

Convergence rate of stable law: Stein's method approach

Lihu Xu

University of Macau

July 20, 2017

13th Workshop on Markov Process and Related Topics

Beijing Normal University and Wuhan University

Outline

- Central limit theorem (CLT), Stein's method: convergence rate of CLT
- Stable law and Convergence rate of Stable law
- An inequality of stable law with an example
- Proof of the convergence rate of stable law
- Future work

Let $\xi_{n,1}, \dots, \xi_{n,n}$ be a sequence of independent random variables such that

$$\mathbb{E}\xi_{n,k} = 0 \quad \forall k, \quad \sum_{k=1}^n \mathbb{E}\xi_{n,k}^2 = 1.$$

Denote

$$S_n = \sum_{k=1}^n \xi_{n,k}.$$

Then $S_n \Rightarrow N(0, 1)$ iff the following Lindeberg's condition holds: for every $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{E} \left[\xi_{n,k}^2 1_{\{|\xi_{n,k}| > \epsilon\}} \right] = 0. \quad (1)$$

CLT: Berry-Esseen bound

Moreover, if $\mathbb{E}|\xi_{n,k}|^3 < \infty$ for every k , we further have

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(S_n \leq x) - \Phi(x)| \leq C \sum_{k=1}^n \mathbb{E}|\xi_{n,k}|^3. \quad (2)$$

An simple example: Let X_1, \dots, X_n, \dots be i.i.d. random variables with $\mathbb{E}X_k = 0$, $\mathbb{E}X_k^2 = 1$ and $\mathbb{E}|X_k|^3 < \infty$ for each k . Then, taking $\xi_{n,k} = \frac{X_k}{\sqrt{n}}$, we have

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{X_1 + \dots + X_n}{\sqrt{n}} \leq x \right) - \Phi(x) \right| \leq \frac{C\mathbb{E}|X_1|^3}{\sqrt{n}}. \quad (3)$$

Stein's method

Charles Stein put forward a new method to study (3). The main ingredient of his method are the following two points:

- Stein's characterized equation:

$$\mathbb{E}[f'(Z) - Zf(Z)] = 0 \quad \forall f \in C_b^1(\mathbb{R}) \quad \iff \quad Z \sim N(0, 1).$$

- Stein's equation: for some class of function h ,

$$f'(x) - xf(x) = h(x) - \mathbb{E}[h(Z)]. \quad (4)$$

Stein's method (Ctd)

After solving Stein's equation:

$$f'(x) - xf(x) = h(x) - \mathbb{E}[h(Z)], \quad (5)$$

we get

$$|\mathbb{E}h(S_n) - \mathbb{E}h(Z)| = |\mathbb{E}[f'(S_n)] - \mathbb{E}[S_n f(S_n)]|. \quad (6)$$

If $S_n \sim N(0, 1)$, we know from (i) that $|\mathbb{E}[f'(S_n)] - \mathbb{E}[S_n f(S_n)]| = 0$.

If S_n is very close to $N(0, 1)$, we expect

$$|\mathbb{E}[f'(S_n)] - \mathbb{E}[S_n f(S_n)]| \quad \text{will be small.}$$

Hence, the key point is to bound $|\mathbb{E}[f'(S_n)] - \mathbb{E}[S_n f(S_n)]|$.

Literature review of Stein's method: Normal and Poisson cases (far from being complete)

- Normal approximation: Stein (1970s, 1980s), Barbour, Goldstein, Shao, Xia, Hsu, Charterjee,...
- Poisson approximation: Chen (1970s), Goldstein (1990s), Xia, ...
- Malliavin calculus with Stein's method: Picatti, Nourdin, Swan, Ledoux,...
- Ultrahigh dimensional Stein's method: Chernozhukoff, Chetverikov, Kato,...
- Xie and X. applied Stein's method to estimate the penalty level of ultrahigh dimensional minimization problem in Statistics (lasso).
- More: <https://sites.google.com/site/steinsmethod/home>

Stable law

Theorem (Theorem 3.7.2, Durrett's book)

Let $\xi_1, \dots, \xi_n, \dots$ be i.i.d. with a distribution that satisfies

$$(i) \lim_{x \rightarrow \infty} \frac{\mathbb{P}(\xi_1 > x)}{\mathbb{P}(|\xi_1| > x)} = \frac{1}{2}, \quad (ii) \mathbb{P}(|\xi_1| > x) = Kx^{-\alpha}L(x),$$

where $\alpha \in (0, 2)$, $K > 0$ and $L : [0, \infty) \rightarrow [0, \infty)$ is a slowly varying function, i.e. $\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1$ for all $t > 0$. Let $T_n = \xi_1 + \dots + \xi_n$,

$$a_n = \inf\{x : \mathbb{P}(|\xi_1| > x) \leq n^{-1}\}, \quad b_n = n\mathbb{E}[\xi_1 1_{(|\xi_1| \leq a_n)}].$$

As $n \rightarrow \infty$, the law of $(T_n - b_n)/a_n$ weakly converges to a symmetric stable law μ with characteristic function $\exp\left(-\frac{\alpha|\lambda|^\alpha}{2c_\alpha}\right)$, $c_\alpha = \left(\int_{-\infty}^{\infty} \frac{1 - \cos y}{|y|^{1+\alpha}} dy\right)^{-1}$.

Some remark about Stable law

- The theorem in the previous slide gives the sufficient and necessary condition such that the stable law holds. That is, (i) and (ii) is a 'Lindeberg condition' of stable law (see the remark below Theorem 3.7.2 in Durrett's book).
- Similar as studying the Berry-Esseen bound of CLT, to obtain the convergence rate of stable law, we need to strengthen (i) and (ii).

Literature review of convergence rate of stable law

There are not many results on the convergence rate of stable law:

- Butzer and Hahn (1970s): Under very complicate conditions.
- Kuste and Keller (1998): convergence rate $n^{-\frac{2-\alpha}{\alpha}}$ for $1 \leq \alpha \leq 2$ and i.i.d. random variables with density function $\frac{K}{|x|^{1+\alpha}}$ ($|x| \leq A$).
- Hall (1981, 1984): convergence rate $n^{-\frac{2-\alpha-\beta}{\alpha}}$ for some β .

All the known work was done by characteristic function, no work by Stein's method. See the the website of Stein's method:

<https://sites.google.com/site/steinsmethod/home>

Metric distance between probabilities

- Wasserstein's distance-1: $d_W(\mu, \nu) = \inf_{(X,Y) \in \mathcal{C}(\mu,\nu)} \mathbb{E}|X - Y|$. Moreover, by a duality,

$$d_W(\mu, \nu) = \sup_{h \text{ Lip}, \|h'\| \leq 1} |\mu(h) - \nu(h)|. \quad (7)$$

- Kolmogorov distance: Let X and Y be the random variables with distributions μ and ν respectively. Then

$$d_{\text{Kol}}(\mu, \nu) = \sup_{x \in \mathbb{R}} |\mathbb{P}(X \leq x) - \mathbb{P}(Y \leq x)|.$$

There are other distances: Total variation distance, W_2 distance, Entropic distance,

An inequality of stable law

Theorem

Let $n \in \mathbb{N}$ and let $\{\zeta_{n,i}\}_{1 \leq i \leq n}$ be a sequence of i.i.d. random variables such that $\mathbb{E}\zeta_{n,1} = 0$. Denote

$$S_n = \zeta_{n,1} + \dots + \zeta_{n,n}.$$

If $\mathcal{L}(S_n) \Rightarrow \mu$ such that μ is an α -stable distribution with characteristic function:

$$\int_{\mathbb{R}} e^{i\lambda x} \mu(dx) = e^{-\sigma|\lambda|^\alpha}$$

for some $\sigma > 0$ and $\alpha \in (1, 2)$, then we have

$$d_W(\mathcal{L}(S_n), \mu) \leq C_{\alpha, \sigma} [\Lambda_1(n, N) + \Lambda_2(n, N)], \quad (8)$$

An inequality of stable law (Ctd)

Theorem

where $N > 0$ is an arbitrary number and

$$\Lambda_1(n, N) = \int_{-N}^N \left| \frac{c_\alpha \sigma}{(\alpha - 1)} \left(\frac{1}{|t|^{\alpha-1}} - \frac{1}{N^{\alpha-1}} \right) - nK_N(t) \right| dt, \quad (9)$$

$$\Lambda_2(n, N) = n\mathbb{E}(|\zeta_{n,1}|1_{\{|\zeta_{n,1}|>N\}}) + n(\mathbb{E}|\zeta_{n,1}|)^2 + \mathbb{E}|\zeta_{n,1}| + N^{1-\alpha}$$

with

$$c_\alpha = \left(\int_{-\infty}^{\infty} \frac{1 - \cos y}{|y|^{1+\alpha}} dy \right)^{-1}, \quad (10)$$

$$K_N(t) = \mathbb{E} \left[\zeta_{n,i} 1_{\{0 \leq t \leq \zeta_{n,i} \leq N\}} - \zeta_{n,i} 1_{\{-N \leq \zeta_{n,i} \leq t \leq 0\}} \right]$$

and arbitrary $N > 0$.

A typical example about stable law (Chung, Durrett)

Assume that $\xi_1, \dots, \xi_n, \dots$ be i.i.d. with a distribution that satisfies

$$\mathbb{P}(\xi_1 \geq x) = \frac{1}{2|x|^\alpha}, \quad x \geq 1,$$

$$\mathbb{P}(\xi_1 \leq x) = \frac{1}{2|x|^\alpha}, \quad x \leq -1,$$

i.e., ξ_1 has a density function $p(x)$

$$p(x) = 0, \quad |x| \leq 1; \quad p(x) = \frac{\alpha}{2|x|^{\alpha+1}}, \quad |x| > 1.$$

Denote $\zeta_{n,i} = n^{-\frac{1}{\alpha}}\xi_i$ for $i = 1, \dots, n$, S_n weakly converges to a stable μ with characteristic function

$$\int_{\mathbb{R}} e^{i\lambda x} \mu(dx) = \exp\left(-\frac{\alpha|\lambda|^\alpha}{2c_\alpha}\right).$$

A typical example about stable law (Ctd)

So, $\sigma = \frac{\alpha}{2c_\alpha}$ in the previous inequality. Moreover, we take $N = n^{\frac{2-\alpha}{\alpha(\alpha-1)}}$, then it is easy to check

$$\Lambda_1(n, N) \asymp n^{-\frac{2-\alpha}{\alpha}}.$$

Let us compute $\Lambda_2(n, N)$. When $t > 0$,

$$\begin{aligned} K_N(t) &= \mathbb{E}(|\zeta_{n,1}| 1_{\{0 \leq t \leq \zeta_{n,1} \leq N\}}) \\ &= n^{-\frac{1}{\alpha}} \int_{n^{1/\alpha}t}^{n^{1/\alpha}N} xp(x)dx \\ &= \frac{1}{2}n^{-\frac{1}{\alpha}}\alpha \left[\frac{(tn^{1/\alpha})^{-\alpha+1} \vee 1}{\alpha-1} - \frac{(Nn^{1/\alpha})^{-\alpha+1}}{\alpha-1} \right] \quad (11) \\ &= \frac{\alpha}{2n(\alpha-1)} \left[\left(t \vee \frac{1}{n^{1/\alpha}} \right)^{-\alpha+1} - \frac{1}{N^{\alpha-1}} \right]. \end{aligned}$$

A typical example about stable law (Ctd)

When $t \leq 0$,

$$\begin{aligned} K_N(t) &= \mathbb{E}(|\zeta_{n,1}| 1_{\{-N \leq \zeta_{n,1} \leq t < 0\}}) \\ &= n^{-\frac{1}{\alpha}} \int_{-n^{1/\alpha}N}^{n^{1/\alpha}t} |x|p(x)dx \\ &= \frac{1}{2}n^{-\frac{1}{\alpha}}\alpha \left[\frac{(|t|n^{1/\alpha})^{-\alpha+1} \vee 1}{\alpha-1} - \frac{(Nn^{1/\alpha})^{-\alpha+1}}{\alpha-1} \right] \quad (12) \\ &= \frac{\alpha}{2n(\alpha-1)} \left[\left(|t| \vee \frac{1}{n^{1/\alpha}} \right)^{-\alpha+1} - \frac{1}{N^{\alpha-1}} \right]. \end{aligned}$$

A typical example about stable law (Ctd)

Hence,

$$\begin{aligned}\Lambda_2(n, N) &= n \int_{-N}^N \left| \frac{c_\alpha \sigma}{n(\alpha - 1)} \left(\frac{1}{|t|^{\alpha-1}} - \frac{1}{N^{\alpha-1}} \right) - \mathbb{E} (|\zeta_{n,1}| 1_{\{t \leq \zeta_{n,1} \leq N\}}) \right| dt \\ &= n \int_{-N}^N \left| \frac{\alpha}{2n(\alpha - 1)} \left(\frac{1}{|t|^{\alpha-1}} - \frac{1}{N^{\alpha-1}} \right) - \mathbb{E} (|\zeta_{n,1}| 1_{\{t \leq \zeta_{n,1} \leq N\}}) \right| dt \\ &= \frac{\alpha}{2(\alpha - 1)} \int_{-N}^N \left| \frac{1}{|t|^{\alpha-1}} - \left(t \vee \frac{1}{n^{1/\alpha}} \right)^{-\alpha+1} \right| dt \\ &= \frac{\alpha}{2(\alpha - 1)} \int_{-n^{-\frac{1}{\alpha}}}^{n^{-\frac{1}{\alpha}}} \left| \frac{1}{|t|^{\alpha-1}} - n^{1-\frac{1}{\alpha}} \right| dt \\ &= \frac{\alpha}{2(\alpha - 1)} \int_{-n^{-\frac{1}{\alpha}}}^{n^{-\frac{1}{\alpha}}} \left| \frac{1}{|t|^{\alpha-1}} - n^{1-\frac{1}{\alpha}} \right| dt \\ &\leq C_\alpha n^{-\frac{2-\alpha}{\alpha}}.\end{aligned}$$

A typical example about stable law (Ctd)

Hence, we have

$$d_W(S_n, \mu) \leq C_\alpha n^{-\frac{2-\alpha}{\alpha}}. \quad (14)$$

We can study some other examples, e.g. Hall (1981).

The strategy of the proof of the inequality

- Stein's characterized equation

$$\mathbb{E} \left[\Delta^{\alpha/2} f(X) - X f'(X) \right] = 0 \iff X \sim \mu.$$

- Study Stein's equation $\Delta^{\alpha/2} f(x) - x f'(x) = h(x) - \mu(h)$ with h Lipschitz.
- Calculate $\mathbb{E}[S_n f'(S_n)]$.
- Rewrite the fractional Laplacian $\Delta^{\alpha/2} f(x)$ in a new form.

Stein's equation: $\Delta^{\alpha/2} f(x) - x f'(x) = h(x) - \mu(h)$

The solution is

$$f(x) = \int_0^\infty \mathbb{E}[h(X_t(x)) - \mu(h)] dt \quad (15)$$

where $dX_t = -X_t dt + dZ_t$ with Z_t being standard symmetric α -stable process. We have the following regularity result:

$$\|f'\| \leq C\|h'\|, \quad \|f''\| \leq C\|h'\|, \quad |\Delta^{\alpha/2} f'(x)| \leq C(1 + |x|)\|h'\|. \quad (16)$$

Proof of regularity of Stein's equation

- Relation between the X_t and Z_t :

$$X_t \sim p\left(1 - e^{-t}, y - e^{-\frac{t}{\alpha}}x\right), \quad Z_t \sim p(t, x).$$

- To prove the above regularity, we need to use some integration by parts and the following classical heat kernel estimates (Chen, Kim, Kumagai, Song, Zhang):

$$p(t, x) \asymp \frac{t}{(t^{1/\alpha} + |x|)^{1+\alpha}}, \quad (17)$$

$$\partial_x p(t, x) \lesssim \frac{t}{(t^{1/\alpha} + |x|)^{2+\alpha}}, \quad (18)$$

Calculation of $\mathbb{E}[S_n f'(S_n)]$

For every $1 \leq i \leq n$, further denote $S_n(i) = S_n - \zeta_{n,i}$.

Lemma

We have

$$\mathbb{E}[S_n f'(S_n)] = \sum_{i=1}^n I(i) + R_1 \quad (19)$$

where

$$\begin{aligned} I(i) &= \int_{-N}^N \mathbb{E}[f''(S_n(i) + t)] K_N(t) dt, \\ R_1 &= \sum_{i=1}^n \mathbb{E} \left\{ \zeta_{n,i} [f'(S_n) - f'(S_n(i))] 1_{\{|\zeta_{n,i}| > N\}} \right\}, \end{aligned} \quad (20)$$

and $K_N(t) = \mathbb{E} \left[\zeta_{n,i} 1_{\{0 \leq t \leq \zeta_{n,i} \leq N\}} - \zeta_{n,i} 1_{\{-N \leq \zeta_{n,i} \leq t \leq 0\}} \right]$ and $N > 0$.

A remark about the lemma

- For the normal approximation (Chen and Shao), one has

$$\mathbb{E}[S_n f'(S_n)] = \sum_{i=1}^n \int_{-\infty}^{\infty} f''(S_n(i) + t) K(t) dt \quad (21)$$

where

$$K(t) = \mathbb{E} \left[\zeta_{n,i} 1_{\{0 \leq t \leq \zeta_{n,i}\}} - \zeta_{n,i} 1_{\{\zeta_{n,i} \leq t \leq 0\}} \right].$$

- For our case, we can derive the same relation as the above.
- $K_N(t)$ and $K(t)$ can be taken as Stein's kernel (Ledoux, Picatti, Nourdin).
- Due to the heavy tail property of $\Delta^{\alpha/2}$, we need to do a truncation to meet this property.

A new form of $\Delta^{\alpha/2} f$

Lemma

Let $x \in \mathbb{R}$, we have

$$\Delta^{\frac{\alpha}{2}} f(x) = J[x] + R_2[x] \quad (22)$$

where

$$\begin{aligned} J[x] &= \frac{c_\alpha}{\alpha(\alpha-1)} \int_{-N}^N (|t|^{1-\alpha} - N^{1-\alpha}) f''(x+t) dt, \\ R_2[x] &= \frac{c_\alpha}{\alpha} \int_{|z|>N} \frac{f'(x+z) - f'(x)}{\operatorname{sgn}(z)|z|^\alpha} dz, \end{aligned} \quad (23)$$

and $N > 0$ is an arbitrary number.

Proof of the inequality

$$\mathbb{E}[h(S_n)] - \mu(h) = \mathbb{E}[\Delta^{\frac{\alpha}{2}} f(S_n)] - \mathbb{E}[S_n f'(S_n)] = \Theta_1 + \Theta_2, \quad (24)$$

where

$$\begin{aligned} \Theta_1 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\Delta^{\frac{\alpha}{2}} f(S_n) - \Delta^{\frac{\alpha}{2}} f(S_n(i))], \\ \Theta_2 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\Delta^{\frac{\alpha}{2}} f(S_n(i))] - \frac{1}{\alpha} \mathbb{E}[S_n f'(S_n)]. \end{aligned} \quad (25)$$

Proof of the inequality (Ctd)

$$\begin{aligned} |\Theta_1| &\leq \frac{1}{n} \sum_{i=1}^n \left| \mathbb{E} [\Delta^{\frac{\alpha}{2}} f(S_n) - \Delta^{\frac{\alpha}{2}} f(S_n(i))] \right| \\ &= \frac{1}{n} \sum_{i=1}^n \left| \mathbb{E} \left[\int_0^{\zeta_{n,i}} (\Delta^{\frac{\alpha}{2}} f)'(S_n(i) + t) dt \right] \right| \\ &= \frac{1}{n} \sum_{i=1}^n \left| \mathbb{E} \left\{ \mathbb{E} \left[\int_0^{\zeta_{n,i}} (\Delta^{\frac{\alpha}{2}} f)'(S_n(i) + t) dt \mid \zeta_{n,i} \right] \right\} \right| \\ &= \frac{1}{n} \sum_{i=1}^n \left| \mathbb{E} \left\{ \int_0^{\zeta_{n,i}} \mathbb{E} \left[(\Delta^{\frac{\alpha}{2}} f)'(S_n(i) + t) \right] dt \right\} \right|. \end{aligned} \tag{26}$$

Proof of the inequality (Ctd)

Hence, by the regularity $|\Delta^{\alpha/2} f'(x)| \leq C(1 + |x|)\|h'\|$,

$$\begin{aligned} |\Theta_1| &\leq \frac{1}{n} \sum_{i=1}^n \left| \mathbb{E} \left\{ \int_0^{\zeta_{n,i}} \mathbb{E} \left| (\Delta^{\frac{\alpha}{2}} f)'(S_n(i) + t) \right| dt \right\} \right| \\ &\leq \frac{C_\alpha \|h'\|}{n} \sum_{i=1}^n \left| \mathbb{E} \left\{ \int_0^{\zeta_{n,i}} (\mathbb{E} |S_n(i) + t| + 1) dt \right\} \right| \\ &\leq \frac{C_\alpha \|h'\|}{n} \sum_{i=1}^n \left| \mathbb{E} \left\{ \int_0^{\zeta_{n,i}} (\mathbb{E} |S_n(i)| + \mathbb{E} |\zeta_{n,i}| + 1) dt \right\} \right| \\ &= \frac{C_\alpha \|h'\|}{n} \sum_{i=1}^n \mathbb{E} |\zeta_{n,1}| (\mathbb{E} |S_n(i)| + \mathbb{E} |\zeta_{n,1}| + 1) \\ &\leq \frac{C_\alpha \|h'\|}{n} \sum_{i=1}^n \mathbb{E} |\zeta_{n,1}| (n\mathbb{E} |\zeta_{n,1}| + 1) \\ &= C_\alpha \|h'\| [n(\mathbb{E} |\zeta_{n,1}|)^2 + \mathbb{E} |\zeta_{n,1}|]. \end{aligned} \tag{27}$$

Proof of the inequality (Ctd)

For Θ_2 , by the two lemmas, we have

$$\begin{aligned}\Theta_2 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \{ J[S_n(i)] + R_2[S_n(i)] \} \\ &\quad - \frac{1}{\alpha\sigma} \sum_{i=1}^n I(i) - \frac{1}{\alpha} R_1 \\ &= \sum_{i=1}^n \left\{ \frac{1}{n} \mathbb{E} \{ J[S_n(i)] \} - \frac{1}{\alpha} I(i) \right\} \\ &\quad + \left\{ \frac{1}{n} \sum_{i=1}^n \mathbb{E} \{ R_2[S_n(i)] \} - \frac{1}{\alpha} R_1 \right\}.\end{aligned}\tag{28}$$

Proof of the inequality (Ctd)

Noticing the definition of $J_1[x]$ and I_1 in the two lemmas, we further get

$$\begin{aligned} & \frac{1}{n} \mathbb{E}\{J_1[S_n(i)]\} - \frac{1}{\alpha} I_1(i) \\ &= \frac{1}{\alpha} \int_{-N}^N \mathbb{E}[f''(S_n(i) + t)] \left\{ \frac{c_\alpha}{(\alpha - 1)n} \left(\frac{1}{t^{\alpha-1}} - \frac{1}{N^{\alpha-1}} \right) - K_N(t) \right\} dt. \end{aligned} \quad (29)$$

Hence, by the regularity $\|f''\| \leq C\|h'\|$,

$$\begin{aligned} & \left| \frac{1}{n} \mathbb{E}\{J[S_n(i)]\} - \frac{1}{\alpha} I(i) \right| \\ & \leq C_{\alpha, \|h'\|} \int_{-N}^N \left| \frac{c_\alpha}{(\alpha - 1)n} \left(\frac{1}{t^{\alpha-1}} - \frac{1}{N^{\alpha-1}} \right) - K_N(t) \right| dt. \end{aligned} \quad (30)$$

Proof of the inequality (Ctd)

Collecting all the previous estimates, we immediately obtain the inequality, as desired.

Future work

- $\alpha \leq 1$ case.
- The convergence rate under Kolmogorov distance, total variation distance, W_2 distance.
- The convergence rate of high dimensional stable law.
- Applications to some problems in Statistics, e.g., a group of people are now studying the estimation of parameters of stable distribution.

Thanks a lot for your attention!