Estimating Between Line Correlations Generated by Parameter Uncertainty

Glenn Meyers, FCAS, MAAA
Estimating Between Line Correlations
Generated by Parameter Uncertainty

by

Glenn Meyers
Insurance Services Office, Inc.

Abstract
When applying the collective risk model to an analysis of insurer capital needs, it is crucial to consider the effect of correlation between lines of insurance. Recent work sponsored by the Committee on the Theory of Risk has sparked the development of methods that include correlation in the collective risk model. One of these methods is built around the view that correlation is generated by parameter uncertainty affecting several lines of insurance simultaneously.

This paper uses simulation analyses to explore the properties of both classical and Bayesian methods of quantifying parameter uncertainty. We conclude that in order to get sufficient accuracy to determine the necessary capital, one must use the combined data of several insurers. Using the combined data of several insurers forces us to consider a collective risk model where parameter uncertainty affects several insurers – as well as several lines of insurance – simultaneously.
1. Introduction

The collective risk model has long been one of the primary tools of actuarial science. One can view this model as a computer simulation where one first picks a random number of claims and then sums the random loss amounts for each claim.

The early uses of the collective risk model were mostly theoretical illustrations of the role of insurer surplus and profit margins. Such illustrations are still common today in insurance educational readings such as Bowers, Gerber, Jones, Hickman and Nesbitt [1997, Ch 13].

By the late 1970’s, members of the Casualty Actuarial Society were beginning to use the collective risk model as input for real-life insurance decisions. The early applications of the collective risk model included retrospective rating, e.g. Meyers [1980], and aggregate stop loss reinsurance, e.g. John and Patrik [1980] which is also described by Patrik [1996]. Bear and Nemlick [1990] provide further examples of the use of the collective risk model in the pricing of reinsurance contracts. Meyers [1989] begins to apply the collective risk model to an analysis of insurer capital.

This paper is part of a collective effort to extend the use of the collective risk model to Dynamic Financial Analysis (DFA). One goal of DFA is the management of an insurer’s capital. An insurer requires sufficient capital so that its chance of insolvency is reasonably remote. An insurer can manage its capital needs by structuring its business so that it has an acceptably remote chance of a large loss. This structuring can include the use of reinsurance.

While the collective risk model arose from theoretical exercises in insurer solvency, it has not been widely used in practice for setting solvency standards. The main reason for this has been that it requires that individual lines of insurance be independent. Almost nobody believes this to be true. And as we shall demonstrate below, assuming independence can lead to a significantly understated solvency standard.
Recognizing this problem, the CAS Committee on the Theory of Risk commissioned Dr. Shaun Wang to develop versions of the collective risk model that do not require one to assume independence between lines of insurance. This work led to a paper titled "Aggregation of Correlated Risk Portfolios: Models & Algorithms" which is to appear in the next volume of the *Proceedings of the Casualty Actuarial Society*.

Inspired by Dr. Wang’s work, we followed with a discussion of his paper, Meyers [1999], that focused on a version of the collective risk model where the claim count distribution for each line of insurance was conditionally independent given a parameter \( \alpha \). Treating \( \alpha \) as a random variable leads to a particular kind of dependence between lines of insurance.

In this paper we propose a methodology for estimating the variance of \( \alpha \) and explore the data requirements necessary to provide reliable estimates of this variance.

### 2. The Collective Risk Model

For the \( h \)-th line of insurance let:

\[
\begin{align*}
\mu_h & = \text{Expected claim severity;} \\
\sigma_h^2 & = \text{Variance of the claim severity distribution;} \\
\lambda_h & = \text{Expected claim count;} \text{ and} \\
\lambda_h + c_h \lambda_h^2 & = \text{Variance of the claim count distribution.}
\end{align*}
\]

Following Heckman and Meyers [1983], we call \( c_h \) the contagion parameter. If the claim count distribution is:

- Poisson, then \( c_h = 0 \);
- negative binomial, then \( c_h > 0 \); and
- binomial with \( n \) trials, then \( c_h = -1/n \).
A good way to view the collective risk model is by a Monte-Carlo simulation.

**Simulation Algorithm #1**

**The Collective Risk Model Assuming Independence Between Lines of Insurance**

1. For lines of insurance 1 to n, select a random number of claims, $K_h$, for each line of insurance $h$.
2. For each line of insurance $h$, select random claim amounts $Z_{hk}$, for $k = 1, \ldots, K_h$. Each $Z_{hk}$ has a common distribution $\{Z_k\}$.
3. Set $X_h = \sum_{k=1}^{K_h} Z_{hk}$.
4. Set $X = \sum_{h=1}^{n} X_h$.

The collective risk model describes the distribution of $X$.

Mayers [1999] shows that if $K_h$ is independent of $K_d$ for $d \neq h$, and $Z_h$ is independent of $K_h$ we have:

\[ \text{Var}[X_h] = \lambda_h \mu_h^2 + \mu_h \left( \lambda_h - \mu_h \right)^2; \quad (2.1) \]

and

\[ \text{Cov}[X_d, X_h] = 0 \text{ for } d \neq h. \quad (2.2) \]

We now introduce parameter uncertainty that affects the claim count distribution that affects several lines of insurance simultaneously. We partition the lines of insurance into covariance groups $\{G_i\}$. Our next version of the collective risk model is defined as follows.
Simulation Algorithm #2
The Collective Risk Model with Parameter Uncertainty
in the Claim Count Distributions

1. For each covariance group \( i \), select \( \alpha_i > 0 \) from a distribution with:
   \[
   E[\alpha_i] = 1 \quad \text{and} \quad \text{Var}[\alpha_i] = \gamma_i.
   \]
   \( \gamma_i \) is called the covariance generator for the covariance group \( i \).

2. For line of insurance \( h \) in covariance group \( i \), select a random number of claims \( K_{ih} \)
   from a distribution with mean \( \alpha_i \lambda_{bh} \).

3. For each line of insurance \( h \) in covariance group \( i \), select random claim amounts \( Z_{ih} \)
   for \( k = 1, \ldots, K_{ih} \). Each \( Z_{ih} \) has a common distribution \( \{Z_{ih}\} \).

4. Set \( X_{ih} = \sum_{k=1}^{K_{ih}} Z_{ih} \).

5. Set \( X_{ih} = \sum_{h \in G_i} X_{ih} \).

6. Set \( X = \sum_{i=1}^{I} X_{ih} \).

Meyers [1999] shows that for \( d \neq h \):
\[
\text{Cov}[X_{ih}, X_{ih}] = \gamma_i \lambda_{bh} \cdot \mu_{ih} \cdot \lambda_{bh} \cdot \mu_{bh}.
\] (2.3)

For \( d = h \):
\[
\text{Cov}[X_{ih}, X_{ih}] = \text{Var}[X_{ih}] = \lambda_{bh} \cdot \sigma_{bh}^2 + \mu_{ih} \cdot (\lambda_{bh} + (1 + \gamma_i) \cdot \mu_{bh} \cdot \lambda_{bh}) + \gamma_i \cdot \lambda_{bh} \cdot \mu_{bh}^2.
\] (2.4)

And for \( i \neq j \):
\[
\text{Cov}[X_{ih}, X_{ih}] = 0.
\] (2.5)
The ultimate purpose of this paper is to discuss the estimation of the $g_i$'s from claim count data, so we remove claim severity from the above equations by setting each $\mu_n = 1$ and $\sigma^2_n = 0$. This gives us:

$$\text{Cov}[K_n, K_n] = g_i \cdot \lambda_n.$$  \hfill (2.6)

and for $d - h$:

$$\text{Cov}[K_n, K_n] = \text{Var}[K_n] = \lambda_n + (c_n + g_i + c_n \cdot g_i) \cdot \lambda_n^2.$$  \hfill (2.7)

and for $i \neq j$:

$$\text{Cov}[K_i, K_j] = 0.$$  \hfill (2.8)

3. The Impact of the Covariance Generator on Required Capital

The purpose of this paper is to give some estimators of the covariance generator, $g$. To this end, we give an example on a hypothetical insurer writing four lines of insurance. The insurer expects 1,000 claims in each line, and the contagion parameter for each line is equal to 0.02. The covariance generator is equal to 0.04. The claim severity distributions are given in Meyers [1999]. Tables 3.1 and 3.2 give various summary statistics of the insurer's aggregate loss distribution.

**Table 3.1**

<table>
<thead>
<tr>
<th>Aggregate Summary Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Mean</td>
</tr>
<tr>
<td>Aggregate Std. Dev.</td>
</tr>
</tbody>
</table>

**Table 3.2**

<table>
<thead>
<tr>
<th>Claim Severity and Claim Count Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>GL-$1$M</td>
</tr>
<tr>
<td>GL-$5$M</td>
</tr>
<tr>
<td>AL-$1$M</td>
</tr>
<tr>
<td>AL-$5$M</td>
</tr>
</tbody>
</table>
Table 3.3 and 3.4 give the correlations between each of the lines of insurance for the claim counts, and for the total losses.

### Table 3.3
Claim Count Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>GL-$1M</th>
<th>GL-$5M</th>
<th>AL-$1M</th>
<th>AL-$5M</th>
</tr>
</thead>
<tbody>
<tr>
<td>GL-$1M</td>
<td>1.000</td>
<td>0.647</td>
<td>0.647</td>
<td>0.647</td>
</tr>
<tr>
<td>GL-$5M</td>
<td>0.647</td>
<td>1.000</td>
<td>0.647</td>
<td>0.647</td>
</tr>
<tr>
<td>AL-$1M</td>
<td>0.647</td>
<td>0.647</td>
<td>1.000</td>
<td>0.647</td>
</tr>
<tr>
<td>AL-$5M</td>
<td>0.647</td>
<td>0.647</td>
<td>0.647</td>
<td>1.000</td>
</tr>
</tbody>
</table>

### Table 3.4
Total Loss Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>GL-$1M</th>
<th>GL-$5M</th>
<th>AL-$1M</th>
<th>AL-$5M</th>
</tr>
</thead>
<tbody>
<tr>
<td>GL-$1M</td>
<td>1.000</td>
<td>0.531</td>
<td>0.453</td>
<td>0.423</td>
</tr>
<tr>
<td>GL-$5M</td>
<td>0.531</td>
<td>1.000</td>
<td>0.440</td>
<td>0.410</td>
</tr>
<tr>
<td>AL-$1M</td>
<td>0.453</td>
<td>0.440</td>
<td>1.000</td>
<td>0.351</td>
</tr>
<tr>
<td>AL-$5M</td>
<td>0.423</td>
<td>0.410</td>
<td>0.351</td>
<td>1.000</td>
</tr>
</tbody>
</table>

We now consider some capital requirement formulas. Let $X$ be a random variable representing the insurer's aggregate loss. Let:

- $F(x) = \Pr(X \leq x)$
- $f(x) = F'(x)$
- $\sigma = $ Standard Deviation of $X$
- $C = $ Required Insurer Capital

Then the required capital can be defined by one of the following equations

1. **Probability of Ruin Formula:** 
   $$ \Pr(C + E[X]) = 1 - \epsilon. $$

2. **Expected Policyholder Deficit Formula:** 
   $$ \int_{C - E[X]}^{\infty} (x - C - E[X]) \cdot f(x) \, dx = \eta. $$

3. **Standard Deviation Formula:** 
   $$ C = T \cdot \sigma. $$
The probability of ruin is a common textbook capital requirement formula in actuarial mathematics. The standard deviation formula is the probability of ruin formula, when applied to a normal approximation of the insurer’s aggregate loss distribution. The expected policyholder deficit formula is more recent, and takes into account the amount of insolvency as well as the probability of insolvency.

We calculated the distribution of $X$ using the Heckman/Meyers algorithm [1983] as modified by Meyers [1999]. We then calculated the capital requirements using the above formulas (with $\varepsilon = 0.01$, $\eta = 0.001$ and $T = 2.32$) for the insurer using various values of $g$. The results are in Tables 3.5 and 3.6.

### Table 3.5
The Effect of $g$ on Capital Requirements

<table>
<thead>
<tr>
<th>$g$</th>
<th>Standard Deviation of Ruin</th>
<th>Probability of Ruin</th>
<th>Expected Policyholder Deficit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>42,388,424</td>
<td>43,179,285</td>
<td>46,210,851</td>
</tr>
<tr>
<td>0.03</td>
<td>48,535,720</td>
<td>52,492,867</td>
<td>49,606,674</td>
</tr>
<tr>
<td>0.04</td>
<td>53,987,534</td>
<td>57,818,856</td>
<td>55,052,911</td>
</tr>
<tr>
<td>0.05</td>
<td>58,937,183</td>
<td>62,516,435</td>
<td>59,858,191</td>
</tr>
<tr>
<td>0.06</td>
<td>63,502,198</td>
<td>66,763,256</td>
<td>64,205,165</td>
</tr>
</tbody>
</table>

### Table 3.6
The Effect of $g$ on Capital Requirements
% Deviations from the Base $g = 0.04$

<table>
<thead>
<tr>
<th>$g$</th>
<th>Standard Deviation of Ruin</th>
<th>Probability of Ruin</th>
<th>Expected Policyholder Deficit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>-21.5%</td>
<td>-25.3%</td>
<td>-16.1%</td>
</tr>
<tr>
<td>0.03</td>
<td>-10.1%</td>
<td>-9.2%</td>
<td>-9.9%</td>
</tr>
<tr>
<td>0.04</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.05</td>
<td>9.2%</td>
<td>8.1%</td>
<td>8.7%</td>
</tr>
<tr>
<td>0.06</td>
<td>17.6%</td>
<td>15.5%</td>
<td>16.6%</td>
</tr>
</tbody>
</table>

The above tables show that the value of $g$ can have a significant effect on the required surplus.

205
4. The Likelihood Function for a Multivariate Claim Count Distribution

From this point forward, we shall assume there is only one covariance group and drop the subscripts \(i\) and \(j\) in Simulation Algorithm #2.

As we estimate the \(g\) parameter across different lines in a covariance group, we will be estimating the parameters, \(\lambda_h\) and \(c_h\), of each claim count distribution simultaneously. In effect, we will be estimating the parameters of a multivariate distribution on the random vector \(\tilde{K} = \{K_h\}\).

At this point, it is helpful to adopt the vector notation \(\tilde{c} = \{c_h\}\) and \(\tilde{\lambda} = \{\lambda_h\}\).

The negative binomial claim count distribution, conditional on \(\alpha\), will be obtained from the standard negative binomial distribution by multiplying its mean, \(\lambda_h\), by \(\alpha\).

Following Meyers [1999], we shall use the following form of the negative binomial distribution for the probability of \(K_h = k_h\) conditional on \(\alpha\).

\[
Pr\{K_h = k_h | \alpha\} = \frac{\Gamma(1/c_h + k_h)}{\Gamma(1/c_h) \cdot \Gamma(k_h + 1)} \cdot \frac{(c_h \alpha \lambda_h)^k_h}{(1 + c_h \alpha \lambda_h)^{k_h + 1}}
\]  

(4.1)

Given \(g \geq 0\), define \(^1^\):

\[
\alpha_i = 1 - \sqrt{3g}, \quad \alpha_j = 1, \quad \text{and} \quad \alpha_s = 1 + \sqrt{3g},
\]

and

\[
Pr\{\alpha = \alpha_i\} = 1/6, \quad Pr\{\alpha = \alpha_j\} = 2/3, \quad \text{and} \quad Pr\{\alpha = \alpha_s\} = 1/6.
\]

(4.2)

One can easily verify that \(E[\alpha] = 1\) and \(\text{Var}(\alpha) = g\).

The conditional likelihood of a claim count vector \(\tilde{K}\{|\alpha\} = \{K_h|\alpha\}\) is given by:

\[
\ell(\tilde{K}; \tilde{\lambda}, \tilde{c}|\alpha) = \prod_h Pr\{K_h = k_h | \alpha\}.
\]

(4.3)

\(^1^\) As pointed out in Meyers [1999], this discrete distribution for \(\alpha\) was motivated by the Gauss-Hermite numerical integration formula. One can easily derive similar distributions with more points.
The unconditional likelihood of a claim count vector \( k = \{ k_n \} \) is given by:

\[
\ell(\tilde{\lambda}, \tilde{\epsilon}, \tilde{\gamma}) = \frac{\ell(\tilde{\lambda}, \tilde{\epsilon}, \tilde{\gamma} | x)}{6} + \frac{2 \cdot \ell(\tilde{\lambda}, \tilde{\epsilon}, \tilde{\gamma} | x)}{3} + \frac{\ell(\tilde{\lambda}, \tilde{\epsilon}, \tilde{\gamma} | x)}{6}
\]  

(4.4)

As we go about the computational efforts described below, we will work with the log-likelihood functions:

\[
\ell(\tilde{\lambda}, \tilde{\epsilon} | x) = \ln\left(\frac{\ell(\tilde{\lambda}, \tilde{\epsilon} | x)}{6}\right) \; \text{and} \; \\
\ell(\tilde{\lambda}, \tilde{\gamma} | x) = \ln\left(\frac{\ell(\tilde{\lambda}, \tilde{\gamma} | x)}{6} + \frac{2 \cdot \ell(\tilde{\lambda}, \tilde{\gamma} | x)}{3} + \frac{\ell(\tilde{\lambda}, \tilde{\gamma} | x)}{6}\right)
\]  

(4.5)

(4.6)

5. Maximum Likelihood Estimation

Under the assumption that claims are generated by the process described in Simulation Algorithm #2, an insurer wishing to estimate the parameters \( \tilde{\lambda} \), \( \tilde{\epsilon} \) and \( \tilde{\gamma} \) might gather data like that in the following table from its own claims experience.

**Table 5.1**

**Insurance Data for Estimating \( \tilde{\epsilon} \) and \( \tilde{\gamma} \)**

<table>
<thead>
<tr>
<th>Year</th>
<th>Line 1</th>
<th>Line 2</th>
<th>Line 3</th>
<th>Line 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>100</td>
<td>80</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>1997</td>
<td>100</td>
<td>80</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>1996</td>
<td>100</td>
<td>80</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>1995</td>
<td>100</td>
<td>80</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>1994</td>
<td>100</td>
<td>80</td>
<td>40</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Line 1</th>
<th>Line 2</th>
<th>Line 3</th>
<th>Line 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>153</td>
<td>131</td>
<td>53</td>
<td>31</td>
</tr>
<tr>
<td>1997</td>
<td>96</td>
<td>77</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>1996</td>
<td>53</td>
<td>89</td>
<td>45</td>
<td>16</td>
</tr>
<tr>
<td>1995</td>
<td>92</td>
<td>72</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>1994</td>
<td>92</td>
<td>90</td>
<td>43</td>
<td>16</td>
</tr>
</tbody>
</table>

Estimated Frequency:

|          | 0.9720 | 1.1475 | 1.1350 | 1.1300 |
We estimated the insurer's frequency by line of insurance by dividing the total claim count by the total exposure. We then assumed that $c_h = c$ for all $h$.

Let $\tilde{k}_y$ and $\tilde{\lambda}_y$ be respectively, an observed claim count vector and an estimated expected claim count vector for the year $y$.

In Table 5.1 the observed claim count vector, $\tilde{k}_{1998}$, is equal to $(153, 131, 53, 31)^T$. The expected claim count vector, $\tilde{\lambda}_{1998}$, is equal to $(100.09720, 80.11475, 40.11350, 20.11300)^T$ which is equal to $(97.2, 81.8, 45.4, 22.6)^T$. The parameter vector, $\tilde{c}$, is equal to $(c, c, c, c)^T$. The maximum likelihood estimates $\hat{c}$ and $\hat{g}$ of $c$ and $g$ are the values of $c$ and $g$ that maximizes:

$$\sum_i L (\tilde{k}_i, \tilde{\lambda}_i, \hat{c}, \hat{g})$$

(5.1)

Using Excel Solver, we found the maximum likelihood estimate (MLE) $\hat{c}$ of $c$ to be 0.0134 and the maximum likelihood estimate $\hat{g}$ of $g$ to be 0.0226.

We should note that the data in Table 5.1 was not generated from actual insurer data. It was taken from five random drawings from Simulation Algorithm #2 with the “true” frequencies set equal to 1.0000 for each line of insurance, the “true” value of $c$ set equal to 0.0200, and the “true” value of $g$ set equal to 0.0400. We repeated the simulation 100 times with the following results.

| Table 5.2 Property of MLE’s for $c$ and $g$ Derived from 100 Simulations of a Single Insurer’s Data |
|-----------------|-------|-------|
| $c$ True Value  | 0.0200| 0.0400|
| Average MLE     | 0.0134| 0.0226|
| Std. Dev. of the MLE | 0.0126| 0.0208|

One can see from Tables 3.5 and 3.6 that the estimation errors can lead to a significant understating of the required surplus.
Based on this and other similar simulations we conclude that estimating $c$ and $g$ in this manner can lead to biased and highly volatile results.

We now examine some other estimation methods.

The first alternative is to combine the data of several "similar" insurers. Let $A$ be the set of insurers and let $a \in A$. We created 40 nearly identical "copies" of our insurer and simulated the MLE's for $c$ and $g$. Table 5.3 below shows the exposures and claim counts for the first two insurers in a typical simulation.

When combining the data of several insurers we maximize the log-likelihood expression:

$$
\sum_{i} L\left(\hat{\kappa}, \hat{\lambda}, \hat{c}, \hat{g}\right).
$$

(5.2)
### Table 5.3
Multi-Insurer Data for Estimating c and g

<table>
<thead>
<tr>
<th>Insurer #1</th>
<th>Exposure by Line and Year</th>
<th>Claim Count by Line and Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year</td>
<td>Line 1</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>1997</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>1996</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>1995</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>1994</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Insurer #2</th>
<th>Exposure by Line and Year</th>
<th>Claim Count by Line and Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year</td>
<td>Line 1</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1997</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1996</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1995</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1994</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Insurer #3</th>
<th>Exposure by Line and Year $^2$</th>
<th>Estimated Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year</td>
<td>Line 1</td>
</tr>
</tbody>
</table>

210
We ran 100 simulations of data like that in Table 5.3 and calculated the maximum likelihood estimators for \( c \) and \( g \) with the following results.

### Table 5.4
Properties of MLE's for \( c \) and \( g \)
Derived from 100 Simulations of 40 Insurers' Data

<table>
<thead>
<tr>
<th></th>
<th>( c )</th>
<th>( g )</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Value</td>
<td>0.0200</td>
<td>0.0400</td>
</tr>
<tr>
<td>Average MLE</td>
<td>0.0199</td>
<td>0.0399</td>
</tr>
<tr>
<td>Std. Dev. of the MLE</td>
<td>0.0022</td>
<td>0.0030</td>
</tr>
</tbody>
</table>

Based on this and other similar simulations we conclude that we can obtain accurate estimates of \( c \) and \( g \) — if we can get the combined results of several "similar" insurers.

The existence (or non-existence) of similar insurers opens up a host of issues. We now explore a few of these issues.

### 6. Bayesian Estimation

We suspect few insurers would agree that they are sufficiently "similar" to any other group of insurers to fully accept the results of an analysis like that given above. They might accept the results because they have no quantitative alternative, and then judgmentally modify the results. Since we consider it likely that judgment will enter the picture, we consider a Bayesian approach to the problem.

Consider a grid \((c_{i}, g_{i})\) of possible values of \( c \) and \( g \). Let \( \{k_{i}\} \) be a set of observations needed to calculate the likelihood function for each point \((c_{i}, g_{i})\). Let \( p_{i} \) be the "prior" probability of each point \((c_{i}, g_{i})\).

---

2 We varied the exposure for the lines in the patterns: 100, 80, 40, 20; 20, 100, 80, 40; 40, 20, 100, 80, and 80, 40, 20, 100

3 The reader may observe that the expected claim counts for the insurer in this simulated sample were significantly smaller than the insurer discussed in Section 3 above. We also did a simulation where the insurers were 10 times as large. We obtained Std Dev[\(c\)] = 0.0011 and Std Dev[\(g\)] = 0.0022.
Then according to Bayes' Theorem, the posterior likelihood of each \((c, g)\) will be proportional to

\[
\prod_y \ell(\hat{k}_y, \bar{x}_y, c, g) \cdot p_y. \tag{6.1}
\]

As an illustration, suppose that we choose a prior so that the \(p_y\)'s are equally likely. For one simulated \([k_y]\) based on a single insurer's exposure we obtained the following posterior distribution of \((c, g)\), which we show (part of) graphically.

---

\(\text{Graph 6.1}

Posterior Likelihood for a Single Insurer
with a Uniform Prior Distribution

---

\(^4\) For the time being we are assuming that the expected claim count is known. We will address this problem below.
As an example, we construct a prior distribution so that

\[ p_\gamma \propto \prod_{i=1}^{n} \ell(\bar{K}_i^*, \bar{\lambda}, c_i, g_i) \],

where \( \{\bar{K}_i^*\} \) comes from the (simulated here, but in practice real) data of the 40 "peer group" insurers given above. We obtained the following posterior distribution for the same insurer that we show graphically.

Below, we will show how to use the posterior distribution as input into the collective risk model, as described in Simulation Algorithm #2.

**Graph 6.2**

*Posterior Likelihood for a Single Insurer with a Prior Distribution Based on Industry Data*
7. Industry Drivers of Correlation

The likelihood Equation 3.6 was derived under the assumption that the “driver” of the correlation, i.e., the random variable $\alpha$, was independent for each individual insurer. This section considers the consequences of the random variable $\alpha$ being common to all insurers. To this end, we replace Steps 1 and 2 of Simulation Algorithm #2 with the more complicated process.

Simulation Algorithm #3

The Collective Risk Model with Parameter Uncertainty in the Claim Count Distributions Driven by Industry and Insurer Parameter Uncertainty

1. For each covariance group $i$, select $\alpha_i^A$ and $\alpha_i^I$ as follows.
   1.1. Select $\alpha_i^A$ from a distribution with $E(\alpha_i^A) = 1$ and $\text{Var}(\alpha_i^A) = g_i^A$. $g_i^A$ is called the industry covariance generator for covariance group $i$.
   1.2. Select $\alpha_i^I$ from a distribution with $E(\alpha_i^I) = 1$ and $\text{Var}(\alpha_i^I) = g_i^I$. $g_i^I$ is called the insurer covariance generator for covariance group $i$.

2. For line of insurance $h$ in covariance group $i$, select a random number of claims $K_{hi}$ from a distribution with mean $\alpha_i^A \cdot \alpha_i^I \cdot K_{hi}$.

3. For each line of insurance $h$ in covariance group $i$, select random claim amounts $Z_{hki}$ for $k = 1, \ldots, K_{hi}$. Each $Z_{hki}$ has a common distribution $\{Z_{hi}\}$.

4. Set $X_{hi} = \sum_{k=1}^{K_{hi}} Z_{hki}$.

5. Set $X_{si} = \sum_{h \in \text{set }i} X_{hi}$.

6. Set $X = \sum_{i=1}^{k} X_{si}$.
We now calculate the moments of the aggregate loss distribution described by Simulation Algorithm #3

$$E[\alpha^\lambda \cdot \alpha_1] = E_{\alpha^\lambda}[\alpha^\lambda \cdot \alpha_1] = E_{\alpha^\lambda}[\alpha^\lambda] - 1 \quad (7.1)$$

$$\text{Var}[\alpha^\lambda \cdot \alpha_1] = E_{\alpha^\lambda}[\text{Var}[\alpha^\lambda \cdot \alpha_1] + \text{Var}[E[\alpha^\lambda \cdot \alpha_1]]]$$

$$= E_{\alpha^\lambda}[\alpha^\lambda \cdot \text{Var}[\alpha_1]] + \text{Var}[E_{\alpha^\lambda}[\alpha^\lambda]]$$

$$= (1 - g_1) \cdot g_1 + g_1$$

$$= g_1 + g_1^2 + g_2 \cdot g_1 \quad (7.2)$$

To calculate the variances and covariances analogous to Simulation Algorithm #2, we simply replace the variance \( g_1 \) in Equations 2.3, 2.4, 2.6 and 2.7 with the expression \( g_1 + g_1^2 + g_2 \cdot g_1 \).

Let \( \hat{k}_i^\lambda \) be a vector of observed claim counts for the "industry" in year \( y \). An example of such a vector based on Table 5.3 is \( \hat{k}_i^\lambda = (69, 69, 53, 20, 25, 108, 64, 45, \ldots) \).

Similarly let \( \hat{x}_i^\lambda \) be a vector of expected claim counts for the "industry" in year \( y \).

The likelihood function of \( \hat{k}_i^\lambda \) conditional on \( \alpha^\lambda \) is given by:

$$\ell(\hat{k}_i^\lambda, \hat{x}_i^\lambda, \bar{c}, g | \alpha^\lambda) = \prod_i \ell(\hat{k}_i^\lambda, \hat{x}_i^\lambda, \bar{c}, g | \alpha^\lambda) \quad (7.3)$$

The associated log-likelihood function is given by:

$$L(\hat{k}_i^\lambda, \hat{x}_i^\lambda, \bar{c}, g | \alpha^\lambda) = \sum \ell(\hat{k}_i^\lambda, \hat{x}_i^\lambda, \bar{c}, g | \alpha^\lambda) \quad (7.4)$$
Given \( g^A \geq 0 \) define

\[
\alpha_i^A = 1 - \sqrt{3g^A}, \quad \alpha_j^A = 1, \quad \text{and} \quad \alpha_k^A = 1 + \sqrt{3g^A},
\]

and

\[
\Pr\{\alpha_i^A = \alpha_j^A\} = 1/6, \quad \Pr\{\alpha_i^A = \alpha_k^A\} = 2/3, \quad \text{and} \quad \Pr\{\alpha_i^A = \alpha_k^A\} = 1/6.
\] (7.5)

The unconditional log-likelihood function is then given by:

\[
L(\tilde{K}^A, \tilde{X}^A, \tilde{c}, g, g^A) = \ln\left(\frac{e^{\frac{1}{6}(\tilde{K}^A, \tilde{X}^A, \tilde{c}, g, g^A)}}{6} + \frac{2e^{\frac{1}{3}(\tilde{K}^A, \tilde{X}^A, \tilde{c}, g, g^A)}}{6} + \frac{e^{\frac{1}{6}(\tilde{K}^A, \tilde{X}^A, \tilde{c}, g, g^A)}}{6}\right)\] (7.6)

8. Maximum Likelihood Estimation Revisited

Consider the following two situations.

1. \( g = r > 0 \) and \( g^A = 0 \).
2. \( g = 0 \) and \( g^A = r > 0 \).

From the insurer's point of view, the two situations are identical. Its expected claim counts are multiplied by a random number each year.

But from the point of view of one who is trying to estimate the variance of the random multiplier, the situations are different. In the first situation, a new \( \alpha \) is picked for each insurer for each year. In the second situation, \( \alpha^A \) is picked once each year for all insurers.

The estimator should use the log-likelihood function in Equation 4.6. In the second situation the estimator should use the log-likelihood function in Equation 7.6.

We did 100 simulations of our 40 insurers where the claim counts are generated by Simulation Algorithm #3, with \( c = 0.02 \), \( g = 0 \) and \( g^A = 0.04 \). We then estimated \( c \) and "\( g \)" using maximum likelihood on Equation 4.6, with the following results.
We next did 100 simulations of our 40 insurers where the claim counts are generated by Simulation Algorithm #3, with $c = 0.02$, $g = 0.01$ and $g^A = 0.03$. We then estimated $c$, $g$ and $g^A$ using maximum likelihood on the “correct” Equation 7.6, with the following results.

If you used the estimated $g$ and $g^A$ in equation 7.2 instead of the true value of $g$ and $g^A$, you could significantly understate your capital requirements.

It may occur to one that the reason for this downward bias is due to the fact that we use estimated frequencies, rather than true frequencies. To test this we repeated the simulation using the “true” frequency rather than the estimated frequency and obtained the following results.

Table 8.1
Properties of MLE’s for $c$ and $g$
Derived from 100 Simulations of 40 Insurers’ Data with Industrywide Parameter Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$g$</th>
<th>$g^A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Value</td>
<td>0.0200</td>
<td>0.0000</td>
<td>0.0400</td>
</tr>
<tr>
<td>Average MLE</td>
<td>0.0218</td>
<td>0.0249</td>
<td>—</td>
</tr>
<tr>
<td>Std. Dev. of the MLE</td>
<td>0.0039</td>
<td>0.0158</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 8.2
Properties of MLE’s for $c$, $g$ and $g^A$
Using Estimated Frequencies
Derived from 100 Simulations of 40 Insurers’ Data with Industrywide Parameter Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$g$</th>
<th>$g^A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Value</td>
<td>0.0200</td>
<td>0.0100</td>
<td>0.0300</td>
</tr>
<tr>
<td>Average MLE</td>
<td>0.0201</td>
<td>0.0114</td>
<td>0.0213</td>
</tr>
<tr>
<td>Std. Dev. of the MLE</td>
<td>0.0023</td>
<td>0.0026</td>
<td>0.0090</td>
</tr>
</tbody>
</table>

Table 8.3
Properties of MLE’s for $c$, $g$ and $g^A$
Using “True” Frequencies
Derived from 100 Simulations of 40 Insurers’ Data with Industrywide Parameter Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$g$</th>
<th>$g^A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Value</td>
<td>0.0200</td>
<td>0.0100</td>
<td>0.0300</td>
</tr>
<tr>
<td>Average MLE</td>
<td>0.0200</td>
<td>0.0104</td>
<td>0.0298</td>
</tr>
<tr>
<td>Std. Dev. of the MLE</td>
<td>0.0023</td>
<td>0.0029</td>
<td>0.0033</td>
</tr>
</tbody>
</table>
This simulation indicates that the bias is indeed caused by using estimated frequencies in the MLE. However, in practice the "true" mean is not known.

9. Bayesian Estimation Revisited

Consider a grid \((\hat{\lambda}^{\lambda}, c, g, g^\lambda)\) of possible values of \(\hat{\lambda}^{\lambda}, c, g\) and \(g^\lambda\). Let \(\{k_i\}\) be a set of observations needed to calculate the likelihood function for each point \((\hat{\lambda}^{\lambda}, c, g, g^\lambda)\).

Let \(\pi_i\) be the "prior" probability of each point \((\hat{\lambda}^{\lambda}, c, g, g^\lambda)\).

Then according to Bayes' Theorem, the posterior likelihood of each \((\hat{\lambda}^{\lambda}, c, g, g^\lambda)\) is proportional to:

\[
\prod_y \ell(k_i, \hat{\lambda}^{\lambda}, c, g, g^\lambda) \cdot \pi_i \tag{9.1}
\]

Let \(\hat{e}_i^\lambda\) be a vector of exposures for the set of insurers, \(A\), in year \(y\). Let \(\hat{f}_i^\lambda\) be vector of claim frequencies. Then each coordinate of the expected claim count vector \(\hat{\lambda}^{\lambda}\) is equal to the product of the corresponding coordinates of \(\hat{e}^\lambda\) and \(\hat{f}_i^\lambda\). Since the exposures are known and the claim frequencies are unknown, we should put a prior distribution on the grid \((\hat{\lambda}^{\lambda}, c, g, g^\lambda)\).

Let \(\pi_i\) be the posterior probability of each point in the grid \((\hat{f}_i^\lambda, c, g, g^\lambda)\). Then one can obtain estimates of \(\hat{f}_i^\lambda, c, g,\) and \(g^\lambda\) by the following formulas

\[
\hat{f}_i^\lambda = \sum_i f_i^\lambda \cdot \pi_i
\]
\[
\hat{c} = \sum_i c_i \cdot \pi_i
\]
\[
\hat{g} = \sum_i g_i \cdot \pi_i
\]
\[
\hat{g}^\lambda = \sum_i g_i^\lambda \cdot \pi_i \tag{9.2}
\]
We then tested the variability of these estimators on our simulated set of 40 insurers. The grid was constructed by varying $\hat{f}^\lambda, c, g$, and $g^\lambda$ in the following manner.

1. Each component of $\hat{f}^\lambda$ was set equal to 0.9875. Each component of $\tilde{f}^\lambda$ was set equal to 1.0125. The components for $i = 1, 2$ and 3 were equally spaced in between.

2. $c_0$ was set equal to 0.0100, $c_1$ was set equal to 0.0300. The components for $i = 1, 2$ and 3 were equally spaced in between.

3. $g_0$ was set equal to 0.0020, $g_1$ was set equal to 0.0180. The components for $i = 1, 2$ and 3 were equally spaced in between.

4. $g^\lambda_0$ was set equal to 0.0700, $g^\lambda_1$ was set equal to 0.0400. The components for $i = 1, 2$ and 3 were equally spaced in between.

In total, the grid had $5^4 = 625$ points. We assumed all points in the grid were equally likely.$^5$

We made 100 simulated estimates with the following results.

Table 8.4

| Properties of Bayesian Estimates for $c, g$ and $g^\lambda$ Using “True” Frequencies Derived from 100 Simulations of 40 Insurers’ Data with Industrywide Parameter Uncertainty |
|---------------------------------|-----------------|-----------------|-----------------|
| True Value                     | 0.0200          | 0.0100          | 0.0300          |
| Average Estimate               | 0.0201          | 0.0105          | 0.0303          |
| Std Dev of the Estimate        | 0.0021          | 0.0020          | 0.0027          |

Here we see that one can obtain stable and unbiased (in the classic statistical sense) by an appropriate use of Bayes’ Theorem.

---

$^5$ This “equally likely” is as subjective as any other assumption that one can make. The spacing of the grid is one part of the subjectivity. Another subjective assumption is that the frequencies for the four lines of insurance move together.
9. Using Real Data

This paper has taken a version of the collective risk model, in which the lines of insurance are correlated and explored some methods of estimating parameters of the claim count distributions. The data used in these methods consisted of both exposures and claim counts that span several years.

We explored the use of maximum likelihood on a single insurer's data to estimate the parameters and concluded that the random variation of the estimates were too large to derive a reliable estimate of the insurer's required surplus. One can obtain more stable estimates of the parameters by combining the data of several insurers.

We drew these conclusions from experiments performed on simulated "data."

We now raise some of the issues that one must address when estimating these parameters of the collective risk model with real data from several insurers.

1. Claim Count Development

When analyzing several years of claim count data, one must take care to distinguish the random variation from the systematic claim count development that occurs because of delays in reporting claims.

2. Insurer Class Differences

Different insurers can focus on different classes of business. When analyzing the data of several insurers, one must take care to distinguish the random variation from the systematic differences that occur because of the different classes of business that insurers write.

3. Insurer Strategy Changes

When analyzing the data of several insurers, one must take care to note that planned changes in insurer strategy that result in changes in claim counts. This can be difficult because insurers usually keep their strategy changes to themselves.

We are in the process of fitting this model to the data of several insurers. We are not yet in a position to say how we are addressing these and other issues. Suffice it to say that
we are using our judgment, and we anticipate that the ultimate users of this information will want to impose their own judgment. The Bayesian methodology provides a framework for making these judgments.

In spite of the judgments that one must make, we do feel that parameter estimates using the combined data of several insurers provides a useful starting point for insurers as they go about doing their Dynamic Financial Analysis.
References


222