Bootstrap Methods in Reserving

Michal Pešta

Charles University in Prague & Generali PPF Holding

Actuarial Seminar, Prague
December 9, 2011

Outline

“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

What is bootstrapping ?

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

Outline

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

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

Bootstrap example

- random sampling with replacement from the original dataset $\rightsquigarrow$ 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): $35, 34, 13, 33, 27, 30, 19, 31, 10, 33 \rightsquigarrow mean = 26.5,\ sd = 9.168182$
- bootstrap sample 1 (1st draw with replacement): $30, 27, 35, 35, 13, 35, 33, 34, 35, 33 \rightsquigarrow mean^*_1 = 31.0,\ sd^*_1 = 6.847546$

$\vdots$

- bootstrap sample 1000 (1000th draw with replacement): $19, 19, 31, 19, 33, 34, 31, 34, 34, 10 \rightsquigarrow mean^*_{1000} = 26.4,\ sd^*_{1000} = 8.771165$
- $mean^*_1, \ldots, mean^*_{1000}$ 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 $\rightsquigarrow$ 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): $35, 34, 13, 33, 27, 30, 19, 31, 10, 33 \rightsquigarrow mean = 26.5,\ sd = 9.168182$
- bootstrap sample 1 (1st draw with replacement): $30, 27, 35, 35, 13, 35, 33, 34, 35, 33 \rightsquigarrow mean^*_1 = 31.0,\ sd^*_1 = 6.847546$

⋮

- bootstrap sample 1000 (1000th draw with replacement): $19, 19, 31, 19, 33, 34, 31, 34, 34, 10 \rightsquigarrow mean^*_{1000} = 26.4,\ sd^*_{1000} = 8.771165$
- $mean^*_1, \ldots, mean^*_{1000}$ 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 $\rightsquigarrow$ 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): $35, 34, 13, 33, 27, 30, 19, 31, 10, 33 \rightsquigarrow mean = 26.5,\ sd = 9.168182$
- bootstrap sample 1 (1st draw with replacement): $30, 27, 35, 35, 13, 35, 33, 34, 35, 33 \rightsquigarrow mean^*_1 = 31.0,\ sd^*_1 = 6.847546$

⋮

- bootstrap sample 1000 (1000th draw with replacement): $19, 19, 31, 19, 33, 34, 31, 34, 34, 10 \rightsquigarrow mean^*_{1000} = 26.4,\ sd^*_{1000} = 8.771165$
- $mean^*_1, \ldots, mean^*_{1000}$ 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 $\rightsquigarrow$ 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):
  $35, 34, 13, 33, 27, 30, 19, 31, 10, 33 \rightsquigarrow mean = 26.5,\ sd = 9.168182$
- bootstrap sample 1 (1st draw with replacement):
  $30, 27, 35, 35, 13, 35, 33, 34, 35, 33 \rightsquigarrow mean^*_1 = 31.0,\ sd^*_1 = 6.847546$

⋮

- bootstrap sample 1000 (1000th draw with replacement):
  $19, 19, 31, 19, 33, 34, 31, 34, 34, 10 \rightsquigarrow mean^*_{1000} = 26.4,\ sd^*_{1000} = 8.771165$
- $mean^*_1, \ldots, mean^*_{1000}$ provide bootstrap empirical distribution for $mean$ and $sd^*_1, \ldots, sd^*_{1000}$ provide bootstrap empirical distribution for sd (REALLY!?)

Outline

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

Mack's chain ladder (no tail factor)

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

Assumptions

[1] $\mathbb{E}[C_{i,j+1} | C_{i,1}, \ldots, C_{i,j}] = f_j C_{i,j}, \quad 1 \le i \le n, 1 \le j \le n-1$

[2] $\mathbb{V}ar\,[C_{i,j+1} | C_{i,1}, \ldots, C_{i,j}] = \sigma_j^2 C_{i,j}, \quad 1 \le i \le n, 1 \le j \le n-1$

[3] accident years $[C_{i,1}, \ldots, C_{i,n}], \quad 1 \le i \le n$ are independent

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

Main Goals

- reasonable estimate $\hat{f}_j$ for development factors (is unbiased, but consistent?)
- estimate conditional s.e. of estimates of ultimates and reserves

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

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] $\mathbb{V}ar\,[C_{i,j+1}|C_{i,1},\ldots,C_{i,j}] = \sigma_j^2 C_{i,j}, \quad 1 \leq i \leq n,\ 1 \leq j \leq n-1$

[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 $\widehat{f}_j$ for development factors (is unbiased, but consistent?)

[b] estimate conditional s.e. of estimates of ultimates and reserves
$\mathbb{E}[(\widehat{C}_{i,n} - C_{i,n})^2|\{C_{i,j} : i+j \leq n+1\}] = \mathbb{E}[(\widehat{R}_i - R_i)^2|\{C_{i,j} : i+j \leq n+1\}]$

[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] $\mathbb{E}[C_{i,j+1}|C_{i,1},\ldots,C_{i,j}] = f_j C_{i,j}, \quad 1 \le i \le n,\ 1 \le j \le n-1$

[2] $\mathbb{V}ar\,[C_{i,j+1}|C_{i,1},\ldots,C_{i,j}] = \sigma_j^2 C_{i,j}, \quad 1 \le i \le n,\ 1 \le j \le n-1$

[3] accident years $\quad [C_{i,1},\ldots,C_{i,n}], \quad 1 \le i \le n \quad$ 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 $\widehat{f}_j$ for development factors (is unbiased, but consistent?)

[b] estimate conditional s.e. of estimates of ultimates and reserves
$\mathbb{E}[(\widehat{C}_{i,n} - C_{i,n})^2|\{C_{i,j} : i+j \le n+1\}] = \mathbb{E}[(\widehat{R}_i - R_i)^2|\{C_{i,j} : i+j \le n+1\}]$

[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] $\mathbb{V}ar\,[C_{i,j+1}|C_{i,1},\ldots,C_{i,j}] = \sigma_j^2 C_{i,j}, \quad 1 \leq i \leq n,\ 1 \leq j \leq n-1$

[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 $\widehat{f}_j$ for development factors (is unbiased, but consistent?)

[b] estimate conditional s.e. of estimates of ultimates and reserves
$\mathbb{E}[(\widehat{C}_{i,n} - C_{i,n})^2|\{C_{i,j} : i+j \leq n+1\}] = \mathbb{E}[(\widehat{R}_i - R_i)^2|\{C_{i,j} : i+j \leq n+1\}]$

[c] conditional distribution of reserves given data?

Bootstrapping the chain ladder I

Algorithm 1 (Part I)

[1] estimate development factors

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

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

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

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

$$\widehat{C}_{i,n+1-i} \equiv C_{i,n+1-i}; \qquad \widehat{C}_{i,j} = \frac{C_{i,n+1-i}}{\widehat{f}_{n-i} \times \ldots \times \widehat{f}_j}, \quad i+j \leq n$$

Bootstrapping the chain ladder I

Algorithm 1 (Part I)

[1] estimate development factors

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

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

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

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

$$\widehat{C}_{i,n+1-i} \equiv C_{i,n+1-i}; \qquad \widehat{C}_{i,j} = \frac{C_{i,n+1-i}}{\widehat{f}_{n-i} \times \ldots \times \widehat{f}_j}, \quad i+j \leq n$$

Bootstrapping the chain ladder I

Algorithm 1 (Part I)

[1] estimate development factors

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

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

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

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

$$\widehat{C}_{i,n+1-i} \equiv C_{i,n+1-i}; \qquad \widehat{C}_{i,j} = \frac{C_{i,n+1-i}}{\widehat{f}_{n-i} \times \ldots \times \widehat{f}_j}, \quad i+j \leq n$$

Bootstrapping the chain ladder II

Algorithm 1 (Part II)

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

$$r_{i,j} = \frac{(C_{i,j} - C_{i,j-1}) - \left(\widehat{C}_{i,j} - \widehat{C}_{i,j-1}\right)}{\sqrt{\widehat{C}_{i,j} - \widehat{C}_{i,j-1}}}, \quad i + j \leq n + 1$$

- [1]–[4] are just Mack chain ladder

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

[6] construct B incremental bootstrap triangles

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

Bootstrapping the chain ladder II

Algorithm 1 (Part II)

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

$$r_{i,j} = \frac{(C_{i,j} - C_{i,j-1}) - \left(\widehat{C}_{i,j} - \widehat{C}_{i,j-1}\right)}{\sqrt{\widehat{C}_{i,j} - \widehat{C}_{i,j-1}}}, \quad i+j \leq n+1$$

- [1]–[4] are just Mack chain ladder

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

[6] construct B incremental bootstrap triangles

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

Bootstrapping the chain ladder II

Algorithm 1 (Part II)

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

$$r_{i,j} = \frac{(C_{i,j} - C_{i,j-1}) - \left(\widehat{C}_{i,j} - \widehat{C}_{i,j-1}\right)}{\sqrt{\widehat{C}_{i,j} - \widehat{C}_{i,j-1}}}, \quad i + j \leq n + 1$$

- [1]–[4] are just Mack chain ladder

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

[6] construct B incremental bootstrap triangles

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

Algorithm 1 (Part III)

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

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

[8] perform chain ladder on each bootstrap cumulative triangle ⇝ 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 ⇝ empirical (estimated) mean, s.e., quantiles, ...

Bootstrapping the chain ladder III

Algorithm 1 (Part III)

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

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

[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, ...

Bootstrapping the chain ladder III

Algorithm 1 (Part III)

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

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

[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, ...

Outline

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 $\mathbf{Y}$ from a particular distribution in the overdispersed exponential family, i.e.,

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

where τ is dispersion parameter

2. linear predictor (mean structure)

$$\boldsymbol{\eta} = \mathbf{X}\beta$$

3. link function g (element-wise)

$$\mathbb{E}\mathbf{Y} = \boldsymbol{\mu} = g^{-1}(\boldsymbol{\eta})$$

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 $\mathbf{Y}$ from a particular distribution in the overdispersed exponential family, i.e.,

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

   where τ is dispersion parameter
2. linear predictor (mean structure)

$$\boldsymbol{\eta} = \mathbf{X}\beta$$

3. link function g (element-wise)

$$\mathbb{E}\mathbf{Y} = \boldsymbol{\mu} = g^{-1}(\boldsymbol{\eta})$$

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 $\mathbf{Y}$ from a particular distribution in the overdispersed exponential family, i.e.,

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

where τ is dispersion parameter

2. linear predictor (mean structure)

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$$

3. link function g (element-wise)

$$\mathbb{E}\mathbf{Y} = \boldsymbol{\mu} = g^{-1}(\boldsymbol{\eta})$$

Overdispersed exponential family

- 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

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

- distribution $\longleftrightarrow$ link function (element-wise)
  - normal ... identity: $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$
  - gamma (exponential) ... inverse: $(\boldsymbol{\mu})^{-1} = \mathbf{X}\boldsymbol{\beta}$
  - Poisson ... logarithm: $\log(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$
  - binomial (multinomial) ... logit: $\log\left(\frac{\boldsymbol{\mu}}{1-\boldsymbol{\mu}}\right) = \mathbf{X}\boldsymbol{\beta}$
- estimation of the parameters via maximum likelihood, quasi-likelihood or Bayesian techniques

Generalized Linear Models (GLM) II

- overdispersed exponential family

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

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

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

- conditional weighted normal GLM

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

- Mack's model was not derived/designed as a GLM, but a conditional weighted normal GLM gives the same estimates
- NO distribution-free approach !

Mack's model as GLM

- reformulate Mack's model as a model of ratios

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

- conditional weighted normal GLM

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

- Mack's model was not derived/designed as a GLM, but a conditional weighted normal 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}

$$\mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \mathbb{V}ar\left[X_{ij}\right] = \phi m_{ij}$$

  - Gamma distributed X_{ij}

$$\mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \mathbb{V}ar\left[X_{ij}\right] = \phi m_{ij}^2$$

- 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}

$$\mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \mathbb{V}ar\,[X_{ij}] = \phi m_{ij}$$

  - Gamma distributed X_{ij}

$$\mathbb{E}[X_{ij}] = m_{ij} \quad \text{and} \quad \mathbb{V}ar\,[X_{ij}] = \phi m_{ij}^2$$

- logarithmic link function

$$\log(m_{ij}) = \gamma + \alpha_i + \beta_j, \quad \alpha_1 = \beta_1 = 0$$

GLM for triangles II

- overdispersed Poisson with log link provides asymptotically same parameter estimates, predicted values and prediction errors
- possible extensions:
  - Hoerl curve

$$\log(m_{ij}) = \gamma + \alpha_i + \beta_j \log(j) + \delta_j j$$

  - smoother (semiparametric)

$$\log(m_{ij}) = \gamma + \alpha_i + s_1(\log(j)) + s_2(j)$$

Bootstrap GLM

Algorithm 2

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

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

[2] scaled Pearson residuals

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

[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

Bootstrap GLM

Algorithm 2

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

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

[2] scaled Pearson residuals

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

[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

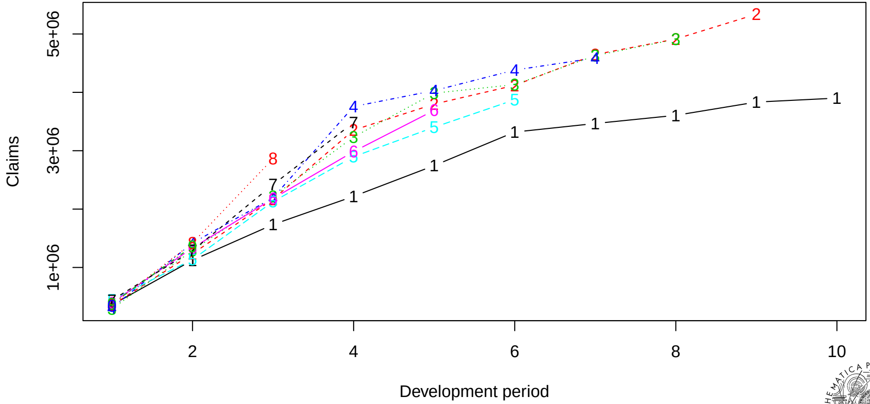
Taylor and Ashe (1983) data

- incremental triangle

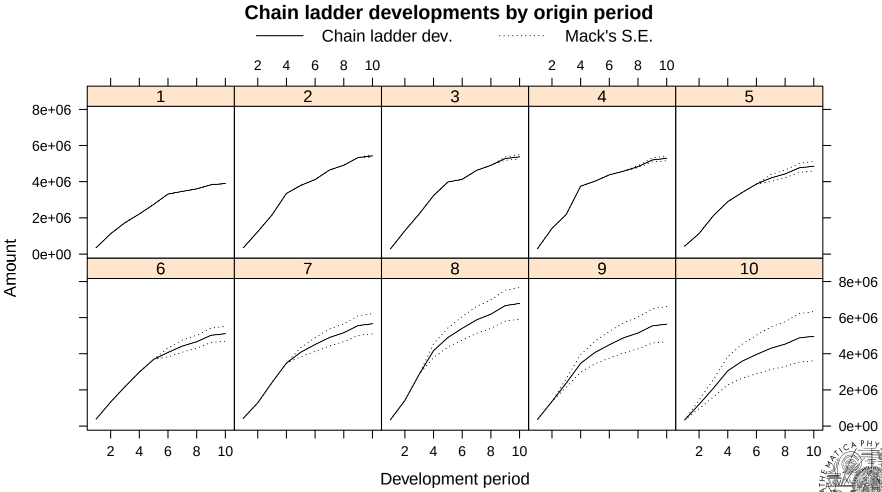
357 848	766 940	610 542	482 940	527 326	574 398	146 342	139 950	227 229	67 948
352 118	884 021	933 894	1 183 289	445 745	320 996	527 804	266 172	425 046	
290 507	1 001 799	926 219	1 016 654	750 816	146 923	495 992	280 405		
310 608	1 108 250	776 189	1 562 400	272 482	352 053	206 286			
443 160	693 190	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

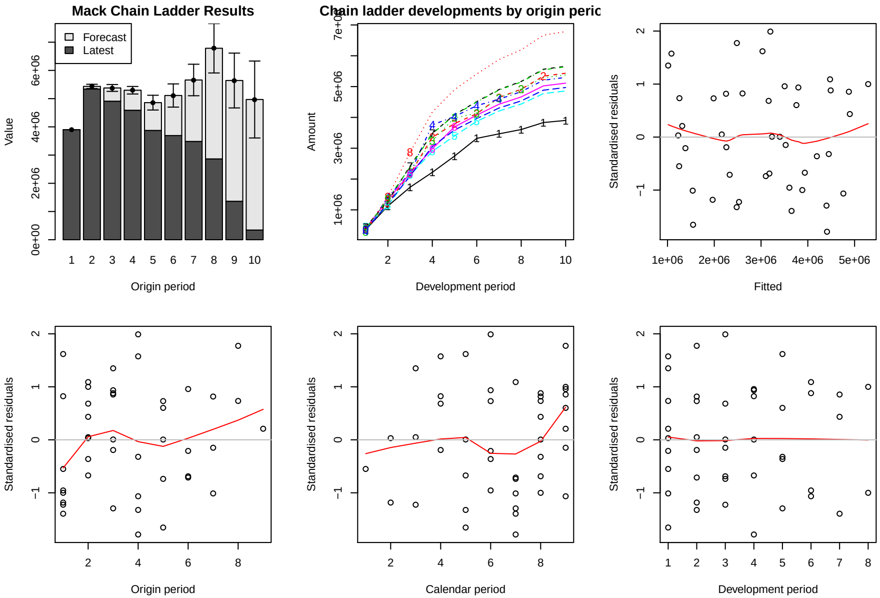
Development of claims



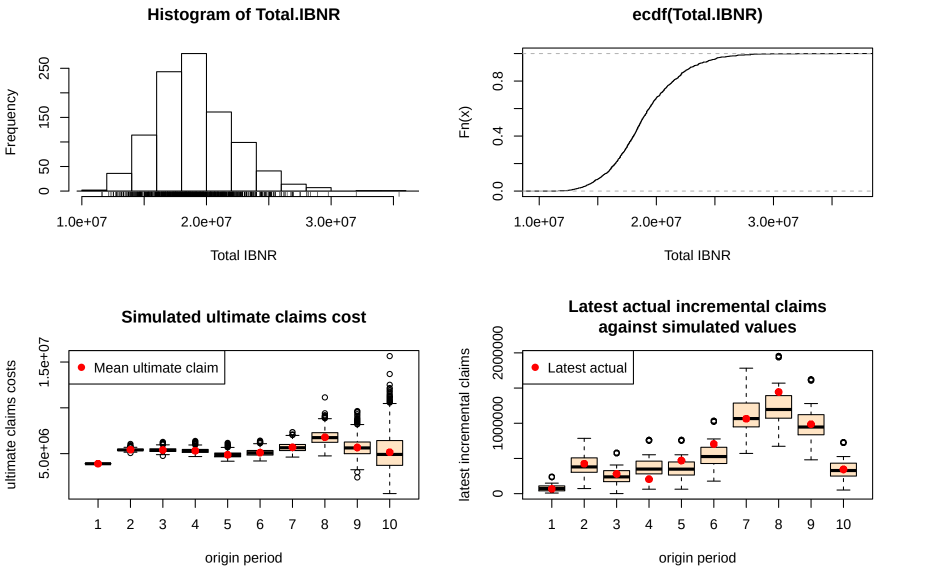
Claims development by ChL with Mack's s.e.



Chain ladder diagnostics



Bootstrap results



Mack Chain Ladder vs. Bootstrap GLM

Accident year	Chain Ladder Ultimate	Chain Ladder IBNR	Chain Ladder S.E.	Bootstrap Ultimate	Bootstrap IBNR	Bootstrap 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

Conclusions

- chain ladder based reserving techniques $\rightsquigarrow$ 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 $\longleftrightarrow$ do not provide distributional properties
- mean and variance do not contain full information about the distribution $\rightsquigarrow$ 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 $\rightsquigarrow$ 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 $\longleftrightarrow$ do not provide distributional properties
- mean and variance do not contain full information about the distribution $\rightsquigarrow$ 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 $\rightsquigarrow$ 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 $\longleftrightarrow$ do not provide distributional properties
- mean and variance do not contain full information about the distribution $\rightsquigarrow$ 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

Mack, T. (1993)
Distribution-free calculation of the standard error of chain ladder reserve estimates.
Astin Bulletin, **23**(2), 213–225.

Mack, T. (1999)
The standard error of chain ladder reserve estimates: Recursive calculation and inclusion of a tail factor.
Astin Bulletin, **29**(2), 361–366.

England, P. D. and Verrall, R. J. (1999)
Analytic and bootstrap estimates of prediction errors in claims reserving.
Insurance: Mathematics and Economics, **25**, 281–293.

England, P. D. and Verrall, R. J. (2002)
Stochastic claims reserving in general insurance (with discussion).
British Actuarial Journal, **8**, III.