A Comparison of Stochastic Loss Reserving Methods

Ezgi Nevruz , Yasemin Gençtürk

Department of Actuarial Sciences
Hacettepe University
Ankara/TURKEY

02.04.2014

Outline

- Introduction
- Loss Process
- Simulation
- Loss Reserving Methods
- Application
- Conclusion

Our Aim

- Insurers should allocate an adequate reserve.
- Insurers need to select a suitable reserving method which estimates the expected liabilities as truely as possible.
- Profit of the companies does not only depend on the paid losses, but also the estimation of the future losses.

We aim to:

- estimate reserves using several loss reserving methods
- decide the suitable method for various scenarios by taking into account different performance criteria

Process of Loss Reserving

Steps of a stochastic loss reserving process:

1. Defining the model structure for the loss
2. Preparing the loss data in accordance with the loss development triangle (upper-left triangle)
3. Obtaining the goal triangle (lower-right triangle) by means of the loss development triangle and suitable reserve estimation method

Units of a Loss Development Triangle

Assumption: Claims are settled at accident year or within the next n development years.

$S_{i,j}$: Incremental losses for accident year i, development year j
$L_{i,j}$: Cumulative losses for accident year i, development year j

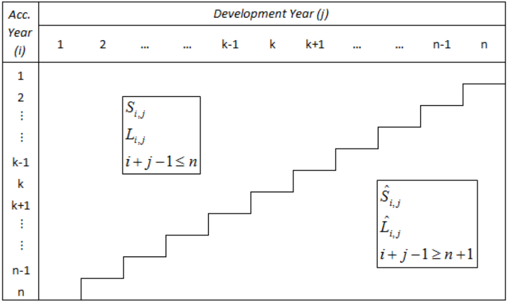
$$L_{i,j} = \sum_{k=1}^{j} S_{i,k}$$

By assumption, incremental and cumulative losses are

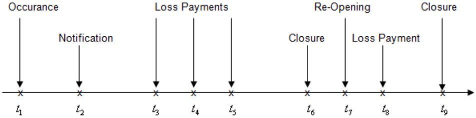
- observable for calendar year $i + j - 1 \leq n$
- not observable for calendar year $i + j - 1 \geq n + 1$

Process of Loss

Loss development triangle



The time of a single claim



Simulation of a Loss Development Triangle Using Indivial Losses with Changing Severity Method

Algorithm:

Step 1: Generate claim numbers N_i and individual claim amounts $\{C_{i,k}; k = 1, 2, ..., N_i\}$ for each accident year.

- Claim numbers: Poisson distribution
- Individual claim amounts: Pareto, Gamma and Lognormal distributions

Step 2: Obtain $\{U_{i,k}\}$ percentiles of each individual claim amounts $C_{i,k}$.

$$U_{i,k} = F(C_{i,k})$$

Simulation of a Loss Development Triangle (continued)

Step 3: For each $C_{i,k}; k = 1, 2, ..., N_i$ generate

- $X_{i,k,1}$ is the occurance date
- $X_{i,k,2}$ is the reporting delay
- $X_{i,k,3}$ is the settlement delay

Here we define

$$r_{i,k} = min\{\lfloor (X_{i,k,1} + X_{i,k,2}) \rfloor, n\}$$

$$R_{i,k} = min\{\lfloor (X_{i,k,1} + X_{i,k,2} + X_{i,k,3}) \rfloor, n\}$$

Thus, the kth individual loss at the accident year i is reported in $i + r_{i,k}$ calendar year and settled in $i + R_{i,k}$ calendar year (Narayan and Warthen, 2000).

Simulation of a Loss Development Triangle (continued)

Step 4: Obtain the developed loss amounts $\hat{C}_{i,k,j}$ increasingly with increasing j for each individual loss amount by means of the inverse of the distribution function, i.e. $F^{-1}(U_{i,k})$:

If $C_{i,k}$ has lognormal distribution with log-scale parameter μ_j and shape parameter σ_j^2

$$\hat{C}_{i,k,j} = \begin{cases} 0 & ;\ j = 1, 2, ..., r_{i,k} \\ exp\{\sqrt{2}\sigma_j[erf^{-1}(2U_{i,k} - 1] + \mu_j\} & ;\ j = r_{i,k} + 1, ..., R_{i,k} \\ exp\{\sqrt{2}\sigma_{R_{i,k}+1}[erf^{-1}(2U_{i,k} - 1] + \mu_{R_{i,k}+1}\}] & ;\ j = R_{i,k} + 1, ..., n \end{cases}$$

Here, erf^{-1} is the inverse of the error function which can be shown as

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

Simulation of a Loss Development Triangle (continued)

Step 4: Continued

If $C_{i,k}$ has Pareto distribution with location parameter α_j and shape parameter θ_j:

$$\hat{C}_{i,k,j} = \begin{cases} 0 & ;\ j = 1, 2, ..., r_{i,k} \\ \theta_j[\frac{1}{(1-U_{i,k})^{1/\alpha_j}} - 1] & ;\ j = r_{i,k}+1, ..., R_{i,k} \\ \theta_{R_{i,k}+1}[\frac{1}{(1-U_{i,k})^{1/\alpha_{R_{i,k}+1}}} - 1] & ;\ j = R_{i,k}+1, ..., n \end{cases}$$

Simulation of a Loss Development Triangle (continued)

Step 4: Continued

If $C_{i,k}$ has Gamma distribution with scale parameter θ_j and shape parameter α_j:

$$\hat{C}_{i,k,j} = \begin{cases} 0 & ;\ j = 1, 2, ..., r_{i,k} \\ \theta_j P^{-1}(\alpha_j, U_{i,k}) & ;\ j = r_{i,k}+1, ..., R_{i,k} \\ \theta_{R_{i,k}+1} P^{-1}(\alpha_{R_{i,k}+1}, U_{i,k}) & ;\ j = R_{i,k}+1, ..., n \end{cases}$$

Here, P is the lower regularized Gamma function that

$$P(a, x) = \frac{\gamma(a, x)}{\Gamma(a)} = \frac{\int_0^x t^{a-1} e^{-t} dt}{\Gamma(a)}$$

Simulation of a Loss Development Triangle (continued)

Step 5: Calculate the cumulative losses by the equation

$$L_{i,j} = (1+e)^{i-1} \sum_{k=1}^{N_i} \hat{C}_{i,k,j}$$

Here, because we obtain the cumulative losses increasingly, the incremental losses will be positive.

Inflation-Adjusted Chain Ladder (IACL) Method

In the setting of reserves on the basis of information obtained from past years, one should be aware of the fact that inflation may have affected the values of claims.

1. The incremental losses $S_{i,j}$'s are calculated for $j = 2, 3, ..., n$ and each accident year.
2. The loss amounts are accumulated by accident and development year as
$$S_{i,j}(1+e)^{n-i-j+1}$$
3. The inflation-adjusted cumulative losses $L_{i,j}$ are obtained from the inflation-adjusted incremental losses.

Inflation-Adjusted Chain Ladder Method (continued)

4. The development factor estimations $\hat{f}_j$ for $j = 2, 3, ..., n$ are obtained by

$$\hat{f}_j = \frac{\sum_{i=1}^{n-j+1} L_{i,j}}{\sum_{i=1}^{n-j+1} L_{i,j-1}}$$

5. Ultimate losses for each accident year are estimated as

$$\hat{L}_i = \hat{L}_{i,n} = L_{i,n-i+1} \prod_{j=n-i+2}^{n} \hat{f}_j$$

6. Finally, the reserve estimation for the ith accident year is calculated by the equation

$$\hat{R}_i = \hat{L}_i - L_{i,n-i+1}$$

We are dealing with the loss development triangles that include positive incremental losses.

When the expected value of an incremental loss is $\theta_{i,j}$, we will obtain the unbiased estimate of $\theta_{i,j}$'s for $i = 1, 2, ..., n$ and $j = n - i + 2, ..., n$.

Under the assumption that the incremental losses are positive, the regression model is

$$Z_{i,j} = ln(S_{i,j}) = \mu + \alpha_i + \beta_j + \varepsilon_{i,j}$$

where $\varepsilon_{i,j}$'s are iid $N(0, \sigma^2)$ distributed.

Regression Methods (continued)

Under the assumption that the $\{S_{i,j}\}$ r.v.s are independent and lognormally distributed, the $\{Z_{i,j}\}$ r.v.s. are independent and normally distributed where $i = 1, 2, ..., n; j = 1, 2, ..., n - i + 1$.

$$\mathbb{E}[Z_{i,j}] = X_{i,j}\beta, Var(Z_{i,j}) = \sigma^2$$

Therefore,

$$\mathbb{E}[S_{i,j}] = \theta_{i,j} = exp(X_{i,j}\beta + \frac{1}{2}\sigma^2)$$

where $X_{i,j}$ is the row vector of explanatory variables and β is a column vector of parameters.

Regression Methods (continued) Model 1 & Model 2 & Model 3

Model 1:

$$Z_{i,j} = \mu + \alpha_i + \beta_j + \varepsilon_{i,j} \Rightarrow \beta = [\mu, \alpha_2, ..., \alpha_n, \beta_2, ..., \beta_n]$$

Model 2:

$$Z_{i,j} = \mu + (i-1)\alpha + \beta_j + \varepsilon_{i,j} \Rightarrow \beta = [\mu, \alpha, \beta_2, ..., \beta_n]$$

Model 3:

$$Z_{i,j} = \mu + (i-1)\alpha + (j-1)\beta + \gamma ln(j) + \varepsilon_{i,j} \Rightarrow \beta = [\mu, \alpha, \beta, \gamma]$$

Here, there is a usual assumption that $\alpha_1 = 0$ and $\beta_1 = 0$ to make the model full rank (Verrall, 1991).

After β's are estimated by the method of least squares, the unbiased estimations of $\theta_{i,j}$'s will be obtained for $i = 2, ..., n$ and $j = n - i + 2, ..., n$.

1. After $\hat{\beta} = (X'X)^{-1}X'z$ is estimated,

   variance of the error is calculated as

$$\hat{\sigma}^2 = \frac{1}{r-p}(z - X\hat{\beta})'(z - X\hat{\beta})$$

   where $r = \frac{1}{2}n(n+1)$ is the number of observations, p is the number of parameters, X is the $(r \times p)$-dimensional design matrix and $z = [Z_{1,1}, Z_{1,2}, ..., Z_{1,n}, Z_{2,1}, ..., Z_{n,1}]'$ is the vector of observed losses.

2. Unbiased estimation of $\theta_{i,j}$ is obtained by

$$\hat{\theta}_{i,j} = exp(X_{i,j}\hat{\beta})g_m[\frac{1}{2}(1 - X_{i,j}(X'X)^{-1}X'_{i,j})s^2]$$

where

- the biased estimate of σ^2 is $\hat{\sigma}^2$ and the unbiased estimate of σ^2 is $s^2 = \frac{r}{r-p}\hat{\sigma}^2$
- $m = r - p$ is the degree of freedom
- if the df of $\hat{\sigma}^2$ is m, then

$$g_m(t) = \sum_{k=0}^{\infty} \frac{m^k(m+2k)}{m(m+2)...(m+2k)} \frac{t^k}{k!}$$

- $n = 11 \Rightarrow i = 1, 2, ..., 11; j = 1, 2, ..., 11$
- 10000 iterations
- Calculation of ultimate losses and actual reserves by the simulation of (11×11)-dimensional loss squares
- Estimation of reserves from the upper-left loss triangles
- Calculation of deviations and testing the performance of the reserving methods

Scenarios

System parameters:

- Inflation Rate (3 groups):
  Low (6%) & Med (8%) & High (10%)
- Individual loss amount rv sample (4 groups)
  1. Low Mean (500), Low Variance (150^2)
  2. Low Mean (500), High Variance (1000^2)
  3. High Mean (5000), Low Variance (1500^2)
  4. High Mean (5000), High Variance (10000^2)
- Individual loss amount rv distribution (3 groups):
  Pareto & Gamma & Lognormal

Thus, number of scenarios for each loss reserving method is $3 \times 4 \times 3 = 36$

Scenarios (summary)

Scenarios	Inflation Rate			Distribution of Individual Loss Amount rv			Individual Loss Amount Sample				Scenarios	Inflation Rate			Distribution of Individual Loss Amount rv			Individual Loss Amount Sample			
	Low: %6	Med: %8	High: %10	Lognormal	Pareto	Gamma	Mean: Low/Var: Low	Mean: Low/Var: High	Mean: High/Var: Low	Mean: High/Var: High		Low: %6	Med: %8	High: %10	Lognormal	Pareto	Gamma	Mean: Low/Var: Low	Mean: Low/Var: High	Mean: High/Var: Low	Mean: High/Var: High
S1	X			X			X				S19		X			X				X	
S2	X			X				X			S20		X			X					X
S3	X			X					X		S21		X				X	X			
S4	X			X						X	S22		X				X		X		
S5	X				X		X				S23		X				X			X	
S6	X				X			X			S24		X				X				X
S7	X				X				X		S25			X	X			X			
S8	X				X					X	S26			X	X				X		
S9	X					X	X				S27			X	X					X	
S10	X					X		X			S28			X	X						X
S11	X					X			X		S29			X		X		X			
S12	X					X				X	S30			X		X			X		
S13		X		X			X				S31			X		X				X	
S14		X		X				X			S32			X		X					X
S15		X		X					X		S33			X			X	X			
S16		X		X						X	S34			X			X		X		
S17		X			X		X				S35			X			X			X	
S18		X			X			X			S36			X			X				X

Table: Performance criteria of IACL

Sce	RMSE	MAPE	Sce	RMSE	MAPE	Sce	RMSE	MAPE
S1	231468	0.1268	S13	344257	0.1626	S25	482287	0.1951
S2	175684215	0.2497	S14	232986131	0.3263	S26	301075611	0.3909
S3	5217130	0.1848	S15	7610891	0.2361	S27	10446852	0.2821
S4	21064×10^7	**0.0051**	S16	8045642	0.4510	S28	7296275	0.3870
S5	513449	0.1524	S17	760505	0.1946	S29	1055637	0.2325
S6	185034	0.1521	S18	258931	0.2054	S30	283731	0.2474
S7	8144046	0.1922	S19	11884373	0.2458	S31	1634321	0.2949
S8	15983160	0.0391	S20	9430356	0.0526	S32	10709149	0.0143
S9	271503	0.1316	S21	404390	0.1688	S33	567182	0.2028
S10	**90643**	0.1440	S22	116253	0.1883	S34	149313	0.2310
S11	4495501	0.1803	S23	6552893	0.2300	S35	9003001	0.2749
S12	16781989	1.9814	S24	19190289	1.8623	S36	21716332	1.6306

Results of the IACL Method (continued)

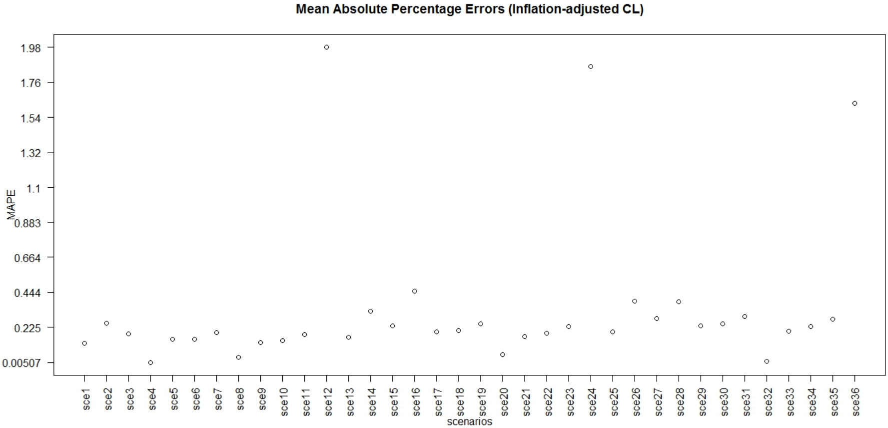
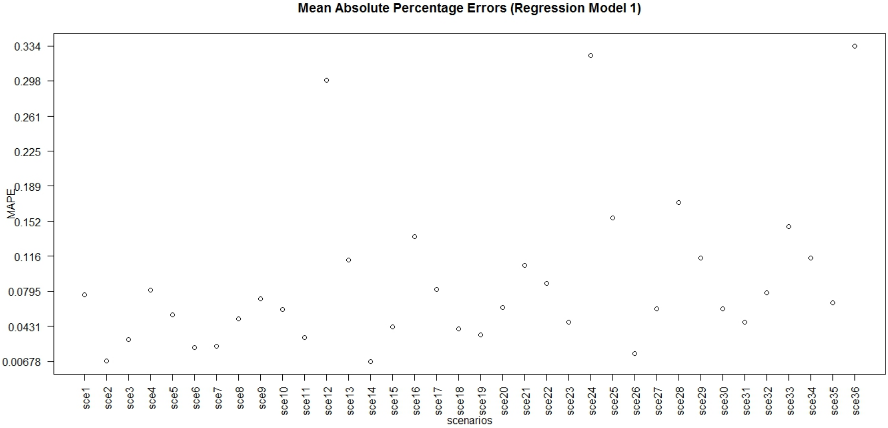


Table: Performance criteria of Reg.Model 1

Sce	RMSE	MAPE	Sce	RMSE	MAPE	Sce	RMSE	MAPE
S1	156007	0.0761	S13	251518	0.1120	S25	397975	0.1554
S2	160773809	0.0076	S14	183099933	**0.0068**	S26	207919528	0.0148
S3	1815472	0.0292	S15	2398084	0.0430	S27	3292199	0.0613
S4	9214×10^7	0.0810	S16	1319829	0.1361	S28	1716437	0.1718
S5	220734	0.0550	S17	349412	0.0813	S29	546526	0.1140
S6	146470	0.0211	S18	182668	0.0405	S30	225181	0.0615
S7	2483165	0.0227	S19	3213342	0.0343	S31	4229056	0.0475
S8	4927896	0.0511	S20	2334869	0.0626	S32	2482616	0.0782
S9	168473	0.0721	S21	271546	0.1062	S33	427905	0.1470
S10	**107060**	0.0606	S22	136029	0.0876	S34	172851	0.1144
S11	1658207	0.0316	S23	2236729	0.0475	S35	3082662	0.0676
S12	2077116	0.2987	S24	2439234	0.3244	S36	2943919	0.3339

Results of the Regression Model 1 (continued)

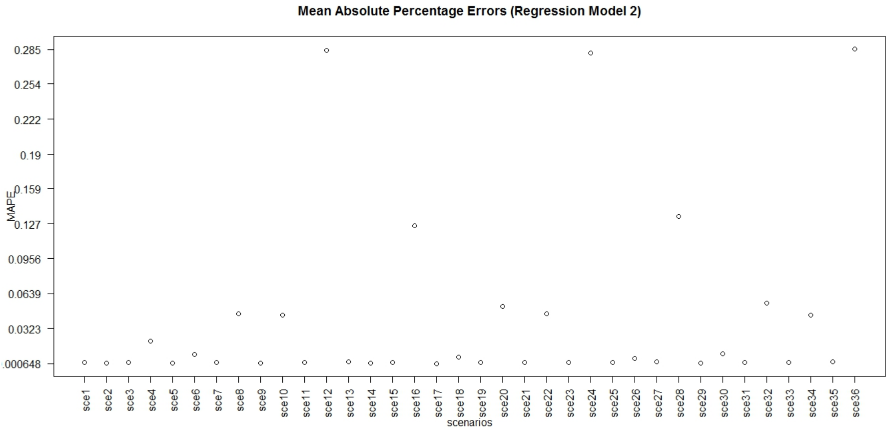


Results of the Regression Model 2

Table: Performance criteria of Reg.Model 2

Sce	RMSE	MAPE	Sce	RMSE	MAPE	Sce	RMSE	MAPE
S1	**60497**	0.0019	S13	70827	0.0020	S25	83526	0.0019
S2	172421446	0.0011	S14	194496253	0.0013	S26	218521613	0.0052
S3	1261596	0.0018	S15	1470149	0.0018	S27	1700560	0.0022
S4	8818×10^7	0.0209	S16	908302	0.1257	S28	1085860	0.1340
S5	99861	0.0012	S17	117533	**0.0006**	S29	137858	0.0011
S6	179460	0.0089	S18	196338	0.0064	S30	173524	0.0095
S7	1875667	0.0018	S19	2173718	0.0014	S31	2521990	0.0019
S8	4761533	0.0458	S20	2268173	0.0523	S32	2185048	0.0551
S9	64271	0.0013	S21	77189	0.0016	S33	90530	0.0016
S10	71075	0.0447	S22	83570	0.0455	S34	98912	0.0444
S11	1114531	0.0018	S23	1308376	0.0020	S35	1523648	0.0026
S12	1562025	0.2840	S24	1753841	0.2819	S36	2058645	0.2854

Results of the Regression Model 2 (continued)

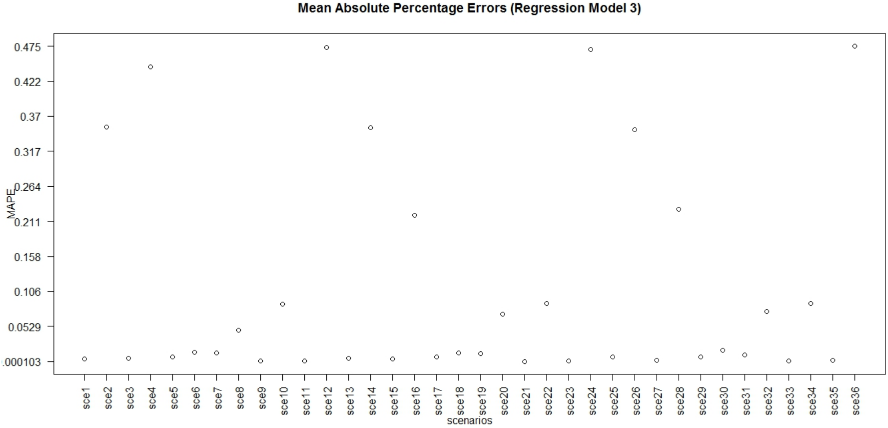


Results of the Regression Model 3

Table: Performance criteria of Reg.Model 3

Sce	RMSE	MAPE	Sce	RMSE	MAPE	Sce	RMSE	MAPE
S1	**61173**	0.0037	S13	71947	0.0051	S25	85430	0.0067
S2	216667078	0.3537	S14	243112338	0.3527	S26	273237566	0.3489
S3	1221086	0.0048	S15	1428751	0.0038	S27	1650780	0.0023
S4	10415×10^7	0.4439	S16	1084199	0.2200	S28	1290691	0.2296
S5	101873	0.0072	S17	119499	0.0066	S29	140802	0.0070
S6	123818	0.0138	S18	175276	0.0129	S30	166999	0.0172
S7	1859678	0.0125	S19	2148445	0.0117	S31	2473910	0.0101
S8	4492516	0.0470	S20	2240185	0.0714	S32	2155219	0.0750
S9	64619	0.0011	S21	**77590**	0.0001	S33	90999	0.0011
S10	77513	0.0869	S22	90831	0.0877	S34	107509	0.0871
S11	1080735	0.0006	S23	1271817	0.0005	S35	1485463	0.0019
S12	2123385	0.4732	S24	2384650	0.4703	S36	2795760	0.4752

Results of the Regression Model 3 (continued)



Closing Comments

- Regression models give better results. They do not provide the best answers in all situations but they are consistent and give not only the point estimation but also a confidence interval.
- Actuaries do not apply the CL method blindly. This method is efficient when the development factors are reasonable.
- Inflation rate affects the performance of a loss reserving method.
- We assume that the inflation rate is constant over the accounting period, however it could have been modelled by time series analysis.

References

- Boles T., Staudt A.,*On the Accuracy of Loss Reserving Methodology*, CAS E-Forum Fall 2010, pp. 43-104, 2010.
- Boles D., Mundlak Y.,*Estimation in Lognormal Lineer Models*, JASA Vol. 6, pp. 198-211, 1970.
- Brown R. L.,*Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance*, ACTEX Publications, 1993.
- Bühlmann H.,*Estimation of IBNR Reserves by the Methods of Chain-Ladder, Cape-Cod, and Complementary Loss Ratio*, International Summer School 1983, Unpublished, 1983.
- Bühlmann H., Schnieper R., Straub E.,*Claims Reserves in Casualty Insurance Based on a Probabilistic Model*, Mitt. SVVM 80, pp. 21-45, 1980.
- Choy S. T. B., Chan J. S. K., Makov U. E.,*Model Selection for Loss Reserves: The Growing Triangle Technique*, Life&Pensions, 2007.
- England P. D., Verrall R. J.,*Standard Errors of Prediction in Claims Reserving: A Comparison of Methods*, General Insurance Convention & Astin Colloquium, London: Institute of Actuaries, Vol. 1, pp. 459-478, 2002.
- Finney D. J.,*On the Distribution of a Variate Whose Logarithm is Normally Distributed*, JRSS Suppl. 7, pp. 155-161, 1941.

References (continued)

- Friedland J.,*Estimating Unpaid Claims Using Basic Techniques*, CAS-FCAS, KPMG LLP, Version II, 2009.
- Jing Y., Lebens J., Lowe S.,*Claim Reserving: Performance Testing and the Control Cycle*, CAS: Variance Advancing the Science of Risk, Vol. 3 Issue 2, 2009.
- Khury C. K. ,*Loss reserves: Performance Standards*, Proc. CAS 67, pp. 1-21, 1980.
- Klugman S. A., Panjer H. H., Willmot G. E.,*Loss Models: From Data to Decisions*, John Willey & Sons Ltd, 2008.
- Kremer E.,*IBNR Claims and the Two-way Model of ANOVA*, Scand. Actuar. J., pp. 47-55, 1982.
- Mack T.,*Distribution-free Calculation of the Standard Error of Chain-Ladder Reserve Estimates*, ASTIN Bull. 23, pp. 213-225, 1993.
- Mack T.,*The Standard Error of Chain-Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor*, ASTIN Bull. 29, pp. 361-366, 1999.
- Mahon M. J.,*The Scorecard System*, CAS Forum Summer 1997, Vol. 1, pp. 97-135, 1997.
- Miller I., Miller M.,*John E. Freunds Mathematical Statistics with Applications*, 7th Edition, Prentice Hall, NJ, 2004.
- Narayan P., Warthen T. V.,*A Comparative Study of the Performance of Loss Reserving Methods Through Simulation*, Journal of Actuarial Practice, Vol. 8, 2000.

References (continued)

- Pentikinen T., Rantala J.,*A Simulation Procedure for Comparing Different Claims Reserving Methods*, ASTIN Bull, Vol. 22, No. 2, pp. 191-216, 1992.
- Renshaw A. E., Verrall R. J.,*A Stochastic Model Underlying the Chain-Ladder Technique*, British Actuarial Journal, Vol. 4, pp. 903-923, 1998.
- Stanard J. N.,*A Simulation Test of Prediction Errors of Loss Reserving Techniques*, CAS Proceedings May 1985, Vol. 72, pp. 124-148, 1985.
- Taylor G. C.,*Loss Reserving : An Actuarial Perspective*, Boston-Dordrecht-London: Kluwer, 2000.
- Verrall R. J.,*On the Unbiased Estimation of Reserves from Loglinear Models*, Insurance: Mathematics and Economics, 10, 1, pp. 75-80, 1991.
- Verrall R. J.,*Statistical Methods for the Chain Ladder Technique*, CAS Forum Spring 1994, pp. 393-446, 1994.
- Verrall R. J.,*An Investigation into Stochastic Claims Reserving Models and the Chain-ladder Technique*, Insurance Math. Econom. 26, pp. 91-99, 2000.
- Wiser R. F.,*Loss reserving, Revised And Updated By Jo Ellen Cockley and Andrea Gardner*, FCAS, New York, pp. 197-285, 2001.
- Wüthrich M. V, Merz M.,*Stochastic Claims Reserving Methods in Insurance*, John Willey & Sons Ltd, 2008.

Thank you.

http://www.aktuerya.hacettepe.edu.tr/pers/EzgiNevruz.php

ezginevruz@hacettepe.edu.tr