Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

Ji Yao, PhD FIA

GIRO 2009, Edinburgh
6-9 October 2009

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

INTRODUCTION

- Prediction error of ultimate claim is very important
  - Understand the volatility of ultimate estimation
  - Risk management
  - Regulation
- However, all analysis on prediction error depends on a set of assumption and assumed models
  - Explicit: model setting, distribution of claim
  - Implicit: prior knowledge of triangle development?
- Need to understand how these assumption affect the results: does prediction error sensitive to these assumptions?
- Bayesian approach explicitly introduce the assumptions on prior

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

BASIC IDEA OF BAYESIAN APPROACH

- Try to understand the impact of prior knowledge
  - For example, the development factor (five years) shown in data are

    5, 2, 1.3, 1.1, 1
  - Considering two prior knowledge
    - Prior 1: 5, 2, 1.3, 1.1, 1 (exactly same as experience)
    - Prior 2: 10, 3, 1.1, 1, 1
  - Does the best estimate of ultimate claim same?
  - Dose the prediction error of ultimate claim same?
- Intuitively they are different: Bayesian approach try to answer the question theoretically.

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

MODEL SETTINGS

- The claim triangle: fixed and known

$$\begin{array}{ccccc} x_{1,1} & x_{1,2} & \cdots & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,N-1} & \\ \vdots & & \iddots & & \\ \vdots & \iddots & & & \\ x_{N,1} & & & & \end{array}$$

- Ladder chain estimation:
  - Development factor $\hat{f}_j \middle| \mathbf{x} = \sum_{i=1}^{N-j} x_{i,j+1} \Big/ \sum_{i=1}^{N-j} x_{i,j}$
  - Ultimate claim $\hat{x}_{i,N} \middle| \mathbf{x} = x_{i,N-i+1} \prod_{j=N-i+1}^{N-1} \hat{f}_j$
  - Prediction error $MSE_i = E\left[\hat{x}_{i,N} - X_{i,N} \;^2 \middle| \mathbf{x} \right]$

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

MEAN SQUARE ERROR

- Bayesian approach MSE

$$MSE_i = E\left[\hat{x}_{i,N} - X_{i,N} \;^2 \middle| \mathbf{x} \right]$$

$$= E\left[\hat{x}_{i,N} - E \;\; X_{i,N} \;^2 - 2 \;\hat{x}_{i,N} - E \;\; X_{i,N} \quad X_{i,N} - E \;\; X_{i,N} \;+\; X_{i,N} - E \;\; X_{i,N} \;^2 \middle| \mathbf{x} \right]$$

$$= \;\hat{x}_{i,N} \middle| \mathbf{x} - E \;\; X_{i,N} \middle| \mathbf{x} \;^2 + E\left[X_{i,N} - E \;\; X_{i,N} \;^2 \middle| \mathbf{x} \right] \geq \operatorname{var} \;\; X_{i,N} \middle| \mathbf{x}$$

- Frequentist approach MSE

$$MSE_i = \operatorname{var} \;\; X_{i,N} \middle| \mathbf{x} \;+\; E \;\; X_{i,N} \middle| \mathbf{x} \;-\hat{x}_{i,N} \;^2$$

Process error Parameter error

  - But more accurately

$$MSE_i = \operatorname{var} \;\; X_{i,N} \middle| MLE \;\, parameters \;+\; E \;\; X_{i,N} \middle| \mathbf{x} \;-\hat{x}_{i,N} \;^2$$

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

MACK's MODEL

- Key assumptions

$$E \;\; X_{i,j+1} \middle| X_{i,1}, \ldots, X_{i,j} \;\; = f_j X_{i,j}$$

$$X_{i,1}, \ldots, X_{i,N} \qquad , \; X_{k,1}, \ldots, X_{k,N}$$ are independent $i \neq k$

$$\operatorname{var} \;\; X_{i,j+1} \middle| X_{i,1}, \ldots, X_{i,j} \;\; = \sigma_j^2 X_{i,j}$$

- One more Gaussian assumption

$$X_{i,j+1} \middle| \;\; X_{i,1}, \ldots, X_{i,j} \;\; \sim N \;\; f_j X_{i,j}, \sigma_j^2 X_{i,j}$$

- Other models are possible

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

BAYESIAN APPROACH

- Basically,

$$p(\mathbf{f},\boldsymbol{\sigma}^2 \mid \mathbf{x}) \propto p(\mathbf{x} \mid \mathbf{f},\boldsymbol{\sigma}^2) \cdot p(\mathbf{f},\boldsymbol{\sigma}^2)$$

  where $\mathbf{f} = (f_1, f_2, \ldots, f_{N-1})$ and $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N-1}^2)$

- $p(\mathbf{f},\boldsymbol{\sigma}^2)$ is joint prior distribution
- $p(\mathbf{f},\boldsymbol{\sigma}^2 \mid \mathbf{x})$ is joint posterior distribution, which we want to calculate
- $p(\mathbf{x} \mid \mathbf{f},\boldsymbol{\sigma}^2)$ is determined by the model used (Mack's model)

$$p(\mathbf{x} \mid \mathbf{f},\boldsymbol{\sigma}^2) = \prod_{i=1}^{N} p(x_{i,1}, x_{i,2}, \ldots, x_{i,N-i+1} \mid \mathbf{f},\boldsymbol{\sigma}^2) \propto \prod_{j=1}^{N-1}\left\{\prod_{i=1}^{N-j}\left\{\frac{1}{\sqrt{\sigma_j^2}}\exp\left[-\frac{(y_{i,j}-f_j)^2}{2(\sigma_j^2/x_{i,j})}\right]\right\}\right\}$$

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

BAYESIAN APPROACH

- Assume prior distributions pair (f_j,σ_j^2) are independent

$$p(\mathbf{f},\boldsymbol{\sigma}^2 \mid \mathbf{x}) \propto \prod_{j=1}^{N-1}\left\{\prod_{i=1}^{N-j}\left\{\frac{1}{\sqrt{\sigma_j^2}}\exp\left[-\frac{(y_{i,j}-f_j)^2}{2(\sigma_j^2/x_{i,j})}\right]\right\}\cdot p(f_j,\sigma_j^2)\right\}$$

  which shows the posterior pair $(f_j,\sigma_j^2 \mid \mathbf{x})$ are independent as well

$$p(f_j,\sigma_j^2 \mid \mathbf{x}) \propto \prod_{i=1}^{N-j}\left\{\frac{1}{\sqrt{\sigma_j^2}}\exp\left[-\frac{(y_{i,j}-f_j)^2}{2(\sigma_j^2/x_{i,j})}\right]\right\}\cdot p(f_j,\sigma_j^2)$$

- Marginal posterior distribution

$$p(f_j \mid \mathbf{x}) = \int_0^{+\infty} p(f_j \mid \sigma_j^2,\mathbf{x})\, p(\sigma_j^2 \mid \mathbf{x})\, d\sigma_j^2 = \int_0^{+\infty} p(f_j,\sigma_j^2 \mid \mathbf{x})\, d\sigma_j^2$$

$$p(\sigma_j^2 \mid \mathbf{x}) = \int_0^{+\infty} p(\sigma_j^2 \mid f_j,\mathbf{x})\, p(f_j \mid \mathbf{x})\, df_j = \int_0^{+\infty} p(f_j,\sigma_j^2 \mid \mathbf{x})\, df_j$$

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

RECURSIVE MSE FORMULA

- Some algebra could show MSE is

$$\mathrm{var}(X_{i,j+1} \mid \mathbf{x}) = E(X_{i,j+1}^2 \mid \mathbf{x}) - E^2(X_{i,j+1} \mid \mathbf{x})$$

$$= \mathrm{var}(f_j \mid \mathbf{x})\, E^2(X_{i,j} \mid \mathbf{x}) + E(f_j^2 \mid \mathbf{x})\, \mathrm{var}(X_{i,j} \mid \mathbf{x}) + E(\sigma_j^2 \mid \mathbf{x})\, E(X_{i,j} \mid \mathbf{x})$$

  with initial value $E(X_{i,N-i+1} \mid \mathbf{x}) = x_{i,N-i+1}$

$$\mathrm{var}(X_{i,N-i+1} \mid \mathbf{x}) = 0$$

- There is no split of process and parameter error in Bayesian approach.
- To compare with Frequentist approach results, this have to artificially be split into process and parameter components
  - The idea is to keep the process component same as Frequentist approach and compare the balancing parameter component

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

COMPARISON WITH FREQUENTIST APPROACH

- Bayesian approach

$$\mathrm{var}_{par}\ X_{i,j+1}|\mathbf{x}\ = \mathrm{var}\ f_j|\mathbf{x}\ E^2\ X_{i,j}|\mathbf{x}\ + \mathrm{var}\ f_j|\mathbf{x}\ \mathrm{var}_{pro}\ X_{i,j}|\mathbf{x}\ + E\ f_j^2|\mathbf{x}\ \mathrm{var}_{par}\ X_{i,j}|\mathbf{x}$$

- Mack's formula

$$\mathrm{var}_{par}\ X_{i,j+1}|\mathbf{x}\ = \mathrm{var}\ f_j|\mathbf{x}\ E^2\ X_{i,j}|\mathbf{x}\ + E^2\ f_j|\mathbf{x}\ \mathrm{var}_{par}\ X_{i,j}|\mathbf{x}$$

- Murphy's formula & BBMW's formula

$$\mathrm{var}_{par}\ X_{i,j+1}|\mathbf{x}\ = \mathrm{var}\ f_j|\mathbf{x}\ E^2\ X_{i,j}|\mathbf{x}\ + E\ f_j^2|\mathbf{x}\ \mathrm{var}_{par}\ X_{i,j}|\mathbf{x}$$

- Bayesian approach is always larger, but may be marginal.

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

COMPARISON-II

- Parameter estimation also takes different philosophy
  - Frequentist approach use MLE σ_j^2
  - Bayesian approach use $E\ \sigma_j^2|\mathbf{x}$
- As will be shown, this difference is large in the tail.

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

PARAMETER ESTIMATION

- Parameter are estimated under different prior distributions.
- Prior 1: $p\ f\ = 1$ and σ^2 is fixed and known
- Prior 2: $p\ f, \sigma^2\ \propto 1/\sigma^2$
- Prior 3: $p\ f, \sigma^2\ = p\ f|\sigma^2\ p\ \sigma^2$
  where $\sigma^2 \sim IG\ \varpi_0/2, \sigma_0^2/2$ and $f|\sigma^2 \sim N\ \mu_0, \sigma^2/\eta_0$
- Prior 4: $p\ f, \sigma^2\ = p\ f\ p\ \sigma^2$
  where $f \sim N\ \mu_0, \varepsilon_0^2$ and $\sigma^2 \sim IG\ \varpi_0/2, \sigma_0^2/2$

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

PARAMETER ESTIMATION - PRIOR 2

- Marginal posterior distribution of f is

$$p\left(f\,|\,\mathbf{x}\right) \propto \left\{1+\frac{\left(f-\hat{f}\right)^{2}\sum_{i=1}^{K}x_i}{\left(K-1\right)s^{2}}\right\}^{-K/2} \quad \text{where} \quad s^{2}=\frac{1}{K-1}\sum_{i=1}^{K}x_i\left(y_i-\hat{f}\right)^{2}$$

  which is a shifted and scaled t-distribution.
- Marginal posterior distribution of σ^2 is

$$p\left(\sigma^{2}\,|\,\mathbf{x}\right) \propto \left(\sigma^{2}\right)^{-\left(K+1\right)/2}\exp\left[-\frac{\left(K-1\right)s^{2}}{2\sigma^{2}}\right]$$

  which is a inverse Gamma distribution.
- Parameters can be calculated from these distributions, for example

$$\operatorname{var}\left(f\,|\,\mathbf{x}\right)=\left(\frac{K-1}{K-3}s^{2}\right)\Big/\sum_{i=1}^{K}x_i$$

  only define for $K > 3$

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

NUMERICAL EXAMPLE - Parameter estimation

- $E\left(f_j\,|\,\mathbf{x}\right)$ are almost same
- $\operatorname{var}\left(f_j\,|\,\mathbf{x}\right)$ are different

j	Prior 1	Prior 2	Prior 3	Prior 4
1	0.04817026	0.06422701	0.05504437	0.04816468
2	0.00368120	0.00515367	0.00429406	0.00368071
3	0.00278879	0.00418318	0.00334590	0.00278834
4	0.00082302	0.00137170	0.00102854	0.00082287
5	0.00076441	0.00152882	0.00101890	0.00076424
6	0.00051306	0.00153917	0.00076923	0.00051291
7	0.00003505	**0.00010514**	0.00007011	0.00003507
8	0.00013466	**0.00040399**	**0.00026932**	0.00013466
9	0.00011650	**0.00034951**	**0.00023301**	0.00027045

- $E\left(\sigma_j^{2}\,|\,\mathbf{x}\right)$ shows similar pattern as to $\operatorname{var}\left(f_j\,|\,\mathbf{x}\right)$

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

NUMERICAL EXAMPLE – different formula

- Use the same parameter estimation (from Prior 1), MSE of Bayesian approach is marginally higher

Year	Mack	Murphy/BBMW	Bayesian
2	75,535	75,535	75,535
3	121,699	121,700	121,703
4	133,549	133,551	133,556
5	261,406	261,412	261,436
6	411,010	411,028	411,111
7	558,317	558,356	558,544
8	875,328	875,430	875,921
9	971,258	971,385	972,234
10	1,363,155	1,363,385	1,365,456
Total	2,447,095	2,447,618	2,449,345

- Frequentist approach is almost equivalent to Bayesian approach with Prior 1, which is a very strong prior assuming that is fixed and known

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

NUMERICAL EXAMPLE – different prior

- MSE with different prior

Year	Prior 1	Prior 2	Prior 3	Prior 4
2	75,535	130,831	106,823	115,086
3	121,703	210,810	172,120	149,104
4	133,556	231,348	188,890	158,383
5	261,436	452,921	332,284	273,259
6	411,111	641,245	495,957	419,342
7	558,544	816,905	655,425	565,685
8	875,921	1,184,204	995,294	882,037
9	972,234	1,259,424	1,085,789	976,334
10	1,365,456	1,664,613	1,488,920	1,367,860
Total	2,449,345	3,383,619	2,830,505	2,527,166

- Vaguer prior leads to much higher MSE

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

CONCLUSION

- Frequentist approach is equivalent to Bayesian approach with very strong prior knowledge.
- MSE of Bayesian approach with weak prior knowledge is much larger than that of Frequentist.
- We need prior knowledge to reasonably estimate the development factors in the tail.

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

Questions?

Bayesian Approach for Prediction Error in Chain-Ladder Claims Reserving

Thank You