Brazilian Journal of Probability and Statistics (2005), 19, pp. 1–12 .

Large deviation results for compound Markov renewal processes

Claudio Macci
Università di Roma "Tor Vergata"

Abstract: In this paper we present some large deviation results for compound Markov renewal processes. We start studying the exponential decay of level crossing probabilities as the level goes to infinity. Furthermore we consider a technique called importance sampling to estimate a level crossing probability by Monte Carlo simulations when the level is large; in particular we find an asymptotically efficient simulation law according to a sense specified in other works on the same topic.

Key words: Large deviations, level crossing probabilities, importance sampling.

1 Introduction

The aim of this paper is to present some large deviation results for level crossing probabilities concerning a quite natural extension of the classical compound Poisson model used in insurance (see e.g. the books of Asmussen (1987) and (2000); for some more details in this paper see Remark 2.1.

In order to do that let us start by considering two independent discrete time Markov additive processes $((J_n, S_n))$ and $((\tilde{J}_n, \tilde{S}_n))$ (for the concept of Markov additive process we refer to the presentation in the book of Asmussen (2000)), where the environments (J_n) and $(\tilde{J}_n)$ are irreducible Markov chains with finite state spaces.

Moreover we assume that the random variables $(S_n - S_{n-1})$ are almost surely positive and let (N_t) be the counting process with interarrival times $(S_n - S_{n-1})$; then

$$\tilde{S}_{N_t} := \sum_{n=1}^{N_t} \tilde{S}_n \quad (\text{where} \quad N_t := \sum_{n \geq 1} 1_{S_n \leq t}) \tag{1.1}$$

is a compound Markov renewal process (and (N_t) is a Markov renewal process).

We point out that compound Markov renewal process is a generalization of compound Poisson process (for details see again Remark 2.1 in this paper).

The first result in this paper gives the exponential decay of the level crossing probabilities $(p(b))_{b>0}$ in the fashion of large deviation theory (here b is the level).

The second result deals with a technique called importance sampling for the estimation by Monte Carlo simulation of $p(b)$ when b is large. Roughly speaking the idea of importance sampling is the following. Let us consider an unbiased estimator $[\widehat{p(b)}]_Q$ of $p(b)$ (called importance sampling estimator) which depends on the simulation law Q and, if we keep a fixed relative precision for $[\widehat{p(b)}]_Q$, the number of simulations can grow exponentially with b. Our aim is to find a simulation law Q with respect to which this fact can be avoided; in order to do that we have to check the limit (3.8) and, if (3.8) holds, Q is said to be an asymptotically efficient simulation law. Then the second result in this paper consists of proving that a suitable law is asymptotically efficient for simulation. Here we do not deal with the problem of the uniqueness of the asymptotically efficient simulation law.

The propositions in this paper have some analogies with the results in some other papers in the literature on the same topic. The first works in the literature are the papers of Lehtonen and Nyrhinen (1992a) and (1992b) which deal with the discrete time case. For the continuous time case a similar work is the paper of Baldi and Piccioni (1999), even if the presentation of some items is different. A more recent work of Collamore (2002) deals with a multidimensional level crossing problem. Finally we also cite the papers of Asmussen and Nielsen (1995), Sadowsky (1996) and Macci (2001).

The proofs of the results in this paper are similar to the proofs of the analogous results in other papers on the same topic. In particular some upper bounds are derived by choosing a suitable exponential change of measure in a family depending on one parameter. On the other hand, when one deals with compound Markov renewal processes, it is natural to consider a product between two exponential changes of measures, so that we have a family which depends on two parameters (see eq. (3.4) in this paper). The author thinks that the main interesting fact concerning the proofs presented in this paper consists to consider a suitable subfamily which depends on only one parameter (see (3.5) in this paper).

The results in this paper are presented in Section 3 after some preliminaries in Section 2.

2 Preliminaries

In this section we present some preliminaries. Subsections 2.1 and 2.2 are devoted to recall some preliminaries on Markov additive processes. In Subsection 2.3 we present conditions **(H1)** and **(H2)** which allow to define the compound Markov renewal processes. Moreover in Subsection 2.3 we also present **(H3)**; this condition plays a crucial role to prove the large deviation results in this paper.

2.1 Generalities on Markov additive processes

Set $E = \{1, \dots, m\}$, let $J = (J_n)$ be a E-valued Markov chain called environment and let $(p_{ij})_{i,j \in E}$ be the transition matrix of J. Then $((J_n, S_n))$ is a Markov additive process if (S_n) is a real valued sequence of random variables such that $S_0 = 0$ and

$$S_n = Y_1 + \cdots + Y_n \quad (\forall n \geq 1) \tag{2.1}$$

where, given J, (Y_n) is a sequence of i.i.d. random variables and, in general, the conditional distribution of Y_k given J depends on (J_{k-1}, J_k) only. In view of what follows we denote the conditional distribution of Y_k given $(J_{k-1}, J_k) = (i, j)$ by H_{ij} and the corresponding moment generating function by $\hat{H}_{ij}$.

A reference for the presentation above of Markov additive processes is the book of Asmussen (2000, chapter 2, section 5, discrete time case at page 40). We also cite the paper of Grigelionis (1994, section 2, Corollary 1) in which the sequence (S_n) takes values on a general Hilbert space. Section 5 in chapter 2 in the book of Asmussen (2000) deals with a general presentation of Markov additive processes in which the set E could be not finite; large deviation results for this general situation can be found in the paper of Iscoe et al. (1985).

Assume that J is irreducible. Then, for each fixed $\alpha \in \mathbb{R}$, we can consider the matrices $F(\alpha) = (F_{ij}(\alpha))_{i,j\in E}$ defined by

$$F_{ij}(\alpha) = \mathbb{E}[e^{\alpha S_1} 1_{J_1=j} | J_0 = i] \quad (\forall i, j \in E).$$

The entries of these matrices could be infinite when $\alpha \neq 0$ (when $\alpha = 0$ the entries are finite, indeed the matrix $F(0)$ coincides with the transition matrix of J) and the latter is equivalent to

$$F_{ij}(\alpha) = \hat{H}_{ij}(\alpha) p_{ij} \quad (\forall i, j \in E).$$

Thus, as far as the family of the moment generating functions $(\hat{H}_{ij})$ is concerned, we remark that some pairs (i, j) can be neglected; indeed we can restrict our attention on the pairs (i, j) such that $p_{ij} > 0$. Moreover, if we consider the n-th power $F^n(\alpha) = ((F^n)_{ij}(\alpha))_{i,j\in E}$ of $F(\alpha)$, we have

$$(F^n)_{ij}(\alpha) = \mathbb{E}[e^{\alpha S_n} 1_{J_n=j} | J_0 = i] \quad (\forall i, j \in E).$$

Now let us consider the set

$$\mathcal{D} := \{\alpha \in \mathbb{R} : F_{ij}(\alpha) < \infty \ (\forall i, j \in E)\}$$

and let $n \geq 1$ and $\alpha \in \mathcal{D}$ be arbitrarily fixed (we point out that $\mathcal{D}$ is not empty since $0 \in \mathcal{D}$). Then we have the following facts.

• Perron Frobenius Theorem provides the existence of a simple and positive eigenvalue $e^{\kappa(\alpha)}$ of the matrix $F(\alpha)$ equal to its spectral radius and

$$\sum_{j\in E} (F^n)_{ij}(\alpha) h_j(\alpha) = e^{n\kappa(\alpha)} h_i(\alpha) \quad (\forall i, j \in E)$$

where $(h_i(\alpha))_{i\in E}$ is an eigenvector with positive components, unique up to a positive constant.

• For any initial distribution for J (namely for any distribution of J_0) we have

$$\mathbb{E}[e^{\alpha S_n}] = \sum_{i\in E} \mathbb{E}[e^{\alpha S_n} | J_0 = i] P(J_0 = i) = \sum_{i\in E} (\sum_{j\in E} \mathbb{E}[e^{\alpha S_n} 1_{J_n=j} | J_0 = i]) P(J_0 = i),$$

whence we obtain

$$\frac{\sum_{i\in E} h_i(\alpha) P(J_0 = i)}{\max_{j\in E} h_j(\alpha)} e^{n\kappa(\alpha)} \leq \mathbb{E}[e^{\alpha S_n}] \leq \frac{\sum_{i\in E} h_i(\alpha) P(J_0 = i)}{\min_{j\in E} h_j(\alpha)} e^{n\kappa(\alpha)}.$$

2.2 A class of conjugate laws for Markov additive processes

In this subsection we present a class of conjugate laws $(P_\alpha)_{\alpha:\kappa(\alpha)<\infty}$ defined as follows. For each fixed $\alpha \in \mathcal{D}$ we have that P is absolutely continuous with respect to P_α on each finite time interval $\{0, 1, \ldots, n\}$ and the corresponding density ℓ_n^{P,P_α} is

$$\ell_n^{P,P_\alpha} = e^{-\alpha S_n + n\kappa(\alpha)} \frac{h_{J_0}(\alpha)}{h_{J_n}(\alpha)}.$$

Then (see e.g. Proposition 5.10 and Theorem 5.11 in the book of Asmussen, 2000, pages 44-45) $((J_n, S_n))$ is again a Markov additive process under each P_α with transition matrix $(p_{ij}^{(\alpha)})_{i,j\in E}$ for (J_n) defined by

$$p_{ij}^{(\alpha)} := \frac{F_{ij}(\alpha) h_j(\alpha)}{e^{\kappa(\alpha)} h_i(\alpha)}$$

and distributions $(H_{ij}^{(\alpha)})$ (which play the role of (H_{ij}) under P) defined by

$$H_{ij}^{(\alpha)}(dx) := \frac{e^{\alpha x}}{\hat{H}_{ij}(\alpha)} H_{ij}(dx).$$

Furthermore, by Proposition 4.1 and by the first part of Lemma 4.2 in the paper of Iscoe et al. (1985) applied with respect to each P_α, we have

$$\lim_{n\to\infty} \frac{S_n}{n} = \kappa'(\alpha) \quad P_\alpha \quad \text{a.s.}; \tag{2.2}$$

the first part of Lemma 4.2 in the paper of Iscoe et al. (1985) follows from a result in the book of Revuz (1975, Theorem 3.6, page 123). Finally we point out that $P_0 = P$.

2.3 Compound Markov renewal processes and hypotheses

The following hypotheses play a crucial role in the construction of compound Markov renewal process:
(H1): *$((J_n, S_n))$ and $((\tilde{J}_n, \tilde{S}_n))$ are independent Markov additive processes.*
(H2): *the distributions (H_{ij}) are concentrated on $]0, \infty[$.*

All the items concerning $((\tilde{J}_n, \tilde{S}_n))$ will be denoted with obvious notations: for instance $\tilde{H}_{ij}$ is the distribution which plays the role of H_{ij} when we have $((\tilde{J}_n, \tilde{S}_n))$ in place of $((J_n, S_n))$.

We point out that the random variables (Y_n) in (2.1) can be seen as the interarrival times of a Markov renewal process (N_t); thus, following the lines of the Introduction and by also taking into account **(H1)**, the process $(\tilde{S}_{N_t})$ defined by (1.1) is called compound Markov renewal process.

Remark 2.1 First of all we remark that $(\tilde{S}_{N_t})$ is a compound Poisson process when the distributions (H_{ij}) are all the same exponential distribution and the distributions $(\tilde{H}_{ij})$ are all coincident. Furthermore, in order to explain our interest for $(\tilde{S}_{N_t} - ct)$, we can say that this process is a generalization of the claim surplus process (and $(u + ct - \tilde{S}_{N_t})$ is the risk reserve process, where u is the initial reserve; see Asmussen, 2000, page 1) concerning the compound Poisson model (see Asmussen, 1987, pages 111 and 280); indeed we have the compound Poisson model when $(\tilde{S}_{N_t})$ is a compound Poisson process and the coincident distributions $(\tilde{H}_{ij})$ are concentrated on $]0, \infty[$.

We remark that (S_n) is nondecreasing (almost surely) by **(H2)**; then the restriction of κ on $\mathcal{D}$ is an increasing function and we can consider the inverse function κ^{-1} of κ.

In view of the large deviation results in this paper, we need to present the next condition **(H3)** inspired by Theorem 1 of Glynn and Whitt (1994).

(H3): (J_n) *and* $(\tilde{J}_n)$ *are irreducible*; *let us consider the function* $\Psi_\kappa(\cdot) = -\kappa^{-1}(-(\cdot))$ *and we have*

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}[e^{\alpha N_t}] = \Psi_\kappa(\alpha) \quad (\forall \alpha \in \mathbb{R});$$

$\Psi_\kappa(\tilde{\kappa}(\cdot))$ *is finite in a neighbourhood of the origin* $\alpha = 0$, *essentially smooth and lower semicontinuous.*

When we say that $\Psi_\kappa(\tilde{\kappa}(\cdot))$ is essentially smooth (see e.g. Dembo and Zeitouni, 1993, Definition 2.3.5, page 44) we mean that the set $\{\tilde{\alpha} \in \mathbb{R} : \Psi_\kappa(\tilde{\kappa}(\tilde{\alpha})) < \infty\}^\circ$ is not empty (where the symbol $^\circ$ means the interior), $\Psi_\kappa(\tilde{\kappa}(\cdot))$ is differentiable on this nonempty set and $\Psi_\kappa(\tilde{\kappa}(\cdot))$ is a steep function (namely $\lim_{n\to\infty} |\Psi'_\kappa(\tilde{\kappa}(\alpha_n))\, \tilde{\kappa}'(\alpha_n)| = \infty$ for any sequence (α_n) converging to some boundary point of $\{\tilde{\alpha} \in \mathbb{R} : \Psi_\kappa(\tilde{\kappa}(\tilde{\alpha})) < \infty\}$).

We point out that we have an essentially smooth function when the function is finite (namely assumes finite values) and differentiable. An illustrative class of examples such that **(H3)** holds is the following simplified situation: the distributions (H_{ij}) are all the same exponential distribution (with parameter $\lambda > 0$) so that

$$\kappa(\alpha) = \begin{cases} \log \frac{\lambda}{\lambda - \alpha} & \text{for } \alpha < \lambda \\ \infty & \text{for } \alpha \geq \lambda \end{cases} \quad \text{and } \Psi_\kappa(\alpha) = \lambda(e^\alpha - 1);$$

the distributions $(\tilde{H}_{ij})$ have finite moment generating function (namely $\mathbb{E}[e^{\alpha X}] < \infty$ for all $\alpha \in \mathbb{R}$ when the distribution of the random variable X is $\tilde{H}_{ij}$). For several commonly used distributions (Binomial, Poisson, Gaussian and so on) we have a finite moment generating function; on the contrary this does not happen when we deal with exponential or Gamma distributions. Anyway **(H3)** holds again if some of the distributions $(\tilde{H}_{ij})$ is exponential or Gamma.

3 The results

Our results concern the level crossing probabilities $(p(b))_{b>0}$ of $(\tilde{S}_{N_t} - ct)$, where $c > 0$ is a suitable constant as before (here b is the level); in detail we set

$$p(b) = P(T_b < \infty) \;\text{ where }\; T_b = \inf\{t \geq 0 : \tilde{S}_{N_t} - ct > b\}.$$

Since $\frac{\tilde{S}_{N_t}}{t}$ converges to some finite limit (this will be precised with eq. (3.3) below), we need to consider c large enough in order to avoid the trivial case $p(b) = 1$ for all $b > 0$; indeed, when c is large enough, we have $\tilde{S}_{N_t} - ct \to -\infty$ and then the level crossing can fail with positive probability. More precisely, as a comment of hypothesis **(H4)** below, we shall see that we have to consider $c > \Psi'_\kappa(\tilde{\kappa}(0))\tilde{\kappa}'(0)$. In Subsection 3.1 we provide the exponential decay for $p(b)$ as $b \to \infty$ in the fashion of large deviation theory. In Subsection 3.2 we consider a technique called importance sampling for the estimation by Monte Carlo simulations of $p(b)$ when b is large; more precisely, after recalling all the preliminaries on importance sampling, we prove that an admissible law is asymptotically efficient for simulations.

3.1 Estimates for level crossing probabilities

Let us start with the following Lemma.

Lemma 3.1 *Assume* **(H1)**-**(H3)** *hold. Then* $(\frac{\tilde{S}_{N_t}}{t})$ *satisfies the LDP with rate function* I *defined by*

$$I(x) = \sup_{\alpha \in \mathbb{R}} [\alpha x - \Psi_\kappa(\tilde{\kappa}(\alpha))] \quad (\forall x \in \mathbb{R}).$$

This means that $I : \mathbb{R} \to [0, \infty]$ *is a lower semicontinuous function and we have*

$$\liminf_{t \to \infty} \frac{1}{t} \log P(\frac{\tilde{S}_{N_t}}{t} \in O) \geq - \inf_{x \in O} I(x) \quad (\forall O \text{ open}) \tag{3.1}$$

and

$$\limsup_{t \to \infty} \frac{1}{t} \log P(\frac{\tilde{S}_{N_t}}{t} \in C) \leq - \inf_{x \in C} I(x) \quad (\forall C \text{ closed}).$$

Proof. We want to apply Gärtner Ellis Theorem (see e.g. Dembo and Zeitouni, 1993, Theorem 2.3.6, page 45) so that we have to check

$$\lim_{t \to \infty} \frac{1}{t} \log \mathbb{E}[e^{\alpha \tilde{S}_{N_t}}] = \Psi_\kappa(\tilde{\kappa}(\alpha)) \quad (\forall \alpha \in \mathbb{R}). \tag{3.2}$$

Then let $\alpha \in \mathbb{R}$ and $n \geq 1$ be arbitrarily fixed. Then let us start with the following equality which follows from **(H1)**:

$$\mathbb{E}[e^{\alpha \tilde{S}_{N_t}}] = \mathbb{E}[\mathbb{E}[e^{\alpha \tilde{S}_n}]_{n = N_t}].$$

Moreover, arguing as in Subsection 2.1 for $(\tilde{S}_n)$ in place of (S_n), we have

$$\frac{\sum_{i \in \tilde{E}} \tilde{h}_i(\alpha) P(\tilde{J}_0 = i)}{\max_{j \in \tilde{E}} \tilde{h}_j(\alpha)} e^{n\tilde{\kappa}(\alpha)} \leq \mathbb{E}[e^{\alpha \tilde{S}_n}] \leq \frac{\sum_{i \in \tilde{E}} \tilde{h}_i(\alpha) P(\tilde{J}_0 = i)}{\min_{j \in \tilde{E}} \tilde{h}_j(\alpha)} e^{n\tilde{\kappa}(\alpha)}$$

(with the obvious change of notation), whence we obtain

$$\frac{\sum_{i \in \tilde{E}} \tilde{h}_i(\alpha) P(\tilde{J}_0 = i)}{\max_{j \in \tilde{E}} \tilde{h}_j(\alpha)} \mathbb{E}[e^{N_t \tilde{\kappa}(\alpha)}] \leq \mathbb{E}[e^{\alpha \tilde{S}_{N_t}}] \leq \frac{\sum_{i \in \tilde{E}} \tilde{h}_i(\alpha) P(\tilde{J}_0 = i)}{\min_{j \in \tilde{E}} \tilde{h}_j(\alpha)} \mathbb{E}[e^{N_t \tilde{\kappa}(\alpha)}]$$

and (3.2) follows **(H3)**. $\diamondsuit$

Now let us consider the family of product measures $(P_\alpha \otimes \tilde{P}_{\tilde{\alpha}})$, varying $(\alpha, \tilde{\alpha}) \in \mathbb{R}^2$ such that $\kappa(\alpha), \tilde{\kappa}(\tilde{\alpha}) < \infty$. We pointed out above (2.2) whence we obtain

$$\lim_{t\to\infty} \frac{N_t}{t} = \frac{1}{\kappa'(\alpha)} \quad P_\alpha \text{ a.s. ;}$$

this can be seen as a consequence of the LDP for the inverse processes (under each P_α) in the paper of Glynn and Whitt (1994). Then, by also taking into account (2.2) for $(\tilde{S}_n)$ in place of (S_n), the trivial equality $\frac{\tilde{S}_{N_t}}{t} = \frac{\tilde{S}_{N_t}}{N_t}\frac{N_t}{t}$ and $\lim_{t\to\infty} N_t = \infty$ (almost surely), we obtain

$$\lim_{t\to\infty} \frac{\tilde{S}_{N_t}}{t} = \frac{\tilde{\kappa}'(\tilde{\alpha})}{\kappa'(\alpha)} \quad P_\alpha \otimes \tilde{P}_{\tilde{\alpha}} \text{ a.s..} \tag{3.3}$$

The following densities

$$\ell_n^{P,P_\alpha} \ell_n^{P,\tilde{P}_{\tilde{\alpha}}} = e^{-(\alpha S_n + \tilde{\alpha}\tilde{S}_n) + n(\kappa(\alpha)+\tilde{\kappa}(\tilde{\alpha}))} \frac{h_{J_0}(\alpha)}{h_{J_n}(\alpha)} \frac{\tilde{h}_{\tilde{J}_0}(\tilde{\alpha})}{\tilde{h}_{\tilde{J}_n}(\tilde{\alpha})} \tag{3.4}$$

of $P \otimes P$ with respect to $P_\alpha \otimes \tilde{P}_{\tilde{\alpha}}$ play a crucial role. This family of densities depends on the pair $(\alpha, \tilde{\alpha})$ and, in order to use the methods in the proofs of similar results in the literature, we need to consider a family of densities which depends on only one parameter. Then we shall consider the family of conjugate laws $(Q_{\tilde{\alpha}})_{\tilde{\alpha}:\tilde{\kappa}(\tilde{\alpha})<\infty}$ defined as follows:

$$Q_{\tilde{\alpha}} := P_{\kappa^{-1}(-\tilde{\kappa}(\tilde{\alpha}))} \otimes \tilde{P}_{\tilde{\alpha}} = P_{-\Psi_\kappa(\tilde{\kappa}(\tilde{\alpha}))} \otimes \tilde{P}_{\tilde{\alpha}}. \tag{3.5}$$

When we consider each law of the family $(Q_{\tilde{\alpha}})_{\tilde{\alpha}:\tilde{\kappa}(\tilde{\alpha})<\infty}$ in (3.5), the density in (3.4) will be denoted by $\ell_n^{P,Q_{\tilde{\alpha}}}$ and we have

$$\ell_n^{P,Q_{\tilde{\alpha}}} = e^{-(-\Psi_\kappa(\tilde{\kappa}(\tilde{\alpha}))S_n + \tilde{\alpha}\tilde{S}_n)} \frac{h_{J_0}(-\Psi_\kappa(\tilde{\kappa}(\tilde{\alpha})))}{h_{J_n}(-\Psi_\kappa(\tilde{\kappa}(\tilde{\alpha})))} \frac{\tilde{h}_{\tilde{J}_0}(\tilde{\alpha})}{\tilde{h}_{\tilde{J}_n}(\tilde{\alpha})}.$$

There are some (almost sure) inequalities which are useful in what follows. If we set

$$M(\tilde{\alpha}) = \max\{\frac{h_i(-\Psi_\kappa(\tilde{\kappa}(\tilde{\alpha})))}{h_j(-\Psi_\kappa(\tilde{\kappa}(\tilde{\alpha})))} \frac{\tilde{h}_{\tilde{i}}(\tilde{\alpha})}{\tilde{h}_{\tilde{j}}(\tilde{\alpha})} : i,j \in E,\ \tilde{i},\tilde{j} \in \tilde{E}\},$$

we have

$$\ell_n^{P,Q_{\tilde{\alpha}}} \leq M(\tilde{\alpha}) e^{-(-\Psi_\kappa(\tilde{\kappa}(\tilde{\alpha}))S_n + \tilde{\alpha}\tilde{S}_n)}. \tag{3.6}$$

Moreover, by construction, we have $S_{N_t} \leq t$ which implies

$$-(\tilde{S}_{N_t} - cS_{N_t}) \leq -(\tilde{S}_{N_t} - ct) \quad (\text{for any } c > 0). \tag{3.7}$$

Finally let us consider the following hypothesis which is the analogous of (A3) of Lehtonen and Nyrhinen (1992b):
(H4): *there exists* $w > 0$ *such that* $\Psi_\kappa(\tilde{\kappa}(w)) - cw = 0$ *and* $\Psi'_\kappa(\tilde{\kappa}(w))\tilde{\kappa}'(w) - c > 0$.
We point out that **(H4)** implies the existence of a positive zero w for the convex function

$$\Lambda_c(\alpha) = \Psi_\kappa(\tilde{\kappa}(\alpha)) - c\alpha$$

and moreover $\Lambda'_c(w) > 0$. Thus, since $\Lambda_c(0) = 0$ follows from $\tilde{\kappa}(0) = 0$ and $\Psi_\kappa(0) = 0$, the derivative of Λ_c at the origin is negative; then we have $0 > \Lambda'_c(0) = \Psi'_\kappa(0)\tilde{\kappa}'(0) - c$, which is equivalent to $c > \Psi'_\kappa(0)\tilde{\kappa}'(0)$. We also remark that, if **(H4)** holds, the level crossing can fail with positive probability; indeed we have $\tilde{S}_{N_t} - ct \to -\infty$ since $\frac{\tilde{S}_{N_t}}{t} - c$ converges to $\Lambda'_c(0) < 0$. Finally we point out the following simple equality concerning Q_w, namely $Q_w = P_{-cw} \otimes \tilde{P}_w$.

Proposition 3.2 *Assume* **(H1)**-**(H4)** *hold. Then* $\lim_{b\to\infty} \frac{1}{b} \log p(b) = -w$.

Proof. The limit follows from the lower bound

$$\liminf_{b\to\infty} \frac{1}{b} \log p(b) \geq -w$$

and the upper bound

$$\limsup_{b\to\infty} \frac{1}{b} \log p(b) \leq -w.$$

For the lower bound we use a standard procedure based on the lower bound for the open sets (3.1) for the LDP of $(\frac{\tilde{S}_{N_t}}{t})$; for instance see e.g. Lemma 2.1 of Duffield and O'Connell (1995) which deals with a more general situation. For all $x > 0$ we have

$$\liminf_{b\to\infty} \frac{1}{b} \log p(b) \geq \liminf_{b\to\infty} \frac{1}{b} \log P(\tilde{S}_{N_{xb}} - cxb > b) \geq$$

$$\geq x \liminf_{b\to\infty} \frac{1}{xb} \log P(\frac{\tilde{S}_{N_{xb}}}{xb} > \frac{1}{x} + c) \geq -x \inf_{y > \frac{1}{x}+c} I(y) = -\frac{I(\frac{1}{x}+c)}{\frac{1}{x}}$$

where the latter equality follows from the continuity of I. Then the lower bound holds by taking $\inf_{x>0} \frac{I(\frac{1}{x}+c)}{\frac{1}{x}}$ which is equal to w and it is attained when $\frac{1}{x} = \Lambda'_c(w)$ (this is the same situation in the proof of Theorem 1 of Lehtonen and Nyrhinen (1992a) and in the proof of Theorem 3.1 of Lehtonen and Nyrhinen (1992b)).

For the upper bound let $b > 0$ be arbitrarily fixed. Then

$$\begin{aligned}
p(b) &= \mathbb{E}_P[1_{T_b<\infty}] = \mathbb{E}_{Q_w}[\ell^{P,Q_w}_{N_{T_b}} 1_{T_b<\infty}] \leq^{(3.6)} \\
&\leq \mathbb{E}_{Q_w}[M(w)e^{-(-\Psi_\kappa(\tilde{\kappa}(w))S_{N_{T_b}} + w\tilde{S}_{N_{T_b}})} 1_{T_b<\infty}] = \\
&= M(w)\mathbb{E}_{Q_w}[e^{(\Psi_\kappa(\tilde{\kappa}(w)) - cw)S_{N_{T_b}} - w(\tilde{S}_{N_{T_b}} - cS_{N_{T_b}})} 1_{T_b<\infty}] \leq^{\textbf{(H4)} \text{ and } (3.7)} \\
&\leq M(w)\mathbb{E}_{Q_w}[e^{-w(\tilde{S}_{N_{T_b}} - cT_b)} 1_{T_b<\infty}] \leq M(w)\mathbb{E}_{Q_w}[e^{-wb} 1_{T_b<\infty}] = \\
&= M(w)e^{-wb} Q_w(T_b < \infty).
\end{aligned}$$

In conclusion we have $p(b) \leq M(w)e^{-wb}$ for all $b > 0$ and the upper bound holds. $\diamondsuit$

Remark 3.3 In view of what follows we need to show that $Q_w(T_b < \infty) = 1$ for all $b > 0$; this means that Q_w is an admissible law in some sense (this will be explained in the next subsection). In order to do that we point out that (3.5), (3.3) and $\Psi'_\kappa(\alpha) = \frac{1}{\kappa'(\kappa^{-1}(-\alpha))}$ provide

$$\lim_{t\to\infty} \frac{\tilde{S}_{N_t}}{t} = \frac{\tilde{\kappa}'(w)}{\kappa'(\kappa^{-1}(-\tilde{\kappa}(w)))} = \Psi'_\kappa(\tilde{\kappa}(w))\tilde{\kappa}'(w) \quad Q_w \text{ a.s.},$$

where $\Psi'_\kappa(\tilde{\kappa}(w))\tilde{\kappa}'(w) > c$ by **(H4)**. Thus $\tilde{S}_{N_t} - ct \to \infty$ under Q_w.

3.2 Importance sampling

In this subsection we show that Q_w is an asymptotically efficient simulation law when we use a technique called importance sampling for the estimation by Monte Carlo simulations of $p(b)$ when b is large (Proposition 3.3). In view of this we need to recall some preliminaries on importance sampling adapted to the framework of our interest and, in particular, the concept of asymptotically efficient simulation law.

Let us suppose we want to estimate $p(b)$ (for a fixed $b > 0$) by Monte Carlo simulation. If $T_b^{(1)}, \ldots, T_b^{(R)}$ are the values of T_b concerning R independent replications of $(\tilde{S}_{N_t} - ct)$ under the law P, an unbiased estimator is the relative frequency $\widehat{p(b)}$ of the level crossings:

$$\widehat{p(b)} = \frac{1}{R}\sum_{i=1}^{R} 1_{T_b^{(i)}<\infty}.$$

By Proposition 3.2 this Monte Carlo approach needs R growing exponentially with b to keep a fixed relative precision; indeed the relative precision of $\widehat{p(b)}$ is

$$\frac{1}{p(b)}\sqrt{\frac{p(b)(1-p(b))}{R}}.$$

Furthermore the simulation time under P is not finite when the level crossing does not occur.

Thus, in order to overcome these troubles, the idea is to consider R independent replications under another law Q chosen in a suitable way. First of all Q is such that P is absolutely continuous with respect to Q locally on the event $\{T_b < \infty\}$ and an unbiased estimator is

$$[\widehat{p(b)}]_Q = \frac{1}{R}\sum_{i=1}^{R} \ell^{P,Q}_{N_{T_b^{(i)}}} 1_{T_b^{(i)}<\infty}$$

where in general $\ell_{N_{T_b}}^{P,Q}$ is the local density of P with respect to Q.

Furthermore we choose Q in a class of admissible laws; this means that $Q(T_b < \infty) = 1$ for all $b > 0$, so that we have finite time in simulation almost surely. The choice of Q will be done in order to minimize in some sense $\mathrm{Var}_Q[[\widehat{p(b)}]_Q]$, which is equal to

$$\mathrm{Var}_Q[[\widehat{p(b)}]_Q] = \frac{\mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2 1_{T_b<\infty}] - p^2(b)}{R}.$$

In conclusion we can restrict our attention on the second moment $\mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2 1_{T_b<\infty}]$ and, since $Q(T_b < \infty) = 1$, this second moment can be rewritten as $\mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2]$.

The problem of minimizing $\mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2 1_{T_b<\infty}]$ for a fixed b is often intractable. Moreover, since in the applications we are interested in large values of b, in order to use standard features on large deviation theory we concentrate our attention on

$$\liminf_{b\to\infty} \frac{1}{b} \log \mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2 1_{T_b<\infty}];$$

then the next lower bound easily follows from Jensen inequality and Proposition 3.2:

$$\liminf_{b\to\infty} \frac{1}{b} \log \mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2 1_{T_b<\infty}] \geq$$

$$\geq \liminf_{b\to\infty} \frac{1}{b} \log \mathbb{E}_Q^2[\ell_{N_{T_b}}^{P,Q} 1_{T_b<\infty}] = \liminf_{b\to\infty} \frac{1}{b} \log(p(b))^2 = -2w.$$

In conclusion an admissible law Q is said to be an asymptotically efficient simulation law if

$$\lim_{b\to\infty} \frac{1}{b} \log \mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2 1_{T_b<\infty}] = -2w; \tag{3.8}$$

indeed, if R is chosen to guarantee a fixed relative precision

$$\frac{1}{p(b)} \sqrt{\frac{\mathbb{E}_Q[(\ell_{N_{T_b}}^{P,Q})^2 1_{T_b<\infty}] - p^2(b)}{R}}$$

of $[\widehat{p(b)}]_Q$, R has chance of growing less than exponentially if and only if (3.8) holds.

Proposition 3.3 *Assume* **(H1)-(H4)** *hold. Then* Q_w *is an asymptotically efficient simulation law.*

Proof. First of all Q_w is an admissible law by Remark 3.3. Then we have to check the limit (3.8) with $Q = Q_w$, which follows from the lower bound

$$\liminf_{b\to\infty} \frac{1}{b} \log \mathbb{E}_{Q_w}[(\ell_{N_{T_b}}^{P,Q_w})^2 1_{T_b<\infty}] \geq -2w$$

and the upper bound

$$\limsup_{b\to\infty} \frac{1}{b} \log \mathbb{E}_{Q_w}[(\ell^{P,Q_w}_{N_{T_b}})^2 1_{T_b<\infty}] \leq -2w.$$

The lower bound holds because this lower bound has been proved above for any admissible law. The upper bound is proved below and the procedure has some analogies with the proof of the upper bound in Proposition 3.2; indeed for all $b > 0$ we have

$$\mathbb{E}_{Q_w}[(\ell^{P,Q_w}_{N_{T_b}})^2 1_{T_b<\infty}] \leq^{(3.6)} \mathbb{E}_{Q_w}[(M(w))^2 e^{-2(-\Psi_\kappa(\tilde{\kappa}(w))S_{N_{T_b}}+w\tilde{S}_{N_{T_b}})} 1_{T_b<\infty}] =$$

$$= (M(w))^2 \mathbb{E}_{Q_w}[e^{2(\Psi_\kappa(\tilde{\kappa}(w))-cw)S_{N_{T_b}}-2w(\tilde{S}_{N_{T_b}}-cS_{N_{T_b}})} 1_{T_b<\infty}] \leq^{(\mathbf{H4})\text{ and }(3.7)}$$

$$\leq (M(w))^2 \mathbb{E}_{Q_w}[e^{-2w(\tilde{S}_{N_{T_b}}-cT_b)} 1_{T_b<\infty}] \leq (M(w))^2 \mathbb{E}_{Q_w}[e^{-2wb} 1_{T_b<\infty}] =$$

$$= (M(w))^2 e^{-2wb} Q_w(T_b < \infty) \leq (M(w))^2 e^{-2wb}$$

and the upper bound immediately follows. $\diamondsuit$

Acknowledgments

The author thanks the referee for pointing out an inexactness in an earlier version of this article. His suggestions led to an improvement in both the content and the readability of the paper. This work has been partially supported by Murst Project "Processi Stocastici e Applicazioni a Filtraggio, Controllo, Simulazione e Finanza Matematica".

(Received December, 2003. Accepted October, 2004.)

References

Asmussen, S. (1987). *Applied Probability and Queues.* New York: Wiley and Sons.

Asmussen, S. and Nielsen, H. M. (1995). Ruin probabilities via local adjustment coefficients. *J. Appl. Probab.*, **32**, 736–755.

Asmussen, S. (2000). *Ruin Probabilities.* London: World Scientific.

Baldi, P. and Piccioni, M. (1999). Importance sampling for continuous time Markov chains and applications to fluid models. *Methodol. Comput. Appl. Probab.*, **1**, 375–390.

Collamore, J. F. (2002). Importance sampling techniques for multidimensional ruin problem for general Markov additive sequences of random vectors. *Ann. Appl. Probab.*, **12**, 382–421.

Dembo, A. and Zeitouni, O. (1993). *Large Deviations Techniques and Applications.* Boston: Jones and Bartlett.

Duffield, N. G. and O'Connell, N. (1995). Large deviations and overflow probabilities for a general single server queue, with applications. *Math. Proc. Camb. Phil. Soc.*, **118**, 363–374.

Glynn, P. and Whitt, W. (1994). Large deviations behavior of counting processes and their inverses. *Queueing Systems Theory Appl.*, **17**, 107–128.

Grigelionis, B. (1994). Conditionally exponential families and Lundberg exponents of Markov additive processes. In: *Probability Theory and Mathematical Statistics*, (Vilnius, 1993), pp. 337–350. B. Grigelionis et al. Eds., TEV.

Iscoe, I., Ney, P. and Nummelin, E. (1985). Large deviations of uniformly recurrent Markov additive processes. *Adv. Appl. Math.*, **6**, 373–412.

Lehtonen, T. and Nyrhinen, H. (1992a). Simulating level crossing probabilities by importance sampling. *Adv. Appl. Prob.*, **24**, 858–874.

Lehtonen, T. and Nyrhinen, H. (1992b). On asymptotically efficient simulation of ruin probabilities in a Markovian environment. *Scand. Actuarial J.*, 60–75.

Macci, C. (2001). Simulating level crossing probabilities by importance sampling for non-decreasing compound Poisson process with bounded jumps and a negative drift. *Statist. Decisions*, **19**, 191–202.

Sadowsky, J. S. (1996). On Monte Carlo estimation of large deviations probabilities. *Ann. Appl. Probab.*, **6**, 399–422.

Revuz, D. (1975). *Markov Chains.* Amsterdam: North-Holland.

Claudio Macci
Dipartimento di Matematica
Università di Roma "Tor Vergata"
Via della Ricerca Scientifica
I-00133 Roma, Italia
E-mail: macci@mat.uniroma2.it