APPROXIMATION OF LÉVY PROCESSES

GIUSEPPE DI BENEDETTO & MARCIN MIDER & ANDI WANG

08/03/2016

Lévy processes are a diverse and mathematically interesting collection of stochastic processes, generalising both Brownian motion and the Poisson process. In this report we introduce Lévy processes, describe the approach of [1] to approximate a general real-valued Lévy process, and show some simulated examples.

1 INTRODUCTION

Lévy processes constitute a very wide class of stochastic processes which contain well-known processes such as Poisson and Brownian processes and they are defined as follows:

Definition 1 (Lévy Process)**.** A stochastic process $X = \{X(t)\}_{t \geqslant 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a Lévy process if it satisfies the following properties:

1. $\mathbb{P}(X(0) = 0) = 1$
2. For $0 \leqslant s \leqslant t$, $X(t) - X(s) \sim X(t - s)$
3. For $0 \leqslant s \leqslant t$, $X(t) - X(s)$ is independent of $\{X(u) : u < s\}$
4. X is stochastically continuous: $\forall \epsilon > 0$, $\forall s > 0$

$$\lim_{t \to s} \mathbb{P}\Big(|X(t) - X(s)| > \epsilon\Big) = 0.$$

This definition does not include any restriction on the distribution of the marginals. The apparently innocuous lack of any distributional assumptions makes this class of processes huge, in the sense that it comprises processes whose sample paths are very different. It is not clear from the definition how the paths of a Lévy process should look. For instance, let us just think about the above mentioned examples: Brownian motion has $\mathbb{P}$-almost surely continuous paths of unbounded variation, whereas Poisson processes have non decreasing (and therefore of bounded variation) $\mathbb{P}$-almost surely discontinuous paths. In fact it can be shown that for any Lévy process X there exists a version of X whose paths are $\mathbb{P}$-almost surely right-continuous with left limits and we will work with this version henceforth.

One characterization of Lévy processes which also gives insight into their path properties is provided by the strict relationship between these processes and the infinitely divisible distributions, defined as follows.

Definition 2 (Infinitely divisible distribution)**.** A real-valued random variable Y is said to have infinitely divisible distribution if for each $n \in \mathbb{N}$ there exist independent and identically distributed random variables $Y_{1,n}, \dots, Y_{n,n}$ such that $Y \sim Y_{1,n} + \cdots + Y_{n,n}$.

If $\{X(t)\}_{t \geqslant 0}$ is a Lévy process then for each $n \in \mathbb{N}$, X_t can be seen as sum of n i.i.d. random variables: $X(t) = X(\frac{t}{n}) + (X(\frac{2t}{n}) - X(\frac{t}{n})) + \cdots + (X(t) - X(\frac{(n-1)t}{n}))$, therefore its distribution is infinitely divisible because of independent and stationary increments. Defining the characteristic exponent of a random variable Y to be $\Psi_Y(\theta) = -\log \mathbb{E}(e^{i\theta Y})$, it follows that for every $t > 0$ the characteristic exponent $\Psi_t(\theta) = t\Psi_1(\theta)$. [1] Hence a Lévy process is characterized by the characteristic exponent of X_1, which is called characteristic exponent of the Lévy process and it is denoted by $\Psi(\theta)$.

The next result shows the importance of the link between Lévy processes and infinitely divisible distributions, in fact the characteristic exponent of the latter has a specific form given by the Lévy–Khintchine formula.

Theorem 1 (Lévy–Khintchine formula)**.** *The probability distribution μ of a real-valued random variable is infinitely divisible if and only if there exists a triple (a, b, Q) where $a \in \mathbb{R}$, $b \geqslant 0$, Q is a measure concentrated on $\mathbb{R} \setminus \{0\}$ (called Lévy measure) satisfying $\int_{\mathbb{R}} (1 \wedge x^2)\, Q(dx) < \infty$, such that for every $\theta \in \mathbb{R}$*

$$\Psi(\theta) = ia\theta + \frac{1}{2} b^2 \theta^2 + \int_{\mathbb{R}} \left(1 - e^{i\theta x} + i\theta x \mathbf{1}_{(|x|<1)}\right) Q(dx). \tag{1}$$

Let us remark that the integral condition on the Lévy measure Q ensures the convergence of the integral. The relationship is even closer since a Lévy process can be defined from an infinitely divisible distribution.

Theorem 2. *For each triple (a, b, Q) such that $a \in \mathbb{R}$, $b \geqslant 0$, Q is a measure concentrated on $\mathbb{R} \setminus \{0\}$ satisfying $\int_{\mathbb{R}} (1 \wedge x^2)\, Q(dx) < \infty$, there exists a probability space on which a Lévy process with characteristic exponent Ψ defined as in (1) is defined.*

In order to better understand the Lévy–Khintchine formula let us just write down Ψ for two well known Lévy processes: the compound Poisson process and linear Brownian motion.

1.1 Compound Poisson process

Let $\{\beta_i\}_{i \geqslant 1}$ be i.i.d. random variables with distribution F, $N(t) \sim \text{Poisson}(\lambda t)$ and $X(t) = \sum_{i=1}^{N(t)} \beta_i$ for each $t \geqslant 0$. Then it is easy to see that X is a Lévy process with

$$\Psi(\theta) = -\log \mathbb{E}\left(e^{i\theta \sum_{i=1}^{N(1)} \beta_i}\right) = \lambda \int_{\mathbb{R}} (1 - e^{i\theta x}) F(dx).$$

Therefore $Q(dx) = \lambda F(dx)$, $b = 0$, $a = -\lambda \int_{0<|x|<1} x F(dx)$. It is worth noting that F represents the distribution of the size of the jumps and λ is the rate of the jumps of the process; intuitively $\lambda dt\, Q(dx)$ is the probability that a jump of size dx occurs in a time interval of length dt.

1 It is sufficient to prove the equality for $t \in \mathbb{Q}_+$ and extend the result to $\mathbb{R}_+$ taking $t_n \downarrow t$ where $(t_n)_n \subset \mathbb{Q}$, exploiting the right continuity of the Lévy process.

1.2 Linear Brownian motion

A linear Brownian motion is defined as $X(t) = \sigma W(t) + ct$ for each $t \geqslant 0$, where $\sigma \geqslant 0$, $c \in \mathbb{R}$ and $\{W(t)\}_{t\geqslant 0}$ is standard Brownian motion. Then

$$\Psi(\theta) = -ic\theta + \frac{1}{2}\sigma^2\theta^2$$

therefore $Q(dx) = 0$, $b = \sigma$, $a = -c$.

It seems that the characteristic exponent of the most general Lévy process comprises of terms of these forms. It can be shown that the linear combination of independent Lévy processes is again a Lévy process and we can rewrite the Lévy–Khintchine formula as follows [2]

$$\Psi(\theta) = \left(ia\theta + \frac{1}{2}b^2\theta^2\right) + \left(Q(\mathbb{R}\setminus(-1,1))\int_{|x|>1}(1-e^{i\theta x})\frac{Q(dx)}{Q(\mathbb{R}\setminus(-1,1))}\right) + \left(\int_{0<|x|<1}(1-e^{i\theta x}+i\theta x)Q(dx)\right).$$

Therefore a general Lévy process can be written as a sum of independent linear Brownian motion (first addend), a compound Poisson process (second addend) whose rate is $Q(\mathbb{R}\setminus(-1,1))$ and jump distribution is $F(A) = \int_A I_{(|x|>1)}Q(dx)/Q(\mathbb{R}\setminus(-1,1))$ for every Borel set A, and a third addend which can be written as a countable sum of independent compound Poisson processes with different rates and drifts splitting $[0,1]$ into $\bigcup_{n\geqslant 0}[2^{-(n+1)}, 2^{-n})$.

It is worth remarking that the integral condition on Q implies that for each Borel set B such that $0 \in B \setminus \partial B$ we have $Q(\mathbb{R}\setminus B) < \infty$, whereas $Q(B)$ can be finite or infinite. In the latter case we would have that $Q(-\epsilon,\epsilon) = \infty$ for each $\epsilon > 0$, which means that there are infinitely many jumps of arbitrarily small size. After this remark it should be clear that the choice of $B = (-1,1)$ in the formula above is just for convention, and we could choose $B = (-\epsilon,\epsilon)$ for an arbitrary $\epsilon > 0$.

2 APPROXIMATIONS

Let X be a real-valued Lévy process with Lévy–Khintchine representation as in (1). We have seen that Lévy processes are mathematically interesting and diverse. Now, from a practical point of view, can we simulate X? The drift and Brownian components are easy to simulate, so the question is whether we can simulate the pure-jump process. When the Lévy measure Q is finite, $\|Q\| < \infty$, this is straightforward, as the jump process is a compound Poisson process. On the other hand, when Q is infinite, $\|Q\| = \infty$, the "rate" of jumps becomes infinite and some approximation becomes necessary.

As we have seen, a Lévy measure can only fail to be finite in a neighbourhood of 0. Hence, one approach is to remove the jumps of magnitude below some threshold $\epsilon > 0$ [1]. Writing $\tilde{Q} := Q_{\{|x|\geqslant\epsilon\}}$, this leads to an approximation of X of the form

$$X_1^\epsilon(t) = \mu_\epsilon t + bW(t) + N^\epsilon(t) \tag{2}$$

where

$$\mu_\epsilon = a - \int_{\epsilon\leqslant|x|\leqslant 1} x\,Q(dx)$$

2 See [2] (section 3) for more details.

and $N^\epsilon(t)$ is a Poisson process with (finite) rate $\|\tilde{Q}\|$ and jump distribution $\tilde{Q}/\|\tilde{Q}\|$ independent of the standard Brownian motion W. Heuristically, we can think of N^ϵ as the jumps bigger than ϵ,

$$\sum_{s \leqslant t} \Delta X(s) \mathbf{1}(|\Delta X(s)| \geqslant \epsilon).$$

In the finite variation case, in the representation (2) we are obtaining X_1^ϵ from X by replacing small jumps by their expectation.

We would expect that as $\epsilon \to 0$, X_1^ϵ should approximate X better and better. So a natural question to ask is, what does the error process $X_\epsilon = \{X_\epsilon(t) : t \geqslant 0\}$ given by

$$X_\epsilon(t) = X(t) - X_1^\epsilon(t) \tag{3}$$

look like as $\epsilon \to 0$? Intuitively, we might expect this process, suitable scaled, to look more and more like a standard Brownian motion. And indeed, this can be the case on compacts.

Let $\sigma^2(\epsilon) := \int_{|x|<\epsilon} x^2 Q(dx)$.

Theorem 1. *[1] Restricting our processes to $t \in [0,1]$,*

$$\sigma^{-1}(\epsilon) X_\epsilon \xrightarrow{D} W \text{ as } \epsilon \to 0^{3}$$

if and only if for each $\kappa > 0$

$$\sigma(\kappa\sigma(\epsilon) \wedge \epsilon) \sim \sigma(\epsilon) \text{ as } \epsilon \to 0. \tag{4}$$

This condition is rather mysterious, and a more straightforward condition is given by the following.

Theorem 2. *[1] A sufficient condition for (4) is*

$$\lim_{\epsilon \to 0} \frac{\sigma(\epsilon)}{\epsilon} = \infty. \tag{5}$$

Moreover, if Q does not have atoms in some neighbourhood of 0, (5) is also necessary for (4).

This says that we have convergence to Brownian motion if the variance of the small jumps of X converge to 0 more slowly than the truncation level ϵ.

When this error is approximately Brownian, we can include this to obtain a better approximation of X given by

$$X_2^\epsilon(t) = \mu_\epsilon t + (b^2 + \sigma^2(\epsilon))^{1/2} W(t) + N^\epsilon(t). \tag{6}$$

In [1] the authors also proved a Berry–Esseen-type estimate for the speed of convergence of the distribution functions of $X_2^\epsilon(1)$ to $X(1)$.

Theorem 3. *[1]*

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(X_2^\epsilon(1) \leqslant x) - \mathbb{P}(X(1) \leqslant x)| \leqslant 0.7975 \sigma^{-3}(\epsilon) \int_{|x|<\epsilon} |x|^3 \, Q(dx). \tag{7}$$

3 This weak convergence is in the space $D[0,1]$ of cadlag functions $\omega : [0,1] \to \mathbb{R}$ equipped with the uniform norm; that is, for each $f : D[0,1] \to \mathbb{R}$ that is continuous with respect to the uniform norm, bounded, and measurable with respect to the cylinder σ-algebra, we have $\mathbb{E} f(\sigma^{-1}(\epsilon) X_\epsilon) \to \mathbb{E} f(W)$.

2.1 Simulation

In principle, for any $t \geqslant 0$ both $X_1^\epsilon(t)$ and $X_2^\epsilon(t)$ are straightforward to simulate. The drift and Brownian parts are trivially simulated, and since N^ϵ is a Poisson process, it is easily simulated in the following way.

First, simulate the number of jumps in $[0,t]$, $M \sim \text{Poisson}(t\|\tilde{Q}\|)$. By properties of the Poisson process, we know that conditional on the number of jumps M in $[0,t]$ the locations of these jumps $\{t_i\}_{i=1}^M$ are independent and uniform on $[0,t]$. For each jump time t_i, we simulate the size of the jump itself independently according to the probability measure $\tilde{Q}/\|\tilde{Q}\|$, assuming this can be done.

For error comparisons later, we will also want to compare our approximation X_1^ϵ to X, the *true* process. But of course we cannot simulate X (otherwise we would not be doing this!), so instead we compare against $X_1^{\tilde{\epsilon}}$ where $\tilde{\epsilon} > 0$ is many orders of magnitude smaller than ϵ.

3 RESULTS AND DISCUSSION

In this section we illustrate some of the previously presented theorems on three examples of simple Lévy processes. These are

- stable process of index $\alpha \in (0,1)$:

  $$Q(dx) = c_1|x|^{-1-\alpha}\mathbf{1}_{\{x<0\}}dx + c_2x^{-1-\alpha}\mathbf{1}_{\{x\geqslant 0\}}dx, \quad c_1, c_2 \geqslant 0, \quad c_1 + c_2 > 0$$

- gamma process:

  $$Q(dx) = cx^{-1}e^{-\tau x}\mathbf{1}_{\{x>0\}}dx, \quad c, \tau > 0$$

- generalised gamma process:

  $$Q(dx) = \frac{1}{\Gamma(1-\alpha)}x^{-1-\alpha}e^{-\tau x}\mathbf{1}_{\{x>0\}}dx, \quad \alpha \in (0,1), \quad \tau > 0$$

where throughout dx denotes Lebesgue measure on $\mathbb{R}$ and $a = b = 0$ in (1). For each of the above we show whether conditions for the normality of the errors as defined in (3) are satisfied, and we support the formal arguments with the empirical comparisons to (appropriately re-scaled) normal densities. We then examine the validity of the Berry–Esseen bound (7).

3.1 Stable process

We start with the stable process, for which all of the results discussed in this report hold.

For fixed $\epsilon > 0$, simulation of X_1^ϵ as defined in (2) in the stable case is not difficult, using the procedure outlined in Section 2. We can analytically calculate the rate of N^ϵ, $\|\tilde{Q}\| = (c_1 + c_2)\epsilon^{-\sigma}/\sigma$, and its jump distribution $\tilde{Q}/\|\tilde{Q}\|$ can be simulated as a Pareto random variable, with sign chosen randomly depending on the ratio of c_1 to c_2.

3.1.1 *Normal approximation to errors* $X_\epsilon(t)$

It is easy to see that:

$$\sigma(\epsilon) = \left[(c_1 + c_2)\int_0^\epsilon x^{1-\alpha}dx\right]^{1/2} = \left[\frac{(c_1 + c_2)}{2-\alpha}\right]^{1/2}\epsilon^{1-\alpha/2},$$

from which $\sigma(\epsilon)/\epsilon \to \infty$ as $\epsilon \to 0$ is a trivial consequence. Therefore validity of a normal approximation to the errors follows by a direct application of Theorem 2. To see how an approximation from Theorem 1 holds in practice, we compare the empirical distribution of $\sigma^{-1}(\epsilon)X_\epsilon(1)$ with the standard normal distribution and present the results in a Q-Q plot. Figure 1 illustrates the case of a stable process with parameters $\sigma = 0.5$, $c_1 = 1$, $c_2 = 0$, $\epsilon \in \{0.1, 0.01, 0.001, 0.0001\}$. Even for the relatively crude cutoff point

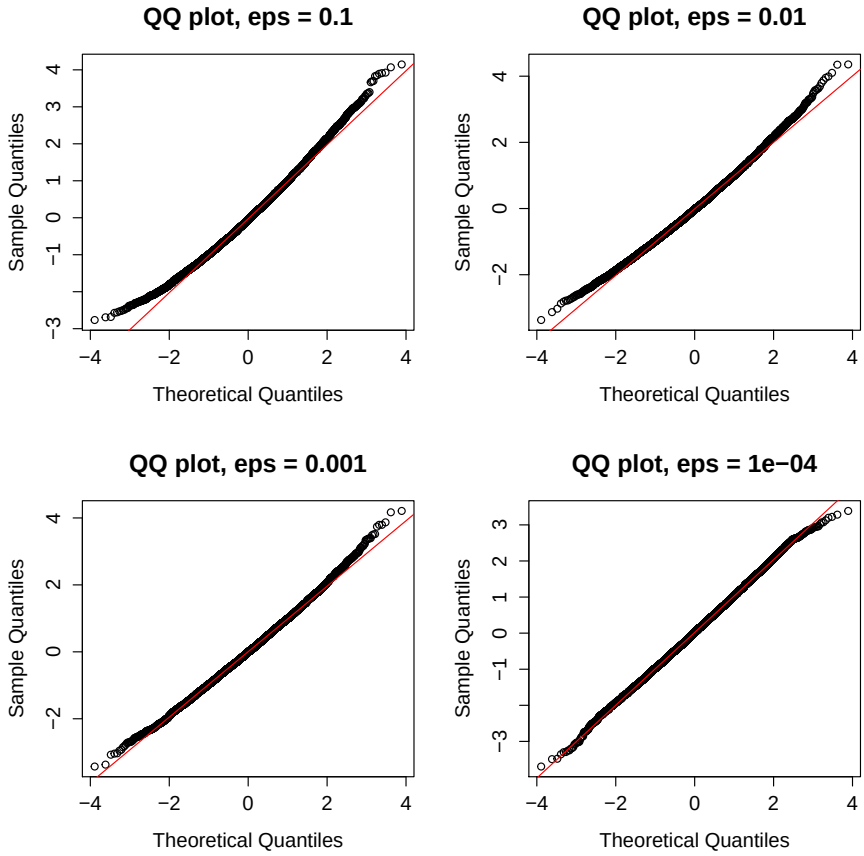


Figure 1: Q-Q plots for the error term of a stable process with parameters $\sigma = 0.5$, $c_1 = 1$, $c_2 = 0$.

of $\epsilon = 0.1$ normal approximation to the error term seems to be very good. For smaller values of ϵ the results become even more convincing.

3.1.2 *Speed of convergence*

Having established that $X_2^\epsilon(t)$ is an improved approximation to $X(t)$, we want to compare the marginal distributions of the two. Consider figure 2, which gives four realisations of a stable process. By simulating multiple such paths we compute the empirical marginal distribution of the approximate simulator $X_2^\epsilon(1)$ at time $t = 1$ and compare it to the theoretical marginal distribution of $X(1)$, known to have the law of a stable random variable. Figure 3 gives the results of such comparison. The two densities are almost identical. To give a more quantitative comparison we look at $E_\epsilon(x) := |\hat{P}(X_2^\epsilon(1) \leqslant x) - \hat{P}(X(1) \leqslant x)|$ (with $\hat{P}(X < \cdot)$ denoting an empirical c.d.f.), plotted in figure 4 and also verify that it is below the Berry–Esseen

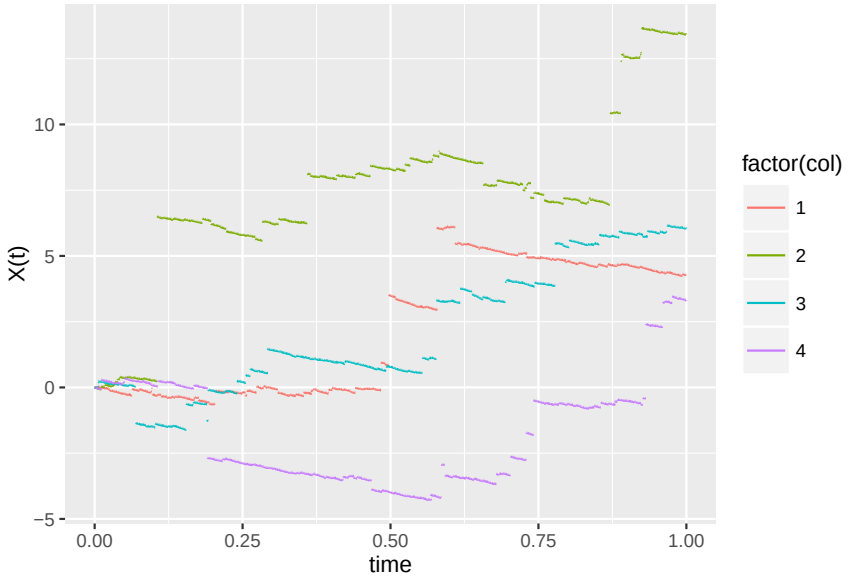


Figure 2: Four sample paths of a stable process with parameters $\sigma = 0.9$, $c_1 = 2$, $c_2 = 0.4$.

bound Function $E_\epsilon(x)$ gives the discrepancy between the two c.d.f.'s, which for this case appears to be always less than 0.01. Note that it is well below the Berry–Esseen bound. Finally, we plot $\sup_x E_\epsilon(x)$ for various values of ϵ together with the Berry–Esseen bounds and give the results in figure 5. Overall, the empirical results are in line with the theory, confirming convergence in distribution of the simulated approximate sample paths.

3.2 Gamma and Generalised Gamma processes

To simulate gamma and generalised gamma processes we used an adaptive thinning method, described in [3]. This generates both the number of jumps and the jumps sizes, so there is no need to use the Poisson process approach as in Section 2. Some sample paths of the generalised gamma process are shown in figure 6.

3.2.1 *Normal approximation to errors* $X_\epsilon(t)$

The gamma process is an example of a Lévy process for which the Brownian approximation does not hold. Notice:

$$\sigma^2(\epsilon) = \int_0^\epsilon x^2 c x^{-1} e^{-\tau x} dx = \int_0^\epsilon c x(1 + o(x))dx = \frac{c\epsilon^2}{2} + o(\epsilon^3),$$

hence $\sigma(\epsilon)/\epsilon \sim \sqrt{c/2} \not\to \infty$. Since $Q(dx)$ of the gamma process has no atoms, it follows by Theorem 2 that the distribution of the errors defined in (3) can no longer be approximated by a normal law. Figure 7 is a Q-Q plot for a gamma process which demonstrates how the normal approximation can fail.

On the other hand, for the generalised gamma process, we have:

$$\sigma^2(\epsilon) = \int_0^\epsilon x^2 \Gamma(1-\alpha)^{-1} x^{-1-\alpha} e^{-\tau x} dx = \int_0^\epsilon \Gamma(1-\alpha) x^{1-\alpha}(1 + O(x))dx =$$
$$= \Gamma(2-\alpha)^{-1}\epsilon^{2-\alpha} + O(\epsilon^{3-\alpha}),$$

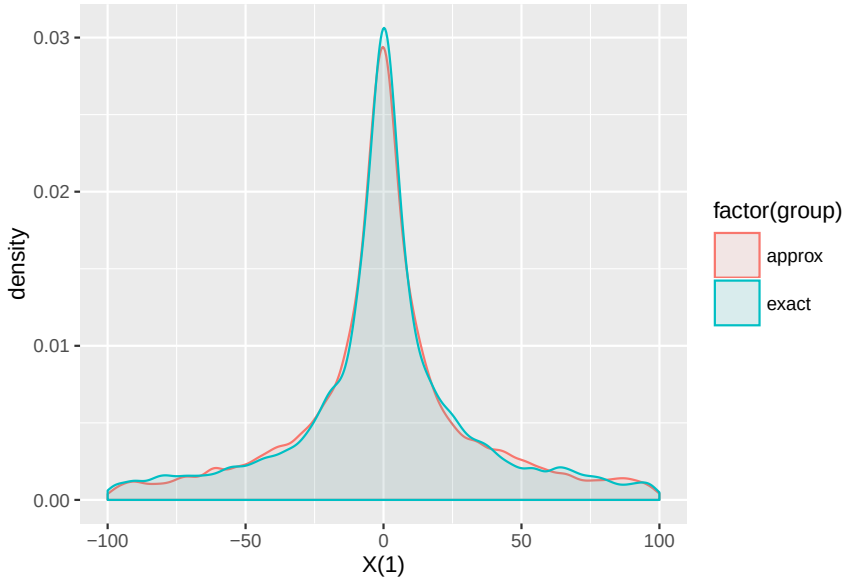


Figure 3: Marginal densities of the approximate process X_2^ϵ and stable distribution, for parameters: $\sigma = 0.5$, $c_1 = 1$, $c_2 = 1$, $\epsilon = 0.01$ (these are for the Lévy process, parameters for the stable distribution are determined implicitly).

thus $\sigma(\epsilon)/\epsilon \sim \Gamma(2-\alpha)^{-1}\epsilon^{-\alpha/2} \to \infty$. Therefore, by Theorem 2 the distribution of the re-scaled errors tends to standard Gaussian. Figure 8 shows Q-Q plots for the errors of two generalised gamma processes against the appropriately re-scaled normal distribution. Though the approximation is not as good as in the case of a stable process, shrinking ϵ to zero definitely leads to an improvement in the approximation.

3.2.2 *Speed of convergence*

In order to assess the speed of convergence in distribution of an approximated process $X_2^\epsilon(t)$ to the exact one $X(t)$ we once again compare the marginal distributions of the processes at time $t = 1$, using Kolmogorov–Smirnov distance as a measure of discrepancy and we verify that it agrees with the theoretical Berry–Esseen bound. We apply this methodology only to the generalised gamma case, since - as outlined in the previous subsection - the approximation is invalid for a regular gamma process. The results are plotted in figure 9. The empirical results seem consistent with the theory.

3.3 Further comments

We have showed through simulations that if X does not satisfy the condition in 1, the normal approximation of the error $\sigma(\epsilon)^{-1}X_\epsilon(1)$ does not hold (cf. Gamma process example). In [1] the authors describe how big the set of possible limits can be. Denoting by $\mathcal{L}(Y)$ the law of a r.v. Y, and by $\mathcal{I}_0$ the set of infinitely divisible distributions with zero mean and variance less of equal than one, we define [4]

$$\mathcal{I}(Q) = \bigcap_{\delta>0} \overline{\{\mathcal{L}(\sigma(\epsilon)^{-1}X_\epsilon(1)) : 0 < \epsilon < \delta\}} \subset \mathcal{I}_0.$$

4 The closure is w.r.t. weak convergence.

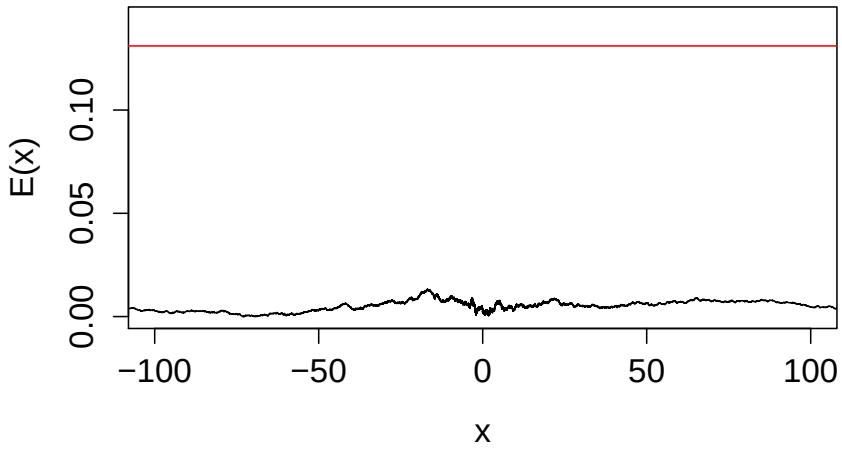


Figure 4: Absolute difference between the empirical cdf's of approximate and exact stable process, with parameters $\sigma = 0.5$, $c_1 = 1$, $c_2 = 1$, $\epsilon = 0.01$.

Theorem 4.1 in [1] actually proves the existence of a Lévy measure Q such that $\mathcal{I}(Q) = \mathcal{I}_0$, i.e. there exists a Lévy process X with Lévy measure Q such that for every other Lévy process Y such that $\mathcal{L}(Y(1)) \in \mathcal{I}_0$ there exists a sequence $\epsilon_n \downarrow 0$ such that $\sigma(\epsilon_n)^{-1} X_{\epsilon_n}(1) \xrightarrow{D} Y(1)$.

REFERENCES

[1] Jan Rosiński Søren Asmussen. Approximations of small jumps of lévy processes with a view towards simulation. *Journal of Applied Probability*, 38(2):482–493, 2001.

[2] Andreas Kyprianou. Lévy processes and continuous-state branching processes: part i. *http://www.maths.bath.ac.uk/~ak257/LCSB/part1.pdf*, *accessed March 2016*.

[3] Stefano Favaro, Yee Whye Teh, et al. Mcmc for normalized random measure mixture models. *Statistical Science*, 28(3):335–359, 2013.

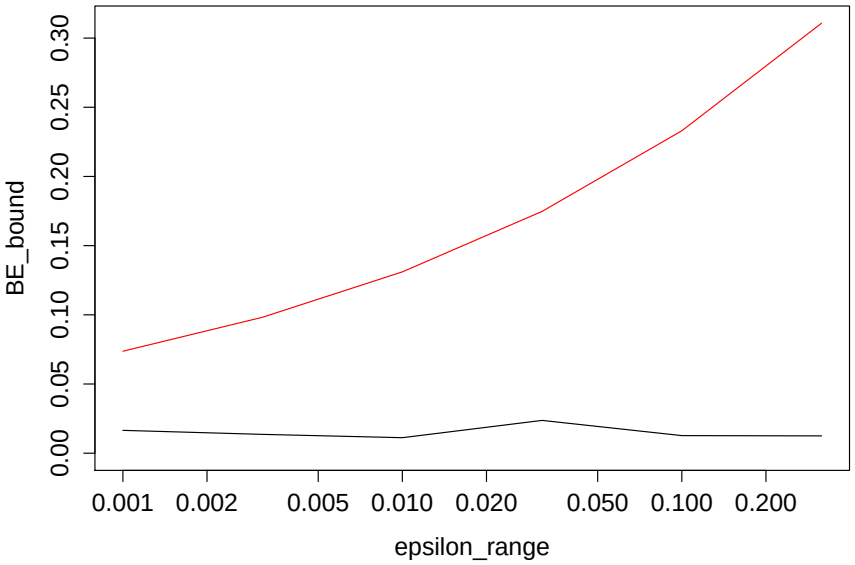


Figure 5: Comparison of Berry–Esseen bound (red line) and empirical Kolmogorov–Smirnov distance between the exact and approximate marginal distribution (black line), when parameters are $\sigma = 0.5$, $c_1 = 1$, $c_2 = 1$.

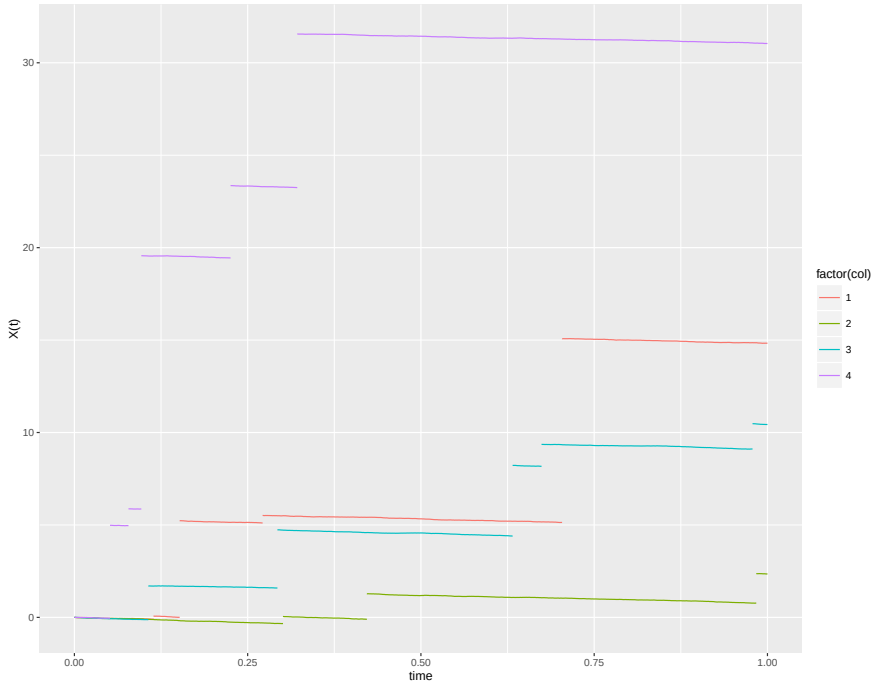


Figure 6: Four sample paths of a generalised gamma process with parameters $\sigma = 0.5$, $\tau = 0.5$.

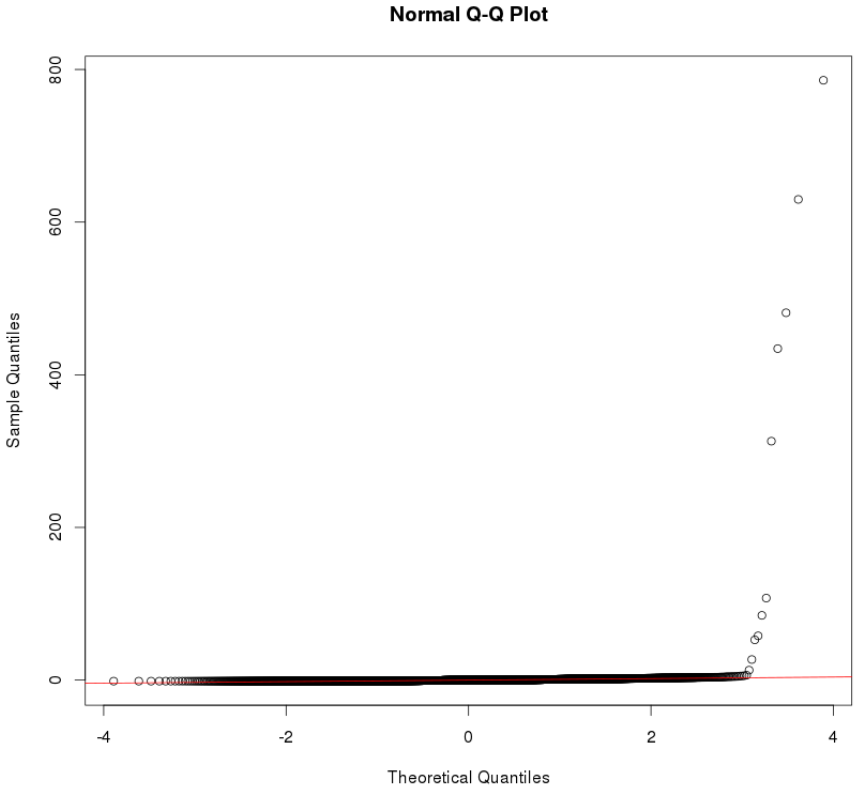
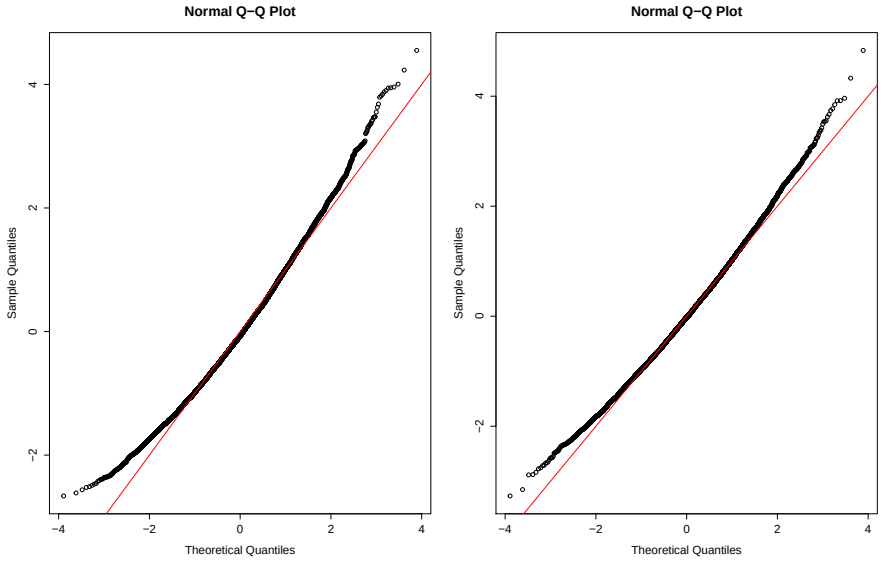


Figure 7: Q-Q plot for the error term of a gamma process with parameters $c = 1$, $\tau = 0.5$ for $\epsilon = 0.01$.



(a) $\epsilon = 0.1$ **(b)** $\epsilon = 0.01$

Figure 8: Q-Q plots for the error term of a generalised gamma processes with parameters $\alpha = 0.5, \tau = 0.5$

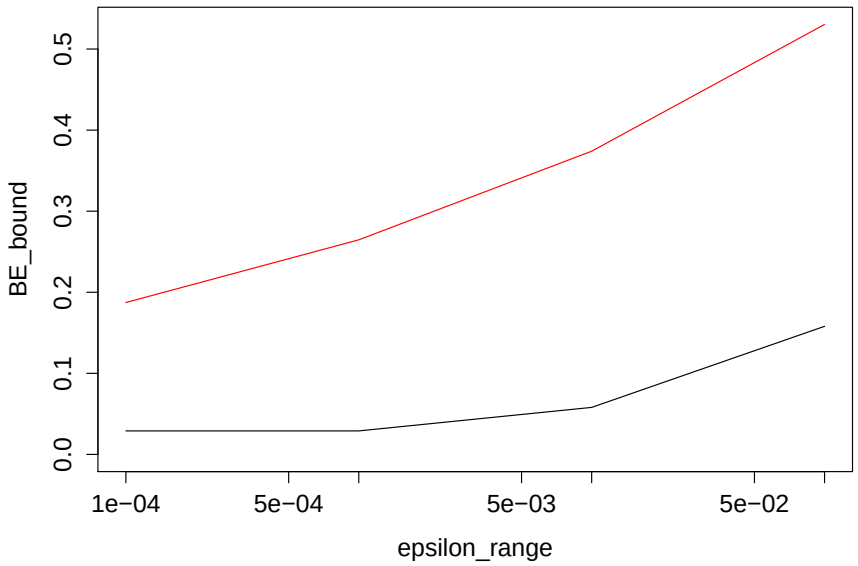


Figure 9: Comparison of Berry–Esseen bound (red line) and empirical Kolmogorov–Smirnov distance between the exact and approximate marginal distribution (black line), when parameters are $\sigma = 0.3$, $\tau = 0.2$.