde{\sigma}^2$, with

$$\frac{1}{2}\tilde{\sigma}^2 \sim \text{Gamma} \left(\frac{n-p}{2}, \frac{\sigma^2}{n-p} \right) \quad (104)$$

For example, we have for the upperbound of $\hat{W}_{DD}$

$$\hat{W}_{DD}^u = \sum_{i=2}^t \sum_{j=t+2-i}^t \exp \left(F_{(\mathbf{B}\tilde{\beta})_{ij} - Y(i+j-t-1)}^{-1}(U) \right) \exp \left(F_{\frac{1}{2}\tilde{\sigma}^2}^{-1}(V) \right) \quad (105)$$

From previous results, we get for the distribution function of $\hat{W}_{DD}^u$,

$$F_{\hat{W}_{DD}^u}(z) = \int_0^1 F_G \left(\ln(z) - \ln \left(\sum_{i=2}^t \sum_{j=t+2-i}^t e^{(\mathbf{B}\tilde{\beta})_{ij} - (\mu + \frac{1}{2}\delta^2)(i+j-t-1) + \sqrt{\sigma^2(\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}')_{ij} + \delta^2(i+j-t-1)}\Phi^{-1}(u)} \right) \right) du, \quad (106)$$

with $F_G(x)$ the cdf of $\text{Gamma}(\frac{n-p}{2}, \frac{\sigma^2}{n-p})$.

4 A bootstrap approach

The bootstrap as an inferential statistical computer intensive device was introduced by Efron (1979) as a quite intuitive and simple way of making approximations to the distributions of which are very hard or even impossible to compute analytically. This technique has proved to be a very useful tool in many statistical applications and can be particularly interesting to assess the variability of the claim reserving predictions and to construct upper limits at an adequate confidence level. Its popularity is due to a combination of available computing power and theoretical development. One advantage of the bootstrap is that the technique can be applied to any data set without having to assume an underlying distribution. Most computer packages can handle very large numbers of repeated samplings, and this should not limit the accuracy of the bootstrap estimates.

Our goal is to construct a prediction interval for the IBNR reserve in which the population distribution is not known. If we do not know the distribution of the population, then our best guess at the distribution is provided by the data. The main idea in bootstrapping is that we (a) pretend that the data constitute the population and (b) take samples from this pretended population (which we call "resamples"). Substituting the sample for the population means that we are interested in the frequency with which the observed values occurred. Thus, if values 930 or larger occur in 2% of the samples, we want it to occur in 2% of our resamples, on average. This is done by sampling with replacement. From the re-sample, we calculate the statistic we are interested in. This is called a "bootstrap statistic". After storing this value, one repeats the above steps collecting a large number (B) of bootstrap statistics. The general idea is that the relationship of the bootstrap statistics to the observed statistic is the same as the relationship of the observed statistic to the true value. Under mild regularity conditions, the bootstrap yields an approximation to the distribution of an estimator or test statistic that is at least as accurate as the approximation obtained from first-order asymptotic theory. Presentation of the bootstrap technique could be easily found in the literature (see for instance Efron and Tibshirani (1993)).

As already mentioned above, with bootstrapping, we treat the obtained data as if they are an accurate reflection of the parent population, and then draw many bootstrapped samples by sampling, with replacement, from a pseudo-population consisting of the obtained data. Technically, this is called "non-parametric bootstrapping", because we are sampling from the actual data and we have made no assumptions about the distribution of the parent population, other than that the raw data adequately reflect the population's shape. If we were willing to make more assumptions, such as an assumption that the parent population follows a normal distribution, then we could do our sampling, with replacement, from a normal distribution. This is called "parametric bootstrapping". These bootstrap methods will be used here in order to obtain confidence bounds for (high quantiles of) the distribution of the *discounted* IBNR reserve.

The lower bound S_l turns out to perform very well. A final method to obtain a confidence bound for the predictive distribution is a combination of the power of this lower bound and bootstrapping. We will bootstrap a high percentile of the distribution of the lower bound. This is done as follows:

1. The preliminaries:

- Estimate the model parameters $\vec{\beta}$ and σ^2 .
- Calculate the fitted values: $\hat{\mu}_{ij} = e^{(\mathbf{X}\vec{\beta})_{ij}}$ ($i = 1, \dots, t; j = 1, \dots, t+1-i$).
- Calculate the residuals: $r_{ij} = Z_{ij} - \ln \hat{\mu}_{ij}$ ($i = 1, \dots, t; j = 1, \dots, t+1-i$).

2. Bootstrap loop (to be repeated B times):

- Generate a set of residuals r_{ij}^* by sampling with replacement from the original residuals (r_{ij}) ($i = 1, \dots, t; j = 1, \dots, t+1-i$).
- Create a new upper triangle Y_{ij}^* :
 - non-parametric bootstrap (NPB)
 $Y_{ij}^* = e^{\ln(\hat{\mu}_{ij}) + r_{ij}^*}$ ($i = 1, \dots, t; j = 1, \dots, t+1-i$).
 - parametric bootstrap (PB)
 $Y_{ij}^* = e^{(\mathbf{X}\vec{\beta})_{ij} + \tilde{\sigma}N(0,1)}$ ($i = 1, \dots, t; j = 1, \dots, t+1-i$).

Now we have bootstrapped a run-off triangle.

- Calculate for this bootstrapped triangle the parameters $\hat{\vec{\beta}}^*$ and $(\tilde{\sigma}^2)^*$.
- Calculate the percentile k of the distribution of S_l , $S_{l(k)}^*$, using $\hat{\vec{\beta}}^*$ and $(\tilde{\sigma}^2)^*$.
- Return to the beginning of step 2 until the B repetitions are completed.

3. Analysis of the bootstrap data:

- Apply the percentile method to the bootstrap observations to obtain the required prediction interval.

Other references concerning the use of the bootstrap methodology in claims reserving are [10] and [21]. In these papers the bootstrap technique is used to obtain prediction errors for different claim reserving methods, namely methods based on the chain ladder technique and on generalised linear models. Applications of the bootstrap technique to claim reserving can also be found in Lowe (1994), in Taylor (2000) and in England and Verrall (2002).

5 Numerical illustrations

This section is structured as follows. In the first part of this section we illustrate the bounds derived for the discounted IBNR reserve S numerically and graphically. Additionally, we perform the bootstrap on the lower bound S_l , in order to describe the statistical error involved.

In the first part, we investigate the accuracy of the proposed bounds, by comparing their cdf to the empirical cdf obtained with Monte Carlo simulation. To demonstrate graphically the precision of the derived bounds (and not the choice of the stochastic model), we build a non-cumulative run-off triangle ourselves starting from a given set of parameters and an given stochastic model. In other words: 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 empirical distribution of the total IBNR reserve S , we have to simulate a triangle. So, the maximum likelihood estimation procedure of the parameters is skipped because we know their correct values. In this example we will choose the probabilistic trend family of models from Zehnwirth (see (4)) with the following parameter values:

Parameters	Values	Parameters	Values
$\alpha_1 = \alpha_2 = \alpha_3$	12	β_2	-0.5
α_4	11.8	$\beta_3 = \beta_4$	-0.55
α_5	11.6	$\beta_5 = \beta_6 = \beta_7$	-0.5
α_6	11.4	$\beta_8 = \beta_9 = \beta_{10}$	-0.4
$\alpha_7 = \alpha_8$	11.2	$\gamma_1 = \dots = \gamma_7$	0.15
$\alpha_9 = \alpha_{10}$	11.0	γ_8	0.20
α_{11}	10.8	$\gamma_9 = \dots = \gamma_{20}$	0.25
β_1	0.07	σ^2	0.0324

Table 1: Parameter values Probabilistic Trend Family

Given the stochastic model (4) and the parameter values in Table 1, the triangle in Table 2 could be a possible run-off triangle.

In order to be able to compare the distribution functions of the stochastic bounds S_l , S_u and S'_u with the distribution function of S , we will completely specify the multivariate distribution function of the random vector $(Y_1, Y_2, \dots, Y_{t-1})$. In particular, we will assume that the random variables Y_i are i.i.d. and $\text{normal}(\mu + \frac{1}{2}\delta^2, \delta^2)$. This will enable us to simulate the cdf's in case there is no way to compute them analytically. The conditioning random variable Z is defined as before

$$Z = \sum_{i=2}^t \sum_{j=t+2-i}^t \nu_{ij} Y_{i+j-t-1}, \quad (107)$$

	1	2	3	4	5	6	7	8	9	10	11
1	204717	227507	183238	121488	54519	29881	44001	24109	18465	14719	10339
2	196781	183982	142761	107199	85482	56254	29081	35837	23763	28472	
3	176731	237786	168028	126843	72702	51353	29448	30307	31771		
4	202179	252667	179682	126795	75341	53803	52129	34346			
5	207159	208416	170027	133789	70243	65843	47930				
6	186543	220898	114236	114024	111391	75912					
7	168067	198976	173653	114802	82857						
8	211075	343066	135456	175479							
9	174303	341114	144884								
10	257741	373418									
11	261338										

Table 2: Run-off triangle with non-cumulative claim figures

with ν_{ij} as in (63). In our numerical illustrations, we will choose the parameters of the normal distribution involved as follows:

$$\mu = 0.08, \quad \delta = 0.11 \quad (108)$$

Fig. 2 shows the cdfs of the upper and lower bounds, compared to the empirical distribution based on 100000 randomly generated, normally distributed vectors. Since $S_l \leq_{cx} S \leq_{cx} S'_u \leq_{cx} S_u$, the same ordering holds for the tails of their distribution functions which can be observed to cross only once, and so we can easily identify the cdfs. The (simulated) cdf of S is the dotted line. We see that the cdf of S_l is very close to the distribution of S , which was expected because of the choice of Z . The closeness can be illustrated by the fact that the first moments are equal and the second moments are almost equal. The "real" standard deviation equals 556165 whereas the standard deviation of the lower bound equals 526786. An estimate for the 95th percentile is given by 4554959. The comonotonous upper bound S_u performs very badly in this case. Here, the only information used to compute the distribution of the sum are the marginal distribution functions of the respective cells. Given the marginal distribution functions, comonotonicity is the dependence structure of the vector $\mathbf{X}$ which leads to the most risky sum S (in the sense of convex order). The standard deviation of the upper bound is given by 1037390 which is much higher than the real standard deviation. The estimate for the 95th percentile now equals 5567299. This estimate is of course much higher than the estimate for the 95th percentile of the cdf of S_l . This comes from the fact that in order to determine S_l , we make use of the (estimated values of the) correlations between the cells of the lower triangle, whereas in the case of S_u , the distribution is an upper bound (in the sense of convex order) for any possible dependence structure between the components of the vector $\mathbf{X}$. The improved upper bound performs better, as could be expected.

Figure 2: Lower, upper and improved upper bound (solid lines) vs. empirical distribution of the discounted IBNR reserve (dotted line) for the run-off triangle in Table 2.

ear	F_S			F_{S_l}			F_{S_u}			$F_{S'_u}$		
	95%	mean	st. dev.	95%	mean	st. dev.	95%	mean	st. dev.	95%	mean	st. dev.
2	22708	16418	3509	18297	16477	1075	11867	9020	1606	21067	16310	2697
3	57763	43926	7817	50790	44013	3945	34398	25605	4927	56745	43682	7385
4	94584	72577	12373	86426	72609	7955	60110	43589	9191	95057	72239	12851
5	139176	106940	18099	130631	107077	13448	93464	65714	15331	143489	106767	20582
6	199548	153059	26111	190236	153426	20896	139279	95775	23898	209461	153246	31357
7	281013	215091	36907	270243	215609	30895	200726	136472	35170	296911	215651	45191
8	477490	364683	63327	461778	365788	54147	348667	234258	62430	509386	366263	79326
9	666706	509794	88361	647783	511669	76681	497352	330962	90587	720246	512691	114690
10	1143586	874306	151494	1112158	877859	131928	864339	572766	158548	1241926	880002	199606
11	1584387	1212375	209922	1559617	1229482	185840	1286448	808480	258355	1851357	1232984	337894
R	4555332	3569168	556165	4554959	3594009	526786	5567299	3659097	1037390	5145645	3599834	851405

Table 3: 95th percentiles, means and standard deviations of the distributions of S_l , S_u and S'_u vs. S . ($\mu = 0.08, \delta = 0.11$)

Table 3 summarizes the numerical values of the means and the standard deviations of the 3 approximations S_l , S_u and S'_u vs. S , as well as for the overall reserve as for the row totals

$$S_i = \sum_{j=t+2-i}^t Y_{ij} e^{(\mathbf{B}\hat{\beta})_{ij} - Y(i+j-t-1) + \epsilon_{ij}}, \quad i = 2, \dots, t. \quad (109)$$

The first column of each distribution function gives the upper limits for a confidence level of 95%. We can conclude that the lower bound approximates the "real discounted

reserve" very good. Remark that each year the upper bound S_u underestimates the row totals heavily and overestimates finally the overall reserve ($\mathbf{R}$).

Figure 3 shows the results of the obtained bounds for the different row totals.

Figure 3: Lower and improved upper bound (solid lines) vs. empirical distribution (dotted line) for the different discounted row totals.

As can be seen, the performance of the lower bound increases with time during more recent years, while the improved upper bound performs better in the first years.

In order to describe the statistical error involved, we apply the non-parametric bootstrap procedure (as described in section 4) to the run-off triangle in Table 2. Table 4 and Figure 4 show the results of this method for a bootstrap loop of 5000 times.

	$S_{l(95)}^B$	Simulation
1 st percentile	2798705	2909915
2.5 th percentile	2981332	3069365
5 th percentile	3149503	3222480
10 th percentile	3370466	3417358
25 th percentile	3750665	3773876
50 th percentile	4254657	4223183
75 th percentile	4850316	4815883
90 th percentile	5520199	5464421
95 th percentile	5872907	5694252
97.5 th percentile	6168048	5956964
99 th percentile	6580555	6330745

Table 4: Percentiles of the bootstrapped 95th percentile of the distribution of the lower bound $S_{l(95)}^B$ vs. the simulation.

Figure 4: Bootstrapped 95th percentile of the distribution function of the lower bound S_l ($S_{l(95)}^B$: dotted line) vs. the simulation (broken line) and the lower bound (solid line).

In order to check the practical validity of the bootstrap method we simulated 5000 triangles from model (4), calculating for each triangle the distribution of the lower bound S_l . The distribution of the corresponding 95th percentiles is also given in Table 4 and Figure 4.

We conclude that the bootstrap distribution yields appropriate confidence bounds. In fact, this plot confirms the problem of under- and overestimating the total reserve.

6 Conclusions and possibilities for future research

In this paper, we considered the problem of deriving the distribution function of the present value of a triangle of claims payments that are discounted using some given stochastic return process. It is known that it is impossible to find an explicit expression for the distribution function, even when starting from a classical loglinear regression model. 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.

Extensions could be the generalizations of the underlying regression model for instance to the error structure of the exponential family of distributions. Incorporating stable laws in order to model the discount factor is another way to generalize the results given above.

The assumption of a normal distribution for $\tilde{\epsilon}$ presents one advantage over that of other distributions. When the reserves are to be discounted for interest, we can still find the distribution of the present value of the future payments. If the force of interest is constant over a year, it follows from a property of the lognormal distribution that the joint distribution of the discounted value of the future payments is also multivariate lognormal. Stochastic interest rates could also be built in the model and the reserve estimated by simulation. This also immediately imposes a limitation on this class of models in that incremental claim amounts must be positive.

As already mentioned, the sum of lognormal random variables is not lognormally distributed. However in practice it is often claimed to be approximately lognormal. Perhaps it is useful to quantify the distance between the distribution of S and the lognormal family of distributions by means of the so-called Kullback-Leibler information. The Hellinger distance will allow us to measure the closeness of the derived lower and upper bounds as well as how far these bounds are from S in the distributional sense. So, we can identify which parameters have influence on this distance. This could also be a topic for a next paper.

References

- [1] Barnett, G. and Zehnwirth, B., 1998. Best estimates for reserves. *Casualty Actuarial Society Forum*, 1-54.
- [2] Christofides, S., 1990. Regression models based on log-incremental payments. *Claims Reserving Manual*, Vol 2, Institute of Actuaries, London.

- [3] Davison, A. C. and Hinkley, D.V., 1997. *Bootstrap Methods and their Application*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.
- [4] 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.
- [5] Dhaene, J., Wang, S., Young, V. and Goovaerts, M. J., 1997. Comonotonicity and Maximal Stop-Loss Premiums. Research Report 9730, Department of Applied Economics K.U.Leuven.
- [6] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D., 2002. (a) The concept of comonotonicity in actuarial science and finance: theory. to be published in *Insurance Mathematics and Economics*.
- [7] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D., 2002. (b) The concept of comonotonicity in actuarial science and finance: applications. to be published in *Insurance Mathematics and Economics*.
- [8] Doray, L.G., 1996. UMVUE of the IBNR reserve in a lognormal linear regression model. *Insurance Mathematics and Economics*, **18(1)**, 43-58.
- [9] Efron, B., Tibshirani, R. J., 1993. *An introduction to the bootstrap*. Chapman and Hall, New York.
- [10] 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.
- [11] England, P. D. and Verrall, R. J., 2001. A flexible framework for stochastic claims reserving. Proceedings of the Casualty Actuarial Society. to appear.
- [12] England, P. D. and Verrall, R. J., 2002. Stochastic claims reserving in general insurance. Institute of Actuaries and faculty of Actuaries.
- [13] Goovaerts, M.J. and Redant, H., 1999. On the distribution of IBNR reserves. *Insurance Mathematics and Economics*, **25(1)**, 1-9.
- [14] Goovaerts, M.J., Kaas, R., Dhaene, J. and Denuit, M., 2001. *Modern Actuarial Risk Theory*. Kluwer Academic, The Netherlands.
- [15] 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.
- [16] Kaas, R., Dhaene, J., Vyncke, D., Goovaerts, M.J. and Denuit, M. 2001. A simple Geometric proof that comonotonic risks have the convex-largest sum, *ASTIN Bulletin*, accepted.

- [17] Kremer, E., 1982. IBNR-Claims and the Two-way Model of ANOVA. *Scandinavian Actuarial Journal*, 47-55.
- [18] 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.
- [19] Mack, T., 1993. Distribution free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, **23**(2), 213-225.
- [20] Mack, T., 1994. Measuring the variability of chain-ladder reserve estimates. Casualty Actuarial Society, Spring Forum.
- [21] Pinheiro, P. J. R., Andrade e Silva, J. M. and Centeno, M., 2002. Bootstrap Methodology in Claim Reserving.
- [22] Renshaw, A.E., 1989. Chain ladder and interactive modelling (claims reserving and GLIM). *Journal of the Institute of Actuaries*, **116**(III), 559-587.
- [23] 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.
- [24] Renshaw, A.E. and Verrall, R.J., 1994. A stochastic model underlying the chain-ladder technique. *Proceedings XXV Astin Colloquium*, Cannes.
- [25] Taylor, G.C. and Ashe, F.R., 1983. Second moments of estimates of outstanding claims. *Journal of Econometrics*, **23**, 37-61.
- [26] Taylor, G.C., 2000. *Loss reserving: an actuarial perspective*. Kluwer Academic Publishers.
- [27] Verrall, R. J., 1989. A state space representation of the chain-ladder linear model. *J.I.A.*, **116**, 589-610.
- [28] Verrall, R. J., 1991. On the unbiased estimation of reserves from loglinear models. *Insurance Mathematics and Economics*, **10**, 75-80.
- [29] Zehnwirth, B., 1985. *ICRFS version 4 manual and users guide*. Benhar Nominees PTy Ltd, Turramurra, NSW, Australia
- [30] Zehnwirth, B., 1989. The chain ladder technique - A stochastic model. *Claims Reserving Manual*, Vol 2, Institute of Actuaries, London.