分块 Bootstrap 方法在非寿险准备金评估中的应用

刘乐平　高磊

天津财经大学统计系

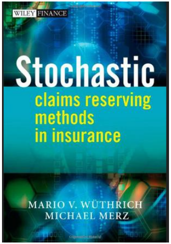
2015 年 6 月 23 日

Stochastic claims reserving

- This has become a new academic dicipline
- Numerous papers appeare in academic journals
- A book has appeared
- There is a Wikipedia page

Google scholar search for "Stochastic claims reserving"

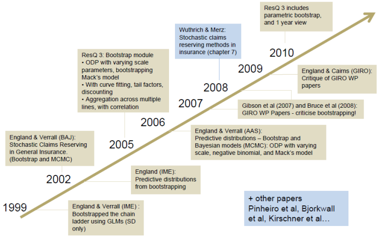
Mario V. Wüthrich

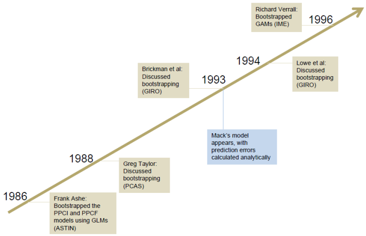


Wüthrich & Merz(2008)

Bootstrapping: the last 20 years(England,2010)

The holy grail of stochastic reserving is to obtain a predictive distribution of outstanding liabilities.





One method that has been proposed to produce a predictive distribution is Bootstrapping.

Content

Content

The over-dispersed Poisson model

- One of the most popular stochastic claims reserving models.
- Model assumption(*Renshaw & Verall,1998*)

$X_{i,j}$,incremental payments.
$X_{i,j} \sim ODP(m_{i,j}, \phi_j)$
$E[X_{i,j}] = m_{i,j} = x_i * y_j \qquad Var[X_{i,j}] = \phi_j * m_{i,j}$

$$log(m_{i,j}) = c + \alpha_i + \beta_j.$$

R function:glm()

Bootstrapping ODP model

- Step of bootstrapping ODP model(*England & Verall, 1999, 2002, 2006*).
  1. Fit the model and obtain fitted values $\hat{m}_{ij}$
  2. Calculate the residuals
  $$r_{i,j} = \frac{X_{ij} - \hat{m}_{i,j}}{\sqrt{\phi_j \hat{m}_{ij}}}.$$
  3. Resample residual r_{ij}^*
  4. Obtain pseudo data
  $$X_{ij}^* = r_{ij}^* \sqrt{\phi_j \hat{m}_{ij}} + \hat{m}_{ij}.$$
  5. Refit ODP model to estimate the future incremental payments
  6. Simulate forecast incremental payments from process distribution
  7. Repeate many times and store the simulated forecast payment

A practical problem:violation of independence assumption

- A motivating example(*Joseph,2011*).

	0	1	2	3	4	5	6	7
0	1167	6544	16689	33506	57307	84796	116127	146842
1	13639	47608	117523	213809	328127	457809	602945	
2	11392	53394	130296	248022	401575	588795		
3	20546	72208	159786	287992	448246			
4	22147	77021	163717	282129				
5	23313	97398	215608					
6	34009	103645						
7	21972							

- Fit ODP model and obtain fitted incremental payments $\hat{m}_{ij}$.

Violation of independence assumption

- Calculate the Scaled Pearson Residuals: $r_{i,j} = \frac{X_{ij} - \hat{m}_{i,j}}{\sqrt{\phi_j \hat{m}_{ij}}}$.
- Scaled Pearson Residuals.

	0	1	2	3	4	5	6	7
0	-1.22	-0.76	-0.72	-0.17	0.44	0.22	0.73	0
1	0.07	-0.32	1.40	0.93	-0.19	-1.04	-0.32	
2	-2.00	-0.77	-0.50	0.17	0.42	0.83		
3	0.47	0.31	-0.02	0.10	-0.40			
4	1.03	1.01	-0.06	-0.99				
5	-0.78	0.88	-0.30					
6	1.41	-0.87						
7	0.00							

Violation of independence assumption

- '+':residual is greater than 0; '−': residual is less than 0.

	0	1	2	3	4	5	6	7
0	-	-	-	-	+	+	+	o
1	+	-	+	+	-	-	-	
2	-	-	-	+	+	+		
3	+	+	-	+	-			
4	+	+	-	-				
5	-	+	-					
6	+	-						
7	o							

- '+' and '−' appear consecutively, which means the residuals are non-random or non-independent.

Content

Runs-test of residual sequence

- P-value of runs-test *(Mendenhall,1982)*.

	0.937	0.113	0.819	0.331	0.89	1		
0.011	-	-	-	-	+	+	+	o
0.358	+	-	+	+	-	-	-	
0.034	-	-	-	+	+	+		
0.063	+	+	-	+	-			
0.11	+	+	-	-				
0.24	-	+	-					
	+	-						
	o							

- This process is called multiple testing.FWER:

$$1 - (1 - 0.05)^{12} = 0.46 \gg 0.05$$

Multiple testing:BH's FDR Control Algorithm

- *Benjamini & Hochberg(1995)*
  1. Order p-values in increasing order and denote them by
  $$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(i)} \leq \cdots \leq p_{(N)}.$$
  2. For a fixed value of α, find the largest k_{max} for which
  $$p_{(k)} \leq \frac{k}{N}\alpha.$$
  3. Then reject the null hypothesis corresponding to $p_{(k)}$, if $k \leq k_{max}$.

Multiple testing:BH's FDR Control Algorithm

- The result of FDR control.

	0	1	2	3	4	5	6	7
0	-	-	-	-	+	+	+	o
1	+	-	+	+	-	-	-	
2	-	-	-	+	+	+		
3	+	+	-	+	-			
4	+	+	-	-				
5	-	+	-					
6	+	-						
7	o							

- The accident year $i = 2$ is excluded.

Block bootstrapping

- The residuals triangle is divided into some blocks.
- Bootstrap residuals r_{ij}^* are selected from the corresponding block.

	0	1	2	3	4	5	6	7
0	-	-	-	-	+	+	+	0
1	+	-	+	+	-	-	-	
2	-	-	-	+	+	+		
3	+	+	-	+	-			
4	+	+	-	-				
5	-	+	-					
6	+	-						
7	0							

- The following procedure are as same as the original bootstrap method.

Content

3 A real example

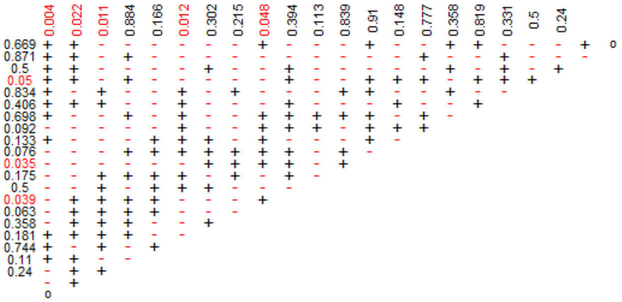
A real data example

- The data are from Verall & Wüthrich(2012).
- '$+$':residual is greater than 0; '$-$': residual is less than 0.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	+	+	-	-	-	-	-	-	+	-	-	-	+	-	-	+	+	-	-	-	+	o
1	+	+	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	
2	+	+	-	-	-	-	+	-	-	+	-	-	-	-	-	+	-	+	-	+		
3	+	+	-	+	-	-	-	-	-	+	-	-	+	+	+	-	+	+	+			
4	+	-	+	-	-	+	-	+	-	-	-	+	+	-	-	+	-	-				
5	+	+	+	-	-	+	-	-	-	+	-	-	-	+	-	-	+					
6	+	-	-	+	-	+	-	-	+	+	+	+	+	-	+	-						
7	-	-	-	-	-	+	-	-	+	+	+	-	+	+	+							
8	+	-	-	-	+	+	+	-	+	-	-	-	+	-								
9	-	-	-	+	+	+	+	+	+	+	-	+	-									
10	-	-	-	-	-	-	+	+	+	+	-	+										
11	-	-	+	+	+	+	-	+	-	+	-											
12	-	-	+	-	+	+	+	-	-	-												
13	-	+	+	+	+	-	-	-	+													
14	-	+	+	+	+	-	-	-														
15	-	+	+	+	-	-	+															
16	+	+	+	+	-	-																
17	+	-	+	-	+																	
18	+	+	-	-																		
19	-	+	+																			
20	-	+																				
21	o																					

Runs-test of residual sequence

- P-value of runs-test.



Multiple testing:BH's FDR Control Algorithm

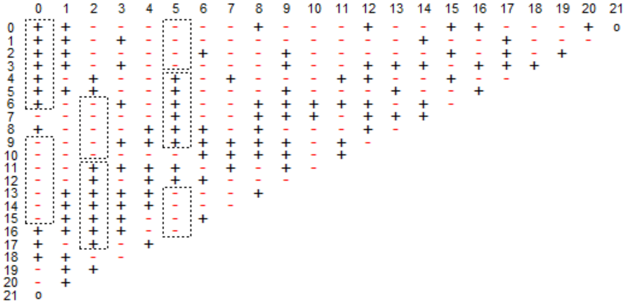
- The result of FDR control.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	+	+	-	-	-	-	-	-	+	-	-	-	+	-	-	+	+	-	-	-	+	o
1	+	+	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	
2	+	+	-	-	-	-	+	-	-	+	-	-	-	-	-	+	-	+	-	+		
3	+	+	-	+	-	-	-	-	-	+	-	-	+	+	+	-	+	+	+			
4	+	-	+	-	-	+	-	+	-	-	-	+	+	-	-	+	-	-				
5	+	+	+	-	-	+	-	-	-	+	-	-	-	+	-	-	+					
6	+	-	-	+	-	+	-	-	+	+	+	+	+	-	+	-						
7	-	-	-	-	-	+	-	-	+	+	+	-	+	+	+							
8	+	-	-	-	+	+	+	-	+	-	-	-	+	-								
9	-	-	-	+	+	+	+	+	+	+	-	+	-									
10	-	-	-	-	-	-	+	+	+	+	-	+										
11	-	-	+	+	+	+	-	+	-	+	-											
12	-	-	+	-	+	+	+	-	-	-												
13	-	+	+	+	+	-	-	-	+													
14	-	+	+	+	+	-	-	-														
15	-	+	+	+	-	-	+															
16	+	+	+	+	-	-																
17	+	-	+	-	+																	
18	+	+	-	-																		
19	-	+	+																			
20	-	+																				
21	o																					

- The accident years $i = 3, 10, 13$ and development years $j = 1, 8$ are excluded.

Block bootstrapping

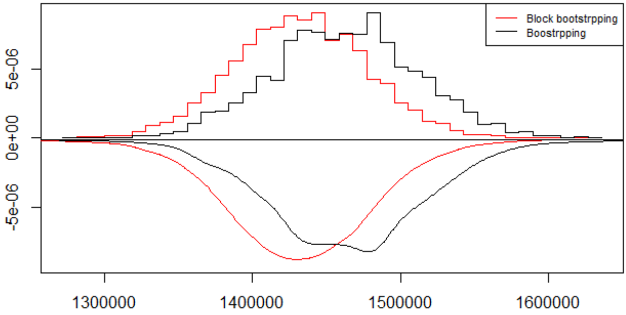
- The residuals triangle is divided into some blocks.
- Bootstrap residuals r_{ij}^* are selected from the corresponding block.



- The following procedure are as same as the original bootstrap method.

Numerical result:predictive distribution and prediction error

- Histogram and density chart



- Five-number summary, mean and standard error

	Min.	1st Qu.	Median	3rd Qu.	Max.	Mean	Std.
bootstrap	1298000	1430000	1463000	1494000	1636000	1463000	48972
block bootstrap	1282000	1402000	1432000	1463000	1569000	1432000	44977

References

1. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing[J]. Journal of the Royal Statistical Society. Series B (Methodological), 1995: 289-300.
2. England P, Verrall R. Analytic and bootstrap estimates of prediction errors in claims reserving[J]. Insurance: mathematics and economics, 1999, 25(3): 281-293.
3. England P D, Verrall R J. Stochastic claims reserving in general insurance[J]. British Actuarial Journal, 2002, 8(03): 443-518.
4. England P D, Verrall R J. Predictive distributions of outstanding liabilities in general insurance[J]. Annals of Actuarial Science, 2006, 1(02): 221-270.
5. Wüthrich M V, Merz M. Stochastic claims reserving methods in insurance[M]. John Wiley & Sons, 2008.
6. Joseph Lo A. Extending the Mack Bootstrap: Hypothesis Testing and Resampling Techniques[J]. The Actuarial Profession, 2011.
7. Verrall R J, Wüthrich M V. Reversible jump Markov chain Monte Carlo method for parameter reduction in claims reserving[J]. North American Actuarial Journal, 2012, 16(2): 240-259.

Thank you !

A/Q?