Development Pattern and Prediction Error for the Stochastic Bornhuetter-Ferguson Claims Reserving Method

Annina Saluz, Alois Gisler, Mario V. Wüthrich

ETH Zurich

ASTIN Colloquium Madrid, June 2011
Overview

1. Notation

2. Bornhuetter-Ferguson Method (BF)

3. Normal Model

4. Estimation of the Parameters and their Correlations

5. Prediction Uncertainty

6. Conclusions and Remarks
Notation

- Accident years \( i, 0 \leq i \leq I \)

Table: Claims development triangle

<table>
<thead>
<tr>
<th>Accident year ( i )</th>
<th>Development year ( j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0  ( \ldots )  ( j )  ( \ldots )  ( J )</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>observations ( \mathcal{D}_I )</td>
</tr>
<tr>
<td>( I )</td>
<td>( \mathcal{D}^c_I )</td>
</tr>
</tbody>
</table>

\( \mathcal{D}_I \) observations
Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$

Table: Claims development triangle

<table>
<thead>
<tr>
<th>accident year $i$</th>
<th>development year $j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$0 \ldots j \ldots J$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>observations $\mathcal{D}_I$</td>
</tr>
<tr>
<td>$i$</td>
<td>$\mathcal{D}^c_I$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td></td>
</tr>
</tbody>
</table>

Goal: Predict $\mathcal{D}_c^I = \{ X_{i,j}; i + j > I, i \leq I, j \leq J \}$
Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$

**Table: Claims development triangle**

<table>
<thead>
<tr>
<th>accident year $i$</th>
<th>development year $j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 \ldots j \ldots J</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>$i$</td>
<td>observations $\mathcal{D}_I$</td>
</tr>
<tr>
<td>\vdots</td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td>$\mathcal{D}_I^c$</td>
</tr>
</tbody>
</table>
**Notation**

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$
- Cumulative claims $C_{i,j}$

**Table: Claims development triangle**

<table>
<thead>
<tr>
<th>accident year $i$</th>
<th>development year $j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 \ldots j \ldots J</td>
</tr>
<tr>
<td>\vdots</td>
<td>observations $\mathcal{D}_I$</td>
</tr>
<tr>
<td>$i$</td>
<td>$\mathcal{D}^c_I$</td>
</tr>
<tr>
<td>\vdots</td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td></td>
</tr>
</tbody>
</table>
Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$
- Cumulative claims $C_{i,j}$
- Observed claims at time $I$: $\mathcal{D}_I = \{X_{i,j}; i + j \leq I\}$

Table: Claims development triangle

<table>
<thead>
<tr>
<th>accident year $i$</th>
<th>development year $j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 \ldots \ j \ldots \ J</td>
</tr>
<tr>
<td>\vdots</td>
<td>observations $\mathcal{D}_I$</td>
</tr>
<tr>
<td>$i$</td>
<td>$\mathcal{D}_I^c$</td>
</tr>
<tr>
<td>\vdots</td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td></td>
</tr>
</tbody>
</table>
Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$
- Cumulative claims $C_{i,j}$
- Observed claims at time $I$: $\mathcal{D}_I = \{X_{i,j}; \ i + j \leq I\}$
- Goal: Predict $\mathcal{D}_I^c = \{X_{i,j}; \ i + j > I, i \leq I, j \leq J\}$

Table: Claims development triangle

<table>
<thead>
<tr>
<th>accident year $i$</th>
<th>development year $j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 ... $j$ ... $J$</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>$i$</td>
<td>observations $\mathcal{D}_I$</td>
</tr>
<tr>
<td>\vdots</td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td>$\mathcal{D}_I^c$</td>
</tr>
</tbody>
</table>
Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) $R_i = C_{i,J} - C_{i,I-i}$ at time $I$

$\hat{R}_i = \hat{\mu}_i (1 - \hat{\beta}_{I-i})$, where

- $\hat{\mu}_i$ a prior estimate for $\mu_i = E[C_{i,J}]$
Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) \( R_i = C_{i,J} - C_{i,I-i} \) at time \( I \)

\[
\hat{R}_i = \hat{\mu}_i (1 - \hat{\beta}_{I-i}),
\]

where

- \( \hat{\mu}_i \) a prior estimate for \( \mu_i = \mathbb{E}[C_{i,J}] \)
- \( 1 - \hat{\beta}_{I-i} \) estimated still to come factor at the end of development year \( I - i \),

\( \hat{\beta}_J = 1, \ (\hat{\beta}_j)_{j=0,...,J} \) estimated cumulative development pattern)
Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) $R_i = C_{i,J} - C_{i,I-1}$ at time $I$

$$\hat{R}_i = \hat{\mu}_i (1 - \hat{\beta}_{I-i}),$$

where

- $\hat{\mu}_i$ a prior estimate for $\mu_i = E[C_{i,J}]$
- $1 - \hat{\beta}_{I-i}$ estimated still to come factor at the end of development year $I - i$, $\hat{\beta}_J = 1$, $((\hat{\beta}_j)_{j=0,...,J}$ estimated cumulative development pattern)
- $\hat{\gamma}_0 = \hat{\beta}_0$, $\hat{\gamma}_j = \hat{\beta}_j - \hat{\beta}_{j-1}$, $1 \leq j \leq J$, (estimated incremental development pattern)
Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) $R_i = C_{i,J} - C_{i,I-i}$ at time $I$

$$\hat{R}_i = \hat{\mu}_i (1 - \hat{\beta}_{I-i}),$$

where

- $\hat{\mu}_i$ a prior estimate for $\mu_i = E[C_{i,J}]$
- $1 - \hat{\beta}_{I-i}$ estimated still to come factor at the end of development year $I - i$, $\hat{\beta}_J = 1$, $((\hat{\beta}_j)_{j=0,...,J}$ estimated cumulative development pattern)
- $\hat{\gamma}_0 = \hat{\beta}_0$, $\hat{\gamma}_j = \hat{\beta}_j - \hat{\beta}_{j-1}$, $1 \leq j \leq J$, (estimated incremental development pattern)
- Assumption of a cross-classified model $E[X_{i,j}] = \mu_i \gamma_j$
Motivation for the Estimation of the Pattern

- Often the chain ladder (CL) development pattern is used for $\hat{\gamma}_j$

$$\hat{\gamma}_j^{CL} = \prod_{k=j}^{J-1} \hat{f}_k^{-1} - \prod_{k=j-1}^{J-1} \hat{f}_k^{-1}, \quad \hat{f}_k = \frac{\sum_{i=0}^{I-k-1} C_{i,k+1}}{\sum_{i=0}^{I-k-1} C_{i,k}}.$$

**Issue:** In the BF method we are given a priori estimates of the $\mu_i$. The CL pattern ignores any prior information.

**Goal:** Incorporate the a priori estimates in the estimation of the development pattern.
Motivation for the Estimation of the Pattern

- Often the chain ladder (CL) development pattern is used for $\hat{\gamma}_j$

$$\hat{\gamma}^{CL}_j = \prod_{k=j}^{J-1} \hat{f}_k^{-1} - \prod_{k=j-1}^{J-1} \hat{f}_k^{-1}, \quad \hat{f}_k = \frac{\sum_{i=0}^{I-k-1} C_{i,k+1}}{\sum_{i=0}^{I-k-1} C_{i,k}}.$$  

Issue: In the BF method we are given a priori estimates of the $\mu_i$. The CL pattern ignores any prior information.

Goal: Incorporate the a priori estimates in the estimation of the development pattern.

- If the $\mu_i$ were known then the best linear unbiased estimate of $\gamma_j$ would be

$$\gamma_j^{(0)} = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \mu_i}, \quad 0 \leq j \leq J.$$  

⇒ A first candidate is the raw estimate

$$\hat{\gamma}_j^{(0)} = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \hat{\mu}_i},$$  

However, they do not sum up to 1.
Intuitive Estimation of the Pattern

- If the full rectangle was known an obvious estimate is given by

\[ \hat{\gamma}_j = \frac{\sum_{i=0}^{I} X_{i,j}}{\sum_{i=0}^{I} C_{i,J}}, \quad 0 \leq j \leq J. \]
Intuitive Estimation of the Pattern

- If the full rectangle was known an obvious estimate is given by

\[ \hat{\gamma}_j = \frac{\sum_{i=0}^{I} X_{i,j}}{\sum_{i=0}^{I} C_{i,J}}, \quad 0 \leq j \leq J. \]

- If only the upper trapezoid and the \( \mu_i \) are known, replace the unknown \( X_{i,j} \) by predictors \( \mu_i \hat{\gamma}_j \)

\[ \hat{\gamma}_j = \frac{\sum_{i=0}^{I-j} X_{i,j} + \sum_{i=I-j+1}^{I} \mu_i \hat{\gamma}_j}{\sum_{i=0}^{I-J} C_{i,J} + \sum_{i=I-J+1}^{I} \left( C_{i,I-i} + \sum_{l=I-i+1}^{J} \mu_i \hat{\gamma}_l \right)}, \quad 0 \leq j \leq J. \]
Intuitive Estimation of the Pattern

- If the full rectangle was known an obvious estimate is given by

\[ \hat{\gamma}_j = \frac{\sum_{i=0}^{I} X_{i,j}}{\sum_{i=0}^{I} C_{i,J}}, \quad 0 \leq j \leq J. \]

- If only the upper trapezoid and the \( \mu_i \) are known, replace the unknown \( X_{i,j} \) by predictors \( \mu_i \hat{\gamma}_j \)

\[ \hat{\gamma}_j = \frac{\sum_{i=0}^{I-j} X_{i,j} + \sum_{i=I-j+1}^{I} \mu_i \hat{\gamma}_j}{\sum_{i=0}^{I} C_{i,J} + \sum_{i=I-J+1}^{I} \left( C_{i,I-I_i} + \sum_{l=I-I_i+1}^{J} \mu_l \hat{\gamma}_l \right)}, \quad 0 \leq j \leq J. \]

- In an over-dispersed Poisson model the maximum likelihood estimates (MLE’s) satisfy these equations.
Model Assumptions (Normal Model)

N1 The $X_{i,j}$ are independent and normally distributed and there exist parameters $\mu_0, \mu_1, \ldots, \mu_I$ and $\gamma_0, \gamma_1, \ldots, \gamma_J$ with $\sum_{j=0}^{J} \gamma_j = 1$ and $\sigma^2_0, \ldots, \sigma^2_J$ such that

$$E[X_{i,j}] = \mu_i \gamma_j,$$
$$\text{Var}(X_{i,j}) = \mu_i \sigma^2_j,$$

where $\sigma^2_j > 0$, $0 \leq j \leq J$.

N2 The a priori estimates $\hat{\mu}_i$ for $\mu_i = E[C_{i,J}]$ are unbiased and independent of all $X_{l,j}$.
Maximum Likelihood Estimation (MLE) of the $\gamma_j$'s

- We calculate the MLE's assuming that the true $\mu_i$'s and $\sigma^2_j$'s are known and then replace the $\mu_i$'s by the a priori estimates $\hat{\mu}_i$ and the $\sigma^2_j$'s by estimates $\hat{\sigma}^2_j$. 

\[
\hat{\gamma}_j = \frac{\sum_{i} X_{i,j} - \hat{\mu}_i}{\sum_{i} \tilde{I}_{-j,i}} + \hat{\sigma}^2_j \left( \frac{\sum_{i} X_{i,j} \hat{\mu}_i - \sum_{i} \tilde{I}_{-j,i} X_{i,j} \hat{\mu}_i}{\sum_{i} \tilde{I}_{-j,i}} \right) \]

where $\hat{\sigma}^2_j = \frac{1}{I_j - j} \sum_{i = 0} \hat{\mu}_i \, \left( \sum_{i} X_{i,j} \hat{\mu}_i - \sum_{i} \tilde{I}_{-j,i} X_{i,j} \hat{\mu}_i \right)^2$, $0 \leq j \leq J$, $j \neq I$. 

A. Saluz, A. Gisler, M. V. Wüthrich (ETH Zurich) 
June, 2011 8 / 12
Maximum Likelihood Estimation (MLE) of the $\gamma_j$’s

- We calculate the MLE’s assuming that the true $\mu_i$’s and $\sigma_j^2$’s are known and then replace the $\mu_i$’s by the a priori estimates $\hat{\mu}_i$ and the $\sigma_j^2$’s by estimates $\hat{\sigma}_j^2$.

- In the Normal Model we obtain

$$
\hat{\gamma}_j = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \hat{\mu}_i} + \frac{\hat{\sigma}_j^2}{\sum_{i=0}^{I-j} \hat{\mu}_i} \left( 1 - \sum_{l=0}^{J} \left( \frac{\hat{\sigma}_l^2}{\sum_{i=0}^{I-l} \hat{\mu}_i} \right) \right),
$$

where

$$
\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=0}^{I-j} \hat{\mu}_i \left( \frac{X_{i,j}}{\hat{\mu}_i} - \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \hat{\mu}_i} \right)^2, \quad 0 \leq j \leq J, \ j \neq I.
$$
Covariance matrix of the $\hat{\gamma}_j$’s

We use the asymptotic properties of MLE’s (Fisher Information matrix) to estimate the covariance matrix of the $\hat{\gamma}_j$’s:

$$\widehat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_k) = \frac{\hat{\sigma}^2_j}{\sum_{i=0}^{I-j} \hat{\mu}_i} \left( 1_{\{j=k\}} - \frac{\left( \hat{\sigma}^2_k / \sum_{i=0}^{I-k} \hat{\mu}_i \right)}{\sum_{l=0}^{J} \left( \hat{\sigma}^2_l / \sum_{i=0}^{I-l} \hat{\mu}_i \right)} \right).$$

Remarks:

- For the best linear unbiased estimate $\gamma_j^{(0)}$ (in the case of known $\mu_i$) we have

$$\text{Cov}(\gamma_j^{(0)}, \gamma_k^{(0)}) = 1_{\{j=k\}} \frac{\sigma^2_j}{\sum_{i=0}^{I-j} \mu_i},$$

(compare with first summand).

- Because of the side constraint $\sum_{j=0}^{J} \gamma_j = 1$ the off-diagonal correlations must be negative (compare with second summand).
Prediction Uncertainty

Given all information $\mathcal{I}_I$ (i.e. $\mathcal{D}_I$ and all $\hat{\mu}_i$), the conditional MSEP of the predictor $\hat{R}_i = \hat{\mu}_i(1 - \hat{\beta}_{I-i})$ is given by

$$\text{mse}_{R_i|\mathcal{I}_I}(\hat{R}_i) = E \left[ \left( R_i - \hat{R}_i \right)^2 \middle| \mathcal{I}_I \right]$$

$$= E \left[ \left( \sum_{j=I-i+1}^J X_{i,j} - \hat{\mu}_i(1 - \hat{\beta}_{I-i}) \right)^2 \middle| \mathcal{I}_I \right]$$

$$= \sum_{j=I-i+1}^J \text{Var}(X_{i,j}) + \left( \hat{\mu}_i(1 - \hat{\beta}_{I-i}) - \mu_i(1 - \beta_{I-i}) \right)^2.$$

The term inside the summation represents the process variance ($PV_i$), and the squared term represents the estimation error ($EE_i$).
We estimate the conditional MSEP given $\mathcal{I}_I$ as follows

$$
\hat{\text{msep}}_{R_i|\mathcal{I}_I}(\hat{R}_i) = \sum_{j=I-i+1}^{J} \hat{\text{Var}}(X_{i,j}) + \hat{\text{Var}}(\hat{\mu}_i)(1 - \hat{\beta}_{I-i})^2
$$

$$
+ \hat{\mu}_i^2 \left( \sum_{j=0}^{I-i} \hat{\text{Var}}(\hat{\gamma}_j) + 2 \sum_{0 \leq j < k \leq I-i} \hat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_k) \right),
$$

$$
\hat{\text{msep}}_{\sum_{i=I-J+1}^{I} R_i|\mathcal{I}_I} \left( \sum_{i=I-J+1}^{I} \hat{R}_i \right) = \sum_{i=I-J+1}^{I} \hat{\text{msep}}_{R_i|\mathcal{I}_I}(\hat{R}_i)
$$

$$
+ 2 \sum_{I-J+1 \leq i < k \leq I} \left( (1 - \hat{\beta}_{I-i})(1 - \hat{\beta}_{I-k}) \hat{\text{Cov}}(\hat{\mu}_i, \hat{\mu}_k)
$$

$$
+ \hat{\mu}_i \hat{\mu}_k \sum_{j=0}^{I-i} \sum_{l=0}^{I-k} \hat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_l) \right).
$$
Conclusions and Remarks

Under distributional assumptions we have derived

- estimates for the development pattern taking all relevant information into account
- formulas for the smoothing from the raw estimates $\hat{\gamma}_j^{(0)}$ to the final estimates $\hat{\gamma}_j$
- estimates for the correlations of these estimates.

We recommend to use these formulas also in the distribution-free case (currently there are no estimators available from which we know that they perform better).