9th Seminar on Current Issues in General Insurance



Bayesian MCMC Reserving in General Insurance

Presented by:

Prasun Sarkar, Appointed Actuary and Chief Actuarial Officer, ICICI Lombard General Insurance Co Ltd

Ritu Kotnala, Head of Actuarial Practices, ICICI Lombard General Insurance Co Ltd

Kushagra Patawari, Actuarial Analyst, ICICI Lombard General Insurance Co Ltd

Date: October 28, 2026

BASICS OF BAYESIAN STATISTICS

Institute of Actuaries of India

Frequentist Statistics

- Parameters are fixed
- Data is random and has a distribution

Bayesian Statistics

- Parameters are random and has a distribution
- Data is also random and has a distribution

The Big Question: What is Probability?
For ex: What is the probability it will rain tomorrow?

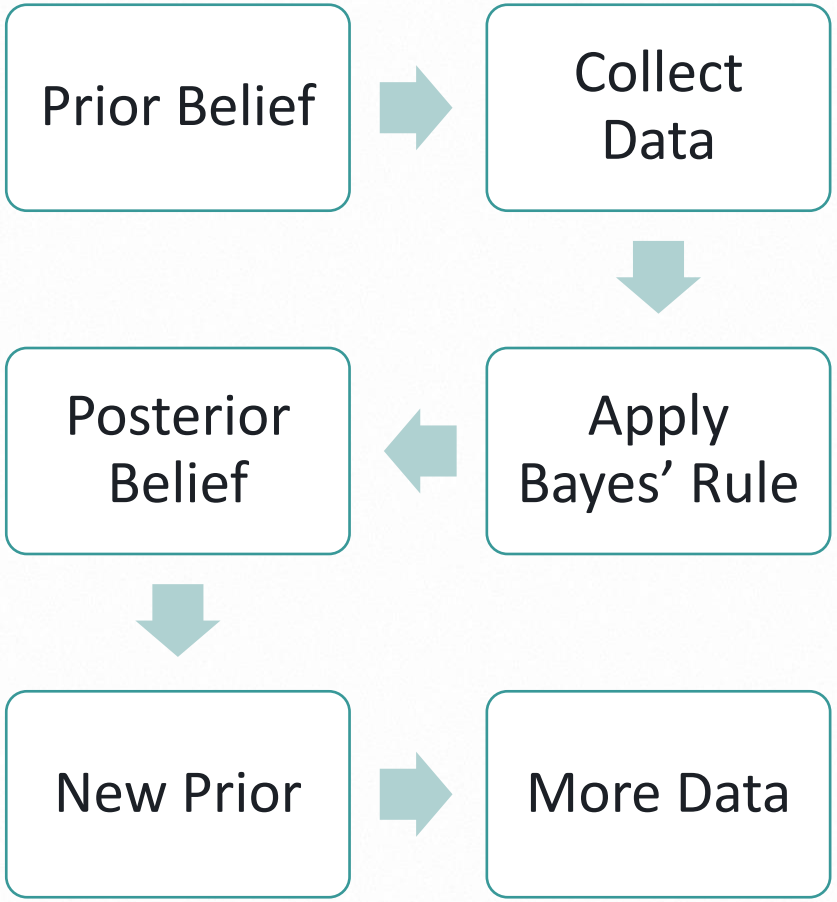
Frequentist Statistics

- Probability = Long-run Frequency of Events
- If we could repeat tomorrow's weather again and again, rain would occur in 70% of those worlds
- $P(Data|\theta)$

Bayesian Statistics

- Probability = Degree of belief given long-run data and prior information.
- Based on everything I know, I believe there's a 70% chance of rain tomorrow.
- The more natural $P(\theta|Data)$

THE BAYESIAN ENGINE



and the process continues …

Bayes' Rule:

$$P(Belief|Data) = \frac{P(Data|Belief) * P(Belief)}{P(Data)}$$

Simplified to:

$$New\ Belief \propto Data * Old\ Belief$$

Practitioner's Problems:

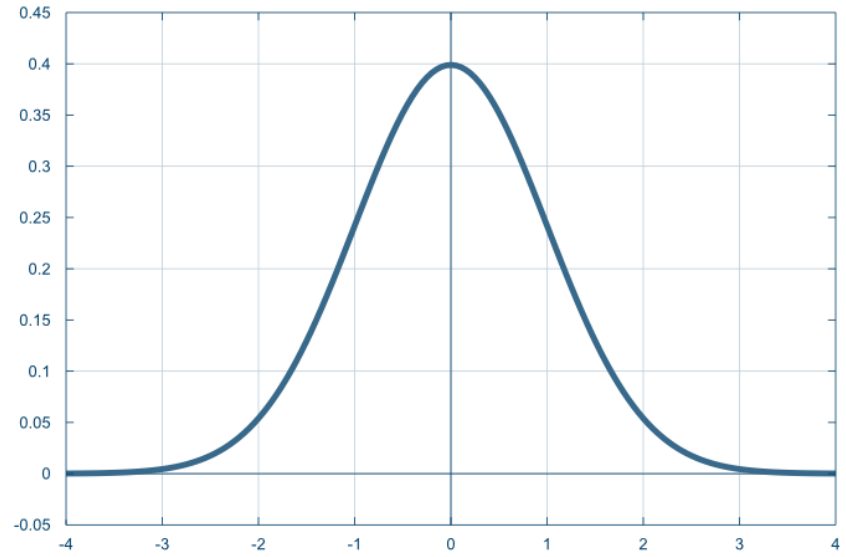
- In some situations, the updated belief (posterior) can be easily calculated from initial belief (prior) and the information provided by the data (likelihood)*.
- However in many real-world applications, the calculations are highly complex and it may not always be possible to obtain a closed-form solution.
- **Bayesian MCMC** helps address these challenges by providing a way to approximate the posterior even when direct computation is difficult.

* When Prior Distribution and Likelihood are Conjugate

BAYESIAN MCMC

Bayesian Inference gives us a posterior, however that distribution is often impossible to calculate exactly. So, we require a clever way of approximating it.

- Imagine you're trying to sketch a mountain in the dark.
- Instead of measuring every single point in the mountain, you can approximate by:
  - Wandering around the mountain
  - Moving more readily towards higher regions than lower ones.
  - Recording your location at regular intervals.
  - Counting how often each location is visited.
- Overtime, you'll collect thousands of recorded locations – peaks that are visited often will have more recorded points than the flat regions visited rarely.
- By plotting all the recorded locations, a picture of the mountain gradually emerges.



This is what MCMC does! Regions that are more plausible are visited more frequently, causing more samples to accumulate there and gradually revealing the posterior distribution. It is often done through algorithms like **Metropolis-Hastings**.

M-H ALGORITHM

In Metropolis-Hastings Algorithm, our main objective is to draw samples from our target posterior distribution (complicated in nature) using some proposal density (easy to compute).

The heart of MH Algorithm is the Acceptance Ratio

Acceptance Ratio = min(1, R)

where $R = R_1 * R_2$

$$= \frac{f(x|y^*).p(y^*)}{f(x|y_{t-1}).p(y_{t-1})} * \frac{J(y_{t-1}|y^*)}{J(y^*|y_{t-1})}$$

Prior distribution of parameter = $p(y)$

Conditional distribution= $f(x|y)$

Proposal Distribution = $J(y^*|y_{t-1})$

Proposed State = y^*

Current State = y_{t-1}

Bayes Rule

Posterior ∝ Likelihood * Prior

$\Rightarrow \pi(y) \propto f(x|y) * p(y)$

This R_1 part comes directly from Bayes Rule. This represents the ratio of the posterior probability of y^*(proposed state), to the posterior probability of y_{t-1}(current state). The higher the value of R_1, the more likely that proposed will be accepted in the chain.

ROLE OF R2 IN M-H

$$R_2 = \frac{J(y_{t-1}|y^*)}{J(y^*|y_{t-1})}$$

This part plays a crucial role to satisfy the detailed balance condition which ensures that the Markov Chain possesses the posterior distribution as its stationary distribution

Now under MH Algorithm,

Acceptance ratio from current state y_{t-1} to proposed state y^*,

$$A(y^*, y_{t-1}) = \min(1, R)$$

Let assume $R < 1$, proposed state is less likely than current state

Now, $\pi(y_{t-1})T(y_{t-1} \to y^*) = \pi(y_{t-1}) * J(y^*|y_{t-1}) * A(y^*, y_{t-1})$

$= \pi(y_{t-1}) * J(y^*|y_{t-1}) * R$

$= \pi(y_{t-1}) * J(y^*|y_{t-1}) * \frac{\pi(y^*)}{\pi(y_{t-1})} * \frac{J(y_{t-1}|y^*)}{J(y^*|y_{t-1})}$

$= \pi(y^*) * J(y_{t-1}|y^*)$

$= \pi(y^*) * J(y_{t-1}|y^*) * A(y_{t-1}, y^*)$

, [Since $A(y^*, y_{t-1}) = R < 1$, $A(y_{t-1}, y^*)$ must be 1]

$= \pi(y^*)T(y^* \to y_{t-1})$

Detailed Balance Condition

$$\pi(y_{t-1})T(y_{t-1} \to y^*) = \pi(y^*)T(y^* \to y_{t-1})$$

Where $\pi(y)$ = posterior density

$T(y_{t-1} \to y^*)$ = transition probability from state y_{t-1} to y^*

Under MH Setup,

$T(y_{t-1} \to y^*) = J(y^*|y_{t-1}) * A(y^*, y_{t-1})$

Thus, Detailed Balance Condition is satisfied under MH setup. In case of $R \geq 1$, staring from RHS the proof can be done in similar way. Hence, MH algorithm satisfies stationarity ensuring the posterior is a stationary distribution.

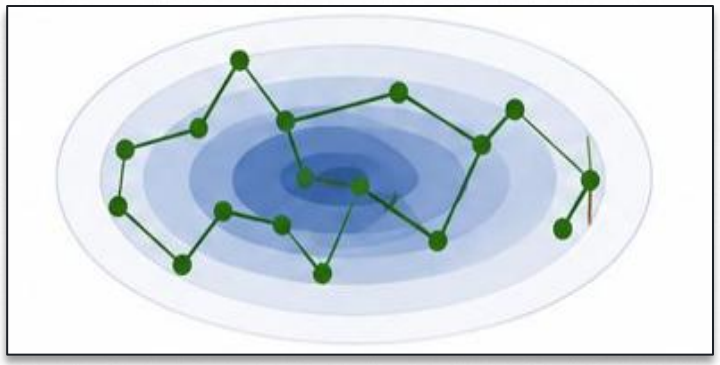
* for sake of calculation we assume $f(x|y^*).p(y^*) = \pi(y^*)$ and $f(x|y_{t-1}).p(y_{t-1}) = \pi(y_{t-1})$.

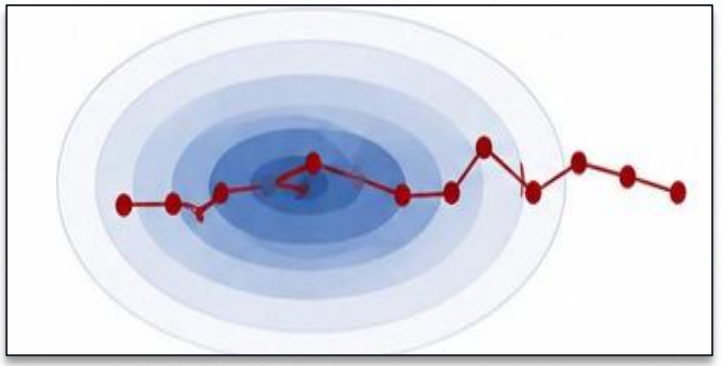
WHAT IF PROPOSAL BIAS (R2) IS NOT CORRECTED?

The Metropolis-Hastings acceptance rule corrects any imbalance in how often different states are proposed. It ensures that the Markov Chain moves in a way that reflects the true target distribution.

But if the proposal mechanism favors some states more than the others then:

An unbiased proposal will move across all directions and explore the distribution well.





A biased proposal which is more likely to move right than explore the distribution well.

- Proposal bias disrupts the true posterior; so, the chain spends too much time in certain region simply because they are proposed more often.
- Detailed balance breaks – unequal forward and backward movement means the probability flow becomes unbalanced, so the target distribution cannot remain stationary.
- The chain either fails to converge, or it converges to a distribution shaped by proposal bias, and not to the true posterior.

HOW M-H CONVERGENCE TO POSTERIOR DISTRIBUTION

Irreducibility –

If the proposal density $J(y^*|y)$ allows proposing any state with non-zero posterior:

$$J(y^*|y) > 0 \quad \text{whenever } \pi(y^*) > 0,$$

Then eventually,

$$T^{(n)}(y \to y^*) > 0 \text{ ,for some n.}$$

Every state can eventually be reached from any other state. **The chain is free to explore the entire parameter space.**

Aperiodicity -

MH has a positive probability of staying at the same state

$$T(y \to y) = 1 - \int J(y^*|y)\, A(y^*, y)\, dy^* > 0$$

A self-transition breaks cycles, giving aperiodicity.

Ergodicity

Irreducibility and aperiodicity together ensure ergodicity.

$$T^{(n)}(y_0 \to y) \xrightarrow{n\to\infty} \pi(y)$$

M-H guarantees the chain will forget its starting point (y_0) and converges to the posterior

Long run Convergence

Once stationarity and ergodicity hold, the ergodic theorem applies. By this theorem empirical frequencies converge to posterior probabilities.

$$\frac{1}{N}\sum_{t=1}^{N} 1(Y_t \in A) \to \int_A \pi(y)\, dy \quad \text{,for any region A}$$

Burn in Period: The burn-in period removes early iterations influenced by the starting value, ensuring the chain reaches the stationary posterior region before valid sampling begins.

When stationarity, ergodicity, long-run convergence, and a proper burn-in are all satisfied, the Metropolis–Hastings algorithm yields samples that accurately represent the posterior distribution.

SOME APPLICATIONS OF BAYESIAN STATISTICS

Pricing – Sparse Data

- For new products, we often have very little historical data.
- Pure experience rating becomes volatile or unusable.
- Bayesian approach lets us start with a prior – for ex: based on industry frequency or severity – and then update the prior as claims emerge.



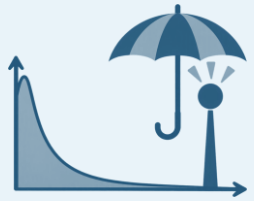
Reserving – Thin Triangles

- Triangles by design are sparse in data.
- Use MCMC to get posterior distribution of reserves
- It stabilizes estimates when data is volatile.
- Can also insert subjective judgement into Bayesian Models.



EVT – Rare Large Losses

- For modelling large losses using EVT, datasets are naturally tiny.
- MLE for the GPD tail index becomes unstable.
- Bayesian EVT can use industry benchmarks to set priors and update as per company experience to stabilize tail estimates and reliable quantiles.



LIMITATIONS OF TRADITIONAL STOCHASTIC MODELS

- In this case study, we are focusing on a long-tailed line of business.

Core Assumptions of Mack:

1. Development Factors are Uncorrelated across Development Periods
2. Independence of Accident Years
3. The conditional variance of the next cumulative claim amount is proportional to the current cumulative claim amount

These assumptions may not always hold true.

Focusing on assumption 2, there may be various reasons that Accident Years may not be independent, such as:

- An UWY is split across multiple AYs – One adverse UWY can affect multiple AYs together. This effect increases for lines with multi-year policies
- Calendar Year Effects, such as:
  - Landmark Judgements
  - Change in Case OS Philosophy.

LIMITATIONS OF TRADITIONAL STOCHASTIC MODELS

- Mack realized these limitations and proposed a statistical test for practitioners to test the presence of Calendar Year effects (slide 23).
- The test examines whether the triangle diagonals contain ATAs that are consistently too high or too low.
- If they are, then it concludes the presence of Calendar Year effects.

Results of the test on a long-tailed triangle:

Confidence Interval	
Lower	52.73
Upper	68.12
Z_j	35

There is significant Calendar Year effect if Z_j does not lie within the confidence interval

- The test statistic (Z_j) falls outside the confidence interval, clearly indicating a significant Calendar Year effect in the triangle.
- This failure shows that **key patterns/trends** like legal judgements that **move across calendar years** are not being captured by the model.
- And once the calendar-year test fails, it becomes evident that Mack is not correctly estimating variances – not accounting for correlation across AYs.

Hence, this is a crucial trend that these stochastic models miss out on

Reference: https://www.casact.org/sites/default/files/2021-03/7_Mack_1994.pdf

MCMC MODEL SPECIFICATION

- Meyers has specified a MCMC model in his CAS Monograph: "CAS Monograph No. 1 –Stochastic Loss Reserving Using Bayesian MCMC Models"
- The Model is specified as:

$$C_{w,d} \sim LogN(\mu_{w,d}, \sigma_d)$$

$$\sim N(\ln(ELR) + \ln(Prem_w), var)$$

$$\mu_{w,d} = \alpha_w + \beta_d + \boxed{\rho}\left(\ln(C_{w-1,d}) - \mu_{w-1,d}\right) \qquad \sigma_d^2 = \sigma_{d+1}^2 + Uniform(0,1)$$

This Rho Parameter models Correlations between successive AYs

Links the current Accident Year to the Previous one

Enforces that $\sigma_1^2 > \sigma_2^2 > \cdots > \sigma_n^2$. i.e. that volatility decreases the more development occurs.

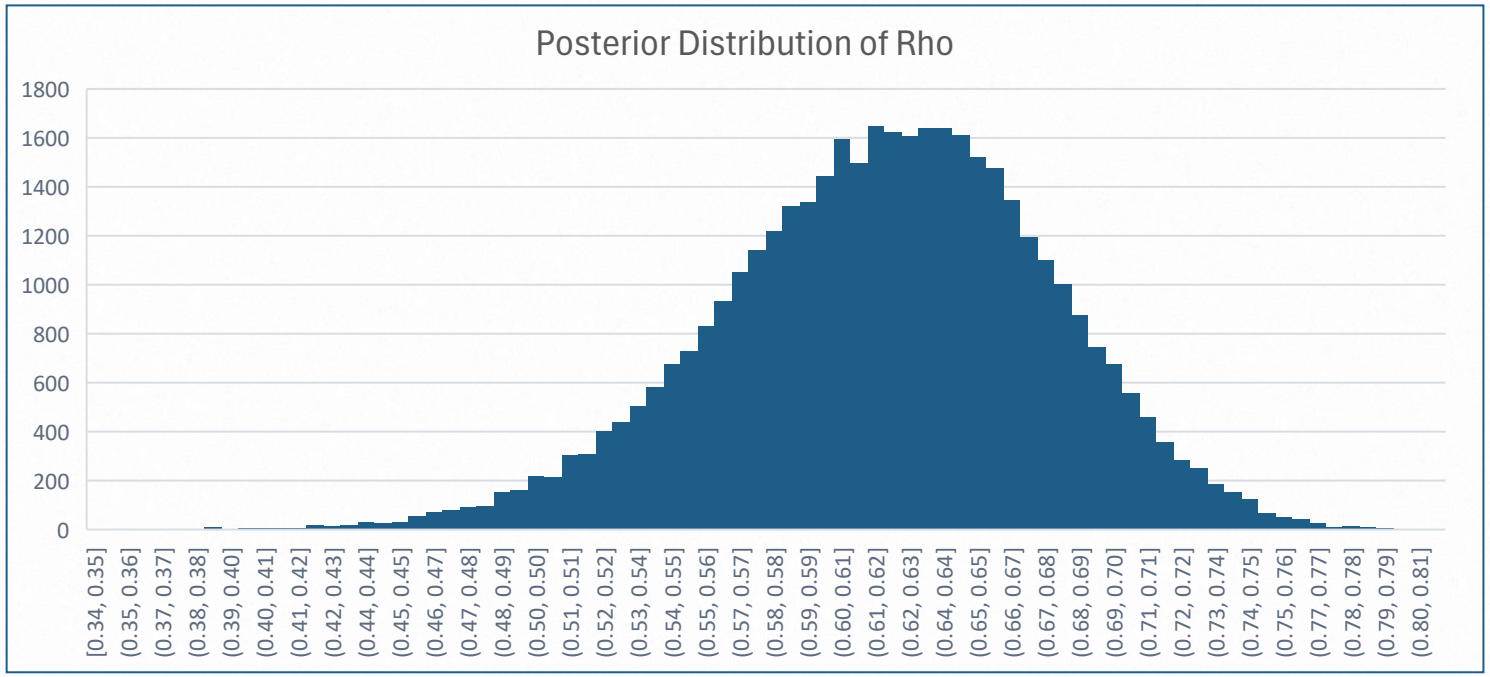
$\mu_{w,d}$ … *Log of Mean Cumulative Losses for AY w at Delay d*
$C_{w,d}$ … *Actual Cumulative Losses for AY w at Delay d*
α_w … *Expected Ultimate Cumulative losses for AY w*
β_d … *Delay Effect*
ρ … *Correlation Parameter*
σ_d … *Variance at Delay d*
$Prem_w$ … *Earned Premium for AY w*
ELR … Expected Loss Ratio for the Line of Business

Reference: https://www.casact.org/monograph/monograph-no-1

THE CORRELATION PARAMETER "RHO"

- The Rho Parameter ρ provides a way to explicitly model correlations across AYs

Posterior Distribution of Rho from a long-tailed triangle:



Prior Distribution of ρ

$$\rho \sim Uniform(-1,1)$$

Credible Interval of Posterior ρ	
Lower Limit	0.4976
Upper Limit	0.7233

Credible Interval of ρ not including zero suggests statistically significant AY correlation. This is explained by the CY effect as shown by Mack's CY effect test

- The posterior distribution of ρ is centered well above zero, and the credible interval ranges from 0.4976 to 0.7233
- Here we can see that, every plausible value of ρ is positive. This means that the data, combined with the prior, strongly supports positive correlation across the AYs.

OVERVIEW OF FURTHER MODEL ITERATIONS

Institute of Actuaries of India

Objective:

All iterations made are with the objective of better modelling trends that practitioners observe in their reserving exercise.

Deterministic Models

Includes models like CLM, BF, Berquist-Sherman, etc

Give accurate Best Estimates, but no indication of variability.

Stochastic Models

Includes models like Mack, ODP Bootstrap, etc.

Give measure of Variability along with BE.

However,

1) Do not capture all trends like correlation between AYs.

2) Practitioners rely on Scaling to align output with selections.

Meyers' MCMC Model

An improvement over Traditional Stochastic Models like Mack as it takes into account trends not considered by other models.

However, assumes:

1) Decreasing Volatility across Development Lags

2) Common ELR Patterns for all Accident Years

MCMC Iteration 1

Fixing assumption of decreasing volatility, which does not hold for all triangles

MCMC Iteration 2

Fixing the Mean ($\mu_{w,d}$) Structure to better align Best Estimates with Selections without relying on Scaling.

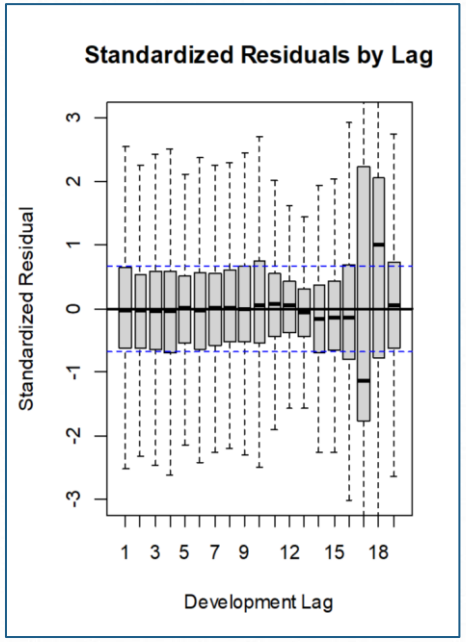
MCMC iterations have occurred over multiple models with trial and error. Details shown in further slides.

ITERATIONS TO MEYERS' MODEL

Institute of Actuaries of India

Model 1

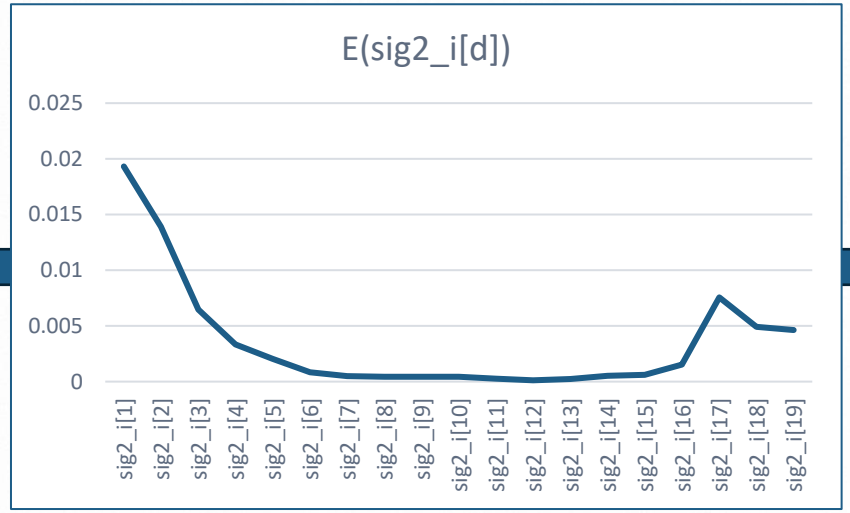
Model Structure as Specified by Meyers.



Goodness of Fit fails for Later Delays….
Hypothesis that variance is underestimated in later delays.

Testing the assumption of
$\sigma_1^2 > \sigma_2^2 > \cdots > \sigma_n^2$

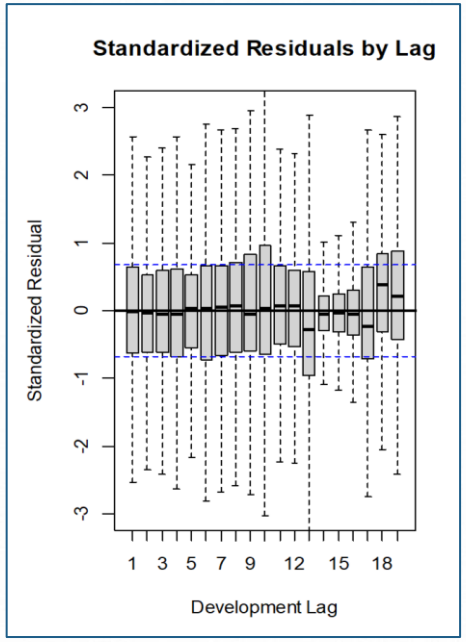
If we remove the decreasing constraint from σ_d^2 , we get this shape:



There isn't a constant decrease, but a decrease till the middle and then a sharp increase in the tail
This is in line with our expectations, that there will be **sudden jumps in volatility** due to **high-value, hard to settle** claims in the tail!

Model 2

New Variance Structure
$\sigma_1^2 > \sigma_2^2 > \cdots > \sigma_{n-k}^2$
$and\ \sigma_{n-k+1}^2 > \cdots > \sigma_n^2$



With this change in Variance, the model fits much better in the later delays.

ITERATIONS TO MEYERS' MODEL

Model 1 & 2

Wide Priors as defined by Meyers.

AY	Incurred CLM	MCMC Model 1	MCMC Model 2
2006-07	1.0000	1.0000	1.0000
2007-08	1.0000	0.9800	0.9783
2008-09	1.0000	0.9636	0.9461
2009-10	1.0000	1.0360	1.0237
2010-11	1.0000	1.0891	1.0760
2011-12	1.0000	1.0928	1.0795
2012-13	1.0000	1.0732	1.0594
2013-14	1.0000	1.0745	1.0545
2014-15	1.0000	1.1048	1.0746
2015-16	1.0000	1.1332	1.0999
2016-17	1.0000	1.1663	1.1269
2017-18	1.0000	1.1819	1.1428
2018-19	1.0000	1.1870	1.1547
2019-20	1.0000	1.1897	1.1638
2020-21	1.0000	1.2005	1.1767
2021-22	1.0000	1.2295	1.2065
2022-23	1.0000	1.2494	1.2258
2023-24	1.0000	1.3815	1.3570
2024-25	1.0000	1.2859	1.2649

In both models, our best estimates differ from best estimates derived from Incurred CLMs

Inputting Chain-Ladder LDFs as Priors

We can use Chain-Ladder LDFs as Priors for Beta to guide the model to our Best Estimates

Let LDF at delay d be:

$$LDF_d$$

Meyers has defined very wide priors in his model for all delay effects.

$$\beta_d \sim N(0, \sqrt{10})$$

Selected LDFs can roughly be transformed into the model's β_d parameter as:

$$E(\beta_d) = \ln\left(\frac{1}{LDF_d}\right)$$

Hence, the model is regularized to our expected development pattern. Unless the data heavily suggests something different, it will follow our suggestions.

Model 3

Priors defined by Incurred CLM selected LDFs

AY	Incurred CLM	MCMC Model 1	MCMC Model 2	MCMC Model 3
2006-07	1.0000	0.9919	0.9919	1.0000
2007-08	1.0000	0.9800	0.9783	0.9636
2008-09	1.0000	0.9636	0.9461	0.8976
2009-10	1.0000	1.0360	1.0237	0.9707
2010-11	1.0000	1.0891	1.0760	1.0192
2011-12	1.0000	1.0928	1.0795	1.0231
2012-13	1.0000	1.0732	1.0594	1.0044
2013-14	1.0000	1.0745	1.0545	0.9988
2014-15	1.0000	1.1048	1.0746	1.0162
2015-16	1.0000	1.1332	1.0999	1.0404
2016-17	1.0000	1.1663	1.1269	1.0652
2017-18	1.0000	1.1819	1.1428	1.0786
2018-19	1.0000	1.1870	1.1547	1.0878
2019-20	1.0000	1.1897	1.1638	1.0925
2020-21	1.0000	1.2005	1.1767	1.0999
2021-22	1.0000	1.2295	1.2065	1.1187
2022-23	1.0000	1.2494	1.2258	1.1246
2023-24	1.0000	1.3815	1.3570	1.2179
2024-25	1.0000	1.2859	1.2649	1.1063

Model outputs are much closer to CLM derived Best Estimates.
However, since we apply the latest trends to even the oldest cohorts, we see a significant dip in model performance of oldest AYs.

ITERATIONS TO MEYERS' MODEL

Model 1, 2 & 3

Meyers defines a common ELR prior for all AYs. However, that may not always hold true.

There might be:

- UW Cycles
- Rate Changes
- Trends/Changes in Underlying Loss Cost

Hence, we have to adjust for these changes and trends

These adjustments can be made to "scaling" the common ELR process for each AY

New Mean Structure:

$$\mu_{w,d} = \alpha_w + \beta_d + \ln(ELR) + \ln(Prem_w) + \ln(ELR\ Factor_w) + \rho(\dots)$$

We introduce a new term, the ELR Factor.

It is defined as:

$$ELR\ Factor_w = \frac{ELR_w}{ELR_1}$$

This is the ratio of practitioner's expectation of ELR of AY w v/s first AY to take into account these trends

Model 4

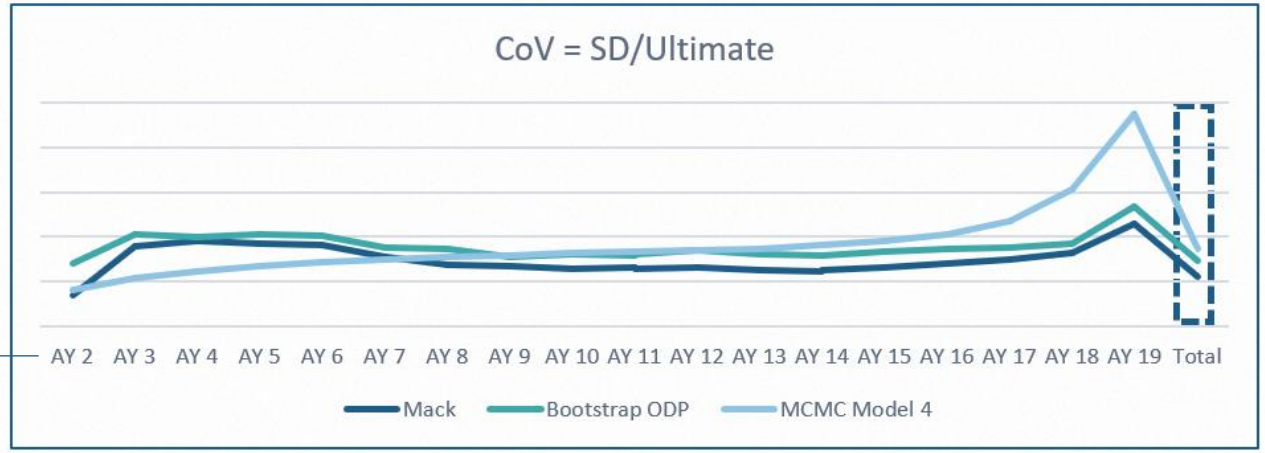
With Prior LDFs and ELR Factor

AY	Incurred CLM	MCMC Model 1	MCMC Model 2	MCMC Model 3	MCMC Model 4
2006-07	1.0000	1.0000	1.0000	1.0000	1.000
2007-08	1.0000	0.9800	0.9783	0.9636	0.9812
2008-09	1.0000	0.9636	0.9461	0.8976	0.9541
2009-10	1.0000	1.0360	1.0237	0.9707	1.0194
2010-11	1.0000	1.0891	1.0760	1.0192	1.0682
2011-12	1.0000	1.0928	1.0795	1.0231	1.0669
2012-13	1.0000	1.0732	1.0594	1.0044	1.0446
2013-14	1.0000	1.0745	1.0545	0.9988	1.0356
2014-15	1.0000	1.1048	1.0746	1.0162	1.0549
2015-16	1.0000	1.1332	1.0999	1.0404	1.0696
2016-17	1.0000	1.1663	1.1269	1.0652	1.0837
2017-18	1.0000	1.1819	1.1428	1.0786	1.0841
2018-19	1.0000	1.1870	1.1547	1.0878	1.0673
2019-20	1.0000	1.1897	1.1638	1.0925	1.0543
2020-21	1.0000	1.2005	1.1767	1.0999	1.0574
2021-22	1.0000	1.2295	1.2065	1.1187	1.0866
2022-23	1.0000	1.2494	1.2258	1.1246	1.1113
2023-24	1.0000	1.3815	1.3570	1.2179	1.2102
2024-25	1.0000	1.2859	1.2649	1.1063	1.1228
Total	1.0000	1.2171	1.1921	1.0986	1.0961

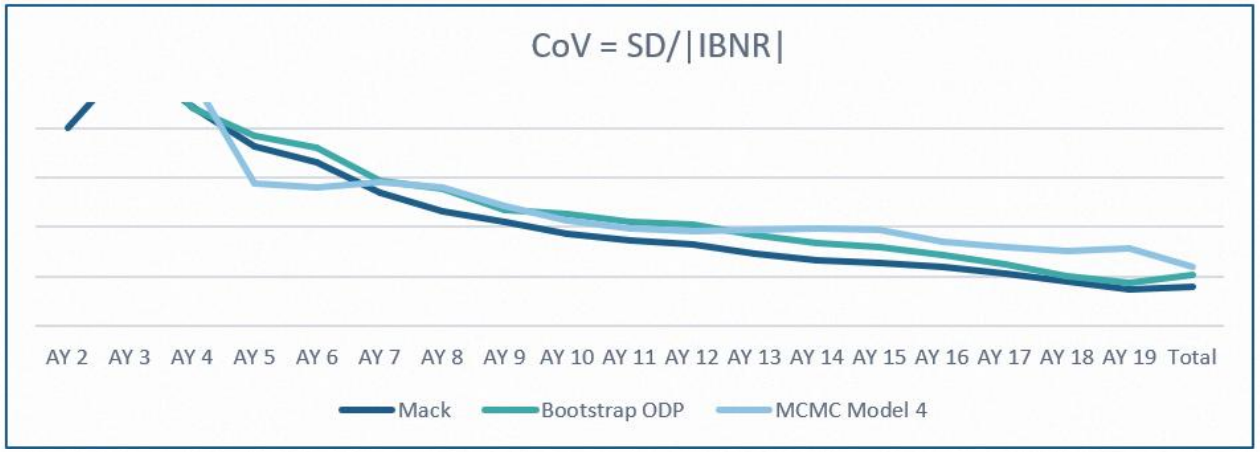
Better Performance in the Oldest AYs and the Recent AYs, however it comes at the cost of AYs in the middle.

At an overall level, Model 4 performs the best

RESULTS FROM MCMC VS MACK & ODP



Oldest AY



Ultimate CoV:

- **Cross Year Correlations:**

ρ Parameter links successive AYs and captures trends better. It transfers uncertainty from older to newer AYs.

- **AY level Uncertainty:**

Fully captures parameter variance of each Accident Year (α_w) via its Bayesian Posterior Distribution.
(More info in the next slide.)

Hence, MCMC shows a smooth increasing trajectory while the other two show a "humped" Ultimate CoV.

Increased Total CoV in MCMC due to:

- Better modelling of volatility provides higher CoV to recent accident years
- Diversification "Benefit" implied in Mack and ODP as correlation assumed to be zero.

MORE ON AY UNCERTAINTY

How Various Models deal with AY Level Parameter

Mack

No AY Level Parameter.

Essentially assumes last diagonal is fixed and no noise in baseline AY level.

Ignores reality that early-stage reported losses are volatile estimates subject to case-reserve shifts and reporting noise.

Bootstrap ODP

Introduces parameter α_w for each Accident Year.

Bootstrap ODP works by resampling residuals across the entire triangle and refitting parameters in every simulation

However, issues occur for the latest AY. The latest AY has only one data point and the parameter ends up overfitting due to a lack of data.

Bootstrap ODP also depends on non-parametric bootstrapping. Given the limited number of data points in the triangle, tail risk may be understated due to limited historical residuals.

MCMC

In MCMC, this problem is solved as we provide an exposure-based prior to guide the model.
When an AY has only one or two data points, the sparse data does not override the prior. It uses the Earned Premium and ELR Prior baseline to inform α_w.

The strength in the latest AY will come from:

1. How stable the loss ratios are in your book. The more stable they are, the more information it can feed the latest AY and get a narrower estimate.
2. How informative your Priors are.

Variability is derived from a full posterior rather than limited non-parametric residual resampling.

BENCHMARKING RESULTS AGAINST CLM, MACK AND BERQUIST-SHERMAN

	BE Ultimate						
AY	Incurred CLM	MCMC Model 1	MCMC Model 2	MCMC Model 3	MCMC Model 4	Berquist-Sherman	Mack
2006-07	1.0000	1.0000	1.0000	1.0000	1.000	0.941	1.000
2007-08	1.0000	0.9800	0.9783	0.9636	0.9812	0.9305	1.0312
2008-09	1.0000	0.9636	0.9461	0.8976	0.9541	0.8851	1.0751
2009-10	1.0000	1.0360	1.0237	0.9707	1.0194	0.9379	1.0979
2010-11	1.0000	1.0891	1.0760	1.0192	1.0682	0.9434	1.1125
2011-12	1.0000	1.0928	1.0795	1.0231	1.0669	0.9770	1.1152
2012-13	1.0000	1.0732	1.0594	1.0044	1.0446	0.9332	1.1055
2013-14	1.0000	1.0745	1.0545	0.9988	1.0356	0.9630	1.0965
2014-15	1.0000	1.1048	1.0746	1.0162	1.0549	0.9633	1.1013
2015-16	1.0000	1.1332	1.0999	1.0404	1.0696	0.9564	1.1028
2016-17	1.0000	1.1663	1.1269	1.0652	1.0837	0.9504	1.1124
2017-18	1.0000	1.1819	1.1428	1.0786	1.0841	0.9347	1.1215
2018-19	1.0000	1.1870	1.1547	1.0878	1.0673	0.9314	1.1282
2019-20	1.0000	1.1897	1.1638	1.0925	1.0543	0.9388	1.1373
2020-21	1.0000	1.2005	1.1767	1.0999	1.0574	0.9901	1.1502
2021-22	1.0000	1.2295	1.2065	1.1187	1.0866	0.9939	1.1531
2022-23	1.0000	1.2494	1.2258	1.1246	1.1113	1.1577	1.1544
2023-24	1.0000	1.3815	1.3570	1.2179	1.2102	1.0234	1.1806
2024-25	1.0000	1.2859	1.2649	1.1063	1.1228	1.0201	1.2110
Total	1.0000	1.2171	1.1921	1.0986	1.0961	0.9987	1.1469

Retroactive Trend Application: MCMC Models project newer development trends onto older cohorts

Alignment with Berquist-Sherman: Hence, Produces results consistent with the deterministic Berquist-Sherman adjustments.

Most pronounced in Model 3, where we essentially guide the model to listen to latest trends more.

Closest Unscaled Fit: MCMC Model 3 and 4 track closer to the baseline than the unscaled Mack Model due to data weighted with priors, whereas Mack relies purely on unweighted historical data.

While MCMC models are closer aligned to a practitioner's Best Estimate, scaling is still required to fully reconcile output to Practitioner's BE.
Even if the Best Estimates are not aligned, the objective of capturing variability is better fulfilled by MCMC.

SUMMARY

Institute of Actuaries of India

BENEFITS

Full Distribution of Ultimate Estimates

Posterior predictive results support tail percentiles, risk margins, capital and scenarios—not only point estimates or moments.

Flexible structure

Represent accident-year, calendar year effect e.g. Inflation, and other trends e.g. changing settlement rates.

Credibility for sparse cells

Priors help stabilize parameter estimates in immature cohorts. Hierarchical priors share information across related segments, improving robustness and reflecting the intuition behind credibility theory.

LIMITATIONS

Prior and structure risk

Priors, likelihoods and dependency choices influence results; document rationale and stress-test credible alternatives.

Diagnostics and compute

Correlated draws reduce effective information; use multiple chains, R-hat, Gelman-Ruben, ESS and trace plots. Give chains as much run time as practically possible.

Local fit and calibration

One model may not fit every accident year or line; back-test, segment and calibrate or scale to trusted benchmarks.

Let’s Discuss

Thank You!

Contact Us

Prasun Sarkar
Appointed Actuary and Chief Actuarial Officer
ICICI Lombard General Insurance
prasun.sarkar@icicilombard.com

Ritu Kotnala
Head of Actuarial Practices
ICICI Lombard General Insurance
ritu.kotnala@icicilombard.com

Kushagra Patawari
Actuarial Analyst
ICICI Lombard General Insurance
kushagra.patawari@icicilombard.com

Annexures

CALENDAR YEAR EFFECTS

Example: Without Landmark Judgment in latest FY

AY/Delay	1	2	3	4	5
1	2.00	3.00	4.00	5.00	6.00
2	2.20	3.30	4.40	5.50	
3	2.42	3.63	4.84		
4	2.66	3.99			
5	2.93				

Example: With Landmark Judgment in latest FY

AY/Delay	1	2	3	4	5
1	2.00	3.00	4.00	5.00	6.60
2	2.20	3.30	4.40	6.05	
3	2.42	3.63	5.32		
4	2.66	4.39			
5	3.22				

A simple example of a Calendar Year Effect:

- Assuming the triangles are for a long-tailed Liability LoB
- A landmark court judgement occurs, and Incurred Estimates for all claims increase by 10%
- This affects the entire diagonal all at once. All Accident Years move upward together.

These are the type of correlations that traditional stochastic models like Mack do not take into account, and hence do not correctly capture volatility.

Slide 11