Loss Reserving Curves Shape Analysis

Class Project – STAT 651

Zach Horton

December 11, 2018

Insurance Reserving: Annual Cost Prediction

Reserving is a huge financial burden

Sources of variation in annual cost:

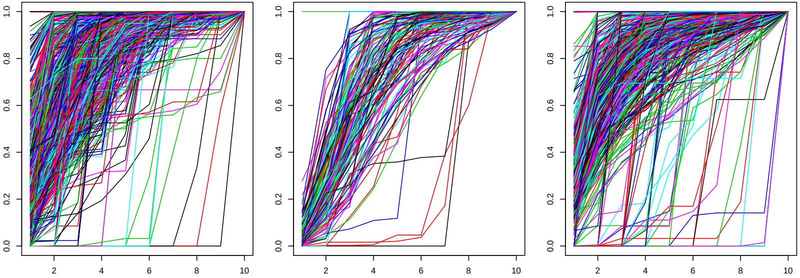
- Standard business costs
- Claims made this year
- Payouts for claims made in past years

Terms:

- Accident year – Year the claim is made (incurred cost)
- Payment year – Year of claim payment (lagged or raw)
- Ultimate – Total payout for a given accident year

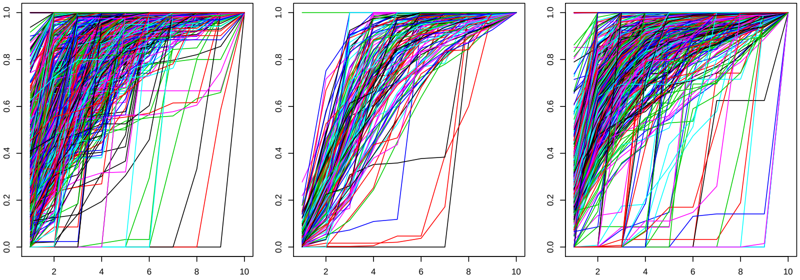
Data

- Commercial Auto, Medical Malpractice, Workers Compensation
- 10 payment years per accident year (866, 130, 639)
- Cumulative percent “trajectory”



Goals and Questions

Trajectory Curve Shape Differences



- How do trajectories differ by business line?
- Reason: accurate annual predictions (cheaper)

Model

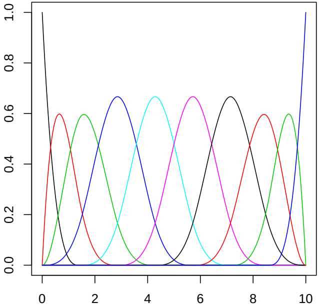
$$
\begin{aligned}
y_{ik}(t)|\boldsymbol{\beta}_i, \alpha, \gamma &\sim N(\mathbf{H}(t)\boldsymbol{\beta}_i, \alpha\gamma^{t-1}) \\
\boldsymbol{\beta}_i|\tau^2 &\sim N_{10}(\mathbf{0}, \tau^2\mathbf{P}^{-1}) \\
\alpha &\sim \text{InvGamma}(a_\alpha = 0.1, b_\alpha = 0.1) \\
\gamma &\sim \text{Unif}(m = 0, M = 1) \\
\tau^2 &\sim \text{InvGamma}(a_\tau = 0.1, b_\tau = 0.1)
\end{aligned}
$$

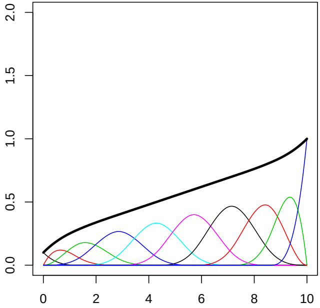
- $i \in (1, \ldots, b = 3)$ denote line of business
- $k \in (1, \ldots, n_i)$ denote an individual accident year
- $t \in (1, \ldots, 10) \subset [1, 10]$ denotes payment year lag
- $y_{ik}(t)$ is percent of ultimate paid off at time t

Model

$$y_{ik}(t)|\boldsymbol{\beta}_i, \alpha, \gamma \sim N(\mathbf{H}(t)\boldsymbol{\beta}_i, \alpha\gamma^{t-1})$$

- $\mathbf{H}(t)$ evaluates to a row vector
- $\boldsymbol{\beta}_i$ weighted combination of basis functions forms curve
- Business line specific coefficients (one per year)

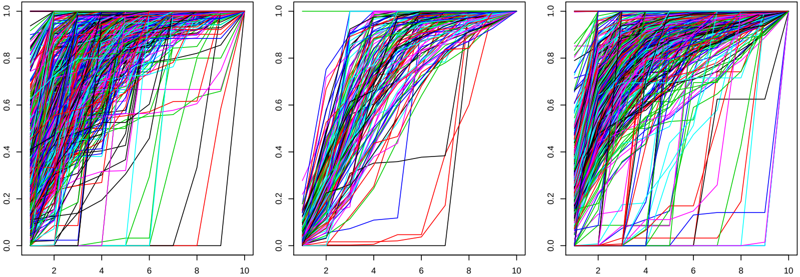




Model

$$y_{ik}(t) | \boldsymbol{\beta}_i, \alpha, \gamma \sim N(\mathbf{H}(t)\boldsymbol{\beta}_i, \alpha\gamma^{t-1})$$

- α baseline or initial variance
- γ geometric variance decay
- $t \in (1, \ldots, 10) \subset [1, 10]$ denotes payment year lag



Model

$$y_{ik}(t)|\boldsymbol{\beta}_i, \alpha, \gamma \sim N(\mathbf{H}(t)\boldsymbol{\beta}_i, \alpha\gamma^{t-1})$$
$$\boldsymbol{\beta}_i|\tau^2 \sim N_{10}(\mathbf{0}, \tau^2\mathbf{P}^{-1})$$

$$\mathbf{P} = \begin{bmatrix} 2 & -1 & 0 & \dots & \dots & 0 \\ -1 & 2 & -1 & \ddots & & \vdots \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 & 0 \\ \vdots & & \ddots & -1 & 2 & -1 \\ 0 & \dots & \dots & 0 & -1 & 1 \end{bmatrix}$$

P imposes smoothness and shrinkage.

Model Summary

$$
\begin{aligned}
y_{ik}(t)|\boldsymbol{\beta}_i, \alpha, \gamma &\sim N(\mathbf{H}(t)\boldsymbol{\beta}_i, \alpha\gamma^{t-1}) \\
\boldsymbol{\beta}_i|\tau^2 &\sim N_{10}(\mathbf{0}, \tau^2\mathbf{P}^{-1}) \\
\alpha &\sim \text{InvGamma}(a_\alpha = 0.1, b_\alpha = 0.1) \\
\gamma &\sim \text{Unif}(m = 0, M = 1) \\
\tau^2 &\sim \text{InvGamma}(a_\tau = 0.1, b_\tau = 0.1)
\end{aligned}
$$

- Model trajectory curves as noisy versions of business line averages
- Variance decays since all curves approach ultimate
- Fairly diffuse priors

Model Fitting

Alternate Formulation:

$$\mathbf{y}_{ik}|\boldsymbol{\beta}_i, \alpha, \gamma \sim N\left(\mathbf{H}\boldsymbol{\beta}_i, \alpha \begin{bmatrix} \gamma^0 & 0 & \dots & 0 \\ 0 & \gamma^1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \gamma^9 \end{bmatrix}\right)$$

- Gibbs sampler, 33 total parameters
- Conditional conjugacy with $\boldsymbol{\beta}_i$, α, τ^2
- Metropolis–Hastings for γ
- Utilize vectorization and stacking (Matrix package)
- Identifiability brought on by $\gamma^0 = 1$

Conjugacy Notes

Normal-Normal Case

$$\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\Sigma} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$$
$$\boldsymbol{\beta} \sim N_p(\boldsymbol{\mu}, \mathbf{V})$$

$$\boldsymbol{\beta}|- \sim N_p((\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X} + \mathbf{V}^{-1})^{-1}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y} + \mathbf{V}^{-1}\boldsymbol{\mu}), (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X} + \mathbf{V}^{-1})^{-1})$$

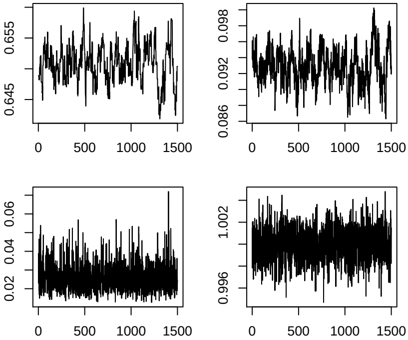
Normal-Inverse Gamma Case

$$\mathbf{y}|\boldsymbol{\mu}, \sigma^2 \sim N_n(\boldsymbol{\mu}, \sigma^2\mathbf{R})$$
$$\sigma^2 \sim IG(a, b)$$

$$\sigma^2|- \sim IG(\frac{n}{2} + a, \frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'R^{-1}(\mathbf{y} - \boldsymbol{\mu}) + b)$$

Convergence

- $\hat{R}$: minimum = 0.9996, maximum = 1.054
- 10 chains, 2000 iterations, 500 burnin
- 2 iterations per second, 16x speed increase with Matrix package
- Lowest effective sample size: 800 for γ, 1000 for α (out of 15000)



Sensitivity and Model Fit

Considerations:

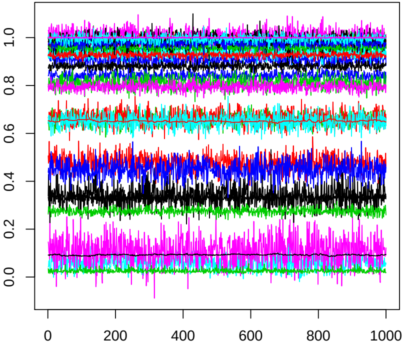
- $IG(2, 20)$ variance priors impose large expected value
- Remove γ
- Use DIC with alternate pD computation to compare fits

	Original Model	Inflated Variance Prior	Removed γ
DIC	-24041.3	-24008.2	-16271.5

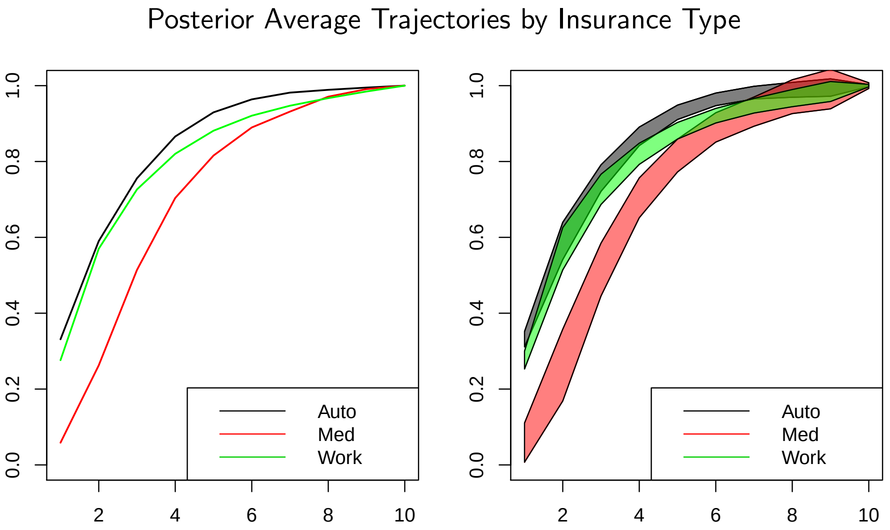
Results

Auto	$\beta_{1,1}$	$\beta_{1,2}$	$\beta_{1,3}$	$\beta_{1,4}$	$\beta_{1,5}$	$\beta_{1,6}$	$\beta_{1,7}$	$\beta_{1,8}$	$\beta_{1,9}$	$\beta_{1,10}$	α
Bayes	0.332	0.475	0.659	0.841	0.931	0.971	0.988	0.99	1.001	1	0.098
Freq	0.332	0.475	0.659	0.841	0.931	0.971	0.989	0.99	1.001	1	0.022
Med	$\beta_{2,1}$	$\beta_{2,2}$	$\beta_{2,3}$	$\beta_{2,4}$	$\beta_{2,5}$	$\beta_{2,6}$	$\beta_{2,7}$	$\beta_{2,8}$	$\beta_{2,9}$	$\beta_{2,10}$	γ
Bayes	0.059	0.096	0.337	0.67	0.807	0.913	0.936	1.017	0.971	1	0.645
Freq	0.059	0.094	0.338	0.67	0.807	0.913	0.936	1.018	0.971	1	
Work	$\beta_{3,1}$	$\beta_{3,2}$	$\beta_{3,3}$	$\beta_{3,4}$	$\beta_{3,5}$	$\beta_{3,6}$	$\beta_{3,7}$	$\beta_{3,8}$	$\beta_{3,9}$	$\beta_{3,10}$	τ^2
Bayes	0.277	0.448	0.651	0.795	0.88	0.928	0.957	0.98	0.994	1	0.646
Freq	0.277	0.447	0.651	0.795	0.88	0.928	0.957	0.98	0.994	1	

- All variables significant
- Frequentist method ignores hierarchy
- Methods match coefficients
- α reflects heteroskedasticity



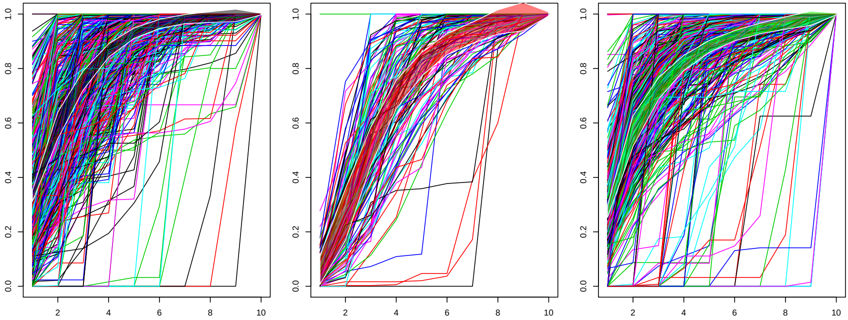
Results



Posterior Average Trajectories by Insurance Type
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10
Auto
Med
Work

Results

Posterior Average Trajectory Bands by Insurance Type



Conclusions

- Medical malpractice takes longer to reach ultimate
- Variance decay over payment years exists
- Matrix library in R
- Improvement: model differences instead of raw confidence regions
- Improvement: Experiment with restricted likelihood support
- Future: Utilize ultimate magnitude