

# Convergence and asymptotic variance of bootstrapped finite-time ruin probabilities with partly shifted risk processes<sup>☆</sup>

Stéphane Loisel<sup>a</sup>, Christian Mazza<sup>b</sup>, Didier Rullière<sup>a,\*</sup>

<sup>a</sup> Université de Lyon, Université Lyon 1, Laboratoire SAF, EA 2429, Institut de Science Financière et d'Assurances, 50 Avenue Tony Garnier, F-69007 Lyon, France

<sup>b</sup> Department of Mathematics, University of Fribourg, Pérolles, Chemin du Musée 23, CH-1700 Fribourg, Switzerland

In the classical risk model, we prove the weak convergence of a sequence of empirical finite-time ruin probabilities. In an earlier paper (see Loisel et al., (2008)), we proved an equivalent result in the special case where the initial reserve is zero, and checked that numerically the general case seems to be true. In this paper, we prove the general case (with a nonnegative initial reserve), which is important for applications to estimation risk. So-called partly shifted risk processes are introduced, and used to derive an explicit expression of the asymptotic variance of the considered estimator. This provides a clear representation of the influence function associated with finite time ruin probabilities and gives a useful tool to quantify estimation risk according to new regulations.

## 1. Introduction

In ruin theory, the surplus of an insurance company is classically represented by the risk process  $(R_t)_{t \geq 0}$  defined as follows: for  $t \geq 0$ ,

$$R_t = u + ct - S_t,$$

where  $u$  is the nonnegative amount of initial reserves and  $c > 0$  is the premium income rate. The cumulated claim amount up to time  $t$  is described by the compound Poisson process

$$S_t = \sum_{i=1}^{N_t} W_i,$$

where the amounts of claims  $W_i$ ,  $i \geq 1$ , are nonnegative independent, identically-distributed random variables, distributed as  $W$ , with the convention that  $S_t = 0$  if  $N_t = 0$ . The number of

claims  $N_t$  until  $t \geq 0$  is modeled by a homogeneous Poisson process  $(N_t)_{t \geq 0}$  of intensity  $\lambda$ . Claim amounts and arrival times are assumed to be independent.

We are interested in the robust estimation of finite-time ruin probabilities. Solvency regulations for insurance companies, called *Solvency II*, impose the control of a certain number of insolvency probabilities. The chosen risk measure to determine the Solvency Capital Requirements (SCR) is more likely to be a 99.5%, one-year Value at Risk than a continuous-time ruin probability. Nevertheless, reserving is expected to be quantified by a best estimate of liabilities, plus a so-called Market-Value Margin (MVM), which is determined by a cost-of-capital approach: this margin corresponds to the cost of maintaining the surplus above the SCR level during the whole period  $[0, t]$ , where  $t$  is typically 10 years. This corresponds to a continuous-time ruin problem in a finite horizon. Let us denote by  $\psi(u, t)$  the probability of ruin before time  $t$  with initial reserve  $u$ :

$$\psi(u, t) = P[\exists s \in [0, t], R_s < 0 \mid R_0 = u], \quad u \geq 0, t > 0,$$

and let

$$\varphi(u, t) = 1 - \psi(u, t)$$

be the probability of non-ruin within time  $t$  with initial reserve  $u$ .

<sup>☆</sup> This work has been partly funded by ANR Research Project ANR-08-BLAN-0314-01.

\* Corresponding author at: Université de Lyon, F-69622, Lyon, France. Tel.: +33 4 37 28 74 38; fax: +33 4 37 28 76 32.

E-mail addresses: [Stephane.Loisel@univ-lyon1.fr](mailto:Stephane.Loisel@univ-lyon1.fr) (S. Loisel), [christian.mazza@unifr.ch](mailto:christian.mazza@unifr.ch) (C. Mazza), [Didier.Rulliere@univ-lyon1.fr](mailto:Didier.Rulliere@univ-lyon1.fr) (D. Rullière).

Algorithms to compute or approximate  $\psi(u, t)$  have been proposed, among others, by Asmussen et al. (2002) by the means of Erlangization, by Picard and Lefèvre (1997) by the means of Appell polynomials, and by Rullière and Loisel (2004) and by the means of a Seal-type argument based on the ballot Lemma (see also Lefèvre and Loisel (2008)). This model has been investigated by Lefèvre and Loisel (2009) and Biard et al. (2008) among others in the case of dependent claim amounts. Sensitivity analysis of finite-time ruin probabilities has been studied in Loisel et al. (2008) and Loisel and Privault (2009). Some estimation problems have been studied in Frees (1986), Hipp (1989), Croux and Veraverbeke (1990), Bening and Korolev (2000). An important feature of Solvency II is that estimation risk should be controlled, particularly if internal models are used. However, most models use a calibrated model and an *a posteriori* proportional loading factor to take into account estimation risk. It would be of course much better to carry out a robustness analysis at the same time. This led us to define in an earlier paper (see Loisel et al. (2008)) reliable ruin probabilities as quantiles of empirical finite-time ruin probabilities, and the Estimation Risk Solvency Margin (ERSM) as the additional capital required to cover estimation risk.

Let  $\psi^N(u, t)$  and  $\varphi^N(u, t)$  respectively be the finite-time ruin and non-ruin probability with claim amounts drawn from the empirical distribution  $F_N$  associated with an i.i.d. sample of distribution  $F$ , where  $F$  is the c.d.f. of  $W$  and  $N \geq 1$  is the size of the sample. Define the *reliable finite-time ruin probability*  $\psi_{1-\varepsilon}^{N, \text{reliable}}(u, t)$  as the  $(1 - \varepsilon)$ -quantile of the (random) bootstrapped finite-time ruin probability  $\psi^N(u, t)$ :

$$\psi_{1-\varepsilon}^{N, \text{reliable}}(u, t) = \inf_{s \geq 0} \{P[\psi^N(u, t) \geq s] \leq \varepsilon\}.$$

If  $u_\eta$  and  $u_{\eta, \varepsilon}$  are respectively defined as the initial capital required to ensure that

$$\psi(u_\eta, t) \leq \eta$$

and

$$\psi_{1-\varepsilon}^{N, \text{reliable}}(u_{\eta, \varepsilon}, t) \leq \eta,$$

the Estimation Risk Solvency Capital  $ERSM_{\eta, 1-\varepsilon}$  can be defined as the additional capital needed to take estimation risk into account:

$$ERSM_{\eta, 1-\varepsilon} = u_{\eta, 1-\varepsilon} - u_\eta.$$

In Loisel et al. (2008), we have shown the convergence of

$$\sqrt{N}(\varphi(u, t) - \varphi^N(u, t))$$

as  $N$  tends to  $+\infty$  in distribution to a centered, Gaussian random variable only for  $u = 0$ . The proof was based on the symmetrical in  $(W_1, \dots, W_n)$  expression

$$\varphi(0, t) = \sum_{n \geq 1} P(N_t = n) E[(u + ct - (W_1 + \dots + W_n))_+],$$

where  $x_+$  denotes the positive part of a real number  $x$ . We also computed the asymptotic variance  $V_u$  of the limit of

$$\sqrt{N}(\varphi(u, t) - \varphi^N(u, t)),$$

and expressed this variance in terms of the variance of a random variable defined from the influence function of finite-time non-ruin probability. Influence functions were introduced in the field of robust statistics to study the impact of data contamination on the estimated quantity (see Huber (1981) and Hampel (1974)). For a functional  $T$  of a distribution  $F$ , the influence function at point  $x \in \mathbf{R}$  is defined as the limit (when it exists)

$$IF_x[T] = \lim_{s \downarrow 0} \frac{T(F^{(s, x)}) - T(F)}{s},$$

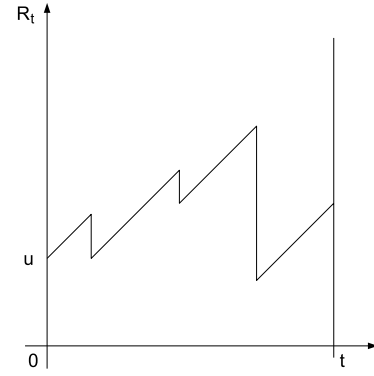


Fig. 1. A sample path of the classical risk process  $R_t$ .

where  $F^{(s, x)}$  is defined for  $x \in \mathbf{R}$  and  $s > 0$  by

$$F^{(s, x)}(u) = s1_{\{x \leq u\}} + (1 - s)F(u).$$

In the sequel, for each quantity related to the contaminated distribution  $F^{(s, x)}$ , we use the exponent  $^{(s, x)}$ . In a recent paper, Marceau and Rioux (2001) provided an algorithm to compute the influence function of the eventual probability of ruin. We obtained in Loisel et al. (2008) that for all  $u \geq 0$ ,

$$V_u = V_Y [IF_Y[\varphi(u, t)]] \quad (1)$$

Nevertheless, it remains to be shown that the limit distribution of

$$\sqrt{N}(\varphi(u, t) - \varphi^N(u, t))$$

is Gaussian, and besides formula (1) involves computation of influence functions described in Loisel et al. (2008), which corresponds to heavy computation times and new convergence issues. In this paper, we prove the convergence in the distribution of

$$\sqrt{N}(\varphi(u, t) - \varphi^N(u, t))$$

toward a Gaussian random variable for all  $u \geq 0$  by the means of  $U$ -statistics and so-called partly shifted risk processes. These processes are defined in Section 2, in which finite-time ruin probabilities for partly shifted risk processes are computed as well. The expression of  $V_u$  in terms of ruin probabilities for modified risk processes derived in Section 3.5 is of fundamental importance from a theoretical and operational point of view: it gives a probabilistic representation of  $V_u$ , which is used to prove the convergence of the empirical ruin probability for arbitrary  $u \geq 0$ . We also give elegant mathematical expressions for influence functions associated to finite time ruin probabilities. Finally, we provide efficient numerical methods to obtain numerical values of  $V_u$ .

## 2. Finite-time ruin probabilities for partly shifted risk processes

### 2.1. Partly shifted risk processes

Given  $x \geq 0$ , we define the  $x$ -partly shifted risk process as the stochastic process given by

$$R_t^x = u + ct - S_t^x,$$

where

$$S_t^x = S_t + x1_{\{U \leq t\}},$$

and  $U$  is a certain positive random variable. After this random delay  $U$ , the sample path of  $R_t^x$  is shifted  $x$  units downwards. It corresponds to add a jump of size  $x$  at a random instant  $U$  (see Figs. 1 and 2).

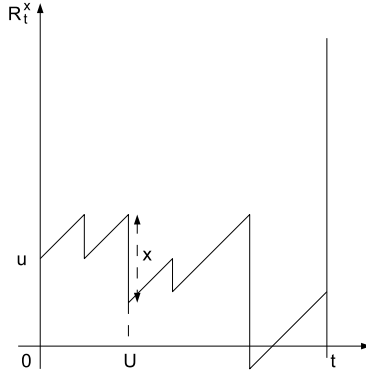


Fig. 2. A corresponding sample path of the  $x$ -partly shifted risk process.

The process  $(R_t^x)_{t \geq 0}$  has no longer stationary and independent increments. Nevertheless, we will show how to adapt results of risk theory to these partly shifted risk processes. This is important since, as we will see, finite-time ruin probabilities for partly shifted processes are directly involved in computations of sensitivities, influence functions and asymptotic variance of finite-time ruin probabilities for classical risk processes.

## 2.2. Finite-time ruin probability for partly shifted risk processes starting from zero

Let us consider the case where claim amounts are integer-valued.

**Proposition 1.** For  $c > 0$ ,  $x > 0$  and  $n \in \mathbf{N} \setminus \{0\}$ , we have

$$P[S_t^x < ct, \forall t < \frac{n}{c}] = P[S_{i/c}^x < i, i = 1..n].$$

This is a direct adaptation of results of [Rullière and Loisel \(2004\)](#). It thus suffices to study the cumulated claim amount process at a finite set of inventory dates.

To this end, consider for  $1 \leq i \leq n$  the random variables defined by  $Y_i = S_{i/c}$ ,  $Y_i^x = S_{i/c}^x$ , and

$$Y_i^x = Y_i + x1_{\{\bar{U} \leq i\}},$$

where  $\bar{U}$  is uniformly distributed on the finite set  $\{1, 2, \dots, n\}$ . The finite-time ruin probability for partly shifted risk processes starting from 0 is obtained by an adaptation of the well known ballot theorem.

**Lemma 1** (Ballot Theorem (See [Takács \(1962a,b\)](#))). For  $n \geq 1$ , let  $Z_i, i = 1..n$  be a process with exchangeable increments. Then, we have

$$P[Z_i < i, i = 1..n, Z_n = j] = \frac{n-j}{n} P[Z_n = j] \quad 0 \leq j \leq n.$$

The ballot theorem applies to the family of random variables  $Y_i, i = 1, \dots, n$ , which has i.i.d. increments, but also to the  $Y_i^x, i = 1, \dots, n$ , which has exchangeable (but not independent) increments.

**Proposition 2** (Ballot Theorem for Partly Shifted Risk Processes). Let  $W_1, \dots, W_n$  be i.i.d., integer-valued random variables. Consider the partial sum process

$$Y_i = \sum_{j=1}^i W_j,$$

and define

$$Y_i^x = Y_i + x1_{\{\bar{U} \leq i\}},$$

where  $\bar{U}$  is uniformly distributed on the finite set  $\{1, 2, \dots, n\}$ . We have

$$P[Y_i^x < i, i = 1..n, Y_n^x = j] = \frac{n-j}{n} P[Y_n^x = j] \quad 0 \leq j \leq n.$$

**Proof.** Takács's result is true for exchangeable random variables (and even for cyclically exchangeable random variables). It can easily be shown that the process  $Y_i^x, 1 \leq i \leq n$  has exchangeable increments  $W_i^x, 1 \leq i \leq n$ , defined by

$$W_i^x = W_i + x1_{\{\bar{U} = i\}}.$$

To this end, consider the distribution function

$$F_{W_{\sigma(1)}^x, \dots, W_{\sigma(n)}^x}$$

associated with the random vector

$$(W_{\sigma(1)}^x, \dots, W_{\sigma(n)}^x),$$

which is also the distribution function of the random vector

$$(W_1^x, \dots, W_n^x),$$

which is denoted by

$$F_{W_1^x, \dots, W_n^x},$$

because for any  $w_1, \dots, w_n \in \mathbf{R}$ , we have

$$F_{W_{\sigma(1)}^x, \dots, W_{\sigma(n)}^x}(w_1, \dots, w_n) = \frac{1}{n} \sum_{i=1}^n F_{W_{\sigma(1)}, \dots, W_{\sigma(n)}}(w_1, \dots, w_{i-1}, w_i - x, w_{i+1}, \dots, w_n) \quad (2)$$

$$= \frac{1}{n} \sum_{i=1}^n F_{W_1, \dots, W_n}(w_1, \dots, w_{i-1}, w_i - x, w_{i+1}, \dots, w_n) \quad (3)$$

$$= F_{W_1^x, \dots, W_n^x}(w_1, \dots, w_n). \quad (4)$$

We used the fact that the random variables  $W_i, 1 \leq i \leq n$ , are exchangeable to get (3) from (2), and the fact that  $\bar{U}$  is uniformly distributed on  $\{1, \dots, n\}$ , and independent of the  $W_i$ 's to write (2), and to get (4) from (3).  $\square$

**Remark 1.** The previous result could also be proved by induction on  $n \geq 1$ .

It remains valid if  $\bar{U}$  is uniformly distributed on  $\{1, \dots, n_{\max}\}$ , with  $n_{\max} \geq n$ . To prove this, distinguish two cases: given that  $\bar{U} \leq n$ , [Proposition 2](#) applies, and given that  $\bar{U} > n$ , the classical ballot lemma applies.

[Propositions 1](#) and [2](#) directly enable us to obtain the finite-time ruin probability for partly shifted risk processes starting from 0:

**Theorem 1.** The finite-time ruin probability for partly shifted risk processes starting from 0 is given by

$$P[R_s^x \geq 0 \forall s < t \mid R_0^x = 0] = E \left[ \frac{(ct - x - S_t)_+}{ct} \right],$$

where  $R_s^x = R_s - x1_{\{\bar{U} \leq s\}}$ , and  $U$  is uniformly distributed on  $[0, t]$ .

## 2.3. Finite-time ruin probability for partly shifted risk processes starting from $u \geq 0$

Conditionally on the last continuous passage of  $R_t$  at 0, the process is located under the barrier  $u + ct$  at time  $t = n/c$  if there is no ruin, or if the last visit of the process at 0 occurred at time  $i/c$ . Let

$T_u^x$  be the first instant of ruin associated with the modified process, when  $R_0 = u$ . Then we have (see [Rulli re and Loisel \(2004\)](#))

$$P[S_{n/c}^x < u + n] = P\left[T_u^x > \frac{n}{c}\right] + \Delta,$$

where

$$\Delta = \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} P[S_{i/c}^x = u + i \cap S_{i+k/c}^x - S_{i/c}^x < i, \\ k = 1, \dots, n - i \cap S_{n/c}^x - S_{i/c}^x = j].$$

If  $\bar{U}$  is uniformly distributed on  $1, \dots, n$ , conditioning on  $\bar{U}$ , we can consider every element of the second sum and get the following result.

**Theorem 2** (Modified Ruin Probability when  $R_0 = u$ ). The ruin probability associated to the modified process is given by

$$P\left[T_u^x > \frac{n}{c}\right] = P[x + S_{n/c} < u + n] - \Delta_1 - \Delta_2,$$

where

$$\Delta_1 = \sum_{i=1}^{n-1} \frac{i}{n} P[x + S_{i/c} = u + i] \sum_{j=1}^{n-i} P[S_{(n-i)/c} = j] \frac{n-i-j}{n-i},$$

and

$$\Delta_2 = \sum_{i=1}^{n-1} \frac{n-i}{n} P[S_{i/c} = u + i] \sum_{j=1}^{n-i} P[x + S_{(n-i)/c} = j] \frac{n-i-j}{n-i}.$$

### 3. Sensitivity analysis

#### 3.1. The ruin probability as a function of the additional claim

In the sequel, we consider the probability of non-ruin before  $t = n/c$ , when starting with an integer valued initial reserve. Conditioning on the last passage time  $i/c$  of  $R_t$  at 0, the probability that the reserve is some integer  $j$  at time  $n/c$ ,  $1 \leq j \leq u + n$ , is given by (see [Loisel et al. \(2008\)](#))

$$P[R_{n/c} = j] = P[T_u > n/c \text{ and } R_{n/c} = j] + \sum_{i=1}^{n-j} P[S_{i/c} = u + i] \\ \times P[S_{(n-i)/c} = j \text{ and } S_{v/c} < v, \forall 1 \leq v \leq n - i],$$

where we assume that the sum vanishes when  $j \geq n$ . Using the classical results of [Tak cs \(1962a\)](#), one obtains that the first term is given by

$$P[S_{n/c} = u + n - j] - 1_{[j \leq n]} \sum_{i=1}^{n-j} P[S_{i/c} = u + i] \\ \times P[S_{(n-i)/c} = j] \frac{n-i-j}{n-i}.$$

For  $j = 0$ , we have  $P[T_u > n/c \text{ and } R_{n/c} = 0] = P[S_{n/c} = u + n] - P[S_{n/c} = u + n] = 0$ .

Conditionally on  $\{N_{n/c} = k\}$ , defining

$$\varphi_{k,j}(u, n) = P[T_u > n/c \text{ and } R_{n/c} = j \mid N_{n/c} = k],$$

and  $\varphi_k(u, n) = P[T_u > n/c \mid N_{n/c} = k]$ , one obtains that

$$\varphi_{k,j}(u, n) = P[W^{*k} = u + n - j] \\ - \sum_{i=1}^{n-j} \sum_{n_0=0}^k \alpha_{i,k}(n_0) P[W^{*n_0} = u + i] P[W^{*(k-n_0)} = j] \frac{n-i-j}{n-i},$$

where  $\alpha_{i,k}(n_0) = P[N_{i/c} = n_0] P[N_{(n-i)/c} = k - n_0] / P[N_{n/c} = k]$ , id est

$$\alpha_{i,k}(n_0) = \binom{n_0}{k} i^{n_0} (n-i)^{k-n_0} / n^k.$$

So we have

$$\varphi_k(u, n) = \sum_{j=0}^{u+n} \varphi_{k,j}(u, n)$$

and

$$\varphi(u, n) = \sum_{k=0}^{u+n} P[N_{n/c} = k] \varphi_k(u, n).$$

In fact, every claim amount  $W$  takes here a positive integer value (one can assume that  $W \neq 0$  by modifying  $\lambda$ ).

#### 3.2. Inference of ruin probabilities

We will try here to express the plugin estimator of the probability of ruin as a  $U$ -statistic. The basic information we use is exposed in [Hoeffding \(1948\)](#) and [Von Mises \(1947\)](#).

Assuming that the observed claims belong to the finite set  $\{w_1, \dots, w_N\}$ ,  $\varphi_{k,j}(u, n)$  can be estimated by using a statistic  $\hat{\varphi}_{k,j}(u, n)$ . We will write  $[N] = \{1, \dots, N\}$ . We are looking for a plugin estimator which corresponds to the probability one gets if the empirical distribution (associated to the observed claim amounts) is used as the claim amount distribution. We thus work with following estimator, for which this link can be obtained:

$$\hat{\varphi}_{k,j}(u, n) = \frac{1}{N^k} \sum_{i_1, \dots, i_k \in \{1, \dots, N\}} \left[ I_1^k(u + n - j) \right. \\ \left. - \sum_{i=1}^{n-j} \sum_{n_0=0}^k \alpha_{i,k}(n_0) I_1^{n_0}(u + i) I_{n_0+1}^k(j) \frac{n-i-j}{n-i} \right],$$

where  $I_x^y(j) = 1_{\{w_{i_1} + \dots + w_{i_y} = j\}}$ ,  $x, y \in \mathbf{N}, x \leq y$ . The indicator functions  $I_x^y(j)$  are defined for each multi-index  $\vec{i} = (i_1, \dots, i_k)$  (omitted here). We recall that  $\alpha_{i,k}(n_0) = P[B_{k,i/n} = n_0]$ , where  $B_{k,i/n}$  is a binomial random variable of parameters  $k$  and  $i/n$ . Given  $u$  and  $n$ , set

$$\hat{\varphi}_{k,j}(u, n) = \frac{1}{N^k} \sum_{i_1, \dots, i_k \in \{1, \dots, N\}} \Phi_{k,j}(i_1, \dots, i_k),$$

where

$$\Phi_{k,j}(i_1, \dots, i_k) = \Phi_{k,j}^{(1)}(i_1, \dots, i_k) - \Phi_{k,j}^{(2)}(i_1, \dots, i_k),$$

and

$$\Phi_{k,j}^{(1)}(i_1, \dots, i_k) = I_1^k(u + n - j), \\ \Phi_{k,j}^{(2)}(i_1, \dots, i_k) = \sum_{i=1}^{n-j} \sum_{n_0=0}^k \alpha_{i,k}(n_0) I_1^{n_0}(u + i) I_{n_0+1}^k(j) \frac{n-i-j}{n-i}.$$

A basic result of [Hoeffding \(1948\)](#) holds when  $\Phi_{k,j}(i_1, \dots, i_k)$  is symmetric as a function of  $i_1, \dots, i_k$ . We shall study questions of symmetry in the next section.

### 3.3. On the symmetry of $\widehat{\varphi}_{k,j}(u, n)$

First,  $\Phi_{k,j}^{(1)}(i_1, \dots, i_k) = I_1^k(u + n - j)$  is clearly symmetric as a function of  $i_1, \dots, i_k$ . Next, one studies the symmetry of  $\Phi_{k,j}^{(2)}(i_1, \dots, i_k)$ . Let  $\sigma_k$  denote all the possible permutations of the set  $\{i_1, \dots, i_k\}$ . Then one obtains that

$$\frac{1}{k!} \sum_{\sigma_k} \sum_{n_0=0}^k \alpha_{i,k}(n_0) I_1^{n_0}(u + i) I_{n_0+1}^k(j) = E_{\delta_1, \dots, \delta_k} [J_k(u + i) \bar{J}(j)],$$

with

$$J_k(u + i) = 1_{\{\delta_1 w_{i_1} + \dots + \delta_k w_{i_k} = u + i\}},$$

and

$$\bar{J}_k(j) = 1_{\{\bar{\delta}_1 w_{i_1} + \dots + \bar{\delta}_k w_{i_k} = j\}},$$

where  $\bar{\delta}_v = 1 - \delta_v$ , and where the random sequence  $\delta_v$  is i.i.d., with Bernoulli distribution with parameter  $i/n$ ,  $v \leq k$ .

We can show that a transposition of two elements  $i_1$  and  $i_v$  has the same effect than a transposition of  $\delta_1$  and  $\delta_v$ . The random variables  $\delta_v$  are i.i.d., so that the mean is invariant under any permutation of the variables  $\delta_v$ , that is

$$E_{\delta_1, \dots, \delta_k} [J_k(u + i) \bar{J}(j)]$$

is symmetric as a function of  $i_1, \dots, i_k$ .

The point here is that  $\sum_{n_0=0}^k \alpha_{i,k}(n_0) I_1^{n_0}(u + i) I_{n_0+1}^k(j)$  is not symmetric. We will use the following lemma:

**Lemma 2.**

$$N^k \widehat{\varphi}_{k,j}(u, n) = \binom{N}{k} U + \sum^* \Phi_{k,j}(i_1, \dots, i_k),$$

where the sum  $\sum^*$  is taken over all  $k$ -uplets for which at least one pair of indices is such that  $i_{v_0} = i_{v_1}$  ( $v_0 \neq v_1$ ), and where  $U$  is a  $U$ -statistic.

**Proof.** Let  $\Theta_k = \{i_1, \dots, i_k\}$  be a subset of  $[N]$  of  $k$  distinct elements  $i_1 < \dots < i_k$ . Then one can write

$$\sum_{i_1, \dots, i_k \in \{1, \dots, N\}} \Phi_{k,j}^{(2)}(i_1, \dots, i_k) = \sum_{\Theta_k \subset [N]} \sum_{\sigma_k} \sum_{i=1}^{n-j} \sum_{n_0=0}^k \alpha_{i,k}(n_0) I_1^{n_0}(u + i) \times I_{n_0+1}^k(j) \frac{n-i-j}{n-i} + \sum^* \dots$$

As stated previously, the expectation  $E_{\delta_1, \dots, \delta_k} [J_k(u + i) \bar{J}(j)]$  is independent of the choice of the subset  $\Theta_k \subset [N]$ . One obtains that

$$\sum_{i_1, \dots, i_k \in \{1, \dots, N\}} \Phi_{k,j}^{(2)}(i_1, \dots, i_k) = \sum_{\Theta_k \subset [N]} \sum_{i=1}^{n-j} k! E_{\delta_1, \dots, \delta_k} [J_k(u + i) \bar{J}(j)] \frac{n-i-j}{n-i} + \sum^* \dots$$

where

$$U = U^{(1)} - U^{(2)}$$

$$U^{(1)} = \frac{1}{\binom{N}{k}} \sum_{\Theta_k \subset [N]} I_1^k(u + n - j),$$

and

$$U^{(2)} = \frac{1}{\binom{N}{k}} \sum_{\Theta_k \subset [N]} \sum_{i=1}^{n-j} k! E_{\delta_1, \dots, \delta_k} [J_k(u + i) \bar{J}(j)] \frac{n-i-j}{n-i},$$

proving the result.  $\square$

### 3.4. Hoeffding's results

The variance of  $\widehat{\varphi}_{k,j}(u, n)$  is bounded and may be expressed as a function of a  $U$ -statistic. We can thus apply a powerful theorem from Hoeffding (Theorem 7.3 on page 308 of [Hoeffding \(1948\)](#)) to get the following theorem.

**Theorem 3** (Asymptotic Normality of the Empirical Ruin Probability). *Let*

$$\widehat{\varphi}_k(u, n) = \sum_{j=0}^{u+n} \widehat{\varphi}_{k,j}(u, n),$$

and set

$$\Delta_k = \sqrt{N} (\widehat{\varphi}_k(u, n) - \varphi_k(u, n)).$$

The random vector  $(\Delta_1, \dots, \Delta_{u+n})$  is an asymptotically centered normal of variance

$$\Gamma = \{\gamma \delta \zeta_1^{(\gamma, \delta)}\}_{\gamma, \delta \in \{1, \dots, u+n\}}.$$

The variance  $\zeta_1^{(\gamma, \delta)}$  is not very explicit. We will discuss its various properties in the next section.

**Remark 2.** We can assume that the number of claims is smaller than  $u + n$ , since, if not, ruin occurs with probability one:  $W \geq 1$  and therefore  $\Delta_k = 0$ .

The variance of  $\sqrt{(N)} (\widehat{\varphi}(u, n) - \varphi(u, n))$  is given by

$$V_u = \sum_{\gamma=0}^{u+n} \sum_{\delta=0}^{u+n} \gamma \delta P[N_{n/c} = \gamma] P[N_{n/c} = \delta] \zeta_1^{(\gamma, \delta)}.$$

If asymptotic normality holds, the limiting variance is given by formula (7) of [Loisel et al. \(2008\)](#), and we have  $V_u = V_Y [IF_Y \varphi(u, t)]$ , where  $IF_Y$  is the related influence function (for more details, see [Hampel \(1974\)](#) and [Hampel et al. \(1986\)](#), or [Huber \(1981\)](#)). Let  $V_Y$  denote the variance of the random variable  $Y$ . We will give equivalent expressions for these variances using the shifted ruin process of Section 2.

### 3.5. Alternative formulas for $V_u$

Given that among  $k$  claims, one of them is given by  $Y$ , one has  $\varphi_{k,j}^Y(u, n) = L_{k,j}(Y) - R_{k,j}(Y)$ . We omit  $Y$  when there is no ambiguity; furthermore, the letters  $L$  et  $R$  are used for *Left* and *Right*. Note that

$$L_{k,j} = P[Y + W^{*k} = u + n - j].$$

Since we do not know a priori whether the claim  $Y$  occurred before the  $n_0$ th claim,  $R_{k,j}$  is obtained by conditioning on  $\delta$ , the Bernoulli random variable of parameter  $i/n$ , which indicates if the claim occurred. When  $j = 0$ , one has  $L_{k,0} = R_{k,0} = P[Y + W^{*k} = u + n]$ . When  $j \geq 1$ , conditioning on  $\delta$ , one obtains that  $R_{k,j}$  is given by

$$\begin{aligned} & \sum_{i=1}^{n-j} \sum_{n_0=0}^k \frac{i}{n} P[B_{k-1, i/n} = n_0 - 1] \\ & \times P[Y + W^{*n_0} = u + i] P[W^{*k-n_0} = j] \frac{n-i-j}{n-i} \\ & + \frac{n-i}{n} P[B_{k-1, i/n} = n_0] P[W^{*n_0} = u + i] \\ & \times P[Y + W^{*k-n_0-1} = j] \frac{n-i-j}{n-i}. \end{aligned}$$

Set

$$L_k(Y) = \sum_{j=0}^{u+n} L_{k,j}(Y),$$

$$R_k(Y) = \sum_{j=0}^{u+n} R_{k,j}(Y),$$

$\bar{L}_k = E_Y [L_k(Y)]$  and  $\bar{R}_k = E_Y [R_k(Y)]$ . Then, following [Hoeffding \(1948, Section 6\)](#), one gets that the covariance term  $\zeta_1^{(k,l)}$  can be written as

$$\zeta_1^{(k,l)} = E_Y [(L_k - R_k) - (\bar{L}_k - \bar{R}_k)] ((L_l - R_l) - (\bar{L}_l - \bar{R}_l)).$$

Hence, using the fact that both  $\bar{L}$  and  $\bar{R}$  are expectations over  $Y$ , we have

$$\zeta_1^{(k,l)} = E_Y [(L_k - R_k)(L_l - R_l)] - (\bar{L}_k - \bar{R}_k)(\bar{L}_l - \bar{R}_l).$$

In the sequel, if no specific indication is given, we consider that all probabilities, expectations and variances are taken given  $Y$ .

The variance  $V_u$  is given by

$$V_u = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} k l P[N_{n/c} = k] P[N_{n/c} = l] \zeta_1^{(k,l)},$$

and it is easy to show that in fact

$$V_u = V_Y \left[ \sum_{k=1}^{+\infty} k P[N_{n/c} = k] (L_k(Y) - R_k(Y)) \right].$$

Using the fact that

$$k P[N_{n/c} = k] = \frac{\lambda n}{c} P[N_{n/c} = k - 1], \quad k \geq 1,$$

one gets that  $V_u$  is the variance of the random variable

$$\frac{\lambda n}{c} \sum_{k=0}^{+\infty} P[N_{n/c} = k - 1] (L_k(Y) - R_k(Y)).$$

The term on the left is given by

$$\sum_{k=0}^{+\infty} P[N_{n/c} = k - 1] L_{k,j}(Y) = P[Y + S_{n/c} = u + n - j]$$

$$\sum_{k=0}^{+\infty} P[N_{n/c} = k - 1] R_{k,j}(Y) = P[Y + S_{n/c} \leq u + n].$$

On the other hand, the term on the right is obtained by using the identity

$$P[N_{n/c} = k] P[B_{k,i/n} = n_0] = P[N_{i/c} = n_0] P[N_{(n-i)/c} = k - n_0],$$

since we can write

$$\sum_{k=0}^{+\infty} P[N_{n/c} = k - 1] R_{k,j}(Y) = \sum_{i=1}^{n-j} \frac{n-i-j}{n-i} (S_{j,1} + S_{j,2}),$$

where  $S_{j,1}$  and  $S_{j,2}$  are given by

$$\sum_{k \in \mathbb{N}} \sum_{n_0 \leq k} \frac{i}{n} P[N_{i/c} = n_0 - 1]$$

$$\times P[N_{(n-i)/c} = k - n_0] P[W^{*n_0-1} + Y = u + i] P[W^{*k-n_0} = j],$$

and

$$\sum_{k \in \mathbb{N}} \sum_{n_0 \leq k} \frac{n-i}{n} P[N_{i/c} = n_0] P[N_{(n-i)/c} = k - 1 - n_0]$$

$$\times P[W^{*n_0} = u + i] P[W^{*k-n_0-1} + Y = j].$$

Hence, we have

$$S_{j,1} = \frac{i}{n} P[S_{i/c} + Y = u + i] P[S_{(n-i)/c} = j],$$

$$S_{j,2} = \frac{n-i}{n} P[S_{i/c} = u + i] P[S_{(n-i)/c} + Y = j].$$

Finally, one gets the following result.

**Theorem 4** (Asymptotic Variance  $V_u$ ). *Let  $Y$  be distributed according to the claim size distribution.  $V_u$  is the variance over  $Y$  of the following function of  $Y$*

$$\frac{\lambda n}{c} P[Y + S_{n/c} < u + n]$$

$$- \frac{\lambda n}{c} \sum_{i=1}^{n-1} \frac{i}{n} P[S_{i/c} + Y = u + i] \sum_{j=1}^{n-i} P[S_{(n-i)/c} = j] \frac{n-i-j}{n-i}$$

$$- \frac{\lambda n}{c} \sum_{i=1}^{n-1} \frac{n-i}{n} P[S_{i/c} = u + i] \sum_{j=1}^{n-i} P[S_{(n-i)/c} + Y = j] \frac{n-i-j}{n-i}.$$

We therefore find a mathematical expression which is similar to the ruin probability. This can be seen intuitively quite clearly since, apart from a factor  $\lambda n/c$ , the random variable of interest corresponds to the ruin probability associated to the process obtained by adding an additional claim  $Y$ . Looking at the various terms of the above expression,  $Y$  is added to  $S_{i/c}$  with probability  $i/n$ , or to  $S_{(n-i)/c}$  with probability  $n-i/n$ . In the special case where  $Y = 0$ , we recover the ruin probability obtained by summing over  $j$ :

$$P[T_u > n/c \text{ and } R_{n/c} = j] = P[S_{n/c} = u + n - j]$$

$$- 1_{\{j \leq n\}} \sum_{i=1}^{n-j} P[S_{i/c} = u + i] P[S_{(n-i)/c} = j] \frac{n-i-j}{n-i}.$$

Notice that for  $u = 0$ , this formula corresponds to the formula for  $V_0$  given in [Loisel et al. \(2008\)](#). We can finally give a more compact version of the formula:

**Theorem 5** (Link with the Partly Shifted Process). *Set*

$$\varphi^x(u, t) = P[R_s^x \geq 0 \forall s < t \mid R_0^x = u],$$

where  $R_s^x = R_s - 1_{\{U < s\}}x$ ,  $s \geq 0$ , and  $U$  is uniform on  $[0, t]$ . Then we have

$$V_u = V_Y [\lambda t \varphi^Y(u, t)].$$

**Remark 3.** Using the notation of [Loisel et al. \(2008\)](#), one can check that  $V_u$  is also the variance of

$$IF_Y[\varphi(u, n)].$$

This is the variance of

$$\lambda \frac{n}{c} \varphi^Y(u, n),$$

where

$$\varphi^x(u, n) = P[R_s^x \geq 0 \forall s < n/c \mid R_0^x = u],$$

and  $\varphi(u, n) = \varphi^0(u, n)$  are the probabilities of non-ruin before time  $n/c$  for the regular and the modified processes.



#### 4. Link with influence functions

The  $k$ th-order derivative of the ruin probability with respect to  $\lambda$  involves  $R_t^x$  since

$$\frac{\partial^k}{\partial \lambda^k} \varphi(0, t) = t^k \sum_{i=0}^k \binom{k}{i} (-1)^{k-i} E_{Y_1, \dots, Y_i} [\varphi^{Y_1 + \dots + Y_i}(0, t)],$$

as a consequence of Eq. (9) in Loisel et al. (2008). In particular,

$$\frac{\partial}{\partial \lambda} \varphi(0, t) = t E_Y [\varphi^Y(0, t) - \varphi(0, t)],$$

as the  $Y_i$  are i.i.d., and distributed as generic claim amount  $W$ .

##### 4.1. IF for the compound Poisson distribution

We use here some basic facts given in Loisel et al. (2008). Let  $\mathbf{T}$  be a functional of the distribution function  $F$ . By definition of the influence function, one has

$$\text{IF}_x[\mathbf{T}] = \lim_{s \downarrow 0} \frac{\mathbf{T}(F^{(s,x)}) - \mathbf{T}(F)}{s},$$

where  $F^{(s,x)}$  is defined for  $x \in \mathbf{R}$  and  $s > 0$  as

$$\text{for } u \in \mathbf{R}, \quad F^{(s,x)}(u) = s 1_{\{x \leq u\}} + (1-s)F(u).$$

**Theorem 6** (IF of  $S_t$ ). For any Borel set  $A \subset \mathbf{R}$ ,

$$\text{IF}_x[\mathbf{P}[S_t \in A]] = \lambda t (\mathbf{P}[x + S_t \in A] - E_Y[\mathbf{P}[Y + S_t \in A]]),$$

where  $Y$  is distributed as  $W$ .

**Proof.** Let  $n$  be the number of claims before time  $t$ , and set  $p_n = \mathbf{P}[N_t = n]$ . Then

$$\mathbf{P}[S_t^{(s,x)} \in A] = \mathbf{P}\left[\sum_{n=0}^{+\infty} p_n \mathbf{P}[W^{(s,x)*n} \in A]\right].$$

But note that

$$\mathbf{P}[W^{(s,x)*n} \in A] = \sum_{k=0}^n \binom{n}{k} s^k (1-s)^{n-k} \mathbf{P}[kx + W^{*n-k} \in A].$$

Taking the derivative of the above expression with respect to  $s$  when  $s = 0$ , one obtains

$$\begin{aligned} \text{IF}_x[\mathbf{P}[S_t \in A]] &= \sum_{n=0}^{\infty} \left( p_n \binom{n}{1} \mathbf{P}[x + W^{*n-1} \in A] \right. \\ &\quad \left. - p_n \binom{n}{0} n \mathbf{P}[W^{*n} \in A] \right), \end{aligned}$$

and it follows that, using  $p_n = \lambda t p_{n-1}/n$  when  $n \geq 1$ ,

$$\begin{aligned} \text{IF}_x[\mathbf{P}[S_t \in A]] &= \lambda t \sum_{n=0}^{\infty} \left( p_n \mathbf{P}[x + W^{*n} \in A] \right. \\ &\quad \left. - p_n \mathbf{P}[Y + W^{*n} \in A] \right), \end{aligned}$$

where  $Y$  is distributed as  $W$ .  $\square$

##### 4.2. IF associated with ruin probabilities

**Theorem 7** (IF Associated with  $\varphi(0, t)$ ). The influence function of the ruin probability starting from  $R_0 = 0$  is given by

$$\text{IF}_x[\varphi(0, t)] = \lambda t (\varphi^x(0, t) - E_Y[\varphi^Y(0, t)]),$$

where  $\varphi^x(0, t)$  is the ruin probability of the modified risk process  $R_t^x$  defined previously, where  $\varphi^x(0, t) = \frac{1}{ct} E[(ct - x - S_t)_+]$ .

**Proof.** Set  $\text{IF}_{x,j}(n) = \text{IF}_x \mathbf{P}[S_{n/c} = j]$ , and consider

$$\text{IF}_x \left[ \varphi \left( 0, \frac{n}{c} \right) \right] = \sum_{j=0}^n \frac{n-j}{n} \text{IF}_{x,j}(n),$$

which can be obtained by using Takac's results (see Loisel et al. (2008), Proposition 9). The result follows then directly from the definition of  $\varphi^x(0, n/c)$  and the interpretation of  $\varphi^x(0, n/c)$  in the first sections.  $\square$

**Remark 4.** Theorem 7 links the derivative with respect to  $\lambda$  of the ruin probability with the influence function since, as we already checked,

$$\frac{\partial}{\partial \lambda} \varphi(0, t) = t E_Y [\varphi^Y(0, t) - \varphi(0, t)].$$

**Theorem 8** (IF Associated with  $\varphi(u, t)$ ). We now turn to the computation of the influence function associated with the ruin probability starting from the initial reserve  $u$ , which is given by

$$\text{IF}_x[\varphi(u, t)] = \lambda t (\varphi^x(u, t) - E_Y[\varphi^Y(u, t)]),$$

where  $\varphi^x(u, t)$  is the probability of ruin associated with the modified risk process  $R_t^x$ , when  $R_0^x = u$ .

**Proof.** We use Proposition 10 of Loisel et al. (2008):

$$\begin{aligned} \text{IF}_x \left[ \varphi \left( u, \frac{n}{c} \right) \right] &= \sum_{i=0}^{u+n} \text{IF}_{x,i}(n) - \sum_{k=1}^n \text{IF}_{x,u+k}(k) \varphi \left( 0, \frac{n-k}{c} \right) \\ &\quad - \sum_{k=1}^n h_{u+k}(k) \text{IF}_x \left[ \varphi \left( 0, \frac{n-k}{c} \right) \right]. \end{aligned}$$

From the two previous theorems, one can transform influence functions as functions of the distribution of  $x + S_t$  and  $Y + S_t$ . One then gets the result by using that

$$\begin{aligned} \mathbf{P} \left[ T_u^x > \frac{n}{c} \right] &= \mathbf{P}[x + S_{n/c} < u + n] \\ &= \sum_{i=1}^{n-1} \frac{i}{n} \mathbf{P}[x + S_{i/c} = u + i] \sum_{j=1}^{n-i} \mathbf{P}[S_{(n-i)/c} = j] \frac{n-i-j}{n-i} \\ &\quad - \sum_{i=1}^{n-1} \frac{n-i}{n} \mathbf{P}[S_{i/c} = u + i] \sum_{j=1}^{n-i} \mathbf{P}[x + S_{(n-i)/c} = j] \frac{n-i-j}{n-i}. \quad \square \end{aligned}$$

**Remark 5.** Note that one gets back logical expressions for the asymptotic variances:

$$\begin{aligned} V_0 &= V_Y [\lambda t \varphi^Y(0, t)] = V_Y [\text{IF}_Y[\varphi(0, t)]], \\ V_u &= V_Y [\lambda t \varphi^Y(u, t)] = V_Y [\text{IF}_Y[\varphi(u, t)]], \quad u > 0, \end{aligned}$$

as well as the following identities:

$$\begin{aligned} E_Y [\text{IF}_Y[\varphi(0, t)]] &= 0, \\ E_Y [\text{IF}_Y[\varphi(u, t)]] &= 0, \quad u > 0. \end{aligned}$$

**Remark 6.** Expressions involving  $Y + S_t$  or  $\varphi^Y(u, t)$  can be linked with the so-called dual risk process (see Mazza and Rullière, 2004).

## References

- Asmussen, S., Avram, F., Usabel, M., 2002. The Erlang approximation of finite time ruin probabilities. *Astin Bulletin* 32, 267–281.
- Bening, V., Korolev, V., 2000. Nonparametric estimation of the ruin probability for generalized risk processes. *Theory of Probability and its Applications* 47-1, 1–16.
- Biard, R., Lefèvre, C., Loisel, S., 2008. Impact of correlation crises in risk theory: Asymptotics of finite-time ruin probabilities for heavy-tailed claim amounts when some independence and stationarity assumptions are relaxed. *Insurance: Mathematics and Economics* 43-3, 412–421.
- Croux, K., Veraverbeke, N., 1990. Nonparametric estimators for the probability of ruin. *Insurance: Mathematics and Economics* 9-2, 3, 127–130.
- Frees, E., 1986. Nonparametric estimation of the probability of ruin. *Astin Bulletin* 16, 81–90.
- Hampel, F., 1974. The influence curve and its role in robust estimation. *Journal of the American Statistical Association* 69, 383–393.
- Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. *Robust Statistics. The Approach Based on Influence Functions*. In: *Wiley Series in Probability and Statistics*. Wiley, New York.
- Hipp, C., 1989. Estimators and bootstrap confidence intervals for ruin probabilities. *Astin Bulletin* 19-1, 57–70.
- Hoeffding, W., 1948. A class of statistics with asymptotically normal distribution. *The annals of Mathematical Statistics* 19 (3), 293–325.
- Huber, P.J., 1981. *Robust Statistics*. In: *Wiley Series in Probability and Statistics*. Wiley, New York.
- Lefèvre, C., Loisel, S., 2008. On finite-time ruin probabilities for classical risk models. *Scandinavian Actuarial Journal* 1, 41–60.
- Lefèvre, C., Loisel, S., 2009. Finite-time ruin probabilities for discrete, possibly dependent, claim severities. *Methodology and Computing in Applied Probability* 11 (3), 425–441.
- Loisel, S., Mazza, C., Rullière, D., 2008. Robustness analysis and convergence of empirical finite-time ruin probabilities and estimation risk solvency margin. *Insurance: Mathematics and Economics* 42-2, 746–762.
- Loisel, S., Privault, N., 2009. Sensitivity analysis and density estimation for finite-time ruin probabilities. *Journal Computational and Applied Mathematics* 230 (1), 107–120.
- Marceau, E., Rioux, J., 2001. On robustness in risk theory. *Insurance: Mathematics and Economics* 29, 167–185.
- Mazza, C., Rullière, 2004. A link between wave governed random motions and ruin processes. *Insurance: Mathematics and Economics* 35, 205–222.
- Picard, P., Lefèvre, C., 1997. The probability of ruin in finite time with discrete claim size distribution. *Scandinavian Actuarial Journal* 1, 58–69.
- Rullière, D., Loisel, S., 2004. Another look at the Picard-Lefèvre formula for finite-time ruin probabilities. *Insurance: Mathematics and Economics* 35, 187–203.
- Takács, L., 1962a. A generalization of the ballot problem and its application in the theory of queues. *Journal of the American Statistical Association* 57, 327–337.
- Takács, L., 1962b. The time dependence of a single-server queue with Poisson input and general service times. *The Annals of Mathematical Statistics* 33, 1340–1348.
- Von Mises, R., 1947. On the asymptotic distribution of differentiable statistical functions. *The Annals of Mathematical Statistics* 18-3, 309–348.