

## A Chain-Ladder Analysis of P& I Claims<sup>1</sup>

Nicholas G. Berketis<sup>2</sup> PhD

---

According to the latest I.G. P&I Annual Review for 2017/18:

"The number and severity of pool claims currently reported for the 2017/18 policy year is similar to that for the 2016/17 policy year with 19 claims notified, five of which are precautionary notifications within the club retention. The severity of claims notified, however, is significantly up, due in the main to the costs associated with the "Kea Trader" grounding in New Caledonia in July 2017, and the "Sanchi/CF Crystal" collision in the East China Sea in January 2018".

In a recent article, an Italian Insurance Broker stated that, despite the trend from most IG Clubs, the expectation is that P&I Clubs are heading towards increases.

The insurance industry, unlike other industries, does not sell products as such but promises. An insurance policy is a promise by the insurer to the policyholder to pay for future claims for an upfront received premium.

As a result, Insurers don't know the upfront cost for their service, but rely on historical data analysis and judgement to predict a sustainable price for their offering. In General Insurance (or Non-Life Insurance, e.g. motor, property and casualty insurance) most Policies run for a period of 12 months. However, the claims payment process can take years or even decades. Therefore, often not even the delivery date of their product is known to Insurers.

In particular, losses arising from casualty insurance can take a long time to settle and even when the claims are acknowledged it may take time to establish the extent of the claims' settlement cost. Claims can take years to materialize. A complex and costly example are the claims from asbestos liabilities, particularly those in connection with mesothelioma and lung damage arising from prolonged exposure to asbestos. A research report by a working party of the Institute and Faculty of Actuaries estimated that the un-discounted cost of UK mesothelioma-related claims to the UK Insurance Market for the period 2009 to 2050 could be around £10bn. The cost for asbestos related claims in the US for the worldwide insurance industry was estimated to be around \$120bn in 2002.

Thus, it should come as no surprise that the biggest item on the liabilities side of an Insurer's balance sheet is often the provision or reserves for future claims payments. Those reserves can be broken down in case reserves (or outstanding claims), which are losses already reported to the insurance company and losses that are incurred but not reported (IBNR) yet.

The analysis is based on R (Version 3.5.3 – 11<sup>th</sup> March 2019), an integrated language and environment for statistical computing and graphics. R provides a wide variety of statistical and graphical techniques.

---

**Keywords:** claim, CL method, run-off triangle, claim reserving, P. & I.

---

<sup>1</sup> A summary of this paper was presented on 08/05/2019 at the 2<sup>nd</sup> Marine and Cargo Insurance Conference, Antwerp Expo, Antwerp, Belgium also on 07/06/2019 at the 2<sup>nd</sup> Pancyprian Statistics Day, Frederick University, Nicosia, Cyprus.

<sup>2</sup> Manager of J.Kouroutis & Co. Ltd. Insurance and Reinsurance Brokers, Piraeus, Greece.

The estimated cost of notified pool claims (in USD 000,000) is as follows:

<u>YEAR</u>	<u>No.OF CLAIMS</u>	<u>12M</u>	<u>24M</u>	<u>36M</u>	<u>48M</u>	<u>60M</u>	<u>72M</u>	<u>84M</u>	<u>96M</u>	<u>108M</u>	<u>120M</u>
2007/08	27	3032	4369	4798	4859	5154	5300	5300	5201	5252	5295
2008/09	14	876	1204	1103	1260	1240	1236	1269	1290	1285	NA
2009/10	22	2338	2294	2313	2270	2544	2759	2679	2646	NA	NA
2010/11	22	1929	2635	2916	2761	2717	2801	2803	NA	NA	NA
2011/12	14	3310	4692	4845	5121	5118	5175	NA	NA	NA	NA
2012/13	22	3753	4602	4858	4838	4641	NA	NA	NA	NA	NA
2013/14	19	2871	3403	3859	3827	NA	NA	NA	NA	NA	NA
2014/15	16	1771	1905	2013	NA	NA	NA	NA	NA	NA	NA
2015/16	15	2516	3537	NA	NA	NA	NA	NA	NA	NA	NA
2016/17	7	756	NA	NA	NA	NA	NA	NA	NA	NA	NA

This triangle shows the known values of loss from each origin year and of annual evaluations thereafter. For example, the known values of loss originating from the 2013/14 exposure period are 2871, 3403, and 3859 as of year ends 2013, 2012 and 2011, respectively. The latest diagonal – i.e., the vector 5295, 1285, . . . 756 from the upper right to the lower left – shows the most recent evaluation available.

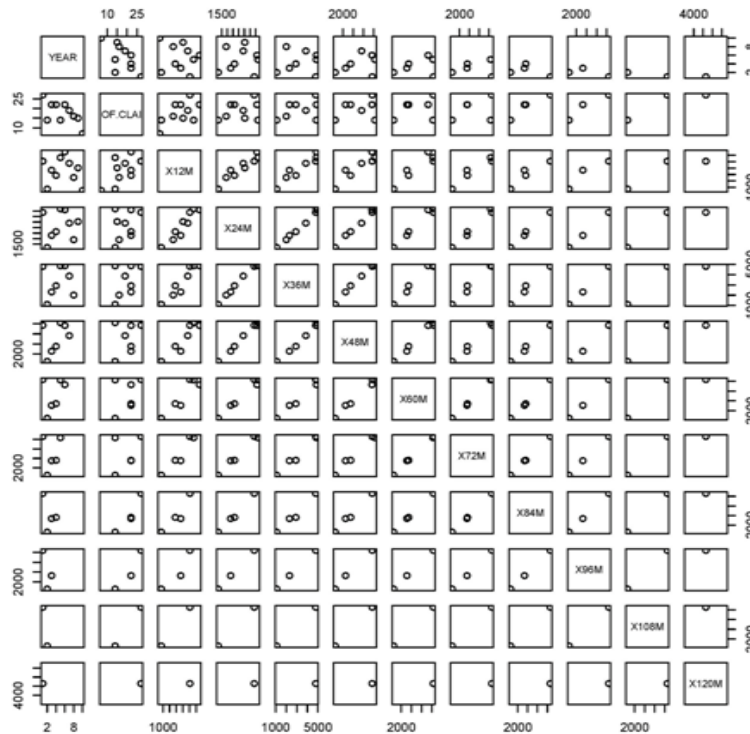
The column headings – 1, 2, . . . , 10 – hold the ages (in years) of the observations in the column relative to the beginning of the exposure period. For example, for the 2014/15 origin year, the age of the 2013 value, evaluated as of 20/02/2017, is three years.

The objective of a reserving exercise is to forecast the future claims development in the bottom right corner of the triangle and potential further developments beyond development age 10. Eventually, all claims for a given origin period will be settled, but it is not always obvious to judge how many years or even decades it will take.

We speak of long and short tail business depending on the time it takes to pay all claims.

In order proceed with our analysis, we first plotted the data to get an overview. Figure 1 that follows shows the claims development chart for the past 10 years.

Figure 1



The classical chain-ladder is a deterministic algorithm to forecast claims based on historical data. It assumes that the proportional developments of claims from one development period to the next are the same for all origin years.

Basic idea

Most commonly as a first step, the age-to-age link ratios are calculated as the volume weighted average development ratios of a cumulative loss development triangle from one development period to the next  $C_{i,k}$ ,  $i, k = 1, \dots, n$ .

$$f_k = \frac{\sum_{i=1}^{n-k} C_{i,k+1}}{\sum_{i=1}^{n-k} C_{i,k}}$$

[1] 1.2788444 1.0637747 1.0098817 1.0144488 1.0296906 0.9962798 0.9879974 1.0070867 1.0081874

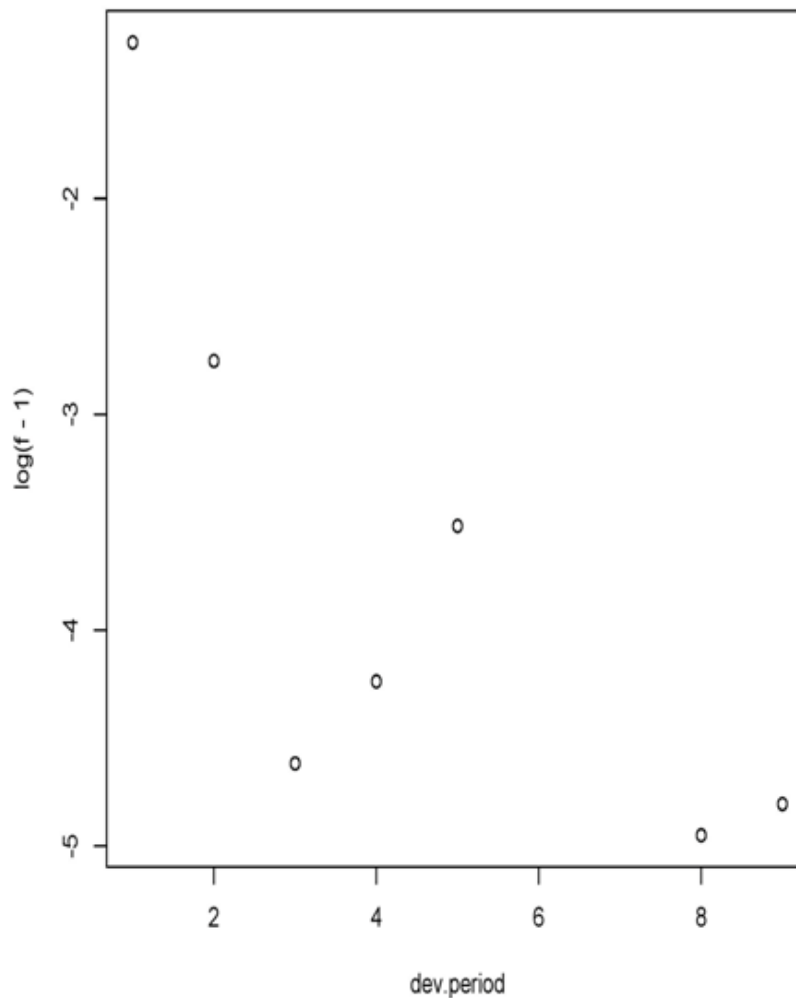
Often it is not suitable to assume that the oldest origin year is fully developed. A typical approach is to extrapolate the development ratios, e.g. assuming a log-linear model.

[1] 1.012789

Figure 2 below shows the Log-linear extrapolation of age-to-age factors.

**Figure 2**

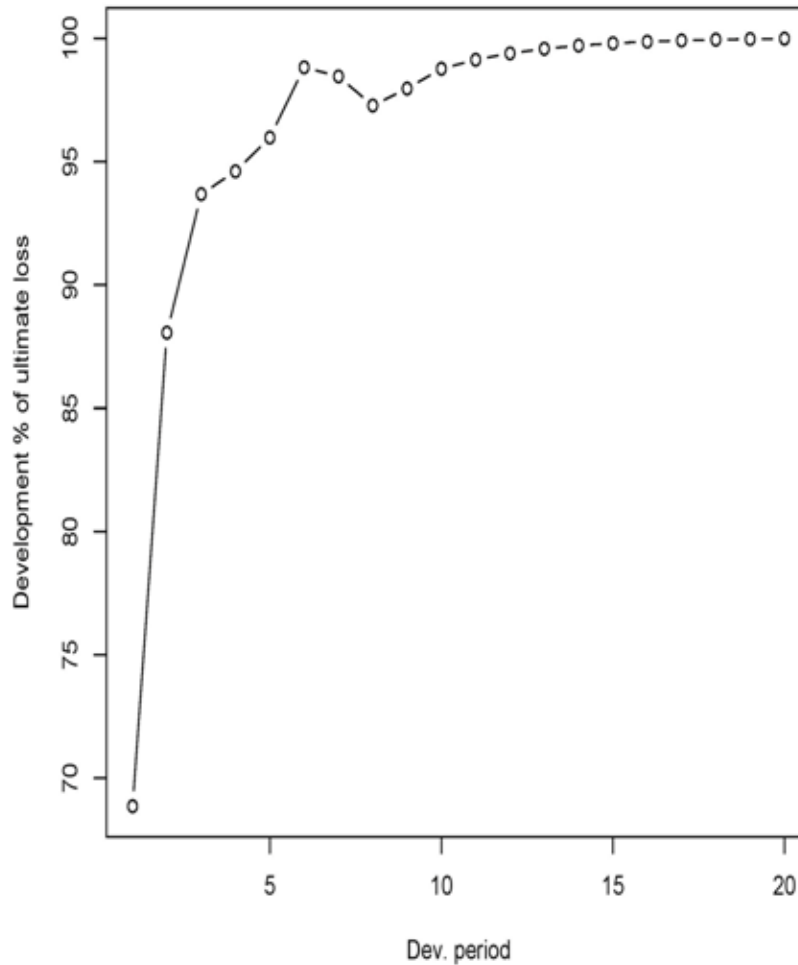
**Log-linear extrapolation of age-to-age factor**



The age-to-age factors allow us to plot the expected claims development patterns.

This is shown on Figure 3:

**Figure 3**  
**Expected claims development pattern**



The link ratios are then applied to the latest known cumulative claims amount to forecast the next development period. The squaring of the triangle is calculated below, where an ultimate column is appended to the right to accommodate the expected development beyond the oldest age (10) of the triangle due to the tail factor (1.012789) being greater than unity.

	X12M	X24M	X36M	X48M	X60M	X72M	X84M	X96M	X108M	X120M	Ult
1	3032	4369	4798	4859	5154	5300	5300	5201	5252	5295	5363
2	876	1204	1103	1260	1240	1236	1269	1290	1285	1296	1312
3	2338	2294	2313	2270	2544	2759	2679	2646	2665	2687	2721
4	1929	2635	2916	2761	2717	2801	2803	2769	2789	2812	2848
5	3310	4692	4845	5121	5118	5175	5156	5094	5130	5172	5238
6	3753	4602	4858	4838	4641	4779	4761	4704	4737	4776	4837
7	2871	3403	3859	3827	3882	3998	3983	3935	3963	3995	4046
8	1771	1905	2013	2033	2062	2123	2116	2090	2105	2122	2149
9	2516	3537	3763	3800	3855	3969	3954	3907	3935	3967	4018
10	756	967	1028	1039	1054	1085	1081	1068	1075	1084	1098

The total estimated outstanding loss under this method is about 33600. In particular, it was calculated 33630.05. This approach is also called Loss Development Factor (LDF) method. More generally, the factors used to square the triangle need not always be drawn from the dollar weighted averages of the triangle. Other sources of factors from which the actuary may select link ratios include simple averages from the triangle, averages weighted toward more recent observations or adjusted for outliers, and benchmark patterns based on related, more credible loss

experience. Also, since the ultimate value of claims is simply the product of the most current diagonal and the cumulative product of the link ratios, the completion of interior of the triangle is usually not displayed in favour of that multiplicative calculation.

Mack chain-ladder

Thomas Mack published in 1993<sup>3</sup> a method which estimates the standard errors of the chain-ladder forecast without assuming a distribution under three conditions.

Following the notation of Mack<sup>4</sup> let  $C_{ik}$  denote the cumulative loss amountsof origin period (e.g. accident year)  $i = 1, \dots, m$ , with losses known for development period (e.g. development year)  $k \leq n + 1 - i$ .

In order to forecast the amounts  $C_{ik}$  for  $k > n+1-i$  the Mack chain-ladder-model assumes:

$$CL1: E[F_{ik}|C_{i1}, C_{i2}, \dots, C_{ik}] = f_k \text{ with } F_{ik} = \frac{C_{i,k+1}}{C_{ik}} \quad (2)$$

$$CL2: Var\left(\frac{C_{i,k+1}}{C_{ik}} | C_{i1}, C_{i2}, \dots, C_{ik}\right) = \frac{\sigma_k^2}{w_{ik} C_{ik}^a} \quad (3)$$

$$CL3: \{C_{i1}, \dots, C_{in}\}, \{C_{j1}, \dots, C_{jn}\}, \text{ are independent for origin period } i \neq j \quad (4)$$

with  $w_{ik} \in [0; 1]$ ,  $a \in \{0, 1, 2\}$ . If these assumptions hold, the Mack chain-ladder model gives an unbiased estimator for IBNR (Incurred But Not Reported) claims.

The Mack chain-ladder model can be regarded as a weighted linear regression through the origin for each development period:  $\ln(y \sim x + 0, \text{ weights}=w/x^{(2-\alpha)})$ , where  $y$  is the vector of claims at development period  $k + 1$  and  $x$  is the vector of claims at development period  $k$ .

The Mack method is implemented in the ChainLadder package via the function Mack ChainLadder.

We therefore applied the MackChainLadder function to our triangle:

	Latest	Dev. To. Date	Ultimate	IBNR	Mack. S.E.	CV (IBNR)
1	5,295	1.000	5,295	0.00	0.00	NaN
2	1,285	0.992	1,296	10.52	9.71	0.923
3	2,646	0.985	2,687	40.57	31.13	0.767
4	2,803	0.997	2,812	8.82	58.49	6.634
5	5,175	1.001	5,172	-3.03	123.16	-40.581
6	4,641	0.972	4,776	134.99	171.94	1.274
7	3,827	0.958	3,995	168.22	277.32	1.649
8	2,013	0.949	2,122	109.25	235.71	2.158
9	3,537	0.892	3,967	429.77	392.49	0.913
10	756	0.697	1,084	328.28	318.80	0.971

Totals	
Latest:	31,978.00
Dev:	0.96
Ultimate:	33,205.38
IBNR:	1,227.38
Mack. S.E.:	782.64
CV(IBNR):	0.64

Executing Mack Chain Ladder will print the following columns of information per accident year (origin period):

1. **Latest:** the claim amount for the last development period
2. **Dev.To.Date:** the development to date or the ratio of the latest over the predicted ultimate
3. **Ultimate:** predicted ultimate claim
4. **IBNR:** the predicted IBNR reserve

<sup>3</sup>Mack, Thomas, (1993), "Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates", ASTIN Bulletin, Vol. 23(2): 213–225.

<sup>4</sup>Mack, Thomas, (1999), "The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor", ASTIN Bulletin, Vol. 29(2): 361-366.

5. **Mack.S.E.:** the standard error, or the standard deviation of the bounds for the predicted ultimate and IBNR since the estimate is unbiased (shown in Mack's 1999 paper). In other words, since the S.E given is equal to one standard deviation, a confidence interval for the true ultimate value can be found using the standard error and the predicted ultimate.
6. **CV(IBNR):** coefficient of variation, or the ratio of the standard error over the predicted IBNR

The bottom output gives a total or sum of the latest, ultimates, IBNRs. It also gives the standard error of the total ultimate (this is not the total of the standard errors). The development to date factor is the ratio of the total latest against the total ultimate, and the CV(IBNR) is the percentage of the total standard error in the total IBNR.

If the CV(absolute value) is greater than 25%, then another model or a log linear regression should be used.

We can access the loss development factors and the full triangle via:

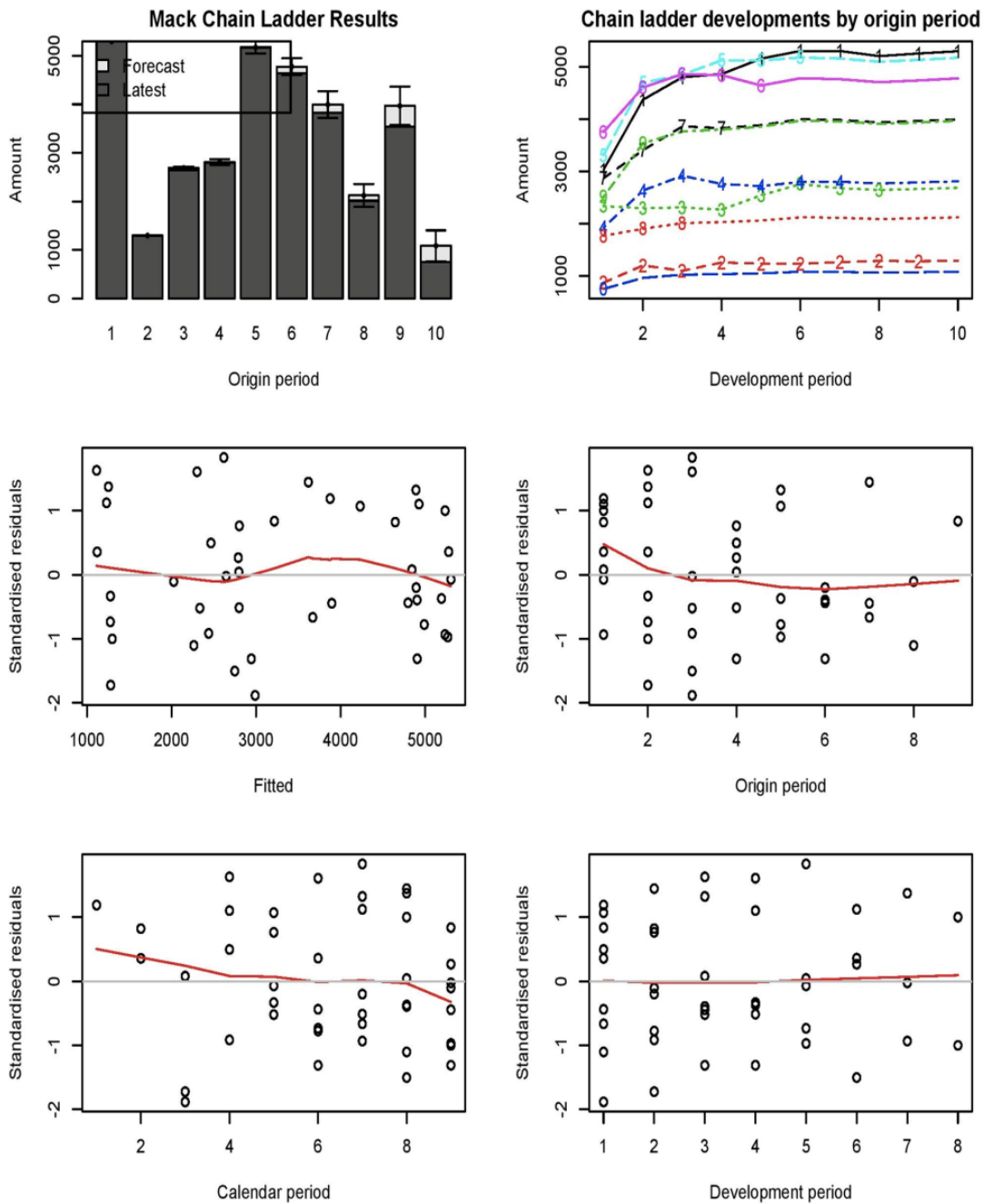
[1] 1.2788444 1.0637747 1.0098817 1.0144488 1.0296906 0.9962798 0.9879974 1.0070867 1.0081874

[10] 1.0000000

origin	X12M	X24M	X36M	X48M	X60M	X72M	X84M	X96M	X108M	X120M
1	3032	4369.0000	4798.000	4859.000	5154.000	5300.000	5300.000	5201.000	5252.000	5295.000
2	876	1204.0000	1103.000	1260.000	1240.000	1236.000	1269.000	1290.000	1285.000	1295.521
3	2338	2294.0000	2313.000	2270.000	2544.000	2759.000	2679.000	2646.000	2664.752	2686.569
4	1929	2635.0000	2916.000	2761.000	2717.000	2801.000	2803.000	2769.357	2788.982	2811.817
5	3310	4692.0000	4845.000	5121.000	5118.000	5175.000	5155.748	5093.865	5129.964	5171.965
6	3753	4602.0000	4858.000	4838.000	4641.000	4778.794	4761.016	4703.871	4737.206	4775.991
7	2871	3403.0000	3859.000	3827.000	3882.296	3997.563	3982.691	3934.889	3962.774	3995.219
8	1771	1905.0000	2013.000	2032.892	2062.265	2123.495	2115.595	2090.202	2105.015	2122.249
9	2516	3537.0000	3762.571	3799.752	3854.655	3969.101	3954.335	3906.872	3934.559	3966.773
10	756	966.8064	1028.464	1038.627	1053.634	1084.917	1080.881	1067.908	1075.476	1084.281

To plot that Mack's assumption are valid review the residual plots, we see no trends in either of them. Please refer to the Figure 4 that follows:

Figure 4



Bootstrap chain-ladder

The BootChainLadder function uses a two-stage bootstrapping / simulation approach following the paper by England and Verrall<sup>5</sup>. In the first stage an ordinary chain-ladder method is applied to the cumulative claims' triangle. From this we calculate the scaled Pearson residuals which we bootstrap R times to forecast future incremental claims payments via the standard chain-ladder method. In the second stage we simulate the process error with the bootstrap value as the mean and using the process distribution assumed. The set of reserves obtained in this way forms the predictive distribution, from which summary statistics such as mean, prediction error or quantiles can be derived.

<sup>5</sup>England, P. D., & Verrall, Richard J., (2002), "Stochastic Claims Reserving in General Insurance", Presented to the Institute of Actuaries, 28 January.

BootChainLadder(Triangle = GGG21, R = 999, process.distr = "gamma")

	Latest Mean	Ultimate Mean	IBNR	IBNR S.E.	IBNR 75%	IBNR 95%
1	5,295	5,295	0.00	0.0	0.0000	0.0
2	1,285	1,293	7.87	87.5	0.0486	59.6
3	2,646	2,699	53.26	222.0	47.2876	445.9
4	2,803	2,812	9.27	256.9	38.2066	387.1
5	5,175	5,177	2.16	426.5	120.6720	639.4
6	4,641	4,793	152.25	430.0	322.6110	952.4
7	3,827	4,010	183.06	426.4	343.0253	945.5
8	2,013	2,127	114.34	316.8	216.1157	648.1
9	3,537	3,989	451.72	593.0	685.7188	1,489.2
10	756	1,075	319.33	461.0	479.4013	1,201.4

Totals	
Latest:	31,978
Mean Ultimate:	33,271
Mean IBNR:	1,293
IBNR S.E.	1,562
Total IBNR 75%:	2,177
Total IBNR 95%:	3,892

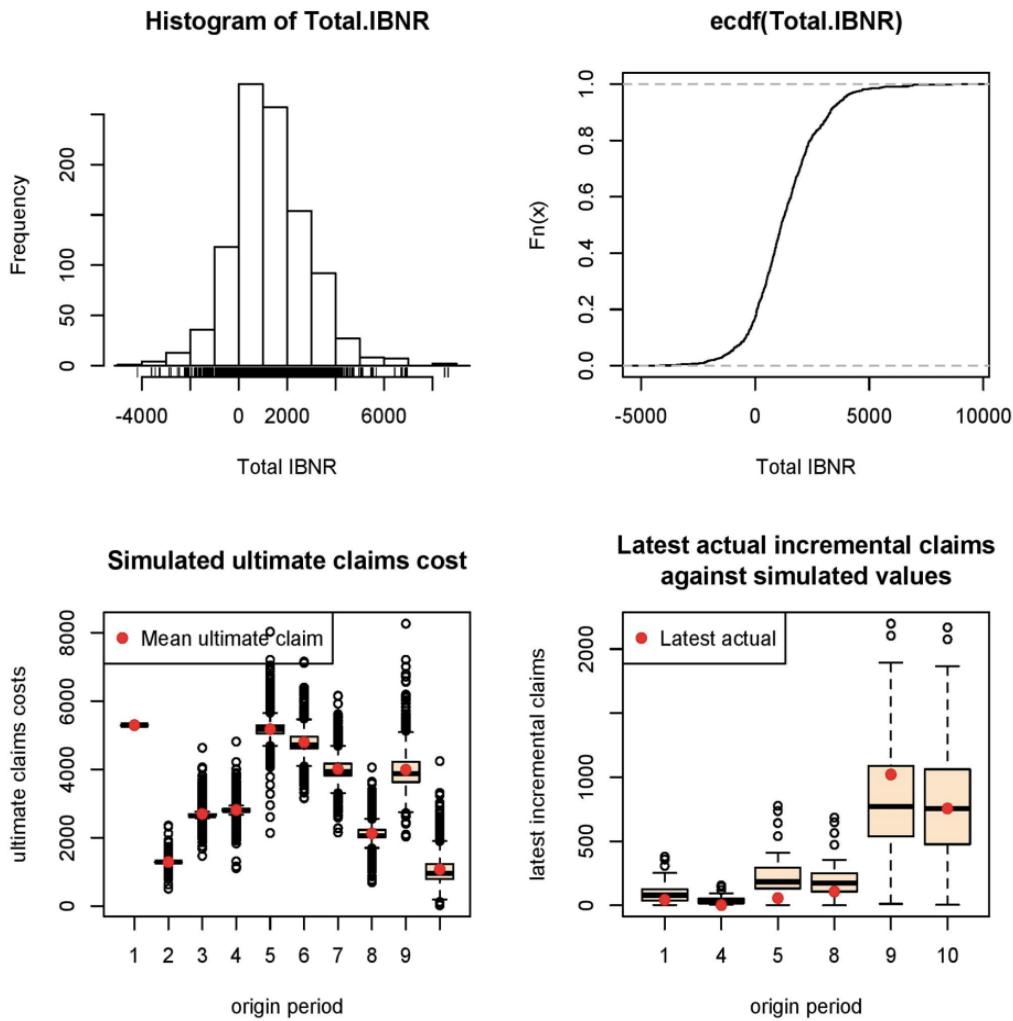
The BootChainLadder is a model that provides a predicted distribution for the IBNR values for a claims' triangle. However, this model predicts IBNR values by a different method than the previous model. First, the development factors are calculated and then they are used in a backwards recursion to predict values for the past loss triangle. Then the predicted values and the actual values are used to calculate Pearson residuals.

Using the adjusted residuals and the predicted losses from before, the model solves for the actual losses in the Pearson formula and forms a new loss triangle. The steps for predicting past losses and residuals are then repeated for this new triangle. After that, the model uses chain ladder ratios to predict the future losses then calculates the ultimate and IBNR values like in the previous Mack model. This cycle is performed R times, depending on the argument values in the model (default is 999 times). The IBNR for each origin period is calculated from each triangle (the default 999) and used to form a predictive distribution, from which summary statistics are obtained such as mean, prediction error, and quantiles.

The output has some of the same values as the Munich and Mack models did. The Mean and SD IBNR is the average and the standard deviation of the predictive distribution of the IBNRs for each origin year. The output also gives the 75% and 95% quantiles of the predictive distribution of IBNRs, in other words 95% or 75% of the predicted IBNRs lie at or below the given values. The above also appear on following Figure 5:



Figure 5



The above Figure 5 shows four graphs, starting with a histogram of the total simulated IBNRs over all origin periods, including a rug plot; a plot of the empirical cumulative distribution of the total IBNRs over all origin periods; a box-whisker plot of simulated ultimate claims costs against origin periods; and a box-whisker plot of simulated incremental claims cost for the latest available calendar period against actual incremental claims of the same period. In the last plot the simulated data should follow the same trend as the actual data, otherwise the original data might have some intrinsic trends which are not reflected in the model.

Quantiles of the bootstrap IBNR can be calculated via the quantile function:

\$By Origin

	<b>IBNR 75%</b>	<b>IBNR 95%</b>	<b>IBNR 99%</b>	<b>IBNR 99.5%</b>
1	0.00000000	0.0000	0.0000	0.0000
2	0.04857565	59.6136	347.7510	451.1056
3	47.28755269	445.8675	961.2352	1150.5122
4	38.20663089	387.1315	857.9082	1080.3987
5	120.67196179	639.3981	1388.2812	1723.2886
6	322.61103376	952.3853	1440.0350	1640.5200
7	343.02532534	945.4920	1526.8586	1715.1438
8	216.11567613	648.1414	1133.8377	1344.0912
9	685.71881492	1489.1964	2565.2205	3046.3018
10	479.40134663	1201.4211	1855.7706	2190.1367
<b>Totals</b>				
IBNR 75%:				2177.064
IBNR 95%:				3891.677
IBNR 99%:				5555.061
IBNR 99.5%:				6873.992

The distribution of the IBNR appears to follow a log-normal distribution, so let's fit it:

meanlog	sdlog
7.08350312	1.05790428
(0.03678694)	(0.02601229)

**Figure 6:**  
ecdf (B\$IBNR.Totals)

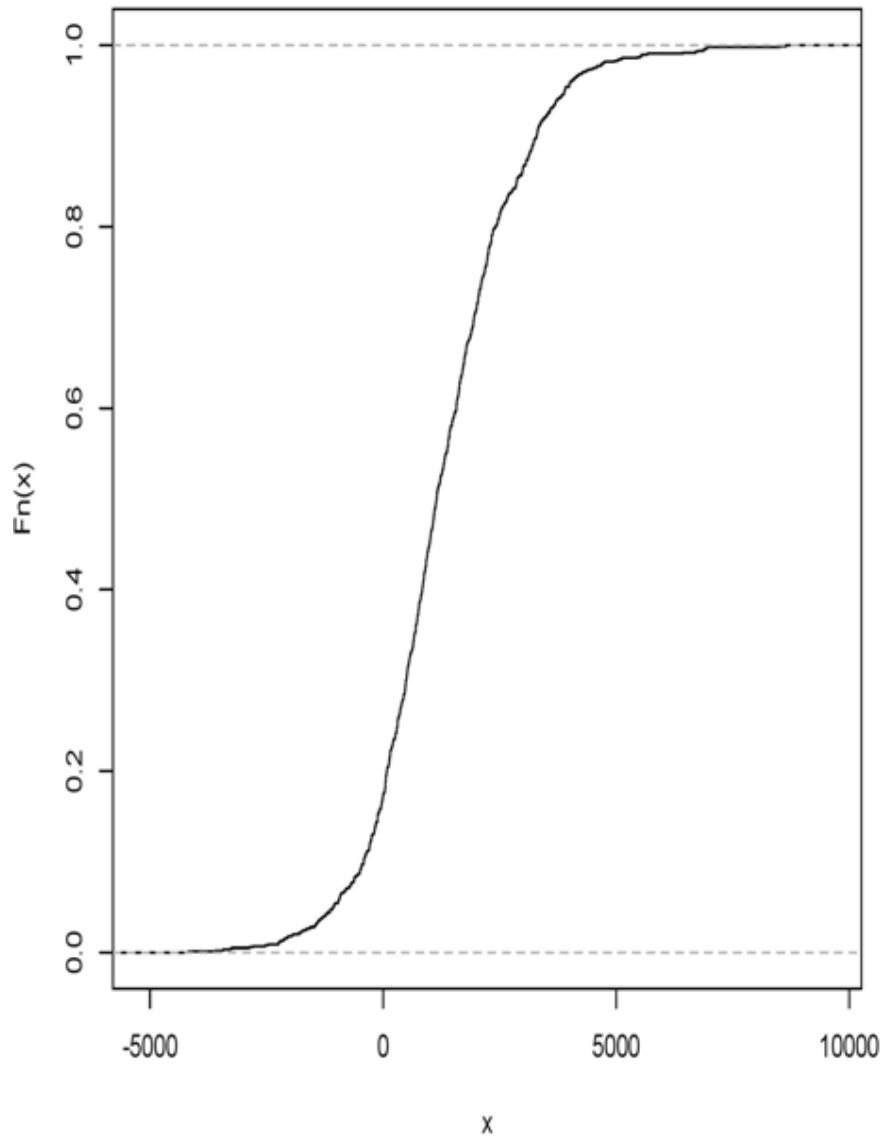
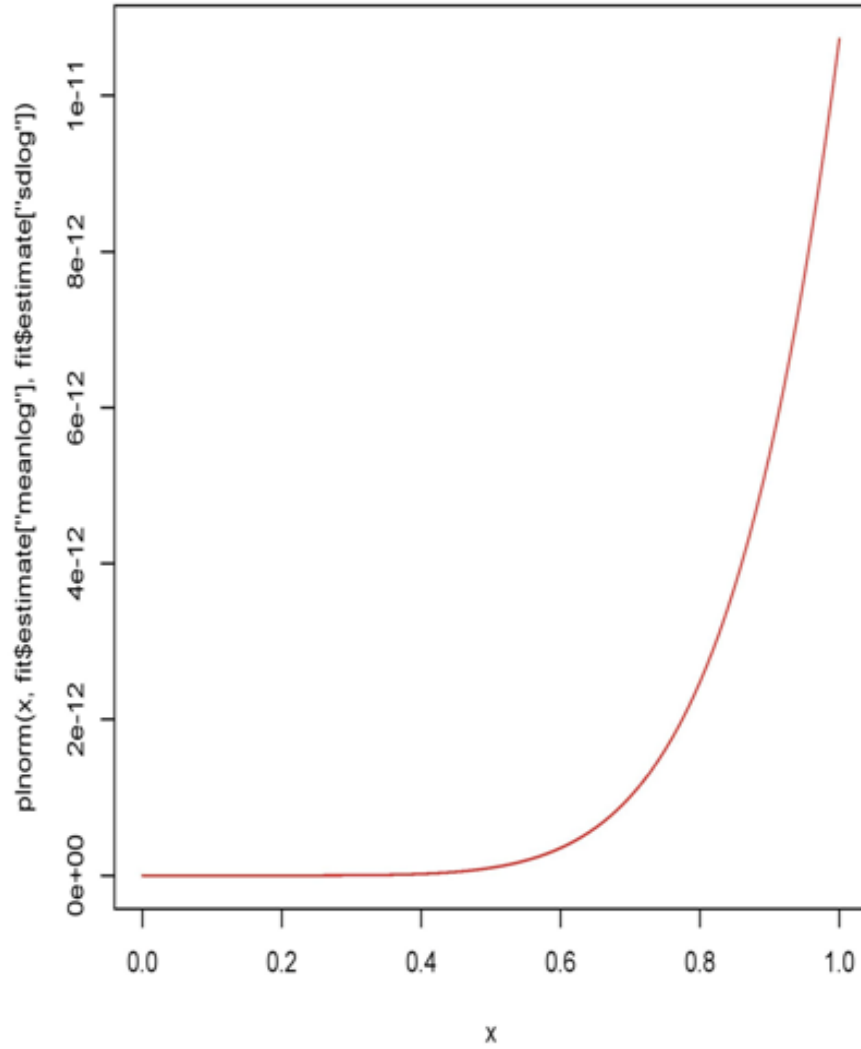


Figure 7:



**CDR function**

The one-year claims development result (CDR) can be estimated via the generic CDR function for objects of MackChainLadder and BootChainLadder.

Further, the tweedie Reserve function offers also the option to estimate the One-year CDR, by setting the argument `reserving=TRUE`.

	<b><u>IBNR</u></b>	<b><u>CDR(1) S.E.</u></b>	<b><u>Mack S.E.</u></b>
<b><u>1</u></b>	0.0	0.0	0.0
<b><u>2</u></b>	10.5	9.7	9.7
<b><u>3</u></b>	40.6	27.4	31.1
<b><u>4</u></b>	8.8	49.8	58.5
<b><u>5</u></b>	-3.0	90.8	123.2
<b><u>6</u></b>	135.0	130.8	171.9
<b><u>7</u></b>	168.2	233.7	277.3
<b><u>8</u></b>	109.2	137.3	235.7
<b><u>9</u></b>	429.8	211.3	392.5
<b><u>10</u></b>	328.3	253.9	318.8
<b><u>TOTAL</u></b>	1,227.4	560.7	782.6

To review the full claims development picture:

	<u>IBNR</u>	<u>CDR(1)</u>	<u>CDR(2)</u>	<u>CDR(3)</u>	<u>CDR(4)</u>	<u>CDR(5)</u>	<u>CDR(6)</u>	<u>CDR(7)</u>	<u>CDR(8)</u>	<u>CDR(9)</u>	<u>CDR(10)</u>	<u>Mack</u>
		<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>	<u>S.E.</u>
<b>1</b>	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0.0
<b>2</b>	10.5	9.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	9.7
<b>3</b>	40.6	27.4	14.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	31.1
<b>4</b>	8.8	49.8	27.0	14.6	0.0	0.0	0.0	0.0	0.0	0.0	0	58.5
<b>5</b>	-3.0	90.8	70.9	38.3	20.8	0.0	0.0	0.0	0.0	0.0	0	123.2
<b>6</b>	135.0	130.8	82.2	64.2	34.8	18.9	0.0	0.0	0.0	0.0	0	171.9
<b>7</b>	168.2	233.7	113.6	71.0	56.1	30.4	16.6	0.0	0.0	0.0	0	277.3
<b>8</b>	109.2	137.3	161.7	78.1	48.7	38.8	21.1	11.6	0.0	0.0	0	235.7
<b>9</b>	429.8	211.3	192.6	227.0	109.5	68.4	54.5	29.7	16.3	0.0	0	392.5
<b>10</b>	328.3	253.9	102.3	94.6	112.6	54.2	33.9	27.1	14.7	8.1	0	318.8
<b>TOTAL</b>	1,227.4	560.7	380.4	309.1	194.5	110.6	73.3	43.1	22.4	8.1	0	782.6

## Conclusions

1. The Loss Development Factor (LDF) is above unity, i.e. 1.012789, which shows an increasingly positive trend for I.B.N.R.'s;
2. The claim amount for the last development period is estimated by both Mack and Bootstrap chain ladder methods at 31,978;
3. The predicted ultimate claim is estimated 33,630.05 under chain ladder method, Mack chain ladder estimated it at 33,205.38, while Bootstrap chain ladder method showed 33,271;
4. The predicted I.B.N.R. reserve was estimated at 1,227.38 under the Mack chain ladder method and 1,293 under Bootstrap chain ladder method;
5. Since the coefficient of variation of I.B.N.R.'s was estimated in absolute value above 25%, i.e. 64%, we followed the Bootstrap chain ladder method, which also justified the increasingly positive trend of I.B.N.R.'s.
6. Hence, the results match the recent Tradewinds article.

## Sources:

1. Carrato, Alessandro, Concina, Fabio, Gesmann, Markus, Murphy, Dan, Wuthrich, Mario & Zhang, Wayne, (2018), "Claims Reserving with R: ChainLadder-0.2.9 Package Vignette", 6, December.
2. Corbett, Adam, (2018), "Insurance Broker Highlights Mounting Pressure on P&I Clubs", Tradewinds, 20, December.
3. Corbett, Adam, (2019), "International Group's Pooled Claims Could Hit 10-year High", Tradewinds, 2, May.
4. England, P. D., & Verrall, Richard J., (2002), "Stochastic Claims Reserving in General Insurance", Presented to the Institute of Actuaries, 28 January.
5. FP Marine Risks, (2018), "P&I Report 2017-2018".
6. International Group of P&I Clubs, (2018), "Annual Review 2017/18", October.
7. Mack, Thomas, (1993), "Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates", ASTIN Bulletin, Vol. 23(2): 213– 225.
8. Mack, Thomas, (1999), "The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor", ASTIN Bulletin, Vol. 29(2): 361-366.
9. Merz, Michael and Mario V. Wüthrich, (2014), "Claims Run-off Uncertainty: The Full Picture", SSRN Manuscript, 2524352.
10. <http://opensourcesoftware.casact.org/chain-ladder>
11. Weindorfer, Bjorn, (2012), "A Practical Guide to the Use of the Chain-Ladder Method for Determining Technical Provisions for Outstanding Reporting Claims in Non-Life Insurance", Working papers series by University of Applied science BFI Vienna, October, No. 77/2012.