Bootstrap Methods in Reserving

Michal Pešta

Charles University in Prague & Generali PPF Holding

Actuarial Seminar, Prague
December 9, 2011
Outline

1 Motivation
   • Origins
     • Prologue for Bootstrap in Statistics
     • Reserving Issue

2 Mathematical Background
   • Bootstrap
   • Stochastics in Insurance
   • Bootstrapping the Chain Ladder
   • Generalized Linear Models

3 Data Analysis
   • Estimation of Distribution

4 Conclusions
   • Discussion
“to pull oneself up by one’s bootstrap”

The Surprising Adventures of Baron Münchausen

recounted in 1785 by Rudolf Erich Raspe

[in Czech: Baron Prášil]

- pulls himself out of a swamp by his pigtail
- the phrase appears to have originated in the early 19th century United States in the sense “pull oneself over a fence by one’s bootstraps” — being an absurdly impossible feat
- the Baron does not, in fact, pull himself out by his bootstraps
“to pull oneself up by one’s bootstrap”

The Surprising Adventures of Baron Münchausen recounted in 1785 by Rudolf Erich Raspe

[in Czech: Baron Prášil]

- pulls himself out of a swamp by his pigtail
- the phrase appears to have originated in the early 19th century United States in the sense “pull oneself over a fence by one’s bootstraps” being an absurdly impossible feat
- the Baron does not, in fact, pull himself out by his bootstraps
Outline

1 Motivation
   • Origins
   • Prologue for Bootstrap in Statistics
   • Reserving Issue

2 Mathematical Background
   • Bootstrap
   • Stochastics in Insurance
   • Bootstrapping the Chain Ladder
   • Generalized Linear Models

3 Data Analysis
   • Estimation of Distribution

4 Conclusions
   • Discussion
What is bootstrapping?

- **computationally intensive** method popularized in 1980s due to the introduction of computers in statistical practice
- a strong mathematical background $\Rightarrow$ bootstrap does not replace or add to the original data
- unfortunately, the name “bootstrap” conveys the impression of “something for nothing” $\Rightarrow$ idly resampling from their samples
What is bootstrapping?

- Computationally intensive method popularized in 1980s due to the introduction of computers in statistical practice.
- A strong mathematical background implies that bootstrap does not replace or add to the original data.
- Unfortunately, the name “bootstrap” conveys the impression of “something for nothing” implying idly resampling from their samples.
What is bootstrapping?

- computationally intensive method popularized in 1980s due to the introduction of computers in statistical practice
- a strong mathematical background → bootstrap does not replace or add to the original data
- unfortunately, the name “bootstrap” conveys the impression of “something for nothing” → idly resampling from their samples
Outline

1 Motivation
   • Origins
   • Prologue for Bootstrap in Statistics
   • Reserving Issue

2 Mathematical Background
   • Bootstrap
   • Stochastics in Insurance
   • Bootstrapping the Chain Ladder
   • Generalized Linear Models

3 Data Analysis
   • Estimation of Distribution

4 Conclusions
   • Discussion
Do we have a problem in reserving?

- consider traditional actuarial approach to reserving risk ... the uncertainty in the outcomes over the lifetime of the liabilities
- bootstrap can be also applied under Solvency II ... outstanding liabilities after 1 year
- distribution-free methods (e.g., chain ladder) only provide a standard deviation of the ultimates/reserves (or claims development result/run-off result)

? another risk measure (e.g., VaR @ 99.5%)
≠ moreover, distributions of ultimate cost of claims and the associated cash flows (not just a standard deviation)?
! claims reserving technique applied mechanically and without judgement
Do we have a problem in reserving?

- Consider traditional actuarial approach to reserving risk... the uncertainty in the outcomes over the lifetime of the liabilities.
- Bootstrap can be also applied under Solvency II... outstanding liabilities after 1 year.
- Distribution-free methods (e.g., chain ladder) only provide a standard deviation of the ultimates/reserves (or claims development result/run-off result).
- Another risk measure (e.g., VaR @ 99.5%).
- Moreover, distributions of ultimate cost of claims and the associated cash flows (not just a standard deviation)?
- Claims reserving technique applied mechanically and without judgement.
Do we have a problem in reserving?

- consider traditional actuarial approach to reserving risk... the uncertainty in the outcomes over the lifetime of the liabilities
- bootstrap can be also applied under Solvency II... outstanding liabilities after 1 year
- distribution-free methods (e.g., chain ladder) only provide a standard deviation of the ultimates/reserves (or claims development result/run-off result)

- another risk measure (e.g., VaR @ 99.5%)
- moreover, distributions of ultimate cost of claims and the associated cash flows (not just a standard deviation)?
- claims reserving technique applied mechanically and without judgement
Outline

1 Motivation
   - Origins
   - Prologue for Bootstrap in Statistics
   - Reserving Issue

2 Mathematical Background
   - Bootstrap
   - Stochastics in Insurance
   - Bootstrapping the Chain Ladder
   - Generalized Linear Models

3 Data Analysis
   - Estimation of Distribution

4 Conclusions
   - Discussion
Bootstrap

- simple (distribution-independent) resampling method
- estimate properties (distribution) of an estimator by sampling from an approximating (e.g., empirical) distribution
- useful when the theoretical distribution of a statistic of interest is complicated or unknown
random sampling with replacement from the original dataset \( \sim \) for \( b = 1, \ldots, B \) resample from \( X_1, \ldots, X_n \) with replacement and obtain \( X_{1,b}^{*}, \ldots, X_{n,b}^{*} \)

Case sampling

- input data (\# of catastrophic claims per year in 10y history):
  
  35, 34, 13, 33, 27, 30, 19, 31, 10, 33 \( \sim \) mean = 26.5, sd = 9.168182

- bootstrap sample 1 (1st draw with replacement):
  
  30, 27, 35, 35, 13, 35, 33, 34, 35, 33 \( \sim \) mean\(^*\) = 31.0, sd\(^*\) = 6.847546

- bootstrap sample 1000 (1000th draw with replacement):
  
  19, 19, 31, 19, 33, 34, 31, 34, 34, 10 \( \sim \) mean\(^*\)\(_{1000}\) = 26.4, sd\(^*\)\(_{1000}\) = 8.771165

- mean\(^*\), \ldots, mean\(^*\)\(_{1000}\) provide bootstrap empirical distribution for mean and sd\(^*\), \ldots, sd\(^*\)\(_{1000}\) provide bootstrap empirical distribution for sd (REALLY!?)
Bootstrap example

- random sampling with replacement from the original dataset $\sim$ for $b = 1, \ldots, B$ resample from $X_1, \ldots, X_n$ with replacement and obtain $X_{1,b}^*, \ldots, X_{n,b}^*$

Case sampling

- input data (# of catastrophic claims per year in 10y history):
  35, 34, 13, 33, 27, 30, 19, 31, 10, 33 $\sim$ mean = 26.5, $sd = 9.168182$

- bootstrap sample 1 (1st draw with replacement):
  30, 27, 35, 35, 13, 35, 33, 34, 35, 33 $\sim$ mean$_1^* = 31.0$, $sd_1^* = 6.847546$

  ...

- bootstrap sample 1000 (1000th draw with replacement):
  19, 19, 31, 19, 33, 34, 31, 34, 34, 10 $\sim$ mean$_{1000}^* = 26.4$, $sd_{1000}^* = 8.771165$

- mean$_1^*, \ldots, mean_{1000}^*$ provide bootstrap empirical distribution for mean and
  $sd_1^*, \ldots, sd_{1000}^*$ provide bootstrap empirical distribution for $sd$ (REALLY!?)
Bootstrap example

- random sampling with replacement from the original dataset \( \sim \) for \( b = 1, \ldots, B \) resample from \( X_1, \ldots, X_n \) with replacement and obtain \( X_{1, b}^*, \ldots, X_{n, b}^* \)

Case sampling

- input data (# of catastrophic claims per year in 10y history): 35, 34, 13, 33, 27, 30, 19, 31, 10, 33 \( \sim \) \( \text{mean} = 26.5, \text{sd} = 9.168182 \)
- bootstrap sample 1 (1st draw with replacement): 30, 27, 35, 35, 13, 35, 33, 34, 35, 33 \( \sim \) \( \text{mean}_1^* = 31.0, \text{sd}_1^* = 6.847546 \)
- bootstrap sample 1000 (1000th draw with replacement): 19, 19, 31, 19, 33, 34, 31, 34, 34, 10 \( \sim \) \( \text{mean}_{1000}^* = 26.4, \text{sd}_{1000}^* = 8.771165 \)
- \( \text{mean}_1^*, \ldots, \text{mean}_{1000}^* \) provide bootstrap empirical distribution for \( \text{mean} \) and \( \text{sd}_1^*, \ldots, \text{sd}_{1000}^* \) provide bootstrap empirical distribution for \( \text{sd} \) (REALLY!?)
Bootstrap example

- random sampling with replacement from the original dataset \( \sim \) for \( b = 1, \ldots, B \) resample from \( X_1, \ldots, X_n \) with replacement and obtain \( X_{1,b}^*, \ldots, X_{n,b}^* \)

Case sampling

- input data (# of catastrophic claims per year in 10y history):
  35, 34, 13, 33, 27, 30, 19, 31, 10, 33 \( \sim \) mean = 26.5, sd = 9.168182
- bootstrap sample 1 (1st draw with replacement):
  30, 27, 35, 35, 13, 35, 33, 34, 35, 33 \( \sim \) mean\(^*_1\) = 31.0, sd\(^*_1\) = 6.847546
  
  ...

- bootstrap sample 1000 (1000th draw with replacement):
  19, 19, 31, 19, 33, 34, 31, 34, 34, 10 \( \sim \) mean\(^*_1000\) = 26.4, sd\(^*_1000\) = 8.771165

- mean\(^*_1\), \ldots, mean\(^*_1000\) provide bootstrap empirical distribution for mean and sd\(^*_1\), \ldots, sd\(^*_1000\) provide bootstrap empirical distribution for sd (REALLY!?)
Outline

1 Motivation
   - Origins
   - Prologue for Bootstrap in Statistics
   - Reserving Issue

2 Mathematical Background
   - Bootstrap
   - Stochastics in Insurance
     - Bootstrapping the Chain Ladder
     - Generalized Linear Models

3 Data Analysis
   - Estimation of Distribution

4 Conclusions
   - Discussion
Stochastics in insurance

- **deterministic** methods $\rightsquigarrow$ reserve estimate (are reasonable?)
- **stochastic** methods (statistical assumptions) $\rightsquigarrow$ prediction of variability (how precise?)
- simulations (resampling methods) $\rightsquigarrow$ predictive distribution
Outline

1 Motivation
   • Origins
   • Prologue for Bootstrap in Statistics
   • Reserving Issue

2 Mathematical Background
   • Bootstrap
   • Stochastics in Insurance
   • Bootstrapping the Chain Ladder
   • Generalized Linear Models

3 Data Analysis
   • Estimation of Distribution

4 Conclusions
   • Discussion
Mack’s chain ladder (no tail factor)

- $C_{ij} \ldots$ cumulative claims in origin year $i$ and development year $j$

**Assumptions**

1. $\mathbb{E}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = f_j C_{i,j}, \quad 1 \leq i \leq n, \ 1 \leq j \leq n-1$
2. $\text{Var}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = \sigma_j^2 C_{i,j}, \quad 1 \leq i \leq n, \ 1 \leq j \leq n-1$
3. accident years $[C_{i,1}, \ldots, C_{i,n}], \quad 1 \leq i \leq n$ are independent

- $C_{i,n} \ldots$ ultimate claims amount
- $R_i = C_{i,n} - C_{i,n+1-i} \ldots$ outstanding claims reserve

reasonable estimate $f_j$ for development factors (is unbiased, but consistent?) estimate conditional s.e. of estimates of ultimates and reserves

$\mathbb{E}[\left(C_{i,n} - C_{i,n+1}|C_{i,j}, 1 \leq i \leq n+1\right)] = \mathbb{E}[R_i - R_i^*|C_{i,j}, 1 \leq i \leq n+1]$ given data?
Mack’s chain ladder (no tail factor)

- \( C_{ij} \) ... cumulative claims in origin year \( i \) and development year \( j \)

### Assumptions

1. \( \mathbb{E}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = f_j C_{i,j}, \quad 1 \leq i \leq n, \ 1 \leq j \leq n - 1 \)
2. \( \text{Var}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = \sigma_j^2 C_{i,j}, \quad 1 \leq i \leq n, \ 1 \leq j \leq n - 1 \)
3. accident years \( [C_{i,1}, \ldots, C_{i,n}], \quad 1 \leq i \leq n \) are independent

- \( C_{i,n} \) ... ultimate claims amount
- \( R_i = C_{i,n} - C_{i,n+1-i} \) ... outstanding claims reserve

### Main Goals

- [a] reasonable estimate \( \hat{f}_j \) for development factors (is unbiased, but consistent?)
- [b] estimate conditional s.e. of estimates of ultimates and reserves
- \( \mathbb{E}[(\hat{C}_{i,n} - C_{i,n})^2|\{C_{i,j} : i+j \leq n+1\}] = \mathbb{E}[(\hat{R}_i - R_i)^2|\{C_{i,j} : i+j \leq n+1\}] \)
- [c] conditional distribution of reserves given data?
Mack’s chain ladder (no tail factor)

- $C_{ij}$ ... cumulative claims in origin year $i$ and development year $j$

### Assumptions

1. $\mathbb{E}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = f_j C_{i,j}, \quad 1 \leq i \leq n, \ 1 \leq j \leq n - 1$
2. $\text{Var}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = \sigma_j^2 C_{i,j}, \quad 1 \leq i \leq n, \ 1 \leq j \leq n - 1$
3. accident years $[C_{i,1}, \ldots, C_{i,n}], \quad 1 \leq i \leq n$ are independent

- $C_{i,n}$ ... ultimate claims amount
- $R_i = C_{i,n} - C_{i,n+1-i}$ ... outstanding claims reserve

### Main Goals

[a] reasonable estimate $\hat{f}_j$ for development factors (is unbiased, but consistent?)

[b] estimate conditional s.e. of estimates of ultimates and reserves

\[
\mathbb{E}[(\hat{C}_{i,n} - C_{i,n})^2|\{C_{i,j} : i+j \leq n+1\}] = \mathbb{E}[(\hat{R}_i - R_i)^2|\{C_{i,j} : i+j \leq n+1\}]
\]

[c] conditional distribution of reserves given data?
Mack’s chain ladder (no tail factor)

- $C_{ij}$ ... cumulative claims in origin year $i$ and development year $j$

Assumptions

1. $\mathbb{E}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = f_j C_{i,j}, \ 1 \leq i \leq n, 1 \leq j \leq n - 1$
2. $\text{Var}[C_{i,j+1}|C_{i,1}, \ldots, C_{i,j}] = \sigma_j^2 C_{i,j}, \ 1 \leq i \leq n, 1 \leq j \leq n - 1$
3. accident years $[C_{i,1}, \ldots, C_{i,n}], \ 1 \leq i \leq n$ are independent

- $C_{i,n}$ ... ultimate claims amount
- $R_i = C_{i,n} - C_{i,n+1-i}$ ... outstanding claims reserve

Main Goals

[a] reasonable estimate $\hat{f}_j$ for development factors (is unbiased, but consistent?)

[b] estimate conditional s.e. of estimates of ultimates and reserves

$$\mathbb{E}[(\hat{C}_{i,n} - C_{i,n})^2|\{C_{i,j} : i+j \leq n+1\}] = \mathbb{E}[(\hat{R}_i - R_i)^2|\{C_{i,j} : i+j \leq n+1\}]$$

[c] conditional distribution of reserves given data?
Algorithm 1 (Part I)

[1] estimate development factors

\[ \hat{f}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}, \quad 1 \leq j \leq n - 1; \quad \hat{f}_n \equiv 1 \quad \text{(no tail)} \]

[2] fit chain ladder to the original data and predict bottom-right triangle

\[ \hat{C}_{i,j} = C_{i,n+1-i} \times \hat{f}_{n+1-i} \times \ldots \times \hat{f}_{j-1}, \quad i + j \geq n + 2 \]

[3] back-fit observed original claims from diagonals \( C_{i,n+1-i} \)

\[ \hat{C}_{i,n+1-i} \equiv C_{i,n+1-i}; \quad \hat{C}_{i,j} = \frac{C_{i,n+1-i}}{\hat{f}_{n-i} \times \ldots \times \hat{f}_j}, \quad i + j \leq n \]
Algorithm 1 (Part I)

[1] estimate development factors

\[ \hat{f}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}, \quad 1 \leq j \leq n - 1; \quad \hat{f}_n \equiv 1 \quad \text{(no tail)} \]

[2] fit chain ladder to the original data and predict bottom-right triangle

\[ \hat{C}_{i,j} = C_{i,n+1-i} \times \hat{f}_{n+1-i} \times \ldots \times \hat{f}_{j-1}, \quad i + j \geq n + 2 \]

[3] back-fit observed original claims from diagonals \( C_{i,n+1-i} \)

\[ \hat{C}_{i,n+1-i} \equiv C_{i,n+1-i}; \quad \hat{C}_{i,j} = \frac{C_{i,n+1-i}}{\hat{f}_{n-i} \times \ldots \times \hat{f}_{j}}, \quad i + j \leq n \]
Algorithm 1 (Part I)

[1] estimate development factors

$$\hat{f}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}, \quad 1 \leq j \leq n - 1; \quad \hat{f}_n \equiv 1 \quad \text{(no tail)}$$

[2] fit chain ladder to the original data and predict bottom-right triangle

$$\hat{C}_{i,j} = C_{i,n+1-i} \times \hat{f}_{n+1-i} \times \ldots \times \hat{f}_{j-1}, \quad i + j \geq n + 2$$

[3] back-fit observed original claims from diagonals $C_{i,n+1-i}$

$$\hat{C}_{i,n+1-i} \equiv C_{i,n+1-i}; \quad \hat{C}_{i,j} = \frac{C_{i,n+1-i}}{\hat{f}_{n-i} \times \ldots \times \hat{f}_j}, \quad i + j \leq n$$
Algorithm 1 (Part II)

[4] calculate unscaled Pearson residuals \( (C_{i,0} = \hat{C}_{i,0} \equiv 0) \)

\[
\begin{align*}
  r_{i,j} &= \frac{(C_{i,j} - C_{i,j-1}) - (\hat{C}_{i,j} - \hat{C}_{i,j-1})}{\sqrt{\hat{C}_{i,j} - \hat{C}_{i,j-1}}}, \quad i + j \leq n + 1 \\
\end{align*}
\]

▶ [1]–[4] are just Mack chain ladder

[5] resample residuals \( \{r_{i,j}\} \) \( B \)-times with replacement \( \rightsquigarrow \) \( B \) triangles of bootstrapped residuals \( \{(b)r_{i,j}^*\}, 1 \leq b \leq B \)

[6] construct \( B \) incremental bootstrap triangles

\[
(b)X_{i,j}^* = (b)r_{i,j}^*\sqrt{\hat{C}_{i,j} - \hat{C}_{i,j-1} + \hat{C}_{i,j} - \hat{C}_{i,j-1}}, \quad i + j \leq n + 1
\]
Algorithm 1 (Part II)

[4] calculate unscaled Pearson residuals \((C_{i,0} = \hat{C}_{i,0} \equiv 0)\)

\[
   r_{i,j} = \frac{(C_{i,j} - C_{i,j-1}) - (\hat{C}_{i,j} - \hat{C}_{i,j-1})}{\sqrt{\hat{C}_{i,j} - \hat{C}_{i,j-1}}}, \quad i + j \leq n + 1
\]

[1]–[4] are just Mack chain ladder

[5] resample residuals \(\{r_{i,j}\}\) \(B\)-times with replacement \(\sim B\) triangles of bootstrapped residuals \(\{(b)r_{i,j}^*\}, 1 \leq b \leq B\)

[6] construct \(B\) incremental bootstrap triangle

\[
   (b)X_{i,j}^* = (b)r_{i,j}^*\sqrt{\hat{C}_{i,j} - \hat{C}_{i,j-1} + \hat{C}_{i,j} - \hat{C}_{i,j-1}}, \quad i + j \leq n + 1
\]
Algorithm 1 (Part II)

[4] calculate unscaled Pearson residuals \((C_{i,0} = \hat{C}_{i,0} \equiv 0)\)

\[
    r_{i,j} = \frac{(C_{i,j} - C_{i,j-1}) - (\hat{C}_{i,j} - \hat{C}_{i,j-1})}{\sqrt{\hat{C}_{i,j} - \hat{C}_{i,j-1}}}, \quad i + j \leq n + 1
\]

- [1]–[4] are just Mack chain ladder

[5] resample residuals \(\{r_{i,j}\}\) \(B\)-times with replacement \(\sim B\) triangles of bootstrapped residuals \(\{(b)r_{i,j}^*\}\), \(1 \leq b \leq B\)

[6] construct \(B\) incremental bootstrap triangles

\[
    (b)X_{i,j}^* = (b)r_{i,j}^*\sqrt{\hat{C}_{i,j} - \hat{C}_{i,j-1} + \hat{C}_{i,j} - \hat{C}_{i,j-1}}, \quad i + j \leq n + 1
\]
Algorithm 1 (Part III)

[7] $B$ cumulative bootstrap triangles \((b) C_{i,0}^* \equiv 0\)

\[
(b) C_{i,j}^* = (b) X_{i,j}^* + (b) C_{i,j-1}^*, \quad i + j \leq n + 1
\]

[8] perform chain ladder on each bootstrap cumulative triangle \(\rightsquigarrow\) reserves \(\left\{ (b) R_i^* \right\}_{i=1}^n, 1 \leq b \leq B \)

- [5]–[8] is a bootstrap loop (repeated \(B\)-times)

[9] empirical distribution of size \(B\) for the reserves \(\rightsquigarrow\) empirical (estimated) mean, s.e., quantiles, ...
Algorithm 1 (Part III)

[7] $B$ cumulative bootstrap triangles \( (b) C_{i,0}^* \equiv 0 \)

\[
(b) C_{i,j}^* = (b) X_{i,j}^* + (b) C_{i,j-1}^*, \quad i + j \leq n + 1
\]

[8] perform chain ladder on each bootstrap cumulative triangle \( \leadsto \) reserves \( \left\{ (b) R_i^* \right\}_{i=1}^{n}, 1 \leq b \leq B \)

\[ \Rightarrow \text{[5]–[8] is a bootstrap loop (repeated $B$-times)} \]

[9] empirical distribution of size $B$ for the reserves \( \leadsto \) empirical (estimated) mean, s.e., quantiles, ...
Algorithm 1 (Part III)

[7] $B$ cumulative bootstrap triangles $(b) C^{*}_{i,0} \equiv 0$

$$(b) C^{*}_{i,j} = (b) X^{*}_{i,j} + (b) C^{*}_{i,j-1}, \quad i + j \leq n + 1$$

[8] perform chain ladder on each bootstrap cumulative triangle $\leadsto$ reserves $\{ (b) R^{*}_{i} \}^{n}_{i=1}, \quad 1 \leq b \leq B$

[5]–[8] is a bootstrap loop (repeated $B$-times)

[9] empirical distribution of size $B$ for the reserves $\leadsto$ empirical (estimated) mean, s.e., quantiles, ...
Outline

1 Motivation
   - Origins
   - Prologue for Bootstrap in Statistics
   - Reserving Issue

2 Mathematical Background
   - Bootstrap
   - Stochastics in Insurance
   - Bootstrapping the Chain Ladder
   - Generalized Linear Models

3 Data Analysis
   - Estimation of Distribution

4 Conclusions
   - Discussion
Generalized Linear Models (GLM) I

- a flexible generalization of ordinary linear regression
- formulated by John Nelder and Robert Wedderburn as a way of unifying various other statistical models, including linear regression, logistic regression and Poisson regression
- GLM consists of three elements:
  1. outcome of the dependent variables $Y$ from a particular distribution in the overdispersed exponential family, i.e.,

$$f_Y(y; \theta, \tau) = h(y, \tau) \exp \left\{ \frac{b(\theta)^\top T(y) - A(\theta)}{d(\tau)} \right\}$$

where $\tau$ is dispersion parameter
  2. linear predictor (mean structure)

$$\eta = X\beta$$

  3. link function $g$ (element-wise)

$$\mathbb{E}Y = \mu = g^{-1}(\eta)$$
Generalized Linear Models (GLM) I

- a flexible generalization of ordinary linear regression
- formulated by John Nelder and Robert Wedderburn as a way of unifying various other statistical models, including linear regression, logistic regression and Poisson regression
- GLM consists of three elements:
  1. outcome of the dependent variables $Y$ from a particular distribution in the overdispersed exponential family, i.e.,

$$f_Y(y; \theta, \tau) = h(y, \tau) \exp \left\{ \frac{b(\theta)^\top T(y) - A(\theta)}{d(\tau)} \right\}$$

  where $\tau$ is dispersion parameter
  2. linear predictor (mean structure)

$$\eta = X\beta$$

  3. link function $g$ (element-wise)

$$EY = \mu = g^{-1}(\eta)$$
a flexible generalization of ordinary linear regression
formulated by John Nelder and Robert Wedderburn as a way of unifying various other statistical models, including linear regression, logistic regression and Poisson regression

GLM consists of three elements:

1. outcome of the dependent variables $Y$ from a particular distribution in the overdispersed exponential family, i.e.,

$$f_Y(y; \theta, \tau) = h(y, \tau) \exp \left\{ \frac{b(\theta)^\top T(y) - A(\theta)}{d(\tau)} \right\}$$

where $\tau$ is dispersion parameter

2. linear predictor (mean structure)

$$\eta = X\beta$$

3. link function $g$ (element-wise)

$$EY = \mu = g^{-1}(\eta)$$
normal, exponential, gamma, chi-squared, beta, Weibull (with known shape parameter), Dirichlet, Bernoulli, binomial, multinomial, Poisson, negative binomial (with known stopping-time parameter), and geometric distributions are all exponential families

- family of Pareto distributions with a fixed minimum bound form an exponential family

- Cauchy and uniform families of distributions are not exponential families

- Laplace family is not an exponential family unless the mean is zero
Generalized Linear Models (GLM) II

- overdispersed exponential family

$$\mathbb{E}(Y) = \mu = g^{-1}(X\beta) \quad \text{and} \quad \text{Var}(Y) = V(\mu) = V(g^{-1}(X\beta))d(\tau)$$

- distribution $\leftrightarrow$ link function (element-wise)
  - normal \ldots identity: $\mu = X\beta$
  - gamma (exponential) \ldots inverse: $(\mu)^{-1} = X\beta$
  - Poisson \ldots logarithm: $\log(\mu) = X\beta$
  - binomial (multinomial) \ldots logit: $\log \left( \frac{\mu}{1-\mu} \right) = X\beta$

- estimation of the parameters via maximum likelihood, quasi-likelihood or Bayesian techniques
Generalized Linear Models (GLM) II

- overdispersed exponential family

\[ \mathbb{E}(Y) = \mu = g^{-1}(X\beta) \quad \text{and} \quad \text{Var}(Y) = V(\mu) = V(g^{-1}(X\beta))d(\tau) \]

- distribution \( \leftrightarrow \) link function (element-wise)
  - normal \ldots identity: \( \mu = X\beta \)
  - gamma (exponential) \ldots inverse: \( (\mu)^{-1} = X\beta \)
  - Poisson \ldots logarithm: \( \log(\mu) = X\beta \)
  - binomial (multinomial) \ldots logit: \( \log\left(\frac{\mu}{1-\mu}\right) = X\beta \)

- estimation of the parameters via maximum likelihood, quasi-likelihood or Bayesian techniques
Mack’s model as GLM

- reformulate Mack’s model as a model of ratios

\[
\mathbb{E} \left[ \frac{C_{i,j+1}}{C_{i,j}} \right] = f_j \quad \text{and} \quad \text{Var} \left[ \frac{C_{i,j+1}}{C_{i,j}} \middle| C_{i,1}, \ldots, C_{i,j} \right] = \frac{\sigma_j^2}{C_{i,j}}
\]

- conditional weighted normal GLM

\[
\frac{C_{i,j+1}}{C_{i,j}} \sim \mathcal{N} \left( f_j, \frac{\sigma_j^2}{C_{i,j}} \right)
\]

- Mack’s model was not derived/designed as a GLM, but a conditional weighted normal normal GLM gives the same estimates

- NO distribution-free approach!
Mack’s model as GLM

- reformulate Mack’s model as a model of ratios

\[ \mathbb{E} \left[ \frac{C_{i,j+1}}{C_{i,j}} \right] = f_j \quad \text{and} \quad \text{Var} \left[ \frac{C_{i,j+1}}{C_{i,j}} \bigg| C_{i,1}, \ldots, C_{i,j} \right] = \frac{\sigma_j^2}{C_{i,j}} \]

- conditional weighted normal GLM

\[ \frac{C_{i,j+1}}{C_{i,j}} \sim \mathcal{N} \left( f_j, \frac{\sigma_j^2}{C_{i,j}} \right) \]

- Mack’s model was not derived/designed as a GLM, but a conditional weighted normal normal GLM gives the same estimates

- NO distribution-free approach!
GLM for triangles

- different (?) view on the triangles and chain ladder
- independent incremental claims \(X_{ij}, i + j \leq n + 1\)
  - overdispersed Poisson distributed \(X_{ij}\)
    \[
    \mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \text{Var}[X_{ij}] = \phi m_{ij}
    \]
  - Gamma distributed \(X_{ij}\)
    \[
    \mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \text{Var}[X_{ij}] = \phi m_{ij}^2
    \]
- logarithmic link function
  \[
  \log(m_{ij}) = \gamma + \alpha_i + \beta_j, \quad \alpha_1 = \beta_1 = 0
  \]
GLM for triangles

- different (?) view on the triangles and chain ladder
- independent incremental claims $X_{ij}$, $i + j \leq n + 1$
  - overdispersed Poisson distributed $X_{ij}$
    \[
    \mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \mathbb{V}[X_{ij}] = \phi m_{ij}
    \]
  - Gamma distributed $X_{ij}$
    \[
    \mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \mathbb{V}[X_{ij}] = \phi m_{ij}^2
    \]
- logarithmic link function
  \[
  \log(m_{ij}) = \gamma + \alpha_i + \beta_j, \quad \alpha_1 = \beta_1 = 0
  \]
overdispersed Poisson with log link provides asymptotically same parameter estimates, predicted values and prediction errors

possible extensions:

- Hoerl curve
  \[ \log(m_{i\cdot}) = \gamma + \alpha_i + \beta_j \log(j) + \delta_{j\cdot} \]

- smoother (semiparametric)
  \[ \log(m_{i\cdot}) = \gamma + \alpha_i + s_1(\log(j)) + s_2(j) \]
Algorithm 2

[1] suitable GLM $\rightarrow$ estimates $\hat{\gamma}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\phi}$ and, consequently, fitted claims

$$\hat{X}_{ij} \equiv \hat{m}_{ij} = \exp\{\hat{\gamma} + \hat{\alpha}_i + \hat{\beta}_j\}$$

[2] scaled Pearson residuals

$$r_{i,j} = \frac{X_{ij} - \hat{X}_{ij}}{\sqrt{\hat{\phi}\hat{X}_{ij}}}$$

[3] resample the residuals many times and fit the GLMs to pseudo triangles

[4] obtain empirical distribution of the reserves from the fitted bootstrapped triangles
Algorithm 2

[1] suitable GLM \( \sim \) estimates \( \hat{\gamma}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\phi} \) and, consequently, fitted claims

\[
\hat{X}_{ij} \equiv \hat{m}_{ij} = \exp\{\hat{\gamma} + \hat{\alpha}_i + \hat{\beta}_j\}
\]

[2] scaled Pearson residuals

\[
r_{i,j} = \frac{X_{ij} - \hat{X}_{ij}}{\sqrt{\hat{\phi}\hat{X}_{ij}}}
\]

[3] resample the residuals many times and fit the GLMs to pseudo triangles

[4] obtain empirical distribution of the reserves from the fitted bootstrapped triangles
Outline

1. Motivation
   - Origins
   - Prologue for Bootstrap in Statistics
   - Reserving Issue

2. Mathematical Background
   - Bootstrap
   - Stochastics in Insurance
   - Bootstrapping the Chain Ladder
   - Generalized Linear Models

3. Data Analysis
   - Estimation of Distribution

4. Conclusions
   - Discussion
## Taylor and Ashe (1983) data

- **Incremental triangle**

|        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|        |
| 357 848| 766 940| 610 542| 482 940| 527 326| 574 398| 146 342| 139 950| 227 229| 67 948 |
| 352 118| 884 021| 933 894| 1 183 289| 445 745| 320 996| 527 804| 266 172| 425 046 |
| 290 507| 1 001 799| 926 219| 1 016 654| 750 816| 146 923| 495 992| 280 405 |
| 310 608| 1 108 250| 776 189| 1 562 400| 272 482| 352 053| 206 286 |
| 443 160| 693 190| 991 983| 769 488| 504 851| 470 639 |
| 396 132| 937 085| 847 498| 805 037| 705 960 |
| 440 832| 847 631| 1 131 398| 1 063 269 |
| 359 480| 1 061 648| 1 443 370 |
| 376 686| 986 608 |
| 344 014 |

- **R software, ChainLadder package**
Development of claims

Development period

Claims

Michal Pešta (MFF UK & GPH)
Bootstrap in Reserving
Claims development by ChL with Mack’s s.e.

Chain ladder developments by origin period

Chain ladder dev.  Mack’s S.E.

Development period

Amount

0e+00  2e+06  4e+06  6e+06  8e+06

1  2  4  6  8  10

Chain ladder development graphs showing development by origin period with chain ladder and Mack's standard error.
Chain ladder diagnostics

Mack Chain Ladder Results

Chain ladder developments by origin period

Standardised residuals

Origin period

Development period

Calendar period

Development period

Standardised residuals

Michal Pešta (MFF UK & GPH)
Bootstrap results

Histogram of Total.IBNR

ecdf(Total.IBNR)

Simulated ultimate claims cost

Latest actual incremental claims against simulated values

Michal Pešta (MFF UK & GPH)

Bootstrap in Reserving

AS 2011
| Accident year | Chain Ladder | | | Bootstrap |
|---------------|--------------|--------------|--------------|
|               | Ultimate     | IBNR         | S.E.         | Ultimate     | IBNR         | S.E.         |
| 1             | 3 901 463    | 0            | 0            | 3 901 463    | 0            | 0            |
| 2             | 5 433 719    | 94 634       | 75 535       | 5 434 680    | 95 595       | 106 313      |
| 3             | 5 378 826    | 469 511      | 121 699      | 5 396 815    | 487 500      | 222 001      |
| 4             | 5 297 906    | 709 638      | 133 549      | 5 315 089    | 726 821      | 265 696      |
| 5             | 4 858 200    | 984 889      | 261 406      | 4 875 837    | 1 002 526    | 313 015      |
| 6             | 5 111 171    | 1 419 459    | 411 010      | 5 113 745    | 1 422 033    | 377 703      |
| 7             | 5 660 771    | 2 177 641    | 558 317      | 5 686 423    | 2 203 293    | 487 891      |
| 8             | 6 784 799    | 3 920 301    | 875 328      | 6 790 462    | 3 925 964    | 789 329      |
| 9             | 5 642 266    | 4 278 972    | 971 258      | 5 675 167    | 4 311 873    | 1 034 465    |
| 10            | 4 969 825    | 4 625 811    | 1 363 155    | 5 148 456    | 4 804 442    | 2 091 629    |
| Total         | 53 038 946   | 18 680 856   | 2 447 095    | 53 338 139   | 18 980 049   | 3 096 767    |
Comparison of distributional properties

- why to bootstrap?
- moment characteristics (mean, s.e., ...) does not provide full information about the reserves’ distribution
- additional assumption required in the classical approach
- 99.5% quantile necessary for VaR
  - assuming normally distributed reserves ... 24 984 154
  - assuming log-normally distributed reserves ... 25 919 050
  - bootstrap ... 28 201 572
Comparison of distributional properties

- **why** to bootstrap?
- moment characteristics (mean, s.e., ...) does not provide full information about the reserves’ distribution
- additional assumption required in the classical approach
- 99.5% quantile necessary for VaR
  - assuming normally distributed reserves ... 24 984 154
  - assuming log-normally distributed reserves ... 25 919 050
  - bootstrap ... 28 201 572
Outline

1 Motivation
   - Origins
   - Prologue for Bootstrap in Statistics
   - Reserving Issue

2 Mathematical Background
   - Bootstrap
   - Stochastics in Insurance
   - Bootstrapping the Chain Ladder
   - Generalized Linear Models

3 Data Analysis
   - Estimation of Distribution

4 Conclusions
   - Discussion
Conclusions

- Chain ladder based reserving techniques $\leadsto$ strong stochastic assumptions even if they do not assume prescribed distribution of claims
- "Distributional-free approaches" is a misleading expression ... do not require distributional assumptions $\leftrightarrow$ do not provide distributional properties
- Mean and variance do not contain full information about the distribution $\leadsto$ cannot provide quantities like VaR
- Assumption of log-normally distributed claims $\leftrightarrow$ log-normally distributed reserves (far more restrictive)
- Bootstrap (simulated) distribution mimics the unknown distribution of reserves (a mathematical proof necessary)
- R software provides a free sufficient actuarial environment for reserving
Conclusions

- chain ladder based reserving techniques $\leadsto$ strong stochastic assumptions even if they do not assume prescribed distribution of claims
- "distributional-free approaches" is a misleading expression . . . do not require distributional assumptions $\leftrightarrow$ do not provide distributional properties
- mean and variance do not contain full information about the distribution $\leadsto$ cannot provide quantities like VaR
- assumption of log-normally distributed claims $\not\leftrightarrow$ log-normally distributed reserves (far more restrictive)
- bootstrap (simulated) distribution mimics the unknown distribution of reserves (a mathematical proof necessary)
- R software provides a free sufficient actuarial environment for reserving
Conclusions

- chain ladder based reserving techniques \(\leadsto\) strong stochastic assumptions even if they do not assume prescribed distribution of claims

- "distributional-free approaches" is a misleading expression \(\ldots\) do not require distributional assumptions \(\leftrightarrow\) do not provide distributional properties

- mean and variance do not contain full information about the distribution \(\leadsto\) cannot provide quantities like VaR

- assumption of log-normally distributed claims \(\not\leftrightarrow\) log-normally distributed reserves (far more restrictive)

- bootstrap (simulated) distribution mimics the unknown distribution of reserves (a mathematical proof necessary)

- R software provides a free sufficient actuarial environment for reserving
Do not forget to . . . bootstrap!

Questions?
References


