Which Stochastic Model is Underlying the Chain Ladder Method?

by Thomas Mack, Ph.D.

WHICH STOCHASTIC MODEL IS UNDERLYING THE CHAIN LADDER METHOD?

BY THOMAS MACK, PH.D., MUNICH RE

Editor's Note: This paper was presented to the XXIV ASTIN Colloquim, in Cambridge in 1993. Also, this paper was awarded the first-ever CAS Charles A. Hachemeister Prize in November 1994.

Abstract:

The usual chain ladder method is a deterministic claims reserving method. In the last years, a stochastic loglinear approximation to the chain ladder method has been used by several authors especially in order to quantify the variability of the estimated claims reserves. Although the reserves estimated by both methods are clearly different, the loglinear approximation has been called "chain ladder," too, by these authors.

In this note, we show that a different distribution-free stochastic model is underlying the chain ladder method; i.e. yields exactly the same claims reserves as the usual chain ladder method. Moreover, a comparison of this stochastic model with the above-mentioned loglinear approximation reveals that the two models rely on different philosophies on the claims process. Because of these fundamental differences the loglinear approximation deviates from the usual chain ladder method in a decisive way and should therefore not be called "chain ladder" any more.

Finally, in the appendix it is shown that the loglinear approximation is much more volatile than the usual chain ladder method.

1. *The usual deterministic chain ladder method*

Let C_{ik} denote the accumulated claims amount of accident year i, $1 \le i \le n$, either paid or incurred up to development year k, $1 \le k \le n$. The values of C_{ik} for $i + k \le n + 1$ are known to us (run-off triangle) and we want to estimate the values of C_{ik} for $i + k > n + 1$, in particular the ultimate claims amount C_{in} of each accident year $i = 2, ..., n$.

The chain ladder method consists of estimating the unknown amounts C_{ik}, $i + k > n + 1$, by

(1) $$\hat{C}_{ik} = C_{i,n+1-i}\hat{f}_{n+1-i} \times \ldots \times \hat{f}_{k-1}, \quad i + k > n + 1,$$

where

(2) $$\hat{f}_k = \sum_{j=1}^{n-k} C_{j,k+1} \Big/ \sum_{j=1}^{n-k} C_{jk}, \quad 1 \le k \le n - 1.$$

For many years this has been used as a self-explaining deterministic algorithm which was not derived from a stochastic model. In order to quantify the variability of the estimated ultimate claims amounts, there have been several attempts to find a stochastic model underlying the chain ladder method. Some of these will be reviewed in the following chapter.

2. *Some stochastic models related to the chain ladder method*

In order to find a stochastic model underlying the chain ladder method we have to cast the central equation (1) of the chain ladder method into stochastic terms. One way of doing this runs along the following lines: We conclude from (1) that

$$\hat{C}_{i,k+1} = \hat{C}_{ik}\,\hat{f}_k, \quad k > n+1-i.$$

This is generalized to the stochastic model

$$E(C_{i,k+1}) = E(C_{ik})\,f_k, \quad 1 \leq k \leq n-1, \tag{3}$$

where all C_{ik} are considered to be random variables and $f_1, \ldots, f_{n-1}$ to be unknown parameters.

Introducing the incremental amounts

$$S_{ik} = C_{ik} - C_{i,k-1}, \quad 1 \leq i, k \leq n,$$

with the convention $C_{i0} = 0$, one can show that model (3) is equivalent to the following model for S_{ik}:

$$E(S_{ik}) = x_i y_k, \quad 1 \leq i, k \leq n, \tag{4}$$

with unknown parameters x_i, $1 \leq i \leq n$, and y_k, $1 \leq k \leq n$, with $y_1 + \ldots + y_n = 1$.

Proof of the equivalence of (3) and (4):

(3)==> (4): Successive application of (3) yields

$$E(C_{in}) = E(C_{ik}) f_k \times \ldots \times f_{n-1}$$

Because

$$E(S_{ik}) = E(C_{i,k}) - E(C_{i,k-1})$$

$$= E(C_{in})((f_k \times \ldots \times f_{n-1})^{-1} - (f_{k-1} \times \ldots \times f_{n-1})^{-1})$$

we obtain (4) by defining

$$x_i = \mathrm{E}(C_{in}), \quad 1 \le i \le n,$$

$$y_i = (f_1 \times \ldots \times f_{n-1})^{-1}$$

$$y_k = (f_k \times \ldots \times f_{n-1})^{-1} - (f_{k-1} \times \ldots \times f_{n-1})^{-1}, \quad 2 \le k \le n-1,$$

$$y_n = 1 - (f_{n-1})^{-1}.$$

This definition fulfills $y_1 + . + y_n = 1$.

(4) ==> (3): we have

$$\mathrm{E}(C_{ik}) = \mathrm{E}(S_{i1}) + \ldots + \mathrm{E}(S_{ik})$$

$$= x_i\,(y_1 + \quad + y_k)$$

and therefore

$$\frac{\mathrm{E}(C_{i,k+1})}{\mathrm{E}(C_{ik})} = \frac{y_1 + \ldots + y_k + y_{k+1}}{y_1 + \ldots + y_k} =: f_k, \quad 1 \le k \le n-1.$$

The stochastic model (4) clearly has $2n-1$ free parameters x_i, y_k. Due to the equivalence of (3) and (4) one concludes that also model (3) must have 2n $-$ 1 parameters. One immediately sees $n - 1$ parameters $f_1, \ . \ . \ , f_{n-1}$. The other n parameters become visible if we look at the proof (3) ==> (4). It shows that the level of each accident year i, here measured by $x_i = E\ (C_{in})$, has to be considered a parameter, too.

Now, one additionally assumes that the variables S_{ik}, $1 \le i,\ k \le n$, arc independent. Then the parameters x_i, y_k of model (4) can be estimated (e.g. by the method of maximum likelihood) if we assume any distribution function for S_{ik}; e.g.. a one-parametric one with expected value x_iy_k or a twoparametric one with the second parameter being constant over all cells (i,k). For example, we can take one of the following possibilities:

(4a) $$S_{ik} \propto \text{Normal } (x_iy_k, \sigma^2)$$

(4b) $$S_{ik} \propto \text{Exponential } (1/(x_iy_k))$$

(4c) $$S_{ik} \propto \text{Lognormal } (x_i + y_k, \sigma^2)$$

(Observe that (4a) and (4c) introduce even a further parameter σ^2). Possibility (4a) has been introduced into the literature by de Vylder 1978 using least squares estimation of the parameters. The fact that claims variables are usually skewed to the right is taken into account by possibilities (4b) and (4c) but at the price that all incremental variables S_{ik} must be positive (which is not the case with the original chain ladder method and often restricts the use of (4b) and (4c) to triangles of paid amounts).

Possibility (4b) has been used by Mack 1991. Possibility (4c) was introduced by Kremer 1982 and extended by Zehnwirth 1989 and 1991. Renshaw 1989, Christofides 1990. Verrall 1990 and 1991. It has the advantage that it leads to a linear model for $\log(S_{ik})$, namely to a two-way analysis of variance, and that the parameters can therefore be estimated using ordinary regression analysis.

Although model (4c) seems to be the most popular possibility of model class (4). we want to emphasize that it is only one of many different ways of stochastifying model (4). Moreover, possibilities (4a), (4b), (4c), yield different estimators for the parameters x_i, y_k, and for the claims reserves and all of these arc different from the result of the original chain ladder method. Therefore this author finds it to be misleading that in the papers by Zehnwirth 1989 and 1991, Renshaw 1989. Christofides 1990. Verrall 1990 and 1991 model (4c) explicitly or implicitly is called "*the* scholastic model underlying the chain ladder" or even directly "chain ladder model." In fact, it is something different. In order to not efface this difference, model (4c) should better lx called "loglinear cross-classified claims reserving method." In the next chapter we show that this difference does not only rely on a different parametric assumption or on different estimators but stems from a different underlying philosophy.

3. A distribution-free stochastic model for the original chain ladder method

The stochastic models (4a). (4b), (4c) described in the last chapter did not lead us to a model which yields the same reserve formula as the original chain ladder method. But we will now develop such a model.

If we compare model (3) with the chain ladder projection (1), we may get the impression that the transition

(A) $$\hat{C}_{i,n+2-i} = C_{i,n+1-i}\hat{f}_{n+1-i}$$

in (1) from the most recent observed amount $C_{i,n+1-i}$ to the estimator for the first unknown amount $C_{i,n+2-i}$ has not been captured very well by model (3) which uses

(B) $$\hat{C}_{i,n+2-i} = \mathrm{E}(C_{i,n+1-i})\, f_{n+1-i}.$$

The crucial difference between (A) and (B) is the fact that (A) uses the actual observation $C_{i,n+1-i}$ itself as basis for the projection whereas (B) takes its expected value. This means that the chain ladder method implicitly must use an assumption which states that the information contained in the most recent observation $C_{i,n+1-i}$ is more relevant than that of the average $\mathrm{E}(C_{i,n+1-i})$ This is duly taken into account by the model

(5) $$\mathrm{E}(C_{i,k+1} \mid C_{i1}, \ldots, C_{ik}) = C_{ik}\, f_k, \quad 1 \le i \le n, \quad 1 \le k \le n-1$$

which is (due to the iterative rule for expectations) more restrictive than (3). Moreover, using (5) we ate able to calculate the conditional expectation $E(C_{ik}|D)$, $i + k > n + 1$, given the data

$$D = \{C_{ik} \mid i + k \leq n + 1\}$$

observed so far, and knowing this conditional expectation is more useful than knowing the unconditional expectation $E(C_{ik})$ which ignores the observation D. Finally, the following theorem shows that using (5) we additionally need only to assume the independence of the accident years, i.e. to assume that

$$\{C_{i1}, \ldots, C_{in}\}, \{C_{j1}, \ldots, C_{jn}\}, i \neq j, \tag{6}$$

are independent, whereas under (4a), (4b), (4c) we had to assume the independence of both. the accident years and the development year increments.

Theorem: Under assumptions (5) and (6) we have fork $> n + 1 - i$

$$E(C_{ik} \mid D) = C_{i,n+1-i} \, f_{n+1-i} \times \ldots \times f_{k-1}. \tag{7}$$

Proof: Using the abbreviation

$$E_i(X) = E(X \mid C_{i1}, \ldots, C_{i,n+1-i})$$

we have due to (6) and by repeated application of (5)

$$\begin{aligned} E(C_{ik}|D) &= E_i(C_{ik}) \\ &= E_i(E(C_{ik}|C_{i1}, \ldots, C_{i,k-1})) \\ &= E_i(C_{i,k-1}) \, f_{k-1} \\ &= \text{etc.} \\ &= E_i(C_{i,n+2-i}) \, f_{n+2-i} \times \ldots \times f_{k-1} \\ &= C_{i,n+1-i} \, f_{n+1-i} \times \ldots \times f_{k-1}. \end{aligned}$$

The theorem shows that the stochastic model (5) produces exactly the same reserves as the original chain ladder method if we estimate the model parameters f_k by (2). Moreover. wc see that the projection basis $C_{i,n+1-i}$ in formulae (7) and (1) is not an estimator of the parameter $E(C_{i,n+1-i})$ but stems from working on condition of the data observed so far. Altogcthcr. model (5) employs only n-l parameters $f_1, \ldots, f_{n-1}$. The price for having less parameters than models (3) or (4) is the fact that in model (5) we do not have a good estimator for $E(C_{in})$ which are the additional parameters of models (3) and (4).

But even models (4) do not use $E(C_{in})$ as estimator for the ultimate claims amount because this would not be meaningful in view of the fact that the knowledge of $E(C_{1n})$ is completely useless (because we already know C_{1n} exactly) and that one might have $E(C_{in}) < C_{i,\,n+1-i}$ (e.g. for $i = 2$) which would lead to a negative claims reserve even if that is not possible. Instead models (4) estimate the ultimate claims amount by estimating

$$C_{i,\,n+1-i} + E(S_{i,\,n+2-i} + \ldots + S_{in}),$$

i.e. they estimate the claims reserve $R_i = C_{in} - C_{i,\,n+1-i} = S_{i,\,n+2-i} + + S_{in}$ by estimating

$$E(R_i) = E(S_{i,\,n+2-i} + \ldots + S_{in}).$$

If we assume that we know the true parameters x_i, y_k of model (4) and f_k of model (5). we can clarify the essential difference between both models in the following way: The claims reserve for model (4) would then be

$$E(R_i) = x_i(y_{n+2-i} + \ldots + y_n)$$

independently of the observed data D, i.e. it will not change if we simulate different data sets D from the underlying distribution. On the other hand, due to the above theorem, model (5) will each time yield a different claims reserve

$$E(R_i \mid D) = C_{i,\,n+1-i}\ (f_{n+1-i} \times \ldots \times f_{n-1} - 1)$$

as $C_{i,\,n+1-i}$ changes from one simulation to the next.

For the practice, this means that we should use the chain ladder method (1) or (5) if we believe that the deviation

$$C_{i,\,n+1-i} - E(C_{i,\,n+1-i})$$

is indicative for the future development of the claims. If not, we can think on applying a model (4) although doubling the number of parameters is a high price and may lead to high instabiiity of the estimated reserves as is shown in the appendix.

4. Final Remark

The aim of this note was to show that the loglinear cross-classiticd model (4c) used by Renshaw. Christotidcs. Verrall and Zehnwirth is *not* a model underlying the usual chain ladder method because it requires independent and strictly positive increments and produces different reserves. We have also shown that model (5) *is* a stochastic model underlying the chain ladder method. Moreover, model (5) has only $n - 1$ parameters-as opposed to 2, - 1 (or even $2n$) in case of model (4c)—and is therefore more robust than model (4c).

Finally. one might argue that one advantage of the loglinear model (4c) is the factthatit allows to calculate the standard errors of the reserve estimators as has been done by Renshaw 1989. Christofides 1990 and Verrall 1991. But this is possible for model (5). too, as is shown in a separate paper (Mack 1993).

Acknowledgement

I first saw the decisive idea to base the stochastic model for the chain ladder method on conditional expectations in Schnieper 1991.

APPENDIX

NUMERAL EXAMPLE WHICH SHOWS THAT THE LOGLINEAR MODEL (4C) IS MORE VOLATILE THAN THE USUAL CHAIN LADDER METHOD

The data for the following example are taken from the "Historical Loss Development Study," 1991 Edition, published by the Reinsurance Association of America (RAA). There, we fmd on page 96 the following run-off triangle of Automatic Facultative business in General Liability (excluding Asbestos & Environmental):

	C_{i1}	C_{i2}	C_{i3}	C_{i4}	C_{i5}	C_{i6}	C_{i7}	C_{i8}	C_{i9}	GO
$i=1$	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
$i=2$	106	4285	5396	10666	13782	15599	15496	16169	16704	
$i=3$	3410	8992	I3873	16141	18735	22214	22863	23466		
$i=4$	5655	11555	15766	21266	23425	26083	27067			
$i=5$	1092	9565	15836	22169	25955	26180				
$i=6$	1513	6445	11702	12935	15852					
$i=7$	557	4020	10946	12314						
$i=8$	1351	6947	13112							
$i=9$	3133	5395								
$i=10$	2063									

The above figures are cumulative incurrcd case losses in $ 1000. Wc have taken the accident years from 1981 ($i=1$) to 1990 ($i=10$). The following table shows the corrcsponding incremental amounts $S_{ik}=C_{ik}-C_{i,k-1}$:

	S_{i1}	S_{i2}	S_{i3}	S_{i4}	S_{i5}	S_{i6}	S_{i7}	S_{i8}	S_{i9}	S_{i10}
$i=1$	5012	3257	2638	898	1734	2642	1828	599	54	172
$i=2$	106	4179	1111	5270	3116	1817	-103	673	535	
$i=3$	3410	5582	488 I	2268	2594	3479	649	603		
$i=4$	5655	5900	4211	5500	2159	2658	984			
$i=5$	1092	8473	6271	6333	3786	225				
$i=6$	1513	4932	5257	1233	2917					
$i=7$	557	3463	6926	1368						
$i=8$	1351	5596	6165							
$i=9$	3133	2262								
$i=10$	2063									

Note that in development year 7 of accident year 2 we have a negative increment $S_{2,7}=C_{2,7}-C_{2,6}=-103$. Because model (4c) works with logarithms of the incremental amounts S_{ik} it cannot handle the negative increments $S_{2,7}$. In order to apply model (4c), we therefore must change $S_{2,7}$ artificially or leave it out. We have tried the following possibilities:

(a) $S_{2,7} = 1$, i.e. $C_{2,7} = 15496 + 104 = 15600$, $C_{2,8} = 16169 + 104$

$= 16273$, $C_{2,9} = 16704 + 104 = 16808$

(b_1) $C_{2,7} = 16000$, i.e. $S_{2,7} = 401$, $S_{2,8} = 169$

(b_2) $S_{2,7}$ = missing value, i.e. $C_{2,7}$ = missing value

When estimating the reserves for these possibilities and looking at the residuals for model (4c), we will identify $S_{2,1} = C_{2,1} = 106$ as an outlier. We have therefore also tried:

C_1 like (b_1) but additionally $S_{2,1} = C_{2,1} = 1500$, i.e. all $C_{2,k}$ are augmented by $1500 - 106 = 1394$

C_2 like (b_2) but additionally $S_{2,1} = C_{2,1}$ = missing value.

This yields the following results (the calculations for model (4c) were done using Ben Zehnwirth's ICRFS, version 6.1):

	Total Estimated Reserves	
Possibility	Chain Ladder	Loglinear Model (4C)
unchanged data	52,135	not possible
(a)	52,274	190,754
(b_1)	51,523	102,065
(b_2)	52,963	107,354
(c_1)	49,720	69,999
(c_2)	51,834	70,032

This comparison clearly shows that the two methods are completely different and that the usual chain ladder method is much less volatile than the loglinear cross-classified method (4c).

For the sake of completeness, the following two tables give the results for the above calculations per accident year:

CHAIN LADDER METHOD—ESTIMATED RESERVES PER ACCIDENT YEAR

Acc. Year	Unchanged	(a)	(b_1)	(b_2)	(c_1)	(c_2)
1981	0	0	0	0	0	0
1982	154	155	154	154	167	154
1983	617	616	617	617	602	617
1984	1,636	1,633	1.382	1,529	I ,348	1529
1985	2,747	2.780	2,664	2.964	2.606	2.964
1986	3.649	3.671	3593	3.795	3,526	3.795
1987	5.435	5,455	5.384	5568	5,286	5,568
1988	10.907	10,935	10.838	11,087	10.622	11,087
1989	10.650	10.668	10.604	10,770	10,322	10.770
1990	16.339	16360	16287	16,477	IS,242	15349
1981-90	52.135	52374	51523	52.963	49.720	51,834

LOGLINEAR METHOD—ESTIMATED RESERVES PER ACCIDENT YEAR

Acc. Year	(a)	(b_1)	(b_2)	(c_1)	(c_2)
1981	0	0	0	0	0
1982	309	249	313	282	387
1983	2.088	949	893	749	674
1984	6.114	2,139	2.683	I.675	1.993
1985	3.773	2,649	3.286	2,086	2.602
1986	6.917	4.658	5,263	3,684	4,097
1987	9.648	6,312	6.780	4.968	5.188
1988	24.790	15,648	16.468	12,000	12.174
1989	36.374	21.429	22.213	15,545	15.343
1990	100,739	48,033	49,454	29,010	27,575
1981-90	190.754	102.065	107.354	69,999	70.032

REFERENCES

Christofides, S. (1990). Regression Models Based on Log-incremental Payments. In: Claims Reserving Manual, Vol. 2. Institute of Actuaries, London.

De Vylder. F. (1978). Estimation of IBNR Claims by Least Squares. Mitteilungen der Vereinigung Schweizerischer Versicherungs-mathematiker, 249-254.

Kremer, E. (1982). IBNR-Claims and the Two-Way Model of ANOVA. Scandinavian Actuarial Journal. 47-55.

Mack, Th. (1991). Claims Reserving: The Direct Method and its Refinement by a Lag-Distribution. ASTIN Colloquium Stockholm 1991.

Mack, Th. (1993). Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. ASTIN Bulletin, to appear.

Renshaw, A. (1989). Chain Ladder and Interactive Modelling. Journal of the Institute of Actuaries 116. 559-587.

Schnieper. R. (1991). Separating True IBNR and IBNER Claims. ASTIN Bulletin 21, 111-127. This paper was presented to the ASTIN Colloquium New York 1989 under the title "A Pragmatic IBNR Method."

Verrall. R.J. (1990). Bayes and Empirical Bayes Estimation for the Chain Ladder Model. ASTIN Bulletin 20.217-243.

Verrall, R.J. (1991). On the Estimation of Reserves from Loglinear Models. Insurance: Mathematics and Economics 10, 75-80.

Zehnwirth, B. (1989). The Chain Ladder Technique-A Stochastic Model. In: Claims Reserving Manual, Vol.2. Institute of Actuaries, London.

Zehnwirth, B. (1991). Interactive Claims Reserving Forecasting System, Version 6.1. Insureware P/L, E. St. Kilda Vic 3 183, Australia.

Address of the author:
Dr. Thomas Mack
Munich Reinsurance Company
Königinstr. 107
D-80791 München