



**HAL**  
open science

# Nonparametric statistical analysis of ruin probability under large claims

Pier Luigi Conti, Esterina Masiello

► **To cite this version:**

Pier Luigi Conti, Esterina Masiello. Nonparametric statistical analysis of ruin probability under large claims. *Extremes*, 2010, 13 (4), pp.439-461. 10.1007/s10687-009-0094-6 . hal-01018874

**HAL Id: hal-01018874**

**<https://hal.science/hal-01018874>**

Submitted on 7 Jul 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Nonparametric statistical analysis of ruin probability under large claims

Pier Luigi Conti, Esterina Masiello  
Dipartimento di Statistica, Probabilità e Statistiche Applicate  
Università di Roma “La Sapienza”  
P.le A. Moro, 5  
Roma, 00185, Italy  
{pierluigi.conti, esterina.masiello}@uniroma1.it

## Abstract

In this paper, the classical Poisson risk model is considered. The claims are supposed to be modeled by heavy-tailed distributions, so that the moment generating function does not exist. The attention is focused on the probability of ruin. We first provide a nonparametric estimator of an upper bound of the ruin probability by Willmot and Lin. Then, its asymptotic behavior is studied. Asymptotic confidence intervals are studied, as well as bootstrap confidence intervals. Results for possibly unstable models are also obtained.

*Key words and phrases.* Poisson risk model, probability of ruin, nonparametric estimation, asymptotics, heavy-tailed distribution.

*AMS 2000 subject classifications.* 62G08, 62G20, 62G32.

## 1 Introduction and preliminaries

The interest of actuarial mathematics in ruin probabilities dates back to the seminal paper by Lundberg (1903). Since then, several contributions have been given, the main trouble being the difficulty in finding a general expression for the probability of ruin (except for a few special cases).

We consider here the classical risk model, where claims occur as a Poisson process. The assumptions on which such a model rests are listed below.

A1. The claim sizes  $X_1, X_2, \dots$  are positive independent and identically distributed (*i.i.d.*) random variables (r.v.'s), with common distribution function (d.f.)  $F(x) = P(X_1 \leq x)$ , finite mean  $\mu = E(X_1)$  and finite, positive, variance  $\sigma^2 = Var(X_1)$ .

A2. The inter-arrival times  $T_1, T_2, \dots$  between consecutive claims are *i.i.d.* exponentially distributed r.v.s, with mean  $\mu_T = E(T_1) = 1/\lambda$ .

A3. The two sequences  $(X_i; i \geq 1)$ ,  $(T_i; i \geq 1)$  are independent.

A4. The initial capital is  $x > 0$ , and the premium income is linear in time with rate  $c$ . For the sake of simplicity, and without loss of generality, in the sequel we will assume that  $c = 1$  (otherwise, it is enough to refer to the transformed claims  $X'_i = X_i/c$ ).

The function of interest is the ruin probability  $\Psi(x)$ , namely the probability that the company will be ruined if the claims will exceed its available resources. Almost all statistical analysis of the ruin probability in risk business are based on

the assumption that the moment generating function (m.g.f.) of the claim size distribution is finite for some positive real argument.

The Lundberg model outline before possess a natural counterpart in queueing theory, where claims are replaced by service times. In that framework, the Lundberg model (with  $c = 1$ ) essentially corresponds to the M/G/1 model, with *i.i.d.* exponential inter-arrival times and general *i.i.d.* service times. The ruin probability  $\Psi(x)$  turns out to be the probability that the equilibrium waiting time is greater than  $x$ . Because of this correspondence, all results obtained in this paper are immediately applicable to queueing models.

The main result in evaluating  $\Psi(x)$  is the celebrated Lundberg's inequality. Assume that there exists a positive  $u$  such that:

$$\int_0^{\infty} e^{ux} dF(x) < \infty$$

Then, Lundberg's inequality states an exponential bound in the initial capital  $x$ :

$$\Psi(x) \leq e^{-Rx} \quad \forall x \geq 0 \quad (1)$$

for an appropriate, positive,  $R$ , such that  $x^{-1} \log \Psi(x) \rightarrow -R$  as  $x$  goes to infinity.

Statistical analysis of  $R$  under assumptions A1-A4 is in Gaver and Jacobs (1988), under the additional assumption that  $\lambda$  is known. See also Pitts *et al.* (1996), as well as the references therein, for the general case of non-exponential inter-arrival times. Sequential analysis of  $R$  is in Conti (2005).

In many practical cases the claim size distribution is well described by d.f.'s for which the above condition of existence of the m.g.f. is not satisfied. As a consequence, Lundberg's inequality does not hold. A case of practical relevance is when claims are modelled by heavy-tailed distributions, such as Pareto, log-gamma, log-normal or heavy-tailed Weibull distributions, that do not have exponential moments. In this case, a condition of "high claims" is met. High claims occur with relatively small probability, and are difficult to predict because they are not generally caused by man-made disasters. Moreover, it has a significant influence on the behavior of the whole model. Under high claims, the ruin probability is (expected to be) considerably larger than the classical case of finite m.g.f.. For this reason, it is of interest to provide a statistical analysis of the ruin probability when the upper bound (1) does not hold.

In this paper, we deal with "high claims" modelled by heavy-tailed distributions. The basic idea consists in considering an appropriate approximation of  $\Psi(x)$  ("for  $x$  large") and estimate this approximation. Clearly, classical risk theory has to be adapted to this new case. Refinements and generalizations of the inequality (1) have been considered by several authors; see, for instance, Gerber (1973), Ross (1974), Stoyan (1983), Willmot and Lin (1994), but they are still based exponential bounds.

We can actually mention not so many works devoted to bounds of ruin probabilities that can be applied to claim size distributions with no exponential moments. The following list, although incomplete, contains the most important papers. An interesting approach to the problem consists in constructing general bound for  $\Psi(x)$  without the Cramér-Lundberg condition or other similar conditions; see, for example, De Vylder and Goovaerts (1984), Broeckx *et al.* (1986). A different approach

simply consists in relaxing the Cramér-Lundberg condition by considering functions different from the exponential function. Following Willmot (1994), Willmot (1996), Willmot (1997), Lin (1997), Cai and Wu (1997), Willmot and Lin (1997), we could consider a “new worse than used” d.f. of claims, and then derive either upper or lower bounds for  $\Psi(x)$ .

An excellent monograph is the volume by Willmot and Lin (2000). It is devoted to approximations and bounds for the tail of compound distributions (*i.e.* distributions of random sums). Random sums are of primary importance in insurance, as well as in other areas of applied probability, such as queueing theory. Bounds in the book by Willmot and Lin (2000) are motivated by the classical Lundberg exponential inequalities, which applies to ruin probabilities. They simply replace the exponential term, involved in the classical bound, by the tail of a distribution which possesses a thicker tail than the exponential one, such as a Pareto-type tail. The connection to compound distributions is through the interpretation of the ruin probability as the tail probability of a compound geometric distribution.

In the sequel, we will consider first the Pareto bound introduced by Willmot (1994), Willmot and Lin (2000), as a suitable approximation of  $\Psi$ . Then, we will study statistical inference problems for such bound.

The paper is structured as follows. We first describe, in Section 2, the Pareto upper bound introduced by Willmot (1994), and adapt it to an insurance context. Then, in Section 3, we introduce a nonparametric estimator for such a bound and study its asymptotic behavior. The behavior of our estimator is also considered in the particular case of unstable model in Section 4.

## 2 Bounds for the ruin probability

Let  $N$  be a r.v. taking non-negative values, let  $p_k = P(N = k)$ , and let  $(Y_i; i \geq 1)$  be a sequence of *i.i.d.* r.v.s (independent of  $N$ ) having common d.f.  $H$ . Denote further by  $a_n$  the quantity (essentially, the survival function of  $N$ ):

$$a_n = P(N > n) = \sum_{k=n+1}^{\infty} p_k \quad n = 0, 1, 2, \dots$$

Our next assumptions refers to the moments of  $H$ , and to the possibility of bounding the (right) tail of  $H$ .

A5. The moments of  $Y_i$ s exist up to order  $m$  (where  $m > 0$ ). In symbols:

$$E(Y_i^j) = \int_0^{\infty} x^j dH(x) < \infty \quad \forall j \leq m.$$

A6. There is  $k_0 > 0$  satisfying

$$\int_0^{\infty} (1 + k_0 x)^{m+1} dH(x) = \frac{1}{\phi} \quad (2)$$

where  $\phi < 1$  is a positive number such that:

$$a_{n+1} \leq \phi a_n \quad n = 0, 1, 2, \dots \quad (3)$$

Define next the random sum  $S_N = Y_1 + \dots + Y_N$ , and let  $\bar{G}(x) = P(S_N \geq x)$ . The following bounds are obtained by combining Corollaries 6.2.1 and 6.2.3 in Willmot and Lin (2000).

**Proposition 1** *Under conditions A5, A6, the following inequality*

$$\bar{G}(x) \leq \frac{1-p_0}{\phi} (1+k_0x)^{-m}, \quad x \geq 0$$

holds.  $\square$

Consider now the classical Poisson risk model. As already explained in the Introduction,  $X_i$  denotes the  $i$ -th claim size and  $T_i$  denotes the  $i$ -th inter-arrival time. The m.g.f. of the d.f.  $F$  of the  $X_i$ 's does not exist. From now on, throughout this section, we assume that the model is *stable*, i.e. that the parameter  $\rho = (\lambda\mu)^{-1}$  is strictly smaller than 1 (recall that  $c$  is assumed equal to 1).

Let

$$F_I(x) = \frac{1}{\mu} \int_0^x y dF(y)$$

be the integrated tail distribution and denote by  $F_I^{*k}$  the  $k$ -fold convolution of the d.f.  $F_I$  (with  $F_I^{*0} = 0$  for every positive  $x$ ). The d.f.  $F_I$  is absolutely continuous w.r.t.  $F$ , with Radon-Nikodym derivative  $dF_I(x)/dF(x) = x/\mu$ .

As well known, under the assumption of the Cramér-Lundberg model and if  $\rho < 1$ , the probability of ruin  $\Psi$  can be expressed as a compound geometric tail probability by mean of the so-called Pollaczec-Khinchine formula (see, e.g., Feller (1971), Grandell (1990) or Embrechts and Veraverbeke (1982) ). In symbols:

$$\Psi(x) = \sum_{k=1}^{\infty} (1-\rho)\rho^k \bar{F}_I^{*k}(x) \quad x \geq 0. \quad (4)$$

Apart for a few special cases, (4) cannot be expressed in a closed form. Hence, it is fairly natural to resort to some (possibly simple) bound of (4). Assume that the moments of  $F_I$  exist up to order  $m > 0$ , i.e. that

$$\int_0^{\infty} x^m dF_I(x) < \infty. \quad (5)$$

Clearly, (5) is fulfilled if and only if the moments of the claim size distribution exist up to order  $m+1$ , i.e. if

$$E(X^{m+1}) = \int_0^{\infty} x^{m+1} dF(x) < \infty$$

Now, for the geometric distribution with  $p_k = (1-\rho)\rho^k$ , ( $k = 0, 1, 2, \dots$ ), we may write

$$a_{n+1} = \sum_{k=n+2}^{\infty} p_k$$

$$\begin{aligned}
&= \sum_{k=n+2}^{\infty} (1-\rho)\rho^k \\
&= \rho \sum_{k=n+2}^{\infty} (1-\rho)\rho^{k-1} \\
&= \rho \sum_{k=n+1}^{\infty} (1-\rho)\rho^k \\
&= \rho a_n
\end{aligned} \tag{6}$$

that clearly implies  $\phi = \rho$ . Furthermore, from elementary properties of the geometric distribution, it is not difficult to see that  $p_0 = 1 - \rho$ .

Finally, suppose there exists a positive  $k_0$  satisfying assumption A6 with  $H = F_I$ , namely:

$$\int_0^{\infty} (1+kx)^m dF_I(x) = \frac{1}{\rho}. \tag{7}$$

From Proposition 1 it is immediate to draw the following result.

**Proposition 2** *Under conditions A1-A4, if assumptions A5, A6 hold when  $H = F_I$ , and if  $\rho < 1$ , then the inequality*

$$\Psi(x) \leq \Psi_U = (1+kx)^{-m} \tag{8}$$

holds for every non-negative  $x$ .  $\square$

Inequality (8) provides a useful approximation of  $\Psi(x)$  (namely,  $\Psi(x) \approx \Psi_U(x)$ ), to be used when the m.g.f. of the claim size d.f. does not exist. The problem we deal with in next section is the construction of a nonparametric point estimator of  $\Psi_U(x)$  in (8).

### 3 Basic statistical analysis

Suppose the observed data  $n$  are *i.i.d.* inter-arrival times  $T_1, \dots, T_n$  and  $n$  *i.i.d.* claim sizes  $X_1, \dots, X_n$ . The two sequences  $(T_i; i \geq 1)$ ,  $(X_i; i \geq 1)$  are assumed to satisfy conditions A1-A4, again with  $c = 1$ . As far as claim size distribution is concerned, we do not assume any specific parametric model for  $F$ . Our unique parametric assumption is that  $T_i$ s are *i.i.d.* exponential r.v.s, with unknown parameter  $\lambda$ .

Our goal is to construct a nonparametric estimator of  $\Psi_U(x)$ . In view of (8), the main problem consists in defining an estimator of  $k_0$ .

To simplify the notation, let  $B(k, m)$  be defined as

$$\begin{aligned}
B(k, m) &= \mu \int_0^{\infty} (1+kx)^m dF_I(x) \\
&= \int_0^{\infty} x (1+kx)^m dF(x) \\
&= E[X_1 (1+kX_1)^m].
\end{aligned} \tag{9}$$

Throughout the present section, we will always suppose that the stability condition  $\rho < 1$  is met. The function

$$G(k, m) = \frac{1}{\mu} B(k, m) - 1/\rho \quad (10)$$

is strictly increasing in  $k$ , with  $G(0, m) = 1 - 1/\rho < 0$  and  $G(k, m) \uparrow \infty$  as  $k \uparrow \infty$ . Hence, as  $H = F_I$ ,  $k_0$  is the unique root of the equation  $G(k, m) = 0$ .

On the basis of our sample data, it is natural to estimate  $\lambda$  by its maximum likelihood estimator (MLE, for short)

$$\hat{\lambda} = \bar{T}^{-1} \quad (11)$$

where  $\bar{T}$  is the sample average of inter-arrival times:

$$\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i.$$

As a “natural” estimator of  $\rho$ , we then take:

$$\hat{\rho} = \bar{X}/\bar{T} \quad (12)$$

where  $\bar{X}$  is the sample average of claim sizes:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Next step consists in constructing an estimator of  $B(k, m)$ . First of all, let  $\hat{F}(x)$  be the empirical distribution function (e.d.f., for short) of the claim sizes:

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n I_{(X_i \leq x)}$$

where

$$I_{(X_i \leq x)} = \begin{cases} 1 & \text{if } X_i \leq x \\ 0 & \text{if } X_i > x \end{cases}. \quad (13)$$

Hence, a “natural” estimator of  $F_I(x)$  is

$$\begin{aligned} \hat{F}_I(x) &= \frac{1}{\bar{X}} \int_0^x y d\hat{F}(y) \\ &= \frac{1}{n \bar{X}} \sum_{i=1}^n X_i I_{(X_i \leq x)} \end{aligned} \quad (14)$$

Note that  $\hat{F}_I(x)$  corresponds to a discrete distributions giving mass  $X_i/(X_1 + \dots + X_n)$  to each sample claim size  $X_i$ ,  $i = 1, \dots, n$ .

On the basis of (14), it is not difficult to construct the following estimator of  $B(k, m)$ :

$$\begin{aligned}\widehat{B}(k, m) &= \bar{X} \int_0^\infty (1 + kx)^m d\widehat{F}_I(x) \\ &= \frac{1}{n} \sum_{i=1}^n X_i (1 + kX_i)^m\end{aligned}\quad (15)$$

as well as the following estimator of (10):

$$\widehat{G}(k, m) = \frac{1}{\bar{X}} \widehat{B}(k, m) - \frac{1}{\widehat{\rho}}. \quad (16)$$

Now,  $\widehat{G}(k, m)$  is strictly increasing in  $k$ , with  $\widehat{G}(0, m) = 1 - 1/\widehat{\rho}$  and  $\widehat{G}(k, m) \uparrow \infty$  as  $k$  increases. Hence, if  $\widehat{\rho} < 1$ , the equation  $\widehat{G}(k, m) = 0$  possesses only one root,  $k_n$ , say. On the other hand, if  $\widehat{\rho} \geq 1$ , the equation  $\widehat{G}(k, m) = 0$  possesses no root. In the sequel, we will then define  $k_n$ , our estimate of  $k_0$ , as

$$k_n = \inf\{k > 0 : G_n(k, m) \geq 0\}. \quad (17)$$

Note that  $k_n > 0$  whenever  $\widehat{\rho} < 1$ , and  $k_n = 0$  whenever  $\widehat{\rho} \geq 1$ .

**Lemma 1** *In assumptions A1-A4 are fulfilled, and if A5, A6 are satisfied by  $F_I$ , then  $G_n(k, m)$  tends to  $G(k, m)$  with probability one, for every fixed, non-negative  $k$ :*

$$G_n(k, m) \xrightarrow{a.s.} G(k, m) \text{ as } n \rightarrow \infty, \quad \forall k \geq 0. \quad (18)$$

**Proof** See Appendix.  $\square$

In next proposition, the strong consistency of the estimator  $k_n$  is shown.

**Proposition 3** *Under the assumptions of Lemma 1, we have*

$$k_n \xrightarrow{a.s.} k_0 \quad (19)$$

*as the sample size  $n$  goes to infinity.*

**Proof** See Appendix.  $\square$

Next step consists in proving the asymptotic normality of  $k_n$ , under appropriate additional assumptions. We begin by some preliminary lemmas, that are worth for proving the main result of the present section (Proposition 5).

**Lemma 2** *Suppose that  $k_n^*$  lies in between  $k_n$  and  $k_0$ , and let  $G^{(j)}(k, m)$  ( $G_n^{(j)}(k, m)$ ) be the  $j$ th derivative of  $G(k, m)$  ( $G_n(k, m)$ ) w.r.t.  $k$  ( $j = 1, 2$ ). Under the assumptions of Lemma 1, we have:*

$$G_n^{(j)}(k_n^*, m) \xrightarrow{a.s.} G^{(j)}(k_0, m) \text{ as } n \rightarrow \infty. \quad (20)$$



**Proof** See Appendix.  $\square$

From the multidimensional central limit theorem for sums of *i.i.d.* random vectors, (see, *e.g.*, Serfling (1980), p. 28) it is then easy to prove the following further lemma, which is of considerable importance for all subsequent developments.

**Lemma 3** *Under the same assumptions of Lemma 1, and if  $\rho < 1$ , and if*

$$C(k_0, 2m) = \int_0^\infty x^2 (1 + k_0 x)^{2m} dF(x) < \infty, \quad (21)$$

*the sequence of (three-dimensional) random vectors*

$$(\sqrt{n}(B_n(k_0, m) - B(k_0, m)), \sqrt{n}(\bar{X} - \mu), \sqrt{n}(\bar{T} - \lambda^{-1}); n \geq 1)$$

*tends in distribution, as  $n$  goes to infinity, to a (three-dimensional) normal random vector, with null mean vector and covariance matrix:*

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{13} \\ \sigma_{11} & \sigma_{23} & \sigma_{33} \end{bmatrix} \quad (22)$$

where  $\sigma_{ij} = \sigma_{ji}$ , and

$$\begin{aligned} \sigma_{11} &= \text{Var}(X_1 (1 + k_0 X_1)^m) = C(k_0, 2m) - B(k_0, m)^2, \\ \sigma_{22} &= \text{Var}(X_1) = \sigma^2, \quad \sigma_{33} = \text{Var}(T_1) = \lambda^{-2} \\ \sigma_{12} &= \text{Cov}(X_1, X_1 (1 + k_0 X_1)^m) = C(k_0, m) - \mu B(k_0, m) \\ \sigma_{13} &= \sigma_{23} = 0. \end{aligned} \quad \square$$

**Proposition 4** *Under the same assumptions of Lemma 1, and if  $\rho < 1$ , the sequence of r.v.s*

$$\sqrt{n}(G_n(k_0, m) - G(k_0, m)), \quad n \geq 1 \quad (23)$$

*tends in distribution, as  $n$  goes to infinity, to a normal variate with zero mean and variance*

$$\sigma_G^2 = \mathbf{a}^T \Sigma \mathbf{a}, \quad (24)$$

where

$$\mathbf{a} = \begin{bmatrix} \mu^{-1} \\ -B(k_0, m)/\mu^2 - \lambda \\ \mu \lambda^2 \end{bmatrix}. \quad (25)$$

**Proof** See Appendix.  $\square$

We are now in a position to prove the main result of the present section, *i.e.* the asymptotic normality of  $k_n$  (if properly rescaled)

**Proposition 5** Under the same assumptions of Lemma 1, if  $\rho < 1$ , and if  $G^{(1)}(k_0, m) \neq 0$ , the sequence of r.v.s

$$\sqrt{n}(k_n - k_0), \quad n \geq 1$$

tends in distribution to a normal variate with zero mean, and variance:

$$\sigma_k^2 = G^{(1)}(k_0, m)^{-2} \sigma_G^2 \quad (26)$$

where  $\sigma_G^2$  is given by (24).

**Proof** See Appendix.  $\square$

On the basis of the estimate  $k_n$ , it is of course possible to construct the following, simple estimator of  $\Psi_U(x)$ , the upper bound of the ruin probability in (8):

$$\widehat{\Psi}_U(x) = (1 + k_n x)^{-m} \quad (27)$$

From Proposition 3 and Proposition 5, using again the delta method, it is easy to obtain the following further result, establishing the asymptotic normality of  $\widehat{\Psi}_U(x)$ .

**Proposition 6** Under the same assumptions of Proposition 5, the sequence of r.v.s

$$\sqrt{n} \left( \widehat{\Psi}_U(x) - \Psi_U(x) \right), \quad n \geq 1 \quad (28)$$

tends in distribution to a normal variate with zero mean and variance

$$\sigma_U^2 = \left( m x (1 + k_0 x)^{-(m+1)} \right)^2 \sigma_k^2 \quad (29)$$

as  $n$  goes to infinity.  $\square$

In order to construct an asymptotic confidence interval for  $\Psi_U(x)$ , we need a consistent estimator of its asymptotic variance  $\sigma_U^2$  in (29).

Let  $\widehat{\sigma}_U^2$  be defined exactly as  $\sigma_U^2$ , except that the unknown quantities  $k_0, \mu, \lambda, B(k_0, m), C(k_0, 2m), G^{(1)}(k_0, m)$  are replaced by their consistent estimators  $k_n, \overline{X}, \widehat{\lambda}, B_n(k_n, m), C_n(k_n, j)$  ( $j = m, 2m$ ),  $G_n^{(1)}(k_n, m)$ , respectively, where

$$C_n(k_n, j) = \frac{1}{n} \sum_{i=1}^n X_i^2 (1 + k_n X_i)^j.$$

In Proposition 7, whose proof is immediate, an asymptotic confidence interval of size  $0 < \alpha < 1$  for  $\Psi_U(x)$  is provided.

**Proposition 7** Under the assumptions of Proposition 5, as  $n \rightarrow \infty$ , the following results hold true.

(i)  $\widehat{\sigma}_U^2 \xrightarrow{a.s.} \sigma_U^2$ ;

(ii)  $\sqrt{n}(\widehat{\Psi}_U(x) - \Psi_U(x))/\widehat{\sigma}_U$  tends in distribution to a normal standard variate;

(iii) If  $z_\alpha$  denotes the  $\alpha$ -th quantile of the normal standard distribution, then a confidence interval for  $\Psi_U(x)$ , of asymptotic size  $(1 - \alpha)$ , is given by

$$\left[ \widehat{\Psi}_U(x) - z_{\alpha/2} \widehat{\sigma}_U / \sqrt{n}, \widehat{\Psi}_U(x) + z_{\alpha/2} \widehat{\sigma}_U / \sqrt{n} \right]. \quad \square \quad (30)$$

## 4 The case of unstable models

We have worked so far under the assumption of stability of the model, namely  $\rho < 1$ . Virtually all papers concerned with statistical analysis of ruin models are actually based on the *a priori* assumption that the model under consideration is stable. Hence, it is of interest to study how robust are the result of statistical analysis when the hypothesis of stability is dismissed. Furthermore, in our opinion the case  $\rho$  “close” to 1 is of special importance, since it corresponds to the possibility of severely high claims. In this section, we deal with statistical analysis of possibly unstable models, where the assumption  $\rho < 1$  is dismissed.

First of all, consider again the function (10), extended now to the case  $\rho \geq 1$ ,

$$G(k, m) = \begin{cases} \frac{1}{\mu} B(k, m) - \frac{1}{\rho} & \text{if } \rho < 1 \\ 0 & \text{if } \rho \geq 1 \end{cases}$$

and denote by  $k_0$  the smallest positive root of the equation  $G(k, m) = 0$ :

$$k_0 = \inf\{k > 0 : G(k, m) \geq 0\}.$$

The value of  $k_0$  is different according to the value of  $\rho$ :  $k_0 > 0$  if  $\rho < 1$  and  $k_0 = 0$  if  $\rho \geq 1$ . Furthermore, it is not difficult to see that in the case  $k_0 = 0$ , (24) reduces to:

$$\sigma_G^2 = \frac{\lambda^2}{\rho^4} (\sigma^2 + \mu^2). \quad (31)$$

We first deal with the case  $\rho = 1$ , which is more difficult (and more interesting, as well) than the case  $\rho > 1$ . The following proposition contains the main result of the present section.

**Proposition 8** *Assume that  $G^{(1)}(0, m) \neq 0$ , and that all hypotheses of Lemma 3 are met, except the stability assumption  $\rho < 1$ . Define in this case  $\sigma_k^2$  as*

$$\begin{aligned} \sigma_k^2 &= G^{(1)}(0, m)^{-2} \sigma_G^2 \\ &= G^{(1)}(0, m)^{-2} \frac{\lambda^2}{\rho^4} (\sigma^2 + \mu^2). \end{aligned} \quad (32)$$

If  $\rho = 1$ , we then have

$$\lim_{n \rightarrow \infty} P(\sqrt{n} k_n \geq x) = \begin{cases} 1 & \text{if } x < 0 \\ 1 - \Phi(x/\sigma_k) & \text{if } x > 0 \end{cases}$$

where  $\Phi(\cdot)$  is the standard normal distribution function.

**Proof** See Appendix.  $\square$

If  $\rho > 1$ , then  $\sqrt{n} k_n$  tends in law to a random variate degenerate at the point zero. It is enough to take into account the consistency of  $\hat{\rho}$  and the relationships:

$$\lim_{n \rightarrow \infty} P(\sqrt{n} k_n = 0) = \lim_{n \rightarrow \infty} P(\hat{\rho} \geq 1) = 1. \quad (33)$$

Next, observe first that when  $\rho = 1$   $\sigma_U$  reduces to

$$\sigma_U^2 = m^2 x^2 \sigma_k^2. \quad (34)$$

As a simple consequence of Proposition 7, using the delta method we have the following further result, whose proof is immediate.

**Proposition 9** *Under the same assumptions as in Proposition 8, the following results hold true.*

(i) *If  $\rho = 1$ , then*

$$\lim_{n \rightarrow \infty} P(\sqrt{n}((1 + k_n x)^{-m} - 1) \geq x) = \begin{cases} 1 & \text{if } x < 0 \\ 1 - \Phi(x/\sigma_U) & \text{if } x > 0 \end{cases}$$

where  $\sigma_U$  is given by (34).

(ii) *If  $\rho > 1$ , then  $\sqrt{n}((1 + k_n x)^{-m} - 1) \geq x$  converges in law to a random variate degenerate at the point zero.  $\square$*

Proposition 8 shows that the asymptotic normality of  $k_n$  only holds under the stability condition  $\rho < 1$ . In order to study how this result affects the behaviour of the confidence interval for  $\Psi_U(x)$  (30), let  $\hat{\sigma}_U^2$  be defined exactly as in Section 3. When  $\rho = 1$ , we have again

$$\hat{\sigma}_U^2 \xrightarrow{a.s.} \sigma_U^2$$

as  $n$  goes to infinity. As a consequence, we may establish the following further results.

**Proposition 10** *Under the assumptions of Lemma 1, and if and if  $G^{(1)}(k_0, m) \neq 0$ , then:*

(i) *If  $\rho < 1$ , the confidence interval (30) has asymptotic level  $1 - \alpha$ , as  $n$  increases;*

(ii) *If  $\rho = 1$ , the confidence interval (30) has asymptotic level  $1 - \alpha/2$ , as  $n$  increases;*

(ii) *If  $\rho > 1$ , the confidence interval (30) has asymptotic level 1, as  $n$  increases.  $\square$*

On the basis of Proposition 10, we may conclude that the confidence interval (30) has asymptotic level not smaller than  $1 - \alpha$  as  $n$  goes to infinity. Hence, it is a conservative confidence interval of level  $1 - \alpha$  for  $\Psi_U(x)$ .

## Appendix

**Proof of Lemma 1** From the strong law of large numbers, it is immediate to see that

$$\bar{T} \xrightarrow{a.s.} \lambda^{-1} \quad (35)$$

$$\bar{X} \xrightarrow{a.s.} \mu \quad (36)$$

$$\frac{1}{n} \sum_{i=1}^n X_i (1 + kX_i)^m \xrightarrow{a.s.} E[X_1 (1 + kX_1)^m] = B(k, m) \quad (37)$$

as  $n$  goes to infinity. Relationship (18) easily follows from (35)-(37).  $\square$

**Proof of Proposition 3** Suppose first that  $\rho < 1$ . The function  $G(\cdot, m)$  is strictly increasing in a neighborhood of  $k_0$ . Hence, there exists  $\delta > 0$  such that

$$G(k_0 - \delta, m) < 0$$

and

$$G(k_0 + \delta, m) > 0$$

Hence, the set of all sequences  $(t_i; i \geq 1)$  and  $(x_i; i \geq 1)$  of inter-arrival times and claims such that

$$G_n(k, m) \rightarrow G(k, m)$$

as  $n \rightarrow \infty$  uniformly in  $[k_0 - \delta, k_0 + \delta]$  does have probability 1, as a consequence of Lemma 1.

Consider now two sequences  $(t_i; i \geq 1)$ ,  $(x_i; i \geq 1)$  in this set. For every  $\epsilon > 0$ , there exists  $n_0$ , which depends on  $\epsilon$ ,  $(t_i; i \geq 1)$  and  $(x_i; i \geq 1)$ , such that

$$G(k, m) - \epsilon \leq G_n(k, m) \leq G(k, m) + \epsilon$$

for every  $n \geq n_0$  and for every  $k \in [k_0 - \delta, k_0 + \delta]$ . The above relationships imply that

$$G_n(k_0 - \delta, m) \leq G(k_0 - \delta, m) + \epsilon$$

and

$$G_n(k_0 + \delta, m) \geq G(k_0 + \delta, m) - \epsilon$$

for every  $n \geq n_0$ . Taking  $\epsilon = \epsilon(\delta) = 1/2 \min(-G(k_0 - \delta, m), G(k_0 + \delta, m))$ , we have

$$G_n(k_0 - \delta, m) < 0$$

and

$$G_n(k_0 + \delta, m) > 0$$

for every  $n \geq n_0$ . These two last relationships imply that, for every  $\delta > 0$ , there exists  $n_0$  (depending on  $\delta$ ,  $(t_i; i \geq 1)$  and  $(x_i; i \geq 1)$ ), such that

$$k_n \in [k_0 - \delta, k_0 + \delta] \quad \forall n \geq n_0$$

which allows to establish (19) as the sample size  $n$  tends to infinity.

If  $\rho = 0$ , then  $G(0, m) = 0$ ,  $G(\delta, m) > 0$  for every positive  $\delta$ . The same technique used in the case  $\rho < 1$  allows us to prove that for every positive  $\delta$  we have  $k_n \in [0, \delta]$  for all  $n$  sufficiently large, with probability one, and this is enough to establish (19). Finally, the case  $\rho > 1$  can be dealt with similarly.  $\square$

**Proof of Lemma 2** For the sake of simplicity, we confine ourselves to the case  $j = 1$ , the case  $j = 2$  being similar.

First of all, observe that  $G^{(1)}(k, m)$  is increasing w.r.t.  $k$ . If  $k_n \in (k_0 - \epsilon, k_0 + \epsilon)$ , where  $0 < \epsilon < k_0$ , because of  $k_n \xrightarrow{a.s.} k_0$ , we may write

$$G_n^{(1)}(k_0 - \epsilon, m) < G_n^{(1)}(k_n^*, m) < G_n^{(1)}(k_0 + \epsilon, m)$$

It suffices to take into account that

$$G_n^{(1)}(k_n^*, m) \leq G_n^{(1)}(\max(k_n, k_0), m) < G_n^{(1)}(k_0 + \epsilon, m)$$

and

$$G_n^{(1)}(k_n^*, m) \geq G_n^{(1)}(\min(k_n, k_0), m) > G_n^{(1)}(k_0 - \epsilon, m)$$

By the strong law of large number, we then obtain

$$\begin{aligned} G_n^{(1)}(k, m) &= \frac{1}{\bar{X}} B_n^{(1)}(k, m) \\ &= \frac{m}{n} \sum_{i=1}^n (1 + kX_i)^{m-1} X_i^2 \\ &\xrightarrow{a.s.} m E[(1 + kX_1)^{m-1} X_1^2] \\ &= \frac{1}{\mu} B^{(1)}(k, m) \\ &= G'(k, m). \end{aligned}$$

As a consequence, it seen that

$$G_n^{(1)}(k, m) \xrightarrow{a.s.} G^{(1)}(k, m)$$

for every non-negative  $k$ , so that for all sufficiently large  $n$

$$G^{(1)}(k_0 - \epsilon, m) < G_n^{(1)}(k_n^*, m) < G^{(1)}(k_0 + \epsilon, m)$$

which implies, in its turn, relationship (20).  $\square$

**Proof of Proposition 4** In the first place, it is possible to write

$$\begin{aligned} \sqrt{n}(G_n(k_0, m) - G(k_0, m)) &= \frac{1}{\bar{X}} \sqrt{n}(B_n(k_0, m) - B(k_0, m)) \\ &+ B(k_0, m) \sqrt{n} \left( \frac{1}{\bar{X}} - \frac{1}{\mu} \right) - \frac{1}{\bar{T}} \sqrt{n}(\bar{X} - \mu) \\ &- \mu \sqrt{n} \left( \frac{1}{\bar{T}} - \lambda \right). \end{aligned} \tag{38}$$

Now, an elementary application of the delta-method (see, *e.g.*, Serfling (1980), p. 118 ff.) shows that

$$\begin{aligned}\sqrt{n} \left( \frac{1}{\bar{X}} - \frac{1}{\mu} \right) &= -\frac{1}{\mu^2} (\bar{X} - \mu) O_p(n^{-1/2}) \\ \sqrt{n} \left( \frac{1}{\bar{T}} - \lambda \right) &= -\lambda^2 (\bar{T} - \lambda^{-1}) + O_p(n^{-1/2})\end{aligned}$$

so that, as a consequence of Slutsky theorem (Serfling (1980), p. 19) we may conclude that the limit distribution of (23) coincide with the limit distribution of

$$\mathbf{a}^T \sqrt{n} \begin{bmatrix} B_n(k_0, m) - B(k_0, m) \\ \bar{X} - \mu \\ \bar{T} - \lambda^{-1} \end{bmatrix}.$$

Finally, the use of Lemma 3 ends the proof.  $\square$

**Proof of Proposition 5** Taking into account that  $G_n(k_n, m) = G(k_0, m) = 0$ , a first order Taylor expansion of  $G_n(k_n, m)$  at the point  $k_0$ , gives the equality

$$G(k_0, m) = 0 = G_n(k_n, m) = G_n(k_0, m) + G'_n(k_n^*, m)(k_n - k_0), \quad (39)$$

where  $k_n^*$  lies in the interval having extremes  $k_0, k_n$ . From (39) we then get

$$\begin{aligned}k_n - k_0 &= -\frac{G_n(k_0, m)}{G_n^{(1)}(k_n^*, m)} \\ &= -\frac{G_n(k_0, m) - G(k_0, m)}{G_n^{(1)}(k_n^*, m)}\end{aligned}$$

and, as a consequence

$$\sqrt{n}(k_n - k_0) = -G_n^{(1)}(k_n^*, m)^{-1} \sqrt{n} (G_n(k_0, m) - G(k_0, m))$$

The conclusion of Theorem 3.2 is now an immediate consequence of Proposition 4 and Lemma 2.  $\square$

**Proof of Proposition 8** The assumption  $\rho = 1$  implies that the equation  $G(k, m) = 0$  possesses only the root  $k_0 = 0$ . Observing that  $k_n$  is equal to zero iff  $\hat{\rho} \geq 1$  (and  $\hat{\rho} < 1$  then  $k_n > 0$ ), we have first

$$\lim_{n \rightarrow \infty} P(\sqrt{n}k_n = 0) = \lim_{n \rightarrow \infty} P(\hat{\rho} \geq 1) = \lim_{n \rightarrow \infty} P(\sqrt{n}(\hat{\rho} - 1) \geq 0) = 1/2.$$

Next, a first order Taylor expansion of  $G_n(k_n, m)$  at the point  $k_0 = 0$  gives

$$0 = G_n(k_n, m) I_{(\hat{\rho} < 1)} = \{G_n(0, m) + G_n^{(1)}(k_n^*, m)(k_n - 0)\} I_{(\hat{\rho} < 1)} \quad (40)$$

where  $k_n^*$  lies in the interval having extremes  $(0, k_n)$ . Taking into account that

$$\begin{aligned}G_n(0, m) &= \begin{cases} B_n(0, m) - \frac{1}{\hat{\rho}} & \text{if } \hat{\rho} < 1 \\ 0 & \text{if } \hat{\rho} \geq 1 \end{cases} \\ &= \left(1 - \frac{1}{\hat{\rho}}\right) I_{(\hat{\rho} < 1)}\end{aligned}$$

from relationship (40) we get the equality

$$\left(\frac{1}{\hat{\rho}} - 1\right) I_{(\hat{\rho} < 1)} = Y_n k_n \quad (41)$$

where

$$\begin{aligned} Y_n &= G_n^{(1)}(0, m) I_{(\hat{\rho} \geq 1)} + G_n^{(1)}(k_n^*, m) I_{(\hat{\rho} < 1)} \\ &= G_n^{(1)}(0, m) (1 - I_{(\hat{\rho} < 1)}) + G_n^{(1)}(k_n^*, m) I_{(\hat{\rho} < 1)} \\ &= G_n^{(1)}(0, m) + I_{(\hat{\rho} < 1)} (G_n^{(1)}(k_n^*, m) - G_n^{(1)}(0, m)). \end{aligned} \quad (42)$$

Now, it is not difficult to prove that  $Y_n$  converges in probability to  $G^{(1)}(0, m)$  as  $n$  increases. It is enough to take into account that  $I_{(\hat{\rho} < 1)}$  converges in probability to a r.v. which takes the values 0, 1 both with probability 1/2, and that  $(G_n^{(1)}(k_n^*, m) - G_n^{(1)}(0, m))$  converges in probability to zero. Hence, we may conclude that  $I_{(\hat{\rho} < 1)} (G_n^{(1)}(k_n^*, m) - G_n^{(1)}(0, m))$  converges to zero in probability, too.

If  $x < 0$ , then  $(1/\hat{\rho} - 1) I_{(\hat{\rho} < 1)}$  is greater than  $x$  with probability 1. Since  $G_n^{(1)}(k_n, m)$  is positive, we then obtain

$$\lim_{n \rightarrow \infty} P(\sqrt{n} k_n \geq x) = 1 \quad \forall x < 0. \quad (43)$$

On the other hand, if  $x > 0$ , by observing that, as  $n$  goes to infinity,  $k_n^*$  converges with probability 1 to  $k_0 = 0$  and that

$$Y_n \xrightarrow{p} G^{(1)}(0, m) > 0, \quad (44)$$

from relationships (41) and (42) we then obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} P(\sqrt{n} k_n > x) &= \lim_{n \rightarrow \infty} P\left(\sqrt{n} \left(\frac{1}{\hat{\rho}} - 1\right) Y_n^{-1} > x\right) \\ &= \lim_{n \rightarrow \infty} P\left(\sqrt{n} \left(\frac{1}{\hat{\rho}} - 1\right) > x G^{(1)}(0, m)\right) \\ &= \lim_{n \rightarrow \infty} P\left(\sqrt{n} \left(\frac{1}{\hat{\rho}} - 1\right) / \sqrt{\text{Var}(1/\hat{\rho})} > x/\sigma_k\right) \\ &= 1 - \Phi(x/\sigma_k) \quad \forall x > 0. \square \end{aligned}$$



## References

- Broeckx, F., Goovaerts, M. J., and De Vylder, F. (1986). Ordering of risks and ruin probabilities. *Insurance: Mathematics and Economics*, **5**, 35–40.
- Cai, J. and Wu, Y. (1997). Some improvements on the lundberg's bound for the ruin probability. *Stat. Probab. Lett.*, **33**, 395–403.
- Conti, P. L. (2005). A nonparametric sequential test with power 1 for the ruin probability in some risk models. *Stat. Probab. Lett.*, **72**, 333–343.
- De Vylder, F. and Goovaerts, M. J. (1984). Bounds for classical ruin probabilities. *Insurance: Mathematics and Economics*, **3**, 121–131.
- Gaver, D. P. and Jacobs, P. A. (1988). Nonparametric estimation of the probability of a long delay in the m/g/1 queue. *Journal of the Royal Statistical Society B*, **50**, 392–402.
- Gerber, H. (1973). Martingales in risk theory. *Vereinigung Schweizerischer Versicherungs Mathematiker Mitteilungen*, **73**, 205–216.
- Lin, X. (1997). Tail of compound distributions and excess time. *Journal of Applied Probability*, **33**, 184–195.
- Lundberg, F. (1903). *Approximerad Framställning av Sannolikhetsfunktioner, II*. Almqvist & Wiksell, Uppsala.
- Pitts, S., Grübel, R., and Embrechts, P. (1996). Confidence bound for the adjustment coefficient. *Advances in Applied Probability*, **28**, 802–827.
- Ross, S. (1974). Bounds on the delay distribution in gi/g/1 queues. *Journal of Applied Probability*, **11**, 417–421.
- Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.
- Stoyan, D. (1983). *Comparison methods for queues and other stochastic models*. Wiley, Chichester.
- Willmot, G. E. (1994). Refinements and distributional generalizations of lundberg's inequality. *Insurance: Mathematics and Economics*, **15**, 49–63.
- Willmot, G. E. (1996). A non-exponential generalization of an inequality arising in queueing and insurance risk. *Journal of Applied Probability*, **33**, 176–183.
- Willmot, G. E. (1997). On the relationship between bounds on the tails of compound distributions. *Insurance: Mathematics and Economics*, **19**, 95–103.
- Willmot, G. E. and Lin, X. S. (1994). Lundberg bounds on the tails of compound distributions. *Journal of Applied Probability*, **31**, 743–756.

Willmot, G. E. and Lin, X. S. (1997). Simplified bounds on the tails of compound distributions. *Journal of Applied Probability*, **34**, 127–133.

Willmot, G. E. and Lin, X. S. (2000). *Lundberg approximations for compound distributions with insurance applications*. Springer-Verlag, Berlin.