AN APPROPRIATE WAY TO SWITCH FROM THE INDIVIDUAL RISK MODEL TO THE COLLECTIVE ONE

BY S KUON, M. RADTKE AND A. REICH

The Cologne Re, Cologne

ABSTRACT

For some time now, the convenient and fast calculability of collective risk models using the Panjer-algorithm has been a well-known fact, and indeed practitioners almost always make use of collective risk models in their daily numerical computations. In doing so, a standard link has been preferred for relating such calculations to the underlying heterogeneous risk portfolio and to the approximation of the aggregate claims distribution function in the individual risk model. In this procedure until now, the approximation quality of the collective risk model upon which such calculations are based is unknown.

It is proved that the approximation error which arises does not converge to zero if the risk portfolio in question continues to grow. Therefore, necessary and sufficient conditions are derived in order to obtain well-adjusted collective risk models which supply convergent approximations. Moreover, it is shown how in practical situations the previous natural link between the individual and the collective risk model can easily be modified to improve its calculation accuracy. A numerical example elucidates this.

KEYWORDS

Individual risk model; collective risk model; modified natural approximation; aggregate claims distribution; Berry-Esséen bound.

INTRODUCTION

For decades one of the central themes of risk theory has been the calculation of the aggregate claims distribution of a portfolio. The aim of this paper is to take this subject and shed a new light on theoretical aspects and practical applications.

In the eighties, with the development of recursive algorithms, a considerable degree of progress was made towards the numerical calculation of the aggregate claims distribution for both the individual and the collective risk model. In particular, the special collective risk models considered by Panjer (1981) are generally accepted by practitioners as being adequate, and the use of Panjer's algorithm has meanwhile become a widespread standard technique of actuaries.
In applied risk theory the $n$ policies $X_i$ of which a portfolio is composed are usually independent but, as a rule, not identically distributed random variables. Instead of the collective risk model, in practice one is initially concerned with the individual risk model, in which the calculation of the distribution function of

$$S^{\text{ind}} = \sum_{i=1}^{n} X_i$$

is a fundamental task. The fact that the above-mentioned collective risk model can be calculated so quickly has led in the practical application to a switch from the individual risk model to a collective risk model, in the hope that the error which inevitably occurs as a result is sufficiently small. So a (hopefully) appropriate collective risk model is linked to the individual risk model.

Until now, when this link was being made, it was not the whole class of collective risk models

$$S^{\text{coll}} = \sum_{i=1}^{N} Z_i$$

(with independent identically distributed random variables $Z_i$ and random claims number $N$ independent of the sequence of single claims amounts $Z_i$, $N$ in the Panjer-class) which was considered with regard to its suitability. Rather, in literature and in practice a "classical link", which is described precisely e.g. in Gerber (1979, p. 50), and, for our purposes, in Section 1, Remark 1.4, has become generally accepted. Here, the $N$ (whether binomial, Poisson or negative binomial distributed) and the $(Z_i)$, both characterizing the collective risk model, are clearly determined by the individual risk model. In practice $N$ is almost always chosen as the Poisson distributed claims number.

For the error

$$\Delta = \sup_{x \in \mathbb{R}} |P(S^{\text{ind}} \leq x) - P(S^{\text{coll}} \leq x)|$$

the paper by Hipp (1985) provides an error estimate for the classical link to the compound Poisson model which is small enough for various practical applications. This sharpens an error estimate given by Gerber (1984).

However, for very large portfolios, these error estimates become so bad that they are unusable — which does not of course rule out the fact that the error $\Delta$ itself may converge to zero for portfolios which are becoming increasingly large. (The meaning of this is to be defined more specifically.)

With regard to the standard link to the compound Poisson model, in Section 1 of this paper, proof is supplied for the (surprising?) results that this error does not in fact converge to zero. This also applies when the Poisson distribution is replaced by the negative binomial distribution. In the binomial case, the situation has proved to be ambiguous (cf. Section 1, Model 1.1). In
short, the methods normally used in practice have proved to be bad for large portfohos.

These results can be derived from the answers to more general questions concerning the connections between the individual and collective risk models. These questions are of interest in their own right and of fundamental significance, and they refer in the first instance to general collective risk models with weak additional conditions. In particular there is no requirement for a collective risk model to emerge from an individual risk model in the standard manner.

The requirement

\[ \Delta \to 0 \]  

(for portfolio size growing to infinity) is obviously a theoretically reasonable (asymptotic) quality criterion for judging whether individual risk models can be adjusted precisely by means of collective risk models. This immediately gives rise to two questions:

With regard to (2), are there equivalent and simple conditions which make it possible to check the validity of (2) in concrete cases? Is the theoretical quality criterion (2) also a relevant measure of quality for practical applications, or, to put it more precisely, is the assumption contained in (2) that \( \Delta \) becomes small equivalent to the assertion that the difference in the two risk premiums does not become overly large?

Both questions are answered in the affirmative with Theorem 2.1, the first question in particular being answered by the fact that (2) is equivalent to the (mostly easily verifiable) condition

\[ \frac{\text{Var} S^\text{coll}}{\text{Var} S^\text{ind}} \to 1. \]  

The more comprehensive result of Theorem 2.2 represents a quantitative sharpening of Theorem 2.1 which is particularly interesting because equivalent conditions are given for situations where the difference in the two portfolio premiums even remains bounded. A useful tool for proving these central statements of the paper is provided by the often neglected paper by von Chossy, R. and G. Rappl (1983); here the possibility of representing stochastic sums by means of deterministic sums is proved. These results and required Berry-Essen bounds are presented separately in the Appendix.

An important point for the practical application is that for good approximations, in addition to the requirement that the expected values should be equal, it would now, in view of (3), be appropriate to seek and construct collective risk models with

\[ \text{Var} S^\text{ind} = \text{Var} S^\text{coll}. \]

In order to ensure that collective risk models can be calculated quickly, only collective risk models belonging to the Panjer-class are suitable. On the other hand—as mentioned at the beginning of this paper—collective risk models
which emerge from an individual risk model via the standard link are ruled out. However, converging approximation models which are simple to construct can be obtained by scaling the range of the single claims amount in the standard link.

If an appropriate scaling factor and the parameter modifications corresponding to it are chosen, then for all three claims number distributions of the Panjer-class, the equality of the first two moments can be achieved in the individual and the collective risk model. Moreover, in Section 3 an analysis is carried out to show that the best adjustment should be reached with the compound binomial model. This is also verified by several numerical examples, which can be taken from the Gerber portfolio (GERBER 1979, p. 53). Thus, in practical applications, instead of the standard link to the compound Poisson model, a modified compound binomial model, which is described precisely in Section 3, should be used (cf. JEWELL and SUNDT (1981)).

1. The link between a given individual risk model and the related collective risk model

In the following $X_i$ denotes the amount of claims produced by risk $i$, $i \in \mathbb{N}$. The single risks are understood to be numbered in a suitable way. Their, in future, undefined claims amounts are understood as random variables. The accidental aggregate claims amount resulting from a segment of $n$ risks, that is the sum of all single claims amounts, is called an individual risk model, if the following is valid:


The individual claims amounts $X_i$, $1 \leq i \leq n$, $n \in \mathbb{N}$, set up a sequence of independent, in general not identically distributed random variables $X_i$ such that $X_i \geq 0$. $X_i = 0$ means that risk $i$ does not produce a claim. The random variable $S^{\text{ind}} = \sum_{i=1}^{n} X_i$ is called the aggregate claims amount of the individual risk model.

We shall write $S^{\text{ind}}_n$ instead of $S^{\text{ind}}$ to indicate the dependency of $S$ on the size of the underlying risk segment. As no misunderstanding is possible below, we will drop the index $n$ there. In view of later considerations and in order to make the model tractable, we shall impose additional conditions.

**Assumption 1.1.** The sequence of random variables $(X_i)$, $i \in \mathbb{N}$, fulfills the inequalities $0 < c \leq EX_i \leq d < \infty$ and $0 < a \leq \text{Var } X_i \leq b < \infty$, where $a, b, c, d$ are real-valued constants

Assumption 1.1 does not impose any restrictions on practical applications, excluding only unrealistic cases. The validity of Assumption 1.1 follows, as a
rule, from the fact that in practical applications the sequence of random variables \( X_i \) are even uniformly bounded, that is \( \sup_{i \in \mathbb{N}} |X_i| < \infty \)

**Remark 1.1.** In accordance with the sequence \((X_i), i \in \mathbb{N}\), we can define a sequence of independent random variables \((I_i), i \in \mathbb{N}\), by setting \( I_i = 1_{\{X_i > 0\}} \). In addition to this we can go on to construct a sequence of independent random variables \((Y_i), i \in \mathbb{N}\), by postulating for their one-dimensional distribution functions:

\[
P(Y_i \leq x) = \frac{P(X_i \leq x) - P(X_i = 0)}{P(X_i > 0)} \quad \text{if} \quad x > 0; \quad P(Y_i = 0) = 0
\]

Thus, we have the representation \( X_i \overset{	ext{dist}}{=} I_i Y_i \) for each \( i \in \mathbb{N} \). \( Y_i \) can be interpreted as the claims amount of risk \( i \), provided that this risk produces a claim. The claims number \( N^* \) in the individual risk model is established by

\[
N^* = \sum_{i=1}^{n} I_i, \quad I_i \overset{	ext{dist}}{=} \beta(1, q_i) \quad \text{with} \quad q_i = P(X_i > 0) \quad \text{and} \quad \beta \quad \text{the binomial distribution.}
\]

The events \( \{I_i = 1\} \) and \( \{Y_i \leq x\}, x > 0 \) arbitrary, are independent.

In many practical situations the calculation of the distribution function \( F_{\text{ind}} \) of the individual aggregate claims amount \( S_{\text{ind}} \) is of fundamental importance. However, its precise numerical computation is in general impossible without the support of a computer and, in spite of recent recursive algorithms \((\text{Kornya (1983)}, \text{Hipp (1985, 1986)}, \text{De Pril (1989)})\), still costly. Therefore, at a very early stage of risk theory, the question of the calculability of \( F_{\text{ind}} \) led to the concept of the collective risk model \((\text{Bowers et al. (1987), p. 317)})\), which is easier to handle when theoretical considerations are made. Its fast numerical calculability \((\text{Panjer-class})\) is another, more recent reason for using it.

In the following we shall denote by \( N \) the random number of claims occurring in a risk portfolio in a given period, and by \( Z_i \) the accidental amount of the \( i \)th-claim, \( i \leq N \). We will then be speaking of a collective risk model, if we state the following


The random collective claims amounts \( Z_i, i \in \mathbb{N} \), set up a sequence of independent, identically distributed random variables such that \( Z_i > 0 \) for each \( i \in \mathbb{N} \). The sequence \( Z_i, i \in \mathbb{N} \), is assumed to be independent of the random claims number \( N \). The random variable \( N \) takes on non-negative integer values.

The random variable \( S_{\text{coll}} = \sum_{i=1}^{N} Z_i \) (with \( S_{\text{coll}} = 0 \) if \( N = 0 \)) is then called the aggregate claims amount of the collective risk model. For \( N \) and \( (Z_i), i \in \mathbb{N} \), we assume in addition: \( 0 < EN < \infty, 0 < \text{Var} \ N < \infty, 0 < E Z_i < \infty, 0 < \text{Var} \ Z_i < \infty \).
Remark 1.2. $S^{\text{coll}}$ satisfies $ES^{\text{coll}} = ENEZ_1$ and $\text{Var } S^{\text{coll}} = EN \text{Var } Z_1 + \text{Var } N(EZ_1)^2$.

The link between Definition 1.1 and Definition 1.2 at once becomes clear when we refer to the individual risk model with independent, identically distributed random variables $X_i$, $1 \leq i \leq n$. This, in turn, brings us to

Remark 1.3. If the individual model satisfies $X_i \sim F$ for each $i$, it follows that $N^* \sim \beta(n, q)$ with $q = P(X_i > 0)$. Put $N^* = N^*$ and $Z_1 \sim G$, where $G(x) = (F(x) - (1 - q))/q$ for $x \geq 0$. Thus we get $S^{\text{ind}} \equiv S^{\text{coll}}$ with claims number distribution $\beta(n, q)$ in the collective risk model.

In general the question arises how the distribution functions of $N$ and $Z_1$ should be chosen such that the distribution function $F^{\text{coll}}$ of $S^{\text{coll}}$ supplies a good approximation to the distribution function $F^{\text{ind}}$ of $S^{\text{ind}}$. The following procedure is usual:

Remark 1.4. Define the distribution function $G$ of $Z_1$ by

$$(1.2) \quad G(x) = \sum_{i=1}^{\infty} \frac{q_i}{nq} G_i(x) \quad \text{with} \quad G_i(x) = \frac{F_i(x) - (1 - q_i)}{q_i}, \quad x \geq 0,$$

and

$$(1.3) \quad q = \frac{1}{n} \sum_{i=1}^{\infty} q_i, \quad q_i = P(X_i > 0), \quad X_i \sim I_i Y_i, \quad X_i \sim F_i, \quad Y_i \sim G_i.$$

In this remark the representation $X_i \sim I_i Y_i$ is such as given in Remark 1.1.

Consequently we have $Z_1 > 0$ and $EZ_1^m = \frac{1}{nq} \sum_{i=1}^{\infty} EX_i^m < \infty, m = 1, 2$.

Assumption 1.1 establishes the existence of real-valued constants $a', b', c'$, $d'$ (independent of $n$) with $0 < c' \leq EZ_1 < d' < \infty$ and $0 < a' \leq \text{Var } Z_1 \leq b' < \infty$. Note that the distribution function $G$ of $Z_1$ depends on $n$.

The last remark results in three different collective risk models, each of them specified by the choice of the claims number distribution (Panjer-class).

Model 1.1. The natural approximation (compound binomial approximation). Let

$$N \sim \beta(n, q) \quad \text{and} \quad Z_1 \sim G,$$

$G$ as defined in (1.2). Then

$$ES^{\text{coll}} = ES^{\text{ind}} \quad \text{and} \quad \text{Var } S^{\text{coll}} = \text{Var } S^{\text{ind}} + \Delta_{b'},$$
where

\[ \Delta_{B_i} = \sum_{i=1}^{n} (EX_i)^2 - \frac{1}{n} \left( \sum_{i=1}^{n} EX_i \right)^2. \]

\[ \Delta_{B_i} \geq 0, \text{ since } n \sum_{i=1}^{n} (EX_i)^2 \geq \left( \sum_{i=1}^{n} EX_i \right)^2. \]

\[ \Delta_{B_i} = 0 \Leftrightarrow EX_i = EX_1 \text{ for each } i = 1, \ldots, n. \]

The natural approximation can also be derived from an individual risk model as follows. Put

\[ S_{\text{coll}} = \sum_{i=1}^{n} Z_i^* \text{ with } Z_i^* \overset{\sim}{\in} G^*, \]

where for \( x \geq 0 \)

\[ (1.5) \quad G^*(x) = \frac{1}{n} \sum_{i=1}^{n} F_i(x) \quad \text{and} \quad X_i \overset{\sim}{\in} F_i. \]

Since

\[ (1.6) \quad G^*(x) = \frac{1}{n} \sum_{i=1}^{n} (1 - q_i) + \frac{1}{n} \sum_{i=1}^{n} q_i \frac{F_i(x) - (1 - q_i)}{q_i} = (1 - q) + qG(x), \]

we conclude from the characteristic function

\[ (1.7) \quad (Ee^{itZ_i^*})^n = \left( (1 - q) + q \int_{0}^{\infty} e^{ux} G(dx) \right)^n \]

\[ = \sum_{k=0}^{n} \binom{n}{k} (1 - q)^{n-k} q^k \int_{0}^{\infty} e^{ux} G^k(dx), \]

that the two approaches lead to the same collective model.

**Model 1.2.** The compound Poisson approximation

Let

\[ N \overset{\sim}{\in} \pi(nq), \pi \text{ the Poisson distribution, and } Z_1 \overset{\sim}{\in} G, \]

\( G \) as defined in (1.2). Then

\[ ES_{\text{coll}} = ES_{\text{ind}} \quad \text{and} \quad \text{Var } S_{\text{coll}} = \text{Var } S_{\text{ind}} + \Delta_{P_0}, \]
where

\[(1.8) \Delta_{P_0} = \sum_{i=1}^{n} (EX_i)^2.\]

**Model 1.3.** The compound negative binomial approximation

Let

\[N \sim \text{NegBin}(n, \frac{1}{1+q}), \quad A \sim \text{NegBin}, \quad G \text{ as defined in (1.2)}.\]

Then

\[ES^{\text{coll}} = ES^{\text{ind}} \quad \text{and} \quad \text{Var} S^{\text{coll}} = \text{Var} S^{\text{ind}} + \Delta_{NB},\]

where

\[(1.9) \Delta_{NB} = \sum_{i=1}^{n} EX_i^2 + \frac{1}{n} (ES^{\text{ind}})^2 \]

\[= \text{Var} S^{\text{ind}} + \sum_{i=1}^{n} (EX_i)^2 + \frac{1}{n} (ES^{\text{ind}})^2.\]

Thus, the three collective risk approximation models correctly adjust the expected claims number \(E \sum_{i=1}^{n} 1_{\{X_i > 0\}} = nq = EN\) and the expected aggregate claims amount, although they overestimate \(\text{Var} S^{\text{ind}}\). Obviously, for the overestimation the following is valid:

\[0 \leq \Delta_{Bi} < \Delta_{P_0} < \Delta_{NB}.\]

In respect of Assumption 1.1 a simple calculation leads to the following result, because \(EX_i\) and \(\text{Var} X_i\) are uniformly bounded.

(i) \(N\) binomial distributed:

\[(1.10) \frac{\text{Var} S^{\text{coll}}}{\text{Var} S^{\text{ind}}} - 1 = \frac{\Delta_{Bi}}{\sum_{i=1}^{n} \text{Var} X_i} \in \left[0, \frac{d^2 - c^2}{a}\right],\]

(ii) \(N\) Poisson distributed:

\[(1.11) \frac{\text{Var} S^{\text{coll}}}{\text{Var} S^{\text{ind}}} - 1 = \frac{\Delta_{P_0}}{\sum_{i=1}^{n} \text{Var} X_i} \in \left[\frac{c^2}{b}, \frac{d^2}{a}\right].\]
(iii) $N$ negative binomial distributed:

$$\frac{\text{Var } S^{\text{coll}}}{\text{Var } S^{\text{ind}}} - 1 = \frac{\Delta_{NB}}{\sum \text{Var } X_i} \in \left[ 1 + \frac{c^2}{b} \left( 1 + \frac{1}{n} \right), 1 + \frac{d^2}{a} \left( 1 + \frac{1}{n} \right) \right].$$

Hence, only in the case of binomial distributed claims number $N$ we can achieve $\text{Var } S^{\text{ind}} = \text{Var } S^{\text{coll}}$. For instance, this is fulfilled if $E X_i = E X_1$ for all $i = 1, \ldots, n$ or if even all of the $X_i$ are identically distributed (cf. Model 1.1). The following example shows that in general the variance ratio in the compound binomial approximation does not equal to 1 either.

**Example 1.1.** Look at a sequence of random variables $(X_i)$, $1 \leq i \leq n$, $X_i \in \{0, 1\}$. Let for each $i$

$$P(X_i = 1) = \begin{cases} 3/4, & \text{i even} \\ 1/4, & \text{i odd}. \end{cases}$$

Therefore we have

$$EX_i = \begin{cases} 3/4, & \text{i even} \\ 1/4, & \text{i odd}. \end{cases} \quad \text{and} \quad \text{Var } X_i = \begin{cases} 3/16, & \text{i even} \\ 1/4, & \text{i odd}. \end{cases}$$

From that

$$\frac{\text{Var } S^{\text{coll}}}{\text{Var } S^{\text{ind}}} = \begin{cases} 4/3, & n \text{ even} \\ (4/3) - 1/(3 n^2), & n \text{ odd} \end{cases}$$

easily is concluded.

Further on we shall analyze the impact to which $\text{Var } S^{\text{ind}} \neq \text{Var } S^{\text{coll}}$ leads in the case of premium calculations which are based on the above-mentioned approximation models instead of the individual risk model. As these assertions depend on the number $n$ of risks underlying the portfolio at issue, we shall now add the dropped index $n$ to our previous notations, thus $S_n^{\text{ind}}$ instead of $S^{\text{ind}}$, $S_n^{\text{coll}}$ instead of $S^{\text{coll}}$, etc.

Assumption 1.1 instantly implies $E S_n^{\text{ind}} \xrightarrow{n \to \infty} \infty$, $\text{Var } S_n^{\text{ind}} \xrightarrow{n \to \infty} \infty$, and, as $E S_n^{\text{ind}} = E S_n^{\text{coll}}$, also $E S_n^{\text{coll}} \xrightarrow{n \to \infty} \infty$. As shown above, the variances of the collective and the individual risk models differ from one another in general. Only in the case of binomial distributed claims number the variance ratio can converge to 1. In particular we have $\text{Var } S_n^{\text{coll}} - \text{Var } S_n^{\text{ind}} \xrightarrow{n \to \infty} \infty$ in most situations.

The consequences of the overestimated actual variance for premium calculation by means of collective risk models is demonstrated using the percentile
premium. Let

\[ \mathcal{J}_n^{\text{ind}}(\alpha) = \inf \{ x \mid P(S_n^{\text{ind}} \leq x) \geq \alpha \} \]

with security level \( \alpha \in (0, 1) \). \( \mathcal{J}_n^{\text{coll}}(\alpha) \) analogously. That is, the premium \( \mathcal{J}_n^{\text{ind}}(\alpha) \) is not exceeded by the aggregate claims amount \( S_n^{\text{ind}} \) with probability \( \alpha \). Of course the difference \( \mathcal{J}_n^{\text{ind}}(\alpha) - \mathcal{J}_n^{\text{coll}}(\alpha) \) is of interest. This heuristic reflection serves to motivate the following. Under assumptions which are always satisfied in practice, we obtain approximately the following result if \( n \) is large enough (cf. Lemma A.2 (i)):

\[ \mathcal{J}_n^{\text{ind}}(\alpha) \approx \mu_n^{\text{ind}} + \Phi^{-1}(\alpha) \sigma_n^{\text{ind}} \]

and

\[ \mathcal{J}_n^{\text{coll}}(\alpha) \approx \mu_n^{\text{coll}} + \Phi^{-1}(\alpha) \sigma_n^{\text{coll}} \]

where \( \mu_n^{\text{ind}} = E S_n^{\text{ind}} \), \( \mu_n^{\text{coll}} = E S_n^{\text{coll}} \), \( \sigma_n^{\text{ind}} = \sqrt{\text{Var} S_n^{\text{ind}}} \), \( \sigma_n^{\text{coll}} = \sqrt{\text{Var} S_n^{\text{coll}}} \). \( \Phi^{-1} \) denotes the inverse function of standardized normal distribution function \( \Phi \).

Thus, as \( \mu_n^{\text{ind}} = \mu_n^{\text{coll}} \) the premium difference \( \mathcal{J}_n^{\text{coll}}(\alpha) - \mathcal{J}_n^{\text{ind}}(\alpha) \) of the risk models under consideration directly depends on the difference of the standard deviation, namely

\[ \mathcal{J}_n^{\text{coll}}(\alpha) - \mathcal{J}_n^{\text{ind}}(\alpha) \approx (\sigma_n^{\text{coll}} - \sigma_n^{\text{ind}}) \Phi^{-1}(\alpha) \geq 0. \]

A further analysis shall show that the difference \( \sigma_n^{\text{coll}} - \sigma_n^{\text{ind}} \) is strictly related to the term \( \sup \{ |F_n^{\text{ind}}(x) - F_n^{\text{coll}}(x)| \} \).

### 2. Approximation of an individual risk model by a collective risk model

In this section, at first we focus our analysis on the approximation of individual risk models \( S_n^{\text{ind}} = \sum_{i=1}^{n} X_i \) of growing size by a sequence of so-called homogeneous collective risk models \( S_n^{\text{coll}} = \sum_{i=1}^{N_n} Z_i \). We shall deduce our main results in Theorem 2.1 and Theorem 2.2 and then apply these results to a reasonable concept for a portfolio growth which conducts to Corollary 2.1 and Corollary 2.2. We start with

**Definition 2.1.** We call \( (S_n^{\text{coll}})_{n \in \mathbb{N}} \) a sequence of homogeneous collective risk models if, for each \( n \in \mathbb{N} \), \( S_n^{\text{coll}} = \sum_{i=1}^{N_n} Z_i \) is a collective risk model and the distribution function of \( Z \), is independent of \( n \). In addition we assume that \( N_n \) possesses a representation \( N_n \sim L^{*n} \) with arbitrary distribution function \( L \) on \( \mathbb{N}_0 \) (cf. Proposition A.1).
AN APPROPRIATE WAY TO SWITCH FROM THE INDIVIDUAL RISK MODEL

In the following, the notations used are the same as in Section 1. The sequence \((X_i)_{i \in \mathbb{N}}\) is supposed to fulfill Definition 1.1, Assumption 1.1 and \(\sup_{i \in \mathbb{N}} \{EX_i^3\} < \infty\). In addition to Definition 2.1, we assume that \(EZ_i^3 < \infty\), \(EN_i^3 < \infty\) for all \(n \in \mathbb{N}\). Under these conditions, all the results listed in the Appendix are applicable to \(S_n^\text{ind}\) and \(S_n^\text{coll}\) and to their distribution functions \(F_n^\text{ind}\) and \(F_n^\text{coll}\).

Let our analysis start from the supremum norm of the difference of the two distribution functions \(F_n^\text{ind}\) and \(F_n^\text{coll}\), i.e.

\[
\Delta_n = \sup_x |F_n^\text{ind}(x) - F_n^\text{coll}(x)|.
\]

With \(F_n^\text{ind}(x) = P(S_n^\text{ind} \leq x)\), \(\Phi_n^\text{ind}(x) = \Phi((x - \mu_n^\text{ind})/\sigma_n^\text{ind})\); \(F_n^\text{coll}\) and \(\Phi_n^\text{coll}\) analogously, note that

\[
\begin{align*}
F_n^\text{ind}(x) - F_n^\text{coll}(x) &= (F_n^\text{ind}(x) - \Phi_n^\text{ind}(x)) + (\Phi_n^\text{ind}(x) - \Phi_n^\text{coll}(x)) \\
&\quad + (\Phi_n^\text{coll}(x) - F_n^\text{coll}(x)).
\end{align*}
\]

Since

\[
\sqrt{\sum_{i=1}^{n} \left|EX_i - EX_i^3\right|^3} \leq \sqrt{n} \sup_{i \in \mathbb{N}} \left\{EX_i^3\right\} \left(\sum_{i=1}^{n} \text{Var} X_i\right)^{3/2} \leq \sqrt{n} \inf_{i \in \mathbb{N}} \left\{\text{Var} X_i\right\} \quad n \to \infty,
\]

the central limit theorem for \(S_n^\text{ind}\) (cf. Theorem A.1) is applicable to the first term of (2.2) and Proposition A 2 can be applied to the third term of (2.2), the following assertion is valid.

**Proposition 2.1.** Under the assumptions stipulated at the beginning of this section we have

\[
\Delta_n \xrightarrow{n \to \infty} 0 \iff \frac{\sigma_n^\text{ind}}{\sigma_n^\text{coll}} \xrightarrow{n \to \infty} 1 \quad \text{and} \quad \frac{(\mu_n^\text{ind} - \mu_n^\text{coll})/\sigma_n^\text{coll}}{n \to \infty} 0.
\]

**Proof.** Referring to (2.2) it remains to be shown that

\[
\Phi_n^\text{ind}(x) - \Phi_n^\text{coll}(x) \xrightarrow{n \to \infty} 0 \quad \text{uniformly in} \ x
\]

is equivalent to the right side of the assertion.
With \( r_n = \sigma_{n, \text{ind}}/\sigma_{n, \text{coll}} \), we have

\[
\Phi_n^{\text{coll}}(x) = \Phi \left( \frac{x - \mu_{n, \text{coll}}}{\sigma_{n, \text{coll}}} \right) = \Phi \left( \frac{x - \mu_{n, \text{ind}}}{\sigma_{n, \text{ind}}} r_n + \frac{\mu_{n, \text{ind}} - \mu_{n, \text{coll}}}{\sigma_{n, \text{coll}}} \right).
\]

"\( \Leftarrow \)". The assertion is true, as \( \Phi \) is a uniformly continuous function.

"\( \Rightarrow \)". Because \( \Phi \) is a strictly increasing function, \( A_n \to 0 \) supplies

\[
\frac{x - \mu_{n, \text{ind}}}{\sigma_{n, \text{ind}}} - \frac{x - \mu_{n, \text{ind}}}{\sigma_{n, \text{ind}}} r_n - \frac{\mu_{n, \text{ind}} - \mu_{n, \text{coll}}}{\sigma_{n, \text{coll}}} \to 0 \text{ uniformly in } x
\]

\( \frac{x - \mu_{n, \text{ind}}}{\sigma_{n, \text{coll}}} - \frac{x - \mu_{n, \text{ind}}}{\sigma_{n, \text{coll}}} r_n - \frac{\mu_{n, \text{ind}} - \mu_{n, \text{coll}}}{\sigma_{n, \text{coll}}} \to 0 \text{ uniformly in } x
\]

By replacing \( x \) with \( \mu_{n, \text{ind}} \) we get \( (\mu_{n, \text{ind}} - \mu_{n, \text{coll}})/\sigma_{n, \text{coll}} \to 0 \).

Then by replacing \( x \) with \( \mu_{n, \text{ind}} + \sigma_{n, \text{ind}} \), we obtain \( r_n \to 1 \).

Furthermore, for the difference \( \mathcal{S}_n^{\text{coll}}(\alpha) - \mathcal{S}_n^{\text{ind}}(\alpha) \) of the percentile premiums related to \( S_{n, \text{ind}} \) and \( S_{n, \text{coll}} \), we obtain a result which corresponds to Proposition 2.1.

**Proposition 2.2.** Under the assumptions stipulated at the beginning of this section we have

\[
\mathcal{S}_n^{\text{ind}}(\alpha) - \mathcal{S}_n^{\text{coll}}(\alpha) = o\left( \sigma_{n, \text{ind}} \right)
\]

\( \Rightarrow \sigma_{n, \text{ind}}/\sigma_{n, \text{coll}} \to 1 \) and \( (\mu_{n, \text{ind}} - \mu_{n, \text{coll}})/\sigma_{n, \text{coll}} \to 0 \).

**Proof.** For all \( \alpha \in (0, 1) \) the following identity is true:

\[
\frac{\mathcal{S}_n^{\text{coll}}(\alpha) - \mathcal{S}_n^{\text{ind}}(\alpha)}{\sigma_{n, \text{ind}}} = \left( \frac{\mathcal{S}_n^{\text{coll}}(\alpha) - \mu_{n, \text{coll}}}{\sigma_{n, \text{coll}}} \right) \times \frac{\sigma_{n, \text{coll}}}{\sigma_{n, \text{ind}}} - \frac{\mathcal{S}_n^{\text{ind}}(\alpha) - \mu_{n, \text{ind}}}{\sigma_{n, \text{ind}}} + \frac{\mu_{n, \text{coll}} - \mu_{n, \text{ind}}}{\sigma_{n}}.
\]

Then, Lemma A.2 (i) supplies

\[
\frac{\mathcal{S}_n^{\text{coll}}(\alpha) - \mu_{n, \text{coll}}}{\sigma_{n, \text{coll}}} \to \Phi^{-1}(\alpha), \frac{\mathcal{S}_n^{\text{ind}}(\alpha) - \mu_{n, \text{ind}}}{\sigma_{n, \text{ind}}} \to \Phi^{-1}(\alpha).
\]
Thus, the conclusion from the right side to the left side follows directly from the above identity. The converse is also true, since \( \alpha \in (0, 1) \) is arbitrary. So is \( \Phi^{-1}(\alpha) \). As \( (\mu_n - \mu_{\text{coll}})/\sigma_{\text{coll}} \) is independent of \( \alpha \), the argumentation is complete.

The following theorem can be gathered directly from the last two propositions.

**Theorem 2.1.** Under the assumptions above the following assertions are equivalent:

(i) \( \sup_x |F_{n, \text{ind}}(x) - F_{n, \text{coll}}(x)| \xrightarrow{n \to \infty} 0 \),

(ii) \( \frac{\text{Var} S_{n, \text{ind}}}{\text{Var} S_{n, \text{coll}}} \xrightarrow{n \to \infty} 1 \) and \( \frac{(ES_{n, \text{ind}} - ES_{n, \text{coll}})}{\text{Var} S_{n, \text{coll}}} \xrightarrow{n \to \infty} 0 \),

(iii) \( \mathcal{G}_{n, \text{coll}}(\alpha) - \mathcal{G}_{n, \text{ind}}(\alpha) = o\left(\sqrt{\text{Var} S_{n, \text{coll}}(\alpha)}\right), \quad \alpha \in (0, 1), \)

(iv) \( \mathcal{G}_{n, \text{coll}}(\alpha) - \mathcal{G}_{n, \text{ind}}(\alpha) = o\left(\sqrt{\text{Var} S_{n, \text{ind}}(\alpha)}\right), \quad \alpha \in (0, 1). \)

This result can be sharpened to a "bounded version" of (iv), i.e. the difference of the two portfolio premiums even remains bounded under certain conditions.

**Theorem 2.2.** Under the assumptions above the following assertions are equivalent:

(i) \( \sup_x |F_{n, \text{ind}}(x) - F_{n, \text{coll}}(x)| = O\left(1/\sqrt{\text{Var} S_{n, \text{ind}}(\alpha)}\right), \)

(ii) \( \sup_x |F_{n, \text{ind}}(x) - F_{n, \text{coll}}(x)| = O\left(1/\sqrt{\text{Var} S_{n, \text{coll}}(\alpha)}\right), \)

(iii) \( \sqrt{\text{Var} S_{n, \text{ind}}(\alpha)} - \sqrt{\text{Var} S_{n, \text{coll}}(\alpha)} = O(1) \) and \( ES_{n, \text{ind}}(\alpha) - ES_{n, \text{coll}}(\alpha) = O(1), \)

(iv) \( \mathcal{G}_{n, \text{ind}}(\alpha) - \mathcal{G}_{n, \text{coll}}(\alpha) = O(1), \quad \alpha \in (0, 1). \)

**Proof.** Let us sharpen our argumentation with regard to the equation (2.2). Since the Berry-Esséen bounds from Theorem A.1 and Proposition A.3 are applicable to the first and the third term, Theorem 2.1 yields the following equivalence:

\[
\Delta_n = O\left(1/\sqrt{\text{Var} S_{n, \text{ind}}(\alpha)}\right) \iff \tilde{\Delta}_n = O\left(1/\sqrt{\text{Var} S_{n, \text{coll}}(\alpha)}\right)
\]

\[
\iff \Delta_n = O\left(1/\sqrt{\text{Var} S_{n, \text{coll}}(\alpha)}\right),
\]

\[
\iff \Delta_n = O\left(1/\sqrt{\text{Var} S_{n, \text{coll}}(\alpha)}\right).
\]
where $\tilde{\Delta}_n = \sup_x |\Phi_n^{\text{ind}}(x) - \Phi_n^{\text{coll}}(x)|$. According to Lemma A.1, (i) $\iff$ (ii) $\iff$ (iii) has been proved.

With $\gamma_n^{\text{ind}}(\alpha) = (\gamma_n^{\text{ind}}(\alpha) - \mu_n^{\text{ind}})/\sigma_n^{\text{ind}}$, $\gamma_n^{\text{coll}}(\alpha)$ analogously, $\alpha \in (0, 1)$, we have

\begin{equation}
(2.12) \quad \gamma_n^{\text{ind}}(\alpha) - \gamma_n^{\text{coll}}(\alpha) = (\mu_n^{\text{ind}} - \mu_n^{\text{coll}}) + \left(\gamma_n^{\text{ind}}(\alpha) - \gamma_n^{\text{coll}}(\alpha)\right)\sigma_n^{\text{ind}} + \gamma_n^{\text{coll}}(\alpha)\left(\sigma_n^{\text{ind}} - \sigma_n^{\text{coll}}\right).
\end{equation}

Consequently, Lemma A.2 and (i) $\iff$ (iii) supply (i) $\iff$ (iv).

It remains to prove (iv) $\Rightarrow$ (i). For all $\alpha \in (0, 1)$ we have

\begin{equation}
(2.13) \quad O(1) = \gamma_n^{\text{ind}}(\alpha) - \gamma_n^{\text{coll}}(\alpha) = (\mu_n^{\text{ind}} - \mu_n^{\text{coll}}) + \gamma_n^{\text{coll}}(\alpha)\left(\sigma_n^{\text{ind}} - \sigma_n^{\text{coll}}\right) + O(1),
\end{equation}

again in respect of Lemma A.2 (ii).

By choosing $\alpha_1 \neq \alpha_2$, we can conclude that

\begin{equation}
(2.14) \quad O(1) = (\gamma_n^{\text{coll}}(\alpha_1) - \gamma_n^{\text{coll}}(\alpha_2))\left(\sigma_n^{\text{ind}} - \sigma_n^{\text{coll}}\right).
\end{equation}

Again using $\gamma_n^{\text{coll}}(\alpha) \xrightarrow[n \to \infty]{} \Phi^{-1}(\alpha)$, we have

\begin{equation}
(2.15) \quad \sigma_n^{\text{ind}} - \sigma_n^{\text{coll}} = O(1), \text{ hence } \mu_n^{\text{ind}} - \mu_n^{\text{coll}} = O(1).
\end{equation}

Thus (iv) $\Rightarrow$ (iii) has been proved, and therefore (iv) $\Rightarrow$ (i) since (iii) $\iff$ (i).

It should be noted that the statements of Theorem 2.2 (i) and (ii) can be specified by deriving explicit constants in the $O$-estimates from the proofs.

Each assertion of Theorem 2.2 implies the corresponding one of Theorem 2.1, but the converse is false, which becomes obvious in Example 2.1 at the end of this section.

For the rest of the section we consider a concept of a portfolio growth described by an appropriately chosen sequence of homogeneous collective risk models. Therefore at first we have to formulate some additional requirements to the underlying risk portfolio. These concern the mixture ratio of its distinct risk classes.

**Assumption 2.1.** In addition to the previous assumptions the sequence of risks and its random claims amounts $(X_i), i \in \mathbb{N}$, are required to fulfill the following: the set of random variables $(X_i, i \in \mathbb{N})$ consists of $K$ distinct risk classes $\mathcal{H}_k, k = 1, \ldots, K$; thus $\{X_i, i \in \mathbb{N}\} = \bigcup_{k=1}^{K} \mathcal{H}_k$. Each class is represented by a distribution function $F(k)$. $X_i \in \mathcal{H}_k$ means that $X_i \sim F_k$ and $F_k = F(k)$. Correspondingly $q(k) \in (0, 1]$ denotes the representative of $q_i$, if $X_i \in \mathcal{H}_k$, where $q_i = P(X_i > 0)$ is related to $X_i$. It is assumed that the mixture ratio of $n$ risks satisfies the stability criteria below: For each $n \in \mathbb{N}$ define for all $k = 1, \ldots, K$
AN APPROPRIATE WAY TO SWITCH FROM THE INDIVIDUAL RISK MODEL

an integer \( n_k = \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq R_k\}} \); thus \( n = \sum_{k=1}^{K} n_k \). Assume, for each \( k \) there exists a number \( c_k \in (0, 1] \) independent of \( n \), which fulfills \( n_k - n c_k = O(1) \) as \( n \) tends to infinity.

Continuing to use the previous notations, we also introduce some new. Let

\[
q^{(n)} = \frac{1}{n} \sum_{i=1}^{n} q_i \quad \text{and} \quad \bar{q} = \sum_{k=1}^{K} c_k q(k).
\]

This means that \( \bar{q} \in (0, 1] \), since \( c_k \leq 1 \) for all \( k \) and \( c_k \neq 0 \) for at least one \( k \). Furthermore, let

\[
G_n(x) = \sum_{i=1}^{n} \frac{q_i}{nq^{(n)}} G_i(x) \quad \text{with} \quad G_i(x) = \frac{F_i(x) - (1 - q_i)}{q_i}, \quad x \geq 0,
\]

that is

\[
G_n(x) = \sum_{k=1}^{K} \frac{n_k}{n} \frac{q(k)}{q^{(n)}} G(k)(x).
\]

Moreover, let

\[
\bar{G}(x) = \sum_{k=1}^{K} c_k \frac{q(k)}{\bar{q}} G(k)(x) \quad \text{with} \quad G(k)(x) = \frac{F(k)(x) - (1 - q(k))}{q(k)}, \quad x \geq 0.
\]

In the following, the claims number in the collective risk models specified in Section 1 (cf. Model 1.1-1.3) is denoted by \( N_n \) and \( \bar{N}_n \), referring to the parameters \( q^{(n)} \) and \( \bar{q} \). We write \( Z_1, \bar{Z}_1 \) resp. for the collective single claims amount variable, where \( Z_1 \not\asymp G_n, \bar{Z}_1 \not\asymp \bar{G} \) resp.; thus

\[
S_n^{\text{coll}} = \sum_{i=1}^{N_n} Z_i, Z_i \not\asymp Z_1
\]

and

\[
S_n^{\text{coll}} = \sum_{i=1}^{\bar{N}_n} \bar{Z}_i, \bar{Z}_i \not\asymp \bar{Z}_1.
\]

Finally, in the collective risk model \( S_n^{\text{coll}} \) we denote by \( \mu_n^{\text{coll}} \) and \( \sigma_n^{\text{coll}} \) the mean value and the standard deviation resp., in line with \( \mu_n^{\text{ind}} \) and \( \sigma_n^{\text{ind}} \) above.

In this framework the portfolio growth is defined by the corresponding sequence of homogeneous collective risk models \( (S_n^{\text{coll}})_{n \in \mathbb{N}} \) which fulfill Assumption 2.1 and to which Theorem 2.1 and Theorem 2.2 can be applied. Thus, we can prove the following:
Proposition 2.3. Under the assumptions of Section 2 in all the three collective risk models described above (cf. Model 1.1-1.3) the distribution function $F_n^{\text{ind}}$ of $S_n^{\text{ind}}$ and $\bar{F}_n^{\text{coll}}$ of $\bar{S}_n^{\text{coll}}$ resp., fulfill a Berry-Esséen bound, i.e.

(i) $\sup_x |F_n^{\text{ind}}(x) - \Phi_n^{\text{ind}}(x)| = O(1/\sqrt{\text{Var } S_n^{\text{ind}}})$, 

(ii) $\sup_x |\bar{F}_n^{\text{coll}}(x) - \bar{\Phi}_n^{\text{coll}}(x)| = O(1/\sqrt{\text{Var } \bar{S}_n^{\text{coll}}})$,

where $\bar{\Phi}_n^{\text{coll}}(x) = \Phi \left( \frac{x - \bar{\mu}_n^{\text{coll}}}{\bar{\sigma}_n^{\text{coll}}} \right)$.

Proof:

(i) As $E|X_i - EX_i|^3 \leq EX_i^3 + (EX_i)^3$, we have from the assumptions

$max \{ E|X_i - EX_i|^3 \} < \infty$. Assumption 1.1 supplies $\text{Var } S_n^{\text{ind}} \geq na$ and thus we conclude

$$\frac{\sum_{i=1}^n E|X_i - EX_i|^3}{\sum_{i=1}^n \text{Var } X_i} \leq \frac{\max \{ E|X_i - EX_i|^3 \}}{a} < \infty. \quad (2.22)$$

Consequently, from the Berry-Esséen bound for non-identically distributed random variables (cf Theorem A.1) we obtain

$$\sup_x |F_n^{\text{ind}}(x) - \Phi_n^{\text{ind}}(x)| \leq \frac{6}{\sigma_n^{\text{ind}}} \sum_{i=1}^n \text{Var } X_i \quad (2.23)$$

(ii) Since $E\tilde{N}_n^3 < \infty$ and $E\tilde{Z}_1^3 < \infty$, the Berry-Esséen bound for random sums according to Definition 2.1 can be applied (cf. Proposition A.3). When this is done,

$$\text{Var } \bar{S}_n^{\text{coll}} = n(\bar{q} \text{ Var } \bar{Z}_1 + \text{Var } \bar{Z}_1 (E\bar{Z}_1)^2) \quad (2.24)$$

must be taken into consideration.

Consequently $S_n^{\text{ind}}$ and $\bar{S}_n^{\text{coll}}$ fulfill in particular the central limit theorem with the standard normalization and the law of large numbers.

A result such as that in Proposition 2.3 valid for $\bar{S}_n^{\text{coll}}$ and $\bar{\Phi}_n^{\text{coll}}$ resp., $\bar{\Phi}_n^{\text{coll}}(x) = \Phi((x - \bar{\mu}_n^{\text{coll}})/\bar{\sigma}_n^{\text{coll}})$, cannot be directly deduced from the Berry-
Essén bound in Proposition A.3. In fact, the distribution function $G_n$ of $Z_1$ depends on $n$. Thus $G_n$ and the distribution function of $N_n$ do not satisfy the assumptions required in Definition 2.1.

Nevertheless, taking into account the stability of the mixture ratio given in Assumption 2.1 we have

$$E \tilde{S}_n^{\text{coll}} = E \tilde{S}_n^{\text{coll}} + O(1)$$

and

$$\sqrt{\text{Var} \tilde{S}_n^{\text{coll}}} = \sqrt{\text{Var} \tilde{S}_n^{\text{coll}}} + O(1/\sqrt{\text{Var} \tilde{S}_n^{\text{coll}}}).$$

Together with

$$\sup_{x} |\hat{F}_n^{\text{coll}}(x) - \hat{F}_n^{\text{coll}}(x)| = O(1/\sqrt{\text{Var} \tilde{S}_n^{\text{coll}}}),$$

and the Berry-Essèen bound for $\tilde{S}_n^{\text{coll}}$ and $\hat{F}_n^{\text{coll}}$, i.e.

$$\sup_{x} |\hat{F}_n^{\text{coll}}(x) - \hat{F}_n^{\text{coll}}(x)| = O(1/\sqrt{\text{Var} \tilde{S}_n^{\text{coll}}}),$$

we get ultimately the validity of Theorem 2.1 and Theorem 2.2 even for $\tilde{S}_n^{\text{coll}}$ straightforward from the identity

$$F_n^{\text{ind}}(x) - \tilde{F}_n^{\text{coll}}(x) = (F_n^{\text{ind}}(x) - \hat{F}_n^{\text{coll}}(x)) + (\hat{F}_n^{\text{coll}}(x) - \tilde{F}_n^{\text{coll}}(x)).$$

Thus we have proved

**Corollary 2.1.** Under the assumptions of Section 2 these assertions are equivalent:

(i) $\sup_{x} |F_n^{\text{ind}}(x) - \tilde{F}_n^{\text{coll}}(x)| \xrightarrow{n \to \infty} 0$,

(ii) $\text{Var} S_n^{\text{ind}}/\text{Var} \tilde{S}_n^{\text{coll}} \xrightarrow{n \to \infty} 1$,

(iii) $\tilde{S}_n^{\text{coll}}(\alpha) - S_n^{\text{ind}}(\alpha) = o(\sqrt{\text{Var} \tilde{S}_n^{\text{coll}}})$, $\alpha \in (0, 1)$,

(iv) $\tilde{S}_n^{\text{coll}}(\alpha) - S_n^{\text{ind}}(\alpha) = o(\sqrt{\text{Var} S_n^{\text{ind}}})$, $\alpha \in (0, 1)$.

The following result represents a quantitative sharpening of Corollary 2.1.

**Corollary 2.2.** Under the assumptions of Section 2 these assertions are equivalent:

(i) $\sup_{x} |F_n^{\text{ind}}(x) - \tilde{F}_n^{\text{coll}}(x)| = O(1/\sqrt{\text{Var} S_n^{\text{ind}}})$,

(ii) $\sup_{x} |F_n^{\text{ind}}(x) - \tilde{F}_n^{\text{coll}}(x)| = O(1/\sqrt{\text{Var} \tilde{S}_n^{\text{coll}}})$. 
(iii) \( \sqrt{\text{Var} \, S_n^\text{ind}} - \sqrt{\text{Var} \, S_n^\text{coll}} = O(1) \),

(iv) \( \mathcal{S}_n^\text{ind}(\alpha) - \mathcal{S}_n^\text{coll}(\alpha) = O(1), \quad \alpha \in (0, 1) \).

Obviously each condition of Corollary 2.1 follows from the corresponding one of Corollary 2.2. The inverse conclusion is wrong as shown below by Example 2.1. Note that from the proofs given above explicit constants can be derived to replace the \( O \)-constants in Corollary 2.2 (i) and (ii).

Accurate premium calculation or their equivalent, precise approximation of the distribution function in the individual model, depends mainly on well variance fitted collective risk models. The previous collective risk models do not achieve that as proved for Model 1.1-1.3. In the next section we shall look at modifications of these models, which improve the variance fit.

**Example 2.1.** Let us consider a sequence of positive random variables \( (X_i)_{i \in \mathbb{N}} \) with distribution functions

\[
F_i(x) = (1 - q_i) + q_i F(x), \quad q_i \in (0, 1), \quad F(x) = 1 - e^{-x}, \quad x \geq 0.
\]

Therefore,

\[
P(X_i > 0) = q_i, \quad EX_i = q_i, \quad \text{Var} \, X_i = q_i(2 - q_i),
\]

For each \( n \in \mathbb{N} \), let

\[
a_n = \sum_{i=1}^{n} q_i, \quad \text{and} \quad b_n = \sum_{i=1}^{n} q_i^2.
\]

For \( S_n^\text{ind} = \sum_{i=1}^{n} X_i \), this implies that

\[
ES_n^\text{ind} = a_n \quad \text{and} \quad \text{Var} \, S_n^\text{ind} = 2a_n - b_n.
\]

We construct the collective risk model \( S_n^\text{coll} = \sum_{i=1}^{N_n} Z_i \) corresponding to \( S_n^\text{ind} \) in the same way as described in Section 1, by means of the following equation:

\[
G(x) = \sum_{i=1}^{n} \frac{q_i}{nq^{(n)}} G_i(x) \quad \text{with} \quad x \geq 0, \quad q^{(n)} = \frac{1}{n} \sum_{i=1}^{n} q_i,
\]

where \( G_i(x) = (F_i(x) - (1 - q_i))/q_i = F(x) \), \( x \geq 0 \). Thus, we have \( G(x) = F(x) \). Assuming \( Z_i \nless G \), we obtain \( EZ_1 = \text{Var} \, Z_1 = 1 \).

Moreover, we stipulate that \( N_n \) is distributed as \( \beta(n, q^{(n)}) \) Hence (cf. Model 1.1),

\[
ES_n^\text{coll} = a_n \quad \text{and} \quad \text{Var} \, S_n^\text{coll} = 2a_n - \frac{a_n^2}{n}.
\]
Now, for all $\alpha \in (0, 1/2)$, with a suitable choice of $q_i \in (0, 1)$, $i \in \mathbb{N}$,

\begin{equation}
(2.36) \quad a_n \sim n - n^{1-\alpha} \quad \text{and} \quad b_n \sim n - 2n^{1-\alpha} + \frac{(1-\alpha)^2}{1-2\alpha} n^{1-2\alpha}
\end{equation}

is fulfilled. For instance, $q_i = 1 - (1-\alpha)/i$ is appropriate.

For proof of this, note that

\begin{equation}
(2.37) \quad \sum_{i=1}^{n} (1-\beta) i^{-\beta} \sim \int_1^n (1-\beta) x^{-\beta} \, dx \sim n^{1-\beta}, \quad \beta \in (0, 1).
\end{equation}

All these definitions supply

\begin{equation}
(2.38) \quad \frac{\text{Var } S_{\text{coll}}}{\text{Var } S_{\text{ind}}} = \frac{2 - \frac{a_n}{n}}{2 - \frac{b_n}{a_n}} \quad \text{as } n \to \infty
\end{equation}

since $a_n/n \xrightarrow[n \to \infty]{} 1$ and $b_n/a_n \xrightarrow[n \to \infty]{} 1$.

However,

\begin{equation}
(2.39) \quad \sqrt{\text{Var } S_{\text{ind}}} - \sqrt{\text{Var } S_{\text{coll}}} = \left(\frac{\sqrt{a_n^3} - \frac{b_n}{\sqrt{a_n}}}{n} \right) \left(2 - \frac{a_n}{n}\right)^{-1} \left(1 + \frac{\text{Var } S_{\text{coll}}}{\text{Var } S_{\text{ind}}}\right)^{-1}
\end{equation}

\begin{align*}
&\sim \sqrt{n^{1-4\alpha}} \left(1 - \frac{(1-\alpha)^2}{1-2\alpha}\right) \frac{1}{2} \\
&\quad \left\{ \begin{array}{l}
-\infty \quad \text{if} \quad \alpha \in (0, \frac{1}{4}), \\
\frac{1}{16} \quad \text{if} \quad \alpha = \frac{1}{4}, \\
0 \quad \text{if} \quad \alpha \in (\frac{1}{4}, \frac{1}{2}),
\end{array} \right.
\end{align*}

since

\begin{equation*}
\frac{a_n^2 - nb_n}{n \sqrt{a_n}} \sim \sqrt{n^{1-4\alpha}}.
\end{equation*}

3. Modified collective risk models with variance adjusted to that of the underlying individual risk model

In this section the assumptions of Section 2 are stipulated. The notations used below are the same as stated previously. We drop the index $n$ because there is no misunderstanding possible.
Corollary 2.1 and Corollary 2.2 proved above suggest to adjust not only $E_{\text{ind}}$ but also $\text{Var } S_{\text{ind}} = \text{Var } S_{\text{coll}}$ for all $n$. In this case condition (iii) and therefore all conditions of Corollary 2.2 are valid. The classical approaches, which derive collective risk models from an individual one, do not fulfill the two conditions (equality of the mean values and the variances) simultaneously in general (cf. Model 1.1, 1.2, 1.3). JEWELL and SUNDT (1981) deal with this problem in their paper too. They discuss two different modifications of the compound binomial approximation (Model 1.1) by using modified counting distributions. In addition SUNDT (1985) studies an approach with an "average" collective claims amount distribution.

We shall now derive a similar modification of our in Section 1 constructed collective risk models which ensures the equality of their first two moments with those in the individual risk models given. In view of practical applications (i.e. numerical computation by the Panjer algorithm) we presume the range of the collective claims amounts to be discrete and arithmetic. For the purpose of modeling a new collective claims distribution function we define a random variable $Z_{\text{mod}}$ with discrete range \{\(k\gamma|k\in\mathbb{N}\}, \gamma > 0\) fixed, by setting

\[
P(Z_{\text{mod}} = k\gamma) = g(k), \quad k \in \mathbb{N},
\]

where (cf. Remark 1.4)

\[
g(k) = \sum_{i=1}^{n} \frac{q_i}{nq} g_i(k), \quad g_i(k) = \frac{P(X_i = k)}{q}, \quad q = \frac{1}{n} \sum_{i=1}^{n} q_i, \quad q_i = P(X_i > 0).
\]

\(Z_{\text{mod}}\) differs from \(Z_1\) as constructed in the models provided above only by a simple transformation of the range. Obviously we have (cf. Remark 1.4)

\[
E(Z_{\text{mod}}) = \gamma E_{Z_1} = \frac{\gamma}{nq} \sum_{i=1}^{n} EX_i
\]

and

\[
E(Z_{\text{mod}})^2 = \gamma^2 E_{Z_1}^2 = \frac{\gamma^2}{nq} \sum_{i=1}^{n} EX_i^2.
\]

If one considers $S_{\text{coll}} = \sum_{i=1}^{N} Z_{\text{mod}}$, the basic requirement $E_{\text{ind}} = E_{\text{coll}}$ results in $EN = nq/\gamma$, because $E_{\text{coll}} = ENE_{Z_{\text{mod}}}$. The following is also valid in this case.

\[
\text{Var } S_{\text{coll}} = ENE(Z_{\text{mod}})^2 + (\text{Var } N - EN)(E_{Z_{\text{mod}}})^2
\]

\[
= \frac{\gamma^2}{nq} \sum_{i=1}^{n} EX_i^2 + \left(\frac{\gamma}{nq}\right)^2 (\text{Var } N - EN)(E_{\text{ind}})^2
\]
AN APPROPRIATE WAY TO SWITCH FROM THE INDIVIDUAL RISK MODEL

\[\begin{align*}
&= \gamma \sum_{i=1}^{n} EX_i^2 + \frac{1}{EN} \left( \frac{\text{Var } N}{EN} - 1 \right) (ES_{\text{ind}})^2 \\
&= \text{Var } S_{\text{ind}} + \Delta(\gamma),
\end{align*}\]

where

\[(3.6) \quad \Delta(\gamma) = (\gamma - 1) \sum_{i=1}^{n} EX_i^2 + \sum_{i=1}^{n} (EX_i)^2 + \frac{1}{EN} \left( \frac{\text{Var } N}{EN} - 1 \right) (ES_{\text{ind}})^2.\]

Therefore the following equivalence holds:

\[(3.7) \quad \text{Var } S_{\text{coll}} = \text{Var } S_{\text{ind}} \iff \Delta(\gamma) = 0.\]

\(\Delta(\gamma) = 0\) cannot be fulfilled with \(\gamma = 1\) (cf. Model 1.1-1.3), i.e. the original range of the collective claims variables \(Z_i, i \in \mathbb{N}\), must be transformed.

**Model 3.1.** The modified natural approximation (modified compound binomial approximation; cf JEWELL and SUNDT (1981))

Let

\[(3.8) \quad \gamma = 1 - \frac{\sum_{i=1}^{n} (EX_i)^2 - n^{-1} (ES_{\text{ind}})^2}{\sum_{i=1}^{n} EX_i^2} \quad \text{and} \quad N \sim \beta(n, q/\gamma).\]

Then \(\gamma \in (0, 1]\), since \((ES_{\text{ind}})^2 \leq n \sum_{i=1}^{n} (EX_i)^2\) and

\[\gamma = \left( \sum_{i=1}^{n} \text{Var } X_i + n^{-1} (ES_{\text{ind}})^2 \right) / \sum_{i=1}^{n} EX_i^2.\]

If \(q/\gamma \geq 1\) we modify the parameters \(n, q, \gamma\), see below. Obviously we have \(\Delta(\gamma) = 0\) and, hence, \(\text{Var } S_{\text{ind}} = \text{Var } S_{\text{coll}}\). However, with this stipulation

\(EN = nq/\gamma\) differs from \(E \sum_{i=1}^{n} 1_{\{X_i > 0\}} = nq\).

A simple manipulation of the parameters \(n, q, \gamma\) facilitates obtaining in addition \(EN = E \sum_{i=1}^{n} 1_{\{X_i > 0\}}\). For this purpose, we set \(N \sim \beta(n', q'/\gamma')\) and adjust \(n', q', \gamma'\) accordingly. The condition \(ES_{\text{ind}} = ES_{\text{coll}}\) implies that \(n'q' = nq\); consequently \(q' = \frac{n}{n'} q\).
From the equivalence \( \text{Var } S^{\text{ind}} = \text{Var } S^{\text{coll}} \Leftrightarrow \mathcal{A}(\gamma) = 0 \) we deduce

\[
(3.9) \quad \gamma' = 1 - \frac{\sum_{i=1}^{n} (EX_i)^2 - (n')^{-1} (ES^{\text{ind}})^2}{\sum_{i=1}^{n} EX_i^2}
\]

On the other hand \( \mathcal{E} N = E \sum_{i=1}^{n} l_{(X_i > 0)} \Leftrightarrow \gamma' = 1. \)

This leads to the choice

\[
(3.10) \quad n' = \left[ \frac{(ES^{\text{ind}})^2}{\sum_{i=1}^{n} (EX_i)^2} \right], \quad [x] = \text{greatest integer } m \text{ with } m \leq x,
\]

from that we have \( \gamma' \approx 1 (\gamma' \geq 1) \) and \( n' \leq n \) as

\[
(ES^{\text{ind}})^2 \leq n \sum_{i=1}^{n} (EX_i)^2.
\]

However, note that possibly

\[
\frac{q'}{\gamma'} < 1, \quad \text{where} \quad q' = \frac{n}{n'},
\]

is no longer valid with such a choice of \( n' \). Clearly, increasing \( n' \) ultimately guarantees \( q'/\gamma' < 1 \). Note that \( n' = n \Leftrightarrow q' = q \). However, \( \gamma' \) is then more and more different from 1.

To show that possibly \( q'/\gamma' \geq 1 \), let \( n \geq 2 \). Choose \( X_1, \ldots, X_n \) such that \( q_i = q_0 > 1/2 \) for each \( i = 1, \ldots, n \) and \( EX_2 = \ldots = EX_n = 1 \).

The ratio

\[
(3.11) \quad \frac{(ES^{\text{ind}})^2}{\sum_{i=1}^{n} (EX_i)^2} = \frac{(1 + (n-1) (EX_2/EX_1))^2}{1 + (n-1) (EX_2/EX_1)^2}
\]

takes on values near \( n \) for \( EX_1 \) close to 1, and tends to 1 if \( EX_1 \to \infty \).

Therefore, we can choose \( EX_1 \) such that

\[
(3.12) \quad \frac{(ES^{\text{ind}})^2}{\sum_{i=1}^{n} (EX_i)^2} = \frac{n}{2}.
\]
If \( n \) is even we have \( n' = n/2 \) and \( \gamma' = 1 \), however.

\[
\frac{q'}{\gamma'} = \frac{nq}{n' \gamma'} = \frac{nq}{n/2} > 1.
\]

**Model 3.2. The modified compound Poisson approximation.**

Let

\[
\frac{\sum_{i=1}^{n} (EX_i)^2}{\sum_{i=1}^{n} EX_i^2} = \frac{\text{Var} S^{\text{ind}}}{\sum_{i=1}^{n} EX_i^2} \quad \text{and} \quad N \sim \pi \left( \frac{nq}{\gamma} \right).
\]

In this case we have \( \Delta (\gamma) = 0 \), that is \( \text{Var} S^{\text{ind}} = \text{Var} S^{\text{coll}} \).

With this choice of parameters, \( EN = nq/\gamma \) differs from \( E \sum_{i=1}^{n} 1_{\{X_i > 0\}} = nq \).

The harmonization of these two quantities fails in this case, because we can select only two parameters.

**Model 3.3. The modified compound negative binomial approximation.**

Let

\[
\gamma = \frac{\sum_{i=1}^{n} \text{Var} X_i}{\sum_{i=1}^{n} EX_i^2} \left( \frac{1}{n} \frac{(ES^{\text{ind}})^2}{\sum_{i=1}^{n} \text{Var} X_i} \right) \quad \text{and} \quad N \sim \mathcal{G} \left( n, \frac{1}{1 + q/\gamma} \right)
\]

Obviously \( \gamma < \left( \frac{\sum_{i=1}^{n} \text{Var} X_i}{\sum_{i=1}^{n} EX_i^2} \right) < 1 \).

Hence, \( \gamma = 1 \) is impossible, that is equivalent to the assertion, that \( EN = nq/\gamma \) differs from \( E \sum_{i=1}^{n} 1_{\{X_i > 0\}} = nq \). However, we have achieved \( \Delta (\gamma) = 0 \).

**Application 3.1** In order to verify, whether our modified collective risk models lead to good results also in the case of small portfolios, we have calculated the stop loss premium (without any loading) by means of the distribution functions of the discussed standard and modified collective risk models (Model 1.1-1.3, Model 3.1-3.3 resp.). The calculations are based on the Gerber-Portfolio (cf
GERBER (1979), p. 53) and the 100-fold Gerber-Portfolio. Comparison was made between the different models by the relative error, that is the absolute error in percentage of the "true" risk premium, which was exactly calculated by convolution. It can be easily seen from the figures below, that the modified collective risk models lead almost always to smaller errors than in the case of the standard approximations. Obviously the absolute relative error depends on the underlying priority, i.e. the stop loss point.

### Gerber — Portfolio of 31 Policies

<table>
<thead>
<tr>
<th>$p_i$</th>
<th>Amount at Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0.03</td>
<td>2</td>
</tr>
<tr>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>0.06</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0.06</td>
</tr>
</tbody>
</table>

### Gerber — Portfolio of 3100 Policies

<table>
<thead>
<tr>
<th>$p_i$</th>
<th>Amount at Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0.03</td>
<td>200</td>
</tr>
<tr>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>0.06</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0.06</td>
</tr>
</tbody>
</table>

#### Error of Stop-Loss Premiums
for the Gerber-Portfolio
related to the Standard and Modified Collective Models

- Negative-binomial
- Modified negative-binomial
- Poisson
- Modified Poisson
- Binomial
- Modified binomial

Aggregate claim of the portfolio
expected value: 4.49
variance: 15.3883

Stop-loss point
AN APPROPRIATE WAY TO SWITCH FROM THE INDIVIDUAL RISK MODEL

Error of Stop-Loss Premiums for the 100-fold Gerber-Portfolio related to the Standard and Modified Collective Models

![Graph showing the error of stop-loss premiums for different security levels]

**TABLE 1**

**Standard Collective Models for the Gerber Portfolio**

**Error of the Stop Loss Premium**

(WITHOUT ANY LOADING)

<table>
<thead>
<tr>
<th>Security Level of Percentile Premium</th>
<th>Stop Loss Point</th>
<th>Stop Loss Premium in the Ind Mod</th>
<th>Error in % of the Stop Loss Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Binomial</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>1.776</td>
<td>0.16</td>
</tr>
<tr>
<td>60</td>
<td>5</td>
<td>1.340</td>
<td>0.37</td>
</tr>
<tr>
<td>70</td>
<td>6</td>
<td>1.001</td>
<td>0.54</td>
</tr>
<tr>
<td>80</td>
<td>8</td>
<td>0.515</td>
<td>1.25</td>
</tr>
<tr>
<td>90</td>
<td>10</td>
<td>0.251</td>
<td>2.35</td>
</tr>
<tr>
<td>95</td>
<td>12</td>
<td>0.113</td>
<td>4.28</td>
</tr>
<tr>
<td>99</td>
<td>16</td>
<td>0.019</td>
<td>9.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Poisson</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>492.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1725.27</td>
</tr>
</tbody>
</table>

*Expected Value: 448.00, Variance: 1536.63*
### TABLE 2
**Modified Collective Models for the Gerber Portfolio**
**Error of the Stop Loss Premium**
(without any loading)

<table>
<thead>
<tr>
<th>Security Level of Percentile Premium</th>
<th>Stop Loss Point</th>
<th>Stop Loss Premium in the Ind Mod</th>
<th>Error in % of the Stop Loss Premium</th>
<th>Error in % of the Stop Loss Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Binomial</td>
<td>Poisson</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>1.776</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.340</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.001</td>
<td>0.12</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.515</td>
<td>0.06</td>
<td>1.85</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.251</td>
<td>0.44</td>
<td>3.71</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.113</td>
<td>1.42</td>
<td>6.81</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>0.019</td>
<td>4.31</td>
<td>15.89</td>
</tr>
</tbody>
</table>

### TABLE 3
**Standard Collective Models for the 100-fold Gerber Portfolio**
**Error of the Stop Loss Premium**
(without any loading)

<table>
<thead>
<tr>
<th>Security Level of Percentile Premium</th>
<th>Stop Loss Point</th>
<th>Stop Loss Premium in the Ind Mod</th>
<th>Error in % of the Stop Loss Premium</th>
<th>Error in % of the Stop Loss Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Binomial</td>
<td>Poisson</td>
</tr>
<tr>
<td>50</td>
<td>448</td>
<td>16.10</td>
<td>0.44</td>
<td>2.46</td>
</tr>
<tr>
<td></td>
<td>458</td>
<td>11.57</td>
<td>0.61</td>
<td>3.38</td>
</tr>
<tr>
<td></td>
<td>469</td>
<td>7.70</td>
<td>0.84</td>
<td>4.66</td>
</tr>
<tr>
<td></td>
<td>482</td>
<td>4.49</td>
<td>1.19</td>
<td>6.56</td>
</tr>
<tr>
<td></td>
<td>499</td>
<td>1.99</td>
<td>1.80</td>
<td>9.81</td>
</tr>
<tr>
<td></td>
<td>514</td>
<td>0.88</td>
<td>2.47</td>
<td>13.48</td>
</tr>
<tr>
<td></td>
<td>543</td>
<td>0.14</td>
<td>4.22</td>
<td>23.19</td>
</tr>
</tbody>
</table>

### TABLE 4
**Modified Collective Models for the 100-fold Gerber Portfolio**
**Error of the Stop Loss Premium**
(without any loading)

<table>
<thead>
<tr>
<th>Security Level of Percentile Premium</th>
<th>Stop Loss Point</th>
<th>Stop Loss Premium in the Ind Mod</th>
<th>Error in % of the Stop Loss Premium</th>
<th>Error in % of the Stop Loss Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Binomial</td>
<td>Poisson</td>
</tr>
<tr>
<td>50</td>
<td>448</td>
<td>16.10</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>458</td>
<td>11.57</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>469</td>
<td>7.70</td>
<td>0.02</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>482</td>
<td>4.49</td>
<td>0.04</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>499</td>
<td>1.99</td>
<td>0.09</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>514</td>
<td>0.88</td>
<td>0.16</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>543</td>
<td>0.14</td>
<td>0.38</td>
<td>1.51</td>
</tr>
</tbody>
</table>
We start with a selection of results which are contained in a paper written by von Chossy, R. and G. Rappl (1983). Let

(A.1) \[
S_n^\text{coll} \sum_{i=1}^{N_n} Y_i, \quad n \in \mathbb{N},
\]

be a random sum where \((Y_i)_{i \in \mathbb{N}}\) is a sequence of real-valued, independent, identically distributed random variables, and \((N_n)_{n \in \mathbb{N}}\) is a sequence of integer-valued random variables, \(N_n \geq 0\). \((N_n)_{n \in \mathbb{N}}\) and \((Y_i)_{i \in \mathbb{N}}\) are supposed to be independent for each \(n \in \mathbb{N}\). Furthermore, the second moments of \(Y_i\) and \(N_n\) may exist in the proper sense (cf. Definition 1.2).

Von Chossy, R. and G. Rappl (1983, p. 252) proved that, in certain cases it is possible to represent random sums as deterministic sums.

**Proposition A.1.** Let \(K\) be a distribution function on \(\mathbb{N}_0\) such that for each \(n \in \mathbb{N}\)

(A.2) \[
N_n \leq K^n.
\]

Further, let

(A.3) \[
\bar{F} = \int_{\mathbb{N}_0} F^k K(dk), \quad Y_1 \leq F.
\]

Then there exists a sequence \((\bar{Y}_1, \bar{Y}_2, \ldots)\) of independent and, according to \(\bar{F}\), identically distributed random variables with

(A.4) \[
S_n^\text{coll} \leq \sum_{i=1}^{n} \bar{Y}_i \quad \text{(i.e.} \quad \bar{Y}_i \leq \sum_{i=1}^{N_n} Y_i) \quad \text{for all} \quad n \in \mathbb{N}.
\]

**Definition A.1.** The central limit theorem (with standard normalization) is said to be valid for a sequence of random variable \((S_n)_{n \in \mathbb{N}}\) if \((S_n - ES_n)/\sqrt{\text{Var} S_n}\) converges in distribution to a standard normal distributed random variable as \(n \to \infty\), i.e. \(|F_n(x) - \Phi_n(x)| \to 0\) uniformly in \(x\) with

\[
F_n(x) = P(S_n \leq x) \quad \text{and} \quad \Phi_n(x) = \Phi((x - ES_n)/\sqrt{\text{Var} S_n}), \quad \Phi \text{ the standard normal distribution function}.
\]

**Proposition A.2.** Under the assumptions of Proposition A.1 with $\text{Var} \hat{Y}_1 > 0$, the central limit theorem is valid for the sequence $(S_{\text{coll}} - E S_{\text{coll}})/\sqrt{\text{Var} S_{\text{coll}}}$. Using the standard Berry-Esseen inequality (cf. FELLER (1971), p. 542), both authors proved, in addition, a Berry-Esseen bound for special random sums.

**Proposition A.3.** Let the assumptions of Proposition A.1 be fulfilled; further, let $\text{Var} \hat{Y}_1 > 0$, $E|Y_1|^3 < \infty$, $E N_n^3 < \infty$. Denote by $F_{n}^{\text{coll}}$ the distribution function of $S_{\text{coll}}$ and put

$$
(A.5) \quad \phi_{n}^{\text{coll}}(x) = \Phi \left( \frac{(x - E S_{n}^{\text{coll}})}{\sqrt{\text{Var} S_{n}^{\text{coll}}}} \right),
$$

$\Phi$ the standard normal distribution function. Then, for all $n \in \mathbb{N}$, we have

$$
(A.6) \quad \sup_{x} |F_{n}^{\text{coll}}(x) - \phi_{n}^{\text{coll}}(x)| \leq \frac{3}{\sqrt{n}} \frac{E|\hat{Y}_1 - E \hat{Y}_1|^3}{(\text{Var} \hat{Y}_1)^{3/2}}
$$

Furthermore, it holds

$$
(A.7) \quad \text{Var} \hat{Y}_1 = \mu_1 \text{Var} Y_1 + \sigma^2 (E Y_1)^2,
$$

where

$$
(A.8) \quad \mu_1 = \int_{N_0} k K(dk), \quad \sigma^2 = \int_{N_0} (k - \mu_1)^2 K(dk),
$$

and

$$
(A.9) \quad E|\hat{Y}_1 - E \hat{Y}_1|^3 \leq 4 \left( \rho_3 - 3 \rho_2 \rho_1 + 2 \rho_1^3 \right) \mu_1 + (3 \rho_2 \rho_1 - 3 \rho_1^3) \mu_2
$$

$$
+ \rho_1^3 |E Y_1|^3 \int_{N_0} |k - \mu_1|^3 K(dk),
$$

where

$$
(A.10) \quad \mu_i = \int_{N_0} k^i K(dk), \quad \rho_i = \int_{\mathbb{R}} |x - E Y_1|^i F(dx), \quad i = 1, 2, 3.
$$

**Remark A.1.** If $N_n$ is Poisson distributed with parameter $n \lambda$, $\lambda > 0$, Proposition A.1 can be applied. The same is true in the case of the binomial distribution with parameters $(n, q)$, $q \in (0, 1)$, and in the case of the negative binomial distribution with parameters $(n, q)$, $q \in (0, 1)$ (cf. VON CHOSSY, R and G. RAPPL (1983), p. 253). Thus the assertions of Proposition A.2 and Proposition A.3 are valid for collective risk models with these distribution functions, if $Y_1$ is appropriate.
Finally, we formulate a Berry-Esseen bound for deterministic sums of independent, not necessarily identically distributed random variables with finite absolute third moments (cf. FELLER (1971), p. 544).

Theorem A.1. Let the sequence of $X_i$ be independent variables and $EX_i = \mu_i$, $E(X_i - \mu_i)^2 = \sigma_i^2$, $E|X_i - \mu_i|^3 = \rho_i$, $i \in \mathbb{N}$.

Put $m_n = \sum_{i=1}^{n} \mu_i$, $s_n^2 = \sum_{i=1}^{n} \sigma_i^2$, $r_n^2 = \sum_{i=1}^{n} \rho_i$ and denote by $F_n$ the distribution function of the sum $\sum_{i=1}^{n} X_i$, $\Phi_n(x) = \Phi((x - m_n) s_n^{-1})$, $\Phi$ the standard normal distribution function. Then for all $n \in \mathbb{N}$

$$\sup_{x} |F_n(x) - \Phi_n(x)| \leq 6 r_n s_n^{-3}.$$  

The next two lemmata state some auxiliary results which are needed in Section 2. Notations and assumptions are such as stated there.

Lemma A.1. Let

$$\tilde{A}_n = \sup_{x} |\Phi_{n}^{\text{ind}}(x) - \Phi_{n}^{\text{coll}}(x)|.$$  

Then we have

$$\tilde{A}_n = O(1/\sigma_n^{\text{coll}}) \Rightarrow \sigma_n^{\text{coll}} - \sigma_n^{\text{ind}} = O(1) \text{ and } \mu_n^{\text{coll}} - \mu_n^{\text{ind}} = O(1).$$  

Proof. Put $a_n = \mu_n^{\text{ind}}$, $a'_n = \mu_n^{\text{coll}}$, $b_n = \sigma_n^{\text{ind}}$, $b'_n = \sigma_n^{\text{coll}}$.

"$\Rightarrow$": By applying the mean value theorem we obtain

$$\tilde{\tilde{A}}_n = \sup_{x} |\Phi'(\xi_n)\left(\frac{x-a_n}{b_n} - \frac{x-a'_n}{b'_n}\right)|,$$

where

$$\xi_n = \alpha_n \frac{x-a_n}{b_n} + (1-\alpha_n) \frac{x-a'_n}{b'_n}, \quad \alpha_n \in (0, 1), \quad x \in \mathbb{R}$$

Choose a sequence $x_n = a_n + cb_n$ for any $0 \neq c \in \mathbb{R}$ and replace $x$ by $x_n$. Hence in view of the assumptions we have

$$\xi_n = \alpha_n c + (1-\alpha_n) \left(\frac{a_n-a'_n}{b'_n} + c \frac{b_n}{b'_n}\right) \xrightarrow{n \to \infty} c.$$
In addition, we have

\[
\frac{x_n - a_n}{b_n} - \frac{x_n - a'_n}{b'_n} = c - \left( \frac{a_n - a'_n}{b_n} + c \frac{b_n}{b'_n} \right) = c \frac{b'_n - b_n}{b'_n} + \frac{a'_n - a_n}{b'_n}.
\]

Consequently, and because \( \tilde{A}_n = O(1/b'_n) \), \( \xi_n \to c \neq 0, \Phi'(c) > 0 \), we get \( b'_n - b_n = O(1) \) and \( a'_n - a_n = O(1) \).

\( \Leftarrow \): Again, by applying the mean value theorem, we have with \( r_n = b_n/b'_n \)

\[
\tilde{A}_n = \sup_x \Phi' \left( \alpha_n \left( \frac{x - a_n}{b_n} + \frac{a_n - a'_n}{b'_n} \right) + (1 - \alpha_n) \frac{x - a_n}{b_n} \right) \times \left( \frac{x - a_n}{b_n} (r_n - 1) + \frac{a_n - a'_n}{b'_n} \right).
\]

Therefore, according to the assumptions we have

\[
b'_n \tilde{A}_n = \sup_x \left( \Phi' \left( \alpha_n \frac{a_n - a'_n}{b'_n} \Phi' \left( \frac{x - a_n}{b_n} (1 - \alpha_n (1 - r_n)) \right) \right) \times \left( \frac{x - a_n}{b_n} (b_n - b'_n) + (a_n - a'_n) \right) \right.
\]

Now the assertion follows from \( \sup_x |x\Phi'(x)| = 1/\sqrt{2\pi} \).

**Lemma A.2.**

(i) Let the central limit theorem be valid for \( F_n^{\text{ind}} \) and \( F_n^{\text{coll}} \). Then for all \( \alpha \in (0, 1) \) we have

\[
\frac{\mathcal{F}_n^{\text{ind}}(\alpha) - \mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}} \rightarrow_{n \rightarrow \infty} \Phi^{-1}(\alpha), \quad \frac{\mathcal{F}_n^{\text{coll}}(\alpha) - \mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}} \rightarrow_{n \rightarrow \infty} \Phi^{-1}(\alpha),
\]

(ii) \( F_n^{\text{ind}} \) fulfills the Berry-Esséen bound from Theorem A.1, \( F_n^{\text{coll}} \) that from Proposition A.3.
If $\sigma_{n,\text{ind}}^2/\sigma_{n,\text{coll}}^2 \xrightarrow{n \to \infty} 1$, then
\begin{equation}
\frac{\gamma_{n,\text{ind}}(\alpha) - \mu_{n,\text{ind}}^*}{\sigma_{n,\text{ind}}^2} - \frac{\gamma_{n,\text{coll}}(\alpha) - \mu_{n,\text{coll}}^*}{\sigma_{n,\text{coll}}^2} = O\left(1/\sigma_{n,\text{ind}}^2\right), \quad \alpha \in (0, 1).
\end{equation}

**Proof.**

(i) Put
\begin{equation}
\gamma_{n,\text{ind}}(\alpha) = \frac{\gamma_{n,\text{ind}}(\alpha) - \mu_{n,\text{ind}}^*}{\sigma_{n,\text{ind}}}, \quad \gamma_{n,\text{coll}}(\alpha) \text{ analogously}.
\end{equation}

Then
\begin{equation}
\Phi\left(\gamma_{n,\text{ind}}(\alpha)\right) = \Phi_{n,\text{ind}} \circ (F_{n,\text{ind}}^{-1}) \circ F_{n,\text{ind}}\left(\inf\{p \in [0, \infty) | F_{n,\text{ind}}^\alpha(p) \geq \alpha\}\right),
\end{equation}
\begin{equation}
F_{n,\text{ind}}\left(\inf\{p \in [0, \infty) | F_{n,\text{ind}}^\alpha(p) \geq \alpha\}\right) \xrightarrow{n \to \infty} \alpha,
\end{equation}
\begin{equation}
\Phi_{n,\text{ind}} \circ (F_{n,\text{ind}}^{-1}) \xrightarrow{n \to \infty} \text{id}
\end{equation}
in view of the assumptions; thus $\gamma_{n,\text{ind}}(\alpha) \xrightarrow{n \to \infty} \Phi^{-1}(\alpha)$.

The assertion for $\gamma_{n,\text{coll}}(\alpha)$ follows from a similar argumentation.

(ii) From $\gamma_{n,\text{ind}}(\alpha) = (F_{n,\text{ind}}^{-1}(\alpha))^{-1}(\alpha)$ we obtain
\begin{equation}
\sup_x |F_{n,\text{ind}}(x) - \Phi_{n,\text{ind}}(x)| = \sup_{y \in \mathcal{M}} |y - \Phi_{n,\text{ind}} \circ (F_{n,\text{ind}}^{-1}(y))|
\end{equation}
where $\mathcal{M} = \{y | y = F_{n,\text{ind}}(x)\}$

Using the mean value theorem, we have with suitably chosen $\alpha_n \in (0, 1)$
\begin{equation}
y - \Phi_{n,\text{ind}} \circ (F_{n,\text{ind}}^{-1}(y)) = \Phi \circ \Phi^{-1}(y) - \Phi \left(\frac{(F_{n,\text{ind}}^{-1}(y) - \mu_{n,\text{ind}}^*)}{\sigma_{n,\text{ind}}^2}\right) = \Phi'(\alpha_n \Phi^{-1}(y) + (1 - \alpha_n) \gamma_{n,\text{ind}}(y))(\Phi^{-1}(y) - \gamma_{n,\text{ind}}(y)).
\end{equation}

Because of (i),
\begin{equation}
\Phi'(\alpha_n \Phi^{-1}(y) + (1 - \alpha_n) \gamma_{n,\text{ind}}(y)) \xrightarrow{n \to \infty} \Phi' \circ \Phi^{-1}(y) \text{ uniformly in } y.
\end{equation}

$\Phi' \circ \Phi^{-1}(y) > 0$ and the Berry-Esseen bound now supply
\begin{equation}
\Phi^{-1}(y) - \gamma_{n,\text{ind}}(y) = O\left(1/\sigma_{n,\text{ind}}^2\right).
\end{equation}
The same argumentation applied to $F_n^{\text{coll}}$ and $\Phi_n^{\text{coll}}$ yields
\begin{equation}
(\text{A.28}) \quad \Phi^{-1}(y) - \gamma_n^{\text{coll}}(y) = O\left(1/\sigma_n^{\text{coll}}\right).
\end{equation}
Finally, the assertion follows from the last two bounds taking $\sigma_n^{\text{ind}}/\sigma_n^{\text{coll}} \to 1$ into account.

REFERENCES


S. KUON, M. RADTKE AND A. REICH
Kölnische Rückversicherungs-Gesellschaft AG,
Theodor-Heuss-Ring 11, Postfach 10 80 16,
D-5000 Köln 1, Germany.