A Multivariate Analysis of Intercompany Loss Triangles

Peng Shi

School of Business
University of Wisconsin-Madison

ASTIN Colloquium
May 21-24, 2013
Outline

• Introduction: background and motivation
• Modeling
  • Data distribution
  • Bayesian hierarchical model
  • Model assessment
• Data analysis
  • NAIC schedule P
  • Estimation and inference
• Predictive applications
  • Reserving variability
  • Reinsurance example
• Concluding remarks
Introduction

Background

- Loss reserves: the technical provisions to support outstanding liabilities of a property casualty insurer
  - Reserves represent the largest balance sheet liability
  - Loss reserving is a classic actuarial problem
  - Improper reserving could be detrimental
### Introduction

### Background

- An example of run-off triangle

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Premiums</th>
<th>Development Lag</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>1988</td>
<td>267,666</td>
<td>33,810</td>
</tr>
<tr>
<td>1989</td>
<td>274,526</td>
<td>37,663</td>
</tr>
<tr>
<td>1990</td>
<td>268,161</td>
<td>40,630</td>
</tr>
<tr>
<td>1991</td>
<td>276,821</td>
<td>40,475</td>
</tr>
<tr>
<td>1992</td>
<td>270,214</td>
<td>37,127</td>
</tr>
<tr>
<td>1993</td>
<td>280,568</td>
<td>41,125</td>
</tr>
<tr>
<td>1994</td>
<td>344,915</td>
<td>57,515</td>
</tr>
<tr>
<td>1995</td>
<td>371,139</td>
<td>61,553</td>
</tr>
<tr>
<td>1996</td>
<td>323,753</td>
<td>112,103</td>
</tr>
<tr>
<td>1997</td>
<td>221,448</td>
<td>37,554</td>
</tr>
</tbody>
</table>
Literature

- **Univariate loss reserving**
  - Chain ladder and others

- **Multivariate loss reserving**
  - Naive approach: the “silhouette” method
  - Additivity issue (see Ajne (1994)) attracts more attention recently
    - EU capital adequacy regime: Solvency II
    - CAS loss reserve dependency working party
Recent literature emphasizes dependencies among triangles

- Distribution-Free approach
  - Multivariate additive model, e.g. Hess et al. (2006), Merz and Wüthrich (2009)

- Parametric approach
  - Parametric distributions, e.g. Shi et al. (2012)
  - Copula approach, e.g. Shi and Frees (2011), de Jong (2012)
Motivation

- Prediction of insurance liabilities often requires aggregating the experiences of loss payment from multiple insurers
  - to borrow information for lines of business from other insurers
  - to identify industry-wide under- or over-reserving problem
  - to predict outstanding liabilities for a reinsurer
- The resulting dataset of intercompany loss triangles displays a multilevel structure of claim development
  - a portfolio consists of a group of insurers
  - each insurer several lines of business
  - and each line various cohorts of claims
- Our goal is to propose a Bayesian hierarchical model to analyze intercompany claim triangles, accommodating association within and between insurers
Some Notations

- **Index**
  - \( n = 1, \ldots, N \) indicates the \( n \)-th insurer;
  - \( l = 1, \ldots, L \) indicates the \( l \)-th line of business
  - \( i = 1, \ldots, I \) indicates the \( i \)-th accident year
  - \( j = 1, \ldots, J(= l) \) indicates the \( j \)-th valuation point \( t_j \)

- **Variable of interest**
  - \( X_{n,l,i}(t_j) \) denotes incremental paid losses
  - \( \omega_{n,l,i} \) denotes the exposure in accident year \( i \)
  - We normalize incremental payments by \( Y_{n,l,i}(t_j) = X_{n,l,i}(t_j)/\omega_{n,l,i} \)

- **Index set** \( \{(i,j) : i + j \leq l + 1\} \) divides data into two subsets
  - \( \mathcal{D}_I \): information available by calendar year \( l + 1 \)
  - \( \mathcal{D}_I^c \): future payments in years \( t = i + j > l + 1 \)
Data Distribution - Marginal

- Assume a parametric distribution for $Y_{n,l,i}(t_j)$, for example:
  - $Y_{n,l,i}(t_j) \sim F_{n,l}(\cdot; \eta_{n,l,i,j}, \phi_{n,l})$
  - $\eta_{n,l,i,j}$ determines the location such that $\eta_{n,l,i,j} = g(\mu_{n,l,i}(t_j))$
  - Vector $\phi_{n,l}$ summarizes additional parameters

- Alternative models for location
  - Parametric with two factors
    $$g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + \zeta_{n,l,j}$$
  - Semiparametric regression
    $$g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + s_{n,l}(t_j)$$
Data Distribution - Joint

- Dependency among multiple lines is accommodated by a copula
- Distribution of \((Y_{n,1,i(t_j)}, \cdots, Y_{n,L,i(t_j)})\) has the following copula representation

\[
F_n(y_{n,1,i,j}, \cdots, y_{n,L,i,j}) = \text{Prob}(Y_{n,1,i(t_j)} \leq y_{n,1,i,j}, \cdots, Y_{n,L,i(t_j)} \leq y_{n,L,i,j}) = H_n(F_{n,1}(y_{n,1,i,j}; \eta_{n,1,i,j}, \phi_{n,1}), \cdots, F_{n,L}(y_{n,L,i,j}; \eta_{n,L,i,j}, \phi_{n,L}); \rho_n)
\]
Multilevel Structure

- Consider a Bayesian hierarchical model
  - allows insurers to learn from each other
  - provides predictive distribution

\[
f(y^{D_i} | y^{D_I}) = \int f(y^{D_i} | \Theta) f(\Theta | y^{D_I}) d\Theta
\]

- For instance \( g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + \zeta_{n,l,j} \)
  - \( \delta_{n,l} \sim N(0, \sigma_\delta^2[l]) \) for \( n = 1, \ldots, N \)
  - \( \alpha_{n,l,i} \sim N(0, \sigma_\alpha^2[l,i]) \) for \( n = 1, \ldots, N \)
  - \( \sigma_\alpha^2[l,i] \sim IG(\psi_\alpha[l], \psi_\alpha[l]) \) for \( i = 1, \ldots, l \)
Model Assessment

- For training data, we consider logarithm of the pseudo-marginal likelihood (LPML) statistic
  - Denote $\mathbf{y}_{n,i,j} = (y_{n,1,i,j}, \ldots, y_{n,L,i,j})$
  - Define $CPO_{n,i,j} = f(y_{n,i,j}|\mathbf{y}^D_{n,i,j}) = \int f(y_{n,i,j}|\Theta)f(\Theta|\mathbf{y}^D_{n,i,j})d\Theta$
  - Calculate $CPO^M = \sum \log CPO_{n,i,j}$

- For validation data, we consider $L$-criterion
  - $L$-measure $= \frac{1}{S-b} \sum_{r=b+1}^{S} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{\{(i,j):i+j>l+1\}} \left( [y_{n,l,i,j}]_r - y_{n,l,i,j} \right)^2$
  - Can be evaluated using the basis of either paid losses or loss ratio
Data

- NAIC Schedule P
- Data preparation
  - Triangles constructed from Schedule P of year 1997
  - Future payments in lower triangles from year 1998-2006

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Premium</th>
<th>Development Lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1989</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1990</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1991</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1992</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1993</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1994</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1995</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1996</td>
<td>XXX</td>
<td>XXX</td>
</tr>
<tr>
<td>1997</td>
<td>XXX</td>
<td>XXX</td>
</tr>
</tbody>
</table>

Development Lag:
- 1998
- 1999
- 2000
- 2001
- 2002
- 2003
- 2004
- 2005
- 2006
Data

- Consider a hypothetical portfolio
  - Five insurers
  - Each with personal and commercial auto lines
- Triangles of paid losses
Model Specification

- **Sampling distribution**
  - Lognormal model for incremental payment, i.e.
    \[ Y_{n,l,i}(t_j) \sim LN(\eta_{n,l,i}(t_j), \sigma_{n,l}^2) \]

- Parametric regression for personal auto
- Penalized regression spline for commercial auto

- \( s_{n,l}(t_j; \theta_{n,l}) = \beta_{n,l} \times t_j + \sum_{k=1}^{K} \lambda_{n,l,k}|t_j - \nu_k|^3 \)

- \( \nu_1 < \nu_2 < \cdots < \nu_K \) are fixed knots, could be the \( k/(K+1) \)th sample quantile of covariate \( t_j \)

- \( \sum_{k=1}^{K} \lambda_{n,l,k}^2 < \tau \) (\( \tau \) is a constant) to penalize the roughness of the fit

- Frank copula to join multiple lines
  \[ H_n(u, v; \rho_n) = -\frac{1}{\rho_n} \ln \left( 1 + \frac{(e^{-\rho_n u} - 1)(e^{-\rho_n v} - 1)}{e^{-\rho_n} - 1} \right) \]
Inference

- Vague priors are used in the inference
- Run 50,000 MCMC iterations in two parallel chains
- First 40,000 iterations in each chain discarded as burn-in sample
- Some selected results
  - Left panel: $\sigma^2$ in log-normal
  - Right panel: $\rho$ in Frank copula
Model Comparison

- Consider three models
  - Model 1: Assume independence among business lines and no learning across insurers
  - Model 2: Allow for dependence among business lines within each insurer but no learning across insurers
  - Model 3: Allow for dependence among business lines within each insurer and information sharing between insurers

<table>
<thead>
<tr>
<th>Model</th>
<th>LPML</th>
<th>Amount</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>51.64</td>
<td>7.23e+06</td>
<td>19.35</td>
</tr>
<tr>
<td>Model 2</td>
<td>60.62</td>
<td>4.89e+06</td>
<td>14.20</td>
</tr>
<tr>
<td>Model 3</td>
<td>66.66</td>
<td>1.42e+06</td>
<td>2.74</td>
</tr>
</tbody>
</table>
Reserving Variability

- Define reserves as $R = g(Y^{Di}_c)$

- Quantities of interest could be:
  - accident year reserves
  - calendar year reserves
  - reserves by business line
  - firm-level reserves
  - ...

- Bayesian model provides predictive distributions of $Y^{Di}_c$, and thus of reserves $R$
Reserving Variability

Company #2

Company #3

Peng Shi (Wisconsin School of Business)
Reserving Variability

- Distribution-free approaches rely on conditional mean squared error of prediction (MSEP)

\[
\text{MSEP}_{R|D_I} = E \left[ (R - \hat{R}^B)^2 | D_I \right]
\]

- Given \( \hat{R}^B = E[E(R|\Theta)|D_I] = E(R|D_I) \)

\[
\text{MSEP}_{R|D_I} = \text{Var}(R|D_I) = E \left[ \text{Var}(R|\Theta)|D_I \right] + \text{Var} \left[ E(R|\Theta)|D_I \right]
\]

variability = process variance + estimation error (PV) (ER)
## Reserving Variability

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Calendar Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\sqrt{ER})</td>
</tr>
<tr>
<td>1989</td>
<td>3,001</td>
</tr>
<tr>
<td>1990</td>
<td>6,544</td>
</tr>
<tr>
<td>1991</td>
<td>10,511</td>
</tr>
<tr>
<td>1992</td>
<td>17,314</td>
</tr>
<tr>
<td>1993</td>
<td>28,482</td>
</tr>
<tr>
<td>1994</td>
<td>57,676</td>
</tr>
<tr>
<td>1995</td>
<td>127,717</td>
</tr>
<tr>
<td>1996</td>
<td>319,228</td>
</tr>
<tr>
<td>1997</td>
<td>820,240</td>
</tr>
</tbody>
</table>

Peng Shi (Wisconsin School of Business)
Reinsurance Example

- Consider two types of reinsurance contract
  - Quota share reinsurance: share risk proportionally
  - Excess-of-loss reinsurance: share risk above threshold
# Reinsurance Example

- Risk capital for the hypothetical reinsurance portfolio
  - Value-at-Risk (VaR): $VaR(p) = Q_R(p)$
  - Conditional Tail Expectation (CTE): $CTE(p) = E[R | R > Q_R(p)]$

<table>
<thead>
<tr>
<th>Quota = 0.25</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>VaR</td>
<td>21,662,280</td>
<td>22,293,481</td>
<td>23,500,626</td>
<td>22,505,966</td>
<td>23,066,293</td>
<td>24,175,243</td>
</tr>
<tr>
<td>Quota = 0.5</td>
<td>14,441,520</td>
<td>14,862,321</td>
<td>15,667,084</td>
<td>15,003,977</td>
<td>15,377,529</td>
<td>16,116,829</td>
</tr>
<tr>
<td>Quota = 0.75</td>
<td>7,220,760</td>
<td>7,431,160</td>
<td>7,833,542</td>
<td>7,501,989</td>
<td>7,688,764</td>
<td>8,058,414</td>
</tr>
<tr>
<td>Retention = 1</td>
<td>27,085,174</td>
<td>27,971,135</td>
<td>29,646,489</td>
<td>28,230,301</td>
<td>28,989,703</td>
<td>30,506,736</td>
</tr>
<tr>
<td>Retention = 5</td>
<td>22,581,757</td>
<td>23,273,150</td>
<td>24,809,406</td>
<td>23,551,479</td>
<td>24,202,636</td>
<td>25,621,111</td>
</tr>
<tr>
<td>Retention = 10</td>
<td>15,205,267</td>
<td>15,922,672</td>
<td>17,681,267</td>
<td>16,526,250</td>
<td>17,525,540</td>
<td>21,394,198</td>
</tr>
</tbody>
</table>
Summary

- Several features of our approach
  - Both parametric and semi-parametric formulations
  - Copula model to associate business lines
  - A hierarchical structure to allow for learning across insurers
  - Predictions are allowed at different levels of interest

- Future research
  - To compare with classical multilevel modeling
  - To incorporate collateral information
    - Firm-level heterogeneity
  - Look at triangles by state
A copula is a multivariate distribution function with uniform marginals. Let $U_1, \ldots, U_T$ be $T$ uniform random variables on $(0,1)$. Their distribution function

$$C(u_1, \ldots, u_T) = \Pr(U_1 \leq u_1, \ldots, U_T \leq u_T)$$

For general applications, consider arbitrary marginal distributions $F_1(y_1), \ldots, F_T(y_T)$. Define a multivariate distribution function using the copula such that

$$F(y_1, \ldots, y_T) = C(F_1(y_1), \ldots, F_T(y_T))$$

Sklar(1959) established the converse: any multivariate distribution function $F$ can be written in the form of the above equation, i.e., using a copula representation.
Penalized Regression Spline

\[ g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + \beta_{n,l} \times t_j + \Gamma'_j \gamma_{n,l} \]

- Denote \( \gamma_{n,l} = (\gamma_{n,l,1}, \cdots, \gamma_{n,l,K})' \)

- Define \( \Gamma = \Gamma_K \Lambda_K^{-1/2} \) and \( \Gamma_j \) is the \( j \)th row of \( \Gamma \)

\[ \Gamma_K = \begin{pmatrix} |t_1 - \nu_1|^3 & |t_1 - \nu_2|^3 & \cdots & |t_1 - \nu_K|^3 \\ |t_2 - \nu_1|^3 & |t_2 - \nu_2|^3 & \cdots & |t_2 - \nu_K|^3 \\ \vdots & \vdots & \ddots & \vdots \\ |t_J - \nu_1|^3 & |t_J - \nu_2|^3 & \cdots & |t_J - \nu_K|^3 \end{pmatrix} \]

\[ \Lambda_K = \begin{pmatrix} 0 & |\nu_1 - \nu_2|^3 & \cdots & |\nu_1 - \nu_K|^3 \\ |\nu_2 - \nu_1|^3 & 0 & \cdots & |\nu_2 - \nu_K|^3 \\ \vdots & \vdots & \ddots & \vdots \\ |\nu_K - \nu_1|^3 & |\nu_K - \nu_2|^3 & \cdots & 0 \end{pmatrix} \]
Data Exploration

- Two scatter plots showing the relationship between Commercial Auto and Personal Auto for different companies.
- Each company is represented by a different color.

Company #1, Company #2, Company #3, Company #4, Company #5.

- Additional scatter plots showing the distribution of Commercial Auto and Personal Auto for each company.
Model Specification

- Some examples on priors
  - $\delta_{n,l} \sim N(0, \sigma^2_{\delta}[l])$ for $n = 1, \ldots, N$
  - $\sigma^2_{\delta}[l] \sim IG(10^{-4}, 10^{-4})$ for $l = 1, \ldots, L$
  - $\sigma^2_{n,l} \sim IG(\psi[l], \psi[l])$ for $n = 1, \ldots, N$
    - $\psi[l] \sim Gamma(10^{-4}, 10^{-4})$ for $l = 1, \ldots, L$
  - $\theta_n \sim Uniform(-100, 100)$ for $n = 1, \ldots, N$
Convergence
Copula Validation

- Use $K$-plot to validate the Frank copula
- Based on function $K(w) = w - \frac{\phi(w)}{\phi'(w)}$
- Visualize parametric and non-parametric estimates
  - Parametric: $\phi(w) = \ln\left(\frac{e^{\rho w} - 1}{(e^\rho - 1)}\right)$ for Frank
  - Non-parametric: use pseudo-observations $W_s = \hat{H}(U_s, V_s)$