Chain Ladder Reserve Risk Estimators

Daniel M. Murphy, FCAS, MAAA

Abstract

Mack (1993) [2] and Murphy (1994) [4] derived analytic formulas for the reserve risk of the chain ladder method. In 1999, Mack [3] gave a recursive version of his formula for total risk. This paper provides the recursive versions of Mack's formulas for process risk and parameter risk and shows that they agree with the formulas in Murphy [4] except for a parameter risk cross-product term. MSE is decomposed into variance and bias components. For the unbiased all-year weighted average link ratios in Mack [2] and Murphy [4] the MSE decomposition in this paper yields formulas that agree with Murphy [4]. For well-behaved triangles the difference between Mack and Murphy parameter risk estimates should be negligible. The concepts are illustrated with an example using data from Taylor and Ashe [5].

Keywords: chain ladder; reserve risk; Mack; mean square error; parameter risk; bias; benchmarks.

Introduction

Mack [1] derived formulas for the chain ladder reserve risk when the age-to-age factors are based on the all-year weighted average. Murphy [4] derived recursive formulas for the chain ladder reserve risk under assumptions that are equivalent to Mack's. The authors' formulas yield different results, for reasons to be discussed herein.

Mack [3] presented a recursive version of the total risk formula. In Section 1 we show recursive formulas for process risk and parameter risk not shown in [3]. We compare them with Murphy's recursive formulas using Mack's notation and note that the difference between the Mack and Murphy reserve risk estimates lies in the parameter risk component.

Mack's reserve risk is measured by the mean square error (MSE). Murphy's reserve risk is measured by total variance. Although MSE is employed in many authors' actuarial research, a mathematically precise definition, particularly as regards reserve risk, is not readily found in the literature. In Section 2 we present a definition of mean square error using the calculus of probability density functions. We will see that MSE can be decomposed into three terms: process risk, parameter risk, and bias. Since total variance is the sum of process variance and parameter variance, the difference between the Mack and Murphy reserve risk measures is bias. A separate mathematical manipulation, this time of parameter risk, yields a recursive formula that agrees with Murphy's. Most of the mathematics will be relegated to the appendix.

Bias is ubiquitous in actuarial practice. When an actuary employs benchmark or industry factors in reserving, there arises a very real potential for bias. Yet biased development factors can yield estimated ultimates with smaller MSE than ultimates based solely on a company's own experience, especially when that experience lacks sufficient credibility. The role that bias plays in estimating reserves and reserve risk has received little attention in the literature.

In Section 3 we illustrate the above with an example using the Taylor/Ashe data analyzed by Mack [2] and elsewhere in the literature. We expand on the discussion by exploring the data a bit more with the regression perspective of [1]. We show how a simple graphical diagnostic leads to a different deterministic method with a not insignificantly smaller MSE.

1 Recursive Reserve Risk Formulas

We start with the model of loss development presented in [2] and [4], employing Mack's notation.

Suppose we are given a triangle of cumulative loss amounts C_{ij} by accident year i and development age j, $1 \le i,j \le I$. The triangle is assumed to be sufficiently large that age I can be considered "ultimate." Note that for a given accident year i the triangle's current diagonal observation has column index $j = I + 1 - i$, a useful fact to keep in mind when reading Mack's formulas. The triangle in hand can be considered a sample from a theoretical set of random variables $D = \{C_{ij} \mid 1 \le i \le I, 1 \le j \le I+1-i\}$.

Under the assumptions[1]

(CL1) $E(C_{i,k+1} \mid D) = C_{ik} f_k$,

(CL2) $Var(C_{i,k+1} \mid D) = C_{ik}\sigma_k^2$ for unknown parameters σ_k^2, $1 \le i \le I$, $1 \le k \le I-1$,

and (CL3) accident years are independent,

Mack derived the following closed-form formula for the estimate of the mean square error (MSE) of the chain ladder estimated ultimate losses:

$$\hat{mse}(\hat{C}_{iI}) = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right) \tag{1}$$

where

[1] Assumptions from Mack [2], pp. 214-217, which agree with those of Model IV in [4]; labeling from Mack [3].

- $\hat{\sigma}_k^2 = \dfrac{1}{I-k-1}\sum_{i=1}^{I-k} C_{ik}\left(\dfrac{C_{i,k+1}}{C_{ik}} - \hat{f}_k\right)^2$ for $1 \le k \le I-2$; (2)
- $\hat{\sigma}_{I-1}^2$ is judgmentally selected[2];
- the link ratio estimates are calculated using the all-year weighted averages

$$\hat{f}_k = \frac{\sum_{j=1}^{I-k} C_{j,k+1}}{\sum_{j=1}^{I-k} C_{jk}};$$

- accident year losses for future ages ($k > I + 1 - i$) are predicted using the chain ladder method

$$\hat{C}_{ik} = C_{i,I+1-i}\hat{f}_{I+1-i}\cdots\hat{f}_{k-1};$$

- and, despite being scalars and not estimates, the current diagonal elements are granted "hats" ($C_{i,I+1-i} = \hat{C}_{i,I+1-i}$), which makes the formula more concise.

Formula (1) is a combination of process risk and parameter risk (a.k.a., "estimation error," but more about that later).

We next look at recursive versions of the process and parameter risk components of equation (1). In the remainder of this paper unless otherwise noted it is understood that all expectations are conditional expectations, conditional on the triangle D. Also, depending on the context, sometimes it will be convenient to refer to "risk" in terms of variance and sometimes in terms of standard deviation.

1.1 Process Risk

It can be seen in [2] that Mack's closed-form estimator[3] for the process risk component of equation (1) is

$$V\hat{a}r(C_{iI}) = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{\hat{C}_{ik}}. \tag{3}$$

Mack based the derivation of equation (3) on the recursive property[4] of process risk

$$\text{Var}(C_{ik}) = E(C_{i,k-1})\sigma_{k-1}^2 + \text{Var}(C_{i,k-1}) f_{k-1}^2 \tag{4}$$

---

[2] Mack suggests $\hat{\sigma}_{I-1}^2 = \min(\hat{\sigma}_{I-2}^4 / \hat{\sigma}_{I-3}^2, \min(\hat{\sigma}_{I-3}^2, \hat{\sigma}_{I-2}^2))$.

[3] p. 218; the hat notation in (3) shows that $\text{V}\hat{\text{a}}\text{r}(C_{iI})$ is an estimator of the variance $\text{Var}(C_{iI})$.

[4] Ibid.

for ages *k* beyond the first future diagonal for the given accident year *i*. For the first future diagonal, (4) reduces to

$$\mathrm{Var}(C_{ik}) = E(C_{i,k-1})\sigma_{k-1}^2 = C_{i,I+1-i}\sigma_{k-1}^2,$$

which is assumption CL2 above.

We obtain a recursive version of Mack's estimator for process risk by substituting estimators of the unknowns in (4):

$$\widehat{ProcessRisk}_{ik} = \begin{cases} \hat{f}_{k-1}^2 \widehat{ProcessRisk}_{i,k-1} + \hat{C}_{i,k-1}\hat{\sigma}_{k-1}^2 & \text{for } k > I+2-i \\ C_{i,I+1-i}\hat{\sigma}_{k-1}^2 & \text{for } k = I+2-i\,. \end{cases} \tag{5}$$

The process risk estimator in (5) has the same form as Murphy's recursive estimator[5]. To demonstrate that the authors' formulas are identical in substance as well as form, it remains to be shown that Mack and Murphy have the same formula for the variance estimator $\hat{\sigma}_k^2$ (both authors' models yield weighted average link ratios).

Mack's formula (2) for the variance estimator[6] can be rewritten as

$$\hat{\sigma}_k^2 = \frac{1}{I-k-1}\sum_{i=1}^{I-k}\left(C_{i,k+1} - \hat{f}_k C_{ik}\right)^2.$$

So $\hat{\sigma}_k^2$ is the sum of the squared deviations of losses at the end of the development period from the chain ladder predictions given the losses at the beginning of the period, all divided by *n*-1, where *n* is the number of terms in the summation. This is the formula for residual variance when the regression line (the paradigm in Murphy [4]) is determined by only a slope parameter, no intercept. Thus, the Mack and Murphy formulas for the variance estimator, and in turn for process risk, are equivalent.

1.2 Parameter Risk

It can be seen in Mack [2] that the author's closed-form estimator for parameter risk[7] is

$$\widehat{ParameterRisk}_{ik} = \hat{C}_{ik}^2 \sum_{j=I+1-i}^{k-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \frac{1}{\sum_{r=1}^{I-j} C_{rj}}. \tag{6}$$

This can be reformulated recursively as follows:

---

[5] Murphy [4], p. 168, under the weighted average development model.

[6] Mack [2], p. 217.

[7] In Mack's derivation of equation (1).

$$Parameter\hat{R}isk_{ik} = \hat{C}_{i,k}^2 \sum_{j=I+1-i}^{k-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \frac{1}{\sum_{r=1}^{I-j} C_{rj}}$$

$$= \hat{f}_{k-1}^2 \hat{C}_{i,k-1}^2 \left(\sum_{j=I+1-i}^{k-2} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \frac{1}{\sum_{r=1}^{I-j} C_{rj}} + \frac{\hat{\sigma}_{k-1}^2}{\hat{f}_{k-1}^2} \frac{1}{\sum_{r=1}^{I-k-1} C_{r,k-1}} \right)$$

$$= \hat{f}_{k-1}^2 \hat{C}_{i,k-1}^2 \sum_{j=I+1-i}^{k-2} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \frac{1}{\sum_{r=1}^{I-j} C_{rj}} + \hat{C}_{i,k-1}^2 \frac{\hat{\sigma}_{k-1}^2}{\sum_{r=1}^{I-k-1} C_{r,k-1}}$$

$$= \hat{f}_{k-1}^2 Parameter\hat{R}isk_{i,k-1} + \hat{C}_{i,k-1}^2 V\hat{a}r(\hat{f}_{k-1}) \ .$$

For k equal to the first future diagonal, the prior parameter risk is zero, and Mack's estimator above reduces to simply the second term.

Murphy's recursive estimator for parameter risk in Mack's notation is[8]

$$Parameter\hat{R}isk_{ik} = \begin{cases} \hat{f}_{k-1}^2 Parameter\hat{R}isk_{i,k-1} + \hat{C}_{i,k-1}^2 V\hat{a}r(\hat{f}_{k-1}) + \\ \qquad V\hat{a}r(\hat{f}_{k-1}) Parameter\hat{R}isk_{i,k-1} & \text{for } k > I+2-i \\ C_{I+1-i}^2 V\hat{a}r(\hat{f}_{k-1}) & \text{for } k = I+2-i \ . \end{cases} \tag{7}$$

Thus, the Mack and Murphy formulas differ only by the third, cross-product term in (7).[9] The derivation in theorem 2 in the appendix also yields the recursive formula (7).

2 Decomposition of the Mean Square Error

2.1 MSE Defined

Dispensing with the subscripts for accident year i and ultimate development age I, the mean square error (MSE) of the predictor $\hat{C}$ is defined[10] as the expected squared deviation of the predictor $\hat{C}$, a random variable, from the value of the random variable C being predicted; in operator notation

$$mse(\hat{C}) = E(\hat{C} - C)^2$$

where the expectation is taken with respect to the joint probability distribution of $\hat{C}$ and C.

[8] Mack [1] p. 167, assuming no constant term in the loss development model.
[9] The missing cross-product term has been noted elsewhere. See Buchwalder [1] for an example.
[10] For an example, see Mack [2], p. 216.

2.1 MSE Decomposed

Theorem 1 in the appendix shows that the MSE can be decomposed into variance and bias terms:

$$mse(\hat{C}) = Var(C) + Var(\hat{C}) + Bias^2(\hat{C}) \ . \tag{8}$$

The bias of the estimator is the difference between its mean and the mean of its target:

$$Bias(\hat{C}) = E(\hat{C}) - E(C) \ .$$

Thus, the MSE is the sum of process risk, parameter risk, and the squared bias of the estimator.

As can be seen from equation (8), it is possible for the MSE of a biased estimator to be smaller than the MSE of an unbiased estimator. For example, when a company's triangle is small or "thin" the resulting link ratios can bounce around too much from one reserve review to the next – high parameter risk. To stabilize the indications between reserve reviews, actuaries often supplement unstable company factors with more stable industry benchmarks. Do those benchmark factors introduce bias? Perhaps. If so, what might be the magnitude of that bias, and how does it compare with the corresponding reduction in MSE? Those questions are beyond the scope of this paper.

The all-year weighted averages in Mack [2] and Murphy [4] are unbiased.

2.2 Estimation Error Decomposed

Equation (12) in Theorem 1 in the appendix shows that an intermediate decomposition of the MSE has two terms, process risk and estimation error:

$$mse(\hat{C}) = Var(C) + E_{\hat{C}}(\hat{C} - \mu_C)^2 \ .$$

Estimation error $E_{\hat{C}}(\hat{C} - \mu_C)^2$ is the expected squared deviation of the estimator, not from its own mean, but from the mean of its target.[11] That expectation can be decomposed into the squared deviation of the estimator from its own mean plus the squared difference between the two means:

$$\begin{aligned} E_{\hat{C}}(\hat{C} - \mu_C)^2 &= E_{\hat{C}}(\hat{C} - \mu_{\hat{C}})^2 + (\mu_{\hat{C}} + \mu_C)^2 \\ &= Var(\hat{C}) + Bias^2(\hat{C}) \ . \end{aligned}$$

Thus, for unbiased estimators, estimation error and parameter risk are synonymous. For biased estimators, they are not.

[11] Contrast this with Mack's formulation of estimation error (Mack [2], p. 217), $(\hat{C} - \mu_C)^2$, a random variable.

2.3 The Magnitude of the Cross-Product Parameter Risk Term

Theorem 2 in the appendix proves (in parameter notation) that an estimator of the parameter risk of losses projected to age *k* is

$$\hat{\sigma}^2_{\hat{C}_k} = \hat{f}^2_{k-1}\hat{\sigma}^2_{\hat{C}_{k-1}} + \hat{C}^2_{k-1}\hat{\sigma}^2_{\hat{f}_{k-1}} + \hat{\sigma}^2_{\hat{f}_{k-1}}\hat{\sigma}^2_{\hat{C}_{k-1}} .$$

The ratio of the cross product term to the parameter risk estimator gives an idea of the relative magnitude of its contribution to the parameter risk estimate:

$$\frac{\hat{\sigma}^2_{\hat{f}_{k-1}}\hat{\sigma}^2_{\hat{C}_{k-1}}}{\hat{\sigma}^2_{\hat{C}_k}} = \frac{\hat{\sigma}^2_{\hat{f}_{k-1}}\hat{\sigma}^2_{\hat{C}_{k-1}}}{\hat{f}^2_{k-1}\hat{\sigma}^2_{\hat{C}_{k-1}} + \hat{C}^2_{k-1}\hat{\sigma}^2_{\hat{f}_{k-1}} + \hat{\sigma}^2_{\hat{f}_{k-1}}\hat{\sigma}^2_{\hat{C}_{k-1}}}$$

$$= \frac{1}{\frac{\hat{f}^2_{k-1}}{\hat{\sigma}^2_{\hat{f}_{k-1}}} + \frac{\hat{C}^2_{k-1}}{\hat{\sigma}^2_{\hat{C}_{k-1}}} + 1} . \qquad (9)$$

As can be seen from equation (9) the contribution of the cross-product term to the parameter risk estimate will be large when the denominator in (9) is small, which can occur when the link ratio variation is large relative to the square of link ratio. So for small triangles or triangles with wildly varying development, it would behoove the actuary not to ignore the cross-product term. In our experience, with reasonably stable triangles the impact of the cross-product term has been negligible.

3 An Example

Mack [1] applied his formulas to the following triangular array of data from Taylor and Ashe [5]:

357848	1124788	1735330	2218270	2745596	3319994	3466336	3606286	3833515	3901463
352118	1236139	2170033	3353322	3799067	4120063	4647867	4914039	5339085	
290507	1292306	2218525	3235179	3985995	4132918	4628910	4909315		
310608	1418858	2195047	3757447	4029929	4381982	4588268			
443160	1136350	2128333	2897821	3402672	3873311				
396132	1333217	2180715	2985752	3691712					
440832	1288463	2419861	3483130						
359480	1421128	2864498							
376686	1363294								
344014									

Given the all-year weighted average link ratios below and the cumulative loss development factors (LDFs)

	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	tail
Link Ratio	3.491	1.747	1.457	1.174	1.104	1.086	1.054	1.077	1.018	1.000
LDF	14.447	4.139	2.369	1.625	1.384	1.254	1.155	1.096	1.018	1.000

the completed triangle is

i/*k*	*k*=1	*k*=2	*k*=3	*k*=4	*k*=5	*k*=6	*k*=7	*k*=8	*k*=9	*k*=10
i=1	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
i=2	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	5,433,719
i=3	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315	5,285,148	5,378,826
i=4	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268	4,835,458	5,205,637	5,297,906
i=5	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311	4,207,459	4,434,133	4,773,589	4,858,200
i=6	396,132	1,333,217	2,180,715	2,985,752	3,691,712	4,074,999	4,426,546	4,665,023	5,022,155	5,111,171
i=7	440,832	1,288,463	2,419,861	3,483,130	4,088,678	4,513,179	4,902,528	5,166,649	5,562,182	5,660,771
i=8	359,480	1,421,128	2,864,498	4,174,756	4,900,545	5,409,337	5,875,997	6,192,562	6,666,635	6,784,799
i=9	376,686	1,363,294	2,382,128	3,471,744	4,075,313	4,498,426	4,886,502	5,149,760	5,544,000	5,642,266
i=10	344,014	1,200,818	2,098,228	3,057,984	3,589,620	3,962,307	4,304,132	4,536,015	4,883,270	4,969,825

The variance estimates are

	k=1	*k*=2	*k*=3	*k*=4	*k*=5	*k*=6	*k*=7	*k*=8	*k*=9
$\hat{\sigma}_k^2$	160,280	37,737	41,965	15,183	13,731	8,186	447	1,147	447
$\hat{\sigma}_{f_k}^2$	0.048170	0.003681	0.002789	0.000823	0.000764	0.00051	0.00004	0.00013	0.00012

Using formula (5) the process risk (variance) estimates of the future losses displayed above are calculated recursively left to right. The variance of the sum is the sum of the variances because years *i*=1…10 are independent.

	k=1	*k*=2	*k*=3	*k*=4	*k*=5	*k*=6	*k*=7	*k*=8	*k*=9	*k*=10
i=1										
i=2										2.38E+09
i=3									5.63E+09	8.19E+09
i=4								2.05E+09	7.92E+09	1.05E+10
i=5							3.17E+10	3.71E+10	4.81E+10	5.19E+10
i=6						5.07E+10	9.32E+10	1.05E+11	1.28E+11	1.34E+11
i=7					5.29E+10	1.21E+11	1.79E+11	2.01E+11	2.39E+11	2.50E+11
i=8				1.20E+11	2.29E+11	3.46E+11	4.53E+11	5.06E+11	5.93E+11	6.17E+11
i=9			5.14E+10	2.09E+11	3.41E+11	4.71E+11	5.93E+11	6.61E+11	7.72E+11	8.02E+11
i=10		5.51E+10	2.14E+11	5.42E+11	7.93E+11	1.02E+12	1.23E+12	1.37E+12	1.59E+12	1.65E+12
Sum		5.51E+10	2.65E+11	8.71E+11	1.42E+12	2.00E+12	2.58E+12	2.88E+12	3.39E+12	3.53E+12

For example, for *i*=8, *k*=6, $3.46 \cdot 10^{11} = 1.104^2 \cdot 2.29 \cdot 10^{11} + 4900545 \cdot 13731$.

Using formula (6) the parameter risk (variance) estimates of the future losses are also calculated recursively left to right. The variance of the sum is calculated using formulas in Murphy [4].

	k=1	*k*=2	*k*=3	*k*=4	*k*=5	*k*=6	*k*=7	*k*=8	*k*=9	*k*=10
i=1										
i=2										3.32E+09
i=3									3.25E+09	6.62E+09
i=4								7.38E+08	4.00E+09	7.30E+09
i=5							7.70E+09	9.17E+09	1.33E+10	1.64E+10
i=6						1.04E+10	2.08E+10	2.38E+10	3.05E+10	3.46E+10
i=7					9.99E+09	2.50E+10	3.99E+10	4.52E+10	5.59E+10	6.16E+10
i=8				2.29E+10	4.59E+10	7.43E+10	1.03E+11	1.15E+11	1.39E+11	1.49E+11
i=9			6.84E+09	3.04E+10	5.18E+10	7.59E+10	9.99E+10	1.12E+11	1.33E+11	1.42E+11
i=10		5.70E+09	2.27E+10	6.06E+10	9.13E+10	1.21E+11	1.51E+11	1.68E+11	1.98E+11	2.08E+11
Sum		5.70E+09	4.16E+10	2.39E+11	4.95E+11	9.20E+11	1.44E+12	1.64E+12	2.12E+12	2.46E+12

For example, for *i*=8, *k*=6,

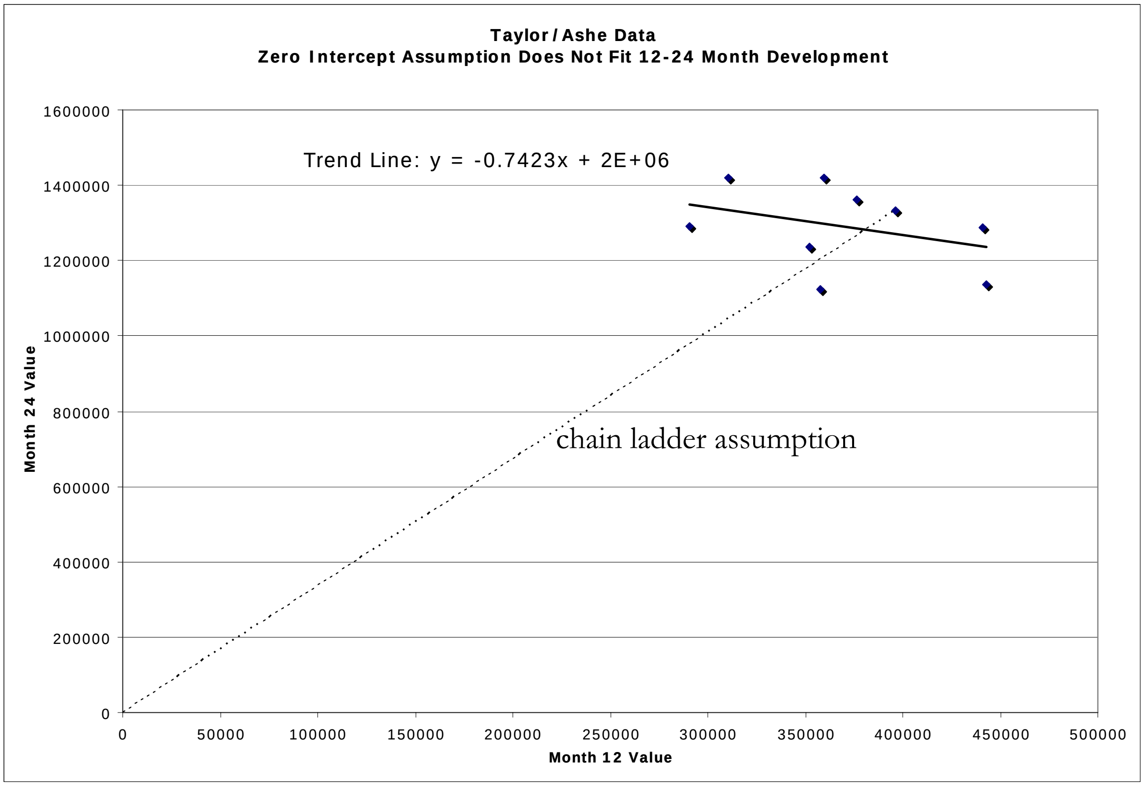
$$7.43 \cdot 10^{10} = 1.104^2 \cdot 4.59 \cdot 10^{10} + 4900545^2 \cdot 0.000764 + 0.000764 \cdot 4.59 \cdot 10^{10}.$$

Comparisons of these Murphy-formula results with the Mack-formula results from Mack [2] are displayed in row detail, and in total, in the following table:

	Reserve Risk Estimates					
Origination Year	**Mack Formula**			**Murphy Formula**		
	Process	**Parameter**	**Total**	**Process**	**Parameter**	**Total**
i=2	48,832	57,628	75,535	48,832	57,628	75,535
i=3	90,524	81,338	121,699	90,524	81,340	121,700
i=4	102,622	85,464	133,549	102,622	85,467	133,551
i=5	227,880	128,078	261,406	227,880	128,091	261,412
i=6	366,582	185,867	411,010	366,582	185,907	411,028
i=7	500,202	248,023	558,317	500,202	248,110	558,356
i=8	785,741	385,759	875,328	785,741	385,991	875,430
i=9	895,570	375,893	971,258	895,570	376,222	971,385
i=10	1,284,882	455,270	1,363,155	1,284,882	455,957	1,363,385
Total:	1,878,292	1,568,532	2,447,095	1,878,292	1,569,349	2,447,618

The Mack and Murphy process risk estimates are identical. Differences in parameter risk occur, at most, only in the 3rd or 4th significant digit.

Continuing with this example, the regression perspective of Murphy [4] provides additional insight into the Taylor/Ashe data. The graphical display below of the historical relationship between 12- and 24-month losses clearly shows that the data violate the first chain ladder assumption (Mack's CL1), i.e., that the expected relationship is a line through the origin.



Although the indicated slope of the trend line is negative, the regression statistics support the statement that it is not significantly different from zero, implying that the 12- and 24-month losses are actually uncorrelated. Therefore, a reasonable estimate of the 24-month losses for year 10 would simply be the average of all of the previous years' 24-month losses, 1,290,505. This estimate would be reasonable not just from a statistical standpoint but from a business standpoint if we knew, for instance, that all losses are on-level and of equal exposure. The standard deviation of those losses is 108,885 = process risk, and the standard deviation of the mean is 38497 = sqrt(108895^2/(9-1)) = parameter risk.

This demonstrates one of the advantages of recursive formulas: flexibility. The recursive formulas (5) and (7) do not know how the predictions and variances are estimated, nor do they care (e.g., see Theorem 2). One need only substitute these two new process risk and parameter risk estimates for year 10 into the corresponding (i=10,k=2) cells in the tables above and the recursive calculations for k>2 carry on as before. The new comparison table is

	Reserve Risk Estimates					
Origination Year	Mack Formula			Murphy Formula		
	Process	Parameter	Total	Process	Parameter	Total
i=2	48,832	57,628	75,535	48,832	57,628	75,535
i=3	90,524	81,338	121,699	90,524	81,340	121,700
i=4	102,622	85,464	133,549	102,622	85,467	133,551
i=5	227,880	128,078	261,406	227,880	128,091	261,412
i=6	366,582	185,867	411,010	366,582	185,907	411,028
i=7	500,202	248,023	558,317	500,202	248,110	558,356
i=8	785,741	385,759	875,328	785,741	385,991	875,430
i=9	895,570	375,893	971,258	895,570	376,222	971,385
i=10	1,284,882	455,270	1,363,155	980,971	390,295	1,055,762
Total:	1,878,292	1,568,532	2,447,095	1,685,041	1,568,504	2,302,079

Thus, after a simple diagnostic of the underlying data and an appropriate adjustment in the actuarial projection, process risk for year 10 is reduced by 22.5%, parameter risk by 14.3%, and total risk by 21.5%, and the total risk estimate for all years combined is 6% lower than that produced by the Mack method. This example also points out how it is not necessary – or even advisable – to use a single reserving method for the entire future development of a given year. In some instances it is beneficial to "change methods in the middle of the development stream."

4 Conclusion

Although Mack's reserve risk formulas omit a parameter risk cross-product term, the understatement should be negligible for reasonably behaved triangles. The advantage of closed-form formulas as in Mack [2] is that they are concise. Recursive formulas by Murphy [4], by Mack [3], and in this paper are not as concise but are more flexible, e.g., allowing for projections based on a shift in model from one development period to the next.

Mean square error is comprised of process risk, parameter risk, and bias. Estimation error and parameter risk are equivalent when the link ratios are unbiased. Within the context of the chain ladder method, utilization of industry benchmark factors might introduce bias into the projections, but in the actuary's judgment the resulting stabilization may outweigh whatever bias might occur. Estimating the magnitude of the potential for bias and reduction in MSE are areas of further actuarial research.

Appendix

The definition of the mean square error (MSE) of the predictor $\hat{C}$ is the expected squared deviation of the (random variable) predictor $\hat{C}$ from the value of the random variable C being predicted:

$$mse(\hat{C}) = E(\hat{C} - C)^2 \tag{10}$$

where the expectation is taken with respect to the joint probability distribution of $\hat{C}$ and C.

Theorem 1: The MSE Decomposition Theorem

$$\text{mse}(\hat{C}) = \text{Var}(C) + \text{Var}(\hat{C}) + \text{Bias}^2(\hat{C}).$$

Proof: Let $f(c,\hat{c})$ represent the joint density of C and $\hat{C}$. Then the MSE is the integral

$$\text{mse}(\hat{C}) = \iint (\hat{c} - c)^2 f(c,\hat{c})\,dc\,d\hat{c}$$

taken over the joint sample space.

To decompose the MSE into variance and bias components, we will use the fact that the joint density of the two random variables can be factored into a conditional density and a marginal density:

$$f(c,\hat{c}) = f(c \mid \hat{c})\, f(\hat{c}).$$

This fact allows us to write equation (10) as

$$\text{mse}(\hat{C}) = E_{\hat{C}}(E((\hat{C} - C)^2 \mid \hat{C})) \tag{11}$$

where the inner expectation is taken with respect to C conditional on the value of $\hat{C}$. We will manipulate the inner expectation first, taking advantage of the "scalar" nature of $\hat{C}$ with respect to that conditional expectation.

We add and subtract the mean μ_C of the predicted random variable inside the quadratic, group the result into two terms, square the binomial, and observe that the cross-product term disappears. To wit

$$\begin{aligned}
E_C((\hat{C} - C)^2 \mid \hat{C}) &= E_C[(\hat{C} - \mu_C + \mu_C - C)^2 \mid \hat{C}] \\
&= E_C[((\hat{C} - \mu_C) + (\mu_C - C))^2 \mid \hat{C}] \\
&= E_C[(\hat{C} - \mu_C)^2 + 2(\hat{C} - \mu_C)(\mu_C - C) + (\mu_C - C)^2 \mid \hat{C}] \\
&= E_C[(\hat{C} - \mu_C)^2 \mid \hat{C}] + 2E_C[(\hat{C} - \mu_C)(\mu_C - C) \mid \hat{C}] + E_C[(\mu_C - C)^2 \mid \hat{C}].
\end{aligned}$$

The third term above is just Var(C), the first term (conditional on $\hat{C}$) is simply $(\hat{C} - \mu_C)^2$, and the middle term disappears because

$$\begin{aligned}
E_C[(\hat{C} - \mu_C)(\mu_C - C) \mid \hat{C}] &= E_C[(\hat{C}\mu_C - \hat{C}C - \mu_C^2 + \mu_C C) \mid \hat{C}] \\
&= \hat{C}\mu_C - \hat{C}E_C[C \mid \hat{C}] - \mu_C^2 + \mu_C E_C[C \mid \hat{C}] \\
&= \hat{C}\mu_C - \hat{C}\mu_C - \mu_C^2 + \mu_C^2 \\
&= 0 .
\end{aligned}$$

Substituting these expressions into (11), we have that

$$\text{mse}(\hat{C}) = Var(C) + E_{\hat{C}}(\hat{C} - \mu_C)^2 \tag{12}$$

which shows that the MSE equals the process variance plus the expected squared deviation between the predictor and the mean of its target.[12] The second term on the right in (12) is called "estimation error."

To continue the decomposition, we address the estimation error term in (12) by adding and subtracting the mean $\mu_{\hat{C}}$ inside the quadratic and proceeding as above:

$$\begin{aligned}
E_{\hat{C}}(\hat{C} - \mu_C)^2 &= E_{\hat{C}}(\hat{C} - \mu_{\hat{C}} + \mu_{\hat{C}} - \mu_C)^2 \\
&= E_{\hat{C}}[(\hat{C} - \mu_{\hat{C}}) + (\mu_{\hat{C}} - \mu_C)]^2 \\
&= E_{\hat{C}}[(\hat{C} - \mu_{\hat{C}})^2] + 2E_{\hat{C}}[(\hat{C} - \mu_{\hat{C}})(\mu_{\hat{C}} - \mu_C)] + E_{\hat{C}}[(\mu_{\hat{C}} - \mu_C)^2] \\
&= Var(\hat{C}) + 2E_{\hat{C}}[(\hat{C}\mu_{\hat{C}} - \hat{C}\mu_C - \mu_{\hat{C}}^2 + \mu_{\hat{C}}\mu_C)] + (\mu_{\hat{C}} - \mu_C)^2 \\
&= Var(\hat{C}) + 2[\mu_{\hat{C}}^2 - \mu_{\hat{C}}\mu_C - \mu_{\hat{C}}^2 + \mu_{\hat{C}}\mu_C] + (\mu_{\hat{C}} - \mu_C)^2 \\
&= Var(\hat{C}) + (Bias(\hat{C}))^2 .
\end{aligned}$$

Substituting this expression for $E_{\hat{C}}(\hat{C} - \mu_C)^2$ into (12), we have

$$mse(\hat{C}) = Var(C) + Var(\hat{C}) + Bias^2(\hat{C})$$

which proves the theorem.

Theorem 2: The Parameter Risk Recursion Theorem

$$V\hat{a}r(\hat{C}_{ik}) = \hat{f}_{k-1}^2 V\hat{a}r(\hat{C}_{i,k-1}) + \hat{C}_{i,k-1}^2 V\hat{a}r(\hat{f}_{k-1}) + V\hat{a}r(\hat{f}_{k-1}) V\hat{a}r(\hat{C}_{i,k-1})$$

Proof: Following a similar path as in equation (4) in Section 1 above:

$$\begin{aligned}
Var(\hat{C}_{ik}) &= E_{\hat{C}_{i,k-1}}(Var(\hat{C}_{ik} \mid \hat{C}_{i,k-1})) + Var_{\hat{C}_{i,k-1}}(E(\hat{C}_{ik} \mid \hat{C}_{i,k-1})) \\
&= E_{\hat{C}_{i,k-1}}(Var(\hat{f}_{k-1}\hat{C}_{i,k-1} \mid \hat{C}_{i,k-1})) + Var_{\hat{C}_{i,k-1}}(E(\hat{f}_{k-1}\hat{C}_{i,k-1} \mid \hat{C}_{i,k-1})) \\
&= E_{\hat{C}_{i,k-1}}(\hat{C}_{i,k-1}^2 Var(\hat{f}_{k-1})) + Var_{\hat{C}_{i,k-1}}(\hat{C}_{i,k-1} E(\hat{f}_{k-1})) \\
&= Var(\hat{f}_{k-1}) E(\hat{C}_{i,k-1}^2) + Var_{\hat{C}_{i,k-1}}(\hat{C}_{i,k-1} f_{k-1}) \\
&= Var(\hat{f}_{k-1})(Var(\hat{C}_{i,k-1}) + E^2(\hat{C}_{i,k-1})) + f_{k-1}^2 Var(\hat{C}_{i,k-1}) \\
&= Var(\hat{f}_{k-1}) Var(\hat{C}_{i,k-1}) + Var(\hat{f}_{k-1}) C_{i,k-1}^2 + f_{k-1}^2 Var(\hat{C}_{i,k-1}) .
\end{aligned}$$

Substituting estimates for the unknown parameters yields the desired result.

---

[12] Contrast this with Mack's expression for the MSE in [2]: $\text{mse}(\hat{C}) = \text{Var}(C) + (\mu_C - \hat{C})^2$.

Acknowledgment

The author wishes to acknowledge Ali Majidi, Doug Collins, and Emmanuel Bardis for their clarifying comments, direction, and support.

References

[1] Buchwalder, M., et al. 2005. "Legal Valuation Portfolio in Non-Life Insurance." Lecture, ASTIN Colloquium, Zurich, Switzerland, September 5-7, 2005.
[2] Mack, Thomas. 1993. "Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates." *ASTIN Bulletin* 23, no. 2:213-225.
[3] Mack, Thomas. 1999. "The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor." *ASTIN Bulletin* 29, no. 2:361-366.
[4] Murphy, Daniel. 1994. "Unbiased Loss Development Factors." *PCAS* 81:154-222.
[5] Taylor, G., and F. Ashe. 1983. "Second Moments of Estimates of Outstanding Claims." *Journal of Econometrics* 23:37-61.

Biography

Daniel Murphy is a consulting actuary with the Tillinghast business of Towers Perrin. He is a Fellow of the CAS, a Member of the American Academy of Actuaries, and a member of the CAS program planning committee.