KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0050

SOME REMARKS ON IBNR EVALUATION TECHNIQUES

by
M. GOOVAERTS
J. DHAENE
E. VANDEN BORRE
H. REDANT

D/2000/2376/50

Some remarks on IBNR evaluation techniques*

Marc J. Goovaerts[†] Jan Dhaene[‡] Eddy Vanden Borre[§] Hendrik Redant[¶]

November 14, 2000

Abstract

In this short note we give some comments and general remarks on the methodology of IBNR computations, as presented at the workshop on IBNR computations at the 2000 ASTIN Meeting, Porto Cervo, Sardinia.

1 IBNR evaluations

The past data (the upper triangle) are the key elements of IBNR calculations. In almost any method, analysing the upper triangle is based on well-known techniques from statistics, see e.g. Neter, Kutner, Nachtsheim & Wasserman (1995). However, the essential problem to be solved is the management of the risk associated with the future (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. Indeed, 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.

*The computations of Section 2 have been performed with the software VACS-LRC.

[†]K.U.Leuven, Universiteit van Amsterdam.

[‡]K.U.Leuven, Universiteit van Amsterdam.

[§]Appointed actuary Fortis Bank Insurance.

[¶]K.U.Leuven.

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 his reserve as a percentile. In practice, the insurer will choose a very high percentile as basis for his reserve. From the viewpoint of the insured, the choice of a high percentile is a safe strategy, and will as such be favored by the control authorities. The determination of the reserve as a percentile makes it possible to compute an explicit safety loading (= reserve minus expectation of the payments). Fiscal authorities tend to prefer explicit (i.e. visible) margins to implicit (i.e. hidden) margins. The choice of a high percentile is also important for the insurers' rating. It is a key element in a Risk Based Capital approach. In the Belgian and Dutch insurance practice, we observe that insurers determine their reserves on percentiles such as 99.75%, or even higher. Hence, only very high time-consuming multivariate simulations will lead to a sufficient number of simulated values in such an extreme tail. Another disadvantage of a simulation technique is that there is no way to measure the distance between the "real" and the "simulated" distribution function. Hence, there is no information available concerning the error that is involved by using a simulation technique. Of course, a multivariate simulation technique will only be possible if the whole dependency structure of the lower triangle is known. In practice, we encounter situations where only the distribution functions of each cell can be estimated with enough accuracy, but where only limited information of the dependency structure can be obtained (because not enough data are available). We can conclude that a multivariate simulation technique is not the 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 which use as much as possible of the available information. Hence, within a certain class of random vectors (with given marginals, and eventually additional information), we propose to look for upper and lower bounds for the sum of the cells of the lower triangle. For details of this technique, we refer to Redant & Goovaerts (1999), Goovaerts, Dhaene & De Schepper (2000) and Kaas, Dhaene & Goovaerts (2000). The upper and lower bounds presented in these papers are bounds in the sense of convex order, which means that the expectations are exact and the stop-loss premiums are ordered. The convex order can of course be interpreted in terms of utility theory. The "total variation distance" can be used as a measure between the true distribution function and each of the bounds.

Moreover, the proposed technique leads to a solution of the IBNR problem which is similar to a"value at risk" approach in finance. The technique described in the above-mentioned papers also allows to calculate the conditional tail expectation or the remaining tail risk for a given percentile, i.e. for a given level of the IBNR reserve.

2 Numerical illustration

The statistical model that will be used to describe the past and the future claim amounts is a loglinear model which looks for trends in the three directions, namely accident year, development year and calendar year. An early reference to the use of such models in the actuarial literature is De Vylder & Goovaerts (1979). Given the statistical model for the claim amounts, the present value S of the future IBNR payments follows from the vectors $\mathbf{X} = (X_1, \cdots, X_n)$ and $\mathbf{Y} = (Y_1, \cdots, Y_n)$ where the vector $\mathbf{X}$ describes the future claim amounts and the vector $\mathbf{Y}$ describes the discount process:

$$S = \sum_{i=1}^{n} X_i \, Y_i.$$

We assume that the vectors $\mathbf{X}$ and $\mathbf{Y}$ are mutually independent and that both have lognormal marginals. Hence, S is a sum of dependent lognormal random variables.

In order to illustrate the technique explained in Redant & Goovaerts (1999), Goovaerts, Dhaene & De Schepper (2000) and Kaas, Dhaene & Goovaerts (2000) for determining IBNR reserves, we use the run-off triangle of Table 1 in Mack (1993), see also Taylor & Ashe (1983) and Verral (1990, 1991) and references therein. First, we will assume that $\mathbf{Y} \equiv (1, \cdots, 1)$, hence we discount at an interest rate equal to 0.

In Figures 1 and 2, we use a loglinear model with 2 parameters in the direction of the accident years (denoted by α_i) and 4 parameters in the direction of the development years (denoted by β_k) as is characteristic for the chain-ladder model. No parameters are used in the calendar year direction. This model is obtained as a conceivable model (given the data), from the software VACS-LRC.We call this model the "6 parameter model":

$$\ln X_{ij} = \alpha_i + \sum_{k=1}^{j} \beta_k + \varepsilon_{ij},$$

where X_{ij} is the claim amount of accident year i and development year j and the ε_{ij} are mutually independent normally distributed random variables (with zero mean and variance equal to 0.069). The parameters α_i and β_k are given by: $\alpha_1 = 12.514$; $\alpha_2 = \cdots = \alpha_{10} = 12.838$; $\beta_1 = 0.938$; $\beta_4 = -0.579$; $\beta_5 = \beta_6 = \beta_7 = -0.219$; $\beta_9 = -1.089$. The remaining β's are non-significant (equal to 0).

Figure 1 shows the probability density function (pdf) of the optimal approximation bound, as explained in Kaas, Dhaene & Goovaerts (2000). This approximation can be shown to be very close to the real distribution function. 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 1,355,969, whereas the standard deviation of the lower bound equals 1,341,161. An estimate for the 99,75% percentile is given by 22,111,049.

Figure 2 shows the pdf of the comonotonic upper bound. Here, the only information used to compute the distribution function of the sum are the marginal distribution functions of the respective cells. Given the marginal distribution functions, comonotonicity is the dependency structure of the vector **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 5,481,136 which is much higher than the real standard deviation, as could be expected. The estimate for the 99.75% percentile now equals 39,779,075. This estimate is of course much higher than the estimate in Figure 1. This comes from the fact that in order to determine the best approximation, we make use of the (estimated values of the) correlations between the cells of the lower triangle, whereas in Figure 2, the distribution function is an upper bound (in the sense of convex order) for any possible dependency structure between the components of the vector **X**.

In Figure 3, we show the pdf of the optimal approximation of S, when we take a stochastic discounting process into account. We assume that the yearly returns are lognormally distributed (with parameters μ and σ) and mutually independent. Three different scenarios are presented: scenario 1 ($\mu = 0.05$ and $\sigma = 0.03$), scenario 2 ($\mu = 0.08$ and $\sigma = 0.1$) and finally, the case of no discounting. We observe that increasing the expected yearly return shifts the pdf to the left, and increasing the variance of the yearly return makes the pdf broader. Remark that in scenario 1 the 95% percentile is given by 18,435,063, whereas in scenario 2 this percentile is given by 19,751,126 and in the case of no discounting the 95% percentile equals 20,274,672. Hence, as could be expected, stochastic discounting will normally diminish the required reserve. Finally, remark that the second scenario leads to a higher reserve than the first one, which means that the effect of the higher expected return (which tends to decrease the reserve) is overshadowed by the effect of the higher variability (which tends to increase the reserve).

3 REFERENCES

De Vylder, F.; Goovaerts, M.J. (editors) (1979). Proceedings of the first meeting of the contactgroup "Actuarial Sciences", K.U.Leuven, nr 7904B, Wettelijk Depot: D/1979/2376/5.

Dhaene, J.; Denuit, M. (1999). "The safest dependency structure among risks", *Insurance: Mathematics & Economics* 25, 11-21.

Dhaene, J.; Goovaerts, M. (1996). "Dependency of risks and stop-loss order", *ASTIN Bulletin* 26, 201-212.

Dhaene, J.; Goovaerts, M.J. (1997). "On the dependency of risks in the individual life model", *Insurance: Mathematics & Economics* 19, 243-253.

Dhaene, J.; Wang, S.; Young, V.; Goovaerts, M.J. (1998). "Comonotonicity and maximal stop-loss premiums", *Bulletin of the Swiss Association of Actuaries*, accepted for publication.

Goovaerts, M.J.; Dhaene, J. De Schepper, A. (2000). "Stochastic Upper Bounds for Present Value Functions", *Journal of Risk and Insurance Theory*, 67.1, 1-14.

Goovaerts, M.J.; Redant, R. (1999). "On the distribution of IBNR reserves", *Insurance: Mathematics & Economics* 25, 1-9.

Mack, T. (1993). "Distribution-calculation of the standard error of chain ladder reserve estimates", *ASTIN Bulletin* 23(2), 213-225.

Neter, J.; Kutner, M.H.; Nachtsheim, C.J.; Wasserman, W. (1995). *Applied Linear Statistical Models,* fourth edition, IRWIN, pp. 1408.

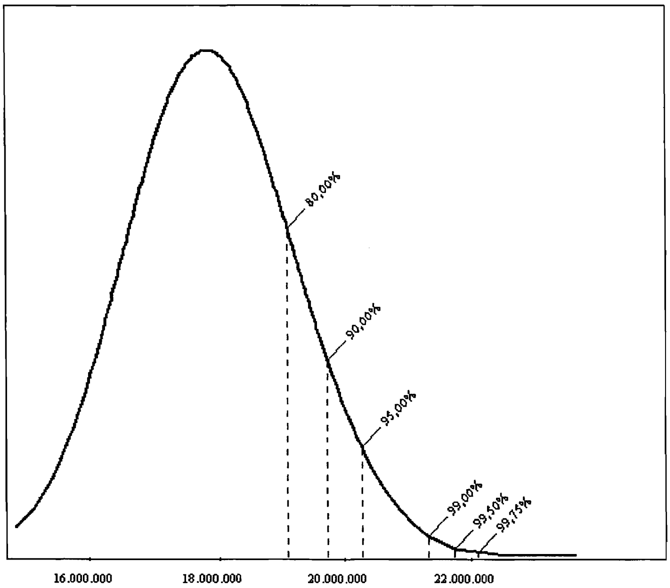
Taylor, G.C. and Ashe, F.R. (1983). "Second moments of estimates of outstanding claims", *Journal of Econometrics* 23, 37-61.

Verral, R.J. (1990). "Bayes and empirical Bayes estimation for the chain ladder model", *ASTIN Bulletin* 20, 217-243.

Verral, R.J. (1991). "On the estimation of reserves from loglinear models", *Insurance: Mathematics & Economics* 10, 75-80.

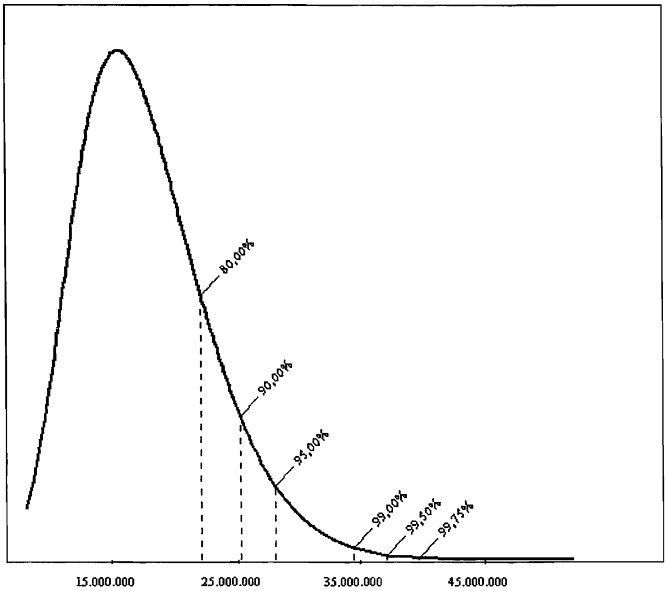
Vyncke, D.; Goovaerts, M.J.; Dhaene, J. (2000): "Convex upper and lower bounds for present value functions", *Applied Stochastic Models in Business and Industry,* accepted for publication.

Wang, S.; Dhaene, J. (1998). "Comonotonicity, correlation order and stop-loss premiums", *Insurance: Mathematics & Economics*, 22(3), 235-243.



Mean :	17.986.481
StDev(Bound) :	1.341.161
StDev(Real) :	1.355.969
Coeff of Var. :	0,07
Mod :	17.835.852
Med :	17.936.083
Percentiles :	
80,00%	19.096.415
90,00%	19.732.893
95,00%	20.274.672
99,00%	21.331.918
99,50%	21.732.764
99,75%	22.111.049

fig1: 6-parameter-model: optimal approximation



Mean :	17.986.481
StDev(Bound) :	5.481.136
StDev(Real) :	1.355.969
Coeff of Var. :	0,30
Mod :	15.767.467
Med :	17.200.696
Percentiles :	
80,00%	22.102.890
90,00%	25.203.503
95,00%	28.092.433
99,00%	34.444.102
99,50%	37.116.172
99,75%	39.779.075

fig2: 6-parameter model: comonotonic upper bound

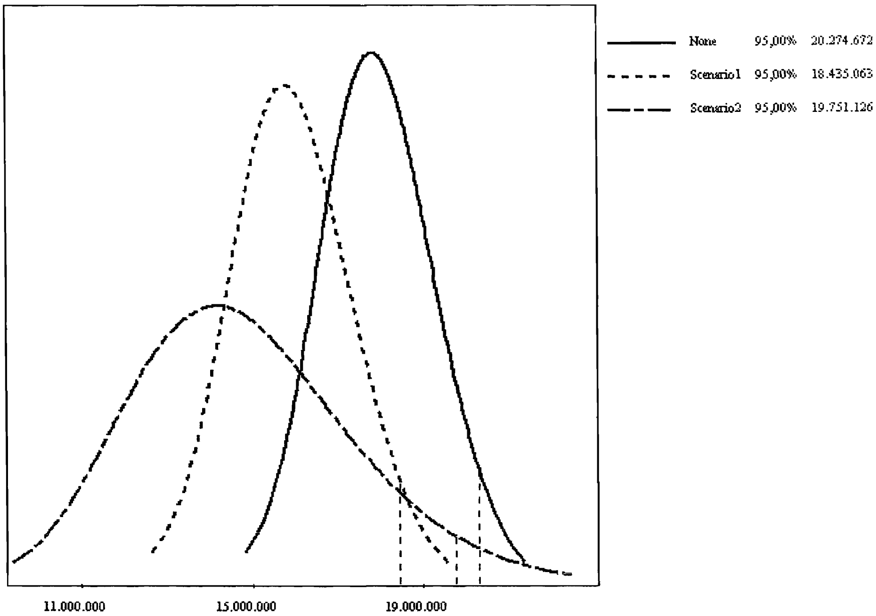


fig3: 6-parameter model: optimal approximation, different discount processes

Scenario1:	$\mu=0.05$, $\sigma=0.03$
Scenario2:	$\mu=0.08$, $\sigma=0.1$
None:	$\mu=\sigma=0$