Stable central limit theorem and its convergence rate

Lihu Xu

University of Macau

The 18th Workshop on Markov Processes and Related Topics

July 29-August 2, 2023

Outline

- Lindeberg principle for stable CLT
- A general approximation framework
- Stable CLT in TV distance

Contents

A simple version of stable CLT

Theorem

Let $\xi_1, ..., \xi_n, ...$ *be i.i.d. with Pareto distribution, i.e.* ξ_1 *has a density function* $p(x) = \frac{K}{|x|^{\alpha+1}} 1_{\{|x| \geq A\}}$ *for some* $A > 0$ *and* $K > 0$ *with* $\alpha \in (0, 2)$*. Let* $S_n = \xi_1 + ... + \xi_n$*. Then*

$$\frac{S_n}{n^{1/\alpha}} \Rightarrow Z,$$

where Z *has a symmetric stable distribution with a characteristic function* $\exp(-c|\lambda|^{\alpha})$*.*

The general stable CLT can be found in the book 'Probability:Theory and Examples' (Durrett), see Theorem 3.7.2.

Lindeberg principle for stable law

Let $X_1, X_2, \ldots$ be i.i.d. r.v. with the following Pareto law:

$$p(x) = \frac{\alpha}{2}|x|^{-(1+\alpha)}\mathbf{1}_{[1,\infty)}(|x|).$$

Denote

$$S_n = n^{-\frac{1}{\alpha}}\left(X_1 + \ldots + X_n\right),$$

we will use Lindeberg principle to prove

$$S_n \Rightarrow Z$$

where Z has a stable distribution with the characteristic function $e^{-|\lambda|^\alpha}$.

Lindeberg principle for stable law

Let $Z_1, Z_2, ...,$ be i.i.d. r.v. with the characteristic function $e^{-|\lambda|^\alpha}$. Denote

$$S_n^{(0)} = n^{-\frac{1}{\alpha}} \left(X_1 + X_2 + ... + X_n\right),$$

$$S_n^{(1)} = n^{-\frac{1}{\alpha}} \left(Z_1 + X_2 + ... + X_n\right),$$

$$..., ...,$$

$$S_n^{(n)} = n^{-\frac{1}{\alpha}} \left(Z_1 + Z_2 + ... + Z_n\right).$$

We have

$$S_n^{(n)} =^d Z.$$

Lindeberg principle for stable law

Set

$$Y_i = n^{-\frac{1}{\alpha}} \left(Z_1 + \cdots + Z_{i-1} + X_{i+1} + \cdots + X_n \right), \quad 1 \leqslant i \leqslant n$$

we have

$$S^{(i)} = Y_i + n^{-\frac{1}{\alpha}} Z_i, \quad S^{(i-1)} = Y_i + n^{-\frac{1}{\alpha}} X_i.$$

Let $f \in C^3$,

$$\begin{aligned} & \mathbb{E}\left[f\left(S_n\right)\right]-\mathbb{E}[f(Z)] \\ = & \sum_{i=1}^n \mathbb{E}\left[f\left(S_n^{(i-1)}\right)-f\left(S_n^{(i)}\right)\right] \\ = & \sum_{i=1}^n\left\{\mathbb{E}\left[f\left(S_n^{(i-1)}\right)-f\left(Y_i\right)\right]-\mathbb{E}\left[f\left(S_n^{(i)}\right)-f\left(Y_i\right)\right]\right\} . \end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}\left[f\left(S_n^{(i-1)}\right) - f\left(Y_i\right)\right] \\
&= \mathbb{E}\left[f\left(Y_i + \frac{X_i}{n^{\frac{1}{\alpha}}}\right) - f\left(Y_i\right)\right] \\
&= \mathbb{E}\left[\frac{\alpha}{2}\int_{|x|\geqslant 1} \frac{f\left(Y_i + \frac{x}{n^{\frac{1}{\alpha}}}\right) - f\left(Y_i\right)}{|x|^{\alpha+1}}\, \mathrm{d}x\right] \\
&= n^{-1}\mathbb{E}\left[\frac{d_\alpha}{2}\int_{|y|\geqslant n^{-\frac{1}{\alpha}}} \frac{f\left(Y_i + y\right) - f\left(Y_i\right)}{|y|^{\alpha+1}}\, \mathrm{d}y\right] \\
&= n^{-1}\mathbb{E}\left[\Delta^{\frac{\alpha}{2}} f\left(Y_i\right) - \frac{d_\alpha}{2}\int_{-n^{-\frac{1}{\alpha}}}^{n^{-\frac{1}{\alpha}}} \frac{f\left(Y_i + y\right) - f\left(Y_i\right)}{|y|^{\alpha+1}}\, \mathrm{d}y\right],
\end{aligned}
$$

By Itô's formula, we have

$$
\begin{aligned}
&\mathbb{E}\left[f\left(S_n^{(i)}\right)-f\left(Y_i\right)\right] \\
&=\int_0^1 \mathbb{E}\left[\Delta_{Z_s}^{\frac{\alpha}{2}} f\left(Y_i+\frac{Z_s}{n^{\frac{1}{\alpha}}}\right)\right] \mathrm{d}s \\
&=n^{-1}\int_0^1 \mathbb{E}\left[\Delta^{\frac{\alpha}{2}} f\left(Y_i+\frac{Z_s}{n^{\frac{1}{\alpha}}}\right)\right] \mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&\left|\mathbb{E}\left[f\left(S_n^{(i-1)}\right)-f\left(Y_i\right)\right]-\mathbb{E}\left[f\left(S_n^{(i)}\right)-f\left(Y_i\right)\right]\right| \\
\leq\; & n^{-1}\left|\int_0^1 \mathbb{E}\left[\Delta^{\frac{\alpha}{2}} f\left(Y_i+\frac{Z_s}{n^{\frac{1}{\alpha}}}\right)-\Delta^{\frac{\alpha}{2}} f\left(Y_i\right)\right] \mathrm{d}s\right| \\
& +\frac{d_\alpha}{2n}\left|\int_{-n^{-\frac{1}{\alpha}}}^{n^{-\frac{1}{\alpha}}} \frac{f\left(Y_i+y\right)-f\left(Y_i\right)}{|y|^{\alpha+1}} \mathrm{d}y\right| \\
\leq\; & C n^{-2/\alpha}
\end{aligned}
$$

Hence, we have

$$
|\mathbb{E} f(S_n)-\mathbb{E} f(Z)| \leq C n^{-\frac{2-\alpha}{\alpha}} .
$$

Contents

A key observation

- We view
  - S_n as the end point of the Markov chain $\{S_k : 0 \le k \le n\}$,
  - Z as the end point of the stable process $\{Z_t : 0 \le t \le 1\}$.
- We can use the Markov semigroup theory to compare

$$
\begin{aligned}
\mathbb{E}f(S_n) - \mathbb{E}f(Z_1) :&= Q_{0,n}f - P_{0,1}f \\
&= \sum_{k=0}^{n-1} Q_{0,k}\left[Q_{k,k+1} - P_{\frac{k}{n},\frac{k+1}{n}}\right]P_{\frac{k+1}{n},1}f.
\end{aligned}
\tag{1}
$$

A general approximation theorem: Chen*, Shao and X. ('23)

Theorem 0 (General framework)

Let $(X_t)_{t\geq 0}$ be an E-valued Markov process with infinitesimal generator $\mathcal{A}^X$, let $(Y_k)_{k\in\mathbb{N}_0}$ be an E-valued Markov chain with infinitesimal generator $\mathcal{A}^Y$. Let $N \geq 2$ be a natural number and let $h : E \to \mathbb{R}$ be a measurable function satisfying a certain condition. Then

$$\mathbb{E}h(X_N) - \mathbb{E}h(Y_N) = \mathcal{I} + \mathcal{II} + \mathcal{III}, \tag{2}$$

where

A general approximation theorem

Theorem 0 (General framework (continued))

$$\mathcal{I} = \sum_{j=1}^{N-1} \mathbb{E}\big[\mathcal{A}^X u_{N-j}(Y_{j-1}) - \mathcal{A}^Y u_{N-j}(Y_{j-1})\big],$$

$$\mathcal{II} = \sum_{j=1}^{N-1} \mathbb{E} \int_0^1 \big[\mathcal{A}^X u_{N-j}(X_s^{Y_{j-1}}) - \mathcal{A}^X u_{N-j}(Y_{j-1})\big] \mathrm{ds},$$

$$\mathcal{III} = \mathbb{E}\big[h\big(X_1^{Y_{N-1}}\big) - h(Y_{N-1})\big] + \mathbb{E}\big[h(Y_N) - h(Y_{N-1})\big],$$

with $u_t(x) := \mathbb{E}[h(X_t^x)]$ for $t > 0$.

Remarks about the theorem

To use the theorem to study $d(\mathcal{L}(X_N), \mathcal{L}(Y_N))$, we need to

- choose the function family of h, e.g.
  - bounded measurable: TV metric
  - Lipschitz: Wasserstein-1 metric
- bound the three terms $\mathcal{I}$, $\mathcal{II}$, $\mathcal{III}$:
  - PDE method
  - heat kernel
  - Malliavin calculus

We have used this theorem or its modification to study the following problems:

- Normal CLT (Stein ('72), Chen et al. ('07), Chen* & Shao & X. ('23))
- Stable CLT (X. ('19), Chen*& Nourdin & X. & Yang ('21, '22), Chen* & Shao & X. ('23))
- Optimal rate of stable CLT in TV distance (Li* & Shao & X. & Yang* ('23+, in progress))
- Minimization problem in machine learning
  - SGD approximation (E et al. ('18), Chen* & Shao & X. ('23))
  - SVRG approximation (Chen* & Lu* & X. ('22))
- Hamiltonian Monte Carlo algorithm (Eberle et al. ('19), Wang* & X. & Yang* ('23+, in progress))
- Steady state approximation for queueing systems (Jin* & Pang & X. & Xu* ('23+))

Contents

Convergence rate of stable CLT

Kolmogorov distance

- Butzer & Hahn ('70s): a convergence rate far from optimal.
- Hall ('81, '84): a rate $n^{-\beta}$ with $\beta \in (0, \frac{2-\alpha}{\alpha} \wedge 1)$.
- Kuste & Keller ('98): optimal rate $n^{-\frac{2-\alpha}{\alpha}}$ with $\alpha \in (1,2)$ for the Pareto distributed ξ_i.

Wasserstein distance: **optimal rate**

- X. ('19): symmetric case, $\alpha \in (1,2)$.
- Chen* & Nourdin & X. ('21): general case, $\alpha \in (1,2)$.
- Chen* & Nourdin & X., Yang, Zhang ('22): general case, $\alpha \in (0,1]$.

CLT in TV distance

Normal CLT:

- One dimensional normal CLT in TV distance: Prohorov ('52), Petrov ('56).
- Multidimensional normal CLT on special sets (e.g. convex sets): Rao ('61), Battacharaya ('68), Chernozhukov et al. ('17)
- Multidimensional normal CLT in TV distance with optimal convergence rate: Bally et al. ('18).

Stable CLT:

- It seems not known whether stable CLT holds in TV distance, not to mention the convergence rate!

A simple version of our result

Theorem (Li*& Shao & X. & Yang*, 2023, in progress)

Let $\xi_1, ..., \xi_n, ...$ be i.i.d. with Pareto distribution, i.e. ξ_1 has a density function $p(x) = \frac{K}{|x|^{\alpha+1}} 1_{\{|x| \geq A\}}$ for some $A > 0$ and $K > 0$. Let $S_n = \xi_1 + ... + \xi_n$. As $n \to \infty$,

$$d_{TV}\left(\mathcal{L}\left(\frac{S_n}{n^{1/\alpha}}\right), \mathcal{L}(Z)\right) \leq C \begin{cases} n^{-\frac{2-\alpha}{\alpha}} & \alpha \in (1,2), \\ n^{-1}\log n & \alpha = 1, \\ n^{-1} & \alpha \in (0,1), \end{cases}$$

where Z has a symmetric stable distribution with a characteristic function $\exp(-c|\lambda|^{\alpha})$.

We prove a general result for ξ_1 with a distribution falling in the domain of normal attraction.

One page for the idea of the proof

- We consider

$$\mathbb{E}f(S_n) - \mathbb{E}f(Z_1) := Q_{0,n}f - P_{0,1}f$$
$$= \sum_{k=0}^{n-1} Q_{0,k}\left[Q_{k,k+1} - P_{\frac{k}{n},\frac{k+1}{n}}\right]P_{\frac{k+1}{n},1}f. \quad (3)$$

- The test functions f are in the bounded measurable function family.
- We need to use the regularity of the semigroups P and Q.

Thanks a lot for your kind attention!