Statistics and Probability Letters 153 (2019) 151–156

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

# Fixed relative precision estimators of growth rate for compound Poisson and Lévy processes

Wojciech Niemiro

*Institute of Applied Mathematics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland*

ARTICLE INFO

*Article history:*
Received 12 October 2018
Received in revised form 14 April 2019
Accepted 14 June 2019
Available online 25 June 2019

*Keywords:*
Lévy processes
Relative error
Sequential estimation
Worst case efficiency
$\varepsilon$-$\alpha$-approximation
Confidence estimation

ABSTRACT

We consider compound Poisson processes or, more generally, Lévy processes $X(t)$ with positive bounded jumps. The problem is to estimate the "growth rate" $\mu = \mathbb{E}X(t)/t$ with fixed relative precision, i.e. to construct an estimator $\hat{\mu}$ such that $\mathbb{P}(|\hat{\mu} - \mu| < \mu\varepsilon) \geq 1 - \alpha$, for a given precision parameter $\varepsilon$ and confidence parameter $\alpha$, given a trajectory $X(t)$ for $0 \leq t \leq T$. Such an estimator must be sequential, i.e. the length $T$ of the observed trajectory must be random and chosen adaptively. Assume that the upper bound on jumps is known (w.l.o.g. equal to 1). We consider the estimator $\hat{\mu}_r = r/T_r$, where $T_r = \min\{t : X(t) \geq r\}$, with a suitably chosen $r = r(\varepsilon, \alpha)$. We show that this estimator is "nearly worst case optimal" in a certain asymptotic sense, for $\varepsilon \to 0$ and $\alpha \to 0$. The "worst case" turns out to be the process with jumps 1, i.e. the Poisson process with intensity $\mu$.



## 1. Introduction

The main motivation of this work is an estimation problem for compound Poisson processes. Let

$$X(t) = \sum_{i=1}^{N(t)} Y_i,$$

where $N(t)$, $t \geq 0$ is a homogeneous Poisson process with intensity $\lambda$ and $Y_i$s are i.i.d.positive random variables with $\mathbb{E}Y_i = m$, independent of $N(\cdot)$. Such processes have various applications, in particular in classical models of insurance mathematics. An important characteristic of $X(\cdot)$ is its "growth rate" or "drift parameter",

$$\mu = \mathbb{E}X(t)/t = \lambda m.$$

In the insurance models $Y_i$s are interpreted as subsequent losses incurred by a company. Therefore $\mu$ is the mean value of losses per unit time. Problem of estimation of this quantity is clearly of great practical interest. We consider estimation with fixed relative precision $\varepsilon$ at a given level of confidence $1-\alpha$. Thus we stipulate that an estimator $\hat{\mu}$ must satisfy the condition $\mathbb{P}(|\hat{\mu} - \mu| < \mu\varepsilon) \geq 1 - \alpha$. (Note that the inequalities are understood in the rigorous, nonasymptotic sense. In the theoretical computer science literature, such an estimator is referred to as "$\varepsilon$-$\alpha$-approximation", see e.g. Dagum et al. (2000).) Clearly, relative precision is important if $\mu$ is small (for example because the "loss" events are rare), but also if $\mu$ is big.

We assume that an estimator $\hat{\mu}$ is based on the trajectory $X(t)$ for $0 \leq t \leq T$. It is rather obvious (and will be proved rigorously in the course of our considerations) that the length $T$ of the observed trajectory cannot be fixed a priori and

*E-mail address:* wniem@mimuw.edu.pl.

https://doi.org/10.1016/j.spl.2019.06.009
0167-7152/

must be a random stopping time. A natural measure of efficiency of an estimator is $\mathbb{E}T$, the mean observation time needed to meet the fixed relative precision requirement.

We focus on the case when the jumps are positive and bounded, i.e. $0 < Y_i \leq h$. If $h$ is known then, without loss of generality, we can and will assume that $h = 1$. In this case, in Section 3 we propose a simple estimator $\hat{\mu}_r = r/T_r$ for $T_r = \inf\{t : X(t) \geq r\}$. (Note that $\hat{\mu}_r$ is close to the average value $X(T_r)/T_r$.) For $\hat{\mu}_r$ we prove exponential inequalities, analogous to the Hoeffding inequalities for bounded random variables. These inequalities allow us to choose $r$ such that $\hat{\mu}_r$ is an $\varepsilon$-$\alpha$-approximation and $\mathbb{E}T_r \sim \log \alpha^{-1}/\mu\varepsilon^2$ for $\varepsilon \to 0$ and $\alpha \to 0$. Details are given in Section 3. On the other hand, we show that the best possible sequential $\varepsilon$-$\alpha$-approximation must, in the least favourable case, use the trajectory of length $T$, where

$$\mathbb{E}T \sim \frac{\log \alpha^{-1}}{\mu\varepsilon^2}.$$

Thus our estimator $\hat{\mu}_r$ is an "asymptotically minimax" $\varepsilon$-$\alpha$-approximation. The least favourable case turns out to be a process with all jumps equal to 1, i.e. the Poisson process with intensity $\mu$. Details are given in Section 4.

Actually, we consider not only compound Poisson processes but also a more general class of Lévy processes. We assume that $X(\cdot)$ is a pure jump time-homogeneous process with independent increments which has positive and bounded jumps. The expected number of jumps of size $x$ at time $t$ is described by a measure $\nu(\mathrm{d}x)\times \mathrm{d}t$, where $\nu$ is called the Lévy measure. In the special case of a compound Poisson process, we have $\nu(\mathrm{d}x) = \lambda\rho(\mathrm{d}x)$, where $\rho$ is the probability distribution of a single summand $Y_i$ and $\lambda$ is the intensity of the Poisson process $N(\cdot)$. Clearly, $\mu = \int x\nu(\mathrm{d}x)$.

In this paper we construct a sequential estimator of $\mu$ with fixed relative precision and show it is in some sense nearly worst case optimal. Our main results, Theorems 3.1, 3.2 and 4.1 are analogues of the theorems established in Gajek et al. (2013) for sums of i.i.d. summands. In the setup of Lévy processes, the proofs are based on an auxiliary result of independent interest (Corollary 3.1, an analogue of Bennett inequality). Considering Levy processes instead of i.i.d. sequences not only allows for new important applications (e.g. to classical models of the risk theory), but also sheds new light on the previous results in Gajek et al. (2013). In particular, we highlight the key role of the Poisson process as the "worst case" and its relation to the Bennett inequality.

## 2. Statement of the problem

Let $X(t)$, $t \geq 0$ be a pure jump Lévy process with Lévy measure $\nu$ concentrated on $[0, 1]$. Suppose that the process $X(\cdot)$ is adapted to filtration $\{\mathcal{F}_t\}$ and $X(t+h) - X(t)$ is independent of $\mathcal{F}_t$ for $h > 0$. We have

$$\mathbb{E}\exp\{sX(t)\} = \exp\left\{t\int_0^1 (\mathrm{e}^{sx} - 1)\nu(\mathrm{d}x)\right\}.$$

It is well-known that

$$\mathbb{E}X(t) = t\mu,$$

where

$$\mu = \int_0^1 x\nu(\mathrm{d}x)$$

must be finite.

Our objective is to estimate $\mu$ (the "growth rate" of the process), given a trajectory $X(t)$, $0 \leq t \leq T$. We will construct an estimator $\hat{\mu}$ which has fixed relative precision in the following sense:

$$\mathbb{P}(|\hat{\mu} - \mu| < \mu\varepsilon) \geq 1 - \alpha, \tag{2.1}$$

for a given precision parameter $\varepsilon$ and a confidence parameter $\alpha$. Every estimator satisfying the above requirement (for all values of $\mu$) has to be sequential, i.e. the length of trajectory $T$ has to be an adaptively chosen random variable. This fact, already noted in the Introduction, will be established later as an easy corollary to Theorem 4.1 in Section 4.

## 3. Construction of an $\varepsilon$-$\alpha$-approximation

### 3.1. Auxiliary results

Our main tool is an exponential inequality for a class of Lévy processes, Corollary 3.1. We begin with the following lemma.

**Lemma 3.1.** *Under our standing assumptions on the process $X(t)$, for all $s > 0$,*

$$\mathbb{E}\exp\{s[X(t) - t\mu]\} \leq \exp\left\{t\mu\left(\mathrm{e}^{s} - s - 1\right)\right\},$$
$$\mathbb{E}\exp\{s[t\mu - X(t)]\} \leq \exp\left\{t\mu\left(\mathrm{e}^{-s} + s - 1\right)\right\}.$$

**Proof.** We will show that the first inequality holds for all real $s$. Clearly,

$$\mathbb{E}\exp\{s[X(t)-t\mu]\} = \exp\left\{t\int_0^1 \left(e^{sx}-sx-1\right)\nu(dx)\right\}.$$

Let $q(dx) = \dfrac{x\nu(dx)}{\mu}$, so that

$$\int_0^1 \left(e^{sx}-sx-1\right)\nu(dx) = \mu\int_0^1 \frac{e^{sx}-sx-1}{x}q(dx).$$

Note that $q$ is a probability measure on $[0, 1]$ and thus the integral above is bounded by the maximum value of the integrand. Now it is enough to notice that the maximum is attained at $x = 1$,

$$\frac{e^{sx}-sx-1}{x} \le e^s - s - 1.$$

Indeed, it is easy to verify that the derivative of the integrand function is positive, because

$$\frac{d}{dx}\left(\frac{e^{sx}-sx-1}{x}\right) > 0 \ \text{ iff } \ e^{-sx} > 1 - sx,$$

and the right hand side inequality is true for all $s$, both positive and negative.

It is worth noting the following interpretation of Lemma 3.1. The moment generating function $\mathbb{E}\exp\{s[X(t)-t\mu]\}$ is maximum if $X$ is the Poisson process with rate $\mu$, i.e. all the jumps are of size 1. This fact sheds some light on our "worst case result", Theorem 4.1.

**Corollary 3.1.** *Under our standing assumptions on the process $X(t)$, if $0 < \varepsilon < 1$,*

$$\mathbb{P}(X(t) \ge t\mu(1+\varepsilon)) \le \exp\{-t\mu[(1+\varepsilon)\log(1+\varepsilon)-\varepsilon]\},$$
$$\mathbb{P}(X(t) \le t\mu(1-\varepsilon)) \le \exp\{-t\mu[(1-\varepsilon)\log(1-\varepsilon)+\varepsilon]\}.$$

**Proof.** Consider the first inequality. By Markov inequality combined with the first part of Lemma 3.1,

$$\begin{aligned}\mathbb{P}(X(t)-t\mu \ge t\mu\varepsilon) &\le \mathbb{E}\exp\{s[X(t)-t\mu]\}\exp\{-st\mu\varepsilon)\\ &\le \exp\left\{t\mu\left(e^s-s-1-s\varepsilon\right)\right\}.\end{aligned}$$

Now it is enough to minimize $e^s - s - 1 - s\varepsilon$ with respect to $s$. It is easy to check that the minimum is obtained for $s = \log(1+\varepsilon)$ and is equal to $-[(1+\varepsilon)\log(1+\varepsilon)-\varepsilon] < 0$.

The proof of the second inequality is analogous, using the second part of Lemma 3.1.

Corollary 3.1 is a continuous time version of the classical Bennett inequality for sums of i.i.d.random variables (Bennett, 1962). Note also that in Gajek et al. (2013, Lemma A.1), a role analogous to our Corollary 3.1 is played by one of the Hoeffding inequalities (Hoeffding, 1963, Theorem 1).

*3.2. The sequential estimator $\hat{\mu}_r$*

Let us fix a number $r > 0$ and put

$$T_r = \inf\{t > 0 : X(t) \ge r\} \tag{3.1}$$

and

$$\hat{\mu}_r = \frac{r}{T_r}. \tag{3.2}$$

It is clear that $T_r$ is a stopping time with respect to the filtration $\{\mathcal{F}_t\}$ and $\hat{\mu}_r$ is $\mathcal{F}_{T_r}$-measurable. The pair $(T_r, \hat{\mu}_r)$ defines a sequential estimation procedure.

**Theorem 3.1.** *For $\hat{\mu}_r$ defined above,*

$$\mathbb{P}(\hat{\mu}_r - \mu \ge \mu\varepsilon) \le \exp\{-rk(\varepsilon)\},$$
$$\mathbb{P}(\mu - \hat{\mu}_r \ge \mu\varepsilon) \le \exp\{-rk(-\varepsilon)\},$$

*where function $k$ is defined (for all real $\varepsilon \ne 0$) as*

$$k(\varepsilon) = \log(1+\varepsilon) - \frac{\varepsilon}{1+\varepsilon}. \tag{3.3}$$

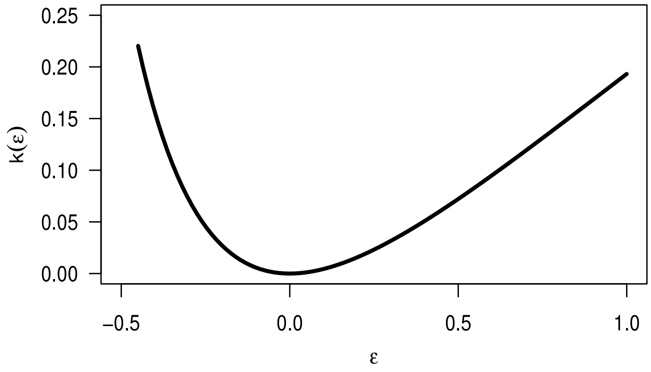


**Fig. 1.** Graph of $k(\varepsilon)$.

**Proof.** Write $T = T_r$ and $\hat{\mu} = \hat{\mu}_r$. To show the first of the inequalities, we use (3.1), (3.2) and the first part of Corollary 3.1, with $t = r/(\mu(1+\varepsilon))$. Indeed, the definition of $T_r$ and the right continuity of sample paths of $X(\cdot)$ imply that

$$\begin{aligned}\mathbb{P}(\hat{\mu} - \mu \geq \mu\varepsilon) &= \mathbb{P}\left(\frac{r}{T} \geq \mu(1+\varepsilon)\right) = \mathbb{P}\left(T \leq \frac{r}{\mu(1+\varepsilon)}\right) \\ &= \mathbb{P}\left(X\left(\frac{r}{\mu(1+\varepsilon)}\right) \geq r\right) \\ &\leq \exp\left\{-\frac{r}{1+\varepsilon}[(1+\varepsilon)\log(1+\varepsilon) - \varepsilon]\right\} = \exp\{-rk(\varepsilon)\}.\end{aligned}$$

To prove the second inequality, just use the second part of Corollary 3.1 in an analogous way.

We should mention that the inequalities in Theorem 3.1 are exactly the same as those derived in Gajek et al. (2013, Th. 2.3) for sums of i.i.d. bounded random variables. However, in the Lévy process setup considered here, Corollary 3.1 leads to a more straightforward proof.

It is easy to see that $k(\varepsilon) > 0$, $k(\varepsilon) \sim \varepsilon^2/2$ for $\varepsilon \to 0$ and $k(\varepsilon) < k(-\varepsilon)$ for $0 < \varepsilon < 1$. The graph of function $k(\varepsilon)$ is displayed in Fig. 1.

Theorem 3.1 gives bounds independent of $\mu$. Thus it can be directly used to find a fixed relative precision estimator. Let

$$r(\varepsilon, \alpha) = \inf\{r : \exp[-rk(\varepsilon)] + \exp[-rk(-\varepsilon)] \leq \alpha\}.$$

It is easy to see that

$$r(\varepsilon, \alpha) \leq \frac{\log(\alpha/2)^{-1}}{k(\varepsilon)}.$$

Indeed, if $\bar{r} = \log(\alpha/2)^{-1}/k(\varepsilon)$ then $\exp[-\bar{r}k(\varepsilon)] + \exp[-\bar{r}k(-\varepsilon)] < \alpha/2 + \alpha/2 = \alpha$, so $r(\varepsilon, \alpha) \leq \bar{r}$. Note also that

$$r \leq \mu\mathbb{E}T_r < r + 1. \tag{3.4}$$

Inequality (3.4) follows from a Lévy process version of the 1st Wald identity,

$$\mathbb{E}X(T) = \mu\mathbb{E}T,$$

applied to the stopping time $T = T_r$. Our assumptions clearly imply $r \leq X(T_r) < r + 1$. Combining Theorem 3.1 with (3.4), we obtain the following result.

**Theorem 3.2.** *If $r = r(\varepsilon, \alpha)$ then estimator $\hat{\mu}_r = r/T_r$ satisfies (2.1) and*

$$\mu\mathbb{E}T_r \leq \frac{\log(\alpha/2)^{-1}}{k(\varepsilon)} + 1.$$

## 4. Worst case lower bound for $\mathbb{E}T$ for an $\varepsilon$-$\alpha$-approximation

Theorem 3.2 gives an upper bound for the expected time of observation $T_r$ required by the $\varepsilon$-$\alpha$-approximation $\mu_r$ constructed in the previous section. We are going to prove that this bound is close to the necessary minimum. In this section we consider an *arbitrary sequential estimation procedure* $(T, \hat{\mu})$, where $T$ is an almost surely finite stopping time and $\hat{\mu}$ is $\mathcal{F}_T$-measurable (depends only on $X(t)$ for $t \le T$). We show that if $\hat{\mu}$ is an $\varepsilon$-$\alpha$-approximation then the expected value of $T$ cannot be much smaller than that of $T_r$.

We consider the case when the process $X(t)$ is simply a homogeneous Poisson process with intensity parameter $\mu$ (any $\varepsilon$-$\alpha$-approximation which works for Lévy processes satisfying our assumptions must work in this particular case). From now on we use the following notations. For emphasis let us write $X(t) = N(t)$. Since we will consider different probability measures, let $\mathbb{P}_\mu$ and $\mathbb{E}_\mu$ correspond to the parameter $\mu$ (i.e. $\mathbb{E}_\mu N(t) = \mu t$ for example).

**Theorem 4.1.** *If $(T, \hat{\mu})$ is a sequential estimation procedure which satisfies (2.1), then in the case of Poisson processes we have*

$$\mu \mathbb{E}_\mu T \ge \frac{\log(2\alpha)^{-1}}{k(-\varepsilon)} \cdot \left( 1 - \frac{\sqrt{2\log(2\alpha)^{-1} + 1} - 1}{\log(2\alpha)^{-1}} \right).$$

Note that both the upper bound in Theorem 3.2 and the lower bound in Theorem 4.1 are of the form $b(\varepsilon, \alpha)$, where $b(\varepsilon, \alpha) \sim \frac{1}{2}\varepsilon^2 \log \alpha^{-1}$ as $\varepsilon \to 0$ and $\alpha \to 0$. In this sense, the procedure given in the previous section is nearly optimal for small $\varepsilon$ and $\alpha$. The worst case (from the viewpoint of efficiency of $\varepsilon$-$\alpha$-approximations) among Lévy processes with jumps in $[0, 1]$ turns out to be Poisson processes with jumps 1.

**Proof.** Fix $\mu$ and put

$$\mu_1 = \frac{\mu}{1+\varepsilon}, \qquad \mu_2 = \frac{\mu}{1-\varepsilon}.$$

Let

$$p(n, t, \mu) = \mathbb{P}_\mu\left[N(t) = n\right] = \mathrm{e}^{-\mu t} \frac{(\mu t)^n}{n!}.$$

Now consider the likelihood ratios

$$\begin{aligned} \frac{p(n, t, \mu_1)}{p(n, t, \mu)} &= \mathrm{e}^{-(\mu_1 - \mu)t} \left( \frac{\mu_1}{\mu} \right)^n \\ &= \exp\left\{ \frac{\varepsilon}{1+\varepsilon} \mu t - n \log(1+\varepsilon) \right\} \\ &= \exp\left\{ -(n - \mu t)\log(1+\varepsilon) - k(\varepsilon)\mu t \right\} \end{aligned}$$

and

$$\begin{aligned} \frac{p(n, t, \mu_2)}{p(n, t, \mu)} &= \mathrm{e}^{-(\mu_2 - \mu)t} \left( \frac{\mu_2}{\mu} \right)^n \\ &= \exp\left\{ -\frac{\varepsilon}{1-\varepsilon} \mu t - n \log(1-\varepsilon) \right\} \\ &= \exp\left\{ -(n - \mu t)\log(1-\varepsilon) - k(-\varepsilon)\mu t \right\}. \end{aligned}$$

The processes

$$M_1(t) = \exp\left\{ -\,(N(t) - \mu t)\log(1+\varepsilon) - k(\varepsilon)\mu t \right\}$$

and

$$M_2(t) = \exp\left\{ -\,(N(t) - \mu t)\log(1-\varepsilon) - k(-\varepsilon)\mu t \right\}$$

are martingales with respect to measure $\mathbb{P}_\mu$, with $\mathbb{E}_\mu M_i(t) = 1$. We have the following "change of measure" formulae: $\mathbb{P}_{\mu_i}(A) = \mathbb{E}_\mu M_i(t)\mathbb{I}(A)$, for $A \in \mathcal{F}_t$, see e.g. (Asmussen and Albrecher, 2010, Prop. 1.2 in Ch. III). These formulae hold also if we replace deterministic $t$ by a stopping time $T$, see (Asmussen and Albrecher, 2010, Th. 1.3 in Ch. III). Consequently, since the events $\{\hat{\mu} \ge \mu\}$ and $\{\hat{\mu} < \mu\}$ belong to $\mathcal{F}_T$, we obtain

$$\begin{aligned} \mathbb{P}_{\mu_1}(\hat{\mu} \ge \mu) &= \mathbb{E}_\mu M_1(T)\mathbb{I}(\hat{\mu} \ge \mu), \\ \mathbb{P}_{\mu_2}(\hat{\mu} < \mu) &= \mathbb{E}_\mu M_2(T)\mathbb{I}(\hat{\mu} < \mu), \end{aligned} \tag{4.1}$$

Clearly, (2.1) implies

$$\mathbb{P}_{\mu_1}\big[\hat{\mu} \ge \mu\big] \le \alpha \quad \text{and} \quad \mathbb{P}_{\mu_2}\big[\hat{\mu} < \mu\big] \le \alpha.$$

It is easy to check that $k(\varepsilon) < k(-\varepsilon)$ and $\log(1+\varepsilon) < -\log(1-\varepsilon) < \sqrt{2k(-\varepsilon)}$. Therefore, (4.1) and Jensen inequality imply

$$\begin{aligned}2\alpha &\geq \mathbb{E}_\mu \min\left[M_1(T), M_2(T)\right]\\ &\geq \mathbb{E}_\mu \exp\left\{-\sqrt{2k(-\varepsilon)}\left|N(T)-\mu T\right| - k(-\varepsilon)\mu T\right\}\\ &\geq \exp\left\{-\sqrt{2k(-\varepsilon)}\mathbb{E}_\mu\left|N(T)-\mu T\right| - k(-\varepsilon)\mu\mathbb{E}_\mu T\right\}.\end{aligned}$$

Equivalently,

$$\begin{aligned}-\log(2\alpha) &\leq \sqrt{2k(-\varepsilon)}\mathbb{E}_\mu\left|N(T)-\mu T\right| + k(-\varepsilon)\mu\mathbb{E}_\mu T\\ &\leq \sqrt{2k(-\varepsilon)\mathbb{E}_\mu\left(N(T)-\mu T\right)^2} + k(-\varepsilon)\mu\mathbb{E}_\mu T.\end{aligned}$$

To complete the proof, notice that a Poisson process version of the second Wald identity gives

$$\mathbb{E}_\mu\left(N(T)-\mu T\right)^2 = \mu\mathbb{E}_\mu T.$$

To simplify notation, let $z = \mu\mathbb{E}_\mu T$, $l = -\log(2\alpha)$ and $k = k(-\varepsilon)$. We have shown that

$$l \leq \sqrt{2kz} + kz.$$

This is in fact a quadratic inequality with respect to $y = \sqrt{2kz}$. If we solve this inequality, we get $y \geq \sqrt{1+2l} - 1$. Hence

$$kz = \frac{y^2}{2} \geq \frac{1}{2}\left(\sqrt{1+2l}-1\right)^2 = l\left(1 - \frac{\sqrt{2l+1}-1}{l}\right),$$

which is equivalent to the conclusion of the theorem.

**Remark 4.1.** It is worth mentioning that $k(\varepsilon)\mu t$ is the Kullback–Leibler divergence between Poisson distributions with parameters $\mu t$ and $\mu_1 t = \mu t/(1+\varepsilon)$,

$$k(\varepsilon)\mu t = -\mathbb{E}_\mu \log\frac{p\left(N(t), t, \mu/(1+\varepsilon)\right)}{p\left(N(t), t, \mu\right)} = -\mathbb{E}_\mu \log M_1(t).$$

Analogously, $k(-\varepsilon)\mu t = -\mathbb{E}_\mu \log M_2(t)$.

Note that the conclusion of Theorem 4.1 is similar to its counterpart for Bernoulli schemes (Gajek et al., 2013, Th. 3.2).

**Remark 4.2.** In this paper we assumed that the upper bound on the size of jumps is known (it was set to 1 without loss of generality). If we drop this assumption then $\varepsilon$-$\alpha$ approximation does not exist. To see this, it is enough to consider a process $X(\cdot)$ with Lévy measure $\nu$ concentrated on $[0, 1]$ and a "perturbed process" $X_1(\cdot) = X(\cdot) + gN(\cdot)$, where $N(\cdot)$ is an independent Poisson process with (small) intensity $\gamma$ and $g$ is a (big) constant. Let $\mu = \mathbb{E}X(1)$ and $\mu_1 = \mathbb{E}X_1(1) = \mu + \gamma g$. It is easy to see that no $\varepsilon$-$\alpha$ approximation can work correctly for both $X(\cdot)$ and $X_1(\cdot)$. Indeed, if $(T, \hat{\mu})$ were an $\varepsilon$-$\alpha$ approximation then for $X(\cdot)$ we would get $\mathbb{P}(\hat{\mu} < \mu(1+\varepsilon)) \geq 1-\alpha$. We could choose $t$ such that $\mathbb{P}(T \leq t) \geq 1-\alpha$, then $\gamma$ such that $\mathbb{P}(N(t) = 0) = \mathrm{e}^{-\gamma t} \geq 1-\alpha$ and finally $g$ so big that $\mu(1+\varepsilon) < \mu_1(1-\varepsilon)$. If we now apply $(T, \hat{\mu})$ to $X_1(\cdot)$ then with probability at least $1-2\alpha$ we would get the same result as for $X(\cdot)$, thus obtaining $\mathbb{P}(\hat{\mu} < \mu_1(1-\varepsilon)) \geq 1-3\alpha$, which contradicts the requirement $\mathbb{P}(\hat{\mu} > \mu_1(1-\varepsilon)) \geq 1-\alpha$.

## References

Asmussen, Søren, Albrecher, Hansjörg, 2010. Ruin Probabilities, second ed. In: Advanced Series on Statistical Science & Applied Probability, vol. 14, World Scientific Publishing Co. Pte. Ltd. Hackensack, NJ, ISBN: 978-981-4282-52-9; 981-4282-52-9, http://dx.doi.org/10.1142/9789814282536.

Bennett, George, 1962. Probability inequalities for the sum of independent random variables. J. Amer. Statist. Assoc. (ISSN: 0162-1459) 57 (297), 33–45, URL http://www.jstor.org/stable/2282438.

Dagum, Paul, Karp, Richard, Luby, Michael, Ross, Sheldon, 2000. An optimal algorithm for Monte Carlo estimation. SIAM J. Comput. (ISSN: 0097-5397) 29 (5), 1484–1496. http://dx.doi.org/10.1137/S0097539797315306.

Gajek, Lesław, Niemiro, Wojciech, Pokarowski, Piotr, 2013. Optimal Monte Carlo integration with fixed relative precision. J. Complexity (ISSN: 0885-064X) 29 (1), 4–26. http://dx.doi.org/10.1016/j.jco.2012.09.001.

Hoeffding, Wassily, 1963. Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc. (ISSN: 0162-1459) 58, 13–30, URL http://links.jstor.org/sici?sici=0162-1459(196303)58:301<13:PIFSOB>2.0.CO;2-D&origin=MSN.