Discussion of 'A Bayesian generalized linear model for the Bornhuetter-Ferguson method of claims reserving'

Katrien Antonio [*†] Jan Beirlant [†] Tom Hoedemakers [†‡]

December 28, 2004

Aim of this comment

Professor Verrall nicely illustrates how Bayesian models can be applied to claims reserving within the framework of Generalized Linear Models (GLMs) and how they lead to posterior predictive distributions of quantities of interest. In this discussion we apply a Bayesian model in the context of discounted loss reserves. The outcomes of this approach are compared with results on approximations for the distribution of the discounted loss reserve when the run-off triangle is modelled by a GLM. These approximations are based on the concepts of comonotonicity and convex order and are explained in full details in Hoedemakers *et al.* (2003) (for lognormal claims reserving models) and Hoedemakers *et al.* (2004) (for claims reserving using GLMs). The comonotonicity approach has been shown to provide elegant solutions to various other actuarial and financial problems involving the distribution of a sum of dependent random variables (check www.kuleuven.ac.be/insurance - research papers for more illustrations). We realize that the Bayesian posterior predictive distribution is a very general workhorse, which takes into account all sources of uncertainty in the model formulation and is applicable to different statistical domains, whereas our approximations originate from a specific actuarial context. We want to illustrate however that the predictive distribution based on the comonotonic bounds provides results that are close to the results obtained via MCMC. The main advantage of the bounds is that several risk measures such as percentiles (VaRs), expected shortfalls (stop-loss premia) and TailVaRs can be calculated easily from it.

Generalized linear models for claims reserving: the likelihood-based approach

We use the notation from Verrall (2004). An insurer is interested in the aggregated value $\sum_{i=2}^{n}\sum_{j=n+2-i}^{n} C_{ij}$ corresponding with the future part of a classical run-off triangle (as shown in Table 1 of Verrall's paper). In a (likelihood-based) GLM framework this reserve will be predicted by

$$\text{reserve} = \sum_{i=2}^{n}\sum_{j=n+2-i}^{n} \hat{\mu}_{ij},$$

[*]Corresponding author: Katrien.Antonio@econ.kuleuven.ac.be

[†]University Center for Statistics, de Croylaan 54, 3001 Heverlee, Belgium

[‡]Department of Applied Economics, Catholic University Leuven, Naamsestraat 69, 3000 Leuven, Belgium

with $\hat{\mu}_{ij} = g^{-1}\left((\mathbf{R}\hat{\vec{\beta}})_{ij}\right)$. Here $\mathbf{R}$ denotes the design matrix, $\vec{\beta}$ is the vector of regression parameters and g the link function of the GLM that is used to model both observed data and future observations. $\hat{\vec{\beta}}$ denotes the maximum likelihood estimator of $\vec{\beta}$.

The reserve is a provision for future payments. Therefore (in this discussion) our estimated loss reserve reflects the time value of money and we consider the *discounted* loss reserve. Assume that the reserve will be invested such that an amount of 1 at time $j-1$ becomes e^{Y_j} at time j. The discount factor for a payment of 1 at time i is then given by $e^{-(Y_1+Y_2+\ldots+Y_i)} := e^{-Y(i)}$. Let us assume for instance that

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

where $B(i)$ is the standard Brownian motion. We extended the WinBugs programs from Appendix B of Verrall (2004) to incorporate this discounting process. The discounted reserve now becomes

$$\begin{aligned} R &:= \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} \hat{\mu}_{ij} e^{-Y(i+j-n-1)} \\ &= \sum_{i=2}^{n} \sum_{j=n+2-i}^{n} \hat{\mu}_{ij} \exp\left(-(\mu - \delta^2/2)(i+j-n-1) - \delta B(i+j-n-1)\right). \end{aligned} \tag{1}$$

In any realistic model for the return process, R will be a sum of strongly dependent random variables. Because one can not rely on traditional risk theory, it becomes very hard or even impossible to compute the cumulative distribution function (cdf) of R analytically. For actuaries however the complete predictive distribution of the reserve is important because different risk measures can be calculated from it. As illustrated by Verrall (2004) (for GLMs) and in earlier work by (for instance) de Alba (2002) (for lognormal models) Bayesian techniques are useful in this area as they provide the posterior predictive distribution of the reserve. This discussion compares the results from Bayesian GLMs (as used in Verrall (2004)) with the results based on upper and lower bounds for R for which the cdf can be easily calculated.

Comonotonicity and convex order applied to loss reserving

Let $S = X_1 + \ldots + X_n$ be an arbitrary sum of random variables. In many actuarial and financial problems (e.g. claims reserving) the distribution of such a sum of random variables is of interest (check Dhaene *et al.* (2002b) for various illustrations). In general, however, this cdf can not be obtained analytically because of the dependency structure involved in the stochastic vector $(X_1, \ldots, X_n)$. Kaas *et al.* (2000) and Dhaene *et al.* (2002a) propose a possible way out by considering upper and lower bounds for (the cdf of) S that allow explicit calculations of various actuarial quantities.

The original dependency structure can be replaced by a simpler one, which leads to a more 'dangerous' random variable, but allows analytical calculations. Decisions (e.g. 'Which amount of money should be held aside?') based on this random variable lead to a safe, conservative strategy. However, in the context of claims reserving, provisions will be held that are too high. Therefore it makes sense that one also considers a less 'dangerous' random variable allowing analytical calculations of various risk measures.

Kaas *et al.* (2000) and Dhaene *et al.* (2002a) provide a mathematical foundation for these intuitive ideas. The concept 'more/less dangerous' is defined by the notion of convex order. $X \leq Y$ in convex order if and only if $E[X] = E[Y]$ and $E[(X-d)_+] \leq E[(Y-d)_+]$ for all real d. Upper and lower bounds for S are derived for which the terms in the involved sums are comonotonic, meaning that they are all monotonic (increasing or decreasing) functions of the same random variable. Precisely the quality of comonotonicity makes it possible to compute various actuarial quantities in an analytical way.

Hoedemakers *et al.* (2003) extend the results mentioned above from ordinary sums of variables to sums of scalar products of independent random variables. This extension allows to derive a lower and upper bound for the discounted loss reserve R. Based on the concept of comonotonicity, they explain how the cdf of these bounds can be derived explicitly.

Numerical illustration

Consider the run-off triangle in Table 1, taken from Taylor & Ashe (1983) and used in various other publications on claims reserving.

	1	2	3	4	5	6	7	8	9	10
1	357,848	766,940	610,542	482,940	527,326	574,398	146,342	139,950	227,229	67,948
2	352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
3	290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
4	310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
5	443,160	693,190	991,983	769,488	504,851	470,639				
6	396,132	937,085	847,498	805,037	705,960					
7	440,832	847,631	1,131,398	1,063,269						
8	359,480	1,061,648	1,443,370							
9	376,686	986,608								
10	344,014									

Table 1: Run-off triangle

These data are modelled using a gamma GLM with logarithmic link function and the following structure for the linear predictor

$$\begin{aligned} \log(\mu_{ij}) &= \alpha_1 I(i=1) + \alpha_2 I(i=2,3) + \alpha_3 I(i=4) + \alpha_4 I(i>4) + \beta_1 I(j>1) \\ &+ \beta_2 I(j>4) + (j-5)\beta_3 I(5<j<9) + 3\beta_3 I(j>8) \end{aligned}$$

whereby $I(.)$ denotes an indicator function.

The discounting process from (1) (with $\mu = 0.08$ and $\delta = 0.11$) is incorporated in the WinBugs code for the gamma GLM. To enable comparisons with the results based on the comonotonic bounds, flat priors were used both for the row and column parameters of the linear predictor and for the scale parameter in the gamma model. Table 2 contains the results for the 95th percentile of the predictive distribution obtained via MCMC simulations with the WinBugs program. A burn-in of 10,000 iterations was allowed, after which another 10,000 iterations were performed.

As mentioned before, an upper (denoted by R_u) and lower bound (denoted by R_l) for R (in convex order) can be derived. R_u and R_l are sums of comonotonic terms. Precisely this quality enables the analytical calculation of various actuarial quantities (such as high percentiles). The bounds for the discounted loss reserve use the maximum likelihood estimates of the parameters in the linear predictor. To incorporate the error arising from the estimation of these parameters Hoedemakers *et al.* (2004) apply bootstrap methodology. For details on this procedure we refer to the paper. Table 2 compares the Bayesian 95th percentile and the bootstrapped 95th percentile of the lower and upper bound for the different reserves. We bootstrapped 1000 times, computed each time (analytically) the 95th percentile of upper and lower bound and report in Table 2 the 95th percentile of the bootstrap samples obtained in this way.

Year	lower bound	upper bound	Bayesian
2	360,725	387,404	436,151
3	700,465	765,451	760,177
4	945,845	982,425	970,535
5	1,441,016	1,513,186	1,448,056
6	1,913,383	1,977,934	1,919,300
7	2,519,292	2,614,564	2,558,208
8	3,557,014	3,702,302	3,641,890
9	4,573,767	4,770,944	4,727,262
10	5,577,925	5,821,804	5,638,301
Overall	20,949,190	21,988,048	20,360,196

Table 2: 95th percentile of the predictive distribution of the different discounted reserves

Compared with the Bayesian results, Table 2 illustrates that the comonotonicity approach indeed provides relevant information concerning the predictive distribution of discounted loss reserves in a GLM framework.

References

- de Alba, E. (2002). *Bayesian estimation of outstanding claims reserves.* North American Actuarial Journal, 6(4), 1-20.
- Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. & Vyncke, D. (2002a). *The concept of comonotonicity in actuarial science and finance: theory.* Insurance: Mathematics and Economics, 31(1), 3-33.
- Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. & Vyncke, D. (2002b). *The concept of comonotonicity in actuarial science and finance: applications.* Insurance: Mathematics and Economics, 31(2), 133-161.
- Hoedemakers, T., Beirlant, J., Goovaerts, M.J. & Dhaene, J. (2003). *Confidence bounds for discounted loss reserves.* Insurance: Mathematics and Economics, 33(2), 297-316.
- Hoedemakers, T., Beirlant, J., Goovaerts, M.J. & Dhaene, J. (2004). *On the distribution of discounted loss reserves using generalized linear models.* Scandinavian Actuarial Journal, *forthcoming.*
- Kaas, R., Dhaene, J. & Goovaerts, M.J. (2000). *Upper and lower bounds for sums of random variables.* Insurance: Mathematics and Economics, 27(2), 151-168.
- Taylor, G.C. & Ashe F.R. (1983). *Second moments of estimates of outstanding claims.* Journal of Econometrics, 23, 37-61.
- Verrall, R.J. (2004). *A Bayesian generalized linear model for the Bornhuetter-Ferguson method of claims reserving.* North American Actuarial Journal, 8(3), 67-89.