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; colletive 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

$$(1) \qquad \Delta = \sup_{x \in \mathbb{R}} |P(S^{\text{ind}} \leq x) - P(S^{\text{coll}} \leq x)|$$

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 Δ 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 portfolios.

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

$$\Delta \to 0 \tag{2}$$

(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 Δ 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

$$\frac{\operatorname{Var} S^{\mathrm{coll}}}{\operatorname{Var} S^{\mathrm{ind}}} \to 1 . \tag{3}$$

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-Esséen 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

$$\operatorname{Var} S^{\mathrm{ind}} = \operatorname{Var} S^{\mathrm{coll}} .$$

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:

Definition 1.1. (Individual risk model, cf BOWERS et al (1987), p 25).

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_n^{ind} instead of S^{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:

$$(1.1) \qquad P(Y_i \leq x) = \frac{P(X_i \leq x) - P(X_i = 0)}{P(X_i > 0)} \quad \text{if} \quad x > 0;\ P(Y_i = 0) = 0$$

Thus, we have the representation $X_i \stackrel{\mathcal{L}}{=} I_i Y_i$ 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

$$N^* = \sum_{i=1}^{n} I_i, \quad I_i \stackrel{\mathcal{L}}{\sim} \beta(1, q_i) \text{ with } q_i = P(X_i > 0) \text{ and } \beta \text{ the binomial distribution.}$$

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^{ind} of the individual aggregate claims amount S^{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 (KORNYA (1983), HIPP (1985, 1986), DE PRIL (1989)), still costly. Therefore, at a very early stage of risk theory, the question of the calculability of F^{ind} led to the concept of the collective risk model (BOWERS et al. (1987), p. 317), which is easier to handle when theoretical considerations are made. Its fast numerical calculability (Panjer-class) is another, more recent reason for using it.

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

Definition 1.2. (Collective risk model, cf. BOWERS et al. (1987), p. 317).

The random collective claims amounts Z_i, $i \in \mathbb{N}$, set up a sequence of independent, indentically 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: $0 < EN < \infty$, $0 < \text{Var } N < \infty$, $0 < EZ_1 < \infty$, $0 < \text{Var } Z_i < \infty$.

Remark 1.2. S^{coll} satisfies $ES^{\text{coll}} = ENEZ_1$ and $\operatorname{Var} S^{\text{coll}} = EN \operatorname{Var} Z_1 + \operatorname{Var} N(EZ_1)^2$.

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, $1 \le i \le n$. This, in turn, brings us to

Remark 1.3. If the individual model satifies $X_i \overset{\mathscr{L}}{\sim} F$ for each i, it follows that $N^* \overset{\mathscr{L}}{\sim} \beta(n, q)$ with $q = P(X_1 > 0)$. Put $N \overset{\mathscr{L}}{=} N^*$ and $Z_1 \overset{\mathscr{L}}{\sim} G$, where $G(x) = (F(x)-(1-q))/q$ for $x \geqslant 0$. Thus we get $S^{\text{ind}} \overset{\mathscr{L}}{=} S^{\text{coll}}$ 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^{coll} of S^{coll} supplies a good approximation to the distribution function F^{ind} of S^{ind}. The following procedure is usual:

Remark 1.4. Define the distribution function G of Z_1 by

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

and

$$(1.3) \qquad q = \frac{1}{n} \sum_{i=1}^{n} q_i, \; q_i = P(X_i > 0), \; X_i \overset{\mathscr{L}}{=} I_i Y_i, \; X_i \overset{\mathscr{L}}{\sim} F_i, \; Y_i \overset{\mathscr{L}}{\sim} G_i.$$

In this remark the representation $X_i \overset{\mathscr{L}}{=} I_i Y_i$ is such as given in Remark 1.1.

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

Assumption 1.1 establishes the existence of real-valued constants a', b', c', d' (independent of n) with $0 < c' \le EZ_1 \le d' < \infty$ and $0 < a' \le \operatorname{Var} Z_1 \le b' < \infty$. 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

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

G as defined in (1.2). Then

$$ES^{\text{coll}} = ES^{\text{ind}} \quad \text{and} \quad \operatorname{Var} S^{\text{coll}} = \operatorname{Var} S^{\text{ind}} + \Delta_{Bi},$$

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

$$S^{\text{coll}} = \sum_{i=1}^{n} Z_i^* \quad \text{with } Z_i^* \overset{\mathcal{L}}{\sim} G^*,$$

where for $x \geq 0$

$$G^*(x) = \frac{1}{n} \sum_{i=1}^{n} F_i(x) \quad \text{and} \quad X_i \overset{\mathcal{L}}{\sim} F_i \tag{1.5}$$

Since

$$\begin{aligned} G^*(x) &= \frac{1}{n} \sum_{i=1}^{n} (1-q_i) + \frac{1}{n} \sum_{i=1}^{n} q_i \frac{F_i(x)-(1-q_i)}{q_i} \\ &= (1-q)+qG(x), \end{aligned} \tag{1.6}$$

we conclude from the characteristic function

$$\begin{aligned} (Ee^{uZ_1^*})^n &= \left((1-q)+q \int_0^\infty e^{ux}\, G(dx)\right)^n \\ &= \sum_{k=0}^{n} \binom{n}{k} (1-q)^{n-k} q^k \int_0^\infty e^{ux}\, G^{*k}(dx), \end{aligned} \tag{1.7}$$

that the two approaches lead to the same collective model.

Model 1.2. The compound Poisson approximation

Let

$$N \overset{\mathcal{L}}{\sim} \pi(nq),\ \pi \text{ the Poisson distribution,} \quad \text{and} \quad Z_1 \overset{\mathcal{L}}{\sim} G,$$

G as defined in (1.2). Then

$$ES^{\text{coll}} = ES^{\text{ind}} \quad \text{and} \quad \operatorname{Var} S^{\text{coll}} = \operatorname{Var} S^{\text{ind}} + \Delta_{Po},$$

where

$$\Delta_{Po} = \sum_{i=1}^{n} (EX_i)^2. \tag{1.8}$$

Model 1.3. The compound negative binomial approximation

Let

$N \sim \mathcal{NB}\left(n, \frac{1}{1+q}\right)$, $\mathcal{NB}$ the negative binomial distribution, and $Z_1 \sim G$,

G as defined in (1.2). Then

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

where

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

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

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 $\operatorname{Var} S^{\text{ind}}$. Obviously, for the overestimation the following is valid:

$$0 \leq \Delta_{Bi} < \Delta_{Po} < \Delta_{NB}.$$

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

(i) N binomial distributed:

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

(ii) N Poisson distributed:

$$\frac{\operatorname{Var} S^{\text{coll}}}{\operatorname{Var} S^{\text{ind}}} - 1 = \frac{\Delta_{Po}}{\sum_{i=1}^{n} \operatorname{Var} X_i} \in \left[\frac{c^2}{b}, \frac{d^2}{a}\right], \tag{1.11}$$

(iii) N negative binomial distributed:

$$(1.12)\quad \frac{\operatorname{Var} S^{\text{coll}}}{\operatorname{Var} S^{\text{ind}}} - 1 = \frac{\Delta_{NB}}{\sum_{i=1}^{n} \operatorname{Var} X_i} \in \left[1 + \frac{c^2}{b}\left(1 + \frac{1}{n}\right), 1 + \frac{d^2}{a}\left(1 + \frac{1}{n}\right)\right].$$

Hence, only in the case of binomial distributed claims number N we can achieve $\operatorname{Var} S^{\text{ind}} = \operatorname{Var} S^{\text{coll}}$. For instance, this is fulfilled if $EX_i = EX_1$ for all $i = 1, \dots, n$ 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), $1 \le i \le n$, $X_i \in \{0, 1\}$. Let for each i

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

Therefore we have

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

From that

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

easily is concluded.

Further on we shall analyze the impact to which $\operatorname{Var} S^{\text{ind}} \neq \operatorname{Var} S^{\text{coll}}$ 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^{ind} instead of S^{ind}, S_n^{coll} instead of S^{coll}, etc.

Assumption 1.1 instantly implies $ES_n^{\text{ind}} \xrightarrow[n \to \infty]{} \infty$, $\operatorname{Var} S_n^{\text{ind}} \xrightarrow[n \to \infty]{} \infty$, and, as $ES_n^{\text{ind}} = ES_n^{\text{coll}}$, also $ES_n^{\text{coll}} \xrightarrow[n \to \infty]{} \infty$. 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 $\operatorname{Var} S_n^{\text{coll}} - \operatorname{Var} S_n^{\text{ind}} \xrightarrow[n \to \infty]{} \infty$ 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

$$\mathscr{S}_n^{\mathrm{ind}}(\alpha) = \inf\{x \mid P(S_n^{\mathrm{ind}} \le x) \ge \alpha\} \tag{1.13}$$

with security level $\alpha \in (0, 1)$, $\mathscr{S}_n^{\mathrm{coll}}(\alpha)$ analogously That is, the premium $\mathscr{S}_n^{\mathrm{ind}}(\alpha)$ is not exceeded by the aggregate claims amount S_n^{ind} with probability α. Of course the difference $\mathscr{S}_n^{\mathrm{ind}}(\alpha) - \mathscr{S}_n^{\mathrm{coll}}(\alpha)$ 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)):

$$\mathscr{S}_n^{\mathrm{ind}}(\alpha) \approx \mu_n^{\mathrm{ind}} + \Phi^{-1}(\alpha)\, \sigma_n^{\mathrm{ind}} \tag{1.14}$$

and

$$\mathscr{S}_n^{\mathrm{coll}}(\alpha) \approx \mu_n^{\mathrm{coll}} + \Phi^{-1}(\alpha)\ \sigma_n^{\mathrm{coll}}, \tag{1.15}$$

where $\mu_n^{\mathrm{ind}} = ES_n^{\mathrm{ind}}$, $\mu_n^{\mathrm{coll}} = ES_n^{\mathrm{coll}}$, $\sigma_n^{\mathrm{ind}} = \sqrt{\mathrm{Var}\ S_n^{\mathrm{ind}}}$, $\sigma_n^{\mathrm{coll}} = \sqrt{\mathrm{Var}\ S_n^{\mathrm{coll}}}$. Φ^{-1} denotes the inverse function of standardized normal distribution function Φ.

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

$$\mathscr{S}_n^{\mathrm{coll}}(\alpha) - \mathscr{S}_n^{\mathrm{ind}}(\alpha) \approx (\sigma_n^{\mathrm{coll}} - \sigma_n^{\mathrm{ind}})\ \Phi^{-1}(\alpha) \ge 0\,. \tag{1.16}$$

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

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 $S_n^{\mathrm{ind}} = \sum_{i=1}^{n} X_i$ of growing size by a sequence of so-called homogeneous collective risk models $S_n^{\mathrm{coll}} = \sum_{i=1}^{N_n} Z_i$. 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 $(S_n^{\mathrm{coll}})_{n\in\mathbb{N}}$ a sequence of homogeneous collective risk models if, for each $n \in \mathbb{N}$, $S_n^{\mathrm{coll}} = \sum_{i=1}^{N_n} Z_i$ is a collective risk model and the distribution function of Z_i is independent of n. In addition we assume that N_n possesses a representation $N_n \overset{\mathscr{L}}{\sim} L^{*n}$ with arbitrary distribution function L on $\mathbb{N}_0$ (cf. Proposition A.1).

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 $EZ_1^3<\infty$, $EN_n^3<\infty$ for all $n\in\mathbb{N}$. Under these conditions, all the results listed in the Appendix are applicable to S_n^{ind} and S_n^{coll} and to their distribution functions F_n^{ind} and F_n^{coll}.

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

$$\Delta_n=\sup_x|F_n^{\text{ind}}(x)-F_n^{\text{coll}}(x)|. \tag{2.1}$$

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

$$\begin{aligned}F_n^{\text{ind}}(x)-F_n^{\text{coll}}(x)&=(F_n^{\text{ind}}(x)-\Phi_n^{\text{ind}}(x))+(\Phi_n^{\text{ind}}(x)-\Phi_n^{\text{coll}}(x))\\&\quad+(\Phi_n^{\text{coll}}(x)-F_n^{\text{coll}}(x)).\end{aligned} \tag{2.2}$$

Since

$$\frac{\sqrt{\sum_{i=1}^{n}E|X_i-EX_i|^3}}{\sqrt{\left(\sum_{i=1}^{n}\operatorname{Var}X_i\right)^3}}\le\frac{\sqrt{n}\sup_{i\in\mathbb{N}}\{E|X_i-EX_i|^3\}}{\sqrt{\left(n\inf_{i\in\mathbb{N}}\{\operatorname{Var}X_i\}\right)^3}}\xrightarrow[n\to\infty]{}0, \tag{2.3}$$

the central limit theorem for S_n^{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

$$\Delta_n\xrightarrow[n\to\infty]{}0\Leftrightarrow\sigma_n^{\text{ind}}/\sigma_n^{\text{coll}}\xrightarrow[n\to\infty]{}1\text{ and }(\mu_n^{\text{ind}}-\mu_n^{\text{coll}})/\sigma_n^{\text{coll}}\xrightarrow[n\to\infty]{}0 \tag{2.4}$$

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

$$\Phi_n^{\text{ind}}(x)-\Phi_n^{\text{coll}}(x)\xrightarrow[n\to\infty]{}0\text{ uniformly in }x \tag{2 5}$$

is equivalent to the right side of the assertion.

With $r_n = \sigma_n^{\text{ind}}/\sigma_n^{\text{coll}}$, we have

$$\Phi_n^{\text{coll}}(x) = \Phi\left(\frac{x-\mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}}\right) = \Phi\left(\frac{x-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}}\, r_n + \frac{\mu_n^{\text{ind}}-\mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}}\right). \tag{2.6}$$

"$\Leftarrow$". The assertion is true, as Φ is a uniformly continuous function.

"$\Rightarrow$". Because Φ is a strictly increasing function, $\Delta_n \xrightarrow[n\to\infty]{} 0$ supplies

$$\frac{x-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}} - \frac{x-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}}\, r_n - \frac{\mu_n^{\text{ind}}-\mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}} \xrightarrow[n\to\infty]{} 0 \text{ uniformly in } x \tag{2.7}$$

By replacing x with μ_n^{ind} we get $(\mu_n^{\text{ind}}-\mu_n^{\text{coll}})/\sigma_n^{\text{coll}} \xrightarrow[n\to\infty]{} 0$.

Then by replacing x with $\mu_n^{\text{ind}}+\sigma_n^{\text{ind}}$, we obtain $r_n \xrightarrow[n\to\infty]{} 1$.

Furthermore, for the difference $\mathscr{S}_n^{\text{coll}}(\alpha)-\mathscr{S}_n^{\text{ind}}(\alpha)$ of the percentile premiums related to S_n^{ind} and S_n^{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

$$\mathscr{S}_n^{\text{ind}}(\alpha)-\mathscr{S}_n^{\text{coll}}(\alpha) = o(\sigma_n^{\text{ind}}) \tag{2.8}$$

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

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

$$\frac{\mathscr{S}_n^{\text{coll}}(\alpha)-\mathscr{S}_n^{\text{ind}}(\alpha)}{\sigma_n^{\text{ind}}} = \left(\frac{\mathscr{S}_n^{\text{coll}}(\alpha)-\mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}}\right) \times \frac{\sigma_n^{\text{coll}}}{\sigma_n^{\text{ind}}} - \frac{\mathscr{S}_n^{\text{ind}}(\alpha)-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}} + \frac{\mu_n^{\text{coll}}-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}}\,. \tag{2.9}$$

Then, Lemma A.2 (i) supplies

$$\frac{\mathscr{S}_n^{\text{coll}}(\alpha)-\mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}} \xrightarrow[n\to\infty]{} \Phi^{-1}(\alpha), \quad \frac{\mathscr{S}_n^{\text{ind}}(\alpha)-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}} \xrightarrow[n\to\infty]{} \Phi^{-1}(\alpha)\,. \tag{2.10}$$

Thus, the conclusion from the right side to the left side follows directly from the above indentity. The converse is also true, since $\alpha \in (0, 1)$ is arbitrary. So is $\Phi^{-1}(\alpha)$. As $(\mu_n^{\text{ind}} - \mu_n^{\text{coll}})/\sigma_n^{\text{coll}}$ is independent of α, 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) $\text{Var}\, S_n^{\text{ind}}/\text{Var}\, S_n^{\text{coll}} \xrightarrow[n\to\infty]{} 1$ and $(ES_n^{\text{ind}} - ES_n^{\text{coll}})/\sqrt{\text{Var}\, S_n^{\text{coll}}} \xrightarrow[n\to\infty]{} 0$,

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

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

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

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

(iii) $\sqrt{\text{Var}\, S_n^{\text{ind}}} - \sqrt{\text{Var}\, S_n^{\text{coll}}} = O(1)$ and $ES_n^{\text{ind}} - ES_n^{\text{coll}} = O(1)$,

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

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:

$$\begin{aligned} \Delta_n = O(1/\sqrt{\text{Var}\, S_n^{\text{ind}}}) &\Leftrightarrow \tilde{\Delta}_n = O(1/\sqrt{\text{Var}\, S_n^{\text{ind}}}) \\ &\Leftrightarrow \tilde{\Delta}_n = O(1/\sqrt{\text{Var}\, S_n^{\text{coll}}}) \\ &\Leftrightarrow \Delta_n = O(1/\sqrt{\text{Var}\, S_n^{\text{coll}}}), \end{aligned} \tag{2.11}$$

where $\tilde{\Delta}_n = \sup_x |\Phi_n^{\text{ind}}(x) - \Phi_n^{\text{coll}}(x)|$. According to Lemma A.1, (i)$\Leftrightarrow$(ii)$\Leftrightarrow$(iii) has been proved.

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

$$\mathscr{S}_n^{\text{ind}}(\alpha) - \mathscr{S}_n^{\text{coll}}(\alpha) = (\mu_n^{\text{ind}} - \mu_n^{\text{coll}}) + (\gamma_n^{\text{ind}}(\alpha) - \gamma_n^{\text{coll}}(\alpha))\,\sigma_n^{\text{ind}} + \gamma_n^{\text{coll}}(\alpha)\,(\sigma_n^{\text{ind}} - \sigma_n^{\text{coll}}). \tag{2.12}$$

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

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

$$\begin{aligned} O(1) &= \mathscr{S}_n^{\text{ind}}(\alpha) - \mathscr{S}_n^{\text{coll}}(\alpha) \\ &= (\mu_n^{\text{ind}} - \mu_n^{\text{coll}}) + \gamma_n^{\text{coll}}(\alpha)\,(\sigma_n^{\text{ind}} - \sigma_n^{\text{coll}}) + O(1), \end{aligned} \tag{2.13}$$

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

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

$$O(1) = (\gamma_n^{\text{coll}}(\alpha_1) - \gamma_n^{\text{coll}}(\alpha_2))\,(\sigma_n^{\text{ind}} - \sigma_n^{\text{coll}}). \tag{2.14}$$

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

$$\sigma_n^{\text{ind}} - \sigma_n^{\text{coll}} = O(1), \text{ hence } \mu_n^{\text{ind}} - \mu_n^{\text{coll}} = O(1). \tag{2.15}$$

Thus (iv)$\Rightarrow$(iii) has been proved, and therefore (iv)$\Rightarrow$(i) since (iii)$\Leftrightarrow$(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 $\mathscr{K}_k$, $k = 1, \ldots, K$; thus $\{X_i | i \in \mathbb{N}\} = \bigcup_{k=1}^{K} \mathscr{K}_k$. Each class is represented by a distribution function $F_{(k)}$. $X_i \in \mathscr{K}_k$ means that $X_i \stackrel{\mathscr{L}}{\sim} F_i$ and $F_i = F_{(k)}$. Correspondingly $q_{(k)} \in (0, 1]$ denotes the representative of q_i, if $X_i \in \mathscr{K}_k$, 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 integer $n_k = \sum_{i=1}^{n} 1_{\{X_i \in \mathscr{X}_k\}}$; 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 $n_k - nc_k = O(1)$ as n tends to infinity.

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

$$q^{(n)} = \frac{1}{n} \sum_{i=1}^{n} q_i \quad \text{and} \quad \bar{q} = \sum_{k=1}^{K} c_k q_{(k)}. \tag{2.16}$$

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

$$G_n(x) = \sum_{i=1}^{n} \frac{q_i}{nq^{(n)}} G_i(x) \quad \text{with} \quad G_i(x) = \frac{F_i(x) - (1 - q_i)}{q_i}, \ x \geq 0, \tag{2.17}$$

that is

$$G_n(x) = \sum_{k=1}^{K} \frac{n_k}{n} \frac{q_{(k)}}{q^{(n)}} G_{(k)}(x). \tag{2.18}$$

Moreover, let

$$\bar{G}(x) = \sum_{k=1}^{K} c_k \frac{q_{(k)}}{\bar{q}} G_{(k)}(x) \text{ with } G_{(k)}(x) = \frac{F_{(k)}(x) - (1 - q_{(k)})}{q_{(k)}}, \ x \geq 0. \tag{2.19}$$

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 $\bar{N}_n$, refering to the parameters $q^{(n)}$ and $\bar{q}$. We write Z_1, $\bar{Z}_1$ resp. for the collective single claims amount variable, where $Z_1 \overset{\mathscr{D}}{\sim} G_n$, $\bar{Z}_1 \overset{\mathscr{D}}{\sim} \bar{G}$ resp.; thus

$$\tilde{S}_n^{\text{coll}} = \sum_{i=1}^{N_n} Z_i, \ Z_i \overset{\mathscr{D}}{\sim} Z_1 \tag{2.20}$$

and

$$\bar{S}_n^{\text{coll}} = \sum_{i=1}^{\bar{N}_n} \bar{Z}_i, \ \bar{Z}_i \overset{\mathscr{D}}{\sim} \bar{Z}_1. \tag{2.21}$$

Finally, in the collective risk model $\bar{S}_n^{\text{coll}}$ we denote by $\bar{\mu}_n^{\text{coll}}$ and $\bar{\sigma}_n^{\text{coll}}$ the mean value and the standard deviation resp., in line with μ_n^{ind} and σ_n^{ind} above.

In this framework the portfolio growth is defined by the corresponding sequence of homogeneous collective risk models $(\bar{S}_n^{\text{coll}})_{n \in \mathbb{N}}$ 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^{ind} of S_n^{ind} and $\bar{F}_n^{\text{coll}}$ of $\bar{S}_n^{\text{coll}}$ resp., fulfill a Berry-Esséen bound, i.e.

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

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

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

Proof:

(i) As $E|X_i - EX_i|^3 \le EX_i^3 + (EX_i)^3$, we have from the assumptions $\max\limits_{i \in \mathbb{N}} \{E|X_i - EX_i|^3\} < \infty$. Assumption 1.1 supplies $\text{Var } S_n^{\text{ind}} \ge na$ and thus we conclude

$$\text{(2.22)} \qquad \frac{\sum_{i=1}^{n} E|X_i - EX_i|^3}{\sum_{i=1}^{n} \text{Var } X_i} \le \frac{\max\limits_{i \in \mathbb{N}} \{E|X_i - EX_i|^3\}}{a} < \infty .$$

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

$$\text{(2.23)} \qquad \sup_x |F_n^{\text{ind}}(x) - \Phi_n^{\text{ind}}(x)| \le \frac{6}{\sigma_n^{\text{ind}}} \frac{\sum_{i=1}^{n} E|X_i - EX_i|^3}{\sum_{i=1}^{n} \text{Var } X_i} \text{ for all } n .$$

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

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

must be taken into consideration.

Consequently S_n^{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 $\tilde{S}_n^{\text{coll}}$ and $\tilde{\Phi}_n^{\text{coll}}$ resp., $\tilde{\Phi}_n^{\text{coll}}(x) = \Phi((x - \tilde{\mu}_n^{\text{coll}})/\tilde{\sigma}_n^{\text{coll}})$, cannot be directly deduced from the Berry-

Esséen 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

$$ES_n^{\widetilde{\text{coll}}} = E\bar{S}_n^{\text{coll}} + O(1) \tag{2.25}$$

and

$$\sqrt{\operatorname{Var} \tilde{S}_n^{\text{coll}}} = \sqrt{\operatorname{Var} \bar{S}_n^{\text{coll}}} + O(1/\sqrt{\operatorname{Var} \tilde{S}_n^{\text{coll}}}). \tag{2.26}$$

Together with

$$\sup_x |\bar{F}_n^{\text{coll}}(x) - \tilde{F}_n^{\text{coll}}(x)| = O(1/\operatorname{Var} \tilde{S}_n^{\text{coll}}). \tag{2.27}$$

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

$$\sup_x |\tilde{F}_n^{\text{coll}}(x) - \tilde{\Phi}_n^{\text{coll}}(x)| = O(1/\sqrt{\operatorname{Var} \tilde{S}_n^{\text{coll}}}), \tag{2.28}$$

we get ultimately the validity of Theorem 2.1 and Theorem 2 2 even for $\tilde{S}_n^{\text{coll}}$ straightforward from the identity

$$F_n^{\text{ind}}(x) - \tilde{F}_n^{\text{coll}}(x) = (F_n^{\text{ind}}(x) - \bar{F}_n^{\text{coll}}(x)) + (\bar{F}_n^{\text{coll}}(x) - \tilde{F}_n^{\text{coll}}(x)). \tag{2.29}$$

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

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

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

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

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

(iii) $\sqrt{\operatorname{Var} S_n^{\text{ind}}} - \sqrt{\operatorname{Var} \tilde{S}_n^{\text{coll}}} = O(1)$,

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

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

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

Therefore,

(2 31) $$P(X_i > 0) = q_i,\ EX_i = q_i,\ \operatorname{Var} X_i = q_i(2-q_i),$$

For each $n \in \mathbb{N}$, let

(2.32) $$a_n = \sum_{i=1}^{n} q_i, \quad \text{and} \quad b_n = \sum_{i=1}^{n} q_i^2$$

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

(2.33) $$ES_n^{\text{ind}} = a_n \quad \text{and} \quad \operatorname{Var} S_n^{\text{ind}} = 2a_n - b_n .$$

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

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

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 $Z_1 \stackrel{\mathcal{L}}{\sim} G$, we obtain $EZ_1 = \operatorname{Var} Z_1 = 1$.
Moreover, we stipulate that N_n is distributed as $\beta(n, q^{(n)})$ Hence (cf. Model 1.1),

(2.35) $$ES_n^{\text{coll}} = a_n \quad \text{and} \quad \operatorname{Var} S_n^{\text{coll}} = 2a_n - \frac{a_n^2}{n} .$$

Now, for all $\alpha \in (0, 1/2)$, with a suitable choice of $q_i \in (0, 1)$, $i \in \mathbb{N}$,

$$(2.36) \qquad a_n \sim n - n^{1-\alpha} \quad \text{and} \quad b_n \sim n - 2n^{1-\alpha} + \frac{(1-\alpha)^2}{1-2\alpha}\, n^{1-2\alpha}$$

is fulfilled. For instance, $q_i = 1-(1-\alpha)/i^{\alpha}$ is appropriate.
For proof of this, note that

$$(2.37) \qquad \sum_{i=1}^{n} (1-\beta)\, i^{-\beta} \sim \int_1^n (1-\beta)\, x^{-\beta}\, dx \sim n^{1-\beta}, \qquad \beta \in (0, 1).$$

All these definitions supply

$$(2.38) \qquad \frac{\operatorname{Var} S_n^{\text{coll}}}{\operatorname{Var} S_n^{\text{ind}}} = \frac{2 - \dfrac{a_n}{n}}{2 - \dfrac{b_n}{a_n}} \xrightarrow[n\to\infty]{} 1,$$

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

However,

$$(2.39) \quad \sqrt{\operatorname{Var} S_n^{\text{ind}}} - \sqrt{\operatorname{Var} S_n^{\text{coll}}} = \left(\frac{\sqrt{a_n^3}}{n} - \frac{b_n}{\sqrt{a_n}}\right)\left(2 - \frac{a_n}{n}\right)^{-1}\left(1 + \sqrt{\frac{\operatorname{Var} S_n^{\text{coll}}}{\operatorname{Var} S_n^{\text{ind}}}}\right)^{-1}$$

$$\sim \sqrt{n^{1-4\alpha}}\left(1 - \frac{(1-\alpha)^2}{1-2\alpha}\right)\frac{1}{2}$$

$$\xrightarrow[n\to\infty]{} \begin{cases} -\infty & \text{if } \alpha \in (0, \frac{1}{4}), \\ -\frac{1}{16} & \text{if } \alpha = \frac{1}{4}, \\ 0 & \text{if } \alpha \in (\frac{1}{4}, \frac{1}{2}), \end{cases}$$

since $\dfrac{a_n^2 - nb_n}{n\sqrt{a_n}} \sim \sqrt{n^{1-4\alpha}}$.

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 $ES^{\text{ind}} = ES^{\text{coll}}$ 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_1^{mod} with discrete range $\{k\gamma | k \in \mathbb{N}\}$, $\gamma > 0$ fixed, by setting

$$P(Z_1^{\text{mod}} = k\gamma) = g(k), \; k \in \mathbb{N}, \tag{3.1}$$

where (cf. Remark 1.4)

$$g(k) = \sum_{i=1}^{n} \frac{q_i}{nq} g_i(k), \; g_i(k) = \frac{P(X_i = k)}{q_i}, \; q = \frac{1}{n} \sum_{i=1}^{n} q_i, \; q_i = P(X_i > 0). \tag{3.2}$$

Z_1^{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)

$$EZ_1^{\text{mod}} = \gamma EZ_1 = \frac{\gamma}{nq} \sum_{i=1}^{n} EX_i \tag{3.3}$$

and

$$E(Z_1^{\text{mod}})^2 = \gamma^2 EZ_1^2 = \frac{\gamma^2}{nq} \sum_{i=1}^{n} EX_i^2 . \tag{3.4}$$

If one considers $S^{\text{coll}} = \sum_{i=1}^{N} Z_i^{\text{mod}}$, the basic requirement $ES^{\text{ind}} = ES^{\text{coll}}$ results in $EN = nq/\gamma$, because $ES^{\text{coll}} = ENEZ_1^{\text{mod}}$. The following is also valid in this case.

$$\begin{aligned} \text{Var } S^{\text{coll}} &= ENE(Z_1^{\text{mod}})^2 + (\text{Var } N - EN)\,(EZ_1^{\text{mod}})^2 \\ &= \frac{\gamma^2 EN}{nq} \sum_{i=1}^{n} EX_i^2 + \left(\frac{\gamma}{nq}\right)^2 (\text{Var } N - EN)\,(ES^{\text{ind}})^2 \end{aligned} \tag{3.5}$$

$$= \gamma \sum_{i=1}^{n} EX_i^2 + \frac{1}{EN}\left(\frac{\operatorname{Var} N}{EN} - 1\right)(ES^{\mathrm{ind}})^2$$

$$= \operatorname{Var} S^{\mathrm{ind}} + \Delta(\gamma),$$

where

$$\Delta(\gamma) = (\gamma - 1)\sum_{i=1}^{n} EX_i^2 + \sum_{i=1}^{n} (EX_i)^2 + \frac{1}{EN}\left(\frac{\operatorname{Var} N}{EN} - 1\right)(ES^{\mathrm{ind}})^2. \tag{3.6}$$

Therefore the following equivalence holds:

$$\operatorname{Var} S^{\mathrm{coll}} = \operatorname{Var} S^{\mathrm{ind}} \Leftrightarrow \Delta(\gamma) = 0. \tag{3.7}$$

$\Delta(\gamma) = 0$ 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

$$\gamma = 1 - \frac{\sum_{i=1}^{n} (EX_i)^2 - n^{-1}(ES^{\mathrm{ind}})^2}{\sum_{i=1}^{n} EX_i^2} \quad \text{and } N \stackrel{\mathcal{L}}{\sim} \beta(n, q/\gamma). \tag{3.8}$$

Then $\gamma \in (0, 1]$, since $(ES^{\mathrm{ind}})^2 \le n \sum_{i=1}^{n} (EX_i)^2$ and

$$\gamma = \left(\sum_{i=1}^{n} \operatorname{Var} X_i + n^{-1}(ES^{\mathrm{ind}})^2\right) \Big/ \sum_{i=1}^{n} EX_i^2$$

If $q/\gamma \ge 1$ we modify the parameters n, q, γ, see below. Obviously we have $\Delta(\gamma) = 0$ and, hence, $\operatorname{Var} S^{\mathrm{ind}} = \operatorname{Var} S^{\mathrm{coll}}$. 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, γ facilitates obtaining in addition $EN = E \sum_{i=1}^{n} 1_{\{X_i > 0\}}$. For this purpose, we set $N \stackrel{\mathcal{L}}{\sim} \beta(n', q'/\gamma')$ and adjust n', q', γ' accordingly. The condition $ES^{\mathrm{ind}} = ES^{\mathrm{coll}}$ implies that $n'q' = nq$; consequently $q' = \dfrac{n}{n'} q$

From the equivalence $\operatorname{Var} S^{\text{ind}} = \operatorname{Var} S^{\text{coll}} \Leftrightarrow \Delta(\gamma) = 0$ we deduce

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

On the other hand $EN = E \sum_{i=1}^{n} 1_{\{X_i > 0\}} \Leftrightarrow \gamma' = 1$.

This leads to the choice

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

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

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

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, γ' 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

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

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

$$\frac{(ES^{\text{ind}})^2}{\sum_{i=1}^{n} (EX_i)^2} = \frac{n}{2} . \tag{3.12}$$

If n is even we have $n' = n/2$ and $\gamma' = 1$, however.

$$\frac{q'}{\gamma'} = \frac{nq}{n'\gamma'} = nq_0 \frac{2}{n} > 1 .$$

Model 3.2. The modified compound Poisson approximation.

Let

$$(3.13) \quad \gamma = 1 - \frac{\sum_{i=1}^{n} (EX_i)^2}{\sum_{i=1}^{n} EX_i^2} = \frac{\operatorname{Var} S^{\text{ind}}}{\sum_{i=1}^{n} EX_i^2} \quad \text{and} \quad N \overset{\mathcal{L}}{\sim} \pi\left(\frac{nq}{\gamma}\right).$$

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

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

$$(3.14) \quad \gamma = \frac{\sum_{i=1}^{n} \operatorname{Var} X_i}{\sum_{i=1}^{n} EX_i^2} \left(1 - \frac{1}{n} \frac{(ES^{\text{ind}})^2}{\sum_{i=1}^{n} \operatorname{Var} X_i} \right) \quad \text{and } N \overset{\mathcal{L}}{\sim} \mathcal{NB}\left(n, \frac{1}{1+q/\gamma}\right).$$

Obviously $\gamma < \left(\sum_{i=1}^{n} \operatorname{Var} X_i \right) \Big/ \left(\sum_{i=1}^{n} EX_i^2 \right) < 1.$

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

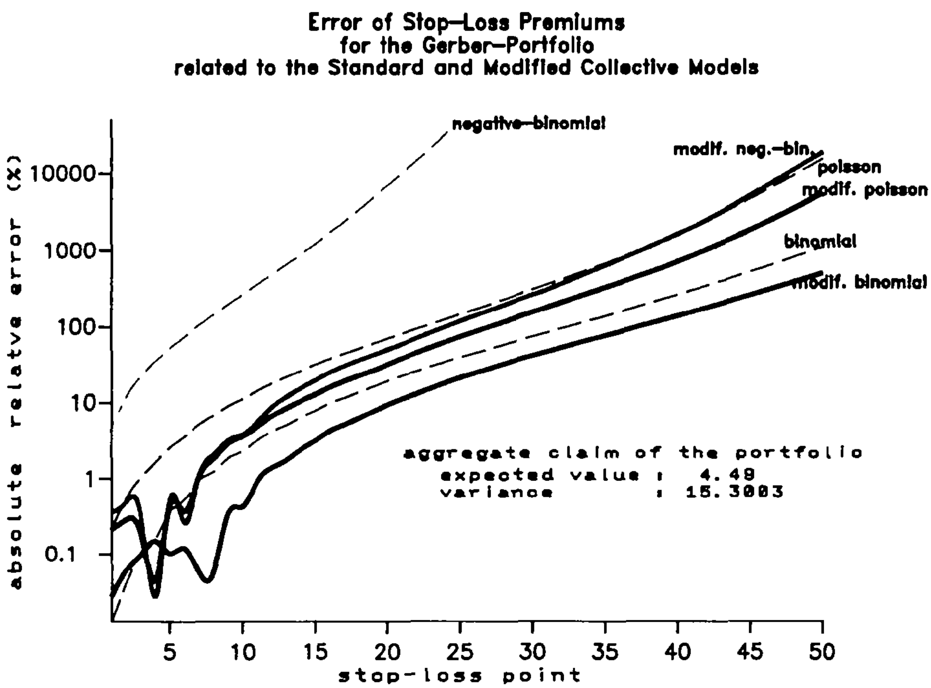
GERBER (1979), p. 53) and the 100-fold Gerber-Portfolio. Comparison was made between the different models by the relative error, that is the absolute error in percentage of the "true" risk premium, which was exactly calculated by convolution. It can be easily seen from the figures below, that the modified 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

p_i	Amount at Risk				
	1	2	3	4	5
0 03	2	3	1	2	0
0 04	0	1	2	2	1
0 05	0	2	4	2	2
0 06	0	2	2	2	1
Total	0 06	0 35	0 43	0 36	0 20

Gerber — Portfolio of 3100 Policies

p_i	Amount at Risk				
	1	2	3	4	5
0 03	200	30	100	200	0
0 04	0	100	200	200	100
0 05	0	200	400	200	200
0 06	0	200	200	200	100
Total	06	35	43	36	20



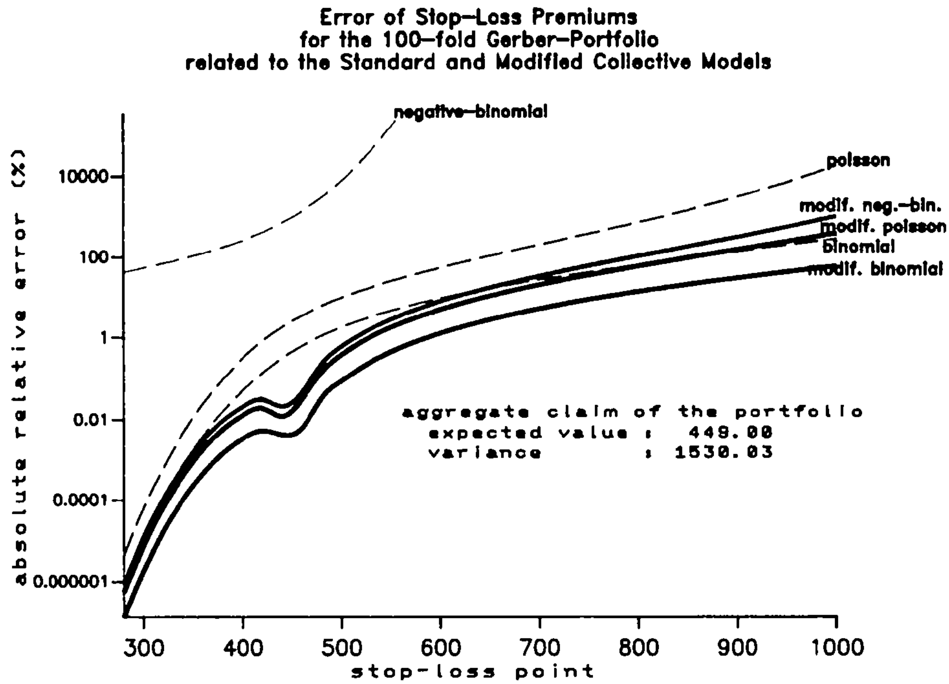


TABLE 1

STANDARD COLLECTIVE MODELS FOR THE GERBER PORTFOLIO
ERROR OF THE STOP LOSS PREMIUM
(WITHOUT ANY LOADING)

Security Level of Percentile Premium	Stop Loss Point	Stop Loss Premium in the Ind Mod	Error in % of the Stop Loss Premium Binomial	Poisson	Neg Binomial
50	4	1 776	0 16	1 68	35 52
60	5	1 340	0 37	2 62	52 95
70	6	1 001	0 54	3 68	75 10
80	8	0 515	1 25	6 92	146 61
90	10	0 251	2 35	11 39	268 31
95	12	0 113	4 28	17 97	492 38
99	16	0 019	9 87	37 51	1725 27

TABLE 2

MODIFIED COLLECTIVE MODELS FOR THE GERBER PORTFOLIO
ERROR OF THE STOP LOSS PREMIUM
(WITHOUT ANY LOADING)

Security Level of Percentile Premium	Stop Loss Point	Stop Loss Premium in the Ind Mod	Error in % of the Stop Loss Premium		
			Binomial	Poisson	Neg Binomial
50	4	1 776	0 15	0 05	0 03
60	5	1 340	0 10	0 45	0 57
70	6	1.001	0 12	0 38	0 27
80	8	0 515	0 06	1 85	2 10
90	10	0 251	0 44	3 71	3 95
95	12	0 113	1 42	6 81	8 90
99	16	0 019	4 31	15 89	24 79

TABLE 3

STANDARD COLLECTIVE MODELS FOR THE 100-FOLD GERBER PORTFOLIO
ERROR OF THE STOP LOSS PREMIUM
(WITHOUT ANY LOADING)

Security Level of Percentile Premium	Stop Loss Point	Stop Loss Premium in the Ind Mod	Error in % of the Stop Loss Premium		
			Binomial	Poisson	Neg Binomial
50	448	16 10	0 44	2 46	951 24
60	458	11 57	0 61	3 38	1332 18
70	469	7 70	0 84	4 66	1999 85
80	482	4 49	1 19	6 56	3406 36
90	499	1 99	1 80	9 81	7503 74
95	514	0 88	2 47	13 48	16554 90
99	543	0 14	4 22	23 18	99879 00

TABLE 4

MODIFIED COLLECTIVE MODELS FOR THE 100-FOLD GERBER PORTFOLIO
ERROR OF THE STOP LOSS PREMIUM
(WITHOUT ANY LOADING)

Security Level of Percentile Premium	Stop Loss Point	Stop Loss Premium in the Ind Mod	Error in % of the Stop Loss Premium		
			Binomial	Poisson	Neg Binomial
50	448	16 10	0 00	0 00	0 01
60	458	11 57	0 00	0 03	0 04
70	469	7 70	0 02	0 08	0 12
80	482	4 49	0 04	0 17	0 28
90	499	1 99	0 09	0 38	0 59
95	514	0 88	0 16	0 67	1 05
99	543	0 14	0 38	1.51	2 44

APPENDIX

We start with a selection of results which are contained in a paper written by VON CHOSSY, R. and G. RAPPL (1983). Let

$$S_n^{\text{coll}} \sum_{i=1}^{N_n} Y_i, \qquad n \in \mathbb{N}, \tag{A.1}$$

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 and $(Y_i)_{i \in \mathbb{N}}$ are supposed to be independent for each $n \in \mathbb{N}$. Furthermore, the second moments of Y_1 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 $n \in \mathbb{N}$

$$N_n \overset{\mathcal{L}}{\sim} K^{*n}. \tag{A 2}$$

Further, let

$$\tilde{F} = \int_{\mathbb{N}_0} F^{*k}\, K(dk), \qquad Y_1 \overset{\mathcal{L}}{\sim} F. \tag{A.3}$$

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

$$S_n^{\text{coll}} \overset{\mathcal{L}}{=} \sum_{i=1}^{n} \tilde{Y}_i \qquad \left(\text{i.e. } \tilde{Y}_1 \overset{\mathcal{L}}{=} \sum_{i=1}^{N_n} Y_i\right) \tag{A.4}$$

for all $n \in \mathbb{N}$.

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 $(S_n - ES_n)/\sqrt{\operatorname{Var} S_n}$ converges in distribution to a standard normal distributed random variable as $n \to \infty$, i e. $|F_n(x) - \Phi_n(x)| \xrightarrow[n \to \infty]{} 0$ uniformly in x with $F_n(x) = P(S_n \leq x)$ and $\Phi_n(x) = \Phi((x - ES_n)/\sqrt{\operatorname{Var} S_n})$, Φ the standard normal distribution function

From Proposition A.1 and the classical central limit theorem (cf. FELLER (1971), p. 515), VON CHOSSY, R. and G. RAPPL (1983, p. 254) deduce directly·

Proposition A.2. Under the assumptions of Proposition A 1 with $\operatorname{Var} \tilde{Y}_1 > 0$, the central limit theorem is valid for the sequence $(S_n^{\text{coll}} - ES_n^{\text{coll}})/\sqrt{\operatorname{Var} S_n^{\text{coll}}}$.

Using the standard Berry-Esséen inequality (cf. FELLER (1971), p 542), both authors proved, in addition, a Berry-Esséen bound for special random sums.

Proposition A.3. Let the assumptions of Proposition A.1 be fulfilled; further, let $\operatorname{Var} \tilde{Y}_1 > 0$, $E|Y_1|^3 < \infty$, $EN_n^3 < \infty$. Denote by F_n^{coll} the distribution function of S_n^{coll} and put

$$\Phi_n^{\text{coll}}(x) = \Phi((x - ES_n^{\text{coll}})/\sqrt{\operatorname{Var} S_n^{\text{coll}}}), \tag{A 5}$$

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

$$\sup_x |F_n^{\text{coll}}(x) - \Phi_n^{\text{coll}}(x)| \leq \frac{3}{\sqrt{n}} \frac{E|\tilde{Y}_1 - E\tilde{Y}_1|^3}{(\operatorname{Var} \tilde{Y}_1)^{3/2}} \tag{A.6}$$

Furthermore, it holds

$$\operatorname{Var} \tilde{Y}_1 = \mu_1 \operatorname{Var} Y_1 + \sigma^2 (EY_1)^2, \tag{A.7}$$

where

$$\mu_1 = \int_{\mathbb{N}_0} kK(dk), \qquad \sigma^2 = \int_{\mathbb{N}_0} (k - \mu_1)^2 K(dk), \tag{A.8}$$

and

$$E|\tilde{Y}_1 - E\tilde{Y}_1|^3 \leq 4 \left((\rho_3 - 3\rho_2\rho_1 + 2\rho_1^3)\mu_1 + (3\rho_2\rho_1 - 3\rho_1^3)\mu_2 + \rho_1^3\mu_3 + |EY_1|^3 \int_{\mathbb{N}_0} |k - \mu_1|^3 K(dk) \right), \tag{A.9}$$

where

$$\mu_i = \int_{\mathbb{N}_0} k^i K(dk), \quad \rho_i = \int_{\mathbb{R}} |x - EY_1|^i F(dx), \quad i = 1, 2, 3. \tag{A.10}$$

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-Esséen 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 $\sum_{i=1}^{n} X_i$, $\Phi_n(x) = \Phi((x - m_n) s_n^{-1})$, Φ the standard normal distribution function. Then for all $n \in \mathbb{N}$

$$\text{(A.11)} \qquad \sup_x |F_n(x) - \Phi_n(x)| \leq 6 r_n s_n^{-3}.$$

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

$$\text{(A 12)} \quad \tilde{\Delta}_n = \sup_x |\Phi_n^{\text{ind}}(x) - \Phi_n^{\text{coll}}(x)|.$$

Then we have

$$\text{(A.13)} \quad \tilde{\Delta}_n = O(1/\sigma_n^{\text{coll}}) \Leftrightarrow \sigma_n^{\text{coll}} - \sigma_n^{\text{ind}} = O(1) \text{ and } \mu_n^{\text{coll}} - \mu_n^{\text{ind}} = O(1).$$

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

$$\text{(A.14)} \qquad \tilde{\Delta}_n = \sup_x \left| \Phi'(\xi_n) \left(\frac{x - a_n}{b_n} - \frac{x - a'_n}{b'_n} \right) \right|,$$

where

$$\text{(A.15)} \quad \xi_n = \alpha_n \frac{x - a_n}{b_n} + (1 - \alpha_n) \frac{x - a'_n}{b'_n}, \qquad \alpha_n \in (0, 1), \qquad x \in \mathbb{R}$$

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

$$\text{(A.16)} \qquad \xi_n = \alpha_n c + (1 - \alpha_n) \left(\frac{a_n - a'_n}{b'_n} + c \frac{b_n}{b'_n} \right) \xrightarrow[n \to \infty]{} c.$$

In addition, we have

$$\text{(A.17)} \qquad \frac{x_n - a_n}{b_n} - \frac{x_n - a'_n}{b'_n} = c - \left(\frac{a_n - a'_n}{b_{n'}} + c\frac{b_n}{b'_n} \right)$$

$$= c\,\frac{b'_n - b_n}{b'_n} + \frac{a'_n - a_n}{b_{n'}}\,.$$

Consequently, and because $\tilde{\Delta}_n = O(1/b'_n)$, $\xi_n \to c \neq 0$, $\Phi'(c) > 0$, 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 $r_n = b_n/b'_n$

$$\text{(A.18)} \qquad \tilde{\Delta}_n = \sup_x \left| \Phi'\left(\alpha_n \left(\frac{x - a_n}{b_n} r_n + \frac{a_n - a'_n}{b'_n} \right) + (1 - \alpha_n) \frac{x - a_n}{b_n} \right) \right.$$

$$\left. \times \left(\frac{x - a_n}{b_n}(r_n - 1) + \frac{a_n - a'_n}{b'_n} \right) \right|.$$

Therefore, according to the assumptions we have

$$\text{(A.19)} \qquad b'_n \tilde{\Delta}_n = \sup_x \left| \Phi'\left(\alpha_n \frac{a_n - a'_n}{b'_n} \right) \Phi'\left(\frac{x - a_n}{b_n} (1 - \alpha_n(1 - r_n)) \right) \right.$$

$$\left. \times \left(\frac{x - a_n}{b_n} (b_n - b'_n) + (a_n - a'_n) \right) \right|$$

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

Lemma A.2.

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

$$\text{(A.20)} \qquad \frac{\mathcal{S}_n^{\text{ind}}(\alpha) - \mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}} \xrightarrow[n \to \infty]{} \Phi^{-1}(\alpha), \quad \frac{\mathcal{S}_n^{\text{coll}}(\alpha) - \mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}} \xrightarrow[n \to \infty]{} \Phi^{-1}(\alpha),$$

(ii) F_n^{ind} fulfills the Berry-Esséen bound from Theorem A.1, F_n^{coll} that from Proposition A.3

If $\sigma_n^{\text{ind}}/\sigma_n^{\text{coll}} \xrightarrow[n\to\infty]{} 1$, then

(A.21) $$\frac{\mathcal{S}_n^{\text{ind}}(\alpha)-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}} - \frac{\mathcal{S}_n^{\text{coll}}(\alpha)-\mu_n^{\text{coll}}}{\sigma_n^{\text{coll}}} = O(1/\sigma_n^{\text{ind}}), \quad \alpha \in (0, 1).$$

Proof.

(i) Put

(A.22) $$\gamma_n^{\text{ind}}(\alpha) = \frac{\mathcal{S}_n^{\text{ind}}(\alpha)-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}}, \; \gamma_n^{\text{coll}}(\alpha) \text{ analogously}.$$

Then

$$\Phi(\gamma_n^{\text{ind}}(\alpha)) = \Phi_n^{\text{ind}} \circ (F_n^{\text{ind}})^{-1} \circ F_n^{\text{ind}}(\inf\{p \in [0, \infty) | F_n^{\text{ind}}(p) \geq \alpha\}),$$

(A.23) $$F_n^{\text{ind}}(\inf\{p \in [0, \infty) | F_n^{\text{ind}}(p) \geq \alpha\}) \xrightarrow[n\to\infty]{} \alpha,$$

$$\Phi_n^{\text{ind}} \circ (F_n^{\text{ind}})^{-1} \xrightarrow[n\to\infty]{} id$$

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 $\mathcal{S}_n^{\text{ind}}(\alpha) = (F_n^{\text{ind}})^{-1}(\alpha)$ we obtain

(A.24) $$\sup_x |F^{\text{ind}}(x) - \Phi_n^{\text{ind}}(x)| = \sup_{y \in \mathcal{M}} |y - \Phi_n^{\text{ind}} \circ (F_n^{\text{ind}})^{-1}(y)|$$

where $\mathcal{M} = \{y | y = F_n^{\text{ind}}(x)\}$

Using the mean value theorem, we have with suitably chosen $\alpha_n \in (0, 1)$

(A.25) $$y - \Phi_n^{\text{ind}} \circ (F_n^{\text{ind}})^{-1}(y) = \Phi \circ \Phi^{-1}(y) - \Phi\left(\frac{(F_n^{\text{ind}})^{-1}(y)-\mu_n^{\text{ind}}}{\sigma_n^{\text{ind}}}\right)$$
$$= \Phi'(\alpha_n \Phi^{-1}(y) + (1-\alpha_n)\gamma_n^{\text{ind}}(y))\,(\Phi^{-1}(y) - \gamma_n^{\text{ind}}(y)).$$

Because of (i),

(A.26) $$\Phi'(\alpha_n \Phi^{-1}(y) + (1-\alpha_n)\gamma_n^{\text{ind}}(y)) \xrightarrow[n\to\infty]{} \Phi' \circ \Phi^{-1}(y) \text{ uniformly in } y.$$

$\Phi' \circ \Phi^{-1}(y) > 0$ and the Berry-Esséen bound now supply

(A 27) $$\Phi^{-1}(y) - \gamma_n^{\text{ind}}(y) = O(1/\sigma_n^{\text{ind}}).$$

The same argumentation applied to F_n^{coll} and Φ_n^{coll} yields

(A.28) $$\Phi^{-1}(y)-\gamma_n^{\text{coll}}(y) = O(1/\sigma_n^{\text{coll}}).$$

Finally, the assertion follows from the last two bounds taking $\sigma_n^{\text{ind}}/\sigma_n^{\text{coll}} \to 1$ into account.

REFERENCES

BOWERS, N L, GERBER, H U, HICKMAN, J C, JONES, D A and NESBITT, C J (1986) *Actuarial Mathematics* Society of Actuaries, Itasca

DE PRIL, N (1989) The Aggregate Claims Distribution in the Individual Model with Arbitrary Positive Claims *ASTIN Bulletin* **19**, 9–24

FELLER, W (1971) *An Introduction to Probability Theory and its Applications* **Vol. 2**, 2nd ed Wiley, New York

GERBER, H U (1979) *An Introduction to Mathematical Risk Theory* Huebner Foundation for Insurance Education, Philadelphia

GERBER, H U (1984) Error Bounds for the Compound Poisson Approximation *Insurance Mathematics and Economics* **3**, 191–194

HIPP, C (1985) Approximation of Aggregate Claims Distributions by Compound Poisson Distributions *Insurance Mathematics and Economics* **4**, 227–232

HIPP, C (1986) Improved Approximation for Aggregate Claims Distributions in the Individual Model *ASTIN Bullletin* **16**, 89–100

JEWELL, W S and SUNDT, B (1981) Improved Approximations for the Distribution of a Heterogeneous Risk Portfolio *Mitt Ver Schweiz Vers Math*, **1981**, 221–240

KORNYA, P S (1983) Distribution of Aggregate Claims in the Individual Risk Theory Model *Transactions of the Society of Actuaries* **35**, 823–836 Discussion 837–858

PANJER, H.H (1981) Recursive Evaluation of a Family of Compound Distributions *ASTIN Bulletin* **12**, 22–26

SUNDT, B (1985) On Approximations for the Distribution of a Heterogenous Risk Portfolio *Mitt Ver Schweiz Vers Math*, **1985**, 189–203

VON CHOSSY, R and RAPPL, G (1983) Some Approximation Methods for the Distribution of Random Sums *Insurance Mathematics and Economics* **2**, 251–270

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