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 α. Treating α 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 α 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:

μ_h = Expected claim severity;

σ_h^2 = Variance of the claim severity distribution;

λ_h = Expected claim count; and

$\lambda_h + c_h \cdot \lambda_h^2$ = Variance of the claim count distribution.

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, ... K_h. Each Z_{hk} has a common distribution $\{Z_h\}$.

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.

Meyers [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:

$$\mathrm{Var}[X_h] = \lambda_h \cdot \sigma_h^2 + \mu_h^2 \cdot (\lambda_h + c_h \cdot \lambda_h^2); \tag{2.1}$$

and

$$\mathrm{Cov}[X_d, X_h] = 0 \text{ for } d \neq h. \tag{2.2}$$

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:

$$E[\alpha_i] = 1 \text{ and } Var[\alpha_i] = g_i.$$

   g_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_{hi} from a distribution with mean $\alpha_i \cdot \lambda_{hi}$.

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

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

5. Set $X_{\bullet i} = \sum_{h \in G_i} X_{hi}$.

6. Set $X = \sum_{i=1}^{n} X_{\bullet i}$.

Meyers [1999] shows that for d ≠ h:

$$Cov[X_{di}, X_{hi}] = g_i \cdot \lambda_{di} \cdot \mu_{di} \cdot \lambda_{hi} \cdot \mu_{hi}. \tag{2.3}$$

For d = h:

$$Cov[X_{hi}, X_{di}] = Var[X_{hi}] = \lambda_{hi} \cdot \sigma_{hi}^2 + \mu_{hi}^2 \cdot (\lambda_{hi} + (1 + g_i) \cdot c_{hi} \cdot \lambda_{hi}^2) + g_i \cdot \lambda_{hi}^2 \cdot \mu_{hi}^2 \tag{2.4}$$

And for i ≠ j:

$$Cov[X_{di}, X_{hj}] = 0. \tag{2.5}$$

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 $\mu_h = 1$ and $\sigma_h^2 = 0$. This gives us:

$$\mathrm{Cov}\left[K_{di}, K_{hi}\right] = g_i \cdot \lambda_{di} \cdot \lambda_{hi}. \tag{2.6}$$

and for d = h:

$$\mathrm{Cov}\left[K_{hi}, K_{di}\right] = \mathrm{Var}\left[K_{hi}\right] = \lambda_{hi} + \left(c_{hi} + g_i + c_{hi} \cdot g_i\right) \cdot \lambda_{hi}^2. \tag{2.7}$$

and for i ≠ j:

$$\mathrm{Cov}\left[K_{di}, K_{hj}\right] = 0. \tag{2.8}$$

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
Aggregate Summary Statistics

Aggregate Mean	101,581,230
Aggregate Std. Dev.	23,270,489

Table 3.2
Claim Severity and Claim Count Statistics

Distribution Name	E[Count]	Std[Count]	E[Severity]	Std[Severity]	E[Total Loss]
GL-$1M	1000	248.60	36,966.16	124,853.59	36,966,160
GL-$5M	1000	248.60	40,348.87	160,218.51	40,348,870
AL-$1M	1000	248.60	11,456.65	76,434.03	11,456,650
AL-$5M	1000	248.60	12,809.55	99,730.27	12,809,550

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

	GL-$1M	GL-$5M	AL-$1M	AL-$5M
GL-$1M	1.000	0.647	0.647	0.647
GL-$5M	0.647	1.000	0.647	0.647
AL-$1M	0.647	0.647	1.000	0.647
AL-$5M	0.647	0.647	0.647	1.000

Table 3.4
Total Loss Correlation Matrix

	GL-$1M	GL-$5M	AL-$1M	AL-$5M
GL-$1M	1.000	0.531	0.453	0.423
GL-$5M	0.531	1.000	0.440	0.410
AL-$1M	0.453	0.440	1.000	0.351
AL-$5M	0.423	0.410	0.351	1.000

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

$$F(x) = \Pr\{X \le x\}$$
$$f(x) = F'(x)$$
$$\sigma = \text{Standard Deviation of X}$$
$$C = \text{Required Insurer Capital}$$

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

1. Probability of Ruin Formula: $F(C + E[X]) = 1 - \varepsilon$.

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

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

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

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

g	Standard Deviation	Probability of Ruin	Expected Policyholder Deficit
0.02	-21.5%	-25.3%	-16.1%
0.03	-10.1%	-9.2%	-9.9%
0.04	0.0%	0.0%	0.0%
0.05	9.2%	8.1%	8.7%
0.06	17.6%	15.5%	16.6%

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

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, λ_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 $\vec{\mathbf{K}} = \{K_h\}$

At this point, it is helpful to adopt the vector notation $\vec{c} = \{c_h\}$ and $\vec{\lambda} = \{\lambda_h\}$

The negative binomial claim count distribution, conditional on α, will be obtained from the standard negative binomial distribution by multiplying its mean, λ_h, by α.

Following Meyers [1999], we shall use the following form of the negative binomial distribution for the probability of k_h conditional on α.

$$\Pr\{K_h = k_h|\alpha\} = \frac{\Gamma(1/c_h + k_h)}{\Gamma(1/c_h)\cdot\Gamma(k_h+1)}\cdot\frac{(c_h\alpha\lambda_h)^{k_h}}{(1+c_h\alpha\lambda_h)^{1/c_h+k_h}} \tag{4.1}$$

Given $g \geq 0$, define[1]:

$$\alpha_1 = 1-\sqrt{3g},\ \alpha_2 = 1,\text{ and } \alpha_3 = 1+\sqrt{3g},$$

and (4.2)

$$\Pr\{\alpha=\alpha_1\} = 1/6,\ \Pr\{\alpha=\alpha_2\} = 2/3,\text{ and } \Pr\{\alpha=\alpha_3\} = 1/6.$$

One can easily verify that $E[\alpha] = 1$ and $Var[\alpha] = g$.

The conditional likelihood of a claim count vector $\vec{\mathbf{k}}|\alpha = \{k_h|\alpha\}$ is given by:

$$\ell(\vec{\mathbf{k}};\vec{\lambda},\vec{c}|\alpha) = \prod_h \Pr(K_h = k_h|\alpha). \tag{4.3}$$

[1] As pointed out in Meyers [1999], this discrete distribution for α 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 $\vec{\mathbf{k}} = \{k_h\}$ is given by:

$$f\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}},g\right) = \frac{f\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha_1\right)}{6} + \frac{2\cdot f\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha_2\right)}{3} + \frac{f\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha_3\right)}{6}. \tag{4.4}$$

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

$$L\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha\right) = \ln\left(f\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha\right)\right);\ \text{and} \tag{4.5}$$

$$L\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}},g\right) = \ln\left(\frac{e^{L\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha_1\right)}}{6} + \frac{2\cdot e^{L\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha_2\right)}}{3} + \frac{e^{L\left(\vec{\mathbf{k}};\vec{\lambda},\vec{\mathbf{c}}|\alpha_3\right)}}{6}\right). \tag{4.6}$$

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 $\vec{\lambda}$, $\vec{\mathbf{c}}$ and g might gather data like that in the following table from its own claims experience.

Table 5.1
Insurer Data for Estimating c and g

	Exposure by Line and Year			
Year	Line 1	Line 2	Line 3	Line 4
1998	100	80	40	20
1997	100	80	40	20
1996	100	80	40	20
1995	100	80	40	20
1994	100	80	40	20
	Claim Count by Line and Year			
	Line 1	Line 2	Line 3	Line 4
1998	153	131	53	31
1997	96	77	41	20
1996	53	89	45	16
1995	92	72	45	30
1994	92	90	43	16
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 $c_h \equiv c$ for all h.

Let $\vec{\mathbf{k}}_y$ and $\vec{\lambda}_y$ 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, $\vec{\mathbf{k}}_{1998}$, is equal to $(153, 131, 53, 31)^T$. The expected claim count vector, $\vec{\lambda}_{1998}$, is equal to $(100 \cdot 0.9720, 80 \cdot 1.1475, 40 \cdot 1.1350, 20 \cdot 1.1300)^T$ which is equal to $(97.2, 91.8, 45.4, 22.6)^T$. The parameter vector, $\vec{\mathbf{c}}$, 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:

$$\sum_y L\left(\vec{\mathbf{k}}_y \vec{\lambda}_y, \vec{\mathbf{c}}, g\right). \tag{5.1}$$

Using Excel Solver™, we found the maximum likelihood estimate (MLE), $\hat{c}$, of c to be 0.0169 and the maximum likelihood estimate $\hat{g}$ of g to be 0.0245.

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
Properties of MLE's for c and g
Derived from 100 Simulations
of a Single Insurer's Data

	c	g
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:

$$\sum_{a,y} L\left(\vec{\mathbf{k}}_y^a; \vec{\lambda}_y^a, \vec{\mathbf{c}}, g\right). \tag{5.2}$$

Table 5.3
Multi-Insurer Data for Estimating c and g

Insurer #1	**Exposure by Line and Year**			
Year	Line 1	Line 2	Line 3	Line 4
1998	100	80	40	20
1997	100	80	40	20
1996	100	80	40	20
1995	100	80	40	20
1994	100	80	40	20
	Claim Count by Line and Year			
	Line 1	Line 2	Line 3	Line 4
1998	69	69	53	20
1997	99	80	51	17
1996	101	78	68	18
1995	129	94	42	17
1994	82	76	30	15
Insurer #2	**Exposure by Line and Year**			
Year	Line 1	Line 2	Line 3	Line 4
1998	20	100	80	40
1997	20	100	80	40
1996	20	100	80	40
1995	20	100	80	40
1994	20	100	80	40
	Claim Count by Line and Year			
	Line 1	Line 2	Line 3	Line 4
1998	25	108	64	45
1997	18	88	75	42
1996	22	87	94	44
1995	22	130	69	47
1994	30	147	111	68
Insurer #3	**Exposure by Line and Year**[2]			
Year	Line 1	Line 2	Line 3	Line 4
⇓	⇓	⇓	⇓	⇓
Estimated Frequency	1.0088	1.0077	1.0088	0.9877

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

	c	g
True Value	0.0200	0.0400
Average MLE	0.0199	0.0399
Std. Dev. of the MLE	0.0022	0.0030

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[3].

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_j) of possible values of c and g. Let $\{\vec{\mathbf{k}}_y\}$ be a set of observations needed to calculate the likelihood function for each point (c_i, g_j). Let p_{ij} be the "prior" probability of each point (c_i, g_j).

---

[2] We varied the exposure for the lines in the pattern: 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 $\text{Std Dev}[\hat{c}] = 0.0011$ and $\text{Std Dev}[\hat{g}] = 0.0022$.

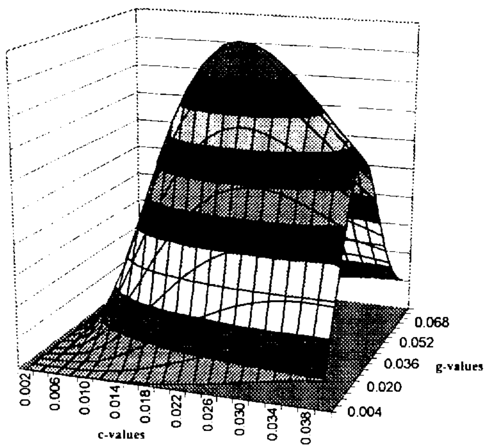
Then according to Bayes' Theorem, the posterior likelihood of each (c_i, g_j) will be proportional to[4]

$$\prod_{y} \ell\left(\vec{\mathbf{k}}_y ; \vec{\lambda}_y, c_i, g_j\right) \cdot p_{ij} . \tag{6.1}$$

As an illustration, suppose that we choose a prior so that the p_{ij}'s are equally likely. For one simulated $\{\vec{\mathbf{k}}_y\}$ based on a single insurer's exposure we obtained the following posterior distribution of (c_i, g_j), which we show (part of) graphically.

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

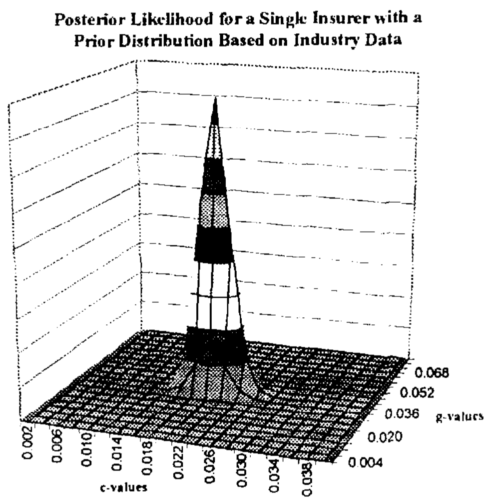
As an example, we construct a prior distribution so that

$$p_{ij} \propto \prod_{a,y} \ell\left(\vec{\mathbf{k}}_y^a ; \vec{\lambda}, c_i, g_j\right), \tag{6.2}$$

where $\{\vec{\mathbf{k}}_y^a\}$ 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



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 α, was independent for each individual insurer. This section considers the consequences of the random variable α 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 $\alpha_i^{\wedge}$ and α_i as follows.
   1.1. Select $\alpha_i^{\wedge}$ from a distribution with $E[\alpha_i^{\wedge}] = 1$ and $\mathrm{Var}[\alpha_i^{\wedge}] = g_i^{\wedge}$. $g_i^{\wedge}$ is called the industry covariance generator for covariance group i.
   1.2. Select α_i from a distribution with $E[\alpha_i] = 1$ and $\mathrm{Var}[\alpha_i] = g_i$. g_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 $\alpha_i^{\wedge} \cdot \alpha_i \cdot K_{hi}$.
3. For each line of insurance h in covariance group i, select random claim amounts Z_{hik} for k = 1, ... K_{hi}. Each Z_{hik} has a common distribution $\{Z_{hi}\}$.
4. Set $X_{hi} = \sum_{k=1}^{K_{hi}} Z_{hik}$.
5. Set $X_{\bullet i} = \sum_{h \in G_i} X_{hi}$.
6. Set $X = \sum_{i=1}^{n} X_{\bullet i}$.

We now calculate the moments of the aggregate loss distribution described by Simulation Algorithm #3.

$$E\left[\alpha_i^A \cdot \alpha_i\right] = E_{\alpha_i^A}\left[\alpha_i^A \cdot \alpha_i \middle| \alpha_i^A\right] = E_{\alpha_i^A}\left[\alpha_i^A\right] = 1 \tag{7.1}$$

$$\begin{aligned}
\operatorname{Var}\left[\alpha_i^A \cdot \alpha_i\right] &= E_{\alpha_i^A}\left[\operatorname{Var}\left[\alpha_i^A \cdot \alpha_i \middle| \alpha_i^A\right]\right] + \operatorname{Var}_{\alpha_i^A}\left[E\left[\alpha_i^A \cdot \alpha_i \middle| \alpha_i^A\right]\right] \\
&= E_{\alpha_i^A}\left[\left(\alpha_i^A\right)^2 \operatorname{Var}\left[\alpha_i\right]\right] + \operatorname{Var}_{\alpha_i^A}\left[\alpha_i^A\right] \\
&= \left(1 + g_i^A\right) \cdot g_i + g_i^A \\
&= g_i + g_i^A + g_i \cdot g_i^A
\end{aligned} \tag{7.2}$$

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

Let $\vec{\mathbf{k}}_y^A$ 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 $\vec{\mathbf{k}}_{1998}^A = (69, 69, 53, 20, 25, 108, 64, 45, \ldots)^T$. Similarly let $\vec{\lambda}_y^A$ be a vector of expected claim counts for the "industry" in year y.

The likelihood function of $\vec{\mathbf{k}}_y^A$ conditional on α^A is given by:

$$\ell\left(\vec{\mathbf{k}}_y^A ; \vec{\lambda}_y^A, \vec{\mathbf{c}}, g \middle| \alpha^A\right) = \prod_a \ell\left(\vec{\mathbf{k}}_y^a ; \alpha^A \vec{\lambda}_y^a, \vec{\mathbf{c}}, g\right). \tag{7.3}$$

The associated log-likelihood function is given by:

$$L\left(\vec{\mathbf{k}}_y^A ; \vec{\lambda}_y^A, \vec{\mathbf{c}}, g \middle| \alpha^A\right) = \sum_a L\left(\vec{\mathbf{k}}_y^a ; \alpha^A \vec{\lambda}_y^a, \vec{\mathbf{c}}, g\right) \tag{7.4}$$

Given $g^A \geq 0$ define

$$\alpha_1^A = 1 - \sqrt{3g^A},\ \alpha_2^A = 1,\text{ and } \alpha_3^A = 1 + \sqrt{3g^A},$$

and (7.5)

$$\Pr\{\alpha^A = \alpha_1^A\} = 1/6,\ \Pr\{\alpha^A = \alpha_2^A\} = 2/3,\text{ and } \Pr\{\alpha^A = \alpha_3^A\} = 1/6.$$

The unconditional log-likelihood function is then given by:

$$L\left(\vec{\mathbf{k}}_y^A;\vec{\lambda}_y^A,\vec{\mathbf{c}},g,g^A\right) = \ln\left(\frac{e^{L\left(\vec{k}_y^A,\vec{\lambda}_y^A,\vec{c},g|\alpha_1^A\right)}}{6} + \frac{2\cdot e^{L\left(\vec{k}_y^A,\vec{\lambda}_y^A,\vec{c},g|\alpha_2^A\right)}}{3} + \frac{e^{L\left(\vec{k}_y^A,\vec{\lambda}_y^A,\vec{c},g|\alpha_3^A\right)}}{6}\right) \tag{7.6}$$

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 α is picked for each insurer for each year. In the second situation, α^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.

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

	c	g	g^A
True Value	0.0200	0.0000	0.0400
Average MLE	0.0218	0.0249	—
Std. Dev. of the MLE	0.0039	0.0158	—

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.

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

	c	g	g^A
True Value	0.0200	0.0100	0.0300
Average MLE	0.0201	0.0114	0.0213
Std. Dev. of the MLE	0.0023	0.0026	0.0090

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 8.3
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

	c	g	g^A
True Value	0.0200	0.0100	0.0300
Average MLE	0.0200	0.0104	0.0298
Std. Dev. of the MLE	0.0023	0.0029	0.0033

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 $\left(\vec{\lambda}_{iy}^{A}, c_i, g_i, g_i^{A}\right)$ of possible values of $\vec{\lambda}^{A}$, c, g and g^{A}. Let $\left\{\vec{\mathbf{k}}_y^{A}\right\}$ be a set of observations needed to calculate the likelihood function for each point $\left(\vec{\lambda}_{iy}^{A}, c_i, g_i, g_i^{A}\right)$. Let p_i be the "prior" probability of each point $\left(\vec{\lambda}_{iy}^{A}, c_i, g_i, g_i^{A}\right)$.

Then according to Bayes' Theorem, the posterior likelihood of each $\left(\vec{\lambda}_{iy}^{A}, c_i, g_i, g_i^{A}\right)$ is proportional to:

$$\prod_{y} \ell\left(\vec{\mathbf{k}}_y^{A}; \vec{\lambda}_{iy}^{A}, c_i, g_i, g_i^{A}\right) \cdot p_i \tag{9.1}$$

Let $\vec{\mathbf{e}}_y^{A}$ be a vector of exposures for the set of insurers, A, in year y. Let $\vec{\mathbf{f}}_i^{A}$ be vector of claim frequencies. Then each coordinate of the expected claim count vector $\vec{\lambda}_{iy}^{A}$ is equal to the product of the corresponding coordinates of $\vec{\mathbf{e}}_y^{A}$ and $\vec{\mathbf{f}}_i^{A}$. Since the exposures are known and the claim frequencies are unknown, we should put a prior distribution on the grid $\left(\vec{\mathbf{f}}_i^{A}, c_i, g_i, g_i^{A}\right)$.

Let $\mathcal{P}_i$ be the posterior probability of each point in the grid $\left(\vec{\mathbf{f}}_i^{A}, c_i, g_i, g_i^{A}\right)$. Then one can obtain estimates of $\vec{\mathbf{f}}_i^{A}, c_i, g_i$ and g_i^{A} by the following formulas

$$\begin{aligned} \hat{\vec{\mathbf{f}}}_i^{A} &= \sum_i \vec{\mathbf{f}}_i^{A} \cdot \mathcal{P}_i \\ \hat{c} &= \sum_i c_i \cdot \mathcal{P}_i \\ \hat{g} &= \sum_i g_i \cdot \mathcal{P}_i \\ \hat{g}^{A} &= \sum_i g_i^{A} \cdot \mathcal{P}_i \end{aligned} \tag{9.2}$$

We then tested the variability of these estimators on our simulated set of 40 insurers. The grid was constructed by varying $\vec{\mathbf{f}}_i^A, c_i, g_i$ and g_i^A in the following manner.

1. Each component of $\vec{\mathbf{f}}_0^A$ was set equal to 0.9875. Each component of $\vec{\mathbf{f}}_4^A$ 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_4 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_4 was set equal to 0.0180. The components for i=1, 2 and 3 were equally spaced in between.

4. g_0^A was set equal to 0.0200. g_4^A 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^A
Using "True" Frequencies
Derived from 100 Simulations of 40 Insurers' Data
with Industrywide Parameter Uncertainty

	c	g	g^A
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

1. Bear, Robert A., and Nemlick, Kenneth J., 1990, "Pricing the Impact of Adjustable Features and Loss Sharing Provisions of Reinsurance Treaties" *PCAS LXXVII*.

2. Bowers, N.L.; Gerber, H.U.; Hickman, J.C.; Jones, D.A.; and Nesbitt, C.J., 1997, *Actuarial Mathematics* (Second Edition), Society of Actuaries.

3. Heckman, Philip E., and Meyers, Glenn G., 1983, "The Calculation of Aggregate Loss Distributions from Claim Severity Distributions and Claim Count Distributions" *PCAS LXX*.

4. John, Russell T., and Patrik, Gary S., 1980, "Pricing Excess-of-Loss Casualty Working Cover Reinsurance Treaties" *Casualty Actuarial Society Discussion Paper Program*.

5. Klugman, Stuart A.; Panjer, Harry H. and Willmot, Gordon E., 1998, Loss Models: From Data to Decisions. John Wiley & Sons.

6. Meyers, Glenn G., 1980, "An Analysis of Retrospective Rating" *PCAS LXVII*.

7. Meyers, Glenn G., 1989, "An Analysis of the Capital Structure of an Insurance Company" *PCAS LXXVI*, 1989.

8. Meyers, Glenn G., 1999, "A Discussion of Aggregation of Correlated Risk Portfolios: Models & Algorithms, by Shaun S. Wang," to appear in the *Proceedings of the Casualty Actuarial Society*.

9. Patrik, Gary S., 1996, "Reinsurance", *Foundations of Casualty Actuarial Science, Third Edition*, Ch.6, Casualty Actuarial Society.

10. Wang, Shaun S., 1998, "Aggregation of Correlated Risk Portfolios: Models & Algorithms," to appear in the 1998 *Proceedings of the Casualty Actuarial Society*.