Stochastic Reserving: Mack and Bootstrapping

Casualty Actuarial Society
Spring Meeting – Puerto Rico
May, 2006

Dave Clark   American Re-Insurance Company

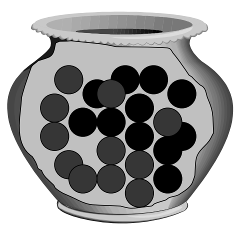
Introduction to Stochastic Reserving

Agenda:
- Refresher on Statistical Models
- Bootstrapping Method
- Mack Method
- Limitations and Caveats

Introduction to Stochastic Reserving

In Classical Statistics, we assume that there is a fixed "population" from which we are sampling.

The world is simple, like an urn with an unknown number of red and black balls.



For Reserving: The historical loss development data is viewed as a sample from a "population" of possible outcomes.

Introduction to Stochastic Reserving

Key idea: A function of random variables is also a random variable.

For example, given a sample = $X_1\ X_2\ X_3 \ldots X_N$

The sample mean is also a random variable with an expected value and variance to be estimated.

Similarly, our estimate of the future payments is a function of the payments to date by year. This is also a random variable with a mean and variance.

Introduction to Stochastic Reserving

Model	Method	Reserve
Mathematical description of the loss-generating phenomenon	The algorithm for using the data to calculate an estimate	The $ amount actually carried in the financial statement

Introduction to Stochastic Reserving

We will look at two models.

- Additive Over-Dispersed Poisson (ODP)
- Multiplicative Chainladder (Mack)

Both of these models lead to the ***same*** chainladder method to estimate ultimate losses, but they include different variability assumptions and so have ***different*** estimates of variability.

See Venter (1998) for ideas on tests to compare models.

Introduction to Stochastic Reserving

Additive Over-Dispersed Poisson Model
(England & Verrall)

Including Bootstrapping

Introduction to Stochastic Reserving

The Bootstrapping Method:

"Bootstrapping" is a method for calculating the standard error of an estimate.

First we need to describe a model.

1 2 3 4

Introduction to Stochastic Reserving

Incremental Paid Loss Model:

- Expected Loss based on accident year (y) and development period (d) factors: $\alpha_y \times \beta_d$
- Incremental paid losses $C_{y,d}$ are independent
- Constant Variance/Mean Ratio σ^2

This can be modeled as an Over-Dispersed Poisson (ODP) distribution

9

1 2 3 4

Introduction to Stochastic Reserving

$\boxed{\text{ODP Model}}$

$$E(C_{y,d}) = \alpha_y \cdot \beta_d$$

$$Var(C_{y,d}) = \sigma^2 \cdot E(C_{y,d})$$

$$\text{MLE for } \hat{\alpha}_y \text{ and } \hat{\beta}_d = \text{chainladder}$$

10

1 2 3 4

Introduction to Stochastic Reserving

The Over-Dispersed Poisson (ODP) model is attractive because:

- The maximum likelihood estimate (MLE) of the expected values equal the chain-ladder estimates.
- We can estimate the process variance as a simple multiple of the estimated reserve.

$$\hat{\sigma}^2 \approx \frac{1}{(n-p)} \cdot \sum_{y,d} \frac{\left(C_{y,d} - \hat{C}_{y,d}\right)^2}{\hat{C}_{y,d}}$$

11

1 2 3 4

Introduction to Stochastic Reserving

But what about the uncertainty in the estimate of the mean (the "parameter variance")?

$$E\left[\left(C_{y,d} - \hat{C}_{y,d}\right)^2\right] = \underbrace{Var\left(C_{y,d}\right)}_{\text{Process Variance}} + \underbrace{Var\left(\hat{C}_{y,d}\right)}_{\text{Parameter Variance}}$$

12

1 2 3 4

Introduction to Stochastic Reserving

The Parameter Variance component can be evaluated in either of two ways:

- Analytically: Using the "delta method"
  - Based on inverting the matrix of second derivatives of the log-likelihood function
- Simulation: Using Bootstrapping
  - Based on creating many "what if" triangles and seeing how the reserve estimates from this differ

13

1 2 3 4

Introduction to Stochastic Reserving

Steps in Bootstrapping:

- Calculate Chainladder Ultimates
- Calculate "Expected" incremental triangle
- Calculate residuals = $(A-E)/(\sigma \cdot E^{1/2})$
- Generate a pseudo-triangle from re-sampled residuals
- Calculate Chainladder Ultimates from pseudo-triangle
- Repeat, Repeat, Repeat

14

1 2 3 4

Introduction to Stochastic Reserving

Two Types of Bootstraps:

- Nonparametric
  - Uses empirical residuals
  - Does not require a distributional assumption
  - Works best for large samples (at least 100 points)
- Parametric
  - Uses simulations from a theoretical distribution (e.g., ODP or Normal) with mean and variance parameters selected from the original data

15

1 2 3 4

Introduction to Stochastic Reserving

Multiplicative / Autocorrelation Model
(Mack)

16

1 2 3 4

Introduction to Stochastic Reserving

The Distribution-Free calculation introduced by Thomas Mack in 1993 is an alternative model that is also consistent with the chainladder method.

Here we do not assume independence of incremental payments. Instead, we assume that each payment is correlated with the earlier payments for that accident year. It is the age-to-age factors that are assumed to be independent.

17

1 2 3 4

Introduction to Stochastic Reserving

The Mack model is attractive because:

- The Best Linear Unbiased Estimator ("BLUE") for the reserves equals the chain-ladder estimates.
  - See Murphy 1994 for further details.
- The model is robust in handling [some] negative development increments.

18

1 2 3 4

Introduction to Stochastic Reserving

Mack Model

Let $D_{y,d} = C_{y,1} + C_{y,2} + \cdots + C_{y,d}$

$$E(D_{y,d} \mid D_{y,d-1}) = \lambda_{d-1} \cdot D_{y,d-1}$$

$$Var(D_{y,d} \mid D_{y,d-1}) = \sigma^2_{d-1} \cdot D_{y,d-1}$$

19

1 2 3 4

Introduction to Stochastic Reserving

Mack Model

Note that λ is an "age-to-age" factor

$$E(D_{y,d} \mid D_{y,d-1}) = \lambda_{d-1} \cdot D_{y,d-1}$$

$$E(D_{y,d} \mid D_{y,d-2}) = \lambda_{d-1} \cdot \lambda_{d-2} \cdot D_{y,d-2}$$

$$E(D_{y,d} \mid D_{y,d-3}) = \lambda_{d-1} \cdot \lambda_{d-2} \cdot \lambda_{d-3} \cdot D_{y,d-3}$$

20

1 2 3 4

Introduction to Stochastic Reserving

Mack Model

Variance can also be stated recursively

$$Var(D_{y,d} \mid D_{y,d-1}) = \sigma_{d-1}^2 \cdot D_{y,d-1}$$

$$\begin{aligned} Var(D_{y,d} \mid D_{y,d-2}) &= \sigma_{d-1}^2 \cdot D_{y,d-1} + Var(\lambda_{d-1} \cdot D_{y,d-1} \mid D_{y,d-2}) \\ &= \sigma_{d-1}^2 D_{y,d-1} + \lambda_{d-1}^2 \cdot \sigma_{d-2}^2 \cdot D_{y,d-2} \end{aligned}$$

1 2 3 4

Introduction to Stochastic Reserving

The variance multiplier in the Mack model is similar to what we saw for the ODP. However, he defines a new sigma (σ) for each development age (d).

$$\hat{\sigma}_{d-1}^2 \approx \frac{1}{(n-p)} \cdot \sum_d \frac{(D_{y,d} - \lambda_{d-1} \cdot D_{y,d-1})^2}{\lambda_{d-1} \cdot D_{y,d-1}}$$

1 2 3 4

Introduction to Stochastic Reserving

The Parameter Variance component can be evaluated in either of two ways:

- Analytically: Using the formulas given in Mack's paper
  - This is not the "delta method" from MLE, since Mack does not explicitly make a distributional assumption
- Simulation: Using Bootstrapping
  - Based on creating many "what if" triangles and seeing how the reserve estimates from this differ

1 2 3 4

Introduction to Stochastic Reserving

<u>Limitations & Caveats:</u>

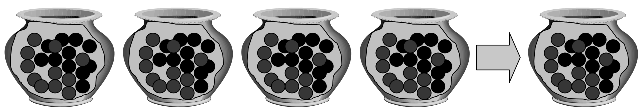
- Assumptions on independence and identical distributions (iid) are weak – an unchanging world is assumed!
- Concern of over-parameterization
- Models handle some zero or negative values, but do not work for very sparse data
- Difficulty in variance of "tail" beyond triangle

1 2 3 4

Introduction to Stochastic Reserving

An Unchanging World:

Assumes that the future payments will be from a distribution identical to the past.



1 2 3 4

Introduction to Stochastic Reserving

Over-Parameterization:

E&V ODP	Mack
10 Years 55 Points	10 Years 45 Points
19 Parameters +1 "sigma" σ	9 Parameters +9 "sigmas" σ_d

1 2 3 4

Introduction to Stochastic Reserving

Tail Factor Extrapolation:

- Select a tail factor and include it as "quasi-data" as though it were part of the original triangle
  - This ignores the additional parameter variance associated with the selection
- Extrapolate a tail-factor from the triangle
  - Need some formula for extrapolated value
  - For bootstrapping, this is done at each iteration

1 2 3 4

Introduction to Stochastic Reserving

References (Bootstrap):

* "Analytic and bootstrap estimates of prediction errors in claims reserving," Insurance: Mathematics and Economics 25 (1999) – Peter England & Richard Verrall.

"Addendum to 'Analytic and bootstrap estimates of prediction errors in claims reserving'," Insurance: Mathematics and Economics 31 (2002) – Peter England.

"Stochastic Claims Reserving in General Insurance," B.A.J. 8.III (2002) – Peter England & Richard Verrall.

1 2 3 4

Introduction to Stochastic Reserving

References (Mack):

"A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserves," ASTIN Bulletin, Vol. 21, No. 1, 1991 – Thomas Mack.

* "Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates," ASTIN Bulletin, Vol. 23, No. 2, 1993 – Thomas Mack.

"The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor," ASTIN Bulletin, Vol. 29, No. 2, 1999 – Thomas Mack.

29

1 2 3 4

Introduction to Stochastic Reserving

References (Other):

"An Introduction to the Bootstrap," Chapman & Hall 1993 – B. Efron & R. Tibshirani.

"Unbiased Loss Development Factors," Proceedings of the CAS, 1994 – Daniel Murphy.

"Testing the Assumptions of Age-to-Age Factors," Proceedings of the CAS, 1998 – Gary Venter.

30

Comparison of Standard Errors for Two Reserving Models

Completing the Triangle

ODP - E&V

AY	Diagonal	Reserve	Process	to Res	Parameter	to Res	Prediction	to Res
2	5,339,085	94,634	70,554	74.6%	84,522	89.3%	110,099	116.3%
3	4,909,315	469,511	157,153	33.5%	148,248	31.6%	216,042	46.0%
4	4,588,268	709,638	193,204	27.2%	175,287	24.7%	260,871	36.8%
5	3,873,311	984,889	227,610	23.1%	200,836	20.4%	303,549	30.8%
6	3,691,712	1,419,459	273,250	19.3%	256,843	18.1%	375,012	26.4%
7	3,483,130	2,177,641	338,448	15.5%	361,732	16.6%	495,376	22.7%
8	2,864,498	3,920,301	454,107	11.6%	646,389	16.5%	789,957	20.2%
9	1,363,294	4,278,972	474,426	11.1%	932,791	21.8%	1,046,508	24.5%
10	344,014	4,625,811	493,279	10.7%	1,917,664	41.5%	1,980,091	42.8%
Total	30,456,627	18,680,856	991,281	5.3%	2,773,841	14.8%	2,945,646	15.8%

"Distribution Free" Chainladder - Mack

AY	Diagonal	Reserve	Process	to Res	Parameter	to Res	Prediction	to Res
2	5,339,085	94,634	48,832	51.6%	57,628	60.9%	75,535	79.8%
3	4,909,315	469,511	90,524	19.3%	81,338	17.3%	121,699	25.9%
4	4,588,268	709,638	102,622	14.5%	85,464	12.0%	133,549	18.8%
5	3,873,311	984,889	227,880	23.1%	128,078	13.0%	261,406	26.5%
6	3,691,712	1,419,459	366,582	25.8%	185,867	13.1%	411,010	29.0%
7	3,483,130	2,177,641	500,202	23.0%	248,023	11.4%	558,317	25.6%
8	2,864,498	3,920,301	785,741	20.0%	385,759	9.8%	875,328	22.3%
9	1,363,294	4,278,972	895,570	20.9%	375,893	8.8%	971,258	22.7%
10	344,014	4,625,811	1,284,882	27.8%	455,270	9.8%	1,363,155	29.5%
Total	30,456,627	18,680,856	1,878,292	10.1%	1,568,532	8.4%	2,447,095	13.1%

Comparison of Standard Errors for Two Reserving Models

Next CY Diagonal

ODP - E&V

AY	Diagonal	Next CY	Process	to Res	Parameter	to Res	Prediction	to Res
2	5,339,085	94,634	70,554	74.6%	84,522	89.3%	110,099	116.3%
3	4,909,315	375,833	140,604	37.4%	117,373	31.2%	183,155	48.7%
4	4,588,268	247,190	114,029	46.1%	75,063	30.4%	136,517	55.2%
5	3,873,311	334,148	132,577	39.7%	78,431	23.5%	154,039	46.1%
6	3,691,712	383,287	141,991	37.0%	81,651	21.3%	163,793	42.7%
7	3,483,130	605,548	178,473	29.5%	112,238	18.5%	210,832	34.8%
8	2,864,498	1,310,258	262,529	20.0%	225,048	17.2%	345,786	26.4%
9	1,363,294	1,018,834	231,500	22.7%	229,139	22.5%	325,725	32.0%
10	344,014	856,804	212,295	24.8%	358,489	41.8%	416,633	48.6%
Total	30,456,627	5,226,536	524,331	10.0%	532,575	10.2%	747,368	14.3%

"Distribution Free" Chainladder - Mack

AY	Diagonal	Next CY	Process	to Res	Parameter	to Res	Prediction	to Res
2	5,339,085	94,634	48,832	51.6%	57,628	60.9%	75,535	79.8%
3	4,909,315	375,833	75,052	20.0%	56,970	15.2%	94,225	25.1%
4	4,588,268	247,190	45,268	18.3%	27,163	11.0%	52,792	21.4%
5	3,873,311	334,148	178,062	53.3%	87,733	26.3%	198,502	59.4%
6	3,691,712	383,287	225,149	58.7%	102,068	26.6%	247,204	64.5%
7	3,483,130	605,548	229,965	38.0%	99,925	16.5%	250,737	41.4%
8	2,864,498	1,310,258	346,712	26.5%	151,271	11.5%	378,275	28.9%
9	1,363,294	1,018,834	226,818	22.3%	82,715	8.1%	241,429	23.7%
10	344,014	856,804	234,816	27.4%	75,503	8.8%	246,656	28.8%
Total	30,456,627	5,226,536	610,035	11.7%	266,139	5.1%	665,562	12.7%

Some Comments on the Mack Variance Formulas

Dave Clark - May 2006

Recursive definition for expected value: $E\left(D_{ik} \mid D_{i,k-1}\right) = \lambda_{k-1} \cdot E\left(D_{i,k-1}\right)$

$D_{i,k}$ = cumulative loss for AY i at development period k

λ_{k-1} = weighted-average age-to-age factor from development period $k-1$ to k

The "process variance" is likewise modeled in a recursive form. The variance increases when more development periods are included.

$$\begin{aligned}
Var\left(D_{i,k} \mid D_{i,k-1}\right) &= \sigma_{k-1}^2 \cdot E\left(D_{i,k-1}\right) \\
&= E\left(D_{i,k}\right)^2 \cdot \left\{\frac{\sigma_{k-1}^2}{\lambda_{k-1}^2 \cdot E\left(D_{i,k-1}\right)}\right\}
\end{aligned}$$

$$\begin{aligned}
Var\left(D_{i,k} \mid D_{i,k-2}\right) &= Var(D_{i,k} \mid D_{i,k-1}) + Var\left(\lambda_{k-1} \cdot D_{i,k-1} \mid D_{i,k-2}\right) \\
&= \sigma_{k-1}^2 \cdot E\left(D_{i,k-1}\right) + \lambda_{k-1}^2 \cdot \sigma_{k-2}^2 \cdot E\left(D_{i,k-2}\right) \\
&= E\left(D_{i,k}\right)^2 \cdot \left\{\frac{\sigma_{k-1}^2}{\lambda_{k-1}^2 \cdot E\left(D_{i,k-1}\right)} + \frac{\sigma_{k-2}^2}{\lambda_{k-2}^2 \cdot E\left(D_{i,k-2}\right)}\right\}
\end{aligned}$$

$$\begin{aligned}
Var\left(D_{i,k} \mid D_{i,k-3}\right) &= Var(D_{i,k} \mid D_{i,k-1}) + Var\left(\lambda_{k-1} \cdot D_{i,k-1} \mid D_{i,k-2}\right) + Var\left(\lambda_{k-1} \cdot \lambda_{k-2} \cdot D_{i,k-2} \mid D_{i,k-3}\right) \\
&= \sigma_{k-1}^2 \cdot E\left(D_{i,k-1}\right) + \lambda_{k-1}^2 \cdot \sigma_{k-2}^2 \cdot E\left(D_{i,k-2}\right) + \lambda_{k-1}^2 \cdot \lambda_{k-2}^2 \cdot \sigma_{k-3}^2 \cdot E\left(D_{i,k-3}\right) \\
&= E\left(D_{i,k}\right)^2 \cdot \left\{\frac{\sigma_{k-1}^2}{\lambda_{k-1}^2 \cdot E\left(D_{i,k-1}\right)} + \frac{\sigma_{k-2}^2}{\lambda_{k-2}^2 \cdot E\left(D_{i,k-2}\right)} + \frac{\sigma_{k-3}^2}{\lambda_{k-3}^2 \cdot E\left(D_{i,k-3}\right)}\right\}
\end{aligned}$$

This expansion can continue for any number of development periods "n."

$$Var\left(D_{i,k} \mid D_{i,k-n}\right) = E\left(D_{i,k}\right)^2 \cdot \left\{\sum_{d=1}^{n} \frac{\hat{\sigma}_{k-d}^2}{\lambda_{k-d}^2 \cdot E\left(D_{i,k-d}\right)}\right\}$$

The reserve (considering ultimate = age N) is given as: $E\left(R_i\right) = E\left(D_{i,N}\right) - D_{i,N+1-i}$

$$Var\left(R_i\right) = Var\left(D_{i,N} \mid D_{i,N+1-i}\right) = E\left(D_{i,N}\right)^2 \cdot \left\{\sum_{d=1}^{i-1} \frac{\hat{\sigma}_{N-d}^2}{\lambda_{N-d}^2 \cdot E\left(D_{i,N-d}\right)}\right\}$$

And then the mean squared error (MSE), including Parameter Variance is given as below:

$$MSE\left(D_{i,N} \mid D_{i,N+1-i}\right) \;=\; E\left(D_{i,N}\right)^2 \cdot \left\{ \sum_{d=1}^{i-1} \frac{\hat{\sigma}^2_{N-d}}{\lambda^2_{N-d}} \left(\frac{1}{E\left(D_{i,N-d}\right)} + \frac{1}{\sum_{y=1}^{N-d-1} D_{y,N-d}} \right) \right\}$$

Think of this as analogous to: $\approx \hat{\sigma}^2 + \dfrac{\hat{\sigma}^2}{n}$

We remember that $\hat{\lambda}_k$ is the dollar-weighted average age-to-age factor, so the additional term included for parameter variance is the total dollars in the denominator of the estimator of each age-to-age factor $\hat{\lambda}_k$. This represents the variance due to the error in the "sample mean" age-to-age factor.

A modest re-arrangement of this expression is also useful.
If we let $LDF_k = \lambda_k \cdot \lambda_{k+1} \cdots \lambda_{N-1}$, such that $E\left(D_{i,N}\right) = LDF_k \cdot D_{i,k}$, then we can re-write the mean square error (MSE) expression as follows:

$$MSE\left(D_{i,N} \mid D_{i,N+1-i}\right) \;=\; E\left(D_{i,N}\right) \cdot \left\{ \sum_{d=1}^{i-1} \frac{\hat{\sigma}^2_{N-d}}{\lambda^2_{N-d}} \cdot LDF_d \right\} + E\left(D_{i,N}\right)^2 \cdot \left\{ \sum_{d=1}^{i-1} \frac{\hat{\sigma}^2_{N-d}}{\lambda^2_{N-d}} \left(\frac{1}{\sum_{y=1}^{N-d-1} D_{y,N-d}} \right) \right\}$$

This form shows that "process variance" is proportional to the loss dollars in the accident year, implying that the CV decreases for years with greater volume. By contrast, the "parameter variance" is proportional to the loss dollars squared, implying that the CV does not decrease even when loss volume increases.

When we want to calculate the covariance between the reserves for any two accident years (say, i and j), the parameter variance terms becomes:

$$Cov\left(D_{i,N}, D_{j,N} \mid D_{i,N+1-i}, D_{j,N+1-j}\right) \;=\; E\left(D_{i,N}\right) \cdot E\left(D_{j,N}\right) \cdot \left\{ \sum_{d=1}^{\min(i,j)-1} \frac{\hat{\sigma}^2_{N-d}}{\lambda^2_{N-d}} \left(\frac{1}{\sum_{y=1}^{N-d-1} D_{y,N-d}} \right) \right\}$$

The MSE for the reserves overall includes the sum of the matrix of covariances terms.