STOCHASTIC CLAIMS INFLATION IN IBNR

PROFESSOR DR ERHARD KREMER

1998 GENERAL INSURANCE CONVENTION
AND
ASTIN COLLOQUIUM

GLASGOW, SCOTLAND: 7-10 OCTOBER 1998

STOCHASTIC CLAIMS INFLATION IN IBNR

ERHARD KREMER, HAMBURG

SUMMARY

This paper deals with loss reserving under inclusion of stochastic claims inflation, a topic that is of current interest. Note that recently a new paper on it was presented at the international Astin colloquium at Cairns. In the following it is basically assumed that the discounted claims increase follows an autoregressive model of ARCH-type and that the stochastic yearly interest intensity follows a classical autoregressive model. A procedure to estimate adequately the stochastic discounting factors is deduced. This is combined with the link-ratio technique and the classical forecasting procedure for autoregressive processes, giving a new stochastic loss reserving technique. The whole method is perfect and handy. Its practicability is demonstrated in an example.

1. INTRODUCTION.

During the last decades a lot was written on how to calculate adequately loss reserves in nonlife insurance. Certain survey books were published (see e.g. Taylor (1986), Institute of Actuaries (1985) and the topic was included into standard text books (see e.g. Sundt (1983), Kremer (1985)). In the present decade many attempts were made to refine and extend previous methods (see e.g. Kremer (1993), (1996) and Doray (1996)). New aspects were considered (see e.g. Kremer (1997)) and totally new approaches presented (see e.g. Verrall (1995)). Nevertheless the toolskit is not yet totally complete. For example the techniques for coping with claims inflation are not very far developed. In practice one usually adjusts in advance the claims for inflation and applies standard techniques to the adjusted data. The reason for proceeding like this is that the models underlying nearly all standard techniques do not incorporate the claims inflation effects. To the author only one standard technique, that does incorporate the claims inflation in its basic model, is known. It is the method published by Verbeek in 1972 and reconsidered by Taylor in 1977. This method is a handy recursive procedure with which one estimates mean yearly claims growth and the mean yearly claims inflation effect. Since the modelling of the claims inflation is quite simple, one called for more refined models, yielding more refined techniques. Clearly one likes to incorporate models of modern financial mathematics for adequate modelling the claims inflation. Most modern it would be to work with stochastic discounting. Just this is done in a recent article of Goovaerts and de Schepper (1997) and in the following one of the author. Whereas Goovaerts and de Schepper use a pure probability-theoretic approach, the author gives a more mathematical-statistical approach. So in the following a fairly new claims reserving technique is developed which is based on a model that incorporates stochastic claims inflation. In an example it is shown that the method works fairly well in case the claims data is not too irregular. In the author's opinion the following method is more practical than that given by Goovaerts and de Schepper.

2. MODEL.

Denote with the random variable Z_{ij} on $(\Omega,\mathcal{A},P)$ the total claims amount of a (collective of) risk(s) in development year no. j with respect to its accident year no. i. With the claims settlement period n, the set of random variables:

$$Z_{\Delta} = (Z_{ij},\ j = 1,\ldots,n - i + 1,\ i = 1,\ldots,n)$$

is the socalled run-off triangle. For the sequel define the increases:

$$Y_{ij} = Z_{ij} - Z_{i,j-1}$$

(with $Y_{i0} := 0$) and with given volume measures $V_i > 0$, $i = 1,\ldots,n$:

$$X_{ij} = Y_{ij}/V_i$$

Obviously the problem of forecasting the (unknown) random variables Z_{ij}, $j \geq n - i + 2$, $i = 2,\ldots,n$ out of the (known) run-off triangle Z_{Δ} is equivalent with the problem of forecasting the (unknown)

$$X_{ij},\ j \geq n - i + 2,\ i = 2,\ldots,n$$

out of the (known) run-off triangle X_{Δ} (of the X_{ij}). Remember that with a forecast $\hat{Z}_{in}$ of Z_{in} $(i \geq 2)$ the IBNR-reserve for accident year no. i is just

$$R_i = \hat{Z}_{in} - Z_{i,n-i+1}$$

For the present paper we assume that there are random variables D_i, $i = 1,\ldots,2n - 1$ with

$$D_1 = 1 \quad , \quad D_i \in (0,1),\ i = 2,\ldots,2n - 1$$

such that the random variables

$$W_{ij} = X_{ij} \cdot D_{i+j-1}$$

follow the *stochastic recursion:*

(2.1) $$W_{ij} = L_j \cdot W_{i,j-1} + e_{ij}$$

where $L_j > 0$ is an unknown parameter and e_{ij} random error terms, independent of the $W_{i,j-1}$ and with:

(2.2) $$E(e_{ij}) = 0$$

(2.3) $$\mathrm{Var}(e_{ij}) = \sigma_j^2 / V_i$$

$\sigma_j^2 \geq 0$ unknown. Furthermore assume that:

(2.4) the vectors $(W_{i1}, e_{i2},\ldots,e_{in})$, $i = 1,\ldots,n$

are stochastically independent,

(2.5) $E(W_{i1}) = E(W_{j1})$, for all i,j

$Var(W_{i1}) = s^2/V_i$, for all i,

with an unknown $s^2 \geq 0$

Note that (2.1), (2.2), (2.3), (2.5) imply that:

(2.6) $E(W_{ik}) = E(W_{jk}) = \mu_k$, for all i,j,

$V_i \cdot Var(W_{ik}) = V_j \cdot Var(W_{jk}) = \eta_k$, for all i,j,

for all k = 1,...,n. The D_i have to be interpreted as stochastic deflation factors. The inverses $D_i^{-1} = A_i$ describe the stochastic claims inflation, more condretely the random variable:

$$I_i = ((\frac{A_i}{A_{i-1}}) - 1)* 100$$

is the percentage of claims size increase in year no. i due purely to claims inflation. Note that the assumption $D_1 = 1$ means that the deflation is done up to the end of the first period. Lateron the $D_2,\dots,D_n$ will be calculated (estimated) from the run-off triangle X_Δ, resulting in $\hat{D}_2,\dots,\hat{D}_n$. For forecasting the X_{ij}, $j \geq n - i + 2$ one needs forecasts $\hat{A}_{n+1},\dots,\hat{A}_{2n-1}$ of the future inflation factors $A_{n+1},\dots,A_{2n-1}$. For giving a reasonable forecast we assume in addition that:

$$(2.7) \qquad D_i = \exp(-\sum_{j=1}^{i} S_j)$$

with random variables S_j following an autoregressive time series model of order one (=AR(1)):

$$(2.8) \qquad S_j = a \cdot S_{j-1} + (1-a) \cdot b + \varepsilon_j$$

$$\text{with } S_1 = 0,$$

where a,b are unknown parameters and the ε_j are uncorrelated error terms with:

$$E(\varepsilon_j) = 0 \quad , \quad Var(\varepsilon_j) = \sigma^2$$

with an unknown $\sigma^2 \geq 0$.

Note that (2.7) is equivalent with

$$(2.9) \qquad S_i = \ln(D_{i-1}) - \ln(D_i), \quad i = 2,\dots,n$$

$$\text{with } S_1 = 0 \quad ,$$

and that for $i > j$:

$$A_i = A_j \cdot \exp\left(\sum_{k=j+1}^{i} S_k\right)$$

or equivalently:

$$(2.10) \qquad A_{i+1} = A_i \cdot \exp(S_{i+1})$$

The model (2.7), (2.8) is a special case of those used in Dhaene (1989).

3. ESTIMATION.

For carring through the forecasts of the X_{ij}, $j \geq n - i + 2$ from the run-off triangle X_Δ (of the X_{ij}, $j \leq n - i + 1$) one needs estimators of the

a) $D_2,\ldots,D_n$

b) $L_2,\ldots,L_n$

c) a,b ,

based on the data of the run-off triangle X_Δ. From statistics one knows that a good estimator of μ_k is just:

$$\hat{\mu}_k = \left(\frac{1}{V_.^{(k)}}\right) \cdot \sum_{i=1}^{n-k+1} V_i \cdot W_{ik}$$

with:

$$V_.^{(k)} = \sum_{i=1}^{n-k+1} V_i$$

According to consequence (2.6) a reasonable criterion for giving estimators $\hat{D}_i$ of D_i is the minimisation of the sum of squares:

$$\sum_{i=1}^{n} \sum_{j=1}^{n-i+1} V_i \cdot [W_{ij} - \hat{\mu}_j]^2$$

in the D_i, $i = 2,\ldots,n$ (note that $D_1 = 1$). This sum of squares can be rewritten as:

$$\sum_{k=1}^{n} \sum_{j=1}^{k} V_{k-j+1} \cdot [W_{k-j+1,j} - \hat{\mu}_j]^2$$

By differentiating with respect to D_κ and putting the result equal to zero, one arrives after routine manipulations at the equation:

$$D_\kappa \cdot \Lambda_\kappa = \Lambda_{1\kappa}(D_1,\ldots,D_n) + \Lambda_{2\kappa}(D_1,\ldots,D_n) \tag{3.1}$$

with the sums:

$$\Lambda_\kappa = \sum_{j=1}^{\kappa} (1 - \frac{V_{\kappa-j+1}}{V_\cdot^{(j)}}) \cdot V_{\kappa-j+1} \cdot X^2_{\kappa-j+1,j}$$

$$\Lambda_{1\kappa}(D_1,\ldots,D_n) = \sum_{j=1}^{\kappa} (1 - \frac{V_{\kappa-j+1}}{V_\cdot^{(j)}}) \cdot V_{\kappa-j+1} \cdot X^2_{\kappa-j+1} \cdot \hat{\mu}_j$$

$$\Lambda_{2\kappa}(D_1,\ldots,D_n) = \sum_{\substack{k=1 \\ k\neq\kappa}}^{n} \sum_{j=1}^{\min(k,\kappa)} \frac{V_{\kappa-j+1}}{V^{(j)}} \cdot V_{\kappa-j+1} \cdot X_{\kappa-j+1} \cdot (W_{\kappa-j+1} - \mu_j)$$

That stochastic equation system (3.1), with $\kappa = 2,\ldots,n$ and $D_1 = 1$, can not be solved analytically. But one can take an iterative numerical procedure. According to ideas of numerical mathematics one guesses the recursive procedure:

$$D_\kappa^{(p+1)} = \Lambda_\kappa^{-1} \cdot [\Lambda_{1\kappa}(D_1^{(p)}..,D_n^{(p)}) + \Lambda_{2\kappa}(D_1^{(p)},\ldots,D_n^{(p)})]$$

$$\text{with } D_1^{(p)} = 1 \text{ for all } p, \text{ and } \kappa = 2,\ldots,n \quad ,$$

where $D_\kappa^{(1)}$, $1 = 2,\ldots,n$ are chosen suitable start values in $(0,\infty)$. When the $D_\kappa^{(p)}$ have stabilized sufficiently good, one takes them as estimates $\hat{D}_\kappa$ of D_κ $(\kappa = 2,\ldots,n)$.

For estmating $L_2,\ldots,L_n$ one can look into the author's paper Kremer (1984). That paper gives under the model (2.1.) - (2.3) the estimator:

$$\text{(3.2)} \qquad \hat{L}_j = \frac{\sum_{i=1}^{n-j+1} V_i \cdot W_{ij} \cdot W_{i,j-1}}{\sum_{i=1}^{n-j+1} V_i \cdot W_{i,j-1}^2} ,$$

The W_{ij} are not exactly known. Clearly one replaces them in (3.2) by their approximations:

$$\hat{W}_{ij} = X_{ij} \cdot \hat{D}_{i+j-1}$$

the discounted claims(-ratio) increments. For estimating the parameters a,b, one computes according to (2.9) estimators $\hat{S}_i$ of S_i as:

$$(3.3) \qquad \hat{S}_i = \ln(\hat{D}_{i-1}) - \ln(\hat{D}_i)$$

with $\hat{S}_i = 0$.

From times series analysis one knows (see e.g. Fuller (1976)) that for the model (2.8) reasonable estimators of a and b are:

$$\hat{b} = \frac{1}{n} \cdot \sum_{i=1}^{n} S_i$$

$$\hat{a} = \frac{\sum_{i=2}^{n} (S_i - \hat{b}) \cdot (S_{i-1} - \hat{b})}{\sum_{i=2}^{n} (S_{i-1} - \hat{b})^2}$$

Again the S_i are unknown. One clearly replaces them by their approximations $\hat{S}_i$

4. METHOD.

First one estimates the $D_2,\ldots,D_n$ like described in the previous section, resulting in the $\hat{D}_2,\ldots,\hat{D}_n$. With these estimated deflation factors one deflates the X_{ij}, giving the $\hat{W}_{ij}$. From (3.2) one calculates the lagfactors $\hat{L}_j$ (with $\hat{W}_{ij}$ instead of W_{ij}) and completes the run-off triangle W_Δ to a rectangle according to the prediction advice:

$$\hat{W}_{ij} = \hat{L}_j \cdot \hat{W}_{i,j-1}$$

$$\text{for} \quad j = n - i + 2,\ldots,n, \quad i = 2,\ldots,n$$

For having the predictions $\hat{X}_{ij}$ of the X_{ij} for $j \geq n - i + 2$ one needs an estimator $\hat{A}_j$ for A_j with $j \geq n + 1$, since one would take:

$$(4.1) \qquad \hat{X}_{ij} = \hat{W}_{ij} \cdot \hat{A}_{i+j-1}$$

For computing the $\hat{A}_i$ for $i \geq n + 1$ one clearly applies the recursion (2.10):

$$(4.2) \qquad \hat{A}_{i+1} = \hat{A}_i \cdot \exp(\hat{S}_{i+1})$$

with start:

$$\hat{A}_n = \hat{D}_n^{-1}$$

and predictions $\hat{S}_i$, $i \geq n + 1$ of the S_i, $i \geq n + 1$

One knows from time series analysis that an optimal prediction $\hat{S}_i$ of S_i, $i \geq n + 1$ (in the model (2.8)), based on $S_1,\ldots,S_n$, can be calculated as:

$$\hat{S}_{n+j} = a^j \cdot S_n + (1-a^j) \cdot b$$

(see e.g. Fuller (1976)). For the unknown a,b,S_n one inserts $\hat{a},\hat{b}$ (see previous section) and the $\hat{S}_n$ according to the definition (3.3). Having those predictions $\hat{S}_i$, $i \geq n + 1$, one applies (4.1) with (4.2).

5. EXAMPLE.

For $n = 7$ consider the following run-off triangle Z_Δ:

23.20	32.33	35.40	37.69	39.13	39.76	40.16
25.08	35.06	38.61	41.61	43.37	44.56	
29.29	39.34	43.85	47.30	49.45		
31.14	43.99	49.52	53.33			
36.37	49.39	55.47				
37.32	51.25					
44.55						

giving for choice $V_i = 1$ for all i, the run-off triangle X_Δ:

23.20	9.13	3.07	2.29	1.44	0.63	0.40
25.08	9.98	3.55	3.00	1.76	1.19	
29.29	10.05	4.51	3.45	2.15		
31.14	12.85	5.53	3.81			
36.37	13.02	6.08				
37.32	13.93					
44.55						

With 100 iterations and start values:

$$D_2^{(1)} = 0.9 \quad , \quad D_3^{(1)} = 0.8 \quad , \quad D_4^{(1)} = 0.7$$

$$D_5^{(1)} = 0.6 \quad , \quad D_6^{(1)} = 0.5 \quad , \quad D_7^{(1)} = 0.4$$

one gets:

$$\hat{D}_2 = 0.91107 \qquad \hat{D}_3 = 0.78763 \qquad \hat{D}_4 = 0.74467$$

$$\hat{D}_5 = 0.63088 \qquad \hat{D}_6 = 0.61321 \qquad \hat{D}_7 = 0.52406$$

what corresponds to the yearly inflation percentages (in %):

$$\hat{I}_2 = 9.76 \qquad \hat{I}_3 = 15.67 \qquad \hat{I}_4 = 5.77$$

$$\hat{I}_5 = 18.04 \qquad \hat{I}_6 = 2.75 \qquad \hat{I}_7 = 17.01$$

The deflates run-off triangle $\hat{W}_\Delta$ is:

23.20	8.32	2.42	1.71	0.91	0.39	0.21
22.85	7.86	2.64	1.89	1.08	0.62	
23.07	7.48	2.85	2.12	1.16		
23.19	8.11	3.39	2.00			
22.95	7.98	3.19				
22.89	7.30					
23.55						

with what one gets the lagfactors:

$$\hat{L}_2 = 0.341 \qquad \hat{L}_3 = 0.364 \qquad \hat{L}_4 = 0.676$$

$$\hat{L}_5 = 0.545 \qquad \hat{L}_6 = 0.515 \qquad \hat{L}_7 = 0.543$$

with the estimates:

$$\hat{a} = -\ 0.6075 \qquad \hat{b} = 0.0923$$

one computes the predictions:

$$\hat{A}_8 = 2.213 \qquad \hat{A}_9 = 2.346 \qquad \hat{A}_{10} = 2.627$$

$$\hat{A}_{11} = 2.845 \qquad \hat{A}_{12} = 3.144 \qquad \hat{A}_{13} = 3.432$$

giving withg (4.1) the forecasts $\hat{X}_{ij}$, $j > n - i + 1$:

					0.75
				1.28	0.74
			2.41	1.31	0.80
		4.77	2.75	1.59	0.93
	5.88	4.21	2.57	1.43	0.86
17.60	6.79	5.14	3.03	1.72	1.06

and the predicted cumulative values $\hat{Z}_{ij}$, $j > n - i + 1$:

					45.31
				50.73	51.47
			55.74	57.05	57.85
		60.24	62.99	64.58	65.51
	57.13	61.34	63.91	65.34	66.20
62.94	68.94	74.08	77.11	78.83	79.85

what looks quite reasonable.

6. REMARKS.

Clearly one can try a more complicate ARIMA (p,d,q)-model instead of the simple AR(1)-model (2.8), like discussed in Dhaene (1989), e.g. an AR(2)-model. Then parameter-estimation of the S_j-process becomes more elaborate and forecasting of the S_j, $j \geq n + 1$ more complicated. But note that for smaller n the parameter-estimation of these models becomes quite unreliable since one has a too short past time series $D_2, \ldots, D_n$ and consequently it is mostly better to use the AR(1)-model.

REFERENCES:

Dhaene, J. (1989): Stochastic interest rates and autoregressive integrated moving average processes. ASTIN bulletin.

Doray, L. G. (1996): UMVUE of the IBNR reserve in a lognormal linear regression model. INSURANCE: Mathematics & Economics.

Fuller, W. A. (1976): Introduction to statistical time series. John Wiley, New York.

Goovaerts, M. and De Schepper, A. (1997): IBNR reserves under stochastic interest rates. ASTIN colloquium at Cairns,

Institute of Actuaries (1997): Claims reserving manual. Part 1-2. London.

Kremer, E. (1984): A class of autoregressive models for predicting the final claims amount. INSURANCE: Mathematics & Economics.

Kremer, E. (1985): Einführung in die Versicherungsmathematik. Vandenhoeck & Ruprecht, Göttingen & Zürich.

Kremer, E. (1993): Certain extensions of the chainladder technique. Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker. Additional correction note in 1994.

Kremer, E. (1996): Threshold lossreserving. Proceedings of the ASTIN colloquium at Copenhagen.

Kremer, E. (1997): Robust lagfactors. Blätter der deutschen Gesellschaft für Versicherungsmathematik.

Sundt, B. (1985): An introduction to nonlife insurance mathematics. Verlag Versicherungswirtschaft, Karlsruhe,

Taylor, G. C. (1977): Separation of inflation and other effects from the distribution of nonlife insurance claims delays. ASTIN bulletin.

Verbeek, H. G. (1972): An approach to the analysis of claims experience in motor liability excess of loss reinsurance. ASTIN bulletin.

Verrall, R. J. (1995): Claims reserving and generalised additive models. Contribution of the ASTIN colloquium at Leuven.

Author:
E. Kremer
Institut für Mathematische Stochastik
Bundesstrasse 55
20146 Hamburg
Germany