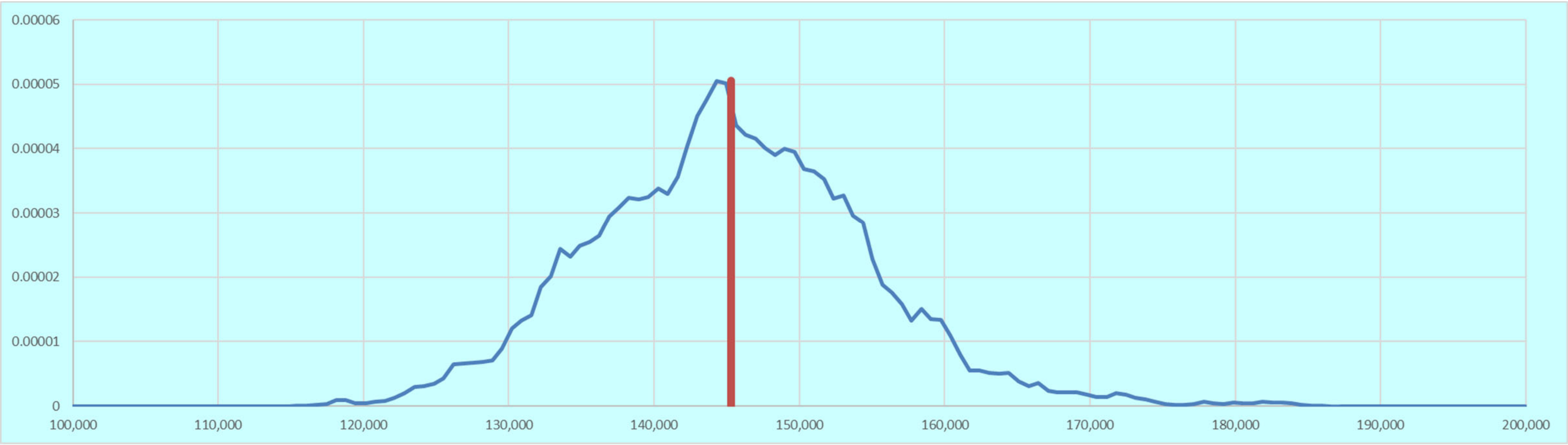


Making Bootstrap Reserve Ranges More Realistic: Including Correlation

Casualty Loss Reserve Seminar
September 2022

Munich Reinsurance America Services David R Clark, Hang Ding

Munich RE

Agenda

Munich RE

1. Overview about bootstrapping method
2. Business Purpose of Bootstrapping
3. Outline of Bootstrapping
4. Example with CAS Schedule P database
5. Including Correlation
6. Results

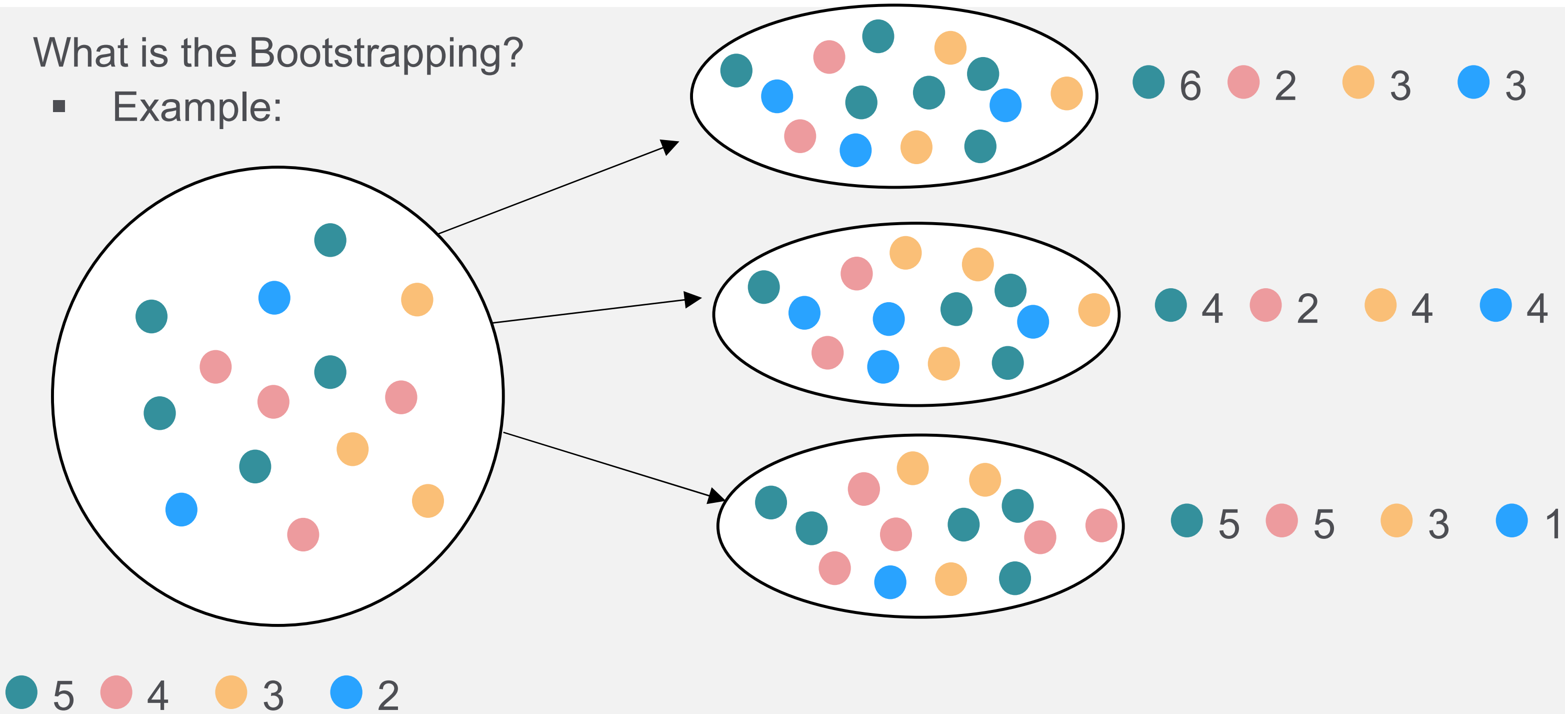
Bootstrapping Introduction

Concept: Statistical technique for estimating the quantity about a population by averaging estimates from multiple small data samples.

1. Choose how many bootstrap samples do we want to use
2. Choose a sample size
3. For each bootstrap sample
    1. Draw a sample with replacement with the chosen size
    2. Calculate the statistic (Mean, Variance, median or even maximum) on the sample
4. Calculate the mean of the calculated sample statistics.

Bootstrapping Introduction

- What is the Bootstrapping?
  - Example:



- Estimate the proportion for each color.
- What is the mean of the proportion for each color? What if we resample it with replacement for 1000 times? It will follow a certain distribution. And Bootstrapping will help us to get the distribution.

Business Purpose

Munich RE

The estimate of future payments on (re)insurance contracts already written is subject to volatility.

The volatility affects many areas of our business:

- Reserving: including Solvency II and IFRS 17
- Loss Portfolio Transfers
- Enterprise Risk Management

Various methods have been used to quantify this uncertainty.

Bootstrapping methods, made popular by England & Verrall (1999, 2002), are commonly used but do not always produce realistic results.

Outline of Bootstrapping

The goal of bootstrapping is to estimate the variance of a statistical estimator based on the data itself.

From the saying: “*lifting yourself up by your own bootstraps*”

We are actually estimating the distribution of the chain ladder estimate of ultimate as the first step.

Key assumptions:

1) Ultimate losses are estimated using the all-year weighted average chain ladder method.
2) Each cell in the incremental development triangle has the same variance/mean ratio, ϕ.
3) Each cell in the incremental development triangle is statistically independent from the other cells.

Any of these assumptions can be relaxed in practice. We will focus on the third.

Outline of Bootstrapping

Munich RE

Actual Incremental Triangle

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120	Sum
1988	358	767	611	483	527	574	146	140	227	68	3,901
1989	352	884	934	1,183	446	321	528	266	425		5,339
1990	291	1,002	926	1,017	751	147	496	280			4,909
1991	311	1,108	776	1,562	272	352	206				4,588
1992	443	693	992	769	505	471					3,873
1993	396	937	847	805	706						3,692
1994	441	848	1,131	1,063							3,483
1995	359	1,062	1,443								2,864
1996	377	987									1,363
1997	344										344
Sum:	3,671	8,287	7,661	6,883	3,207	1,865	1,376	687	652	68	

Fitted Incremental Triangle

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120	Sum
1988	270	673	704	753	417	293	268	182	273	68	3,901
1989	376	937	981	1,049	581	407	374	254	380		5,339
1990	372	927	971	1,039	575	403	370	251			4,909
1991	367	913	957	1,023	567	397	364				4,588
1992	336	838	877	938	520	364					3,873
1993	354	881	923	987	547						3,692
1994	392	976	1,022	1,093							3,483
1995	470	1,170	1,225								2,864
1996	391	973									1,363
1997	344										344
Sum:	3,671	8,287	7,661	6,883	3,207	1,865	1,376	687	652	68	

Bootstrapping starts with an incremental paid (or incurred) triangle.

A “fitted” version of the triangle is created based on chain-ladder or GLM.

The fitted triangle will have the same row and column totals as the original triangle.

Data is the “Taylor / Ashe” example as used in the England and Verrall paper.

Outline of Bootstrapping

Munich RE

Standardized Residuals

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	0.91	0.62	-0.60	-1.68	0.92	2.81	-1.27	-0.53	-0.47	0.00
1989	-0.21	-0.29	-0.26	0.70	-0.96	-0.73	1.36	0.14	0.40	
1990	-0.72	0.42	-0.25	-0.12	1.25	-2.18	1.12	0.32		
1991	-0.50	1.10	-0.99	2.87	-2.11	-0.39	-1.41			
1992	0.99	-0.85	0.66	-0.94	-0.11	0.95				
1993	0.38	0.32	-0.42	-0.99	1.16					
1994	0.42	-0.70	0.58	-0.15						
1995	-0.87	-0.54	1.06							
1996	-0.12	0.08								
1997	0.00									

Average	0.00	Minimum	-2.18	V/M ϕ	34
Std Dev	1.00	Maximum	2.87		

Next a triangle of standardized residuals is created.

This assumes:

- each cell has a mean equal to the "fitted" value
- that there is a constant variance/mean ratio ϕ for all cells in the triangle

The smallest and largest values are marked just for reference.

Outline of Bootstrapping

Munich RE ®

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	-0.73	-0.15	0.99	-0.85	0.42	0.40	-1.41	0.99	0.00	-0.50
1989	0.38	-0.12	0.00	2.87	-1.68	0.00	-0.11	0.70	-0.50	
1990	1.36	0.58	-0.47	0.66	0.00	-0.42	-0.99	0.70		
1991	-0.72	-0.73	2.87	-0.26	0.00	2.87	2.81			
1992	-2.11	-0.47	-0.39	1.06	0.32	-0.85				
1993	-0.11	-0.53	-2.18	-0.99	0.32					
1994	-0.87	0.32	0.70	2.81						
1995	2.87	-1.68	2.87							
1996	-0.21	-1.41								
1997	0.95									

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	0.62	1.25	-0.73	0.32	0.40	2.81	-0.47	0.95	0.99	-0.60
1989	-0.26	-0.42	-0.39	0.42	-1.27	-0.99	-0.99	2.87	-0.11	
1990	-0.99	-0.26	0.40	-0.96	0.91	-1.41	0.99	-1.27		
1991	0.40	-0.87	-2.18	-0.47	2.81	0.99	0.00			
1992	-0.94	0.62	0.08	-0.25	-0.26	-0.21				
1993	-0.60	-0.54	-0.54	0.92	0.38					
1994	-2.18	-0.87	1.25	2.87						
1995	0.62	2.87	0.32							
1996	-1.41	-0.12								
1997	-0.29									

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	0.14	0.00	1.12	0.40	1.06	-0.60	-0.26	1.16	0.99	1.25
1989	-0.99	-0.42	-0.54	0.58	-2.18	0.32	-0.87	0.95	1.06	
1990	-0.73	0.00	0.14	-0.96	-0.72	-0.94	-0.26	-0.96		
1991	1.16	-1.41	0.14	0.58	-0.72	0.08	-0.73			
1992	-0.11	-0.29	0.14	-0.87	-0.99	1.25				
1993	0.08	0.08	-0.54	0.58	-0.53					
1994	-0.29	-2.11	-2.18	-0.50						
1995	-0.15	0.58	0.14							
1996	0.95	-0.99								
1997	-0.53									

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	0.92	0.00	-0.11	0.14	-1.41	-0.29	1.25	0.66	-0.53	-0.15
1989	-1.27	-0.12	0.38	0.00	1.16	1.10	-0.70	0.00	-0.54	
1990	1.06	-0.87	-0.29	1.25	-0.72	-0.15	1.25	-0.50		
1991	-1.41	0.32	-0.96	-0.50	-0.15	-0.54	0.38			
1992	1.36	-0.96	0.66	0.62	0.95	-0.12				
1993	0.92	0.95	-2.11	-0.21	0.40					
1994	-1.27	2.81	-0.70	-0.26						
1995	-2.18	-0.29	-1.41							
1996	-0.50	-0.39								
1997	0.70									

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	0.00	0.00	-0.73	2.81	-0.25	2.81	-0.54	0.66	-2.11	-0.21
1989	-0.70	-0.11	-0.29	1.06	-0.53	0.08	-0.54	-0.50	1.36	
1990	0.32	0.42	0.32	-0.50	-0.70	-0.54	0.99	0.32		
1991	-0.12	0.95	1.06	0.42	0.58	0.00	0.99			
1992	-1.68	-0.12	0.40	-0.15	-0.70	0.62				
1993	-1.68	0.95	0.14	1.06	-2.11					
1994	-0.25	0.95	-0.94	-0.12						
1995	-0.87	1.25	-0.73							
1996	-1.68	0.42								
1997	0.38									

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	0.95	0.70	1.25	2.87	1.25	-0.25	0.92	-0.29	-0.94	0.70
1989	0.62	0.91	0.40	0.32	0.92	-0.29	1.12	-0.99	-0.29	
1990	-0.25	-0.25	-0.94	0.42	-0.21	-0.87	0.91	1.06		
1991	-0.96	0.08	0.14	-0.25	-0.47	-1.27	-2.18			
1992	-0.12	0.42	0.66	1.12	-0.47	-1.41				
1993	0.95	-0.99	2.81	-0.73	-0.60					
1994	0.08	0.99	0.92	0.38						
1995	1.06	-0.87	0.66							
1996	0.91	0.32								
1997	0.70									

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	0.32	0.00	2.87	-2.11	-0.85	-0.85	-0.12	0.92	-0.87	-0.70
1989	1.10	-0.21	0.40	-2.11	1.36	1.25	1.16	0.42	0.08	
1990	0.99	-0.96	0.08	0.70	-0.99	0.32	-1.41	-0.12		
1991	0.42	0.00	-0.25	0.66	0.92	-0.96	0.32			
1992	1.12	-0.26	-0.87	-0.11	0.00	-0.11				
1993	0.42	1.36	1.16	-0.85	0.00					
1994	-0.53	-0.53	0.92	-0.26						
1995	1.10	0.40	-0.47							
1996	2.87	0.32								
1997	0.92									

Year	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
1988	-0.26	-2.18	0.40	-0.54	-0.54	-0.72	0.42	-1.41	-1.41	0.40
1989	0.62	-0.87	0.66	-0.25	0.42	2.81	-0.26	0.38	1.36	
1990	-0.25	-0.73	-0.25	-0.39	2.81	-1.27	-0.26	-0.26		
1991	0.95	-2.11	-0.25	-0.50	0.95	0.92	-0.70			
1992	0.00	-0.21	0.40	-0.11	-0.72	-0.53				
1993	-0.12	-0.53	1.36	2.81	-0.47					
1994	0.92	-1.41	0.95	-0.29						
1995	0.42	1.36	0.08							
1996	-0.60	-0.85								
1997	-0.29									

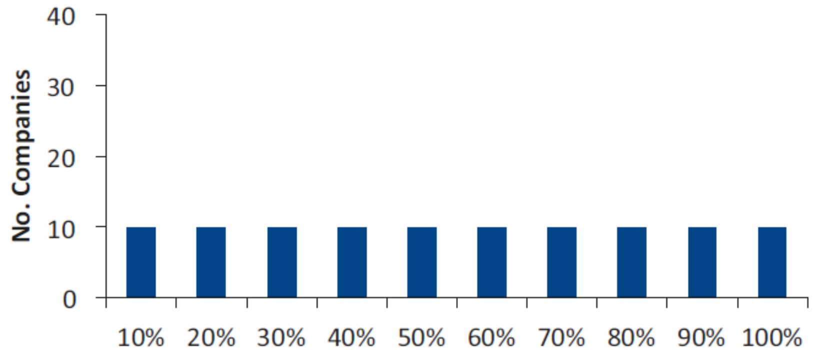
Bootstrap iterations are performed:

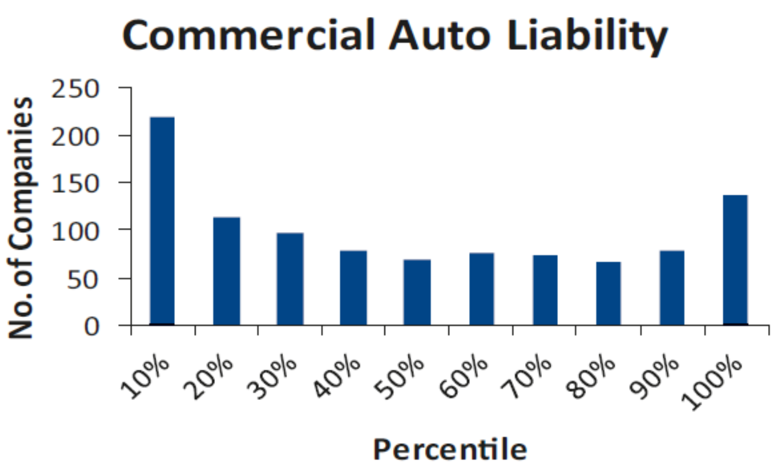
- Resample [with replacement] residuals from the original data
- For each resampled set of residuals, recreate a triangle of loss amounts
- Estimate the chain ladder reserve from each pseudo-triangle
- Include random (process variance) around the chain ladder estimate

Key assumption: each cell is an independent draw from the residuals

Past Back-Testing of Bootstrap (Leong, et al.)

Figure 4. Ideal histogram of percentiles





Back-testing of the bootstrapping method has been performed previously by Leong (2014) and Shapland (2019).

They have compared actual (hold-out) loss outcomes to the quantiles estimated by the bootstrap.

Ideally the actual results would fall evenly across predicted deciles.

In Schedule P examples, more actual losses fell into the tails of the bootstrap ranges than would have been predicted. This indicates that the bootstrap tends to underestimate the range of outcomes.

Data source for back-testing and potential data issues

Munich RE

CAS website schedule P data

- Net paid L&ALAE as of 12-120 months for accident years 1988-1997
- By line of business and by companies
  - CAL (158 companies)
  - OL (239 companies)
  - WC (132 companies)

The CAS website schedule P data is a good toy example, but shows some data problems needing scrubbing

- Negative development (even negative cumulative payments!)
- Operational changes such as LPT or commutation
- Highly skewed data

Data Cleaning

Munich RE

- Excluding companies with fewer than 10 accident years of data
- Excluding companies with data showing obvious abnormality
- No. of companies remained: CAL - 89, OL – 125, WC - 59

38997	Nissan Fire & Marine Ins Co Us Br									Other Liability
AY	12	24	36	48	60	72	84	96	108	120
1988	311	311	311	311	311	311	311	311	311	311
1989	458	457	457	457	457	457	457	457	457	457
1990	81	81	81	81	81	81	81	81	81	81
1991	122	122	122	122	122	122	122	122	122	122
1992	78	78	78	78	78	78	78	78	78	78
1993	112	112	112	112	112	112	112	112	112	112
1994	180	180	180	180	180	180	180	180	180	180
1995	182	182	182	182	182	182	182	182	182	182
1996	178	178	178	178	178	178	178	178	178	178
1997	149	149	149	149	149	149	149	149	149	149

← **No Development**

86	Allstate Ins Co Grp									Workers Compensation
AY	12	24	36	48	60	72	84	96	108	120
1988	70,571	155,905	220,744	251,595	274,156	287,676	298,499	304,873	321,808	325,322
1989	66,547	136,447	179,142	211,343	231,430	244,750	254,557	270,059	273,873	277,574
1990	52,233	133,370	178,444	204,442	222,193	232,940	253,337	256,788	261,166	263,000
1991	59,315	128,051	169,793	196,685	213,165	234,676	239,195	245,499	247,131	248,319
1992	39,991	89,873	114,117	133,003	154,362	159,496	164,013	166,212	167,397	168,844
1993	19,744	47,229	61,909	85,099	87,215	88,602	89,444	89,899	90,446	90,686
1994	20,379	46,773	88,636	91,077	92,583	93,346	93,897	94,165	94,558	94,730
1995	18,756	84,712	87,311	89,200	90,001	90,247	90,687	91,068	91,001	91,161
1996	42,609	44,916	46,981	47,899	48,583	49,109	49,442	49,073	49,161	49,255
1997	691	2,085	2,795	2,866	2,905	2,909	2,908	2,909	2,909	2,909

← **Pooling / Loss Portfolio Transfer ?**

Correlation

Munich RE

The bootstrap method includes an assumption that all cells in the triangle are independent and therefore can be resampled independently.

This assumption can be relaxed by introducing a correlation structure.

And so…

Rho, rho, rho your boot. ♪♪♪♫

Correlation

Munich RE

Independence

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Exchangeable

$$\begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

Unstructured

$$\begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{bmatrix}$$

Autoregressive AR(1)

$$\begin{bmatrix} 1 & \rho^1 & \rho^2 & \rho^3 \\ \rho^1 & 1 & \rho^1 & \rho^2 \\ \rho^2 & \rho^1 & 1 & \rho^1 \\ \rho^3 & \rho^2 & \rho^1 & 1 \end{bmatrix}$$

Correlation in GEE

Generalized Estimating Equations (GEE) is an expansion of GLM to allow for correlation to be included.

Several types of correlation structures are commonly used.

The correlation is ideally estimated from the data itself but can alternatively be “fixed” by the user.

Correlation

The AR(1) autoregressive correlation structure corresponds to a mean-reverting random walk.

$$X_t = \phi \cdot X_{t-1} + C + \varepsilon_t$$

This is a simple time-series structure used, for example, to model inflation.

Possible calendar year effects:

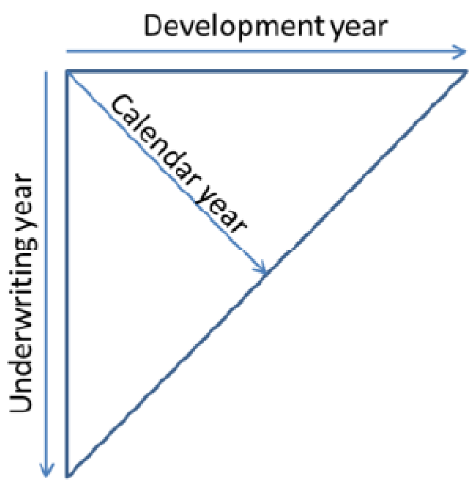
1) Economic inflation
2) Social inflation
3) Shock events such as recent COVID pandemic
4) Changes in claims handling or settlement strategy

Correlation

Munich RE

Calendar year effects operate from one diagonal to another.

We would like losses in the same calendar year to be more correlated with each other than with loss in more “distant” diagonals.



Underwriting year	Development year 1	2	3	4	5	6	7	8	9	10
2005	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
2006	2006	2007	2008	2009	2010	2011	2012	2013	2014	
2007	2007	2008	2009	2010	2011	2012	2013	2014		
2008	2008	2009	2010	2011	2012	2013	2014			
2009	2009	2010	2011	2012	2013	2014				
2010	2010	2011	2012	2013	2014					
2011	2011	2012	2013	2014						
2012	2012	2013	2014							
2013	2013	2014								
2014	2014									

Source for graphic: Lloyd’s Claims Inflation Study 2014

https://assets.lloyds.com/assets/pdf-claims-inflation-discussion-document-mg-20141128/1/Claims-Inflation-Discussion-Document-MG-20141128.pdf

Correlation

The correlation matrix has a row for each “cell” of the triangle. Losses paid in the same calendar year are more strongly correlated. Correlation decreases over time.

Correlation Matrix by Calendar Year

$\rho = 0.5$

AY			1988	1989	1990	1991	1988	1989	1990	1988	1989	1988
	Age		48	36	24	12	36	24	12	24	12	12
		CY	1991	1991	1991	1991	1990	1990	1990	1989	1989	1988
1988	48	1991	1	0.5	0.5	0.5	0.25	0.25	0.25	0.125	0.125	0.0625
1989	36	1991	0.5	1	0.5	0.5	0.25	0.25	0.25	0.125	0.125	0.0625
1990	24	1991	0.5	0.5	1	0.5	0.25	0.25	0.25	0.125	0.125	0.0625
1991	12	1991	0.5	0.5	0.5	1	0.25	0.25	0.25	0.125	0.125	0.0625
1988	36	1990	0.25	0.25	0.25	0.25	1	0.5	0.5	0.25	0.25	0.125
1989	24	1990	0.25	0.25	0.25	0.25	0.5	1	0.5	0.25	0.25	0.125
1990	12	1990	0.25	0.25	0.25	0.25	0.5	0.5	1	0.25	0.25	0.125
1988	24	1989	0.125	0.125	0.125	0.125	0.25	0.25	0.25	1	0.5	0.25
1989	12	1989	0.125	0.125	0.125	0.125	0.25	0.25	0.25	0.5	1	0.25
1988	12	1988	0.0625	0.0625	0.0625	0.0625	0.125	0.125	0.125	0.25	0.25	1

AY: Accident Year
CY: Calendar Year

Correlation

Munich RE

Using this correlation structure with a Gaussian copula allows for simulation of correlated random variables.

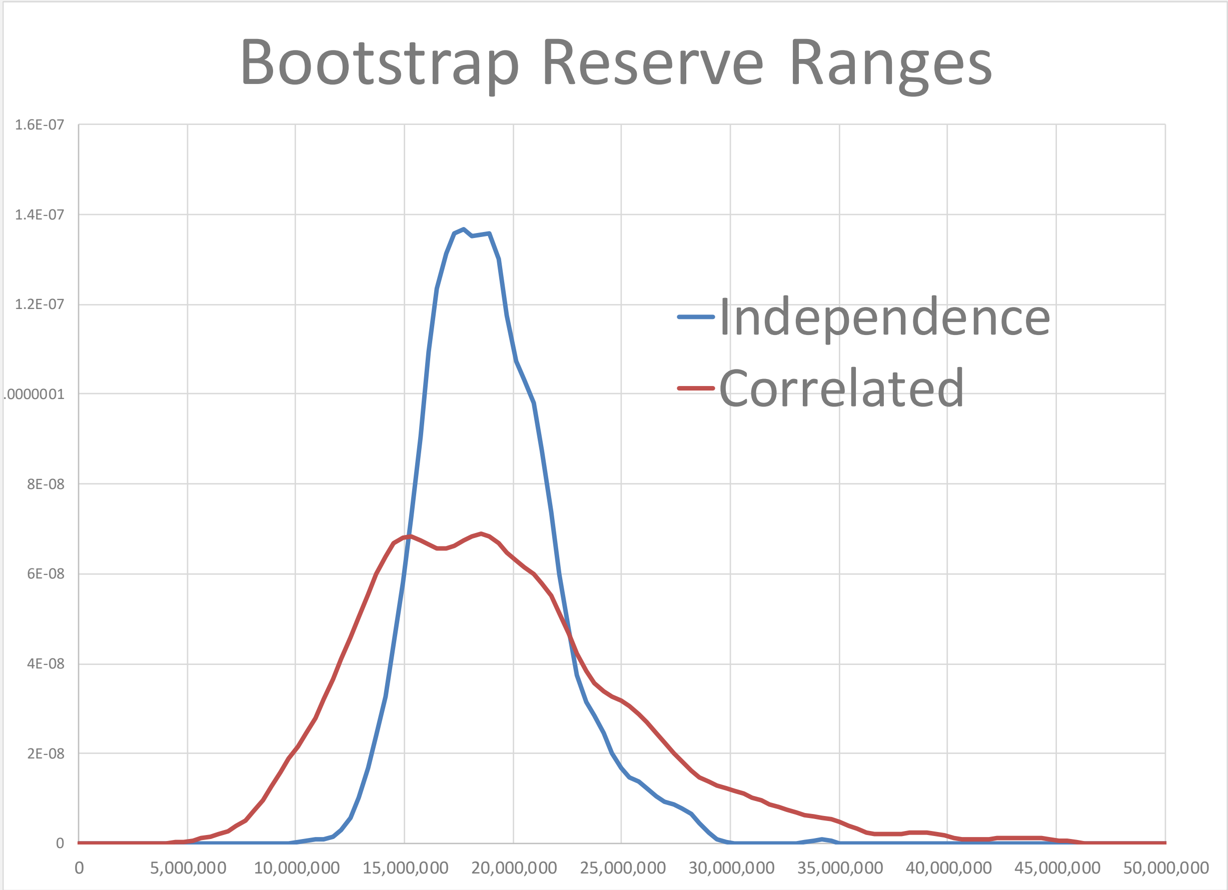
This is illustrated with “heat map” showing diagonal effects.

Simulated Random Numbers (between 0 and 1)

	12	24	36	48	60	72	84	96	108	120
2011	0.98767	0.98965	0.97728	0.57317	0.41200	0.22594	0.70100	0.94418	0.93340	0.84715
2012	0.50823	0.98389	0.73872	0.57661	0.38135	0.58938	0.60015	0.69804	0.96441	
2013	0.95664	0.58871	0.20957	0.16938	0.70855	0.91467	0.96721	0.80541		
2014	0.69993	0.37639	0.48079	0.72381	0.75887	0.92542	0.70153			
2015	0.11218	0.30808	0.88397	0.81062	0.77125	0.90142				
2016	0.13427	0.65575	0.42325	0.63180	0.88701					
2017	0.64204	0.84254	0.90658	0.87774						
2018	0.86747	0.88416	0.96448							
2019	0.70618	0.90368								
2020	0.85393									

Result – Effect of introducing a correlation structure

Munich RE



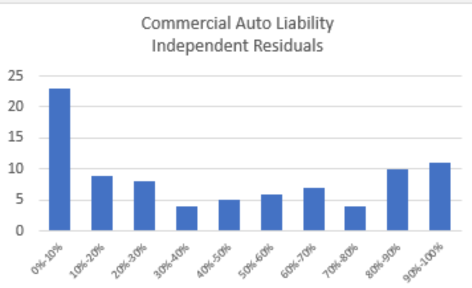
We can run many simulations – perhaps 10,000 – to produce a range of possible outcomes of the final losses.

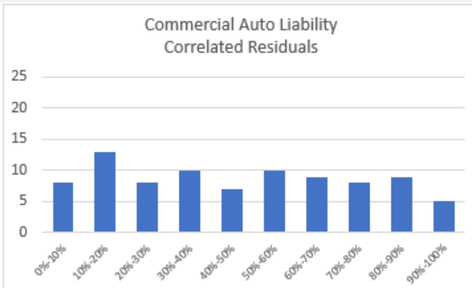
When actual losses emerge, we can see where the losses fell within the range of this estimated distribution.

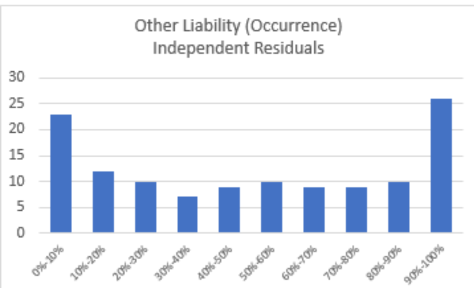
We can also see how the reserve range is expanded when correlation is included.

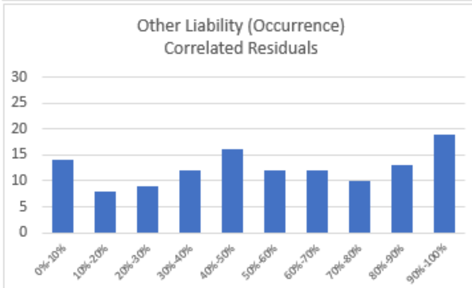
For the US industry data, we can see how this expanded range improves the prediction accuracy.

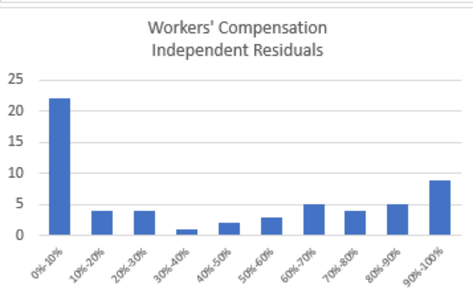
Result – Effect of introducing a correlation structure

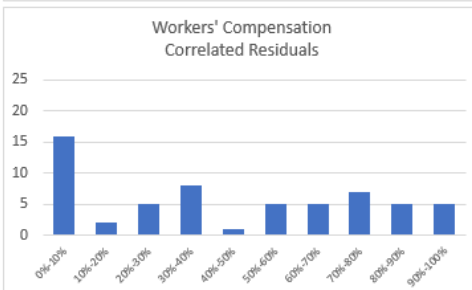












The outcomes previously in the tails of the reserve range are now within the middle deciles.

Ideally, a correct model would have about 10% of the actual outcomes in each decile. The “bathtub” shape, with more outcomes in the tails indicates that the reserve range from the bootstrap with the independence assumption is too narrow. Including a correlation structure brings it closer to a uniform distribution.

Conclusions & Observations

Munich RE

- Bootstrapping models are popular for estimating reserve ranges, but they often show narrow ranges.
- We were able to replicate past studies to confirm this.

- One reason for the narrow range is the independence assumption when resampling.

- We can include correlation in the resampling via a copula.
- This is calibrated on the US industry data sample.

Resources

Peter England, Richard Verrall "Analytic and bootstrap estimate of prediction errors in claims reserving"

Insurance: Mathematics and Economics 25 (1999) 281-293

Peter England "Addendum to 'Analytic and bootstrap estimate of prediction errors in claims reserving'"

Insurance: Mathematics and Economics 31 (2002) 461-466

Jessica Leong, Shaun Wang, Han Chen, "Back-Testing the ODP Bootstrap of the Paid Chain-Ladder Model with Actual Historical Claims Data"

CAS Variance Journal, Volume 8 / Issue 2 (2014) 182-203

David Clark, Hang Ding, Lianmin Zhou "Making Bootstrap Reserve Ranges More Reasonable"

CAS eForum, Fall 2022

Appendix: Summary of Bootstrapping Procedure

1) STAGE 1: Set up model
    a. Estimate chain ladder ultimate losses by year
    b. Create fitted triangle using chain ladder ultimate losses and average development pattern
    c. Estimate variance/mean "dispersion" parameter
    d. Estimate expected mean and variance values for each cell of the incremental triangle
2) STAGE 2: Bootstrap Iterations (repeated multiple times)
    a. Resample [***independent***] random variables for each cell of the triangle
    b. Create pseudo-triangle
    c. Estimate chain ladder unpaid losses from the pseudo-triangle (bootstrap)
    d. Store bootstrap result
3) STAGE 3: Show range of bootstrap results
    a. Calculate standard deviations and quantiles
    b. If possible, compare actual results to the range

Appendix: Summary of Correlated Bootstrapping Procedure

Munich RE

1) STAGE 1: Set up model
   a. Estimate chain ladder ultimate losses by year
   b. Create fitted triangle using chain ladder ultimate losses and average development pattern
   c. Estimate variance/mean "dispersion" parameter
   d. Estimate expected mean and variance values for each cell of the incremental triangle
   e. Calculate Cholesky decomposition of correlation matrix
2) STAGE 2: Bootstrap Iterations (repeated multiple times)
   a. Resample [***correlated***] random variables for each cell of the triangle
   b. Create pseudo-triangle
   c. Estimate chain ladder unpaid losses from the pseudo-triangle (bootstrap)
   d. Store bootstrap result
3) STAGE 3: Show range of bootstrap results
   a. Calculate standard deviations and quantiles
   b. If possible, compare actual results to the range

Imprint

Munich RE

Munich Reinsurance

MRAS – Pricing and Underwriting Services