A simple estimator of a discretely observed stable process

Workshop “SAPS VI” at Université du Maine,

March 23, 2007

Hiroki Masuda

Kyushu University, Japan

Introduction: the model in question and aim

- Let $(P_\theta)_{\theta\in\Theta}$ denote the family of distributions on $\mathcal{D}(\mathbb{R}_+;\mathbb{R})$ of X associated with $\theta\in\Theta$, where:
  - **The underlying process** $X=(X_t)_{t\in\mathbb{R}_+}$:
    a non-Gaussian stable Lévy process on $\mathbb{R}$ s.t.

    $$\mathscr{L}(X_t-\gamma t|P_\theta)=S_\alpha(t^{1/\alpha}\sigma),\quad t\in\mathbb{R}_+,$$

    where $S_\alpha(\sigma)$ has the characteristic function

    $$u\mapsto\exp\{-(\sigma|u|)^\alpha\},\quad u\in\mathbb{R}.$$

  - **The dominating parameter** $\theta:=(\alpha,\sigma,\gamma)^\top\in\Theta\subset\mathbb{R}^3$:
    Θ is a bounded convex domain s.t.

    $$\Theta^-\subset\{(\alpha,\gamma,\sigma):\alpha\in(0,2),\sigma>0,\gamma\in\mathbb{R}\}.$$

Introduction: the model in question and aim

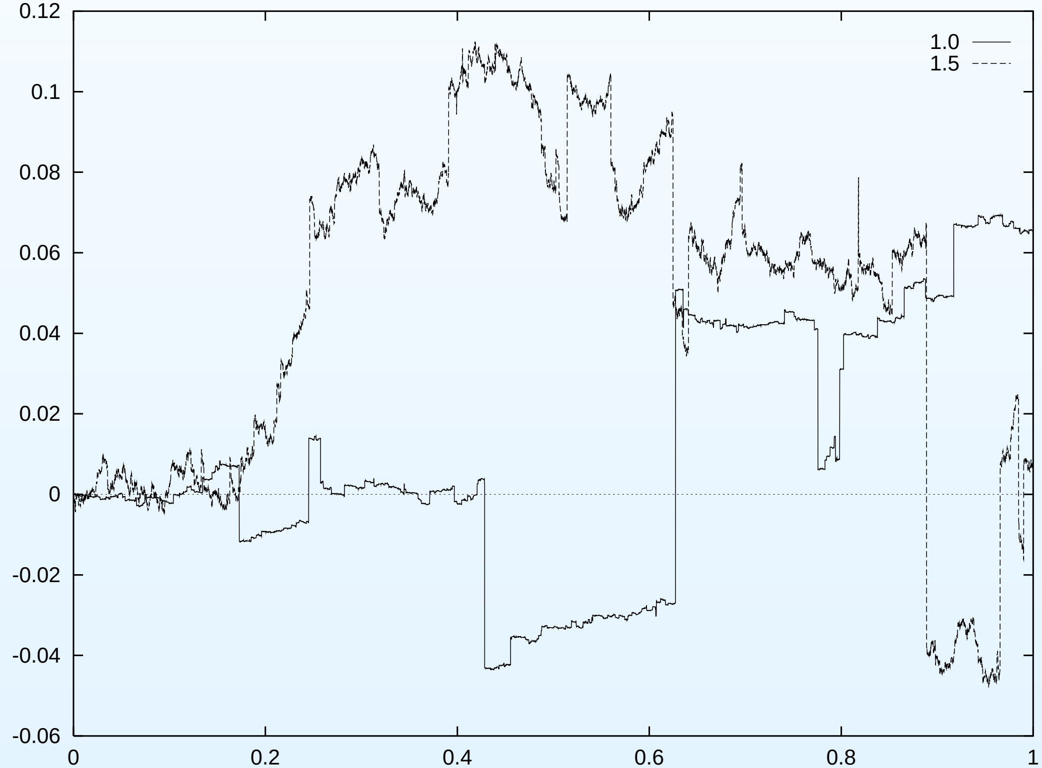
- $(h_n)_{n\in\mathbb{N}} \subset (0,\infty)$ is a non-random bounded sequence s.t. $h_n \to 0$. We write $T_n = nh_n$, the terminal observation time.
- **Aim**:
  Having a discrete observation

$$(X_{h_n}, X_{2h_n}, \dots, X_{nh_n})$$

  drawn under P_θ, we want to estimate $\theta := (\alpha, \sigma, \gamma)$, in the main, when either $\liminf_{n\to\infty} T_n > 0$ or $\lim_{n\to\infty} T_n = \infty$.

Introduction: the model in question and aim

Sample paths of stable Lévy processes for $\alpha = 1.0$ and 1.5 over $[0, 1]$, both having location zero and symmetric Lévy density ($\sigma = 0.3$):



For $P_\theta^T := \mathscr{L}\{(X_t)_{t\in[0,T]}\}$, $T > 0$, we know $P_\theta^T \approx P_{\theta'}^T$ iff $\theta = \theta'$.

Introductory remark: the LAN property

- Write $\Delta_i^n X = X_{ih_n} - X_{(i-1)h_n}$ for $i \le n$.
- The triangular array

$$Y_{ni}(\theta) := \sigma^{-1} h_n^{-1/\alpha} (\Delta_i^n X - \gamma h_n),$$

$i \le n$, forms an i.i.d. sequence for each $n \in \mathbb{N}$, with common law $S_\alpha(1)$.
(i.e., the selfsimilarity)

- The log-likelihood function is written down as

$$\ell_n(\theta) = \sum_{i=1}^{n} \log \left\{ \sigma^{-1} h_n^{-1/\alpha} \phi_\alpha(Y_{ni}(\theta)) \right\},$$

where $y \mapsto \phi_\alpha(y)$ denotes the density of $S_\alpha(1)$.

Introductory remarks: the LAN property

- Suppose $\sqrt{n}h_n^{1-1/\alpha} \to \infty$. (implied by $\liminf_{n\to\infty} T_n > 0$)
- Then we have $(A_n(\alpha), I(\alpha,\sigma))$-LAN property with

$$A_n(\alpha) := \mathrm{diag}\Big\{ \underbrace{\sqrt{n}\log(1/h_n)}_{\alpha}, \underbrace{\sqrt{n}}_{\sigma}, \underbrace{\sqrt{n}h_n^{1-1/\alpha}}_{\gamma} \Big\},$$

$$I(\alpha,\sigma) := \begin{pmatrix} H_\alpha/\alpha^4 & H_\alpha/(\sigma\alpha^2) & 0 \\ H_\alpha/(\sigma\alpha^2) & H_\alpha/\sigma^2 & 0 \\ 0 & 0 & M_\alpha/\sigma^2 \end{pmatrix},$$

$$H_\alpha := \int \frac{\{\phi_\alpha(y) + y\partial\phi_\alpha(y)\}^2}{\phi_\alpha(y)}dy, \quad M_\alpha := \int \frac{\{\partial\phi_\alpha(y)\}^2}{\phi_\alpha(y)}dy.$$

  - Both of H_α and M_α are positive and finite.
  - $\det I(\alpha,\sigma) \equiv 0$.

Introductory remark: the LAN property

Some remarks on the LAN property:

- Any matrix norming of $\partial_\theta \ell_n(\theta)$ may lead to a singular limit information;
- Concerning the optimal rate,
  - $\sqrt{n}\log(1/h_n)$ of α indicates that, high frequency of data may speed up accuracy of estimation of the tail index α.
  - $\sqrt{n}$ of σ coincides with the case of $\alpha = 2$, namely Wiener case;
  - $\sqrt{n}h_n^{1-1/\alpha} = \sqrt{T_n}h_n^{1/2-1/\alpha}$ of γ says that we may consistently estimate it even when $T_n = T > 0$ fixed;

Outline

Construction of simple estimators other than MLE:

- Single estimation of γ, leaving (α, σ) unknown;
- Estimation of (α, σ) with nonsingular limit covariance, when the drift γ is known;
- Joint rate-consistency of (α, σ, γ) for $\alpha \in (1, 2)$.

The estimator we shall consider here are very simple, while the rate-optimality and efficiency are sacrificed.

1. Single estimation of γ with (α, σ) unknown

- Concerning the sample mean $\bar{X}_n := X_{T_n}/T_n$, we have

$$\mathscr{L}\{T_n^{1-1/\alpha}(\bar{X}_n - \gamma) \big| P_\theta\} = S_\alpha(\sigma)$$

for each $n \in \mathbb{N}$, on account of the selfsimilarity of X.

- $\bar{X}_n$ is consistent only when

$$T_n^{1-1/\alpha} \longrightarrow \infty, \quad \text{i.e., } \alpha \in (1, 2),$$

hence useless when $T_n = O(1)$.
Moreover, it has infinite variance, so rather unstable.

1. Single estimation of γ with (α, σ) unknown

- Sample median based estimator may be expected to be robust; *what about its convergence rate*?
- Since $(Y_{ni}(\theta))_{i\le n}$ is a $S_\alpha(1)$-i.i.d. sequence with median 0, using the sample median m_n of $(\Delta_i X)_{i\le n}$ given by

$$m_n = \begin{cases} (\Delta_k^n X + \Delta_{k+1}^n X)/2 & \text{if } n = 2k, \\ \Delta_{k+1}^n X & \text{if } n = 2k+1, \end{cases}$$

we can prove

$$\sqrt{n}\,\underbrace{\{h_n^{-1/\alpha}(m_n - \gamma h_n)\}}_{\text{sample median of } (Y_{ni}(\theta))_{i\le n}} \Rightarrow^{P_\theta} N_1\left(0, \left\{\frac{\sigma\pi}{2\Gamma(1+1/\alpha)}\right\}^2\right)$$

by the standard theory of order statistics.

1. Single estimation of γ with (α, σ) unknown

We obtain the estimator $\hat{\gamma}_n = h_n^{-1} m_n$.

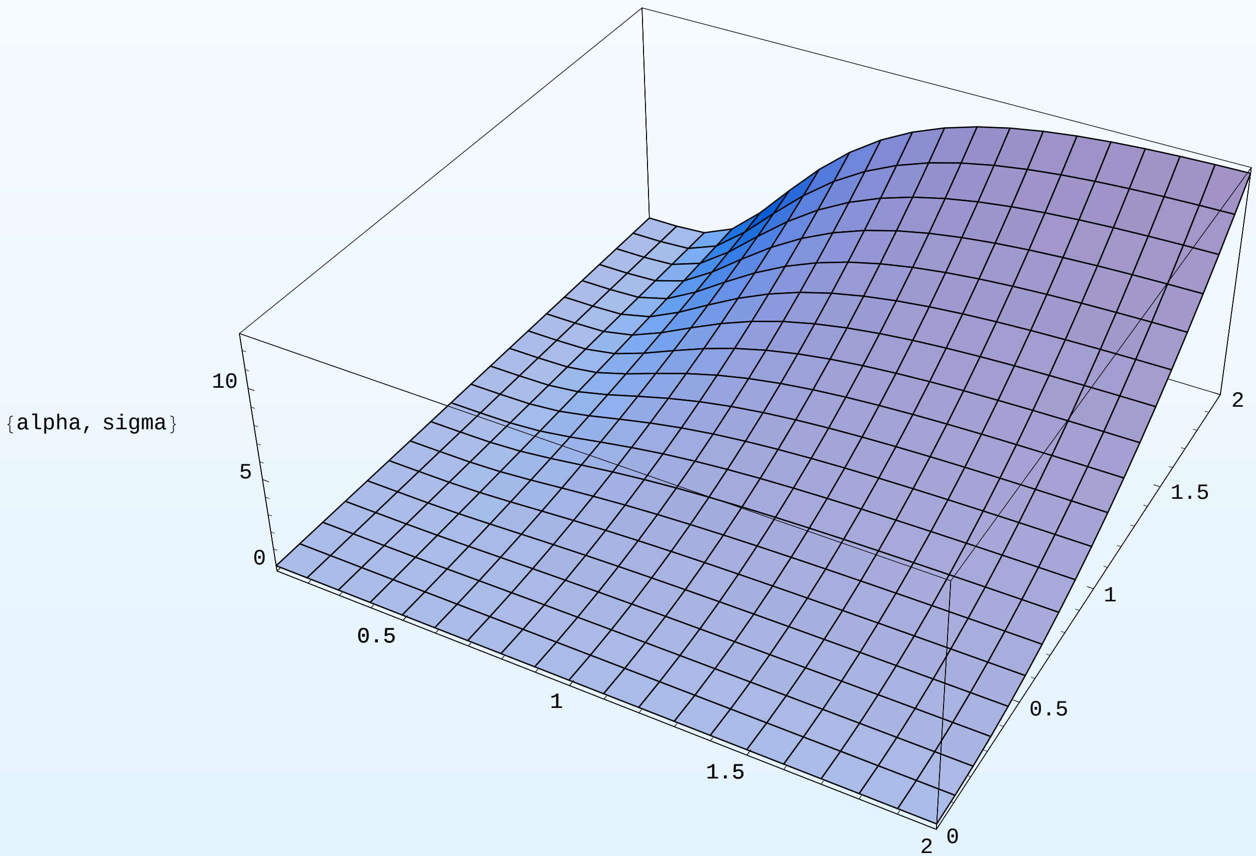
Theorem 1. Fix any $\theta \in \Theta$ and suppose $\sqrt{n} h_n^{1-1/\alpha} \to \infty$. Then $\hat{\gamma}_n$ is a P_θ-unbiased estimator of γ and satisfies

$$\sqrt{T_n} h_n^{1/2-1/\alpha} (\hat{\gamma}_n - \gamma) \Rightarrow^{P_\theta} N_1\left(0, \left\{\frac{\sigma\pi}{2\Gamma(1+1/\alpha)}\right\}^2\right).$$

- We do not need $nh_n \to \infty$, but the performances of $\hat{\gamma}_n$ may become better when $T_n \to \infty$.
- The optimal rate $\sqrt{n} h_n^{1-1/\alpha}$ may become rather slow when, e.g., $h_n = 1/n$ and α is close to 2 (then the rate is $n^{1/\alpha - 1/2}$).

1. Single estimation of γ with (α, σ) unknown

A plot of the asymptotic variance of $\hat{\gamma}_n$:
$(\alpha, \sigma) \mapsto \{\sigma\pi/(2\Gamma(1 + 1/\alpha))\}^2$ over $(0, 2) \times (0, 2)$:

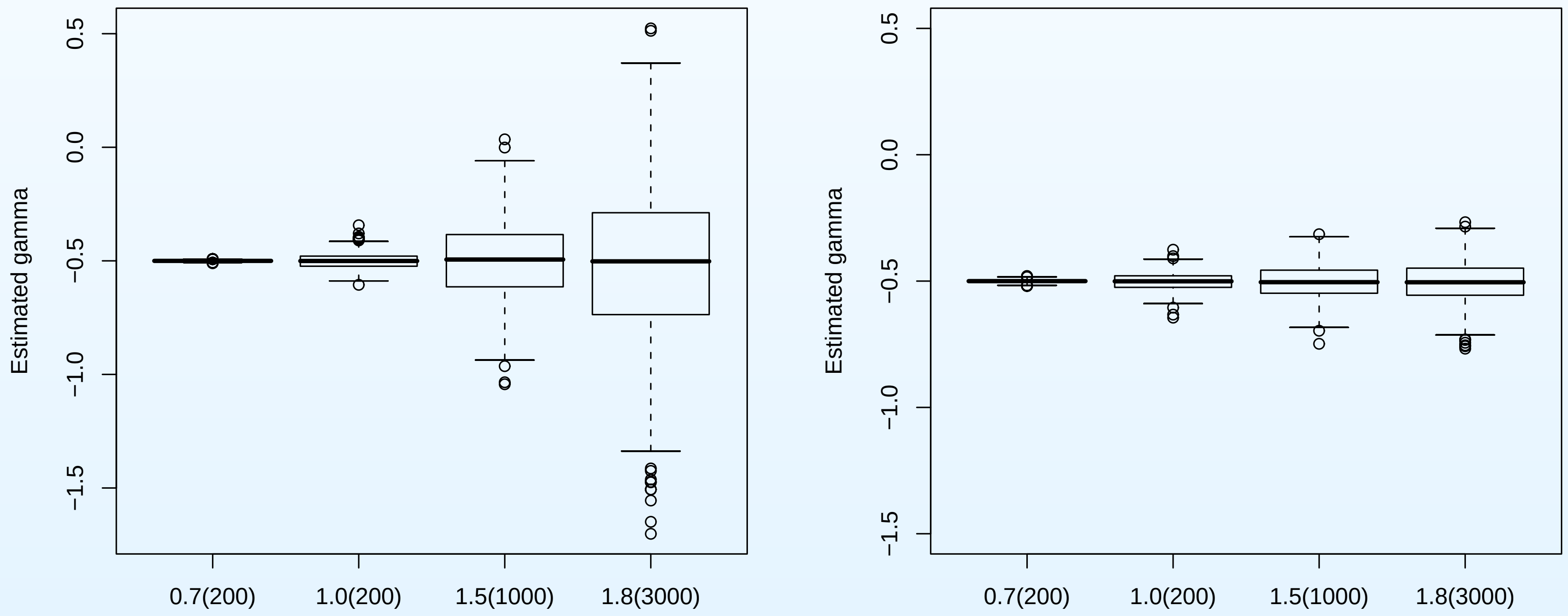


Simulation 1

(True $\gamma = -0.5$)		Case of $h_n = n^{-1}$		Case of $h_n = n^{-3/5}$	
True α	n	T_n	$\hat{\gamma}_n$-mean (s.d.)	T_n	$\hat{\gamma}_n$-mean (s.d.)
0.7	50	1	-0.5003 (0.0112)	4.7818	-0.5003 (0.0214)
	100	1	-0.4999 (0.0057)	6.3096	-0.5000 (0.0120)
	200	1	-0.5002 (0.0028)	8.3255	-0.5003 (0.0066)
1.0	50	1	-0.5036 (0.0675)	4.7818	-0.4998 (0.0716)
	100	1	-0.4982 (0.0470)	6.3096	-0.4989 (0.0462)
	200	1	-0.5009 (0.0341)	8.3255	-0.5017 (0.0336)
1.5	100	1	-0.4912 (0.2375)	6.3096	-0.5087 (0.1335)
	300	1	-0.4942 (0.2069)	9.7915	-0.5005 (0.0964)
	500	1	-0.4991 (0.1862)	12.0112	-0.4987 (0.0815)
	1000	1	-0.4995 (0.1659)	15.8489	-0.5034 (0.0660)
1.8	1000	1	-0.4975 (0.3582)	15.8489	-0.4993 (0.1085)
	1500	1	-0.5106 (0.3565)	18.6396	-0.4987 (0.0962)
	2000	1	-0.4976 (0.3383)	20.9128	-0.4988 (0.0862)

Simulation 1

Boxplots of simulated estimates of γ for $h_n = n^{-1}$ (left panel) and $h_n = n^{-3/5}$ (right panel): in each case, plots for $(\alpha, n) = (0.7, 200), (1.0, 200), (1.5, 1000)$ and $(1.8, 3000)$. We set $(\sigma, \gamma) = (0.3, -0.5)$, and $1,000$ independent trials were used.



2. Estimation of (α, σ) with nonsingular variance

Supposing γ is known, it is possible to perform the method of moments for estimating α and σ. From now on n is even.

- The mean and variance of $\log|S_\alpha(\sigma)|$ are given by $\mathfrak{C}(\frac{1}{\alpha} - 1) + \log \sigma$ and $\frac{\pi^2}{6}(\frac{1}{\alpha^2} + \frac{1}{2})$, respectively. ($\mathfrak{C} \fallingdotseq 0.5772$)
- Using the above, we get simple estimators of (α, σ):

$$\hat{\alpha}_n(\gamma) := \left\{ \frac{6}{\pi^2 n} \sum_{i=1}^{n} \left(\log|\Delta_i^n X - \gamma h_n| - \frac{1}{n} \sum_{i=1}^{n} \log|\Delta_i^n X - \gamma h_n| \right)^2 - \frac{1}{2} \right\}^{-1/2},$$

$$\hat{\sigma}_n(\gamma) := \exp\left\{ \frac{1}{\hat{\alpha}_n} \log(1/h_n) + \frac{1}{n} \sum_{i=1}^{n} \log|\Delta_i^n X - \gamma h_n| - \mathfrak{C}\left(\frac{1}{\hat{\alpha}_n} - 1 \right) \right\}.$$

2. Estimation of (α, σ) with nonsingular variance

Write $\theta' = (\alpha, \sigma)^\top$, with $\Theta' \subset \mathbb{R}^2$ denoting the corresponding parameter space.

Theorem 2. Fix any $\theta \in \Theta$, and suppose $\sqrt{n} h_n^{1-1/\alpha} \to \infty$ and $\gamma \in \mathbb{R}$ is known. Then we have

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}_n(\gamma) - \alpha \\ \hat{\sigma}_n(\gamma) - \sigma \end{pmatrix} \Rightarrow^{P_{\theta'}} N_2(0, \Sigma(\theta')),$$

where $\Sigma(\theta') \in \mathbb{R}^{2 \otimes 2}$ is positive definite.

- $\hat{\sigma}_n(\gamma)$ is rate-efficient, while $\hat{\alpha}_n(\gamma)$ is not.
- Again we do not need $T_n \to \infty$.

Simulation 2

(a) Case of $h_n = n^{-1}$: $\gamma = 0$ is known, and $\sigma = 0.3$.

True α	n	T_n	$\hat{\alpha}_n$-mean (s.d.)	$\hat{\sigma}_n$-mean (s.d.)
0.7	500	1	0.7050 (0.0390)	0.3231 (0.1671)
	1000	1	0.7026 (0.0281)	0.3168 (0.1306)
1.0	500	1	1.0092 (0.0704)	0.3158 (0.1280)
	1000	1	1.0023 (0.0461)	0.3127 (0.0964)
1.3	500	1	1.3195 (0.1168)	0.3113 (0.1200)
	1000	1	1.3080 (0.0763)	0.3087 (0.0892)
1.5	500	1	1.5451 (0.1812)	0.3041 (0.1216)
	1000	1	1.5165 (0.1187)	0.3078 (0.0998)
	1500	1	1.5129 (0.0954)	0.3037 (0.0790)
1.8	500	1	1.8724 (0.3044)	0.3118 (0.1330)
	1000	1	1.8260 (0.1823)	0.3135 (0.1046)
	1500	1	1.8120 (0.1528)	0.3075 (0.0883)

Simulation 2

(b) Case of $h_n = n^{-3/5}$: $\gamma = 0$ is known, and $\sigma = 0.3$.

True α	n	T_n	$\hat{\alpha}_n$-mean (s.d.)	$\hat{\sigma}_n$-mean (s.d.)
0.7	500	12.0112	0.7049 (0.0391)	0.3055 (0.0864)
	1000	15.8489	0.7022 (0.0266)	0.3043 (0.0678)
1.0	500	12.0112	1.0081 (0.0736)	0.3040 (0.0694)
	1000	15.8489	1.0057 (0.0481)	0.3012 (0.0530)
1.3	500	12.0112	1.3145 (0.1147)	0.3027 (0.0593)
	1000	15.8489	1.3120 (0.0845)	0.3004 (0.0507)
1.5	500	12.0112	1.5460 (0.1847)	0.2954 (0.0637)
	1000	15.8489	1.5104 (0.1149)	0.3027 (0.0504)
	1500	18.6396	1.5146 (0.0937)	0.2990 (0.0441)
1.8	500	12.0112	1.8613 (0.2927)	0.3003 (0.0658)
	1000	15.8489	1.8257 (0.1951)	0.3027 (0.0539)
	1500	18.6396	1.8171 (0.1548)	0.3012 (0.0474)

The performances of $\hat{\alpha}_n(\gamma)$ are similar to the case (a), while that of $\hat{\sigma}_n(\gamma)$ is better.

3. Joint rate-consistency when $\alpha \in (1, 2)$

- The optimal rate $\sqrt{T_n} h_n^{1/2-1/\alpha}$ of γ depends on the true α.
- We supposed the true γ is known when defining $(\hat{\alpha}_n(\gamma), \hat{\sigma}_n(\gamma))$.
- Does it hold true that

$$\begin{pmatrix} \sqrt{n}(\hat{\alpha}_n(\hat{\gamma}_n) - \alpha) \\ \sqrt{n}(\hat{\sigma}_n(\hat{\gamma}_n) - \sigma) \\ \sqrt{n} h_n^{1-1/\alpha}(\hat{\gamma_n} - \gamma) \end{pmatrix} \Rightarrow^{P_\theta} \text{Normal}, \quad \text{or}, \; = O_{P_\theta}(1)??$$

...Negative, maybe.

3. Joint rate-consistency when $\alpha \in (1,2)$

Construction of estimators: *neglecting possibly nonzero γ*.

- We continue to use $\hat{\gamma}_n$ for estimating γ.
- At the same time, for estimating (α, γ) we consider

$$\hat{\alpha}_n := \hat{\alpha}_n(0) = \left\{ \frac{6}{\pi^2 n} \sum_{i=1}^{n} \left(\log |\Delta_i^n X| - \frac{1}{n} \sum_{i=1}^{n} \log |\Delta_i^n X| \right)^2 - \frac{1}{2} \right\}^{-1/2},$$

$$\hat{\sigma}_n := \hat{\sigma}_n(0) = \exp \left\{ \frac{1}{\hat{\alpha}_n} \log(1/h_n) + \frac{1}{n} \sum_{i=1}^{n} \log |\Delta_i^n X| - \mathfrak{C} \left(\frac{1}{\hat{\alpha}_n} - 1 \right) \right\}.$$

- Then we set $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\sigma}_n, \hat{\gamma}_n)$.

3. Joint rate-consistency when $\alpha \in (1, 2)$

$$D_n(\alpha) := \mathrm{diag}\left\{ h_n^{-(1-1/\alpha)/2}, \frac{h_n^{-(1-1/\alpha)/2}}{\log(1/h_n)}, \sqrt{n} h_n^{1-1/\hat{\alpha}_n} \right\}.$$

Theorem 3. Fix any $\theta \in \Theta$ and suppose $\sqrt{n} h_n^{1-1/\alpha} \to \infty$. Then $D_n(\hat{\alpha}_n)(\hat{\theta}_n - \theta)$ is bounded in P_θ-probability:

$$h_n^{-(1-1/\hat{\alpha}_n)/2}(\hat{\alpha}_n - \alpha) = O_{P_\theta}(1),$$

$$\frac{h_n^{-(1-1/\hat{\alpha}_n)/2}}{\log(1/h_n)}(\hat{\sigma}_n - \sigma) = O_{P_\theta}(1),$$

$$\sqrt{n} h_n^{1-1/\hat{\alpha}_n}(\hat{\gamma}_n - \gamma) \Rightarrow^{P_\theta} N_1\left(0, \left\{ \frac{\sigma \pi}{2\Gamma(1+1/\alpha)} \right\}^2 \right).$$

Simulation 3

(a) Case of $h_n = n^{-1}$: $(\sigma, \gamma) = (0.3, -0.5)$.

True α	n	T_n	$\hat{\alpha}_n$-mean (s.d.)	$\hat{\sigma}_n$-mean (s.d.)	$\hat{\gamma}_n$-mean (s.d.)
1.3	1000	1	1.3331 (0.0894)	0.2933 (0.0880)	-0.4991 (0.0817)
	1500	1	1.3322 (0.0756)	0.2860 (0.0743)	-0.5008 (0.0704)
	2000	1	1.3244 (0.0664)	0.2911 (0.0723)	-0.5014 (0.0660)
	3000	1	1.3178 (0.0504)	0.2916 (0.0587)	-0.4991 (0.0596)
1.5	1000	1	1.5173 (0.1170)	0.3093 (0.0949)	-0.5079 (0.1677)
	1500	1	1.5096 (0.0929)	0.3088 (0.0982)	-0.4953 (0.1475)
	2000	1	1.5105 (0.0827)	0.3045 (0.0728)	-0.5029 (0.1453)
	3000	1	1.5065 (0.0663)	0.3037 (0.0635)	-0.5015 (0.1380)
1.8	1000	1	1.8350 (0.1878)	0.3065 (0.1022)	-0.5007 (0.3598)
	1500	1	1.8196 (0.1511)	0.3047 (0.0903)	-0.4995 (0.3532)
	2000	1	1.8142 (0.1316)	0.3076 (0.0827)	-0.5072 (0.3464)
	3000	1	1.8150 (0.1049)	0.3010 (0.0667)	-0.4943 (0.3392)

Simulation 3

(b) Case of $h_n = n^{-3/5}$: $(\sigma, \gamma) = (0.3, -0.5)$.

True α	n	T_n	$\hat{\alpha}_n$-mean (s.d.)	$\hat{\sigma}_n$-mean (s.d.)	$\hat{\gamma}_n$-mean (s.d.)
1.3	1000	15.8489	1.4067 (0.1451)	0.2845 (0.0503)	-0.5003 (0.0425)
	1500	18.6396	1.3925 (0.1191)	0.2820 (0.0438)	-0.5003 (0.0368)
	2000	20.9127	1.3890 (0.1119)	0.2793 (0.0431)	-0.5017 (0.0333)
	3000	24.5951	1.3758 (0.0920)	0.2792 (0.0377)	-0.5001 (0.0282)
1.5	1000	15.8489	1.5437 (0.1271)	0.3027 (0.0501)	-0.4996 (0.0659)
	1500	18.6396	1.5375 (0.1065)	0.3012 (0.0460)	-0.5008 (0.0572)
	2000	20.9127	1.5353 (0.0915)	0.2985 (0.0402)	-0.5007 (0.0538)
	3000	24.5951	1.5301 (0.0766)	0.2974 (0.0377)	-0.5009 (0.0468)
1.8	1000	15.8489	1.8410 (0.1881)	0.3025 (0.0533)	-0.5008 (0.1043)
	1500	18.6396	1.8266 (0.1709)	0.3035 (0.0489)	-0.5032 (0.0988)
	2000	20.9127	1.8221 (0.1348)	0.3020 (0.0430)	-0.5011 (0.0886)
	3000	24.5951	1.8165 (0.1047)	0.3011 (0.0371)	-0.5012 (0.0791)

Concluding remarks

We have attempted very simple procedure for estimating stable Lévy processes based on high-frequency discrete-time sampling.

- We have utilized the selfsimilarity of stable Lévy processes; no other Lévy process can exhibit the selfsimilarity.
- There are many other possibilities of estimating θ (quantile based, regression based, etc.), while we do not know specific "asymptotic optimality" of estimators.
- A worth noting point is that all the parameter can be consistently estimated even for bounded time domain data.