Statistical test for some multistable processes

Ronan Le Guével

Joint work in progress with A. Philippe

Journées MAS 2014

First definition : Ferguson-Klass-LePage series

Consider $(E, \mathcal{E}, m)$ a (σ-)finite measure space. Let $\alpha \in (0, 2)$ and $(f_t)_t$ a family of functions such that $\int_E |f_t(x)|^\alpha m(dx) < +\infty$: we can define the stochastic process

$$I(f_t) = \int f_t(x) M(dx)$$

where M is a symmetric α-stable random measure with control measure m: f_t is seen as a limit of simple functions.

First definition : Ferguson-Klass-LePage series

Consider $(E, \mathcal{E}, m)$ a (σ-)finite measure space. Let $\alpha \in (0, 2)$ and $(f_t)_t$ a family of functions such that $\int_E |f_t(x)|^\alpha m(dx) < +\infty$:
we can define the stochastic process

$$I(f_t) = \int f_t(x) M(dx)$$

where M is a symmetric α-stable random measure with control measure m: f_t is seen as a limit of simple functions.

- We use here an other representation [Lévy-Véhel, LG (2012)].

Let

- $(\Gamma_i)_{i\geq 1}$ be a sequence of arrival times of a Poisson process with unit arrival time,
- $(V_i)_{i\geq 1}$ a sequence of i.i.d. random variables with distribution $\hat{m} = m/m(E)$,
- $(\gamma_i)_{i\geq 1}$ a sequence of i.i.d. random variables with distribution $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$.
- with the three sequences $(\Gamma_i)_{i\geq 1}$, $(V_i)_{i\geq 1}$, and $(\gamma_i)_{i\geq 1}$ independent.

Let $\alpha \in (0,2)$ and $c(\alpha) = (2\alpha^{-1}\Gamma(1-\alpha)\cos(\frac{\pi\alpha}{2}))^{-1/\alpha}$:

Let

- $(\Gamma_i)_{i\geq 1}$ be a sequence of arrival times of a Poisson process with unit arrival time,
- $(V_i)_{i\geq 1}$ a sequence of i.i.d. random variables with distribution $\hat{m} = m/m(E)$,
- $(\gamma_i)_{i\geq 1}$ a sequence of i.i.d. random variables with distribution $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$.
- with the three sequences $(\Gamma_i)_{i\geq 1}$, $(V_i)_{i\geq 1}$, and $(\gamma_i)_{i\geq 1}$ independent.

Let $\alpha \in (0,2)$ and $c(\alpha) = (2\alpha^{-1}\Gamma(1-\alpha)\cos(\frac{\pi\alpha}{2}))^{-1/\alpha}$:

$$I(f_t) \stackrel{d}{=} (\frac{\alpha}{2}m(E))^{-1/\alpha} c(\alpha) \sum_{i=1}^{+\infty} \gamma_i \Gamma_i^{-1/\alpha} f_t(V_i).$$

We define the multistable processes on an open interval U:

We define the multistable processes on an open interval U:

let $\alpha : U \to [c, d] \subset (0, 2)$ a continuous function, $f_t(x)$ a family of functions such that $\forall t \in U, \int_E |f_t(x)|^{\alpha(t)} m(dx) < +\infty$:

$$Y(t) := (\frac{\alpha(t)}{2} m(E))^{-1/\alpha(t)} c(\alpha(t)) \sum_{i=1}^{+\infty} \gamma_i \Gamma_i^{-1/\alpha(t)} f_t(V_i)$$

is a symmetric multistable process, with kernel f_t and stability function α.

- Characteristic function: we compute

$$\mathbb{E}[e^{i\theta Y(t)}] = \exp\left(-|\theta|^{\alpha(t)} \int_E |f_t(x)|^{\alpha(t)} m(dx) \right) .$$

Local structure of multistable processes

Let Y be a real stochastic process.

- Property of localisability : Y admits a tangent process. [Falconer (2002,2003)]

Definition

A real stochastic process $Y = \{Y(t) : t \in \mathbb{R}\}$ is $h-$localisable at u if there exists an $h \in \mathbb{R}$ and a non-trivial limiting process Y'_u (the local form) such that

$$\lim_{r \to 0} \frac{Y(u+rt) - Y(u)}{r^h} = Y'_u(t)$$

where convergence occurs in finite dimensional distributions.

Example

The Multistable Lévy Motion:

$$Y(t) = K(\alpha(t)) \sum_{i=1}^{+\infty} \gamma_i \Gamma_i^{-\frac{1}{\alpha(t)}} \mathbf{1}_{[0,t]}(V_i), t \in [0,1]$$

where α is a $\mathcal{C}^1$ function ranging in $(1,2)$, $(V_i)_i$ is i.i.d. uniformly distributed on $(0,1)$, and the kernel $f_t(x) = \mathbf{1}_{[0,t]}(x)$.

Example

The Multistable Lévy Motion:

$$Y(t) = K(\alpha(t)) \sum_{i=1}^{+\infty} \gamma_i \Gamma_i^{-\frac{1}{\alpha(t)}} \mathbf{1}_{[0,t]}(V_i), t \in [0,1]$$

where α is a $\mathcal{C}^1$ function ranging in $(1,2)$, $(V_i)_i$ is i.i.d. uniformly distributed on $(0,1)$, and the kernel $f_t(x) = \mathbf{1}_{[0,t]}(x)$.

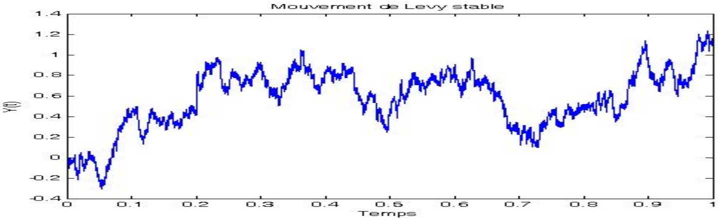
The process Y is $\frac{1}{\alpha(u)}$-localisable at $u \in [0,1]$, with

$$Y_u'(t) = L_{\alpha(u)}(t)$$

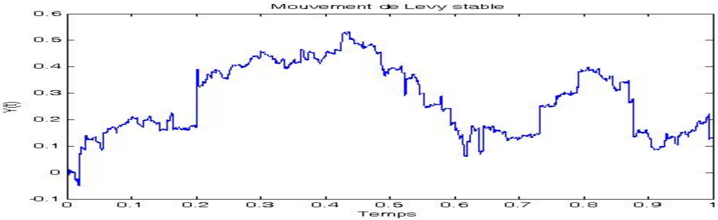
(the Stable Lévy Motion).

Example of a Stable Lévy Motion

- Trajectory of a α-stable Lévy process, $\alpha = 1.9$

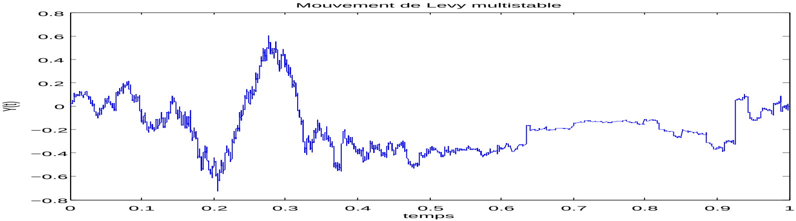


- Trajectory of a α-stable Lévy process, $\alpha = 1.1$

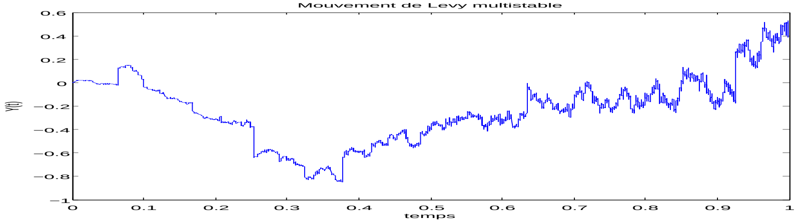


Example of multistable Lévy process

- $\alpha(t) = 1,5 + 0,48 \sin(2\pi t)$, $h(t) = \frac{1}{\alpha(t)}$.



- $\alpha(t) = 1,02 + 0,96t$.



Second definition: multistable measures

[Falconer-Liu (2010)]: Let $\alpha : \mathbb{R} \to (0, 2)$. We define the family of stable integrals $\{I(f), f \in F\}$ as a stochastic process indexed by a set F of functions. We specify its finite-dimensional distributions, and apply the Kolmogorov's existence theorem. Here

$$F = \{f : \int_E |f(x)|^{\alpha(x)} m(dx) < +\infty\}.$$

Second definition: multistable measures

[Falconer-Liu (2010)]: Let $\alpha : \mathbb{R} \to (0,2)$. We define the family of stable integrals $\{I(f), f \in F\}$ as a stochastic process indexed by a set F of functions. We specify its finite-dimensional distributions, and apply the Kolmogorov's existence theorem. Here

$$F = \{f : \int_E |f(x)|^{\alpha(x)} m(dx) < +\infty\}.$$

Given $f_1, ..., f_d \in F$, we define a probability measure $P_{f_1,...,f_d}$ in $\mathbb{R}^d$ by its characteristic function:

$$\phi_{f_1,...,f_d}(\theta_1, ..., \theta_d) := \exp\left(-\int_E |\sum_{j=1}^{d} \theta_j f_j(x)|^{\alpha(x)} m(dx)\right),$$

Second definition: multistable measures

[Falconer-Liu (2010)]: Let $\alpha : \mathbb{R} \to (0, 2)$. We define the family of stable integrals $\{I(f), f \in F\}$ as a stochastic process indexed by a set F of functions. We specify its finite-dimensional distributions, and apply the Kolmogorov's existence theorem. Here

$$F = \{f : \int_E |f(x)|^{\alpha(x)} m(dx) < +\infty\}.$$

Given $f_1, ..., f_d \in F$, we define a probability measure $P_{f_1,...,f_d}$ in $\mathbb{R}^d$ by its characteristic function:

$$\phi_{f_1,...,f_d}(\theta_1, ..., \theta_d) := \exp\left(-\int_E |\sum_{j=1}^{d} \theta_j f_j(x)|^{\alpha(x)} m(dx)\right),$$

$I(f) = \int f(x) M_{\alpha(x)}(dx)$ is called the $\alpha(x)$-multistable integral of f.

- Characteristic function: by definition,

$$\mathbb{E}[e^{i\theta I(f_t)}] = \exp\left(- \int_E |\theta f_t(x)|^{\alpha(x)} m(dx) \right).$$

- Characteristic function: by definition,

$$\mathbb{E}[e^{i\theta I(f_t)}] = \exp\left(-\int_E |\theta f_t(x)|^{\alpha(x)} m(dx)\right).$$

- For the other definition,

$$\mathbb{E}[e^{i\theta I(f_t)}] = \exp\left(-\int_E |\theta f_t(x)|^{\alpha(t)} m(dx)\right).$$

Link between the definitions

We denote by $L(t)$ the Lévy process defined by multistable measures:

$$\mathbb{E}[e^{i\theta L(t)}] = \exp\left(-\int_0^t |\theta|^{\alpha(x)} dx\right).$$

The idea is to obtain the Ferguson-Klass-LePage representation of this process: for $t \in (0,1)$, we have:

$$L(t) = \sum_{i=1}^{\infty} K(\alpha(V_i))\gamma_i \Gamma_i^{-1/\alpha(V_i)} \mathbf{1}_{(V_i \leq t)}.$$

Link between the definitions

We denote by $L(t)$ the Lévy process defined by multistable measures:

$$\mathbb{E}[e^{i\theta L(t)}] = \exp\left(-\int_0^t |\theta|^{\alpha(x)} dx\right).$$

The idea is to obtain the Ferguson-Klass-LePage representation of this process: for $t \in (0,1)$, we have:

$$L(t) = \sum_{i=1}^{\infty} K(\alpha(V_i))\gamma_i \Gamma_i^{-1/\alpha(V_i)} \mathbf{1}_{(V_i \le t)}.$$

L is a pure jump process, with independent increments. It is a Markov process and a semi-martingale.

Link between the definitions

The link is explained by the following decomposition: almost surely, $\forall t \in (0,1)$,

$$\sum_{i=1}^{+\infty} \gamma_i K(\alpha(t)) \Gamma_i^{-\frac{1}{\alpha(t)}} \mathbf{1}_{[0,t]}(V_i) = \sum_{i=1}^{+\infty} \gamma_i K(\alpha(V_i)) \Gamma_i^{-\frac{1}{\alpha(V_i)}} \mathbf{1}_{[0,t]}(V_i) + \varepsilon(t),$$

where $\varepsilon(t) = \int_0^t \sum_{i=1}^{+\infty} \gamma_i g_i'(s) \mathbf{1}_{[0,s[}(V_i) ds$
and $g_i(t) = K(\alpha(t)) \Gamma_i^{-1/\alpha(t)}$.

How to test the multistability

We want to know if the observations come from a stable process or from a multistable one. We want to test if α is varying with time. We consider the statistical test:

$$H_0 : \alpha \text{ is a constant} \qquad \text{vs} \qquad H_1 : \alpha \text{ is varying.}$$

We consider only the case of the Multistable Lévy motion, so $h(t) = \frac{1}{\alpha(t)}$ and $f(t, u, x) = \mathbf{1}_{[0,t]}(x)$.

How to test the multistability

We want to know if the observations come from a stable process or from a multistable one. We want to test if α is varying with time. We consider the statistical test:

$$H_0 : \alpha \text{ is a constant} \qquad \text{vs} \qquad H_1 : \alpha \text{ is varying.}$$

We consider only the case of the Multistable Lévy motion, so $h(t) = \frac{1}{\alpha(t)}$ and $f(t, u, x) = \mathbf{1}_{[0,t]}(x)$.

- Idea : we will test if h is varying.

Estimation of h

We need to estimate the function h, with the observation of one trajectory of Y.

We define the sequence $(Y_{k,N})_{k\in\mathbb{Z},N\in\mathbb{N}}$ by

$$Y_{k,N} = Y(\frac{k+1}{N}) - Y(\frac{k}{N}).$$

Estimation of h

We need to estimate the function h, with the observation of one trajectory of Y.

We define the sequence $(Y_{k,N})_{k\in\mathbb{Z},N\in\mathbb{N}}$ by

$$Y_{k,N} = Y(\frac{k+1}{N}) - Y(\frac{k}{N}).$$

Let $t_0 \in \mathbb{R}$. We introduce an estimate of $h(t_0)$ with

$$\hat{h}_N(t_0) = -\frac{1}{n(N)\log N} \sum_{k=[Nt_0]-\frac{n(N)}{2}}^{[Nt_0]+\frac{n(N)}{2}-1} \log|Y_{k,N}|$$

where $(n(N))_{N\in\mathbb{N}}$ is a sequence taking even integer values.

Estimation of h

Theorem

Under technical conditions on the general kernel $f(t, u, x)$, if

- $\lim_{N \to +\infty} \frac{N}{n(N)} = +\infty$,

then $\forall r > 0$, $\forall t_0$,

$$\lim_{N \to +\infty} \mathbb{E} \left| \hat{h}_N(t_0) - h(t_0) \right|^r = 0.$$

Estimation of h

Theorem

Under technical conditions on the general kernel $f(t,u,x)$, if

- $\lim_{N\to+\infty} \frac{N}{n(N)} = +\infty$,

then $\forall r > 0$, $\forall t_0$,

$$\lim_{N\to+\infty} \mathbb{E}\left|\hat{h}_N(t_0) - h(t_0)\right|^r = 0.$$

Theorem

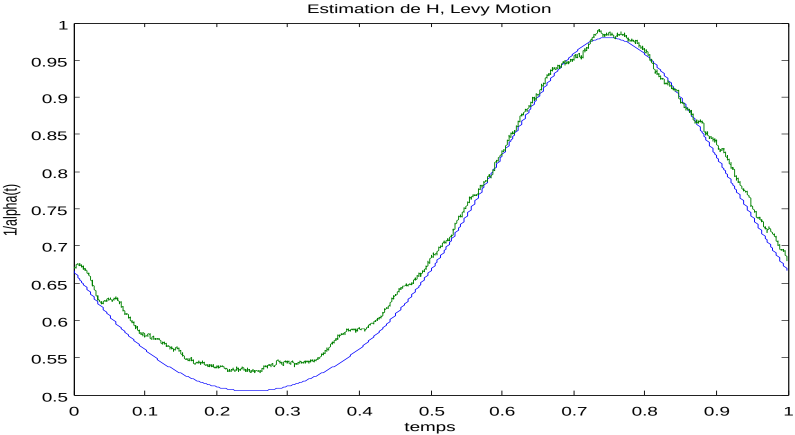
For a Lévy process ($f(t,u,x) = \mathbf{1}_{[0,t]}(x)$)*, when*
$\lim_{N\to+\infty} n(N) = +\infty$ *and* $\lim_{N\to+\infty} \frac{n(N)}{N} = 0$, $\forall t_0 \in (0,1)$,

$$\sqrt{n(N)}\left(\log N(\hat{h}_N(t_0) - h(t_0)) + \mathbb{E}[\ln|S_{\alpha(t_0)}|]\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mathbb{E}[Y^2])$$

where $Y \sim \ln|S_{\alpha(t_0)}| - \mathbb{E}[\ln|S_{\alpha(t_0)}|]$.

Estimation of h, case of the Lévy motion

$\alpha(t) = 1,5 + 0,48\sin(2\pi t)$, $N = 20000$, $n(N) = 500$.



Test of the multistability

The hypotheses are :

$$\begin{aligned} & H_0: \quad \forall (t_1, t_2) \in (0,1)^2, \alpha(t_1) = \alpha(t_2). \\ \text{vs} & \\ & H_1: \quad \exists (t_1, t_2) \in (0,1)^2, \alpha(t_1) \neq \alpha(t_2). \end{aligned}$$

We use the following statistic, for one t_0:

$$I_N = \int_0^1 |\hat{h}_N(t) - \hat{h}_N(t_0)|^2 dt.$$

Under (H_0), $E[I_N] \to 0$.

Under (H_1), $E[I_N] \to \int_0^1 |h(t) - h(t_0)|^2 dt > 0$.

Results under H_0

Let $\sigma^2 = \mathbb{E}[Y^2] = Var(\ln|S_{\alpha(t_0)}| - \mathbb{E}[\ln|S_{\alpha(t_0)}|])$.
We have the following convergence:

$$n(N)(\log N)^2 \int_0^1 |\hat{h}_N(t) - \hat{h}_N(t_0)|^2 dt \xrightarrow{d} \sigma^2(1 + \chi^2(1)).$$

Proof

Let $\mu = -\mathbb{E}[\ln |S_{\alpha(t_0)}|]$.
We have

$$\sqrt{n(N)}\left(\log N(\hat{h}_N(t_0) - h(t_0)) - \mu\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

so with
$Z_N(t) = \sqrt{n(N)}\left(\log N(\hat{h}_N(t) - h(t)) - \mu(t)\right) \to \mathcal{N}(0, \sigma^2)$
and $\Delta h(t) = h(t) - h(t_0)$, $\Delta\mu(t) = \mu(t) - \mu(t_0)$,

$$n(N)(\log N)^2 \int_0^1 |\hat{h}_N(t) - \hat{h}_N(t_0)|^2 dt$$
$$= \int_0^1 \left| Z_N(t) - Z_N(t_0) + \sqrt{n(N)}((\log N)\Delta h(t) + \Delta\mu(t)\ )\right|^2 dt.$$

Proof

Under (H_0), we expand the square,

$$\int_0^1 |Z_N(t) - Z_N(t_0)|^2 \, dt = \int_0^1 |Z_N(t)|^2 dt - 2Z_N(t_0) \int_0^1 Z_N(t) dt + |Z_N(t_0)|^2.$$

We are now able to control each term:

- $\int_0^1 |Z_N(t)|^2 dt \xrightarrow{\mathbb{P}} \sigma^2$,
- $\int_0^1 Z_N(t) dt \xrightarrow{\mathbb{P}} 0$,
- $|Z_N(t_0)|^2 \to \sigma^2 \chi^2(1)$.

Finally,

$$n(N)(\log N)^2 \int_0^1 |\hat{h}_N(t) - \hat{h}_N(t_0)|^2 dt \xrightarrow{d} \sigma^2 (1 + \chi^2(1)).$$

Results under H_1

The rejection region of the test is

$$R_c = \{ \frac{n(N)(\log N)^2}{\hat{\sigma^2}} \int_0^1 |\hat{h}_N(t) - \hat{h}_N(t_0)|^2 dt > \ q_\beta \}$$

with q_β the quantile of the distribution $1 + \chi^2(1)$.

Results under H_1

The rejection region of the test is

$$R_c = \{ \frac{n(N)(\log N)^2}{\hat{\sigma^2}} \int_0^1 |\hat{h}_N(t) - \hat{h}_N(t_0)|^2 dt > \ q_\beta \}$$

with q_β the quantile of the distribution $1 + \chi^2(1)$.

- Statistical power for a Lévy process (multistable measures)

Theorem

$\forall (\varepsilon_N)_N$ such that $\lim_{N\to+\infty} \varepsilon_N = 0$ and $\lim_{N\to+\infty} \varepsilon_N \log N = +\infty$, if $\alpha_* = \min \alpha(u)$ and $\alpha^* = \max \alpha(u)$,

$$\liminf_{N\to+\infty} \frac{-\log \mathbb{P}_1\left(\overline{R_c}\right)}{N\varepsilon_N \log N} \geq 2(\frac{1}{\alpha_*} - \frac{1}{\alpha^*}).$$

Results under H_1

- Statistical power for a Lévy process

$$\liminf_{N \to +\infty} \frac{-\log \mathbb{P}_1\left(\overline{R_c}\right)}{\log \log N} = +\infty.$$