Applications and Applied Mathematics: An International Journal (AAM)

Volume 16 | Issue 2 | Article 5

12-2021

(R1505) A Note on Large Deviations in Insurance Risk

Stefan Gerhold
TU Wien

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Vol. 16, Issue 2 (December 2021), pp. 867 – 880

Applications and Applied Mathematics:
An International Journal
(AAM)

A Note on Large Deviations in Insurance Risk

Stefan Gerhold

TU Wien
Wiedner Hauptstr. 8–10
1140 Vienna, Austria
sgerhold@fam.tuwien.ac.at

Received: May 31, 2021; Accepted: September 1, 2021

Abstract

We study large and moderate deviations for an insurance portfolio, with the number of claims tending to infinity, without assuming identically distributed claims. The crucial assumption is that the centered claims are bounded and that variances are bounded below. From a general large deviations upper bound, we obtain an exponential bound for the probability of the average loss exceeding a threshold. A counterexample shows that a full large deviation principle, including also a lower bound, does not follow from our assumptions. We argue that our assumptions make sense, in particular, for life insurance portfolios and discuss how to apply our upper bound in this context. Finally, we use a moderate deviations result by Petrov (1954) to estimate the probability of exceeding a threshold that depends on portfolio size. In this asymptotic regime, the rate function that determines the asymptotic behavior is explicit and thus very easy to compute numerically without solving an optimization problem.

Keywords: Insurance risk; Individual model; Large deviations; Cramér's theorem; Moderate deviations; Law of large numbers; Life insurance

MSC 2010 No.: 60F10, 91B30

1. Introduction

A basic setting to discuss insurance risk is a portfolio of n claims, modelled by integrable independent random variables. Often, the claims are assumed to be identically distributed. The significance of the law of large numbers (LLN) and the central limit theorem (CLT) in this context has been amply discussed in the literature (Albrecht (1982); Cummins (1974); Smith and Kane (1994)). If a premium that exceeds the expectation by some constant is charged for each claim, then the probability that the aggregated claims cannot be covered tends to zero as portfolio size tends to infinity. This follows immediately from the weak law of large numbers. It is a natural question whether the convergence speed is fast enough to make this probability sufficiently small in practice. Fortunately, Cramér's theorem (Theorem 2.2.3 in Dembo and Zeitouni (1998)) states that this probability decays exponentially fast. The original version of Cramér's theorem, requiring finite exponential moments, was published in 1938 (see Cramér (1938)), and became a cornerstone of the theory of large deviations.

However, the assumption of identical claim distributions is not always satisfied in practice. In life insurance it often makes no sense because the claim distributions depend significantly on several parameters including type of insurance, amount insured, age, and time to expiry. Thus, dividing a portfolio into sub-portfolios with roughly identically distributed claims will usually lead to rather small portfolios that make the limit of portfolio size tending to infinity questionable. The importance of considering non-identical claim distributions is also stressed in Albrecht (1982) and Cummins (1974). The LLN can still be applied, under mild assumptions, e.g., bounded claim variances, but does not say anything about the convergence speed.

To the best of the author's knowledge, the exponential decrease of the average loss probability for non-identical distributions has not been discussed in the insurance literature. This seems worthwhile, because a slow convergence of the loss probability could significantly hamper the practical applicability of the LLN to this problem. The aim of this note is to show that exponential decay persists under mild assumptions, and to discuss the validity of a large deviation principle. The assumption of independent claims is in force throughout this note. In life insurance, independence requires deterministic interest rates and a fixed life table. For risk diversification, and its limits, under stochastic mortality we refer to Milevsky et al. (2006), and for stochastic financial markets to Fischer (2007) and the references therein.

There is a considerable literature on large deviations for compound sums and more sophisticated models in risk theory, but apparently not for the basic individual risk model with non-identical distributions that we consider here. In practice, premia and reserves are calculated for each contract separately, i.e., using an individual model. For computing the distribution of the aggregate loss numerically, e.g., to compute value at risk, a standard approach is to pass to a collective model, which is numerically more tractable. For an asymptotic approximation of the loss probability, which is our goal, such a change of model is not required and seems unnatural. Large deviations for an individual model of credit and insurance risk are also studied in Dembo et al. (2004), but their assumptions and results are quite different from ours.

The rest of this note is structured as follows. From a practical viewpoint, our main result is Theorem 3.1, which shows that a mild assumption on the claim distributions yields an exponentially small upper bound, which is weaker than a full large deviation principle (LDP), but should suffice for practical purposes. In Section 4 we discuss the application of this bound to a life insurance portfolio. In Theorem 5.1, we show that our main assumption does not suffice to establish an LDP for the empirical mean, i.e., lower and upper estimates of the same exponential order. Theorem 5.2 adds a more restrictive assumption, which implies an LDP. Finally, Corollary 6.1 establishes moderate deviation estimates.

2. Basic assumptions

We consider a portfolio of n claims, modelled by integrable independent random variables $Y_1, \dots, Y_n$, and causing a total claim amount

$$S_n = Y_1 + \cdots + Y_n.$$

First, suppose that the claims are identically distributed. If a premium $\mathbb{E}[Y_k] + \delta$, with $\delta > 0$, is charged for the kth claim, then the probability that the aggregated claims cannot be covered tends to zero as portfolio size tends to infinity. Indeed, the weak law of large numbers implies that

$$\lim_{n\to\infty} \mathbb{P}\big[S_n \geq n(\mathbb{E}[Y_1] + \delta)\big] = 0. \tag{1}$$

According to Cramér's theorem (Theorem 2.2.3 in Dembo and Zeitouni (1998)) the probability in (1) decays exponentially fast. Explicitly, with

$$I(x) := \sup_{u\in\mathbb{R}} \big(ux - \log \mathbb{E}[e^{uY_1}]\big),$$

we have

$$\mathbb{P}\big[S_n \geq n(\mathbb{E}[Y_1] + \delta)\big] = \exp\Big(-I(\mathbb{E}[Y_1] + \delta)n + \mathrm{o}(n)\Big). \tag{2}$$

This result shows that, for realistic portfolio sizes, the probability in (1) is extremely small.

As mentioned in the introduction, for life insurance, the assumption that the claims have identical distributions should be dropped. Then, (1) becomes

$$\lim_{n\to\infty} \mathbb{P}\Big[S_n \geq \sum_{k=1}^{n} \mathbb{E}[Y_k] + \delta n\Big] = 0. \tag{3}$$

By the weak law of large numbers, this convergence holds under mild assumptions, e.g., bounded claim variances. It seems that the insurance literature offers no analogue of (2) for non-identical distributions. The aim of this note is to show that exponential decay persists under mild assumptions, and to discuss the validity of a large deviation principle.

Assumption A

(i) $(X_k)_{k\in\mathbb{N}}$ is a sequence of independent centered real random variables,
(ii) there is $c_0 > 0$ such that $|X_k| \leq c_0$ for all k,
(iii) there is $c_1 > 0$ such that $\mathrm{Var}[X_k] = \mathbb{E}[X_k^2] \geq c_1$ for all k.

This assumption will be imposed on the centered claims

$$X_k = Y_k - \mathbb{E}[Y_k].$$

Similar assumptions are made in Albrecht (1982) (p. 515) and Cummins (1974) (p. 153). We now discuss the validity of Assumption A in life insurance. Assuming independence neglects certain risks, such as epidemics and natural disasters, but still seems reasonable for large portfolios. As is customary in the practice of life insurance, we assume that interest rate risk, which affects all contracts simultaneously, is handled by using sufficiently conservative deterministic yield curves for discounting. Thus, no dependence due to stochastic interest rates is introduced. Part (ii) makes sense, as insurers usually prescribe an upper limit on the possible amount insured, and annuity payments due after age 150, say, can be neglected. As for (iii), note that clearly we may assume $\mathrm{Var}[X_k] > 0$, because it makes no sense to include contracts with no remaining random cash flows. Then, since there is usually a lowest possible amount insured, and there are only finitely many value combinations for the parameters age, time to expiry, sex, and type of insurance, a uniform lower bound on the claim variance is natural. Of course, for continuous-time models, which are not widespread in practice anyways, this applies only after time discretization. While our main motivation comes from life insurance, Assumption A makes sense for many non-life insurance portfolios as well.

3. Large deviations: an upper bound

For practical purposes, an upper bound for the probablity in (3) is much more important than a lower bound. We now show that the – rather weak – Assumption A implies an exponential upper estimate. Let

$$M_n = \frac{1}{n}\sum_{k=1}^{n} X_k,$$

denote the empirical mean of the centered claims $X_k = Y_k - \mathbb{E}[Y_k]$.

Theorem 3.1.

Under Assumption A, there exists a positive function $J : (0,\infty) \to (0,\infty)$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}[M_n \geq x] \leq -J(x), \quad x > 0. \tag{4}$$

Proof:

We apply the general LD upper bound from Theorem 4.5.20 in Dembo and Zeitouni (1998). Define

$$\bar{\Lambda}(\lambda) := \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda n M_n}] = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log \mathbb{E}[e^{\lambda X_k}].$$

Since $|M_n| \leq c_0$ is bounded, the sequence of its laws is exponentially tight (definition on p. 8 of Dembo and Zeitouni (1998)). Thus, part (a) of Theorem 4.5.20 in Dembo and Zeitouni (1998) implies

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}[M_n \geq x] \leq - \inf_{y \geq x} \bar{\Lambda}^*(y) =: -J(x),$$

where

$$\bar{\Lambda}^*(x) := \sup_{\lambda \in \mathbb{R}} \big(\lambda x - \bar{\Lambda}(\lambda)\big), \quad x \in \mathbb{R},$$

is the Fenchel-Legendre transform of $\bar{\Lambda}$. The key point now is to show that J is positive under our assumptions, because otherwise (4) would be of little use. By Assumption A,

$$\mathbb{E}[e^{\lambda X_k}] = 1 + \tfrac{1}{2}\mathbb{E}[X_k^2]\lambda^2 + \mathrm{O}(\lambda^3), \quad \lambda \to 0,$$

where the error term $\mathrm{O}(\lambda^3)$ is uniform with respect to k. Hence, using the Taylor expansion

$$\log(1+y) = y + \mathrm{O}(y^2), \quad y \to 0,$$

we obtain

$$\begin{aligned} \log \mathbb{E}[e^{\lambda X_k}] &= \tfrac{1}{2}\mathbb{E}[X_k^2]\lambda^2 + \mathrm{O}(\lambda^3) \\ &\leq \tfrac{1}{2} c_0^2 \lambda^2 + \mathrm{O}(\lambda^3), \end{aligned}$$

and thus $\bar{\Lambda}(\lambda) \leq c_0^2 \lambda^2$ for small λ. Define the function

$$\Theta(\lambda) := \max\{c_0\lambda^2, \bar{\Lambda}(\lambda)\}, \quad \lambda \in \mathbb{R}.$$

Since $\bar{\Lambda}$ is convex (see Theorem 4.5.3 (a) in Dembo and Zeitouni (1998)), Θ is convex, too. Its Fenchel-Legendre transform Θ^* satisfies $\Theta^*(0) = 0$, is non-negative, convex, and strictly convex in a neighborhood of zero. By definition, we have $\Theta^* \leq \bar{\Lambda}^*$, which implies $J(x) > 0$ for $x > 0$. ■

4. Life insurance

Consider a portfolio of n whole life insurance contracts to which we want to apply Theorem 3.1, with similar remarks applying to other kinds of life insurance. Let x_k denote the age of the kth customer at inception of her contract, and t_k the time that has passed since. For simplicity, we suppose that x_k and t_k are integers. With similar notation as in Section 6.3 of Gerber (1997), we write q_x for the probability of death within one year for age x, and ${}_tV_x$ for the reserve after t years. If the sum insured is s_k, payable at the end of the year of death, then the risk premium of the kth contract for the following year is

$$\pi_k^{\mathrm{r}} := (s_k - {}_{t_k+1}V_{x_k}) v q_{x_k+t_k},$$

where v is the one-year discount factor. This is the part of the premium that covers the net risk. Indeed, at the end of the year the savings premium has increased or decreased the reserve from ${}_{t_k}V_{x_k}$ to ${}_{t_k+1}V_{x_k}$, and the risk premium covers the difference of the death benefit to the available funds. The natural definition of the kth claim in this setting is

$$Y_k = \begin{cases} (s_k - {}_{t_k+1}V_{x_k})v, & \text{customer dies in the next year,} \\ 0, & \text{otherwise.} \end{cases}$$

If $Y_k > 0$, then Y_k is the amount not covered by the reserve ${}_{t_k+1}V_{x_k}$. The expection of the claim equals the risk premium,

$$\mathbb{E}[Y_k] = \pi^{\mathrm{r}}_k.$$

Thus, we expect that the risk premium suffices to settle the claim Y_k on average. The event

$$\Big\{ \sum_{k=1}^n Y_k \geq \sum_{k=1}^n \mathbb{E}[Y_k] + n\delta \Big\}, \tag{5}$$

$\delta > 0$ fixed, means that the total claim amount for the next year exceeds the total risk premia by an amount of $n\delta$, so that the portfolio suffers a large loss. Note that the distribution of Y_k depends on the three parameters (s_k, t_k, x_k). In practice, contracts often have a limited term, which adds a fourth parameter. Taking the meaning of these parameters into account, we see that hundreds or rather thousands of different distributions arise for a reasonably sized insurance portfolio. This clearly shows that assuming identical distributions would be very unrealistic.

The moment generating function of the kth centered claim is

$$\mathbb{E}[e^{\lambda X_k}] = q_{x_k+t_k} \exp\big(\lambda \pi^{\mathrm{r}}_k / q_{x_k+t_k} - \lambda \pi^{\mathrm{r}}_k\big).$$

Define $\bar{\Lambda}$ as in the proof of Theorem 3.1, i.e.,

$$\bar{\Lambda}(\lambda) = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \Big(\lambda \pi^{\mathrm{r}}_k (1/q_{x_k+t_k} - 1) + \log q_{x_k+t_k} \Big). \tag{6}$$

The probability of the event (5) equals

$$\mathbb{P}\Big[\sum_{k=1}^n Y_k \geq \sum_{k=1}^n \mathbb{E}[Y_k] + n\delta \Big] = \mathbb{P}[M_n \geq \delta].$$

By Theorem 3.1,

$$\frac{1}{n} \log \mathbb{P}[M_n \geq \delta] \leq -J(\delta) + \mathrm{o}(1), \quad n \to \infty,$$

where J is positive and defined as in the proof of the theorem, using (6). Rearranging yields

$$\begin{aligned} \mathbb{P}\Big[\sum_{k=1}^n Y_k \geq \sum_{k=1}^n \mathbb{E}[Y_k] + n\delta \Big] &= \mathbb{P}[M_n \geq \delta] \\ &\leq \exp\Big(-J(\delta) n \big(1 + \mathrm{o}(1)\big) \Big) = \exp\Big(-J(\delta) n + \mathrm{o}(n) \Big), \end{aligned}$$

which yields the desired exponential decay estimate for the loss probability.

5. Large deviation principle

Recall that a sequence of random variables $(Z_n)_{n\in\mathbb{N}}$ satisfies the LDP (large deviation principle) with good rate function I and speed $s(n)$, if

(i) $I : \mathbb{R} \to [0, \infty]$ is not infinite everywhere, and the level sets $\{x : I(x) \leq c\}$, $c \in [0, \infty)$, are compact. In particular, I is lower semi-continuous.

(ii) $s(n) > 0$ satisfies $\lim_{n\to\infty} s(n) = \infty$.
(iii) For any Borel set G,

$$-I(\operatorname{int}(G)) \le \liminf_{n\to\infty} \frac{1}{s(n)} \log P[Z_n \in G] \le \limsup_{n\to\infty} \frac{1}{s(n)} \log P[Z_n \in G] \le -I(\operatorname{cl}(G)),$$

where $I(A) := \inf_{x\in A} I(x)$ for any $A \subseteq \mathbb{R}$, and int and cl denote interior and closure, respectively.

We first give a counterexample (in Theorem 5.1) that shows that Assumption A does not imply an LDP for the empirical means. In particular, this shows that the Gärtner-Ellis theorem, a standard LDP tool for non-identically distributed sequences, is not applicable here without additional assumptions, such as Assumption B below.

Let $K_1 \subset \mathbb{N}$ be a set of natural numbers with lower density 0 and upper density 1, i.e.,

$$\nu_1(n) := \#\{1 \le k \le n : k \in K_1\},$$

where $\#$ denotes the cardinality of a set, satisfies

$$\liminf_{n\to\infty} \frac{\nu_1(n)}{n} = 0 \quad \text{and} \quad \limsup_{n\to\infty} \frac{\nu_1(n)}{n} = 1.$$

For the existence of such a set, see, e.g., Theorem 3 in Strauch and Tóth (1998). Define $K_2 := \mathbb{N} \setminus K_1$ and $\nu_2(n) := n - \nu_1(n)$.

Theorem 5.1.

Let $X^{(1)}$ be a random variable that takes the values $-1, 1$ with probability $\frac{1}{2}$ each, and $X^{(2)}$ analogously with values $-2, 2$. Let $(X_k)_{k\in\mathbb{N}}$ be a sequence of independent random variables satisfying

$$X_k \overset{\mathrm{d}}{=} X^{(i)}, \quad k \in K_i,\ i = 1, 2.$$

This sequence satisfies Assumption A, and the sequence of empirical means $M_n = \frac{1}{n}\sum_{k=1}^n X_k$ does not satisfy an LDP.

We defer the proof of this theorem to the appendix. Recall that for any exponentially tight sequence, an LDP holds along a subsequence (see Theorem 3.7 in Feng and Kurtz (2006)). Theorem 5.1 provides an example where it is proven that the whole sequence does *not* satisfy any LDP. By Cramér's theorem, in Theorem 5.1 the section means

$$M_n^{(i)} := \frac{1}{\nu_i(n)} \sum_{\substack{k=1 \\ k\in K_i}}^{n} X_k, \quad i = 1, 2, \tag{7}$$

satisfy LDPs with rate functions $I^{(1)}, I^{(2)}$, explicitly given in (14) below, and speeds $\nu_1(n), \nu_2(n)$. Since the moment generating function of $X^{(2)}$ dominates that of $X^{(1)}$, the upper estimate in

$$-I^{(1)}(x) \le \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}[M_n > x] \le \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}[M_n > x] \le -I^{(2)}(x), \quad x > 0, \tag{8}$$

easily follows from the general upper LD bound we used in the proof of Theorem 3.1. The lower estimate (8) then follows from

$$\mathbb{P}[M_n > x] \geq \mathbb{P}\big[M_n^{(1)} > x,\ M_n^{(2)} > x\big] = \mathbb{P}\big[M_n^{(1)} > x\big]\mathbb{P}\big[M_n^{(2)} > x\big].$$

Thus, we have exponential lower and upper bounds, but the highly irregular interlacement of two distributions in Theorem 5.1 precludes a single rate function governing both. When such behavior is explicitly forbidden, we can actually obtain a full LDP, using the Gärtner-Ellis theorem.

Assumption B

(i) There is a partition

$$\mathbb{N} = N_1 \cup \cdots \cup N_p\ ,$$

such that for all $1 \leq i \leq p$ and $k \in N_i$, the law of $X_k \stackrel{\mathrm{d}}{=} X^{(i)}$ is independent of k. We write φ_i for the corresponding moment generating function $\varphi_i(\lambda) = \mathbb{E}[\exp(\lambda X^{(i)})]$.
(ii) For each i, the limit

$$d_i := \lim_{n\to\infty} \frac{1}{n} \#\{1 \leq k \leq n : k \in N_i\},$$

exists.

Theorem 5.2.

Under Assumptions A and B, the sequence of empirical means $(M_n)_{n\in\mathbb{N}}$ satisfies an LDP with good rate function

$$\Lambda^*(x) := \sup_{\lambda\in\mathbb{R}} \big(\lambda x - \Lambda(\lambda)\big), \quad x \in \mathbb{R},$$

the Fenchel-Legendre transform of

$$\Lambda(\lambda) := \sum_{i=1}^{p} d_i \log \varphi_i(\lambda). \tag{9}$$

Proof:

This result is an easy consequence of the Gärtner-Ellis theorem (Theorem 2.3.6 in Dembo and Zeitouni (1998)). Indeed, here the function Λ from Assumption 2.3.2 in Dembo and Zeitouni (1998) equals

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda n M_n}] = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log \mathbb{E}[e^{\lambda X_k}].$$

By Assumption B, $\mathbb{E}[e^{\lambda X_k}] = \varphi_i(\lambda)$ for $k \in N_i$, and so this function further equals

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{p} \sum_{\substack{k=1 \\ k\in N_i}}^{n} \log \varphi_i(\lambda) = \sum_{i=1}^{p} d_i \log \varphi_i(\lambda),$$

which agrees with (9). As the X_k are bounded by Assumption A, the domain of Λ is $\mathbb{R}$. By Remark (c) on p. 45 of Dembo and Zeitouni (1998), it is thus not necessary to verify the so-called steepness of Λ. Since moment generating functions are smooth on their domain, so is Λ. Therefore, all assumptions of the Gärtner-Ellis theorem are satisfied. ■

6. Moderate deviations

As above, we write $M_n = \frac{1}{n}\sum_{k=1}^n X_k$ for the empirical mean of a sequence of centered random variables. When x in $\mathbb{P}[M_n \geq x]$ is allowed to depend on n, and $n^{-1/2} \ll x \ll 1$, we are in a regime in between of the CLT and LD scalings, which is known as moderate deviations regime. This viewpoint allows to deduce lower and upper bounds for $\mathbb{P}[M_n \geq x]$ without the somewhat awkward Assumption B. We need the following result from Petrov (1954), which is also presented in detail as Theorem 1.1 in Petrov and Robinson (2008).

Let $(X_k)_{k\in\mathbb{N}}$ be a sequence of independent centered random variables such that there are positive numbers g, G, H with

$$g \leq \big|\mathbb{E}[e^{hX_k}]\big| \leq G \ \text{ in the complex circle } |h| < H,\ k \in \mathbb{N}. \tag{10}$$

Moreover, suppose that $B_n := \sum_{k=1}^n \mathbb{E}[X_k^2]$ satisfies $\liminf(B_n/n) > 0$. Then, for $1 < y = \mathrm{o}(\sqrt{n})$,

$$\mathbb{P}\Big[B_n^{-1/2}\sum_{k=1}^n X_k > y\Big] = \big(1-\Phi(y)\big)\exp\Big(\frac{y^3}{\sqrt{n}}\lambda_n\Big(\frac{y}{\sqrt{n}}\Big)\Big)\big(1+\mathrm{o}(1)\big), \tag{11}$$

as $n \to \infty$, where Φ is the standard Gaussian cdf, and λ_n is a power series which converges uniformly with respect to n and with coefficients expressible by the cumulants of the X_k.

In Petrov and Robinson (2008), it is mentioned that this is a generalization of Cramér's theorem. Indeed, Theorem 1 in Cramér (1938) treats the scaling on the left hand side of (11) (for the case of identical distributions), whereas the LD scaling result that is nowadays usually called "Cramér's theorem" is Theorem 6 in Cramér (1938).

We now use Petrov's theorem to give a moderate deviations estimate for $\mathbb{P}[M_n \geq x]$. The first estimate, (12), directly follows from the theorem, and thus the scaling involves B_n. The simpler scaling in (13) yields a slightly cruder estimate, in terms of a lower und an upper bound. If the parameter α is close to $\frac{1}{2}$, the regime becomes similar to the LD scaling, which would correspond to $\alpha = \frac{1}{2}$.

Corollary 6.1.

Let $(X_k)_{k\in\mathbb{N}}$ be a sequence of random variables satisfying Assumption A. For $c_2 > 0$, $\alpha \in (0, \frac{1}{2})$, and $B_n = \sum_{k=1}^n \mathbb{E}[X_k^2]$, we have

$$\mathbb{P}\big[M_n > c_2 n^{\alpha-1} B_n^{1/2}\big] = \exp\Big(-\tfrac{1}{2}c_2^2 n^{2\alpha}\big(1+\mathrm{o}(1)\big)\Big). \tag{12}$$

Moreover, with c_0 and c_1 as in Assumption A, the bounds

$$\mathbb{P}\big[M_n > c_2 c_0 n^{\alpha-1/2}\big] \leq \exp\Big(-\tfrac{1}{2}c_2^2 n^{2\alpha}\big(1+\mathrm{o}(1)\big)\Big) \leq \mathbb{P}\big[M_n > c_2 c_1^{1/2} n^{\alpha-1/2}\big], \tag{13}$$

hold.

Proof:

Condition (10) is satisfied with $H = c_0^{-1}$, $g = \frac{1}{2}e^{-c_0 H}$, and $G = e^{c_0 H}$. Indeed, the upper bound is clear, and the lower bound follows from

$$\big|\mathbb{E}[e^{hX_k}]\big| \geq \mathbb{E}\big[e^{\mathrm{Re}(h)X_k}\cos(\mathrm{Im}(h)X_k)\big],$$

and

$$\cos(\mathrm{Im}(h)X_k) \geq 1 - \tfrac{1}{2}\big(\mathrm{Im}(h)X_k\big)^2 \geq 1 - \tfrac{1}{2}(c_0 H)^2 = \tfrac{1}{2}.$$

The condition for B_n follows from part (iii) of Assumption A. We can thus apply Petrov's theorem, with $y = c_2 n^{\alpha}$. The main contribution arises from the factor

$$1 - \Phi(y) = \exp\Big(-\tfrac{1}{2}c_2^2 n^{2\alpha}\big(1 + \mathrm{o}(1)\big)\Big).$$

Since the convergence of λ_n is uniform, we have $\lambda_n(y/\sqrt{n}) = \mathrm{O}(1)$, and thus

$$\frac{y^3}{\sqrt{n}}\lambda_n\Big(\frac{y}{\sqrt{n}}\Big) = \mathrm{O}(n^{3\alpha-1/2}) \ll n^{2\alpha}.$$

This proves (12). For the second assertion, it then suffices to note that Assumption A implies

$$c_1 n \leq B_n \leq c_0^2 n, \quad n \in \mathbb{N}.$$

■

Of course, Petrov's theorem yields further lower order terms in (12), if desired.

7. Conclusion

Our results show that is safe to apply the law of large numbers to the individual model of insurance risk, because the aggregate loss probability converges to zero exponentially fast. This seems to fill a gap in the literature on insurance risk, in which identical claims are usually assumed for assessing this speed. This is of particular importance in life insurance, where identically distributed claims are not a realistic assumption.

We argued that our mathematical assumptions are realistic for life insurance portfolios, except independence, which is only approximately satisfied in practice. Thus, a possible line of future research is to weaken this assumption. A natural extension that introduces dependence would be to make the interest rate stochastic. Presumably, only a much weaker asymptotic result can then be obtained. Moreover, given that interest rates have been kept very low for several years due to fiscal policy, it is currently by no means an easy matter to choose a good stochastic interest rate model. Another possible task is to find a dependence structure that allows a life insurance model to include epidemics, natural disasters and other risk factors that influence many claims simultaneously. Such models have been considered in the literature, but not from the asymptotic viewpoint of the present note.

Acknowledgment:

Financial support from the Austrian Science Fund (FWF) under grant P 30750 is gratefully acknowledged.

REFERENCES

Albrecht, P. (1982). Gesetze der großen Zahlen und Ausgleich im Kollektiv – Bemerkungen zu Grundlagen der Versicherungsproduktion, Zeitschr. f. d. ges. Versicherungsw., Vol. 71, pp. 501–538.

Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités, Colloque consacré à la théorie des probabilités, Actualités scientifques et industrielles, Vol. 736, pp. 2–23. English translation available at arXiv:1802.05988.

Cummins, J.D. (1974). Insurer's risk: A restatement, The Journal of Risk and Insurance, Vol. 41, pp. 147–157.

Dembo, A., Deuschel, J.D. and Duffie, D. (2004). Large portfolio losses, Finance Stoch., Vol. 8, pp. 3–16.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications* (second edition), Springer-Verlag, New York.

Den Hollander, F. (2000). *Large Deviations*, American Mathematical Society, Providence, RI.

Feng, J. and Kurtz, T.G. (2006). *Large Deviations for Stochastic Processes*, American Mathematical Society, Providence, RI.

Fischer, T. (2007). A law of large numbers approach to valuation in life insurance, Insurance Math. Econom., Vol. 40, pp. 35–57.

Gerber, H.U. (1997). *Life Insurance Mathematics* (third edition), Springer-Verlag, Berlin.

Milevsky, M.A., Promislow, S.D. and Young, V.R. (2006). Killing the law of large numbers: Mortality risk premiums and the Sharpe ratio, Journal of Risk and Insurance, Vol. 73, pp. 673–686.

Petrov, V.V. (1954). Generalization of Cramér's limit theorem, Uspehi Matem. Nauk (N.S.), Vol. 9, pp. 195–202.

Petrov, V.V. and Robinson, J. (2008). Large deviations for sums of independent non identically distributed random variables, Comm. Statist. Theory Methods, Vol. 37, pp. 2984–2990.

Smith, M. and Kane, S. (1994). The law of large numbers and the strength of insurance, In *Insurance, Risk Management, and Public Policy* (S.G. Gustavson and S.E. Harrington, Editors), pp. 1–27, Springer, Dordrecht.

Strauch, O. and Tóth, J.T. (1998). Asymptotic density of $A \subset \mathbf{N}$ and density of the ratio set $R(A)$, Acta Arith., Vol. 87, pp. 67–78.

Appendix: Proof of Theorem 5.1

It is obvious that Assumption A is satisfied. By Cramér's theorem, $M_n^{(1)}, M_n^{(2)}$, defined in (7), satisfy LDPs with speed $\nu_1(n)$ (respectively $\nu_2(n)$) and good rate functions

$$\begin{aligned} I^{(1)}(x) &= \begin{cases} \log 2 + \frac{x+1}{2}\log\frac{x+1}{2} + \frac{1-x}{2}\log\frac{1-x}{2}, & x\in[-1,1],\\ \infty, & \text{otherwise,}\end{cases}\\ I^{(2)}(x) &= \begin{cases} \log 2 + \frac{x+2}{4}\log\frac{x+2}{4} + \frac{2-x}{4}\log\frac{2-x}{4}, & x\in[-2,2],\\ \infty, & \text{otherwise,}\end{cases}\end{aligned} \tag{14}$$

where $0\log 0 := 0$. See Theorem I.3 and Exercise I.12 in den Hollander (2000). These functions are strictly convex on $[-1,1]$ and $[-2,2]$, respectively. Let $n_k \to \infty$ be a sequence such that $\nu_2(n_k)/n_k \to 0$. Since

$$M_n = \frac{\nu_1(n)}{n}M_n^{(1)} + \frac{\nu_2(n)}{n}M_n^{(2)},$$

and

$$\mathbb{P}\big[|M_{n_k}^{(2)}| \geq 3\big] = 0,$$

we have, for $x>0$,

$$\begin{aligned}\mathbb{P}\big[M_{n_k}\geq x\big] &= \mathbb{P}[M_{n_k}\geq x,\ |M_{n_k}^{(2)}|<3\big]\\ &= \mathbb{P}\Big[\frac{\nu_1(n_k)}{n_k}M_{n_k}^{(1)} \geq x - \frac{\nu_2(n_k)}{n_k}M_{n_k}^{(2)},\ |M_{n_k}^{(2)}|<3\Big]\\ &\leq \mathbb{P}\Big[\frac{\nu_1(n_k)}{n_k}M_{n_k}^{(1)} \geq x - \frac{3\nu_2(n_k)}{n_k}\Big].\end{aligned}$$

Similarly, we deduce the lower bound

$$\mathbb{P}[M_{n_k}\geq x] \geq \mathbb{P}\Big[\frac{\nu_1(n_k)}{n_k}M_{n_k}^{(1)} \geq x + \frac{3\nu_2(n_k)}{n_k}\Big].$$

Using the LDP for $M_n^{(1)}$ and $\nu_1(n_k)/n_k \to 1$, we obtain

$$\begin{aligned}-I^{(1)}(x+\delta) &\leq \liminf_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}[M_{n_k}\geq x]\\ &\leq \limsup_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}[M_{n_k}\geq x] \leq -I^{(1)}(x-\delta),\end{aligned}$$

for any $\delta>0$, and by taking $\delta\downarrow 0$, we conclude

$$\lim_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}[M_{n_k}\geq x] = -I^{(1)}(x), \quad x>0. \tag{15}$$

Analogously, by choosing a sequence $m_k\to\infty$ satisfying $\nu_1(m_k)/m_k\to 0$, we establish

$$\lim_{k\to\infty}\frac{1}{m_k}\log\mathbb{P}[M_{m_k}\geq x] = -I^{(2)}(x), \quad x>0.$$

Suppose now that M_n satisfies an LDP with good rate function I and speed $s(n)$. For $x>0$ and $N\in\mathbb{N}$, define

$$B_N := (x-1/N, x+1/N).$$

Then, the assumed LDP implies

$$\liminf_{k\to\infty} \frac{1}{s(n_k)} \log \mathbb{P}[M_{n_k} \in B_{N+1}] \geq -I(B_{N+1}), \quad N \in \mathbb{N}. \tag{16}$$

By (15) and the strict convexity of $I^{(1)}$, we have

$$\log \mathbb{P}[M_{n_k} \in B_{N+1}] = -I^{(1)}(B_{N+1}) n_k \big(1 + \mathrm{o}(1)\big), \quad k \to \infty. \tag{17}$$

If $x > 1$, then $I^{(1)}(B_{N+1}) = \infty$ for large N, and (16) and (17) imply $I(B_{N+1}) = \infty$ for large N. By lower semi-continuity, for $N \to \infty$ we obtain

$$I(x) = \infty, \quad x > 1. \tag{18}$$

For $0 < x \leq 1$, $I^{(1)}(B_{N+1})$ is finite, and (16) and (17) imply

$$I^{(1)}(B_{N+1}) \limsup_{k\to\infty} \frac{n_k}{s(n_k)} \leq I(B_{N+1}), \quad N \in \mathbb{N}.$$

Again, by lower semi-continuity, taking $N \to \infty$ yields

$$I^{(1)}(x) \limsup_{k\to\infty} \frac{n_k}{s(n_k)} \leq I(x), \quad 0 < x \leq 1. \tag{19}$$

Analogously, we can use the upper LDP bound

$$\limsup_{k\to\infty} \frac{1}{s(n_k)} \log \mathbb{P}[M_{n_k} \in \mathrm{cl}(B_{N+1})] \leq -I(\mathrm{cl}(B_{N+1})) \leq -I(B_N), \quad N \in \mathbb{N},$$

to prove

$$I^{(1)}(x) \liminf_{k\to\infty} \frac{n_k}{s(n_k)} \geq I(x), \quad 0 < x \leq 1. \tag{20}$$

Putting (19) and (20) together yields

$$I(x) = I^{(1)}(x)\ell_1, \quad 0 < x \leq 1, \tag{21}$$

where

$$\ell_1 := \lim_{k\to\infty} \frac{n_k}{s(n_k)},$$

exists in $[0, \infty]$ and is independent of x. Repeating the same steps with m_k instead of n_k shows

$$I(x) = I^{(2)}(x)\ell_2 := I^{(2)}(x) \lim_{k\to\infty} \frac{m_k}{s(m_k)}, \quad 0 < x \leq 2, \tag{22}$$

and so $I^{(1)}(x)\ell_1 = I^{(2)}(x)\ell_2$ for $0 < x \leq 1$. By coefficient comparison, using the expansions

$$I^{(1)}(x) = \tfrac{1}{2}x^2 + \tfrac{1}{12}x^4 + \tfrac{1}{30}x^6 + \mathrm{O}(x^8),$$
$$I^{(2)}(x) = \tfrac{1}{8}x^2 + \tfrac{1}{192}x^4 + \tfrac{1}{1920}x^6 + \mathrm{O}(x^8), \quad x \downarrow 0,$$

we see that this implies $(\ell_1, \ell_2) = (\infty, \infty)$ or $(\ell_1, \ell_2) = (0, 0)$. The latter is impossible, since (18) and (22) yield

$$\infty = I(\tfrac{3}{2}) = I^{(2)}(\tfrac{3}{2})\ell_2,$$

which requires $\ell_2 = \infty$, as $I^{(2)}(\frac{3}{2})$ is finite. To finish the proof, we must infer a contradiction from $(\ell_1, \ell_2) = (\infty, \infty)$. Indeed, (18) and (21) would then imply $I(x) = \infty$ for all $x > 0$, and so

$$\mathbb{P}[M_n \geq 1] = 0, \quad n \in \mathbb{N}.$$

This is wrong, because $\{M_n \geq 1\}$ contains the event

$$\{X_k = 1 \text{ for } k \leq n, k \in K_1\} \cap \{X_k = 2 \text{ for } k \leq n, k \in K_2\},$$

which has positive probability.