Application of random effects in dependent compound risk model

Himchan Jeong, FSA

University of Connecticut

10 June, 2020

Opening Remarks

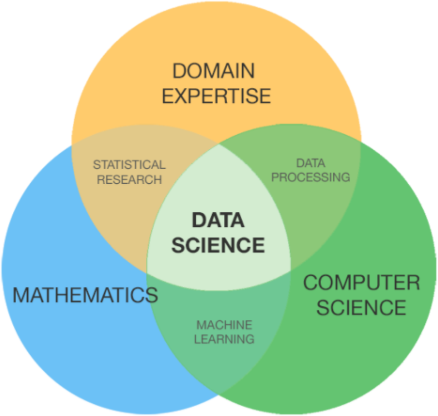
- This presentation is a compilation of
  - "Generalized linear mixed models for dependent compound risk models" (Jeong et al., accepted in *Variance*),
  - "Bayesian credibility premium with GB2 copulas" (Jeong and Valdez, in revision).
  - "Predictive compound risk models with dependence" (Jeong and Valdez, in revision).

Curriculum Vitae

2012.8	B.B.A and B.Sc in Mathematics, Seoul National University
2012.7 – 2015.1	Assistant Manager at DB Insurance
2014.6 – 2019.2	Associate of the Society of Actuaries (ASA)
2016.8	M.Sc in Statistics, Seoul National University
2019.3 – Present	Fellow of the Society of Actuaries (FSA)
2020.8 (expected)	Ph.D in Mathematics with Thesis in Actuarial Science, University of Connecticut
2020.8 – (expected)	Assistant Professor at Simon Fraser University, Department of Statistics & Actuarial Science

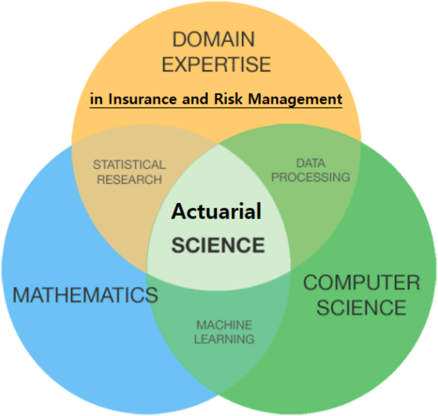
Introduction

What is Actuarial Science?



Actuarial science is "the discipline that applies mathematical and statistical methods to assess risk in insurance, finance and other industries and professions". **(Wikipedia)**

This is Actuarial Science!



Actuarial science is "the discipline that applies mathematical and statistical methods to assess risk in insurance, finance and other industries and professions". **(Wikipedia)**

Considerations in Actuarial Science

- “Transaction of Risk” is naturally involved with “Pricing of Risk’.
- In this regard, actuaries needs well-developed predictive models.

There are some more important considerations other than prediction.

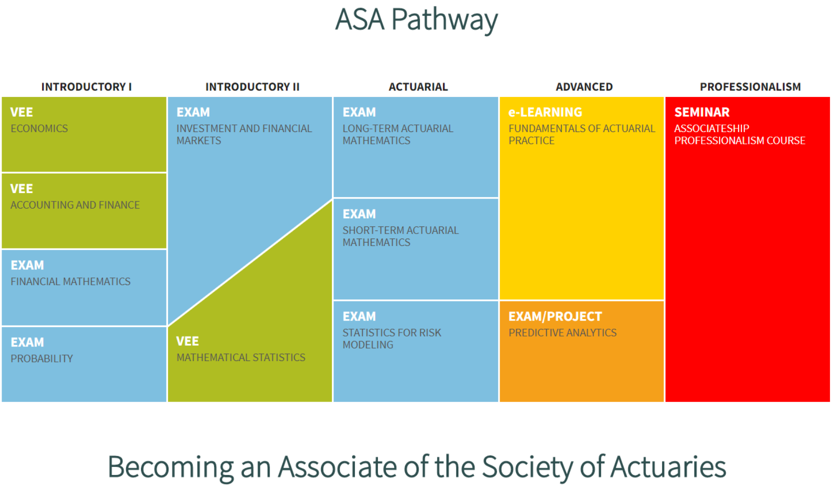
- Tradition
- Internal / External Communication
- Interpretability
- Regulation

How to get prepared as an Actuary?

As an interdisciplinary study, actuarial science is a good combination of

- Statistics,
- Mathematics,
- Economics and Finance,
- and so on.

ASA Pathway


ASA Pathway
INTRODUCTORY I
INTRODUCTORY II
ACTUARIAL
ADVANCED
PROFESSIONALISM
VEE
ECONOMICS
VEE
ACCOUNTING AND FINANCE
EXAM
FINANCIAL MATHEMATICS
EXAM
PROBABILITY
EXAM
INVESTMENT AND FINANCIAL MARKETS
VEE
MATHEMATICAL STATISTICS
EXAM
LONG-TERM ACTUARIAL MATHEMATICS
EXAM
SHORT-TERM ACTUARIAL MATHEMATICS
EXAM
STATISTICS FOR RISK MODELING
e-LEARNING
FUNDAMENTALS OF ACTUARIAL PRACTICE
EXAM/PROJECT
PREDICTIVE ANALYTICS
SEMINAR
ASSOCIATESHIP PROFESSIONALISM COURSE
Becoming an Associate of the Society of Actuaries

ASA Pathway- 7 Preliminary Exams

To be an Associate of the Society of Actuaries (ASA), you need to complete **7 Preliminary Exams** which include

- P (Probability),
- FM (Interest theory and annuity),
- IFM (Corporate finance and derivatives),
- LTAM (Life-contingent insurance pricing and reserving),
- STAM (Frequency/Severity models, Estimation, Bayesian credibility),
- SRM (Statistical learning),
- PA (Data analysis with `R/RMarkdown`).

Exam P (Probability)

- Elementary probability theory, uni/multivariate random varibles.
- *"On a block of ten houses, k are not insured. A tornado randomly damages three houses on the block. The probability that none of the damaged houses are insured is* $1/120$*. Calculate the probability that at most one of the damaged houses is insured."*

ASA Pathway - Exam P

Exam P (Probability)

- Elementary probability theory, uni/multivariate random varibles.
- *"On a block of ten houses, k are not insured. A tornado randomly damages three houses on the block. The probability that none of the damaged houses are insured is* $1/120$*. Calculate the probability that at most one of the damaged houses is insured."*

Hint: $k = 3$.

Exam FM (Financial Mathematics)

- Time value of money, deterministic annuity pricing, immunization
- *A loan of 10,000 is repaid with a payment made at the end of each year for 20 years. The payments are 100, 200, 300, 400, and 500 in years 1 through 5, respectively. In the subsequent 15 years, equal annual payments of X are made. The annual effective interest rate is 5%. Calculate X.*

Exam IFM (Investment and Financial Markets)

- Portfolio theory (Mean-variance optimization, CAPM), derivatives, ...
- *You are given the following information about the annual returns of two stocks, X and Y:*
  - i) *The expected returns of X and Y are $E[R_X] = 10\%$ and $E[R_Y] = 15\%$.*
  - ii) *The volatilities of the returns are $V_X = 18\%$ and $V_Y = 20\%$.*
  - iii) *The correlation coefficient of the returns for these two stocks is 0.25.*
  - iv) *The expected return for a certain portfolio, consisting only of stocks X and Y, is 12%.*

  Calculate the volatility of the portfolio return.

ASA Pathway - Exam LTAM

Exam LTAM (Long-Term Actuarial Mathematics)

- Survival models, premium, reserve, ...
- *You are given that* $\mathbb{P}(T \geq 40 + t | T \geq 40) = {}_t p_{40} = \exp(-0.06(1.12^t - 1))$ *for all* $t \geq 0$. *Calculate the probability that a life age 45 survives 10 years.*
- *How much of a lump-sum payment should be made as of now if a 60-years-old man wants to receive annual payments of* \$12,000 *until he passes away?*

Exam STAM (Short-Term Actuarial Mathematics)

- Loss models, credibility, P&C insurance, ...
- *You are given:*
  - (i) *Losses on a company's insurance policies follow a Pareto distribution with probability density function:*

$$f(x|\theta) = \frac{\theta}{(x+\theta)^2},\ 0 < x < \infty$$

  - (ii) *For half of the company's policies* $\theta = 100$*, while for the other half* $\theta = 300$*.*

  For a randomly selected policy, losses in Year 1 were 500. Calculate the posterior probability that losses for this policy in Year 2 will exceed 800.

Exam SRM (Statistics for Risk Modeling)

- Linear models, PCA, time series, decision tree, clustering
- *For a simple linear regression model the sum of squares of the residuals is* $\sum_{i=1}^{25} e_i^2 = 230$ *and the* R^2 *statistic is 0.64.*
  Calculate the total sum of squares (TSS) for this model.

ASA Pathway - Exam PA

Exam PA (Predictive Analytics)

- Given a dataset, you have to analyze the data, prodive descriptive statistics, fit models using the following pacakages in `R`:

boot	data.table	ggplot2	pdp	rpart
broom	devtools	glmnet	pls	rpart.plot
caret	dplyr	gridExtra	plyr	tidyverse
cluster	e1071	ISLR	pROC	xgboost
coefplot	gbm	MASS	randomForest	lme4

- Based on your analysis, recommend the most suitable model for given dataset and provide a report using `RMarkdown`.
- Both data analysis and writing a report should be finished in 5 hours.

ASA Pathway (CONT'D)

3 Validation with Educational Experiences (VEE) credits which include

- Mathematical Statistics,
- Micro/Macroeconomics,
- Accounting and Finance.

ASA Pathway (CONT'D)

Fundamental of Actuarial Practice (FAP) which consists of

- Case studies,
- Practice of writing actuarial reports (COMMUNICATION).

ASA Pathway (CONT'D)

Associateship Professionalism Course (APC) which addresses

- Professionalism,
- Ethics,
- Continuing Professional Development (CPD).

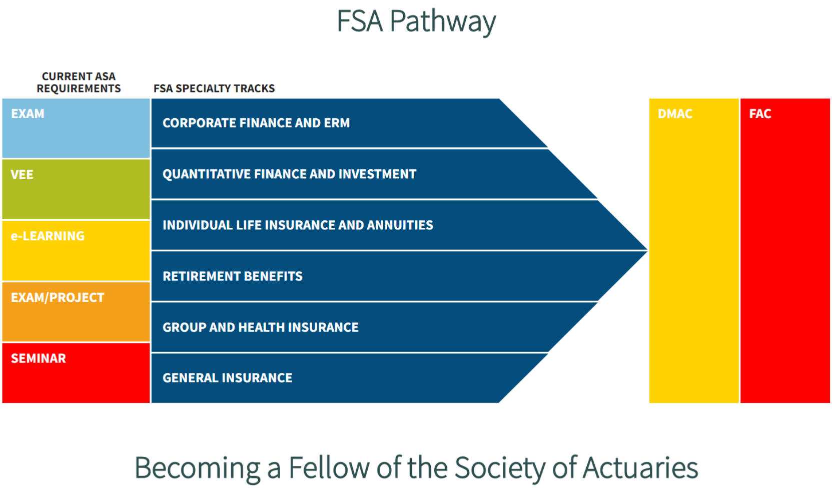
Career Paths as an Actuary

As an actuary, you can work in many fields including but not limited to

- Life insurance and annuities,
- Property and casualty insurance,
- Retirement benefits and pension plan,
- Group and health insurance,
- Enterprise risk management,
- Quantitative finance and investment,
- and Academia,

which would require further specialization with FSA designation.

FSA Pathway


FSA Pathway
CURRENT ASA REQUIREMENTS
FSA SPECIALTY TRACKS
EXAM
VEE
e-LEARNING
EXAM/PROJECT
SEMINAR
CORPORATE FINANCE AND ERM
QUANTITATIVE FINANCE AND INVESTMENT
INDIVIDUAL LIFE INSURANCE AND ANNUITIES
RETIREMENT BENEFITS
GROUP AND HEALTH INSURANCE
GENERAL INSURANCE
DMAC
FAC
Becoming a Fellow of the Society of Actuaries

Actuarial Statistics

Subjects in actuarial science are closely related to statistical methods including but not limited to the following lists:

- Life insurance: Survival analysis, Financial time series
- Property and casualty (P&C) insurance: Categorical data analysis, Longitudinal data analysis, Multivariate analysis
- Risk management: Extreme value theory, Risk measures

Preliminary

Motivating examples

In actuarial ratemaking practice, we need to consider the frequency and severity components together as follows:

- For an annual automobile insurance contract, both the number of accidents in a year and the claim amount for each accident are random.

In this case, the charged premium should be based on the total loss amount or aggregate loss amount.

Aggregate models I

Terminology:

- N: the claim count random variable and refer to its distribution as the claim-count distribution. Another term commonly used is frequency distribution.
- Y_js: the individual or single-loss random variables. Another common term for the Y_js is severity.
- S: the aggregate loss random variable or the total loss random variable.

Aggregate models II

S	N	Y_1	Y_2
0	0		
0	0		
0	0		
1500	1	1500	
0	0		
0	0		
0	0		
0	0		
0	0		
1120	2	321	899
0	0		
0	0		
0	0		
0	0		

- Modeling the distribution of N and the distribution of the Y_js has benefit for using the observed data effectively.
- It is very likely that only one or two policyholders get into accident(s) in a year.
- Therefore, we have less information for modeling the loss amount due to the car accidents.
- However, we can fully utilize observed information on the number of claim.

Aggregate models III

Definition (Collective risk model)

The collective risk model has the following representation

$$S = Y_1 + Y_2 + \cdots + Y_N, \quad N = 0, 1, 2, \ldots,$$

where the Y_js are independent and identically distributed (i.i.d.) random variables, unless otherwise specified.

The collective risk model requires the following independence assumptions:

1. Conditional on $N = n$, the random variables $Y_1, Y_2, \ldots, Y_n$ are i.i.d. random variables.
2. Conditional on $N = n$, the common distribution of the random variables $Y_1, Y_2, \ldots, Y_n$ does not depend on n.
3. The distribution of N does not depend in any way on the values of $Y_1, Y_2, \ldots$.

Aggregate models IV

- Although we assume N and Y_i are independent for computational convenience, in reality, they might not be independent.

Premium for Compound Loss under Independence

- If we assume that N and $Y_1, Y_2, \ldots, Y_n$ are independent, then we can calculate the premium for compound loss as

$$
\begin{aligned}
\mathbb{E}[S] = \mathbb{E}\left[\sum_{k=1}^{N} Y_k\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{N} Y_k | N\right]\right] \\
&= \mathbb{E}\left[N\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{N} Y_k | N\right]\right] = \mathbb{E}\left[N\mathbb{E}\left[Y|N\right]\right] \\
&= \mathbb{E}\left[N\mathbb{E}\left[Y\right]\right] = \mathbb{E}[N]\,\mathbb{E}[Y]
\end{aligned}
$$

  In other words, we just multiply the expected values from frequency model and the severity model.

- In general, $\mathbb{E}[S] \neq \mathbb{E}[N]\,\mathbb{E}[Y]$.

Why is the Dependence Important?

- If there is positive correlation between N and Y, then

$$\mathbb{E}[S] > \mathbb{E}[N]\,\mathbb{E}[Y]$$

  so a company suffers from the higher loss relative to earned premium.

- If there is negative correlation between N and Y, then

$$\mathbb{E}[S] < \mathbb{E}[N]\,\mathbb{E}[Y]$$

  so a company suffers from the loss of market share due to higher premium.

P&C Insurance Claim Data Structure

- Let N be the number of accidents and $Y_1, \cdots, Y_N$ be the size of claim per accident.
- For ratemaking in P&C, we have to predict the cost of claims
  $S_{it} = \sum_{k=1}^{N_{it}} Y_{itk}$.
- Policyholder i is followed over time $t = 1, \ldots, T_i$ years.
- Unit of analysis "*it*" – an insured driver i over time t (year)
- For each "*it*", could have several claims, $k = 0, 1, \ldots, N_{it}$
- Have available information on: number of claims N_{it}, amount of claim Y_{itk}, exposure e_{it} and covariates (explanatory variables) x_{it}
  - covariates often include age, gender, vehicle type, building type, building location, driving history and so forth

Net Premium VS Gross Premium

- Net premium is expected value of a policy's benefits so that it is given as $\mathbb{E}[S_{it}]$.
- Gross premium also account for profit margin, operation costs, and market competitiveness other than the expected value of a policy's benefits. So it is given as $(1+\alpha)\mathbb{E}[S_{it}]$.
- Throughout this presentation, we focus on the net premium, $\mathbb{E}[S_{it}]$.

The Two-Part Frequency-Severity Model

- We will model the pair (N_{it}, C_{it}) where

$$C_{it} = \begin{cases} \frac{1}{N_{it}} \sum_{k=1}^{N_{it}} Y_{itk}, & N_{it} > 0 \\ Undefined, & N_{it} = 0 \end{cases}$$

is the observed average claim size and $S_{it} = \sum_{k=1}^{N_{it}} Y_{itk}$ is the observed aggregate claim size.

- The joint density of the number of claims and the average claim size can be decomposed as

$$\begin{aligned} f(N, C|\mathbf{x}) &= f(N|\mathbf{x}) \times f(C|N, \mathbf{x}) \\ \text{joint} &= \text{frequency} \times \text{average severity}. \end{aligned}$$

Ratemaking with Predictive Analytics

$$
\begin{array}{ll}
\text{[Observed]} & \{x_{it}, N_{it}, C_{it}, e_{it} | i = 1, 2, \ldots, I, t = 1, 2, \ldots, T_i\} \\
& \qquad\quad \Big\Downarrow \ \text{Calibrated Model} \ \Big\Downarrow \\
\text{[Predicted]} & \qquad\quad N_{i,T_i+1}, C_{i,T_i+1}, \text{ and } S_{i,T_i+1}
\end{array}
$$

Premium of S under Traditional Two-Part Model

- If we assume that $\mathbb{E}\left[N_{it}|\mathbf{x}_{it}\right] = \nu_{it}$ and $\mathbb{E}\left[C_{it}|N_{it}, \mathbf{x}_{it}\right] = \mu_{it}$ where $\nu_{it} = e^{\mathbf{x}_{it}\alpha}$ and $\mu_{it} = e^{\mathbf{x}_{it}\beta}$, then we have

$$\begin{aligned}\mathbb{E}\left[S_{it}|\mathbf{x}_{it}\right] &= \mathbb{E}\left[N_{it} \times C_{it}|\mathbf{x}_{it}\right] = \mathbb{E}\left[N_{it} \times \mathbb{E}\left[C_{it}|N_{it}\right]|\mathbf{x}_{it}\right] \\ &= \mathbb{E}\left[N_{it} \times \mu_{it}|\mathbf{x}_{it}\right] = \nu_{it} \times \mu_{it}.\end{aligned}$$

- Although this simple model has been widely used in actuarial practice, we may consider the following issues:
  - Longitudinality of data
  - Possible dependence between N and C

Why Random Effects?

The following hypothetical example shows us that we might capture the unobserved heterogeneity in risk by observing the residual after controlling for the effects of observed covariates, which can be explained in terms of random effects for policyholder A and B.

Year	Gender	Age	Vehicle Size	Policyholder A		Policyholder B	
				# of Claim	Amt of Claim	# of Claim	Amt of Claim
2015	M	45	Medium	0	0	1	500
2016	M	46	Medium	0	0	2	4000
2017	M	47	Large	1	200	1	8000

Note that random effects model has a good connection with credibility theory and bonus-malus system as well.

Empirical Evidences of Dependence between N and C

Recently, evidences showing the necessity to model the dependence between N and C explicitly can be found from the following but not limited to:

- Boudreault et al. (2006),
- Hernandez-Bastida, Fernandez-Sanchez, and Gomez-Deniz (2009),
- Czado et al. (2012).

However, most of relevant works only discussed the possible dependence of N and C with cross-sectional data, ignoring the usual longitudinal property of P&C insurance data.

A Way to Capture Dependence between N and C

- Shi et al. (2015),
- Garrido et al. (2016).

They assumed $\mathbb{E}\left[C|N,\mathbf{x}\right] = e^{\mathbf{x}\beta+\gamma N}$ so that

$$\begin{aligned}\mathbb{E}\left[S|\mathbf{x}\right] &= \mathbb{E}\left[N|\mathbf{x}\right] \times \mathbb{E}\left[C|\mathbf{x}\right] \times D_N(\gamma) \\ &= \text{Freq. mean } \times \text{Sev. mean } \times \text{Dep. Adj.}\end{aligned}$$

Note that if $\gamma = 0$, then $D_N(\gamma) = 1$ so that there is no association between frequency and severity components.

Take-home Messeage

- Credibility premium of S considering the random effects as well as possible dependence between N and C

$$\begin{aligned}\mathbb{E}[S_{T+1}|\mathcal{F}_T] &= \mathbb{E}[N_{T+1}|\mathcal{F}_T] \times \mathbb{E}[C_{T+1}|\mathcal{F}_T] \times D_N(\gamma) \\ &= \text{Cred. Freq. mean} \times \text{Cred. Sev. mean} \\ &\quad \times \text{Adj. factor for dependence}\end{aligned}$$

Credibility premium of compound loss considering dependence between N and C

Frequency Component

In actuarial practice, Poisson distribution has been used for the calibration of the number of claim with the presence of covariates as follows:

$$N_{it}|\mathbf{x}_{it}, e_{it} \overset{indep}{\sim} \mathcal{P}(\nu_{it}) \quad \text{where} \;\; \nu_{it} = e_{it}\exp(\mathbf{x}_{it}\alpha). \tag{1}$$

Due to the longitudinal property of usual claim dataset, we may consider the following extension by incorporating random effects as follows:

$$N_{it}|\theta_i^N \overset{indep}{\sim} \mathcal{P}(\nu_{it}\theta_i^N) \quad \text{where} \;\; \nu_{it} = e_{it}\exp(\mathbf{x}_{it}\alpha), \;\; \theta_i^N \sim \pi_N(\theta). \tag{2}$$

Frequency Model Specifications

We calibrated frequency models with the following specifications:

- For the count of claims (frequency)
  (1) Simple Poisson GLM: $N_{it} \sim \mathcal{P}(e^{\mathbf{x}_{it}\alpha})$ so that $\mathbb{E}\left[N_{it}|\mathbf{x}_{it}\right] = e^{\mathbf{x}_{it}\alpha}$.
  (2) Poisson/Gamma Random Effect Model (= Multivariate NB Model)

Simple Poisson GLM

Data likelihood: $N_{it}|\theta_i^N \overset{indep}{\sim} \mathcal{P}(\nu_{it}\theta_i^N)$

Prior density: $\mathbb{P}(\theta_i^N = \theta) = \mathbb{1}_{\{\theta=1\}}$ for all i.

Posterior density:

$$\mathbb{P}(\theta_i^N = \theta | N_{i1} = n_{i1}, \ldots, N_{iT_i} = n_{iT_i}) \propto$$
$$\mathbb{P}(\{\theta_i^N = \theta\} \cap \{N_{i1} = n_{i1}, \ldots, N_{iT_i} = n_{iT_i}\}) \propto \mathbb{1}_{\{\theta=1\}}.$$

Predictive distribution:

$$p(N_{i,T_i+1}|N_{i1}, N_{i2}, \ldots, N_{iT_i}) =$$
$$\int p(N_{i,T_i+1}|\theta)\pi_N(\theta|N_{i1}, N_{i2}, \ldots, N_{iT_i})d\theta = p(N_{i,T_i+1}|\theta = 1).$$

Therefore, we can see that $N_{i,T_i+1}|N_{i1}, N_{i2}, \ldots, N_{iT_i} \sim \mathcal{P}(\nu_{i,T_i+1})$.

Poisson/Gamma Random Effect Model

Data likelihood: $N_{it}|\theta_i^N \overset{indep}{\sim} \mathcal{P}(\nu_{it}\theta_i^N)$

Prior density: $\pi_N(\theta) \propto \theta^{r-1}e^{-\theta r}$ so that $\theta_i^N \sim \mathcal{G}(r, 1/r)$ and $\mathbb{E}\left[\theta_i^N\right] = 1, \mathrm{V}ar\left(\theta_i^N\right) = \frac{1}{r}$.

- The range of bonus-malus factor on frequency premium is usually from 54% to 200%, according to Lemiare (1998).
- Therefore, it is natural to incorportate this knowledge on choosing the hyperparameter r for our proposed prior so that the 95% highest posterior density (HPD) interval of θ^N can include $(0.54, 2.00)$.
- Thus, $r = 3.8$ is used as the hyperparameter so that 95% HPD interval of θ^N under the proposed prior can be around $(0.16, 2.01)$.

Poisson/Gamma Random Effect Model

Posterior density:
$\theta_i^N | N_{i1}, N_{i2}, \ldots, N_{iT_i} \sim \mathcal{G}(\sum_{t=1}^{T_i} N_{it} + r, [\sum_{t=1}^{T_i} \nu_{it} + r]^{-1})$.

Predictive distribution:

$$N_{i,T_i+1} | N_{i1}, N_{i2}, \ldots, N_{iT_i} \sim \mathcal{NB}\left(\sum_{t=1}^{T_i} N_{it} + r, \frac{\nu_{i,T_i+1}}{\sum_{t=1}^{T_i+1} \nu_{it} + r} \right) \qquad (3)$$

so that $\mathbb{E}\left[N_{i,T_i+1} | N_{i1}, N_{i2}, \ldots, N_{iT_i}\right] = \frac{\sum_{t=1}^{T_i} N_{it} + r}{\sum_{t=1}^{T_i} \nu_{it} + r} \nu_{i,T_i+1}$.

Disclaimer: STATIC Random Effects Model

- In the formulation, we assumed that the random effect θ_i is random but STATIC. In other words, it is hard to capture possible evolution on the unobserved heterogeneity in risks over time with such model.
- This assumption leads to the predictive premium formula which assigns the same weight for every N_{it}, $t = 1, \ldots, T$.
- One can suggest DYNAMIC random effects model to capture possible evolution on the unobserved heterogeneity, for example, Lu (2018, *Journal of Risk and Insurance*).

Implication of Credibility Premium

- We can see that

$$\mathbb{E}\left[N_{T+1}|\mathcal{F}_T\right] = \frac{r + \sum_{t=1}^{T} N_t}{r + \sum_{t=1}^{T} \exp(\mathbf{x}_t \alpha)} \times \exp(\mathbf{x}_{T+1}\alpha)$$

$$= \frac{\text{(actual past \# of claims)}}{\text{(expected past \# of claims)}} \times \text{(expected current \# of claims)}.$$

- In this sense, use of random effects help us to penalize/reward a policyholder based on his/her driving experience.
- What about $\mathbb{E}\left[S_{T+1}|\mathcal{F}_T\right]$?

Credibility Premium of Compound Sum assuming $N \perp C$

Frangos and Vrontos (2001) derived the following formula for credibility premium of S assuming $N \perp C$.

Theorem

Suppose $(N_1, N_2, \ldots, N_t)$ follows MVNB distribution as defined in (3). Moreover, let us assume that $\mathbb{E}\left[C_t | n_t\right] = \theta^C e^{\mathbf{x}_t \beta} = \theta^C \mu_t$. Then credibility premium of $S_{i,T+1}$ given $N_1, \ldots, N_T, C_1, \ldots, C_T$ is given as follows:

$$\begin{aligned}\mathbb{E}[S_{T+1}|\mathcal{F}_T] &= \mathbb{E}[N_{T+1}|N_1, \ldots, N_T] \times \mathbb{E}[C_{T+1}|C_1, \ldots, C_T] \\ &= \textit{Cred. Freq. mean} \times \textit{Cred. Sev. mean}\end{aligned}$$

Credibility Premium of Compound Sum assuming $N \perp C$

Frangos and Vrontos (2001) derived the following formula for credibility premium of S assuming $N \perp C$.

Theorem

Suppose $(N_1, N_2, \ldots, N_t)$ follows MVNB distribution as defined in (3). Moreover, let us assume that $\mathbb{E}\left[C_t | n_t\right] = \theta^C e^{\mathbf{x}_t \beta} = \theta^C \mu_t$. Then credibility premium of $S_{i,T+1}$ given $N_1, \ldots, N_T, C_1, \ldots, C_T$ is given as follows:

$$\begin{aligned}\mathbb{E}[S_{T+1}|\mathcal{F}_T] &= \mathbb{E}[N_{T+1}|N_1, \ldots, N_T] \times \mathbb{E}[C_{T+1}|C_1, \ldots, C_T] \\ &= \textit{Cred. Freq. mean} \times \textit{Cred. Sev. mean}\end{aligned}$$

What if $N \not\perp C$?

Severity Component

Traditionally, gamma distribution has been used for the calibration of the amount of individual claim with the presence of covariates as follows:

$$C_{it}|\mathbf{x}_{it}, N_{it} \overset{indep}{\sim} \mathcal{G}(\psi_{it}, \mu_{it}/\psi_{it}) \quad \text{where} \;\; \mu_{it} = \exp(\mathbf{x}_{it}\beta), \;\; \psi_{it} = N_{it}/\phi. \tag{4}$$

Note that conditioning argument on $\mathbf{x}_{it}, N_{it}$ is suppresed afterwards for notational convenience. Again, the longitudinal property of usual claim dataset allows us to consider the unobserved heterogeneity of the policyholders via random effects as follows:

$$\begin{aligned} &C_{it}|\theta_i^C \overset{indep}{\sim} \mathcal{G}(\psi_{it}, \theta_i^C\mu_{it}/\psi_{it}) \quad \text{where} \;\; \mu_{it} = \exp(\mathbf{x}_{it}\beta + \gamma N_{it}), \\ &\psi_{it} = N_{it}/\phi, \;\; \theta_i^C \sim \pi_C(\theta). \end{aligned} \tag{5}$$

Severity Model Specifications

We calibrated severity models with the following specifications:

- For the average size of claims (severity)
  (1) Simple Gamma GLM: $C_{it}|N_{it} \sim \mathcal{G}(\frac{n_{it}}{\phi}, e^{\mathbf{x}_{it}\beta + n_{it}\gamma}\frac{\phi}{n_{it}})$ so that $\mathbb{E}\left[C_{it}|N_{it}, \mathbf{x}_{it}\right] = e^{\mathbf{x}_{it}\beta + n_{it}\gamma}$ and $\frac{\mathrm{Var}(C_{it}|N_{it}, \mathbf{x}_{it})}{\mathbb{E}[C_{it}|N_{it}, \mathbf{x}_{it}]^2} = \frac{\phi}{n_{it}}$
  (2) Gamma/Normal Random Effect Model (= Gamma GLMM): $C_{it}|N_{it}, \theta_i^C \sim \mathcal{G}(\frac{n_{it}}{\phi}, \theta_i^C e^{\mathbf{x}_{it}\beta + n_{it}\gamma}\frac{\phi}{n_{it}})$ where $\log\theta_i^C \sim \mathcal{N}(-\sigma^2/2, \sigma^2)$ so that $\mathbb{E}\left[C_{it}|N_{it}, \mathbf{x}_{it}, \theta_i^C\right] = \theta_i^C e^{\mathbf{x}_{it}\beta + n_{it}\gamma}$ and $\frac{\mathrm{Var}\left(C_{it}|N_{it}, \mathbf{x}_{it}, \theta_i^C\right)}{\mathbb{E}\left[C_{it}|N_{it}, \mathbf{x}_{it}, \theta_i^C\right]^2} = \frac{\phi}{n_{it}}$
  (3) Gamma/I-gamma Random Effect Model (= Multivariate GP Model)
  (4) G-gamma/GI-gamma Random Effect Model (= Multivariate GB2 Model)

Simple Gamma GLM

Data likelihood: $C_{it}|\theta_i^C \overset{indep}{\sim} \mathcal{G}(\psi_{it}, \theta_i^C \mu_{it}/\psi_{it})$

Prior density: $\mathbb{P}(\theta_i^C = \theta) = \mathbb{1}_{\{\theta=1\}}$ for all i.

Posterior density:

$$\mathbb{P}(\theta_i^C = \theta | C_{i1} = c_{i1}, \ldots, C_{iT_i} = c_{iT_i}) \propto$$
$$\mathbb{P}(\{\theta_i^C = \theta\} \cap \{C_{i1} = c_{i1}, \ldots, C_{iT_i} = c_{iT_i}\}) \propto \mathbb{1}_{\{\theta=1\}}.$$

Predictive distribution:

$$f(C_{i,T_i+1}|C_{i1}, C_{i2}, \ldots, C_{iT_i}) = f(C_{i,T_i+1}|\theta = 1).$$

Therefore, we can see that
$C_{i,T_i+1}|C_{i1}, C_{i2}, \ldots, C_{iT_i} \sim \mathcal{G}(\psi_{i,T_i+1}, \mu_{i,T_i+1}/\psi_{i,T_i+1})$.

Gamma/I-gamma Random Effect Model

Data likelihood: $C_{it}|\theta_i^C \overset{indep}{\sim} \mathcal{G}(\psi_{it}, \theta_i^C \mu_{it}/\psi_{it})$

Prior density: $\pi_C(\theta) \propto \theta^{-k-2} e^{-k/\theta}$ so that $\theta_i^C \sim \mathcal{IG}(k+1, k)$ and $\mathbb{E}\left[\theta_i^C\right] = 1, \mathrm{V}ar\left(\theta_i^C\right) = \frac{1}{k-1}$.

- According to the Lemaire (1998), most of countries except for Korea do not use historically observed claim amount for the construction of penalty or bonus on a policyholder, which support the assertion that there is less variablity on θ^C than on θ^N.
- Therefore, $k = 11$ is used so that the 95% HPD interval of θ^C under the proposed prior can be around $(0.49, 1.61)$, which is narrower than the 95% HPD interval of θ^N under the proposed prior.

Gamma/I-gamma Random Effect Model

Posterior density:
$\theta_i^C | C_{i1}, C_{i2}, \ldots, C_{iT_i} \sim \mathcal{IG}(k + \sum_{t=1}^{T_i} \psi_{it} + 1, \sum_{t=1}^{T_i} \frac{\psi_{it} C_{it}}{\mu_{it}} + k)$.

Predictive distribution:

$$
\begin{aligned}
& C_{i,T_i+1} | C_{i1}, C_{i2}, \ldots, C_{iT_i} \\
& \sim \mathcal{GP}\left(k + \sum_{t=1}^{T_i} \psi_{it} + 1, \left[k + \sum_{t=1}^{T_i} \psi_{i,t} \frac{C_{it}}{\mu_{it}}\right] \frac{\mu_{i,T_i+1}}{\psi_{i,T_i+1}}, \psi_{i,T_i+1}\right),
\end{aligned}
$$

with predictive mean

$$
\mathbb{E}\left[C_{i,T_i+1} | C_{i1}, C_{i2}, \ldots, C_{iT_i}\right] = \frac{\left[k\phi + \sum_{t=1}^{T_i} S_{it}/\mu_{it}\right]}{k\phi + \sum_{t=1}^{T_i} N_{it}} \mu_{i,T_i+1}
$$

since $\psi_{it} C_{it} = N_{it} C_{it}/\phi = S_{it}/\phi$.

G-gamma/GI-gamma Random Effect Model

- G-gamma/GI-gamma random effect model is given as follows by denoting $w = \dfrac{\Gamma(k+1)}{\Gamma(k+1-1/p)}$:

$$C_{it}|N_{it}, \theta_i^C \sim \mathcal{GG}\left(\frac{n_{it}}{\phi}, \theta_i^C \mu_{it} \frac{\Gamma(n_{it}/\phi)}{\Gamma(n_{it}/\phi + 1/p)}, p\right) \text{ and } \theta_i^C \sim \mathcal{GIG}(k+1, w, p) \tag{6}$$

Note that if $p = 1$, then G-gamma and GI-gamma are equivalent to gamma and inverse gamma, respectively.

G-gamma/Gl-gamma Random Effect Model

Predictive distribution:

$$
\begin{aligned}
& C_{i,T_i+1} | C_{i1}, C_{i2}, \ldots, C_{iT_i} \\
& \sim \mathcal{GB}2\left(k_{T_i},\ w^*_{T_i,p} \frac{\mu_{i,T_i+1}}{z_{i,T_i+1}},\ \psi_{i,T_i+1},\ p \right),
\end{aligned}
$$

with predictive mean

$$
\mathbb{E}\left[C_{i,T_i+1} | C_{i1}, C_{i2}, \ldots, C_{iT_i} \right] = w^*_{T_i,p} \frac{\Gamma(k_{T_i} - 1/p)}{\Gamma(k_{T_i})} \mu_{i,T_i+1}
$$

where $k_{T_i} = k + \sum_{t=1}^{T_i} \psi_{it} + 1$, $z_{it} = \dfrac{\Gamma(n_{it}/\phi + 1/p)}{\Gamma(n_{it}/\phi)}$, and
$w^*_{T_i,p} = [w^p + \sum_{t=1}^{T} (z_{it} c_{it} \mu_{it}^{-1})^p]^{1/p}$.

GP and GB2 Distributions

- $Y \sim \mathcal{GP}(a, \xi, \tau)$

$$\Longrightarrow \quad f(y|a, \xi, \tau) = \frac{\Gamma(a+\tau)}{\Gamma(a)\Gamma(\tau)} \frac{\xi^a y^{\tau-1}}{(y+\xi)^{a+\tau}}$$

- $Y \sim \mathcal{GB}2(a, \xi, \tau, p)$

$$\Longrightarrow \quad f(y|a, \xi, \tau, p) = \frac{\Gamma(a+\tau)}{\Gamma(a)\Gamma(\tau)} |p| \frac{\xi^{ap} y^{\tau p-1}}{(y^p+\xi^p)^{a+\tau}}$$

- Note that GP distribution is a special case of GB2 distribution where $p = 1$. (Loss Models, Klugman et al. 2012)

Credibility Premium for the Compound Sum when $N \not\perp C$

Theorem

Suppose $(N_1, N_2, \ldots, N_t)$ follows MVNB distribution as defined in (3). Moreover, let us assume that $\mathbb{E}\left[C_t | n_t\right] = \theta^C e^{\mathbf{x}_t \beta} e^{n_t \gamma} = \theta^C \tilde{\mu}_t e^{n_t \gamma}$. Then, the credibility premium of S_{T+1} is given as follows:

$$
\begin{aligned}
\mathbb{E}\left[S_{T+1} | \mathbf{n}_T, \mathbf{c}_T\right] = \mathbb{E}\left[\tilde{\mu}_{T+1}\theta^C | \mathbf{n}_T, \mathbf{c}_T\right] \times \mathbb{E}\left[\nu_{T+1}\theta^N | \mathbf{n}_T\right] \\
\times e^{\gamma}\left[1 - \left(\frac{\nu_{T+1}}{\hat{r}_T}\right)(e^{\gamma} - 1)\right]^{-r_T - 1}.
\end{aligned}
\tag{7}
$$

If $\gamma = 0$, then (7) is reduced as in Frangos and Vrontos (2001):

$$
\mathbb{E}\left[S_{T+1} | \mathbf{n}_T, \mathbf{c}_T\right] = \mathbb{E}\left[\mu_{T+1}\theta^C | \mathbf{n}_T, \mathbf{c}_T\right] \times \mathbb{E}\left[\nu_{T+1}\theta^N | \mathbf{n}_T\right].
$$

Credibility Premium for the Compound Sum when $N \not\perp C$

- According to the theorem, it can be observed that the credibility premium of S_{T+1} can be expressed as the product of
  - credibility premium of N_{T+1},
  - credibility premium of C_{T+1}, and
  - adjustment factor which accounts for the dependence between frequency and the average severity.
- Under the MVNB distribution defined in (3), we have the following:

$$\mathbb{E}\left[\nu_{T+1}\theta^N \middle| \mathbf{n}_T\right] = \frac{r + \sum_{t=1}^T n_t}{r + \sum_{t=1}^T \nu_t}\nu_{T+1}.$$

- For convenience, let us denote $D_N(\gamma) = e^\gamma \left[1 - \left(\frac{\nu_{T+1}}{\hat{r}_T}\right)(e^\gamma - 1)\right]^{-r_T-1}$ so that $D_N(0) = 1$.

Credibility Premium for the Compound Sum when $N \not\perp C$

Corollary

Suppose $(N_1, N_2, \ldots, N_t)$ follows MVNB distribution as defined in (3). If the average severity component follows Gamma GLM, then the credibility premium of S_{T+1} is given as follows:

$$\mathbb{E}\left[S_{T+1} | \mathbf{n}_T, \mathbf{c}_T\right] = e^{\mathbf{x}_{T+1}\beta} \times \mathbb{E}\left[e^{\mathbf{x}_{T+1}\alpha\theta^N} | \mathbf{n}_T\right] \times D_N(\gamma).$$

If the average severity component follows Gamma GLMM, then we have

$$\mathbb{E}\left[S_{T+1} | \mathbf{n}_T, \mathbf{c}_T\right] \simeq \frac{\sum_{j=1}^{J} \hat{\theta}_j \, f(\mathbf{c}_T | \mathbf{n}_T, \hat{\theta}_j)}{\sum_{j=1}^{J} f(\mathbf{c}_T | \mathbf{n}_T, \hat{\theta}_j)} e^{\mathbf{x}_{T+1}\beta} \times \mathbb{E}\left[e^{\mathbf{x}_{T+1}\alpha\theta^N} | \mathbf{n}_T\right] \times D_N(\gamma),$$

where $\hat{\theta}_j$'s are generated from $\mathcal{LN}\left(-\sigma^2/2, \, \sigma^2\right)$.

Credibility Premium for the Compound Sum when $N \not\perp C$

Corollary

If the average severity component follows MVGP Model, then we have

$$\mathbb{E}\left[S_{T+1}|\mathbf{n}_T, \mathbf{c}_T\right] = \frac{k\phi + \sum_{t=1}^{T} S_t/\mu_t}{k\phi + \sum_{t=1}^{T} n_t} e^{\mathbf{x}_{T+1}\beta} \times \mathbb{E}\left[e^{\mathbf{x}_{T+1}\alpha}\theta^N|\mathbf{n}_T\right] \times D_N(\gamma).$$

Finally, if the average severity component follows MVGB2 Model, then we have

$$\mathbb{E}\left[S_{T+1}|\mathbf{n}_T, \mathbf{c}_T\right] = \left(\sum_{t=1}^{T}\left(\frac{S_t}{\mu_t}\frac{\Gamma(n_t/\phi + 1/p)}{\phi\Gamma(n_t/\phi+1)}\right)^p + w^p\right)^{1/p} \frac{\Gamma(k_T + 1 - 1/p)}{\Gamma(k_T+1)}$$
$$\times\, e^{\mathbf{x}_{T+1}\beta} \times \mathbb{E}\left[e^{\mathbf{x}_{T+1}\alpha}\theta^N|\mathbf{n}_T\right] \times D_N(\gamma).$$

Credibility Premium for the Compound Sum when $N \not\perp C$

Theorem

For $D_N(\gamma)$ in (7), the following are true:

(i) $D_N(\gamma)$ *is well-defined if and only if* $\gamma \leq \log(1 + \tilde{r}_T/\nu_{T+1})$.

(ii) $D_N(\gamma)$ *is a strictly increasing function of* γ.

(iii)

$$D_N(\gamma) = e^{\gamma}\left[1 - \left(\frac{\nu_{T+1}}{\tilde{r}_T}\right)(e^{\gamma} - 1)\right]^{-r_T - 1} \begin{cases} = 1 & \text{if} \quad \gamma = 0 \\ > 1 & \text{if} \quad \gamma > 0 \\ < 1 & \text{if} \quad \gamma < 0 \end{cases}.$$

In this regard, $D_N(\gamma)$ can be interpreted as an informative measure of dependence which considers not only the sign of dependence but also the magnitude of dependence.

Observable Policy Characteristics used as Covariates

Categorical variables	Description		Proportions	
VehType	Type of insured vehicle:	Car	99.27%	
		Motorbike	0.47%	
		Others	0.27%	
Gender	Insured's sex:	Male $= 1$	80.82%	
		Female $= 0$	19.18%	
Cover Code	Type of insurance cover:	Comprehensive $= 1$	78.65%	
		Others $= 0$	21.35%	
Continuous variables		**Minimum**	**Mean**	**Maximum**
VehCapa	Insured vehicle's capacity in cc	10.00	1587.44	9996.00
VehAge	Age of vehicle in years	-1.00	6.71	48.00
Age	The policyholder's issue age	18.00	44.46	99.00
NCD	No Claim Discount in %	0.00	35.67	50.00

- Singapore insurance data (1993–2000: Training, 2001–2002: Test)
- $M = 50,215$ unique policyholder, 162,179 of aggregated total number of observations observed on training set.

Frequency Estimation Results

	Poisson			MVNB		
	Est	s.e	Pr(>|t|)	Est	s.e	Pr(>|t|)
(Intercept)	-4.33	0.40	0.00	-4.93	0.52	0.00
VTypeCar	0.19	0.19	0.33	1.44	0.37	0.00
VTypeMBike	-1.41	0.49	0.00	-1.83	0.92	0.05
log(VehCapa)	0.33	0.03	0.00	0.20	0.04	0.00
VehAge	-0.02	0.00	0.00	-0.02	0.00	0.00
SexM	0.11	0.02	0.00	0.09	0.02	0.00
Comp	0.81	0.04	0.00	0.74	0.04	0.00
Age	-0.03	0.02	0.12	-0.01	0.01	0.37
Age ^{2}	0.00	0.00	0.36	0.00	0.00	0.37
Age ^{3}	0.00	0.00	0.77	0.00	0.00	0.34
NCD	-0.01	0.00	0.00	-0.01	0.00	0.00
Loglikelihood		-49565.37			-49494.62	
AIC		99152.75			99013.24	
BIC		99274.70			98989.24	

Severity Estimation Results

	Gamma GLM		Gamma GLMM		MVGP		MVGB2	
	Est	Pr(>|t|)	Est	Pr(>|t|)	Est	Pr(>|t|)	Est	Pr(>|t|)
(Intercept)	7.61	0.00	6.43	0.00	9.49	0.00	9.74	0.00
VTypeCar	-0.29	0.55	0.12	0.62	-0.42	0.00	-0.20	0.07
VTypeMBike	2.87	0.03	2.32	0.00	5.04	0.02	5.94	0.00
logVehCapa	0.53	0.00	0.33	0.00	0.28	0.00	0.22	0.00
VehAge	-0.03	0.00	-0.01	0.00	-0.02	0.00	-0.01	0.00
SexM	-0.01	0.91	-0.02	0.49	-0.05	0.11	-0.09	0.00
Comp	0.05	0.60	0.19	0.00	0.06	0.00	0.23	0.00
Age	-0.16	0.00	-0.05	0.03	-0.16	0.00	-0.17	0.00
Age ^{2}	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
Age ^{3}	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
NCD	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00	0.00
Count	-0.12	0.01	0.01	0.65	-0.09	0.00	-0.07	0.00
Loglikelihood		-138605.00		-133760.00		-125641.00		-125055.00
AIC		277236.83		267548.24		251309.61		250138.04
BIC		277334.50		267653.42		251407.28		250243.22

Validation of Severity Models

Validation of the fitted models can be performed in terms of

- Root mean squared error (RMSE), Mean absolute error (MAE): For equilibrium on earned premium and actual claim amount in total

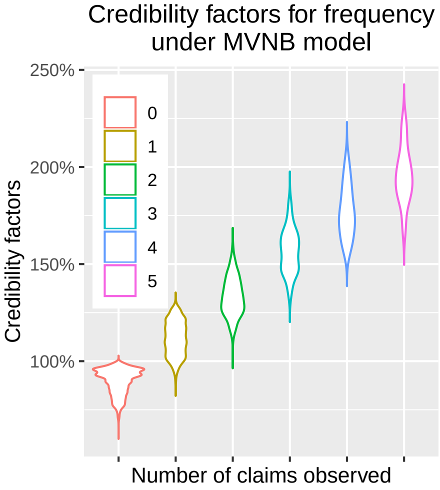
$$\text{RMSE} := \sqrt{\frac{1}{N}\sum_{n=1}^{N}(y_i - \hat{y}_i)^2}, \quad \text{MAE} := \frac{1}{N}\sum_{n=1}^{N}|y_i - \hat{y}_i|.$$

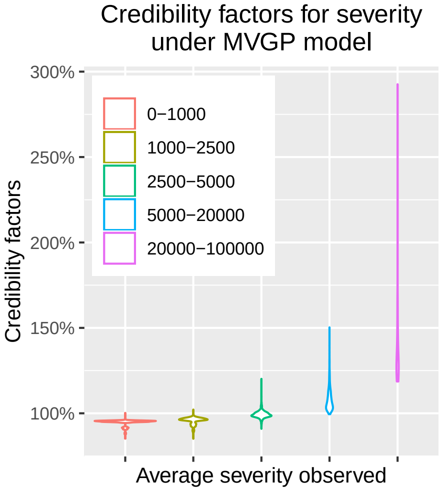
- Predictive distribution: For enterprise risk management

Validation Results for Compound Sum

	Gamma GLM	Gamma GLMM	MVGP	MVGB2
RMSE	1967.3387	1967.1304	1964.729	1975.3112
MAE	475.5807	455.8223	433.621	451.7382

Credibility Factors for Frequency and Severity Components

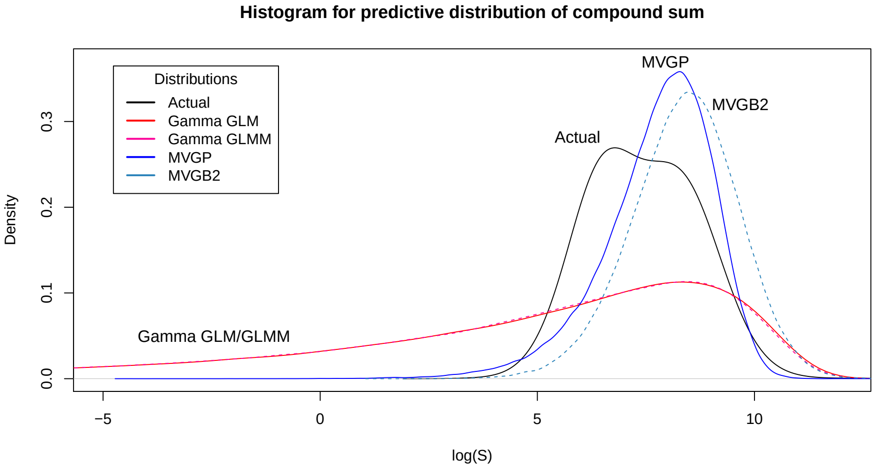




Credibility Factors for Frequency and Severity Components

	Observed claims		Credibility factors		
ID	Counts	Total losses	Frequency	Severity	Total
1614	0	0	83%	100%	83%
371	2	1173	154%	92%	141%
1816	1	32250	114%	148%	169%

Predictive Distribution of Compound Sum



Concluding Remarks

- In this article, we explored a ratemaking framework which considers longitudinality as well as dependence between frequency and severity components.
- The results show us that proposed MVGP distribution outperforms naive Gamma GLM and Gamma GLMM in terms of better prediction and analytical tractability for severity component.
- Furthermore, we could obtain more general formula for credibility premium of the compound sum, incorporating the possible dependence between the frequency and severity conponents.