Recent Developments in Mathematical Finance

in honour of Tomas Björk's 64th birthday, May 9–10, 2011

Ruin probabilities in a random environment

Jan Grandell

1 Introduction

We start with formulating the usual risk model. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space carrying the following independent objects:

(i) a point process $N = \{N(t);\, t \geq 0\}$;

(ii) a sequence $\{Z_j\}_1^\infty$ of independent and identically distributed random variables, having the common distribution function F, with $F(0) = 0$, mean value μ, and variance σ^2.

The *risk process,* X, is defined by

$$X(t) = ct - \sum_{k=1}^{N(t)} Z_k, \quad \left(\sum_{k=1}^{0} Z_k \stackrel{\text{def}}{=} 0 \right),$$

where c is a positive real constant. Let S_k denote the epoch of the kth claim.

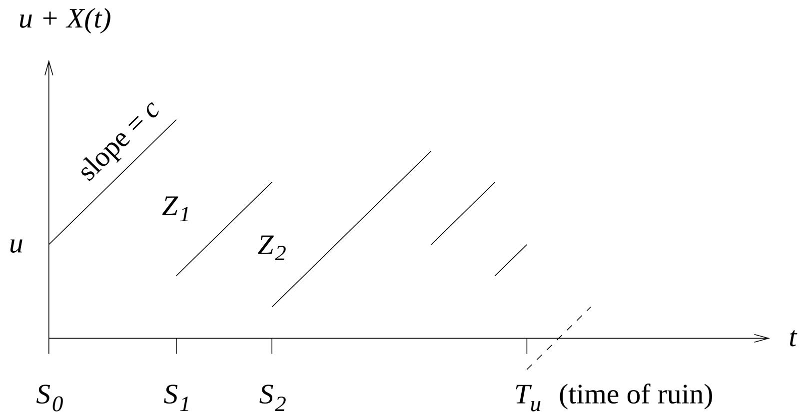

Figure 1: Illustration of notation.

If N is stationary with intensity α, i.e. $E[N(t)] = \alpha t$, we put

$$\rho = \frac{c - \alpha\mu}{\alpha\mu} \quad \text{relative safety loading} \quad \rho > 0.$$

Ruin probability:

$$\Psi(u,t) = P\{u + X(s) < 0 \text{ for some } s \in (0,t]\} = P\{T_u \leq t\}$$

$$\Psi(u) = P\{u + X(t) < 0 \text{ for some } t > 0\} = P\{T_u < \infty\},$$

where T_u the time of ruin.

Let the *integrated tail distribution* F_I be defined by

$$F_I(z) \stackrel{\text{def}}{=} \frac{1}{\mu}\int_0^z \bar{F}(x)\,dx,$$

where $\bar{F}(x) = 1 - F(x)$.

Risk theory goes back to Filip Lundberg (1903).

2 The Poisson case

Let $N = \{N(t);\, t \geq 0\}$ be a Poisson process with intensity α.

We have

$$\Psi(u) = \frac{\rho}{1+\rho}\sum_{n=0}^{\infty}\left(\frac{1}{1+\rho}\right)^n \bar{F}_I^{n*}(u), \tag{1}$$

which, in fact, is the *Pollaczek–Khinchine formula.*

The Pollaczek–Khinchine formula has a natural probabilistic interpretation, in terms of "ascending ladder points" of a random walk. In actuarial literature (1) is often referred to as the *Beekman's convolution formula.*

2.1 Small claims

Put

$$h(r) \stackrel{\text{def}}{=} \int_0^{\infty} (e^{rz} - 1)\, dF_Z(z) \qquad \text{and} \qquad \theta(r) \stackrel{\text{def}}{=} \alpha h(r) - cr.$$

Definition 1 *We talk about* small claims*, or say that F is* light-tailed*, if there exists $r_\infty > 0$ such that $h(r) \uparrow +\infty$ when $r \uparrow r_\infty$ (we allow for the possibility $r_\infty = +\infty$).*

The important part of Definition 1 is that $h(r) < \infty$ for some $r > 0$. This means that the tail of F decreases at least exponentially fast, and thus for instance the lognormal and the Pareto distributions are excluded.

Classical results for the Poisson case which go back to Lundberg (1926) and Cramér (1930):

$$\Psi(0) = \frac{1}{1+\rho}\,;$$

$$\Psi(u) = \frac{1}{1+\rho}e^{-\frac{\rho u}{\mu(1+\rho)}}$$

when the claim costs are exponentially distributed with mean μ;

the *Cramér-Lundberg* approximation

$$\lim_{u\to\infty} e^{Ru}\Psi(u) = C, \tag{2}$$

where the *Lundberg exponent* R is the positive solution of

$$\alpha h(r) = cr \tag{3}$$

and $C = \frac{\rho\mu}{h'(R) - \frac{c}{\alpha}}$;

the *Lundberg inequality*

$$\Psi(u) \leq e^{-Ru}. \tag{4}$$

Techniques:

1. *Differential argument.*

Consider what can happen in $[0, \Delta]$ and use stationary and independent increments.

2.*Renewal argument.*
Consider what happens in $[0, S_1]$ and use renewal structure. Note: Ruin can occur only at epochs of claims.
3. *Martingale method.* (Gerber 1973)
The time of ruin T_u is a $\mathbf{F}^X$-stopping time and

$$M_u(t) = e^{-\theta(r)t}e^{-r(u+X(t))} = \frac{e^{-r(u+X(t))}}{e^{t(\alpha h(r)-cr)}},$$

M_u is an $\mathbf{F}^X$-martingale.

We will first give the main ideas:

By optional stopping (after some technical tricks) we get, since M_u is positive,

$$\begin{aligned}
M_u(0) &= E[M_u(T_u)] \\
&= E[M_u(T_u) \mid T_u < \infty]P\{T_u < \infty\} \\
&\quad + E[M_u(T_u) \mid T_u = \infty]P\{T_u = \infty\} \\
&\geq E[M_u(T_u) \mid T_u < \infty]P\{T_u < \infty\} \\
&= E[M_u(T_u) \mid T_u < \infty]\Psi(u).
\end{aligned}$$

Since $u + X(T_u) \leq 0$ on $\{T_u < \infty\}$ we get

$$\begin{aligned}
\Psi(u) &\leq \frac{M_u(0)}{E[M_u(T_u) \mid T_u < \infty]} \\
&= \frac{e^{-ru}}{E[e^{-\theta(r)T_u}e^{-r(u+X(T_u))}] \mid T_u < \infty]} \\
&\leq \frac{e^{-ru}}{E[e^{-\theta(r)T_u} \mid T_u < \infty]} \\
&\leq \frac{e^{-ru}}{\inf_{0\leq t<\infty} e^{-\theta(r)t}}\,.
\end{aligned}$$

Since we want r as large as possible, it seems natural to define the *Lundberg exponent* R by

$$R = \sup\{r \geq 0 \mid \theta(r) \leq 0\} = \sup\{r \geq 0 \mid \alpha h(r) - cr \leq 0\},$$

which is the same as in (3), and we get the Lundberg inequality.

We will now give the real proof:

Choose $\underline{y}$ and $\overline{y}$ such that $0 \leq \underline{y} \leq \overline{y} < \infty$ and consider $\overline{y}u \wedge T_u \overset{\text{def}}{=} \min(\overline{y}u, T_u)$, which is a bounded $\mathbf{F}^X$-stopping time. Since $\mathcal{F}_0^X$ is trivial and since M_u is positive, it follows, by optional stopping, that

$$\begin{aligned}
M_u(0) &= E[M_u(\overline{y}u \wedge T_u)] \\
&= E[M_u(\overline{y}u \wedge T_u) \mid T_u \leq \underline{y}u]P\{T_u \leq \underline{y}u\} \\
&\quad +\; E[M_u(\overline{y}u \wedge T_u) \mid \underline{y}u < T_u \leq \overline{y}u]P\{\underline{y}u < T_u \leq \overline{y}u\} \\
&\quad +\; E[M_u(\overline{y}u \wedge T_u) \mid T_u > \overline{y}u]P\{T_u > \overline{y}u\} \\
&\geq E[M_u(\overline{y}u \wedge T_u) \mid \underline{y}u < T_u \leq \overline{y}u]P\{\underline{y}u < T_u \leq \overline{y}u\} \\
&= E[M_u(T_u) \mid \underline{y}u < T_u \leq \overline{y}u]P\{\underline{y}u < T_u \leq \overline{y}u\}. \qquad (5)
\end{aligned}$$

Since $u + X(T_u) \le 0$ on $\{T_u < \infty\}$ we get

$$\begin{aligned}
&\Psi(u, \overline{y}u) - \Psi(u, \underline{y}u) \\
&= P\{\underline{y}u < T_u \le \overline{y}u\} \le \frac{M_u(0)}{E[M_u(T_u) \mid \underline{y}u < T_u \le \overline{y}u]} \\
&\le \frac{e^{-ru}}{E[e^{-\theta(r)T_u} \mid \underline{y}u < T_u \le \overline{y}u]} \le \frac{e^{-ru}}{\inf_{\underline{y}u \le t \le \overline{y}u} e^{-\theta(r)t}} \\
&\le e^{-u \min(r - \underline{y}\theta(r), r - \overline{y}\theta(r))}. \qquad (6)
\end{aligned}$$

In order to get (6) as small as possible, we define the *"time-dependent" Lundberg exponent* $R_{\underline{y}}^{\overline{y}}$ by

$$R_{\underline{y}}^{\overline{y}} = \sup_{r \ge 0} \min(r - \underline{y}\theta(r), r - \overline{y}\theta(r)),$$

and we have the *"time-dependent" Lundberg inequality*

$$\Psi(u, \overline{y}u) - \Psi(u, \underline{y}u) \le e^{-R_{\underline{y}}^{\overline{y}} u}. \qquad (7)$$

Note that the "ordinary" Lundberg exponent, see (3), is given by

$$R = \sup\{r \ge 0 \mid \theta(r) \le 0\} = \sup\{r \ge 0 \mid \alpha h(r) - cr \le 0\}.$$

Put

$$f_y(r) = r - y\theta(r),$$

and note that $f_y(R) = R$ and that $f_y(r)$ is concave. Thus $R_{\underline{y}}^{\overline{y}} \ge R$ for all $\underline{y}$ and $\overline{y}$. Since f_y is differentiable and strictly concave for $y \in (0, \infty)$, it follows that

$$\sup_{r \ge 0} f_y(r) = f_y(r_y) \quad \text{where} \quad f'_y(r_y) = 1 - y\theta'(r_y) = 0.$$

Thus $\sup_{r \ge 0} f_y(r) > R$ unless $r_y = R \Leftrightarrow y = 1/\theta'(R)$. The value $y_0 \stackrel{\text{def}}{=} 1/\theta'(R)$ is called the *critical value*. For $y = 0$ we have $f_0(r) = r$ and for $y = \infty$ we put

$$f_\infty(r) \stackrel{\text{def}}{=} \lim_{y \to \infty} f_y(r) = \begin{cases} \infty & \text{for } r < R, \\ R & \text{for } r = R, \\ -\infty & \text{for } r > R. \end{cases}$$

Since

$$r_y \lessgtr R \quad \Leftrightarrow \quad f'_y(R) \lessgtr 0 \quad \Leftrightarrow \quad y \gtrless y_0,$$

it follows that $R_{\underline{y}}^{\overline{y}} = R$ when $\underline{y} \le y_0 \le \overline{y}$. Further

$$f_{\underline{y}}(r) \lessgtr f_{\overline{y}}(r) \quad \text{as} \quad r \lessgtr R.$$

Putting this together, we get

$$R_{\underline{y}}^{\overline{y}} = \begin{cases} R & \text{if } \underline{y} \le y_0 \le \overline{y}, \\ f_{\overline{y}}(r_{\overline{y}}) & \text{if } \overline{y} \le y_0, \\ f_{\underline{y}}(r_{\underline{y}}) & \text{if } y_0 \le \underline{y}. \end{cases} \qquad (8)$$

Thus, in order to get (7) as informative as possible, we shall choose $\overline{y}$ as large as possible when $R_{\underline{y}}^{\overline{y}}$ is determined by $\underline{y}$ and $\underline{y}$ as small as possible when $R_{\underline{y}}^{\overline{y}}$ is determined by $\overline{y}$. This leads to the following three Lundberg inequalities:

For $\underline{y} = 0$ and $\overline{y} = \infty$:

$$\Psi(u) \le e^{-Ru};$$

for $\underline{y} = 0$ and $y \le y_0$:

$$\Psi(u, yu) \le e^{-R_0^y u}; \tag{9}$$

for $\overline{y} = \infty$ and $y \ge y_0$:

$$\Psi(u) - \Psi(u, yu) \le e^{-R_y^\infty u}. \tag{10}$$

For any $\epsilon > 0$ we get, from (9) with $y = y_0 - \epsilon$ and from (10) with $y = y_0 + \epsilon$,

$$P\left\{\left|\frac{T_u}{u} - y_0\right| > \epsilon\right\} \le e^{-R_0^{(y_0-\epsilon)}u} + e^{-R_{(y_0+\epsilon)}^\infty u}.$$

Thus it follows from the Cramér-Lundberg approximation, i.e.,

$$\lim_{u\to\infty} e^{Ru} P\{T_u < \infty\} = C,$$

that

$$\lim_{u\to\infty} P\left\{\left|\frac{T_u}{u} - y_0\right| > \epsilon \mid T_u < \infty\right\} \le \lim_{u\to\infty} \frac{e^{-R_0^{(y_0-\epsilon)}u} + e^{-R_{(y_0+\epsilon)}^\infty u}}{P\{T_u < \infty\}} = 0$$

or, where $\xrightarrow{P}$ means "convergence in probability", that

$$\frac{T_u}{u} \xrightarrow{P} y_0 \quad \text{on } \{T_u < \infty\} \text{ as } u \to \infty. \tag{11}$$

2.2 Large claims

Definition 2 *We talk about* large claims *if F_I belongs to the class $\mathcal{S}$ of* subexponential distributions, *i.e. if*

$$\lim_{z\to\infty} \frac{\overline{F}_I^{2*}(z)}{\overline{F}_I(z)} = 2.$$

It is shown by Embrechts and Veraverbeke (1982, p. 70), cf. Pakes (1975, p. 557) for a queueing setting, that

$$\Psi(u) \sim \frac{1}{\rho}\overline{F}_I(u), \quad u \to \infty. \tag{12}$$

does hold *exactly* when $F_I \in \mathcal{S}$. The approximation (12) has a much slower speed of convergence than (2), see for instance Grandell (1997, p. 222).

Notice that (2) and (12) apply for fixed values of ρ as $u \to \infty$. Thus those approximations may be looked upon as "large deviation" results and it is seen that the asymptotic behaviour of $\Psi(u)$ is very different. However, not only that behaviour is different, but also "the way" ruin occurs is different. The formal way to discuss this is to consider the risk process before ruin occurs, conditioned upon that ruin occurs. These questions are studied by Asmussen (1982)in the case of *small claims*, i.e. when (2) applies, and by Asmussen and Klüppelberg (1996) in the case of *large claim*, i.e. when (12) applies. A nice discussion about these questions can be found in Embrechts, Klüppelberg and Mikosch (1997). Here we will only mention that in the small claim case the risk process before ruin behaves as a risk process with negative drift. A little more precisely the Poisson intensity is $\alpha + cR$ and the claim distribution is

$$\int_0^z e^{Rx} dF(x)/(1 + cR/\alpha)\, dx$$

before ruin, which happens close to the (deterministic) time $u/(\alpha h'(R) - c)$. One may express that as, for large values of u, ruin occurs due to a combination of many and (rather) large claims. In the large claim case it is a little more complicated to describe the situation precisely, but very roughly the risk process behaves normally until a really huge claim occurs which causes ruin. Naturally one may think of a "cooperation" of many large claims also in this case, but for (very) large values of u this possible cooperation is not enough to cause ruin.

3 Generalizations

The classical risk model can be generalized in many ways.

A. The premiums may depend on the result of the risk business.

B. Inflation, interest, and risky investments may be included in the model.

C. The occurrence of the claims may be described by a more general point process than the Poisson process.

Dassios and Embrechts (1989) and Delbaen and Haezendonck (1987) are very readable studies focusing mainly on generalizations A and B, while generalization C is considered by Grandell (1991). In all these studies most results are derived with the help of martingales.

4 Cox models

Intuitively we shall think of a Cox process N as generated in the following way: First a realization $\alpha(t)$ of a non-negative random process $\lambda = \{\lambda(t);\ t \geq 0\}$ is generated and conditioned upon that realization N is a non-homogeneous Poisson process with intensity function $\alpha(t)$. The process λ is called the *intensity process*.

Formally the distribution Π_Λ of a Cox process is given by

$$\Pi_\Lambda\{B\} = \int_{\mathcal{M}} \Pi_\mu\{B\}\ \Pi\{d\mu\} \qquad \text{for } B \in \mathcal{B}(\mathcal{N}),$$

where $\mathcal{M}$ is the set of Borel measures, $\mathcal{N} \subset \mathcal{M}$ is the set of point processes, Λ a random measure with distribution Π, and Π_μ the distribution of a Poisson process with intensity measure μ.

Now we consider a Cox model where N is a Cox process with intensity process $\lambda(t)$. The intensity measure Λ is given by

$$\Lambda(t) = \int_0^t \lambda(s)\ ds.$$

A detailed discussion of Cox processes and their impact on risk theory is found in Grandell (1991).

A natural filtration is $\mathbf{F} = (\mathcal{F}_\infty^\Lambda \vee \mathcal{F}_t^X;\ t \geq 0)$. Note that $\mathcal{F}_0 = \mathcal{F}_\infty^\Lambda$. Then

(i) $N(t)$ has independent increments relative to $\mathcal{F}_\infty^\Lambda$;

(ii) $N(t) - N(s)$ is Poisson distributed with mean $\Lambda(t) - \Lambda(s)$ relative to $\mathcal{F}_\infty^\Lambda$.

We will only consider small claims. It seems very natural to try to find an $\mathbf{F}$-martingale "as close as possible" to the one used in the Poisson case. Therefore we consider

$$M(t) = \frac{e^{-r(u+X(t))}}{e^{\Lambda(t)h(r)-trc}},$$

where we quite simply have replaced αt with $\Lambda(t)$. It is easy to show that M is an $\mathbf{F}$-martingale.

Put $C(r) = E\left[\sup_{t\geq 0} e^{\Lambda(t)h(r)-rct}\right]$ and $R = \sup\{r \mid C(r) < \infty\}$.

Using this martingale Björk and Grandell (1988) gave an extension of Gerber's martingale approach and proved the following Lundberg inequality.

Theorem 1 *For every $\epsilon > 0$ such that $0 < \epsilon < R$ we have*

$$\Psi(u) \leq C(R-\epsilon)e^{-(R-\epsilon)u},$$

where $C(R-\epsilon) < \infty$.

This theorem is rather useless, since in general it is probably very difficult to actually determine R. The condition $\epsilon > 0$ is unpleasant, but it is quite possible, and really natural, that $C(R) = +\infty$. In fact, $C(R) < \infty$ requires a discontinuity in $C(r)$ at $r = R$. In the Poisson case we have such a discontinuity since $C(r) = 1$ for $r \le R$ and $C(r) = \infty$ for $r > R$.

Independent jump intensity

We now consider a class of intensity processes with "independent jumps."

Intuitively an independent jump intensity is a jump process where the jump times form a renewal process and where the value of the intensity between two successive jumps may depend only on the distance between these two jumps. More formally, let Σ_k, $k = 1, 2, \ldots$ denote the epoch of the kth jump of the intensity process and let $\Sigma_0 \stackrel{\text{def}}{=} 0$. Put

$$\begin{aligned} \sigma_n &= \Sigma_n - \Sigma_{n-1} \\ L_n &= \lambda(\Sigma_{n-1}) \end{aligned} \qquad n = 1, 2, 3, \ldots.$$

Here we understand that λ has right-continuous realizations so that $\lambda(t) = L_n$ for $\Sigma_{n-1} \le t < \Sigma_n$.

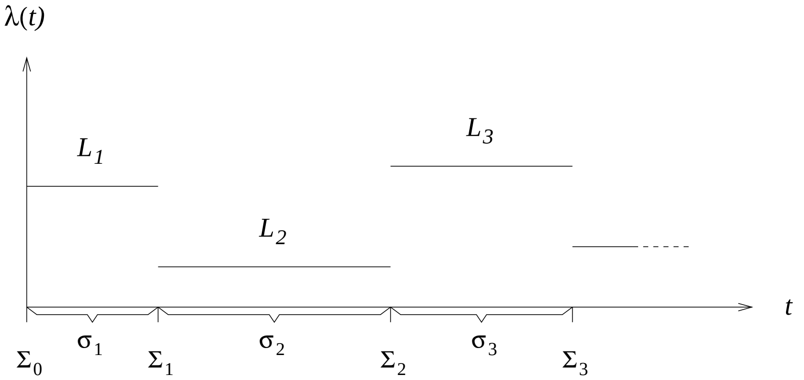

Figure 2: Illustration of notation.

Definition 3 *An intensity process* λ *is called*

(i) an independent jump intensity *if the random vectors*

$$(L_1, \sigma_1), \ (L_2, \sigma_2), \ (L_3, \sigma_3), \ldots$$

are independent and if (L_2, σ_2), $(L_3, \sigma_3), \ldots$ *have the same distribution* p*;*

(ii) an ordinary independent jump intensity *if* (L_1, σ_1) *also has distribution* p*;*

(iii) a stationary independent jump intensity *if the distribution of* (L_1, σ_1) *is chosen such that* λ *is stationary.*

Let (L, σ) be the generic vector for (L_n, σ_n), $n \ge 2$, i.e.,

$$\Pr\{L \in A, \ \sigma \in B\} = p(A \times B) \qquad \text{for } A, \ B \in \mathcal{B}(\mathbb{R}_+).$$

The marginal distribution of L is denoted by p_L, i.e.,

$$p_L(A) = p(A \times \mathbb{R}_+) \qquad \text{for } A \in \mathcal{B}(\mathbb{R}_+).$$

Let q be the distribution of (L_1, σ_1). The intensity λ is stationary if

$$q(A \times B) = \frac{1}{E[\sigma]} \int_B p(A \times (t, \infty)) \, dt. \tag{13}$$

Put

$$\phi(r) \stackrel{\text{def}}{=} E[e^{-rc\sigma + h(r)L\sigma}].$$

A main result in Björk and Grandell (1988) is the following theorem.

Theorem 2 *Lundberg's inequality (in the Cox version) holds with*

$$R = \sup\{r \geq 0 \mid \phi(r) \leq 1\}.$$

When nothing else is said, we assume positive safety loading which here means that $cE[\sigma] > \mu E[L\sigma]$. Consequently we assume that $E[\sigma] < \infty$ and $E[L\sigma] < \infty$. Positive safety loading alone does not imply $R > 0$, but it does if $\phi(r) < \infty$ for some $r > 0$.

The Lundberg exponent R is the "right" exponent in the following sense:

Theorem 3 *Assume that $\phi(r) < \infty$ for some $r > R > 0$. Then*

$$\lim_{u \to \infty} e^{(R+\epsilon)u} \Psi(u) = \infty$$

for every $\epsilon > 0$.

We shall now consider a martingale approach, due to Embrechts, Grandell and Schmidli (1993), which will allow to consider finite-time Lundberg inequalities and to get "$\epsilon = 0$". That approach can be looked upon as a generalization of an approach used by Björk and Grandell (1988) and Grandell (1991) when λ is a Markov process with independent jumps or a finite-state Markov process.

In our applications all Markov processes will be so-called piecewise-deterministic Markov (PD) processes. This class of was introduced by Davis (1984). Dassios and Embrechts (1989) has shown that many important risk processes are naturally handled within the framework of PD processes.

Intuitively a PD process follows a deterministic path, determined by a first-order differential operator χ, until it jumps, according to an *intensity function* $\kappa(y)$ or when it hits the boundary $\partial\mathbb{S}$ of $\mathbb{S}$, and a *jump measure* $K(y, B)$, $y \in \mathbb{S}$, $B \in \mathcal{B}(\mathbb{S})$. The operator χ is the generator "between jumps", $\kappa(y)dt$ is the probability of a jump in the interval $(t, t+dt]$ when $Y(t) = y$, and $K(y, B)$ is the probability that a jump leads to a point in B.

The generator $\mathcal{A}$, see Dassios and Embrechts (1989, pp. 185), is given by

$$\mathcal{A}f(y) = \chi f(y) + \kappa(y) \int_{\mathbb{S}} (f(z) - f(y)) K(y, dz), \tag{14}$$

for all functions f in the domain of $\mathcal{A}$, where $\chi = \sum_{k=1}^{n} c_k(y) \frac{\partial}{\partial y_k}$. If jumps are also caused by hits of the boundary, f must fulfil the condition

$$f(y) = \int_{\mathbb{S}} (f(z) - f(y)) K(y, dz), \quad \text{for all } y \in \partial\mathbb{S},$$

in order to belong to the domain of $\mathcal{A}$.

Thus, if f is in the domain of $\mathcal{A}$ and $\mathcal{A}f \equiv 0$, then, by Dynkin's theorem, $f(Y(t))$ is an $\mathbf{F}^Y$-martingale.

We will, however, need the following rather strong assumptions. Consider the convex sets

$$C_\infty = \{(\vartheta, r) \mid \phi(\vartheta, r) < \infty\} \quad \text{and} \quad C_1 = \{(\vartheta, r) \mid \phi(\vartheta, r) \leq 1\}.$$

Assumption 1 *Assume that $\partial C_1 \subset int C_\infty$, where ∂ means boundary and int stands for interior.*

Assumption 2 *Assume that*

$$\sup_{w \geq 0} E[\sigma - w \mid \sigma > w,\ L = \ell] \leq B < \infty, \quad p_L\text{-}a.s.$$

Example 1 An interesting special case is when σ, conditioned upon L, is exponentially distributed, since then λ is a Markov process with independent jumps. This means that

$$p(d\ell \times ds) = p_L(d\ell)\ \eta_\ell e^{-\eta_\ell s}\ ds.$$

In this case $(X,\,\lambda) = \{(X(t),\,\lambda(t));\ t \geq 0\}$ is a Markov process.

Consider the vector process

$$Y = \{(X(t),\,\lambda(t),\,W(t),\,t);\ t \geq 0\},$$

where $W(t)$ is the time remaining to the next jump of the intensity, and the filtration $\mathbf{F}^Y = (\mathcal{F}_t^X \vee \mathcal{F}_t^\lambda \vee \mathcal{F}_t^W;\ t \geq 0)$. Due to the regenerative structure of $\lambda(t)$, the vector process Y is a PD process.

In the ordinary case there is a jump of the intensity at the origin, i.e., $(\lambda_0,\,W_0)$ has distribution p. In the stationary case $(\lambda_0,\,W_0)$ has distribution q, given by (13). In order to reduce the number of parentheses in the formulas below, we – sometimes – use the notations X_t, λ_t, and W_t.

From (14) if follows that the generator for $(X_t,\,\lambda_t,\,W_t,\,t)$ is given by

$$\begin{aligned}\mathcal{A}f(x,\ell,w,t) =& c\,\frac{\partial f(x,\ell,w,t)}{\partial x} - \frac{\partial f(x,\ell,w,t)}{\partial w} + \frac{\partial f(x,\ell,w,t)}{\partial t} \\ &+ \ell \int_0^\infty (f(x-y,\ell,w,t) - f(x,\ell,w,t))dF(y)\end{aligned}$$

with boundary condition

$$f(x,\ell,0,t) = \int_{\mathbb{R}} f(x,\lambda,w,t)\,p(d\lambda \times dw)\,.$$

For fixed $r < \sup\{\tilde{r} \mid \exists \vartheta \in \mathbb{R},\ (\vartheta,\tilde{r}) \in C_\infty\}$ we look for a positive $\mathbf{F}^Y$-martingale $M_u(t)$ of the form

$$M_u(t) = e^{-\theta(r)t} g(\lambda_t, W_t) e^{-r(u+X_t)},$$

where g, depending on r, is differentiable in its second component.

The following Lemma follows from (14), Dynkin's theorem, and and the implicit function theorem.

Lemma 1 *In the above independent jump intensity model,*

$$M_u(t) = e^{-\theta(r)(t+W_t)}\ e^{-crW_t + \lambda_t h(r) W_t}\ e^{-r(u+X_t)}, \quad t \geq 0,$$

is an $\mathbf{F}^Y$-martingale, where $\theta(r)$, given by

$$E\Big[e^{-\theta(r)\sigma - cr\sigma + h(r)L\sigma}\Big] = 1,$$

is differentiable and convex.

Choose $\underline{y}$ and $\overline{y}$ such that $0 \leq \underline{y} \leq \overline{y} < \infty$. Then, exactly as in (5), we get

$$M_u(0) = E^{\mathcal{F}_0^Y}[M_u(T_u) \mid \underline{y}u < T_u \leq \overline{y}u] P^{\mathcal{F}_0^Y}\{\underline{y}u < T_u \leq \overline{y}u\}$$

and thus, compare (6),

$$\begin{aligned}P^{\mathcal{F}_0^Y}\{\underline{y}u < T_u \leq \overline{y}u\} &\leq \frac{M_u(0)}{E^{\mathcal{F}_0^Y}[M_u(T_u) \mid \underline{y}u < T_u \leq \overline{y}u]} \\ &\leq \frac{g(\lambda_0, W_0)e^{-ru}}{E^{\mathcal{F}_0^Y}[e^{-\theta(r)T_u} g(\lambda(T_u), W(T_u)) \mid \underline{y}u < T_u \leq \overline{y}u]} \\ &\leq \frac{g(\lambda_0, W_0)e^{-u\min(r-\underline{y}\theta(r), r-\overline{y}\theta(r))}}{E^{\mathcal{F}_0^Y}[g(\lambda(T_u), W(T_u)) \mid \underline{y}u < T_u \leq \overline{y}u]}\,.\end{aligned}$$

In contrast to (6), where this posed no problem, here we have to ensure that

$$E^{\mathcal{F}_0^Y}\left[g(\lambda(T_u), W(T_u)) \mid \underline{y}u < T_u \leq \overline{y}u\right]$$

$$= E^{\mathcal{F}_0^Y}\left[\exp\{-W(T_u)(\theta(r) + cr - \lambda(T_u)h(r))\} \mid \underline{y}u < T_u \leq \overline{y}u\right] > 0$$

for all values of u. This is rather technical, and we have to refer to Embrechts, Grandell and Schmidli (1993).

In the ordinary case we have

$$E[g(\lambda_0, W_0)] = 1,$$

whereas in the stationary case, see Grandell (1991, p. 96),

$$\begin{aligned} E[g(\lambda_0, W_0)] &= \frac{E\left[\int_0^\sigma g(L,s)\,ds\right]}{E[\sigma]} = \frac{E\left[\int_0^\sigma e^{-s[\theta(r)+cr-Lh(r)]}\,ds\right]}{E[\sigma]} \\ &\leq \frac{E\left[\sigma\left(1 + e^{-\sigma[\theta(r)+cr-Lh(r)]}\right)\right]}{E[\sigma]} \end{aligned}$$

and it follows from Assumption 1 that $E[g(\lambda_0, W_0)] < \infty$.

The Lundberg inequalities and (11) follows almost as in the Poisson case, since R is the "right" exponent.

A generalization of the Cramér-Lundberg approximation (2) is due to Schmidli (1997). In the case of finite-state Markovian intensity, the Cramér-Lundberg approximation was given by Asmussen (1989).

For large claims we have the following theorem, due to Asmussen, Schmidli and Schmidt (1999) in the ordinary case. Recall that positive safety loading here means that $cE[\sigma] > \mu E[L\sigma]$.

Theorem 4 *Assume that $F, F_I \in \mathcal{S}$ and that there exists a $\delta > 0$ such that $E[e^{\delta L\sigma}] < \infty$. Then*

$$\Psi(u) \sim \frac{\mu E[L\sigma]}{cE[\sigma] - \mu E[L\sigma]} \overline{F}_I(u), \quad u \to \infty.$$

Remark: The natural definition of the safety loading ρ is

$$\rho = \frac{cE[\sigma] - \mu E[L\sigma]}{\mu E[L\sigma]},$$

see Grandell (1991, p. 98). With this definition of ρ it is seen that (12) holds also in this case.

The condition $E[e^{\delta L\sigma}] < \infty$ in Theorem 4 guarantees that $\Psi(u)$ does not become heavy-tailed due to many claims. In fact, if the distribution of $L\sigma$ is heavy-tailed, also $\Psi(u)$ may be heavy-tailed. Such results are given by Asmussen, Schmidli and Schmidt (1999). They are, however, complicated, and will not be given here.

In the case of finite-state Markovian intensity, the correspondence to Theorem 4 is due to Asmussen, Fløe Henriksen and Klüppelberg (1994). Naturally no conditions on the intensity are needed in order to avoid "too many claims".

5 Thinning and the choice of models

A Cox process is a generalization of the Poisson process in the sense that stochastic variation in the intensity is allowed and is therefore very natural as a model for "risk fluctuation". Cox processes are characterized by the fact that they can be obtained by independent p-thinning for all p. An "opposite" class are *"top processes"* which can not be obtained by any p-thinning. It is natural to consider claims as caused by "risk situations" or *incidents.* Then *it is highly unnatural to choose N among top processes.* Furthermore this view, in our opinion, indicates that *it is natural to choose N among Cox processes* in particular when each incident causes a claim with small probability.

Let, as an example, N be a renewal process with Γ-distributed inter-occurrence times with form parameter β. It is shown by Yannaros (1988) that N is *a Cox process if* $0 < \beta \leq 1$ and *a top process if* $\beta > 1$. For $\beta < 1$, the Γ-distribution is a mixture of exponential distributions. Rather much is explicitly known about ruin probabilities in the renewal case when the inter-occurrence time distribution is a mixture of exponential distributions. *Those renewal processes are, in fact, Cox processes.*

REFERENCES

Asmussen, S. (1982). Conditioned limit theorems relating a random walk to its associate, with applications to risk reserve processes and the $GI/G/1$ queue. *J. Appl. Prob. 14*, 143–170.

Asmussen, S. (1989). Risk theory in a Markovian environment. *Scand. Actuarial J.*, 66 - 100.

Asmussen, S., Fløe Henriksen, L. and Klüppelberg, C. (1994). Large claims approximations for risk processes in a Markovian environment. *Stochastic Proc. Applic. 54*, 29–43.

Asmussen, S. and Klüppelberg, C. (1996) Large deviation results for subexponential tails, with applications to insurance risk. *Stoch. Proc. Appl. 64*, 103–125.

Asmussen, S., Schmidli, H. and Schmidt, V. (1999). Tail probabilities for non-standard risk and queueing processes with subexponential jumps. *Adv. Appl. Probab. 31*, 422–447.

Björk, T. and Grandell, J. (1988). Exponential inequalities for ruin probabilities in the Cox case. *Scand. Actuarial J.*, 77–111.

Davis, M. H. A. (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *J. R. Statist. Soc. B 46*, 353–388.

Dassios, A. and Embrechts, P. (1989). Martingales and insurance risk. *Commun. Statist. – Stochastic models 5*, 181–217.

Delbaen, F. and Haezendonck, J. (1987). Classical risk theory in an economic environment. *Insurance: Mathematics and Economics 6*, 85–116.

Cramér, H. (1930). *On the mathematical theory of risk.* Skandia Jubilee Volume, Stockholm.

Embrechts, P., Grandell, J. and Schmidli, H. (1993). Finite-time Lundberg inequalities in the Cox case. *Scand. Actuarial J.*, 17–41

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Springer-Verlag, Berlin.

Embrechts, P. and Veraverbeke, N. (1982). Estimates for the probability of ruin with special emphasis on the possibility of large claims. *Insurance: Mathematics and Economics 1*, 55–72.

Gerber, H. U. (1973). Martingales in risk theory. *Mitt. Ver. Schweiz. Vers. Math. 73*, 205–216.

Grandell, J. (1991). *Aspects of risk theory.* Springer-Verlag, New York.

Grandell, J. (1997). *Mixed Poisson processes.* Chapman & Hall, London.

Lundberg, F. (1903). *I. Approximerad framställning av sannolikhetsfunktionen. II. Återförsäkring av kollektivrisker.* Almqvist & Wiksell, Uppsala.

Lundberg, F. (1926). *Försäkringsteknisk riskutjämning.* F. Englunds boktryckeri AB., Stockholm.

Pakes, A. G. (1975) On the tails of waiting-time distributions. *J. Appl. Prob. 12*, 555–564.

Schmidli, H. (1997). An extension to the renewal theorem and an application to risk theory. *Ann. Appl. Prob. 7*, 121–133.

Yannaros, N. (1988). On Cox processes and gamma renewal processes. *J. Appl. Prob. 25*, 423–427.