THE COMPOUND POISSON APPROXIMATION FOR A PORTFOLIO OF DEPENDENT RISKS *

by

M.J. Goovaerts, Katholieke Universiteit Leuven and Universiteit Amsterdam

and

J. Dhaene, Katholieke Universiteit Leuven

Abstract

A well-known approximation of the aggregate claims distribution in the individual risk theory model with mutually independent individual risks is the compound Poisson approximation. In this paper, we relax the assumption of independency and show that the same compound Poisson approximation will still perform well under certain circumstances.

Keywords

Individual model, dependent risks, compound Poisson approximation

* Work performed under grant OT/93/5 of Onderzoeksfonds K.U. Leuven.

1 Introduction

Consider a portfolio consisting of n risks labelled from 1 to n. Risk i produces a claim amount X_i during a certain reference period. The aggregate claims of the portfolio during the reference period is denoted by

$$S^{ind} = X_1 + X_2 + \ldots + X_n \tag{1.1}$$

The distribution function of S^{ind} is denoted by F^{ind}:

$$F^{ind}(s) = \Pr(S^{ind} \leq s)$$

As in Bowers et al.(1986, chapter 2) we represent each X_i as

$$X_i = I_i . V_i \qquad \text{(i=1,2,...,n)} \tag{1.2}$$

where I_i is a Bernoulli random variable which equals 1 if risk i produces at least one claim during the reference period, Pr $(I_i = 1) = q_i$ = 1 - Pr (I_i = 0), and V_i is the total claim amount produced by risk i. In the sequel we will assume that V_i equals 0 if I_i equals 0 and that V_i is positive if I_i equals 1. The total number of policies producing claims is denoted by I and is given by

$$I = I_1 + I_2 + \ldots + I_n \tag{1.3}$$

Usually the following assumption is made concerning the risks X_i :

assumption (A): *The individual risks X_i are mutually independent.*

A well-known approximation of the distribution of S^{ind} under this assumption, is the compound Poisson approximation, i.e. F^{ind} is approximated by F^{cP} with

$$F^{cP}(s) = \sum_{n=0}^{s} \Pr(K = n)\; F^{*n}(s) \qquad \text{(s=0,1,...)} \tag{1.4}$$

where K is a Poisson distributed random variable with parameter λ given by

$$\lambda = \sum_{i=1}^{n} q_i \tag{1.5}$$

and F(s) is the distribution given by

$$F(s) = \frac{1}{\lambda}\sum_{i=1}^{n} q_i \Pr(V_i \le s | I_i = 1) \tag{1.6}$$

In order to be able to state results for the error related to this approximation, we introduce the following distance between the distributions F_X and F_Y of random variables X and Y

$$d(F_X, F_Y) = \sup_J |\Pr(X \in J) - \Pr(Y \in J)| = \frac{1}{2}\int_{-\infty}^{+\infty} |dF_X(s) - dF_Y(s)| \tag{1.7}$$

where the supremum is taken over all events J.
Remark that, apart from a constant factor, d(.,.) is the total variation distance.

Gerber (1984) proved that, under assumption (A), the following result holds

$$d(F^{ind}, F^{cP}) \le \sum_{i=1}^{n} q_i^2 \tag{1.8}$$

Michel (1987) proved that if the conditional claim amounts $V_i | I_i = 1$ all have the same distribution and if assumption (A) holds, that

$$d(F^{ind}, F^{cP}) \le \frac{1}{\lambda}\sum_{i=1}^{n} q_i^2 \tag{1.9}$$

which is an improvement of Gerber's bound if $\lambda > 1$.

In this paper we will further work within the framework of Michel's "quasi-homogeneous" portfolio but without assuming the mutual independence of the risks involved.
We will use a result of Chen (1975) to show that in certain cases the compound Poisson approximation as defined in (1.4) will still be usable for approximating a portfolio of mutually dependent risks.

2 The Chen-Stein method

From now on, we will relax the independency assumption (A) and consider the following assumption concerning the dependency of the risks $X_i (i = 1,2,\ldots,n)$.

assumption (B): *The conditional claim amounts $V_i | I_i = 1$ are mutually independent.*
However, the indicators I_i are not assumed to be mutually independent.

Every risk X_i can be described by its indicator and by its conditional claim amount. Assumption (B) states that the dependency between the individual risks is caused by the dependency between the indicators.

For a portfolio of insurances which provide a fixed amount in case a claim occurs, the conditional claim amounts are deterministic so that assumption (B) holds in this case. For a portfolio of insurances which compensate the loss incurred after a claim, it will often be the occurrences of claims that will be more or less strongly dependent, while the dependency of the conditional claim amounts will be much weaker. Hence, in this case assumption (B) will often offer a first attempt to describe the dependency between the risks.

In the sequel, we will consider a portfolio which is quasi-homogeneous in the sense that the conditional claim amounts $V_i | I_i = 1$ all have the same distribution F, say. We will consider the compound Poisson approximation F^{cP} defined by (1.4) and (1.5) for the distribution F^{ind} of this portfolio.

The following lemma is an extension of a result of Michel (1987) who proved it if assumption (A) holds.

Lemma 1

Let F^{ind} be the aggregate claims distribution of a portfolio for which assumption (B) holds and where all the conditional claim amounts have the same distribution F. Then the distance between F^{ind} and F^{cP} defined by (1.4) is bounded by

$$d(F^{ind}, F^{cP}) \le d(F_I, F_K)$$

where F_I is the distribution of the number of claims I defined in (1.3) and F_K is the distribution of a Poisson distributed random variable with parameter λ given by (1.5).

<u>Proof:</u>

Let S^{cP} be a random variable with distribution F^{cP} then we have that

$$\begin{aligned}
&\Pr(S^{ind} \in J) - \Pr(S^{cP} \in J) \\
&= \sum_{k=0}^{n} \Pr(S^{ind} \in J, I = k) - \sum_{k=0}^{\infty} \Pr(S^{cP} \in J, K = k) \\
&\le \sum_{k=0}^{n} \left[\Pr(S^{ind} \in J, I = k) - \Pr(S^{cP} \in J, K = k) \right]
\end{aligned}$$

Note that under assumption (B) we have for any possible outcome of k of I

$$\Pr(S^{ind} \in J | I = k) = \Pr(S^{cP} \in J | K = k)$$

so that

$$\Pr(S^{ind} \in J) - \Pr(S^{cP} \in J) \le \sum_{k=0}^{n} (\Pr(I = k) - \Pr(K = k)).\Pr(S^{cP} \in J | K = k)$$

Now let $J_0 = \{k \in \{0,1,\dots,n\} | \Pr(I = k) > \Pr(K = k)\}$

then we find

$$\Pr(S^{ind} \in J) - \Pr(S^{cP} \in J) \le \sum_{k \in J_0} (\Pr(I = k) - \Pr(K = k))$$
$$\le \sup_J (\Pr(I \in J) - \Pr(K \in J))$$

The desired result then follows from

$$d(F^{ind}, F^{cP}) = \sup_J \left| \Pr(S^{ind} \in J) - \Pr(S^{cP} \in J) \right| = \sup_J (\Pr(S^{ind} \in J) - \Pr(S^{cP} \in J))$$

Q.E.D.

For each Bernoulli random variable $I_i (i = 1,2,\dots,n)$ we now define the set of dependence $B_i \subset \{1,2,\dots,n\}$ such that

I_j is independent of I_i if and only if $j \notin B_i$ (j=1,2,...,n)

Further, define

$$b_1 = \sum_{i=1}^{n} \sum_{j \in B_i} q_i q_j \tag{2.1}$$

$$b_2 = \sum_{i=1}^{n} \sum_{i \ne j \in B_i} E(I_i I_j) \tag{2.2}$$

The following result can be found in Chen (1975).

Lemma 2

Using the same notation and assumptions as in Lemma 1 we have that

$$d(F_I, F_K) \le (b_1 + b_2) \frac{1 - e^{-\lambda}}{\lambda}$$

A more general version of this result appears in Arratia et al. (1990). They refer to this approach as the Chen-Stein method and present several applications of it. One of such applications is the study of longest repetitive patterns in random sequences, see e.g. Waterman (1995). In the following theorem we show that the Chen-Stein method can also be useful in individual risk theory.

Theorem 1

Let F^{ind} be the aggregate claims distribution under assumption (B). If all the conditional claim amounts have the same distribution F, then the distance between F^{ind} and F^{cP} defined in (1.4) is bounded by

$$d(F^{ind}, F^{cP}) \leq (b_1 + b_2)\frac{1-e^{-\lambda}}{\lambda}$$

where λ, b_1 and b_2 are defined by (1.5), (2.1) and (2.2) respectively.

Proof:

The proof follows immediately from Lemma 1 and Lemma 2.

Q.E.D.

The bound presented in Theorem 1 will work best for dealing with local dependence, corresponding to situations in which the sets of dependence B_i have only a few elements so that b_1 and b_2 are small. The approximations are useful only if second moments are well behaved. Remark that when the dependence structure is local, finding the Chen-Stein bounds involves the same effort as computing first and second moments of the total number of claims.

Let us now look at the special case where assumption (A) holds. Then we find that $B_i = \{i\} (i = 1,2,\ldots,n)$ and hence $b_1 = \sum_{i=1}^{n} q_i^2$, $b_2 = 0$ so that we obtain from Theorem 1

$$d(F^{ind}, F^{cP}) \leq \sum_{i=1}^{n} q_i^2 \frac{1-e^{-\lambda}}{\lambda} \qquad (2.3)$$

If $\lambda < 1$ we find

$$d(F^{ind}, F^{cP}) \leq \sum_{i=1}^{n} q_i^2$$

which is a special case of Gerber's (1984) more general result.

For $\lambda > 1$ we find from (2.3)

$$d(F^{ind}, F^{cP}) \le \frac{1}{\lambda}\sum_{i=1}^{n} q_i^2$$

which is Michel's (1987) bound.

3 Example

In order to demonstrate the usefulness of Theorem 1 for certain real life situations, we give the following illustration.
Consider a portfolio consisting of (m+n) life insurances providing a death benefit. There are m couples (wife and husband) in the portfolio and all death benefits are equal to 1 (which means that in fact we are looking at the total number of deaths during the reference period). Then we can write the aggregate claims as

$$S^{ind} = \sum_{i=1}^{m}(X_i + X_i^{'}) + \sum_{i=m+1}^{n} X_i$$

We assume that all risks are mutually independent, except for the "coupled" risks. This means that the only dependence that occurs is the dependence between the risks of a wife and her husband.

The sets of dependence are then given by

$$B_i = B_i^{'} = \{i, i^{'}\} \qquad \text{(i=1,2,...,m)}$$

$$B_i = \{i\} \qquad \text{(i=m+1,...,n)}$$

and hence

$$\lambda = \sum_{i=1}^{m}(q_i + q_i^{'}) + \sum_{i=m+1}^{n} q_i$$

$$b_1 = \sum_{i=1}^{m}(q_i + q_i^{'})^2 + \sum_{i=m+1}^{n} q_i^2$$

$$b_2 = 2\sum_{i=1}^{m}(q_i . q_i^{'} + \mathrm{cov}(X_i, X_i^{'}))$$

From Theorem 1 we find the following error bound for the (compound) Poisson approximation of this portfolio.

$$d(F^{ind},F^{cP}) \le \frac{1}{\lambda}\left\{\sum_{i=1}^{m}\left[\left(q_i+q_i^{'}\right)^2+2q_iq_i^{'}+2\,\mathrm{cov}\left(X_i,X_i^{'}\right)\right]+\sum_{i=m+1}^{n}q_i^2\right\} \tag{3.1}$$

We denote Michel's upper bound (1.9), which is valid under the independence assumption (A) by M:

$$M=\frac{1}{\lambda}\left\{\sum_{i=1}^{m}\left(q_i^2+q_i^{'2}\right)+\sum_{i=m+1}^{n}q_i^2\right\}$$

Hence, from (3.1) we find

$$d\left(F^{ind},F^{cP}\right)\le M+\frac{2}{\lambda}\sum_{i=1}^{m}\left(2q_iq_i^{'}+\mathrm{cov}\left(X_i,X_i^{'}\right)\right) \tag{3.2}$$

If we don't have any information about $\mathrm{cov}\left(X_i,X_i^{'}\right)$ then we can use the following upper bound for this covariance:

$$\mathrm{cov}(X_i,X_i^{'})=\sqrt{q_i(1-q_i)q_i^{'}(1-q_i^{'})}\,corr(X_i,X_i^{'})\le\sqrt{q_i(1-q_i)q_i^{'}(1-q_i^{'})}$$

In order to establish the effect on the bound from introducing dependence in the portfolio, assume that all claim probabilities are of the same order, let us say all are equal to q, then we find

$$\lambda=(m+n)q \quad \text{and} \quad M=q$$

so that (3.2) becomes

$$d(F^{ind},F^{cP})\le M+\frac{2m}{m+n}\left[2q+(1-q)\,\mathrm{corr}(X_i,X_i^{'})\right] \tag{3.3}$$

which shows that increasing the relative number of couples or increasing the correlation coëfficients will lead to an increased bound.

As a numerical illustration consider the case that $q\le {}^{10}\!/_{11}$ and

$$\Pr(X_i^{'}=1|X_i=1)=1{,}1q$$

This means that

$$\mathrm{corr}(X_i,X_i^{'})=0{,}1\frac{q}{1-q}$$

Further, let $\frac{m}{m+n} = 0{,}05$, which means that 10 % of the portfolio consists of couples.

Then we find from (3.3)

$$d(F^{ind}, F^{cP}) \leq 1{,}21q$$

which indicates that the bound is increased by ± 20 % if 10 % of the portfolio consists of couples with dependent risks and if the mortality rate of a person is increased by 10 %, given the mortality of his spouse during the reference year.

References

Arratia, R.A.; Goldstein, L. and Gordon, L. (1990). Poisson approximation and the Chen-Stein method. Statistical Science, 5, pp. 403-434.

Bowers, N.L.; Gerber, H.U.; Hickman, J.C.; Jones, D.A. and Nesbitt, C.J. (1986). Actuarial Mathematics. Society of Actuaries, Itacsa, IL.

Chen, L.H.Y. (1975). Poisson approximation for dependent trials. Ann. Probab., 3, pp. 534-545.

Gerber, H.U. (1984). Error bounds for the compound Poisson approximation. Insurance: Mathematics and Economics, vol.3, pp. 191-194.

Kaas, R. ; Gerber, H.U. (1994). Some alternatives for the individual model. Insurance: Mathematics and Economics 15, pp. 127-132.

Michel, R. (1987). An improved error bound for the compound Poisson approximation. ASTIN Bulletin, vol. 17, pp. 165-169.

Waterman, S. (1995). Hearing distant echoes: using extremal statistics to probe evolutionary origins. Chapter 4 of: Calculating the secrets of life. Applications of the mathematical sciences in molecular biology, Lander, E.S. and Waterman, M.S. (editors). National Academy Press, Washington, D.C.