Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: - (1980)

Heft: 1

Artikel: Claims reserves in casualty insurance based on a probabilistic model

Autor: Bühlmann, Hans / Schnieper, René / Straub, Erwin

DOI: https://doi.org/10.5169/seals-571098

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 26.03.2026

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

B. Wissenschaftliche Mitteilungen

HANS BÜHLMANN, RENÉ SCHNIEPER, and ERWIN STRAUB, Zürich

Claims Reserves in Casualty Insurance based on a Probabilistic Model

Dedicated to Dr. Max E. Eisenring – who strongly motivated us for this research – on the occasion of his 70th birthday.

1 The Purpose of this Paper

Before speaking of purpose, we should comment on the title. We have debated whether to say "Reserves in Non Life Insurance" or "Reserves in Casualty Insurance". The approach described here is, indeed, valid – possibly with some minor modifications – for all Non Life Insurance. On the other hand, a solid basis for claims reserving is mostly needed in the long tail business. Hence we have decided to restrict our attention to Casualty Insurance which allows us also – as an additional benefit – to use the standard terminology common in this sector of the insurance industry.

Since the early days of Life Insurance it has been understood that "reserves for future payments of claims (minus future receipts of premiums) had to be calculated from the probabilistic model describing the process of death within a specified population". William Morgan [1] has already made such valutations in 1786. Yes, the calculation of such reserves has since then become *one* central if not *the* central domain of the life actuary. Strangely enough when actuaries were asked to put their skill to work in Non Life Insurance, they did not feel it necessary to have a probabilistic model for the setting of Claims reserves. De Vylder [2] is the exception to the rule since he writes: "In this paper we adopt a rather deterministic approach, but we believe that the whole model can be probabilized . . .". We also just noted the paper by Hachemeister [3] in the Proceedings of the 21st International Congress of Actuaries. The reason for the absence of probabilistic models leading to reserving techniques in Casualty Insurance may be explained (to some extent) by the common fashion in this field of assuming the individual claim amount to "occur" suddenly even if in practice it is practically delayed portionwise over long periods of time. This paper takes exception to this fashion and models the individual claim amount as a random process over time. Claims reserves can then be calculated from the model.

2 The Individual Claim Amount

${}^{m}Z_{ij}^{(k)}$ stands for the *individual claim amount*, originating from claim number k of the accident year j (occurrence year j).
The left upper index (say m) indicates the year (called reporting year) in which the claim has been reported for the first time; the first lower index on the right (say i) gives the development year.
The following convention is used for the numbering:

The "accident year" is an element of the set $\{1, 2, 3, 4, 5, \ldots, n\}$ or in the Standard Example of the set $\{1970, 1971, \ldots, 1979\}$,
The "reporting year" is an element of the set $\{1, 2, 3, \ldots\}$,
The "development year" is also an element of $\{1, 2, 3, \ldots\}$,

with the interpretation, that the numbering of the reporting year as well as the development year starts with 1 in the year of occurrence of the individual claim.
In the accident year j N_j individual claims do occur. Take one of them, e.g. the one with number k. The variable $T_j^{(k)}$ then indicates the year in which this claim is reported for the first time. If $T_j^{(k)}$ takes on the value $m \in N$ then this claim generates a stream of claim amounts $({}^{m}Z_{ij}^{(k)})_{i \geqq m}$. ${}^{m}Z_j^{(k)} = \lim_{i \to \infty} {}^{m}Z_{ij}^{(k)}$ denotes the *final claim amount* of this claim. Obviously the sequence is only written as an infinite sequence for mathematical convenience, and the limit defining the final claim amount is reached after a finite number of years (e.g. 10 years).

3 Derived Quantities: Known Total of Claims Final Total of Claims

At the end of development year i we have then for *Known Total of Claims* (per end of development year i)

$$X_{ij} = \sum_{m=1}^{i} \sum_{k=1}^{N_j} \mathrm{I}_{[T_j^{(k)}=m]} \cdot {}^{m}Z_{ij}^{(k)} \doteqdot \sum_{m=1}^{i} {}^{m}X_{ij}$$

where

$$\mathrm{I}_A \doteqdot \begin{cases} 1 \text{ if } A \text{ occurs} \\ 0 \text{ if } A \text{ does not occur} \end{cases}$$

and where

$${}^{m}X_{ij} = \sum_{k=1}^{N_j} \mathrm{I}_{[T_j^{(k)}=m]} \, {}^{m}Z_{ij}^{(k)}$$

Final Total of Claims

$$X_j = \sum_{m=1}^{\infty} \sum_{k=1}^{N_j} \mathrm{I}_{[T_j^{(k)}=m]}\, {}^m Z_j^{(k)} \doteqdot \sum_{m=1}^{\infty} {}^m X_j$$

where

$${}^m X_j \doteqdot \sum_{k=1}^{N_j} \mathrm{I}_{[T_j^{(k)}=m]}\, {}^m Z_j^{(k)}.$$

Our interest will be focused on the difference between Final Total Claims and Known Total Claims. In a nutshell the whole purpose of the paper is that of "evaluating" this difference. Let us call it the Adjustment for the Total of Claims.

Adjustment for the Total of Claims (per end of development year i)

$$Y_{ij} = X_j - X_{ij} = \underbrace{\sum_{m=1}^{i} ({}^m X_j - {}^m X_{ij})}_{\Gamma_{ij}} + \underbrace{\sum_{m=i+1}^{\infty} \sum_{k=1}^{N_j} \mathrm{I}_{[T_j^{(k)}=m]}\, {}^m Z_j^{(k)}}_{\Delta_{ij}}$$

Γ_{ij} is usually called the adjustment for "incurred but not enough reported" (IBNER),
Δ_{ij} is usually called the adjustment for "incurred but not reported" (IBNR). However, in practice (but not in this paper) "IBNR" is sometimes also just used as synonym for the "Adjustment for the Total of Claims".
The spirit of our description is a probabilistic one (as indicated in the title and in section 1). Hence, all quantities described by capital letters Z, X, A, Γ, Δ, introduced so far, are to be understood as random variables. In particular Y_{ij} for all i and j are random variables. In the following section we are now describing our assumptions regarding the probability laws governing these random variables.

4 The Basic Probabilistic Assumptions

Of course, there are many different probability structures that one might propose. The choice which we have taken is the result of our struggle to combine intuition with mathematical convenience. Some of the basic hypotheses could actually be weakened. It is, however, mainly for reasons of a clear exposition that we prefer to stay with them as stated below:

(H_1) *Distribution of Number of Claims*

N_j, $j=1, 2, 3, \ldots$, are independent and Poisson distributed with parameter $V_j v$ where
V_j is a measure of *volume* for the accident year j, v is a real valued parameter.

(H_2) *Independence of Reporting and Frequency as well as Severity and Frequency*

The T and N variables represent two independent classes of random variables. The same holds for the Z and N variables.

(H_3) *Independence of Accident Year Experiences*

Events defined on different accident years are independent.

(H_4) *Random Variables within an Accident Year*

i) Claim amounts originating from different claims are independent, in particular the sequences of individual claim amounts

$({}^m Z_{ij}^{(k)})_{i \geqq m}$ for $k=1, 2, 3, \ldots$ are independent identically distributed.

For this reason the index k is omitted when we make statements about the distribution of ${}^m Z_{ij}$.

ii) $T_j^{(k)}$, $k=1, 2, 3, \ldots$, are independent and have all the same distribution function $F(t)$, or in different notation $p(m)=F(m)-F(m-1)$.

(H_5) *Stationarity of Growth Rates of Individual Claim Amounts*

$E[{}^m Z_{ij}^{(k)}/\text{given any path leading to } {}^m Z_{i-1\,j}^{(k)}=x]={}^m\lambda_{i-1}x$

i.e. the "growth rate" ${}^m\lambda_{i-1}$ does not depend on the accident year j.

$\mathrm{Var}[{}^m Z_{ij}^{(k)}/\text{given any path leading to } {}^m Z_{i-1\,j}^{(k)}=x]={}^m\sigma_{i-1}^2 f(x)$ for some function $f(x)$, $(f>0)$.

5 The Statistical Information

a) Reserving techniques actually used in the Casualty area start from the *Incurred Claims Triangle*. Say we have reached development year n for the

accident year 1, then the Incurred Claims Triangle has the following form

accident years

development years

$$\begin{array}{ccccc} X_{11} & X_{12} & \dots & X_{1\,n-1} & X_{1\,n} \\ X_{21} & X_{22} & & X_{2\,n-1} & \\ X_{31} & X_{32} & & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \cdot & X_{n-12} & & & \\ X_{n1} & & & & \\ \hline P_1 & P_2 & \dots & P_{n-1} & P_n \end{array}$$

where $X_{ij} = \sum_{m=1}^{i} {}^{m}X_{ij}$ (as defined in section 3) stands for the Known Total of Claims from accident year j per end of development year i, and P_j stands for the Premium Earned in year j.

It is convenient to use the abbreviations

$\triangle_{P}^{X}$ for such a triangle

$\triangle^{X}$ for the Incurred Claims Triangle without the last row (of earned premiums P)

b) Introducing the abbreviation $N_{ij} = \sum_{k=1}^{N_j} \mathrm{I}_{[T_j^{(k)} \leqslant i]}$ i.e. *number* of claims incurred in the year j and known per end of development year n, we can also form an *Incurred Number Triangle*

accident years

development years

$$\begin{array}{ccccc} N_{11} & N_{12} & \dots & N_{1\,n-1} & N_{1\,n} \\ N_{21} & & & N_{2\,n-1} & \\ N_{31} & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & N_{n-12} & & & \\ N_{n1} & & & & \\ \hline V_1 & V_2 & \dots & V_{n-1} & V_n \end{array}$$

where N_{ij} is defined as above and V_j stands for the measure of volume for the accident year j (as defined in section 4)

We use the corresponding abbreviations

$\triangle_{V}^{N}$

and

$\triangle^{N}$

c) Finally we want to split up the $\boxed{X}$ triangle according to reporting years into several $\boxed{{}^{m}X}$ $m=1, 2, 3, \ldots$ triangles, which explicitly written have the following shape

$$
\begin{array}{c|ccccccc}
 & \multicolumn{7}{c}{\overbrace{\hspace{10em}}^{n}} \\
m-1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
 & 0 & 0 & 0 & \ldots & 0 & 0 & \\
 & 0 & 0 & 0 & \ldots & 0 & 0 & \\
 & {}^{m}X_{m1} & {}^{m}X_{m2} & {}^{m}X_{m3} & \ldots & {}^{m}X_{mn+1-m} & & \\
 & {}^{m}X_{m+11} & {}^{m}X_{m+12} & \ldots & & & & \\
 & \cdot & & & & & & \\
 & \cdot & {}^{m}X_{n-12} & & & & & \\
 & \cdot & & & & & & \\
 & {}^{m}X_{n1} & & & & & &
\end{array}
$$

In the following we are considering the situation where all triangles

$$\boxed{X}_{P} \quad \boxed{{}^{m}X}, \; m=1, 2, 3, \ldots, \; \boxed{N}_{V} \quad \text{are given.}$$

Based on this information, we want to evaluate (for each accident year j) the *Adjustment for the Total of Claims*, namely the random variable $Y_{n-j+1,j}$ adjusting for the development from the "diagonal on downwards".

6 The Standard Example

The following explicit numerical example will be followed through the rest of the paper. It relates to the accident years $j=1970, 1971, \ldots, 1979$ with developments until the end of 1979.

a) $\boxed{N}_{V}$ triangle

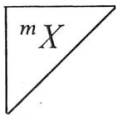
N_j is Poisson with parameter $V_j v$, where $v=0.1128$ and

$$
\begin{aligned}
V_{1970} &= 800 \\
V_{1971} &= 1000 \\
V_{1972} &= 700 \\
V_{1973} &= 600 \\
V_{1974} &= 500
\end{aligned}
$$

$$V_{1975} = 700$$
$$V_{1976} = 900$$
$$V_{1977} = 1200$$
$$V_{1978} = 1600$$
$$V_{1979} = 2000$$

The distribution of the reporting time is given as follows

$p(1) = 0.3$	$p(6) = 0.05$
$p(2) = 0.2$	$p(7) = 0.05$
$p(3) = 0.15$	$p(8) = 0.02$
$p(4) = 0.1$	$p(9) = 0.02$
$p(5) = 0.1$	$p(10) = 0.01$

b) 

triangles

The individual claim amount ${}^mZ_{mj}$ and its development is log normal, to be more precise

$$\log {}^mZ_{mj} \sim \mathcal{N}\,(\mu_m + (j - 1970) \ln (1+\delta),\ \sigma_0^2)$$

and the distribution of ${}^mZ_{ij}$ given the history $({}^mZ_{mj}, \ldots, {}^mZ_{i-1\,j})$ follows

$$\log {}^mZ_{ij} \sim \mathcal{N}\,(\gamma_{i-1} + \log {}^mZ_{i-1\,j},\ \gamma_{i-1}\sigma^2).$$

The conditional mean of ${}^mZ_{ij}$ given the history $({}^mZ_{mj}, \ldots, {}^mZ_{i-1\,j})$ is

$${}^mZ_{i-1\,j}\, e^{\gamma_{i-1}\left(1+\frac{\sigma^2}{2}\right)} \doteqdot {}^mZ_{i-1\,j}\lambda_{i-1}$$

and the conditional variance

$$({}^mZ_{i-1\,j})^2 e^{\gamma_{i-1}(2+\sigma^2)}\,(e^{\gamma_{i-1}\sigma^2} - 1) \doteqdot ({}^mZ_{i-1\,j})^2\,\sigma_{i-1}^2$$

[hence $f(x) = x^2$ in this case cf. section 4]

Observe:

$$\left.\begin{array}{ll} {}^m\lambda_{i-1} \doteqdot \lambda_{i-1} & \text{independent of } m \\ {}^m\sigma_{i-1}^2 \doteqdot \sigma_{i-1}^2 & \text{independent of } m \end{array}\right\} \text{ in the Standard Example.}$$

On the other hand, we have for the initial values in the reporting year m

$$E[{}^mZ_{mj}]=e^{\mu_m+\frac{\sigma_0{}^2}{2}}(1+\delta)^{j-1970}\doteqdot c_m(1+\delta)^{j-1970}$$

and

$$\mathrm{Var}\,[{}^mZ_{mj}]=c_m^2(1+\delta)^{2j-2}(e^{\sigma_0{}^2}-1).$$

c) *Simulation: The following values have been chosen for the parameters under b)*

$\delta=0.05$		$\sigma^2=10$			$\sigma_0^2=1$
m	μ_m	γ_m	λ_m	c_m	$E[{}^mZ_{1970}]$
1	1	0.018	1.114	4.482	7.319
2	1.2	0.015	1.094	5.474	8.024
3	1.3	0.013	1.081	6.050	8.107
4	1.4	0.011	1.068	6.686	8.295
5	1.5	0.009	1.055	7.389	8.576
6	1.6	0.007	1.043	8.166	8.957
7	1.65	0.005	1.030	8.585	9.056
8	1.7	0.003	1.018	9.025	9.243
9	1.73	0.001	1.006	9.300	9.356
10	1.75	—	—	9.488	9.488

i) Observe that from assumption (H_5) we have

$$E[{}^mZ_j]=c_m(1+\delta)^{j-1970}\prod_{i\geq m}\lambda_i.$$

ii) Note that in the above table δ, σ^2, σ_0^2, μ_m, and γ_m can be freely chosen, whereas λ_m, c_m, and $E[{}^mZ_{1970}]$ depend in a unique fashion upon those freely chosen parameters.

With these parameter values the following twelve triangles have been obtained:

N

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	38	37	15	20	13	25	33	47	50	50
2	52	64	28	38	25	35	55	81	80	
3	71	88	43	52	35	47	65	102		
4	82	97	44	57	39	54	80			
5	86	108	54	65	48	62				
6	91	116	60	71	48					
7	96	130	65	72						
8	99	132	68							
9	99	133								
10	99									
	800.00	1000.00	700.00	600.00	500.00	700.00	900.00	1200.00	1600.00	2000.00

X

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	157.46	167.32	69.15	116.93	44.01	118.42	153.78	331.68	438.93	310.47
2	240.25	322.34	136.85	252.90	137.55	164.26	277.68	612.18	860.09	
3	347.35	524.87	276.43	406.32	222.39	275.41	380.70	865.89		
4	407.66	624.92	331.74	524.74	272.82	401.87	573.80			
5	422.61	762.05	421.00	593.48	346.64	507.94				
6	629.02	874.52	501.14	767.00	350.32					
7	664.54	1186.99	583.34	835.84						
8	725.31	1266.42	600.50							
9	826.71	1247.30								
10	776.27									

${}^{m}X$ $m=1$

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	157.46	167.32	69.15	116.93	44.01	118.42	153.78	331.68	438.93	310.47
2	187.64	185.54	59.57	148.18	57.54	114.71	153.77	410.40	562.41	
3	190.88	209.28	83.81	145.23	54.74	120.58	177.49	437.08		
4	197.96	210.75	86.39	180.59	45.83	151.99	177.32			
5	194.91	190.74	80.80	165.17	51.92	155.70				
6	208.98	185.15	81.15	227.99	52.51					
7	235.75	188.23	86.31	296.98						
8	269.10	188.82	82.85							
9	308.00	186.69								
10	297.11									

${}^{m}X$	$m=2$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	52.61	137.30	77.27	104.72	80.00	49.55	123.91	201.77	297.68	
3	64.02	173.55	82.39	133.14	76.79	58.68	143.89	254.20		
4	70.59	239.46	87.86	136.75	93.78	72.54	141.98			
5	85.24	306.60	76.75	165.99	110.39	79.76				
6	85.60	313.05	76.86	185.12	95.44					
7	82.72	328.38	101.33	193.09						
8	79.77	320.95	106.41							
9	87.46	298.10								
10	86.00									

${}^{m}X$	$m=3$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	92.44	142.04	110.23	127.95	90.86	96.15	59.32	174.61		
4	79.41	134.90	121.11	186.71	100.44	99.68	82.03			
5	72.21	136.68	126.76	181.62	100.27	90.18				
6	70.97	115.39	138.12	196.21	122.82					
7	64.08	128.58	142.01	170.20						
8	62.82	143.45	134.08							
9	63.85	145.78								
10	67.10									

${}^{m}X$	$m=4$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0		
4	59.69	39.81	36.38	20.69	32.77	77.67	172.47			
5	55.20	44.65	40.24	28.22	22.16	119.87				
6	56.63	59.07	51.79	32.80	20.44					
7	60.14	83.44	84.69	47.77						
8	69.15	102.05	89.17							
9	73.37	92.96								
10	68.48									

${}^{m}X$	$m=5$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0		
4	0	0	0	0	0	0	0			
5	15.05	83.38	96.45	52.48	61.92	62.43				
6	15.90	103.10	105.62	57.59	59.10					
7	17.94	107.18	101.79	51.91						
8	15.61	128.29	95.50							
9	15.41	128.13								
10	14.61									

${}^{m}X$	$m=6$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0		
4	0	0	0	0	0	0	0			
5	0	0	0	0	0	0				
6	190.93	98.76	47.59	67.31	0					
7	175.13	102.52	37.46	73.94						
8	169.70	116.46	41.56							
9	213.17	100.26								
10	178.85									

${}^{m}X$	$m=7$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0		
4	0	0	0	0	0	0	0			
5	0	0	0	0	0	0				
6	0	0	0	0	0					
7	28.79	248.67	29.75	1.96						
8	35.30	252.93	30.13							
9	39.84	274.01								
10	39.97									

${}^{m}X$	$m=8$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0		
4	0	0	0	0	0	0	0			
5	0	0	0	0	0	0				
6	0	0	0	0	0					
7	0	0	0	0						
8	23.86	7.47	20.80							
9	25.61	9.27								
10	24.15									

${}^{m}X$	$m=9$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0		
4	0	0	0	0	0	0	0			
5	0	0	0	0	0	0				
6	0	0	0	0	0					
7	0	0	0	0						
8	0	0	0							
9	0	12.10								
10	0									

${}^{m}X$	$m=10$									
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0		
4	0	0	0	0	0	0	0			
5	0	0	0	0	0	0				
6	0	0	0	0	0					
7	0	0	0	0						
8	0	0	0							
9	0	0								
10	0									

7 Valuation of the Adjustment for the Total of Claims

Valuation in Life Insurance has become such a standard technique that its meaning may have been forgotten by the practical actuary who does valuations as part of his routine. For this reason we want to remind the reader that valuation means nothing else than taking the expected value over the random variable which describes the (possibly discounted) stream of future payments. Of course, expected values should always take into account the latest information available. In terms of probability theory this means to take the *conditional* expectation, given the latest information.

The valuation of the Adjustment for the Total of Claims on the diagonal

$$Y_{n-j+1,j} = \Gamma_{n-j+1,j} + \Delta_{n-j+1,j}$$

is thus described rather easily as the problem of finding the conditional expectation

$$E\left[Y_{n-j+1,j} \Big/ \begin{array}{l} {}^{m}X \quad (m=1,2,3,\ldots) \\ N \end{array}\right] = E\left[\Gamma_{n-j+1,j} \Big/ \begin{array}{l} {}^{m}X \quad (m=1,2,\ldots) \\ N \end{array}\right]$$

$$+ E\left[\Delta_{n-j+1,j} \Big/ \begin{array}{l} {}^{m}X \quad (m=1,2,\ldots) \\ N \end{array}\right]$$

A straight forward calculation with conditional expectations leads from the definition of $\Gamma_{n-j+1,j}$, $\Delta_{n-j+1,j}$ to the formulae

$$\text{(I)} \quad E\left[\underbrace{\Gamma_{n-j+1,j}}_{\tilde{n}} \Big/ \begin{array}{l} {}^{m}X \quad (m=1,2,\ldots) \\ N \end{array}\right] = \sum_{m=1}^{\tilde{n}} \underbrace{\left(\prod_{i \geq \tilde{n}} {}^{m}\lambda_i - 1\right)}_{{}^{m}H_{\tilde{n}} - 1} {}^{m}X_{\tilde{n}j}$$

$$= \sum_{m=1}^{\tilde{n}} ({}^{m}H_{\tilde{n}} - 1)\, {}^{m}X_{\tilde{n}j}$$

$$\text{(II)} \quad E\left[\underbrace{\Delta_{n-j+1,j}}_{\tilde{n}} \Big/ \begin{array}{l} {}^{m}X \quad (m=1,2,\ldots) \\ N \end{array}\right] = \sum_{m=\tilde{n}+1}^{\infty} p(m) E[{}^{m}Z_j] \cdot V_j \cdot v.$$

I and II are the "valuations" for "IBNER" and "IBNR" respectively. Observe that as in Life Assurance "Valuation" in the sense used here is only establishing the "center of gravity" for future obligations. If one wanted to have information regarding possible fluctuations one should also calculate variances (and possibly higher moments) of the random variables in question. But it seems important to us that such considerations be only made when calculating e.g. catastrophe reserves or safety loadings for ordinary claim reserves but *not* for the ordinary claims reserves themselves.

8 Valuation for "IBNER"

The basic formula has been derived in the previous section.

$$\text{(I)} \quad E\left[\Gamma_{\underbrace{n-j+1}_{\tilde{n}},j} \Big/ \begin{array}{l} \boxed{{}^{m}X} \quad (m=1,2,\ldots) \\ \\ \boxed{N} \end{array}\right] = \sum_{m=1}^{\tilde{n}} ({}^{m}H_{\tilde{n}}-1)\,{}^{m}X_{\tilde{n}j}.$$

It is interesting to note that under the *additional hypothesis* of growth rates of individual claims ${}^{m}\lambda_i$ being *independent of the reporting year m*, i.e. ${}^{m}\lambda_i \equiv \lambda_i$ for all m we can further simplify and obtain with $H_{\tilde{n}} = \prod_{i \geqslant \tilde{n}+1} \lambda_i$

$$\text{(I}'\text{)} \quad E\left[\Gamma_{\tilde{n}j} \Big/ \begin{array}{l} \boxed{{}^{m}X} \quad (m=1,2,\ldots) \\ \\ \boxed{N} \end{array}\right] = (H_{\tilde{n}}-1)X_{\tilde{n}j}.$$

Observe that this last formula corresponds to the most common approach (based on lag factors) for evaluating the Adjustment for the Total of Claims. Our analysis shows that this approach is apparently justifiable within our model provided the adjustment consists of the IBNER component only.
The valuation of IBNER is carried through in section 10 for our Standard Example (where the additional hypothesis ${}^{m}\lambda_i \equiv \lambda_i$ for all m holds). Based on the true parameter values one obtains the exact IBNER Reserve from the formula $(H_{\tilde{n}}-1)X_{\tilde{n}j}$.

9 Valuation for "IBNR"

a) The basic formula has been derived in section 7.

$$\text{(II)}\quad E\left[\Delta_{\tilde{n}j} \Big/ \begin{array}{l} \boxed{{}^{m}X} \quad (m=1, 2, \ldots) \\ \\ \boxed{N} \end{array}\right] = \sum_{m=\tilde{n}+1}^{\infty} p(m) E[{}^{m}Z_j] V_j v.$$

We rewrite it somewhat differently by putting

$$\sum_{m=\tilde{n}+1}^{\infty} p(m) E[{}^{m}Z_j] = \underbrace{E[{}^{>\tilde{n}}Z_j]}_{\substack{\text{expected value} \\ \text{for claims report-} \\ \text{ed after } \tilde{n}}} [1-F(\tilde{n})]$$

Under the additional hypothesis that premiums are correct, i.e.

$$P_j = E[Z_j] \cdot V_j \cdot v$$

we obtain

$$\text{(II)}'\quad E\left[\Delta_{\tilde{n}j} \Big/ \begin{array}{l} \boxed{{}^{m}X} \quad (m=1, 2, \ldots) \\ \\ \boxed{N} \end{array}\right] = \frac{E[{}^{>\tilde{n}}Z_j]}{E[Z_j]} [1-F(\tilde{n})] P_j.$$

If even $E[{}^{m}Z_j] = E[Z_j]$, for all m, is satisfied, we arrive at the simplified formula.

$$\text{(II)}''\quad E\left[\Delta_{\tilde{n}j} \Big/ \begin{array}{l} \boxed{{}^{m}X} \quad (m=1, 2, \ldots) \\ \\ \boxed{N} \end{array}\right] = [1-F(\tilde{n})] P_j.$$

b) In section 10 the valuation of IBNR is explicitly carried through for our standard example. The formula used there is

$$\sum_{m=\tilde{n}+1}^{\infty} p(m) E[{}^{m}Z_j] v V_j.$$

10 True Reserve Valuation for the Standard Example

The great advantage of our approach consists in the fact that for the Standard Example, described in section 6, – on the contrary to the situation encountered in practice – *we know the true parameter values*. This leads to the following *true reserves*.

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
IBNER	0	7.51	14.59	46.38	35.30	82.20	138.34	295.95	402.65	197.33
IBNR	0	11.24	24.55	36.58	63.07	138.00	303.09	582.38	1140.00	2058.78
Total	0	18.74	39.13	82.96	98.36	220.20	441.44	878.33	1542.65	2256.11

These true values should be compared with the estimated values obtained from different estimation methods.
As a first trial let us compare the obtained total with the one calculated by the standard method based on lag factors.

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Total	0	−76.11	−11.57	34.85	80.31	257.76	456.50	1138.74	2203.40	1769.15

The result is rather discouraging! In the following we are trying to do better than the standard method by estimating the components of the formulae I and II according to techniques of mathematical statistics.

11 On the Search for Better Methods of Estimation; the *M*-method

Our aim in this section is that of proposing estimates for the compounds of the formulae (see section 7)

IBNER

$$\text{(I)}\quad E\left[\Gamma_{\tilde{n}j} \Big/ \begin{array}{l} \boxed{{}^{m}X} \quad (m=1, 2, \ldots) \\ \boxed{N} \end{array}\right] = \sum_{m=1}^{\tilde{n}} ({}^{m}H_{\tilde{n}}-1)\,{}^{m}X_{\tilde{n}j},$$

where
$${}^{m}H_{\tilde{n}} = \prod_{i \geqslant \tilde{n}} {}^{m}\lambda_i .$$

IBNR

$$\text{(II)}\quad E\left[\Delta_{\tilde{n}j} \Big/ \begin{matrix} \boxed{{}^{m}X} \\ \boxed{N} \end{matrix} \quad (m=1, 2, \ldots)\right] = \sum_{m=\tilde{n}+1}^{\infty} p(m) E[{}^{m}Z_j] V_j v.$$

As a preparation, we build one more triangle (the $\boxed{U}$ triangle) by defining $U_{mj} = N_{mj} - N_{m-1\,j}$ $(N_{0j}=0)$.

a) *Estimates of Components of IBNER Formula*

All that is needed are the estimates for the ${}^{m}\lambda_i$. We propose

$$\widehat{{}^{m}\lambda_{i-1}} = \frac{\sum_{j=1}^{n+1-i} \frac{{}^{m}X_{ij}\,{}^{m}X_{i-1\,j}}{U_{mj}}}{\sum_{j=1}^{n+1-i} \frac{({}^{m}X_{i-1})^2}{U_{mj}}} \tag{1}$$

Sketch of Derivation of the Estimator:

${}^{m}X_{ij}$ given, the development history until year $i-1$ has according to (H_5)

conditional expectation ${}^{m}\lambda_{i-1}\,{}^{m}X_{i-1\,j}$

conditional variance ${}^{m}\sigma^2_{i-1} \sum_{k=1}^{U_{mj}} f({}^{m}Z^{(k)}_{i-1\,j}) \approx K\,{}^{m}\sigma^2_{i-1} U_{mj}$

(because we do not know f).

Conditionally $\frac{{}^{m}X_{ij}}{{}^{m}X_{i-1\,j}}$ is an unbiased estimator for ${}^{m}\lambda_{i-1}$ with variance $\approx$

$\approx K\ {}^{m}\sigma^2_{i-1} \frac{U_{mj}}{({}^{m}X_{i-1\,j})^2}$.

Our proposed estimate (1) is then $\sum_j a_j \frac{{}^{m}X_{ij}}{{}^{m}X_{i-1\,j}}$ with $\sum_j a_j = 1$, and a_j proportional to $\frac{1}{\text{variance}}$.

P.S. If one knows that ${}^m\lambda_i = \lambda_i$ independent of m, formula (1) might be improved by summing the numerator and denominator of the right hand side also over m.

b) *Estimates of Components of IBNR Formula*

b_1)

$$\left.\begin{aligned} \widehat{p(m)v} &= \frac{\sum_{j=1}^{n+1-m} U_{mj}}{\sum_{j=1}^{n+1-m} V_j} \\ \hat{v} &= \sum_{m=1}^{\infty} \widehat{p(m)v} \end{aligned}\right\} \tag{2}$$

No comment needed.

b_2) Estimate for $E[{}^mZ_j] = E[{}^mZ_{mj}] \prod_{i \geqslant m} {}^m\lambda_i$

Assume: $E[{}^mZ_{mj}] = c_m(1+\delta)^{j-1}$.

The estimates $\hat{\delta}$ and $\hat{c}_m$ are defined as the solutions of the problem

$$Q(\hat{\delta}, \hat{c}_m) = \sum_{\substack{m,j \\ m+j \leqslant n+1}} \left(\frac{{}^mX_{mj}}{U_{mj}} - \hat{c}_m(1+\hat{\delta})^{j-1}\right)^2 U_{mj} = \min!$$

The solutions are obtained as follows:

$$\left.\begin{aligned} &\text{For given } \delta\colon c_m(\delta) = \frac{\sum_{j=1}^{n+1-m} {}^mX_{mj}(1+\delta)^{j-1}}{\sum_{j=1}^{n+1-m} U_{mj}[(1+\delta)^{j-1}]^2} \\ &\text{Choose } \hat{\delta} \text{ such that } Q(\hat{\delta}, c_m(\hat{\delta})) = \min! \\ &\hat{c}_m = c_m(\hat{\delta}) \end{aligned}\right\} \tag{3}$$

The estimation method described here, based on formulae (1), (2), and (3), is called M-method in the rest of the paper.

12 Application of the M-method to the Standard Example

The M-method leads to the following valuation results – to be compared with the true values and the values obtained by the standard method (both exposed in section 10).

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
IBNER	0	−148.29	−34.82	−25.19	15.30	69.10	115.67	302.20	429.39	248.43
IBNR	0	0	4.35	16.84	70.20	197.59	370.30	667.10	1311.77	2246.22
Total	0	−148.29	−30.47	− 8.35	85.50	266.69	485.96	969.30	1741.16	2494.65

In the last 3 years, we seem to be doing substantially better by the M-method than by the standard method (S-method). Of course, this is only a vague indication of quality. The comparison of the quality of estimators *cannot be made on the basis of a single simulation only*. We shall come back to the aspect of quality of different estimators in section 15. It is instructive to compare the M-method with the true values not only by means of the resulting IBNER and IBNR valuation results, but by means of the components of the estimators as well. In our standard example, these comparisons look as follows. (Observe that in the Standard Example we have ${}^m\lambda_i = \lambda_i$ for all m.)

m	a) λ_m		b_1) $p(m)v$		b_2) c_m		
	true	M-method	true	M-method	true	M-method	
1	1.1140	1.2007	0.0338	0.0328	4.482	4.51	$\delta_{\text{true}} = 0.05$
2	1.0942	1.1114	0.0226	0.0225	5.474	5.01	$\hat{\delta} = 0.05$
3	1.0811	1.1227	0.0169	0.0195	6.050	6.07	
4	1.0682	1.0577	0.0113	0.0100	6.686	7.33	
5	1.0555	1.0885	0.0113	0.0166	7.389	6.50	
6	1.0429	1.0761	0.0056	0.0069	8.166	14.58	
7	1.0305	1.0296	0.0056	0.0081	8.585	11.67	
8	1.0182	1.0691	0.0023	0.0032	9.025	6.17	
9	1.0060	0.8811	0.0023	0.0006	9.300	11.53	
10	—	—	0.0011	0.0000	9.988	—	

These comparisons teach us an interesting lesson. The "weak compound" in our estimation formula ((II) for IBNR) is apparently the estimate for c_m in the late reporting years! Here we have rather few claims for estimating the mean of them sufficiently well. To overcome this difficulty, we propose two alterations of the M-method. Both of them use an additional a priori assumption!

13 Alterations of the M-method

The alterations of the M-method only occur in the estimate (3).

a) *The M_1-method* (initial expected values increasing)

This method assumes the additional a priori hypothesis that the parameters c_m are monotone increasing.
Hence we have to solve

Problem:

$$Q(\hat{\delta}, \hat{c}_m) = \sum_{\substack{m,j \\ m+j \leqslant n+1}} \left(\frac{{}^m X_{mj}}{U_{mj}} - \hat{c}_m (1+\hat{\delta})^{j-1} \right)^2 U_{mj} = \min!,$$

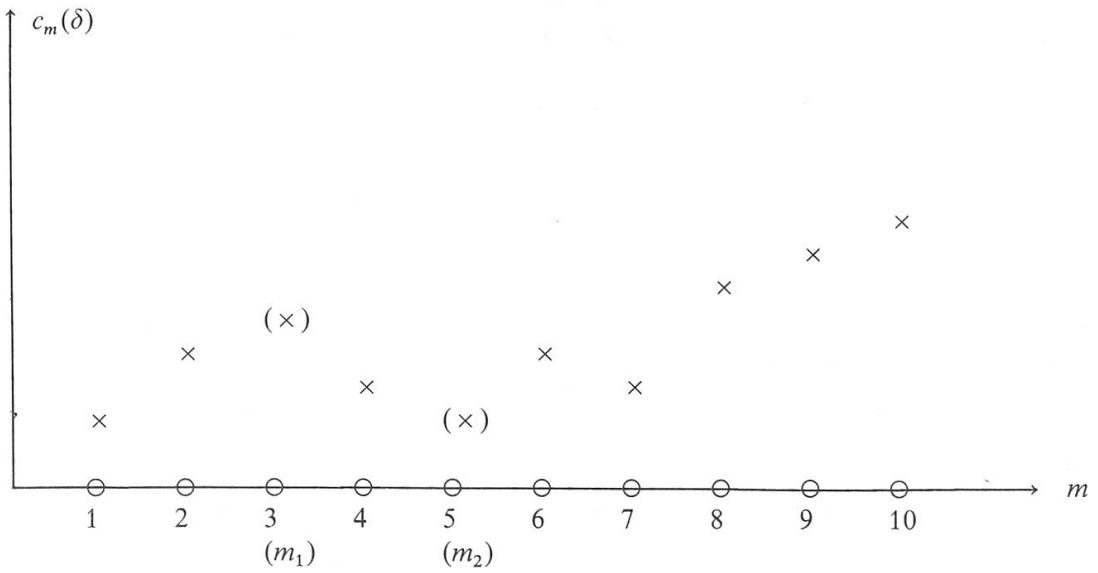
under the side condition that $c_m \leqslant c_{m+1}$ for all m.

Solution:

We proceed as before (section 11) by first assuming that δ is given.
Then

$$c_m(\delta) = \frac{\sum_j {}^m X_{mj} (1+\delta)^{j-1}}{\sum_j U_{mj} [(1+\delta)^{j-1}]^2}.$$

The calculation may (for a given δ) lead to the following result



The first local maximum is reached at m_1, the next local minimum at m_2. In this case, all coefficients $c_m(\delta)$, $m_1 \leqslant m \leqslant m_2$, are replaced by the *same new coefficient* $c_m^*(\delta)$, where

$$c_m^*(\delta) = \frac{\sum_{m=m_1}^{m_2} \sum_j {}^m X_{mj}(1+\delta)^{j-1}}{\sum_{m=m_1}^{m_2} \sum_j U_{mj}[(1+\delta)^{j-1}]^2}$$

and the procedure is repeated until we end with a monotone sequence $c_1(\delta) \leqslant \leqslant c_2(\delta) \leqslant \ldots \leqslant c_m(\delta) \leqslant c_{m+1}(\delta) \leqslant \ldots \leqslant c_n(\delta)$.
The proof that this represents a solution to the problem is left to the reader.

b) *The M_2-method* (final expected values increasing)

Here, the a priori hypothesis is even stronger. We postulate that $E[{}^mZ_j]$ are monotonically increasing for every j. Since $E[{}^mZ_j] = E[{}^mZ_1]\,(1+\delta)^{j-1}$, it suffices to estimate the sequence

$$E[{}^1Z_1],\ E[{}^2Z_1],\ E[{}^3Z_1],\ \ldots,\ E[{}^nZ_1].$$

We find for a given δ

$$E[{}^mZ_1]\,(\delta) = \frac{\sum_j {}^m X_{mj}(1+\delta)^{j-1} \prod_{i \geqslant m} {}^m\lambda_i}{\sum_j U_{mj}[(1+\delta)^{j-1}]^2},$$

and we achieve monotonicity by successively summing over the m groups between m_1 and m_2 exactly in the same way as we did under a).

14 Application of the M_1- and M_2-methods to the Standard Example

Without further comments, the following tables show the results of this application.

M_1-method

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
IBNER	0	−148.29	−34.82	−25.19	15.30	69.10	115.67	302.20	429.39	248.43
IBNR	0	0	4.63	30.09	85.90	209.05	398.43	709.07	1394.87	2392.27
Total	0	−148.29	−30.20	4.89	101.20	278.15	514.10	1011.26	1824.25	2640.70

M_2-method

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
IBNER	0	−148.29	−34.82	−25.19	15.30	69.10	115.67	302.20	429.39	248.43
IBNR	0	0	5.29	32.81	92.07	218.96	421.48	748.84	1475.71	2560.95
Total	0	−148.29	−29.53	7.61	107.38	288.07	537.15	1051.04	1905.10	2809.38

Comparison of Components ($\hat{\lambda}_m$ and $\widehat{p(m) \cdot v}$ as in the M-method)

	c_m		
	true	M_1	M_2
1970	4.482	4.264	3.943
1971	5.474	4.773	4.734
1972	6.050	5.845	5.624
1973	6.686	6.732	6.372
1974	7.389	6.732	6.739
1975	8.166	12.015	11.388
1976	8.585	12.015	12.255
1977	9.025	12.015	12.617
1978	9.300	12.015	13.489
1979	9.988	—	—
		$\hat{\delta}=0.06$	$\hat{\delta}=0.07$

It seems that the transition from M to the modified M_1- and M_2-methods introduces a bias towards overreserving. This will need to be tested in the next section.

15 Quality of Estimates

To get an idea of the quality of the methods S, M, M_1, and M_2 discussed in this paper, we have made 50 simulations of the Standard Example, all with the same model parameters. It could, of course, be possible that under completely different parameter selections, the quality of the estimators might be judged differently. We attach to this theoretical possibility rather little weight, particularly since we believe that our choice of the parameters is typical for that practical situation where the need for good estimators is particularly felt. With higher Poisson parameters and higher volumina, all methods will eventually lead to reasonable results.

For each accident year, we have defined the following measures of deviation:

$$(\text{Estimated Reserve}) - (\text{True Reserve}) = D.$$

D_s stands for this difference obtained from simulation s. We thus define

$$\frac{\sum_{s=1}^{50} D_s}{50} \text{ as bias of the estimate: } B$$

$$\sqrt{\frac{\sum_{s=1}^{50} D_s^2}{49}} \text{ as standard error of the estimate: } SE$$

$$\frac{\sum_{s=1}^{50} (\text{True Reserve})_s}{50} \text{ as true average reserve: } T.$$

The three following tables summarize our results:

Table 1 : True Average Reserves

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
IBNER	0	5.86	15.54	30.32	46.15	90.45	156.36	244.51	336.26	308.10
IBNR	0	11.24	24.55	36.58	63.07	138.00	303.09	582.38	1140.00	2058.78
Total	0	17.09	40.09	66.90	109.22	228.45	459.45	826.89	1476.27	2366.88

Table 2 : BIAS

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
S-method										
Total	0	−2.82	−2.60	−5.69	−6.40	−13.71	−17.55	−29.26	7.78	− 20.56
M-method										
IBNER	0	−3.71	−4.85	−3.77	−2.95	−0.97	8.51	6.52	1.79	−3.02
IBNR	0	0.40	0.30	−2.21	−4.24	−8.06	−13.09	−23.45	−17.76	−31.09
Total	0	−3.31	−4.56	−5.98	−7.19	− 9.03	− 4.58	−16.93	−15.98	−34.11
M_1-method										
IBNER	0	−3.71	−4.85	−3.77	−2.95	−0.97	8.51	6.52	1.79	−3.02
IBNR	0	3.82	5.35	4.77	4.06	3.82	5.57	8.56	34.34	52.85
Total	0	0.11	0.50	0.99	1.11	2.85	14.09	15.08	36.13	49.83
M_2-method										
IBNER	0	−3.71	−4.85	−3.77	−2.95	−0.97	8.51	6.52	1.79	−3.02
IBNR	0	4.15	6.85	7.08	8.49	13.05	23.07	39.70	83.51	126.96
Total	0	0.44	2.00	3.31	5.54	12.08	31.58	46.22	85.29	123.94

Table 3 : Standard Error

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
S-method										
Total	0	31.29	31.99	32.27	43.75	63.63	117.18	159.56	291.25	520.81
M-method										
IBNER	0	31.72	30.09	35.04	40.76	67.54	162.94	156.14	191.80	158.47
IBNR	0	20.31	19.55	20.11	22.66	37.24	60.63	99.21	191.72	336.91
Total	0	40.48	37.68	39.19	48.04	80.09	187.77	205.68	327.92	444.51
M_1-method										
IBNER	0	31.72	30.09	35.04	40.76	67.54	162.94	156.14	191.80	158.47
IBNR	0	21.43	19.76	19.82	22.40	36.73	60.93	106.67	211.39	377.64
Total	0	41.57	36.80	38.48	47.60	80.81	190.69	221.16	349.97	487.26
M_2-method										
IBNER	0	31.72	30.09	35.04	40.76	67.54	162.94	156.14	191.80	158.47
IBNR	0	21.70	20.50	20.90	25.52	44.11	79.84	141.92	268.12	471.25
Total	0	41.81	35.07	40.17	52.06	93.71	221.71	267.21	414.69	588.15

The conclusions from these tables are somewhat surprising:

1. The standard method of estimation is not so bad after all.
2. The standard deviation is quite high for all methods.
3. The *M*-method seems better suited for the very last accident year.

We thus feel that the search for better methods should still go on. Or, is the problem such that the standard deviation of the estimates cannot be substantially improved?

Bibliography

[1] Morgan, William, Valuation (individually) of the assurance contracts in force in 1786, Journal of the Institute of Actuaries, vol. 100, no. 415 (1974).

[2] de Vylder, F., Estimation of IBNR claims by least squares, Bulletin of the Association of Swiss Actuaries, vol. 78/2 (1978).

[3] Hachemeister, Ch., A stochastic model for loss reserving, Transactions 21st International Congress of Actuaries, vol. 1 (1980).

Prof. Dr. H. Bühlmann und R. Schnieper
Abt. Mathematik
ETH-Zentrum
8092 Zürich

Dr. E. Straub
Schweizerische Rück-
versicherungs-Gesellschaft
8022 Zürich

Summary

Based on a detailed stochastic model, IBNR reserves are estimated and in doing so two different procedures are obtained, one for IBNR (incurred but not reported) and one for IBNER (incurred – and reported – but not enough reserved). The application of four distinct methods (one of them being the classical "lag factor" method) on simulated run off figures shows rather high variances of the estimators. Furthermore, it can be seen that the classical methods lead to surprisingly good results on the numerical data under consideration.

Zusammenfassung

Aufgrund eines detaillierten stochastischen Modells werden Spätschadenreserven geschätzt, wobei sich für IBNR (incurred but not reported) und IBNER (incurred – and reported – but not enough reserved) zwei unterschiedliche Schätzmethoden ergeben. Die Anwendung von vier verschiedenen Verfahren (einschliesslich der klassischen "lag factor" Methode) auf simulierten Abwicklungsstatistiken liefert überraschend hohe Varianzen der Schätzwerte; zudem schneidet im verwendeten Beispiel die klassische Methode erstaunlich gut ab.

Résumé

L'auteur propose, sur la base d'un modèle aléatoire détaillé, une procédure d'estimation des provisions IBNR, tant pour la part IBNR (incurred but not reported) proprement dite que pour la part IBNER (incurred – and reported – but not enough reserved). L'application de 4 méthodes (dont la méthode classique du «lag-factor») à des échantillons établis par simulation fait ressortir, pour les estimateurs, des variances assez importantes. De plus, l'article montre que la méthode classique conduit à des résultats étonnamment bons, pour les échantillons considérés.