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

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.

Additive Over-Dispersed Poisson Model
(England & Verrall)

Including Bootstrapping

The Bootstrapping Method:
“Bootstrapping” is a method for calculating the standard error of an estimate.

First we need to describe a model.
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 \( \sigma^2 \)

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

\[
\begin{align*}
\text{ODP Model} \\
E(C_{y,d}) &= \alpha_y \cdot \beta_d \\
\text{Var}(C_{y,d}) &= \sigma^2 \cdot E(C_{y,d}) \\
\text{MLE for } \hat{\alpha}_y \text{ and } \hat{\beta}_d &= \text{chainladder}
\end{align*}
\]

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.

\[
\hat{\sigma}^2 \approx \frac{1}{(n-p)} \sum_{y,d} \frac{(C_{y,d} - \hat{C}_{y,d})^2}{C_{y,d}}
\]

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

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

- Process Variance
- Parameter Variance
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

**Steps in Bootstrapping:**
- Calculate Chainladder Ultimates
- Calculate “Expected” incremental triangle
- Calculate residuals $= (A-E)/(\sigma E^{1/2})$
- Generate a pseudo-triangle from re-sampled residuals
- Calculate Chainladder Ultimates from pseudo-triangle
- Repeat, Repeat, Repeat

**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

**Multiplicative / Autocorrelation Model**
(Mack)
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.

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.

Mack Model

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

\[
E(D_{y,d} | D_{y,d-1}) = \lambda_{d-1} \cdot D_{y,d-1}
\]

\[
\text{Var}(D_{y,d} | D_{y,d-1}) = \sigma_{d-1}^2 \cdot D_{y,d-1}
\]

Note that \( \lambda \) is an "age-to-age" factor

\[
E(D_{y,d} | D_{y,d-1}) = \lambda_{d-1} \cdot D_{y,d-1}
\]

\[
E(D_{y,d} | D_{y,d-2}) = \lambda_{d-2} \cdot D_{y,d-2}
\]

\[
E(D_{y,d} | D_{y,d-3}) = \lambda_{d-3} \cdot D_{y,d-3}
\]
Mack Model

Variance can also be stated recursively

\[
\text{Var}(D_{y,d} | D_{y,d-1}) = \sigma_{d-1}^2 \cdot D_{y,d-1}
\]
\[
\text{Var}(D_{y,d} | D_{y,d-2}) = \sigma_{d-1}^2 \cdot D_{y,d-1} + \text{Var}(\lambda_{d-1} \cdot D_{y,d-1} | D_{y,d-2})
\]
\[
= \sigma_{d-1}^2 D_{y,d-1} + \lambda_{d-1}^2 \cdot \sigma_{d-2}^2 \cdot D_{y,d-2}
\]

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

\[
\hat{\sigma}_{d-1}^2 \approx \frac{1}{(n-p)} \sum_{d} \left( \frac{D_{y,d} - \lambda_{d-1} \cdot D_{y,d-1}}{\lambda_{d-1} \cdot D_{y,d-1}} \right)
\]

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

Limitations & Caveats:
- 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
Introduction to Stochastic Reserving

An Unchanging World:
Assumes that the future payments will be from a distribution identical to the past.

Over-Parameterization:

<table>
<thead>
<tr>
<th>E&amp;V ODP</th>
<th>Mack</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 Years</td>
<td>10 Years</td>
</tr>
<tr>
<td>55 Points</td>
<td>45 Points</td>
</tr>
<tr>
<td>19 Parameters</td>
<td>9 Parameters</td>
</tr>
<tr>
<td>+1 “sigma” σ</td>
<td>+9 “sigmas” σd</td>
</tr>
</tbody>
</table>

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

References (Bootstrap):
Introduction to Stochastic Reserving

References (Mack):


References (Other):


### Comparison of Standard Errors for Two Reserving Models

**Completing the Triangle**

#### ODP - E&V

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

#### "Distribution Free" Chainladder - Mack

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

#### ODP - E&V

<table>
<thead>
<tr>
<th>AY</th>
<th>Diagonal</th>
<th>Next CY</th>
<th>Process</th>
<th>Parameter</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5,339,085</td>
<td>94,634</td>
<td>70,554</td>
<td>84,522</td>
<td>110,099</td>
</tr>
<tr>
<td>3</td>
<td>4,909,315</td>
<td>375,833</td>
<td>140,604</td>
<td>117,373</td>
<td>183,155</td>
</tr>
<tr>
<td>4</td>
<td>4,588,268</td>
<td>247,190</td>
<td>114,029</td>
<td>75,063</td>
<td>136,517</td>
</tr>
<tr>
<td>5</td>
<td>3,873,311</td>
<td>334,148</td>
<td>132,577</td>
<td>78,431</td>
<td>154,039</td>
</tr>
<tr>
<td>6</td>
<td>3,691,712</td>
<td>383,287</td>
<td>141,991</td>
<td>81,651</td>
<td>163,793</td>
</tr>
<tr>
<td>7</td>
<td>3,483,130</td>
<td>605,548</td>
<td>178,473</td>
<td>112,238</td>
<td>210,832</td>
</tr>
<tr>
<td>8</td>
<td>2,864,498</td>
<td>1,310,258</td>
<td>262,529</td>
<td>225,048</td>
<td>345,786</td>
</tr>
<tr>
<td>9</td>
<td>1,363,294</td>
<td>1,018,834</td>
<td>231,500</td>
<td>229,139</td>
<td>325,725</td>
</tr>
<tr>
<td>10</td>
<td>344,014</td>
<td>856,804</td>
<td>212,295</td>
<td>358,489</td>
<td>416,633</td>
</tr>
</tbody>
</table>

#### "Distribution Free" Chainladder - Mack

<table>
<thead>
<tr>
<th>AY</th>
<th>Diagonal</th>
<th>Next CY</th>
<th>Process</th>
<th>Parameter</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5,339,085</td>
<td>94,634</td>
<td>48,832</td>
<td>57,628</td>
<td>75,535</td>
</tr>
<tr>
<td>3</td>
<td>4,909,315</td>
<td>375,833</td>
<td>75,052</td>
<td>56,970</td>
<td>94,225</td>
</tr>
<tr>
<td>4</td>
<td>4,588,268</td>
<td>247,190</td>
<td>45,268</td>
<td>27,163</td>
<td>52,792</td>
</tr>
<tr>
<td>5</td>
<td>3,873,311</td>
<td>334,148</td>
<td>178,062</td>
<td>87,733</td>
<td>198,502</td>
</tr>
<tr>
<td>6</td>
<td>3,691,712</td>
<td>383,287</td>
<td>225,149</td>
<td>102,068</td>
<td>247,204</td>
</tr>
<tr>
<td>7</td>
<td>3,483,130</td>
<td>605,548</td>
<td>229,965</td>
<td>99,925</td>
<td>250,737</td>
</tr>
<tr>
<td>8</td>
<td>2,864,498</td>
<td>1,310,258</td>
<td>346,712</td>
<td>151,271</td>
<td>378,275</td>
</tr>
<tr>
<td>9</td>
<td>1,363,294</td>
<td>1,018,834</td>
<td>226,818</td>
<td>82,715</td>
<td>241,429</td>
</tr>
<tr>
<td>10</td>
<td>344,014</td>
<td>856,804</td>
<td>234,816</td>
<td>75,503</td>
<td>246,656</td>
</tr>
</tbody>
</table>

Total:

**ODP - E&V**: 30,456,627 5,226,536 524,331 10.0% 532,575 10.2% 747,368 14.3%

**"Distribution Free" Chainladder - Mack**: 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:

\[ E(D_{i,k} | D_{i,k-1}) = \lambda_{k-1} \cdot E(D_{i,k-1}) \]

- \( D_{i,k} \) = cumulative loss for AY \( i \) at development period \( k \)
- \( \lambda_{k-1} \) = 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.

\[
\text{Var}(D_{i,k} | D_{i,k-1}) = \sigma_{k-1}^2 \cdot E(D_{i,k-1})
\]

\[
= E(D_{i,k})^2 \cdot \left\{ \frac{\sigma_{k-1}^2}{\lambda_{k-1}^2 \cdot E(D_{i,k-1})} \right\}
\]

\[
\text{Var}(D_{i,k} | D_{i,k-2}) = \text{Var}(D_{i,k} | D_{i,k-1}) + \text{Var}(\lambda_{k-1} \cdot D_{i,k-1} | D_{i,k-2})
\]

\[
= \sigma_{k-1}^2 \cdot E(D_{i,k-1}) + \lambda_{k-1}^2 \cdot \sigma_{k-2}^2 \cdot E(D_{i,k-2})
\]

\[
= E(D_{i,k})^2 \cdot \left\{ \frac{\sigma_{k-1}^2}{\lambda_{k-1}^2 \cdot E(D_{i,k-1})} + \frac{\sigma_{k-2}^2}{\lambda_{k-2}^2 \cdot E(D_{i,k-2})} \right\}
\]

\[
\text{Var}(D_{i,k} | D_{i,k-3}) = \text{Var}(D_{i,k} | D_{i,k-1}) + \text{Var}(\lambda_{k-1} \cdot D_{i,k-1} | D_{i,k-2}) + \text{Var}(\lambda_{k-1} \cdot \lambda_{k-2} \cdot D_{i,k-2} | D_{i,k-3})
\]

\[
= \sigma_{k-1}^2 \cdot E(D_{i,k-1}) + \lambda_{k-1}^2 \cdot \sigma_{k-2}^2 \cdot E(D_{i,k-2}) + \lambda_{k-1}^2 \cdot \lambda_{k-2}^2 \cdot \sigma_{k-3}^2 \cdot E(D_{i,k-3})
\]

\[
= E(D_{i,k})^2 \cdot \left\{ \frac{\sigma_{k-1}^2}{\lambda_{k-1}^2 \cdot E(D_{i,k-1})} + \frac{\sigma_{k-2}^2}{\lambda_{k-2}^2 \cdot E(D_{i,k-2})} + \frac{\sigma_{k-3}^2}{\lambda_{k-3}^2 \cdot E(D_{i,k-3})} \right\}
\]

This expansion can continue for any number of development periods “\( n \).”

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

The reserve (considering ultimate = age \( N \)) is given as:

\[ E(R_i) = E(D_{i,N}) - D_{i,N+1-i} \]

\[
\text{Var}(R_i) = \text{Var}(D_{i,N} | D_{i,N+1-i}) = E(D_{i,N})^2 \cdot \left\{ \sum_{d=1}^{N-N-d} \frac{\hat{\sigma}_{N-d}^2}{\lambda_{N-d}^2 \cdot E(D_{i,N-d})} \right\}
\]
And then the mean squared error (MSE), including Parameter Variance is given as below:

\[
MSE(D_{i,N} \mid D_{i,N+1-i}) = E(D_{i,N})^2 \cdot \left\{ \rho_{\sigma_i^2} \lambda_{N-d}^2 \left\{ \frac{1}{E(D_{i,N})} \sum_{d=1}^{i-1} \frac{\sigma_{N-d}^2}{\lambda_{N-d}^2} \left( \frac{1}{\sum_{y=1}^{N-d-1} D_{y,N-d}} \right) \right\} \right\}
\]

Think of this as analogous to:

\[
\approx \hat{\sigma}^2 + \frac{\sigma^2}{n}
\]

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 \( LDF_k = \hat{\lambda}_k \cdot \hat{\lambda}_{k+1} \cdots \hat{\lambda}_{N-1} \), such that \( E(D_{i,N}) = LDF_k \cdot D_{i,k} \), then we can re-write the mean square error (MSE) expression as follows:

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

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:

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

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