Stochastic Loss Reserving: A New Perspective from a Dirichlet Model

Peng Shi, University of Wisconsin-Madison

2020 Casualty Loss Reserve Seminar

Joint work with

Karthik Sriram, Indian Institute of Management Ahmedabad

Outline

Outline

Background

Loss reserving is a classical actuarial problem:

- Balance sheet item $\rightarrow$ solvency
- Ratemaking $\rightarrow$ profitability

Background

Loss reserving is a classical actuarial problem:

- Balance sheet item $\rightarrow$ solvency
- Ratemaking $\rightarrow$ profitability

Two components

- To estimate the outstanding liability
- To quantify the associated variability

Motivation

Two prominent macro-level stochastic loss reserving methods:

- Chain-Ladder method: development factors or age-to-age ratios
- Bornhuetter-Ferguson method: expected ultimate and unpaid quota

Motivation

Two prominent macro-level stochastic loss reserving methods:

- Chain-Ladder method: development factors or age-to-age ratios
- Bornhuetter-Ferguson method: expected ultimate and unpaid quota

Connections and disconnections:

- BF can be viewed as a credibility weighted average of CL and expected ultimate loss
- Different data generating process for the same run-off triangle

Contribution

We propose a new stochastic loss reserving method based on a Dirichlet distribution

Contribution

We propose a new stochastic loss reserving method based on a Dirichlet distribution

New perspective to view the relation between CL and BF:

- Both CL and BF prediction are derived from the same stochastic model
- The choice between the two depends on the type of information available
- It becomes an inference problem rather than model selection

Contribution

We propose a new stochastic loss reserving method based on a Dirichlet distribution

New perspective to view the relation between CL and BF:

- Both CL and BF prediction are derived from the same stochastic model
- The choice between the two depends on the type of information available
- It becomes an inference problem rather than model selection

Nice Properties:

- Same age-to-age factors as CL
- Generalize the credibility weight in the BF

Outline

Data Structure

Micro-level reserving: triangular aggregate data

- A triangle consists of m accident years and n development years $(m > n)$
- Let X_{ij} denote incremental paid loss in accident year i development year j
- We work with Y_{ij}/E_i, where E_i is the exposure for accident year i

Micro-level reserving: triangular aggregate data

- A triangle consists of m accident years and n development years ($m > n$)
- Let X_{ij} denote incremental paid loss in accident year i development year j
- We work with Y_{ij}/E_i, where E_i is the exposure for accident year i

	Accident Year	Development Year 1	2	$\cdots$	$n-1$	n
Fully developed	1	Y_{11}	Y_{12}	$\cdots$	Y_{1n-1}	Y_{1n}
	$\vdots$	$\vdots$				$\vdots$
	$m-n$	$Y_{m-n,1}$	$Y_{m-n,2}$	$\cdots$	$Y_{m-n,n-1}$	$Y_{m-n,n}$
Run-off Triangle	$m-n+1$	$Y_{m-n+1,1}$	$Y_{m-n+1,2}$	$\cdots$	$Y_{m-n+1,n-1}$	$Y_{m-n+1,n}$
	$m-n+2$	$Y_{m-n+2,1}$	$Y_{m-n+2,2}$	$\cdots$	$Y_{m-n+2,n-1}$	
	$\vdots$	$\vdots$		$\iddots$		
	$m-1$	$Y_{m-1,1}$	$Y_{m-1,2}$			
	m	$Y_{m,1}$				

An Example

Year	Accident Year	Earned Premium	Incremental Losses by Development Year 1	2	3	4	5	6	7	8	9	10
1	1989	1,65,339	41,891	32,156	20,520	15,256	8,170	5,317	3,415	2,504	1,967	940
2	1990	1,68,293	44,050	37,311	22,339	14,356	8,419	6,258	3,545	2,981	1,468	1,265
3	1991	1,83,529	47,778	39,354	21,232	16,132	10,632	6,754	4,311	2,407	1,620	993
4	1992	1,92,991	49,191	42,325	22,731	16,959	11,056	6,972	4,317	2,431	2,016	1,106
5	1993	2,22,666	47,035	38,662	20,081	15,923	10,621	6,266	3,552	2,744	1,513	1,308
6	1994	2,40,844	51,538	33,518	19,964	16,713	11,076	7,526	4,835	4,450	2,273	2,155
7	1995	2,58,703	46,934	31,827	21,236	15,846	11,288	6,317	5,615	4,261	2,798	2,150
8	1996	2,37,131	43,432	32,768	21,697	16,150	10,230	8,056	6,250	4,455	3,417	2,421
9	1997	2,08,179	38,915	28,463	19,494	13,361	10,211	7,176	5,401	3,453	2,551	1,844
10	1998	1,69,361	34,596	28,089	16,409	13,813	8,966	6,333	4,913	4,196	2,670	
11	1999	1,50,912	32,580	24,468	17,672	13,418	7,881	6,616	5,246	3,655		
12	2000	1,75,101	39,248	30,647	19,059	14,599	10,220	7,725	5,126			
13	2001	1,94,483	42,433	32,981	21,082	17,274	12,151	8,309				
14	2002	2,22,002	45,309	36,483	25,777	18,746	12,266					
15	2003	2,44,749	54,589	41,491	26,295	19,207						
16	2004	2,79,994	59,399	47,007	26,169							
17	2005	3,13,808	68,185	54,385								
18	2006	3,41,973	66,827									

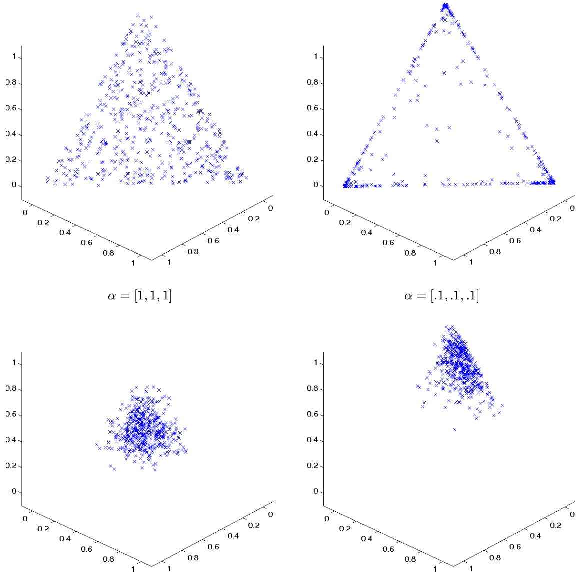
Consider an example of $K = 3$:

$$f(p_1, p_2, p_3) = \frac{\Gamma(p_1 + p_2 + p_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1},$$

where $p_1 + p_2 + p_3 = 1$.

The support is the hyperplane $p_1 + p_2 + p_3 = 1$, and parameters $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ determines the values of p_1, p_2, p_3

Dirichlet Distribution



$\alpha = [1, 1, 1]$ $\alpha = [.1, .1, .1]$

$\alpha = [10, 10, 10]$ $\alpha = [2, 5, 15]$

The Dirichlet distribution can be thought of as a distribution over probability mass functions of length K:

The Dirichlet distribution can be thought of as a distribution over probability mass functions of length K:

Let $P = (P_1, \ldots, P_K)$ be a random vector with $K \geq 2$ components. Then P is said to follow the Dirichlet distribution of order $K \geq 2$, which we denote by $P = (P_1, \ldots, P_K) \sim \text{Dir}(\alpha_1, \ldots, \alpha_K)$, if its density is given by:

$$f(p; \alpha_1, \ldots, \alpha_K) = \frac{\Gamma\left(\sum_{k=1}^K \alpha_k\right)}{\prod_{k=1}^K \Gamma(\alpha_k)} \prod_{k=1}^K p_k^{\alpha_k - 1},$$

where $\alpha_1, \ldots, \alpha_K$ are parameters of the distribution with $\alpha_k > 0$ for each k, and $p = (p_1, \ldots, p_K)$ is on the $(K-1)$-dimensional probability simplex, i.e. $\sum_{k=1}^K p_k = 1$ and $p_k \geq 0$ for $k = 1, \ldots, K$.

Reserving Model

We propose to model the incremental loss ratios ($Y_{ij} = X_{ij}/E_i$) using a scaled Dirichlet distribution.

For accident year i, we assume:

$$\left(\frac{Y_{i1}}{\phi_i}, \cdots, \frac{Y_{in}}{\phi_i}, 1 - \frac{\sum_{j=1}^{n} Y_{ij}}{\phi_i}\right) \sim \text{Dir}(a_1, \ldots, a_n, b_n), \text{ with } 0 < \sum_{j=1}^{n} Y_{ij} < \phi_i,$$

where ϕ_i, $a_1, \ldots, a_n$, and b_n are parameters to be estimated.

Define cumulative paid loss $S_{i,1:n} = Y_{i1} + \ldots Y_{in}$ and $a_0 = a_1 + \ldots + a_n$. For interpretation, we show an equivalent of representation of model as:

$$\begin{cases} \left(\frac{Y_{i1}}{S_{i,1:n}}, \frac{Y_{i2}}{S_{i,1:n}}, \cdots, \frac{Y_{in}}{S_{i,1:n}} \right) \Big| S_{i,1:n} \sim \text{Dir}(a_1, \ldots, a_n) \\ \frac{S_{i,1:n}}{\phi_i} \sim \text{Beta}(a_0, b_n) \end{cases}$$

Define cumulative paid loss $S_{i,1:n} = Y_{i1} + \ldots Y_{in}$ and $a_0 = a_1 + \ldots + a_n$. For interpretation, we show an equivalent of representation of model as:

$$\begin{cases} \left(\frac{Y_{i1}}{S_{i,1:n}}, \frac{Y_{i2}}{S_{i,1:n}}, \cdots, \frac{Y_{in}}{S_{i,1:n}} \right) \Big| S_{i,1:n} \sim \text{Dir}(a_1, \ldots, a_n) \\ \frac{S_{i,1:n}}{\phi_i} \sim \text{Beta}(a_0, b_n) \end{cases}$$

The Dirichlet model is about allocation given $S_{i,1:n}$:

$$x_1 = \frac{Y_{i1}}{S_{i,1:n}}, x_2 = \frac{Y_{i2}}{S_{i,1:n}}, \cdots, x_{n-1} = \frac{Y_{in-1}}{S_{i,1:n}}$$

$$x_n = \frac{Y_{in}}{S_{i,1:n}} = 1 - \sum_{j=1}^{n-1} x_j$$

Define cumulative paid loss $S_{i,1:n} = Y_{i1} + \ldots Y_{in}$ and $a_0 = a_1 + \ldots + a_n$. For interpretation, we show an equivalent of representation of model as:

$$\begin{cases} \left(\frac{Y_{i1}}{S_{i,1:n}}, \frac{Y_{i2}}{S_{i,1:n}}, \cdots, \frac{Y_{in}}{S_{i,1:n}} \right) \Big| S_{i,1:n} \sim \mathrm{Dir}(a_1, \ldots, a_n) \\ \frac{S_{i,1:n}}{\phi_i} \sim \mathrm{Beta}(a_0, b_n) \end{cases}$$

The Dirichlet model is about allocation given $S_{i,1:n}$:

$$x_1 = \frac{Y_{i1}}{S_{i,1:n}}, x_2 = \frac{Y_{i2}}{S_{i,1:n}}, \cdots, x_{n-1} = \frac{Y_{in-1}}{S_{i,1:n}}$$

$$x_n = \frac{Y_{in}}{S_{i,1:n}} = 1 - \sum_{j=1}^{n-1} x_j$$

Mathematical properties of Dirichlet distribution implies that past allocation does not inform future allocation.

The model implies both past and future allocations given cumulative payments:

$$\left(\frac{Y_{i1}}{S_{i,1:k}},\cdots,\frac{Y_{ik}}{S_{i,1:k}}\right)|S_{i,1:k} \sim \mathrm{Dir}(a_1,\ldots,a_k)$$
$$\left(\frac{Y_{ik+1}}{\phi_i - S_{i,1:k}},\cdots,\frac{Y_{in}}{\phi_i - S_{i,1:k}},\frac{\phi_i - S_{i,1:n}}{\phi_i - S_{i,1:k}}\right)|S_{i,1:k} \sim \mathrm{Dir}(a_{k+1},\ldots,a_n,b_n).$$

Properties

Age-to-age factor and unpaid percentage:

In CL and BF:

$$\gamma_{k:k+1} = \frac{\mathrm{E}(S_{i,1:k+1})}{\mathrm{E}(S_{i,1:k})}, \quad \text{and} \quad \eta_k = \frac{\mathrm{E}(S_{i,1:k})}{\mathrm{E}(S_{i,1:n})}.$$

Properties

Age-to-age factor and unpaid percentage:

In CL and BF:

$$\gamma_{k:k+1} = \frac{\mathrm{E}(S_{i,1:k+1})}{\mathrm{E}(S_{i,1:k})}, \quad \text{and} \quad \eta_k = \frac{\mathrm{E}(S_{i,1:k})}{\mathrm{E}(S_{i,1:n})}.$$

Dirichlet model:

$$\gamma_{k:k+1} = \frac{a_1 + \cdots + a_{k+1}}{a_1 + \cdots + a_k}$$
$$\eta_k = \frac{a_1 + \cdots + a_k}{a_1 + \cdots + a_n}$$

Properties

Reserve prediction at development year k is

$$\widehat{R}_i = \widehat{S}_{i,1:n} - S_{i,1:k} = \mathrm{E}(S_{i,1:n}|S_{i,1:k}) - S_{i,1:k}$$

Properties

Reserve prediction at development year k is

$$\widehat{R}_i = \widehat{S}_{i,1:n} - S_{i,1:k} = \mathrm{E}(S_{i,1:n}|S_{i,1:k}) - S_{i,1:k}$$

Recall the well-known results:

$$\begin{aligned}
\widehat{R}_i^{EX} &= \mathrm{E}(S_{i,1:n}) - S_{i,1:k} \\
\widehat{R}_i^{CL} &= \left(\prod_{j=k}^{n-1} \gamma_{j:j+1} - 1\right) S_{i,1:k} \\
\widehat{R}_i^{BF} &= \eta_k \widehat{R}_i^{CL} + (1-\eta_k)\widehat{R}_i^{EX}
\end{aligned}$$

Properties

At development year k, the Dirichlet model predicts:

$$\widehat{R}_i^D = v_k \widehat{R}_i^{CL} + (1 - v_k)\widehat{R}_i^{EX},$$

where

$$v_k = \left\{ \frac{\mathrm{CV}(S_{i,1:n})}{\mathrm{CV}(S_{i,1:k})} \right\}^2 = \frac{\mathrm{Var}(S_{i,1:n})}{\mathrm{Var}(S_{i,1:k})} \left\{ \frac{\mathrm{E}(S_{i,1:k})}{\mathrm{E}(S_{i,1:n})} \right\}^2$$

Properties

At development year k, the Dirichlet model predicts:

$$\widehat{R}_i^D = v_k \widehat{R}_i^{CL} + (1 - v_k)\widehat{R}_i^{EX},$$

where

$$v_k = \left\{ \frac{\mathrm{CV}(S_{i,1:n})}{\mathrm{CV}(S_{i,1:k})} \right\}^2 = \frac{\mathrm{Var}(S_{i,1:n})}{\mathrm{Var}(S_{i,1:k})} \left\{ \frac{\mathrm{E}(S_{i,1:k})}{\mathrm{E}(S_{i,1:n})} \right\}^2$$

Compare with BF prediction:

$$\widehat{R}_i^{BF} = \eta_k \widehat{R}_i^{CL} + (1 - \eta_k)\widehat{R}_i^{EX}$$

where

$$\eta_k = \frac{\mathrm{E}(S_{i,1:k})}{\mathrm{E}(S_{i,1:n})}$$

Properties

At development year k, the Dirichlet model predicts:

$$\widehat{R}_i^D = v_k \widehat{R}_i^{CL} + (1 - v_k)\widehat{R}_i^{EX},$$

where

$$v_k = \left\{ \frac{\mathrm{CV}(S_{i,1:n})}{\mathrm{CV}(S_{i,1:k})} \right\}^2 = \frac{\mathrm{Var}(S_{i,1:n})}{\mathrm{Var}(S_{i,1:k})} \left\{ \frac{\mathrm{E}(S_{i,1:k})}{\mathrm{E}(S_{i,1:n})} \right\}^2$$

Compare with BF prediction:

$$\widehat{R}_i^{BF} = \eta_k \widehat{R}_i^{CL} + (1 - \eta_k)\widehat{R}_i^{EX}$$

where

$$\eta_k = \frac{\mathrm{E}(S_{i,1:k})}{\mathrm{E}(S_{i,1:n})}$$

We note the limiting case: $v_k \to \eta_k$ when $\frac{\sum_{j=k+1}^{n} a_j}{b_n} \to 0$

Outline

Estimation

To predict the outstanding payments, one needs the unknowns

$$\boldsymbol{\theta} = (a_1, a_2, \ldots, a_n, b_n, \phi_1, \phi_2, \ldots, \phi_m).$$

Two types of data:

- Fully developed accident year: for $1 \leq i \leq m-n$,

$$\left(\frac{Y_{i1}}{\phi_i}, \cdots, \frac{Y_{in}}{\phi_i}, 1 - \frac{S_{i,1:n}}{\phi_i}\right) \sim \mathrm{Dir}(a_1, \ldots, a_n, b_n)$$

To predict the outstanding payments, one needs the unknowns

$$\boldsymbol{\theta} = (a_1, a_2, \ldots, a_n, b_n, \phi_1, \phi_2, \ldots, \phi_m).$$

Two types of data:

- Fully developed accident year: for $1 \leq i \leq m - n$,

$$\left(\frac{Y_{i1}}{\phi_i}, \cdots, \frac{Y_{in}}{\phi_i}, 1 - \frac{S_{i,1:n}}{\phi_i}\right) \sim \mathrm{Dir}(a_1, \ldots, a_n, b_n)$$

- Not fully developed accident year: for $m - n + 1 \leq i \leq m$,

$$\left(\frac{Y_{i1}}{\phi_i}, \cdots, \frac{Y_{im+1-i}}{\phi_i}, 1 - \frac{S_{i,1:m+1-i}}{\phi_i}\right) \sim \mathrm{Dir}\left(a_1, \ldots, a_n, a_0 + b_n - \sum_{j=1}^{m+1-i} a_j\right)$$

Estimation

Frequentst estimation:

- Likelihood-based estimation
- When $a_0/(a_0+1) \approx 1$, we obtain CL estimates

Estimation

Frequentst estimation:

- Likelihood-based estimation
- When $a_0/(a_0+1) \approx 1$, we obtain CL estimates

Bayesian estimation:

- It allows for a hierarchical extension: $\phi_1, \phi_2, \ldots, \phi_n \overset{iid}{\sim} uniform(0, \phi)$ with a flat hyper prior $p(\phi) \propto 1$ for $\phi \in (0, \infty)$
- Expert knowledge on development pattern could be incorporated into inference via informative priors
- It is straightforward to blend in collateral information in the model inference

Estimation with a variance constraint: $b_n > a_0 (= a_1 + \ldots + a_n)$

This constraint ensures $\mathrm{Var}(S_{i,1:k})$ increases in k:

$$\mathrm{Var}(S_{i,1:k}) = \frac{\left(\sum_{j=1}^{k} a_j\right)\left(a_0 + b_n - \sum_{j=1}^{k} a_j\right)}{(a_0 + b_n)^2 (a_0 + b_n + 1)} \phi_i^2 .$$

This constraint mimics a condition implied by the Chain-Ladder method: the variance in the cumulative paid loss ratio is increasing by development age.

Outline

Data

Data from the NAIC schedule P:

- We examine the paid losses for worker's compensation

Data

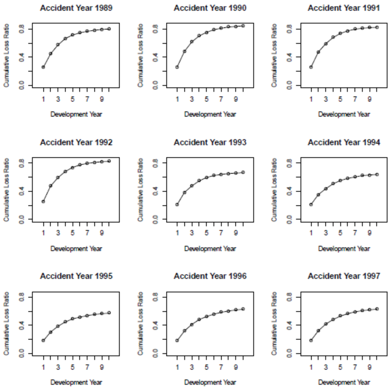
Data from the NAIC schedule P:

- We examine the paid losses for worker's compensation
- The data of each individual company contain incremental paid losses for 18 accident years ($m = 18$) from to 1989 to 2006, and for each accident year, losses are developed for the period of 10 years ($n = 10$).

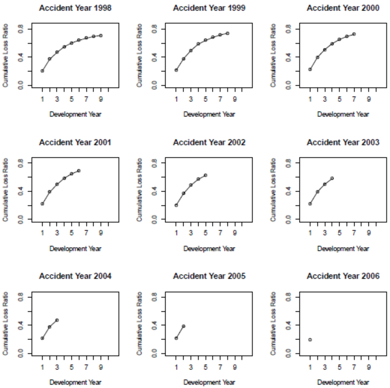
Data

Data from the NAIC schedule P:

- We examine the paid losses for worker's compensation
- The data of each individual company contain incremental paid losses for 18 accident years ($m = 18$) from to 1989 to 2006, and for each accident year, losses are developed for the period of 10 years ($n = 10$).
- Data are split into two parts
  - Upper triangle is used to develop model
  - Lower triangle is for validation



(a) Fully developed claims



(b) Partially developed claims

Frequentist approach: MLE Dirichlet $\hat{a}_0/(\hat{a}_0+1) \approx 1$

Development Factor	10 accident years				18 accident years			
	Dirichlet Model		Mack Chain-Ladder		Dirichlet Model		Mack Chain-Ladder	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\gamma_{1:2}$	1.778	0.0138	1.779	0.0087	1.775	0.0175	1.781	0.013
$\gamma_{2:3}$	1.280	0.0055	1.281	0.0082	1.269	0.0072	1.269	0.006
$\gamma_{3:4}$	1.169	0.0040	1.169	0.0039	1.161	0.0049	1.160	0.003
$\gamma_{4:5}$	1.098	0.0029	1.098	0.0024	1.091	0.0034	1.090	0.003
$\gamma_{5:6}$	1.066	0.0025	1.066	0.0012	1.057	0.0026	1.057	0.002
$\gamma_{6:7}$	1.046	0.0023	1.046	0.0019	1.037	0.0020	1.037	0.003
$\gamma_{7:8}$	1.033	0.0021	1.033	0.0026	1.025	0.0017	1.025	0.002
$\gamma_{8:9}$	1.021	0.0020	1.022	0.0013	1.016	0.0014	1.016	0.002
$\gamma_{9:10}$	1.014	0.0023	1.014	0.0008	1.011	0.0012	1.010	0.001

Comparison of reserve prediction:

		MLE Dirichlet Model (18 accident years data)				Hierarchical Bayes Dirichlet Model (18 accident years data)			
Acc Year	Actual Loss Ratio	95% Interval	|Actual-Predicted|	Interval contains actual	Interval Length	95% Interval	|Actual-Predicted|	Interval contains actual	Interval Length
1997	0.629	[0.629,0.629]	0.000	1	0.000	[0.629,0.629]	0.000	1	0.000
1998	0.719	[0.709,0.720]	0.004	1	0.011	[0.718,0.719]	0.000	1	0.001
1999	0.763	[0.749,0.765]	0.005	1	0.016	[0.763,0.765]	0.001	1	0.002
2000	0.767	[0.747,0.772]	0.006	1	0.025	[0.764,0.768]	0.001	1	0.004
2001	0.765	[0.736,0.767]	0.012	1	0.031	[0.757,0.763]	0.005	0	0.006
2002	0.741	[0.698,0.739]	0.022	0	0.041	[0.722,0.730]	0.015	0	0.008
2003	0.722	[0.699,0.752]	0.005	1	0.053	[0.723,0.735]	0.007	0	0.012
2004	0.705	[0.657,0.723]	0.014	1	0.066	[0.682,0.699]	0.014	0	0.017
2005	0.729	[0.676,0.771]	0.006	1	0.095	[0.715,0.743]	0.001	1	0.028
2006	0.629	[0.576,0.708]	0.013	1	0.132	[0.624,0.667]	0.016	1	0.043
Average			0.009	0.900	0.047		0.006	0.600	0.012

Reserve prediction with variance constraints:

		MLE Dirichlet Model (18 accident years data)				Hierarchical Bayes Dirichlet Model (18 accident years data)			
Acc Year	Actual Loss Ratio	95% Interval	|Actual-Predicted|	Interval contains actual	Interval Length	95% Interval	|Actual-Predicted|	Interval contains actual	Interval Length
1997	0.629	[0.629,0.629]	0.000	1	0.000	[0.629,0.629]	0.000	1	0.000
1998	0.719	[0.712,0.722]	0.003	1	0.010	[0.713,0.722]	0.002	1	0.009
1999	0.763	[0.752,0.768]	0.004	1	0.016	[0.754,0.769]	0.002	1	0.015
2000	0.767	[0.751,0.774]	0.006	1	0.023	[0.754,0.774]	0.004	1	0.020
2001	0.765	[0.740,0.768]	0.012	1	0.028	[0.745,0.771]	0.008	1	0.026
2002	0.741	[0.702,0.739]	0.021	0	0.037	[0.710,0.740]	0.016	0	0.030
2003	0.722	[0.705,0.754]	0.006	1	0.049	[0.712,0.751]	0.009	1	0.039
2004	0.705	[0.661,0.724]	0.013	1	0.063	[0.671,0.721]	0.010	1	0.050
2005	0.729	[0.682,0.770]	0.005	1	0.088	[0.698,0.776]	0.006	1	0.078
2006	0.629	[0.577,0.716]	0.014	1	0.139	[0.603,0.718]	0.027	1	0.115
Average			0.008	0.900	0.047		0.008	0.900	0.038

Some observations for this specific insurer:

- Both point and interval predictions are comparable between the CL and MLE Dirichlet model. Recall that $\hat{a}_0/(\hat{a}_0+1)\approx 1$.

Case Study Using One Insurer

Some observations for this specific insurer:

- Both point and interval predictions are comparable between the CL and MLE Dirichlet model. Recall that $\hat{a}_0/(\hat{a}_0+1) \approx 1$.
- Additional 8 years fully developed claims data help improve prediction in terms of coverage probability. It is not true in general.
  - The insurer has stable underwriting criterion and business mix.
  - It focuses on assigned risk market.

Case Study Using One Insurer

Some observations for this specific insurer:

- Both point and interval predictions are comparable between the CL and MLE Dirichlet model. Recall that $\hat{a}_0/(\hat{a}_0+1) \approx 1$.
- Additional 8 years fully developed claims data help improve prediction in terms of coverage probability. It is not true in general.
  - The insurer has stable underwriting criterion and business mix.
  - It focuses on assigned risk market.
- Prediction interval from the MLE is wider than the Bayesian approach, and variance constraint further improves the hierarchical model performance.

Comparative Study Using Many Insurers

We focus on large insurers. Limit to earned premium ≥ 10 million dollars. This leaves us 139 insurers in the analysis.

Comparative Study Using Many Insurers

We focus on large insurers. Limit to earned premium ≥ 10 million dollars. This leaves us 139 insurers in the analysis.

Performance of prediction is assessed using out-of-sample validation, and we compute three metrics:

- Root mean squared error (RMSE)
- Coverage probability of the 95% prediction interval (Coverage)
- Average length of the 95% prediction interval (Length)

MLE Dirichlet Model:

- The usage of additional 8 years of claims data provide worse prediction
- Variance constraint improves coverage at the price of RMSE and interval length

	10-year Triangle					
	Unconstrained			Constrained		
AY	Coverage	Length	RMSE	Coverage	Length	RMSE
1998	0.950	0.037	0.007	0.942	0.061	0.015
1999	0.871	0.084	0.052	0.914	0.111	0.067
2000	0.871	0.102	0.059	0.892	0.130	0.083
2001	0.813	0.109	0.053	0.784	0.149	0.068
2002	0.763	0.114	0.057	0.799	0.152	0.080
2003	0.791	0.118	0.050	0.784	0.157	0.075
2004	0.791	0.135	0.059	0.849	0.182	0.081
2005	0.806	0.203	0.120	0.856	0.260	0.146
2006	0.871	0.417	0.219	0.849	0.464	0.225
Overall	0.836	0.147	0.075	0.852	0.185	0.093

Bayesian Inference with informative priors:

Bayesian Inference with informative priors:

We employ prior knowledge on the expected unpaid loss ratio for the ith accident year:

$$\mathrm{E}(S_{i,m+2-i:n}|S_{i,1:m+1-i}=s_{i,1:m+1-i})$$
$$=\frac{\sum_{j=m+2-i}^{n}a_j}{\sum_{j=m+2-i}^{n}a_j+b_n}(\phi_i-s_{i,1:m+1-i}),\ i=m-n+2,\ldots,m.$$

Comparative Study Using Many Insurers

Bayesian Inference with informative priors:

We employ prior knowledge on the expected unpaid loss ratio for the ith accident year:

$$\mathrm{E}(S_{i,m+2-i:n}|S_{i,1:m+1-i}=s_{i,1:m+1-i}) \\ = \frac{\sum_{j=m+2-i}^{n} a_j}{\sum_{j=m+2-i}^{n} a_j + b_n}(\phi_i - s_{i,1:m+1-i}),\ i=m-n+2,\ldots,m.$$

Consider two different levels of uncertainty on the information:

(a)

$$\mathrm{E}(S_{i,m+2-i:n}|S_{i,1:m+1-i}) \in [0.5(CLR_i - S_{i,m+2-i:n}), 1.5(CLR_i - S_{i,m+2-i:n})]$$

(b)

$$\mathrm{E}(S_{i,m+2-i:n}|S_{i,1:m+1-i}) \in [0.9(CLR_i - S_{i,m+2-i:n}), 1.1(CLR_i - S_{i,m+2-i:n})]$$

Scenario #1: CLR is the actual cumulative loss ratio

	Scenario #1(a)			Scenario #1(b)		
AY	Coverage	Length	RMSE	Coverage	Length	RMSE
1998	1.000	0.045	0.005	1.000	0.035	0.001
1999	1.000	0.089	0.013	1.000	0.070	0.002
2000	1.000	0.116	0.019	1.000	0.088	0.003
2001	1.000	0.145	0.028	1.000	0.110	0.004
2002	1.000	0.172	0.036	1.000	0.129	0.005
2003	1.000	0.186	0.045	1.000	0.138	0.006
2004	1.000	0.247	0.062	1.000	0.182	0.007
2005	1.000	0.317	0.075	1.000	0.232	0.007
2006	1.000	0.543	0.069	1.000	0.374	0.004
Overall	1.000	0.207	0.039	1.000	0.151	0.004

Scenario #2: CLR is estimated from industry-level loss development factor Sherman and Diss (2005).

	Scenario #2(a)			Scenario #2(b)		
AY	Coverage	Length	RMSE	Coverage	Length	RMSE
1998	0.964	0.060	0.011	0.971	0.051	0.009
1999	0.942	0.105	0.024	0.950	0.087	0.018
2000	0.871	0.135	0.036	0.878	0.111	0.026
2001	0.863	0.165	0.064	0.906	0.133	0.054
2002	0.827	0.175	0.067	0.835	0.142	0.055
2003	0.842	0.182	0.062	0.914	0.147	0.043
2004	0.914	0.205	0.057	0.899	0.165	0.048
2005	0.906	0.226	0.069	0.655	0.181	0.081
2006	0.942	0.375	0.140	0.719	0.293	0.161
Overall	0.897	0.181	0.059	0.859	0.146	0.055

Scenario #3: CLR is estimated using growth curve model Clark (2003).

AY	Scenario #3(a) Coverage	Length	RMSE	Scenario #3(b) Coverage	Length	RMSE
1998	1.000	0.051	0.009	0.993	0.042	0.007
1999	0.957	0.092	0.021	0.950	0.076	0.017
2000	0.957	0.122	0.033	0.935	0.100	0.026
2001	0.914	0.146	0.054	0.892	0.119	0.048
2002	0.906	0.165	0.055	0.878	0.134	0.046
2003	0.957	0.186	0.059	0.986	0.151	0.038
2004	0.928	0.235	0.070	0.928	0.189	0.045
2005	0.914	0.314	0.115	0.935	0.250	0.091
2006	0.957	0.549	0.204	0.906	0.413	0.195
Overall	0.943	0.207	0.069	0.934	0.164	0.057

Comparison with existing methods:

- Mack Chain-Ladder
- Bootstrap Chain-Ladder
- Clark's growth curve
- GLM: Poisson, gamma, and Tweedie

Comparative Study Using Many Insurers

	Mack Chain-Ladder			Bootstrap Chain-Ladder			Clark		
AY	Coverage	Length	RMSE	Coverage	Length	RMSE	Coverage	Length	RMSE
1998	0.446	0.011	0.007	0.971	0.042	0.014	0.921	0.032	0.007
1999	0.496	0.029	0.062	0.899	0.081	0.088	0.835	0.049	0.016
2000	0.619	0.045	0.068	0.885	0.092	0.097	0.799	0.063	0.024
2001	0.712	0.061	0.056	0.799	0.091	0.064	0.770	0.075	0.047
2002	0.691	0.074	0.056	0.827	0.098	0.071	0.885	0.087	0.045
2003	0.770	0.092	0.042	0.849	0.109	0.053	0.899	0.101	0.034
2004	0.856	0.134	0.061	0.878	0.133	0.074	0.906	0.125	0.044
2005	0.871	0.186	0.116	0.871	0.196	0.144	0.871	0.174	0.090
2006	0.835	0.582	0.219	0.856	0.398	0.267	0.899	0.353	0.196
Overall	0.699	0.135	0.076	0.871	0.138	0.097	0.865	0.118	0.056
	Tweedie $p = 1$			Tweedie $p = 2$			Tweedie $1 < p < 2$		
AY	Coverage	Length	RMSE	Coverage	Length	RMSE	Coverage	Length	RMSE
1998	0.871	0.036	0.007	0.604	0.019	0.009	0.777	0.025	0.008
1999	0.813	0.066	0.062	0.511	0.042	0.056	0.669	0.049	0.053
2000	0.899	0.077	0.068	0.633	0.056	0.081	0.705	0.060	0.071
2001	0.806	0.083	0.056	0.633	0.061	0.066	0.727	0.065	0.060
2002	0.784	0.092	0.056	0.734	0.075	0.066	0.727	0.072	0.060
2003	0.885	0.102	0.042	0.806	0.093	0.060	0.784	0.084	0.050
2004	0.849	0.128	0.061	0.871	0.137	0.085	0.799	0.112	0.071
2005	0.878	0.183	0.116	0.957	0.239	0.158	0.863	0.180	0.136
2006	0.871	0.361	0.219	0.935	0.524	0.248	0.863	0.378	0.229
Overall	0.851	0.125	0.076	0.743	0.138	0.092	0.768	0.114	0.082

Outline

Conclusion

We proposed a stochastic loss reserving method based on Dirichlet distribution

- A new perspective to view CL and BF as inference issue
- Prediction from the Dirichlet model is a credibility weighted average of CL and expected payments
- Good performance was supported by comparison with existing methods using out-of-sample validation

Conclusion

We proposed a stochastic loss reserving method based on Dirichlet distribution

- A new perspective to view CL and BF as inference issue
- Prediction from the Dirichlet model is a credibility weighted average of CL and expected payments
- Good performance was supported by comparison with existing methods using out-of-sample validation

For more details, check out our paper:

Sriram, K. and Shi, P. (2020+) Stochastic loss reserving: A new perspective from a Dirichlet model, *Journal of Risk and Insurance*.

Conclusion

Thank you for your attention!!!

Contact Info:

Peng Shi can be reached by pshi@bus.wisc.edu.