The CLT for stationary Markov chains with trivial tail sigma field

Magda Peligrad, University of Cincinnati

HDP conference, Będlewo, Poland.

June, 2023

Markov chains

We consider a sequence $(\xi_n)_{n\geq 1}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $(S, \mathcal{A})$, (S a nice space, R, or Banach separable) and specify its **finite dimensional distributions**.
We give a probability measure $m : \mathcal{A} \to [0,1]$ and a kernel $P(x,A) : S \times \mathcal{A} \to [0,1]$.
For each x fixed $P(x,A)$ is a measure on $\mathcal{A}$.
For each A fixed $P(x,A)$ is a measurable function.
Define

$$P(\xi_0 \in A_0, \xi_1 \in A_1, ..., \xi_n \in A_n) = \int_{A_0} m(dx_0) \int_{A_1} P(x_0, dx_1) ... \int_{A_n} P(x_{n-1}, dx_n).$$

Stationary Markov chains

$(\xi_n)_{n\in Z}$, such that for each $i \in Z$ and some probability measure π on $\mathcal{A}$

$$P(\xi_i \in A_1, \xi_{i+1} \in A_2, ..., \xi_{i+n} \in A_n) = \int_{A_1} \pi(dx_1) \int_{A_2} P(x_1, dx_2) ... \int_{A_n} P(x_{n-1}, dx_n).$$

Then π is invariant in the sense that

$$\int_S P(x, A)\pi(dx) = \pi(A).$$

Additive functionals

Consider now

$$f : (S, \mathcal{A}, \pi) \to (R, \mathcal{B}).$$

Denote by $\mathbb{L}_0^2(\pi)$ the set of measurable functions on S such that $\int f^2 d\pi < \infty$ and $\int f d\pi = 0$.

For a function $f \in \mathbb{L}_0^2(\pi)$ let

$$X_i = f(\xi_i), \;\; S_n = \sum_{i=1}^{n} X_i.$$

Annealed and quenched central limit theorem

Annealed CLT: (Markov chain started from π i.e. stationary)

$$P(S_n/\sqrt{n} \le t) \to \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{u^2}{2\sigma^2}} du, \text{ for } \sigma > 0 \text{ and to } 0 \text{ if } \sigma = 0.$$

We denote this convergence

$$\frac{S_n}{\sqrt{n}} \Rightarrow \sigma N(0,1).$$

Quenched CLT: Denote by P^x the measure associated with the Markov chain started from x and the same kernel. (i.e. $m(A) = 1$ if $x \in A$, and $m(A) = 0$ if $x \notin A$). For $\pi-$almost all x in S,

$$P^x(S_n/\sqrt{n} \le t) \to \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{u^2}{2\sigma^2}} du \text{ for } \sigma > 0 \text{ and to } 0 \text{ if } \sigma = 0.$$

We denote this convergence,

$$\frac{S_n}{\sqrt{n}} \Rightarrow \sigma N(0,1) \text{ under } P^x, \ \pi - \text{a.s.}$$

Variance of partial sums

The variance of partial sums: $Var\ S_n = E(|S_n|^2)$,

gives some information on the strength of dependence.
In some situations we have either

$$\sup_n \frac{Var\ S_n}{n} \leq C$$

or

$$\lim_{n\to\infty} \frac{Var\ S_n}{n} = \sigma^2.$$

This convergence resembles the i.i.d. case, when we have the CLT.
Is this a sufficient condition for CLT in the stationary Markov case? In general, the answer is **NO**.

Motivation: CLT for reversible Markov chains

Reversible Markov chains.

Definition: (ξ_i, ξ_{i+1}) is distributed as (ξ_{i+1}, ξ_i) for all i.

Gordin and Lifshitz, Kipnis and Varadhan (1986):

Theorem

Assume that the Markov chain is reversible stationary and ergodic and satisfies

$$\sup_n \frac{E(S_n^2)}{n} < \infty.$$

Then the functional CLT holds ($W(t)$, standard Brownian motion)

$$\frac{S_n}{\sqrt{n}} \Rightarrow \sigma N(0,1). \quad \textit{Furthermore} \quad \frac{S_{[nt]}}{\sqrt{n}} \Rightarrow \sigma W(t)$$

Identification of σ^2 : $\lim_{n\to\infty} \frac{E(S_n^2)}{n} = \sigma^2$. ($W(t)$, standard Brownian motion).

The validity of the quenched CLT is still an open problem.

Double tail sigma field.

Denote by $\mathcal{F}_n = \sigma(\xi_k, k \leq n)$ and by $\mathcal{F}^n = \sigma(\xi_k, k \geq n)$ completed with the sets of measure 0 with respect to P.

We define the **two-sided** tail sigma field by

$$\mathcal{T}_d = \cap_{n\geq 1}(\mathcal{F}_{-n} \vee \mathcal{F}^n).$$

We say that $\mathcal{T}_d$ is trivial if for any $A \in \mathcal{T}_d$ we have $P(A) = 0$ or 1. This property is called **regularity.**

An annealed CLT for Markov chains with trivial tail sigma field

P-(2023)

Theorem

Assume

$$\sup_n \frac{E(S_n^2)}{n} < \infty.$$

and (ξ_n) *has trivial double tail sigma field* $\mathcal{T}_d$. *Then, for some* $c \geq 0$, *the following limit exists*

$$\lim_{n\to\infty} \frac{E(|S_n|)}{\sqrt{n}} = \frac{c}{\sqrt{2\pi}} \geq 0 \ ,$$

and

$$\frac{S_n}{\sqrt{n}} \Rightarrow N(0, c^2) \text{ as } n \to \infty.$$

A quenched CLT for Markov chains with trivial double tail sigma field

Theorem

P-(2023). Assume (X_n) *and* (S_n) *are as before,* $\mathcal{T}_d$ *is trivial and* S_n^2/n *is uniformly integrable. Then there is* $\sigma \geq 0$ *such that*

$$\frac{S_n}{\sqrt{n}} \Rightarrow \sigma N(0,1) \text{ and } \frac{E(S_n^2)}{n} \to \sigma^2.$$

Furthermore, the following are equivalent:

$$(a) \qquad \limsup_{n\to\infty} \frac{E^x\left(S_n^2\right)}{n} \leq \sigma^2 \quad \pi - a.s.$$

$$(b) \qquad \frac{E^x(S_n^2)}{n} \text{ converges } \pi - a.s.$$

(c) *The quenched CLT holds under* P^x *for* $\pi-$ *almost all* x.

Necesarry and sufficient conditions for regularity.

Theorem

The following are equivalent:
(a) The double tail sigma field $\mathcal{T}_d$ is trivial.
(b) For any $D \in \mathcal{F}$

$$\sup_{A \in (\mathcal{F}_{-n} \vee \mathcal{F}^n)} |P(A \cap D) - P(A)P(D)| \to 0 \text{ as } n \to \infty.$$

(c) For any $J \leq L$ and $D \in \sigma(\xi_j : J \leq j \leq L)$ the above convergence holds.

Remark. If (ξ_j), (η_j) are two independent sequences, each with double tail sigma field trivial then $\zeta_j = f(\xi_j, \eta_j)$ also has its double tail sigma field trivial.

Example with double tail sigma field trivial: Absolutely regular Markov chains

For a stationary Markov chain $(\xi_k)_{k\in Z}$, with values in a separable Banach space, the coefficient of absolute regularity is

$$\beta_n = \beta(\xi_0, \xi_n) = \sup_{C\in\mathcal{B}^2} |P((\xi_0, \xi_n)\in C) - P((\xi_0, \xi_n^*)\in C)|,$$

where (ξ_0, ξ_n^*) are independent and identically distributed. This coefficient was introduced by Volkonskii and Rozanov (1959) and was attributed there to Kolmogorov.

A stationary Markov chain is absolutely regular iff $\beta_n \to 0$.
A stationary Markov chain is Harris recurrent and aperiodic iff $\beta_n \to 0$.
A stationary, countable state Markov chain is irreducible and aperiodic iff $\beta_n \to 0$.

CLT for absolutely regular Markov chain

Chen (1999) and P. (2020)

Corollary

If $(\xi_n)_{n\in\mathbb{Z}}$ is a stationary Markov chain which is absolutely regular and

$$\sup_{n\geq 1}\frac{E(S_n^2)}{n}<\infty,$$

then there is a constant $c\geq 0$ such that

$$\frac{S_n}{\sqrt{n}}\Rightarrow cN(0,1),$$

and

$$c=\lim_{n\to\infty}\frac{\sqrt{\pi}\,(E|S_n|)}{\sqrt{2n}}.$$

Interlaced mixing

For $\mathcal{A}, \mathcal{B}$ two sub σ-algebras of $\mathcal{F}$ define the **maximal coefficient of correlation**

$$\rho(\mathcal{A}, \mathcal{B}) = \sup_{f \in \mathbb{L}_0^2(\mathcal{A}),\ g \in \mathbb{L}_0^2(\mathcal{B})} \frac{|E(fg)|}{||f|| \cdot ||g||} \ ,$$

For a sequence of random variables, $(\xi_k)_{k \in \mathbb{Z}}$, we define

$$\rho_n^* = \sup \rho(\sigma(\xi_i, i \in S), \sigma(\xi_j, j \in T)) \ ,$$

T, S such that $\min\{|t - s| : t \in T,\ s \in S\} \geq n$.

We call the sequence ρ^*-**mixing** if $\rho_n^* \to 0$ as $n \to \infty$.

This class also has the **double tail sigma field trivial**.

Tool use in the proof: notion of total ergodicity

Total ergodicity

Let Q denotes the Markov **transition operator**. For f measurable and bounded

$$(Qf)(x) = \int_S f(s)P(x, \mathrm{d}s).$$

The Markov chain is called **totally ergodic** if for every $m \in N$

$$Q^m f = f \text{ implies } f \text{ is a constant.}$$

Trivial double tail sigma field implies the Markov chain is **totally ergodic**.

Basic tool: A CLT with random centering.

The proof is based on two results. One is a CLT with random centering. P. (2020)

Theorem

If $(\xi_n)_{n\in\mathbb{Z}}$ is a stationary and ergodic Markov chain totally ergodic such that

$$\sup_{n\geq 1}\frac{E(S_n^2)}{n}<\infty,$$

then, the following limit exists

$$\lim_{n\to\infty}\frac{1}{n}||S_n-E(S_n|\xi_0,\xi_n)||^2=c^2$$

and

$$\frac{S_n-E(S_n|\xi_0,\xi_n)}{\sqrt{n}}\Rightarrow cN(0,1).$$

Tools: Lemma for dealing with the random centering

The CLT above will be combined with a lemma which takes care of the random centering.

Lemma

Let $(\xi_n)_{n\in Z}$ be a stationary sequence not necessarily Markov with trivial tail sigma field $\mathcal{T}_d$. Let $(X_n)_n$, $(S_n)_n$, $\mathcal{F}_0$ and $\mathcal{F}^n$ as defined above and

$$\sup_{n\geq 1} \frac{E(S_n^2)}{n} < \infty.$$

Then

$$E\left| E\left(\frac{S_n}{\sqrt{n}} | \mathcal{F}_0 \vee \mathcal{F}^n\right)\right| \to 0 \text{ as } n \to \infty.$$

Steps in the proof of the CLT with random centering

1. A blocking argument
2. Martingale construction with "a special compensator"
3. Martingale approximation
4. CLT
5. The characterization of the limiting variance.

Blocking argument

Blocking
Fix m ($m < n$) a positive integer and make consecutive blocks of size m.
Let $u = u_n(m) = [n/m]$.

Denote by Y_k the sum of variables in the k'th block.

So, for $k = 0, 1, ..., u - 1$, we have

$$Y_k = Y_k(m) = (X_{km+1} + ... + X_{(k+1)m}).$$

We also have a last block

$$Y_u = Y_u(m) = (X_{um+1} + ... + X_n).$$

Martingale construction

For $k = 0, 1, ..., u - 1$ let us consider the random variables

$$D_k = D_k(m) = \frac{1}{\sqrt{m}}(Y_k - E(Y_k|\xi_{km}, \xi_{(k+1)m})).$$

Let $\mathcal{F}_n = \sigma(\xi_j; j \leq n)$ (past) and $\mathcal{F}^n = \sigma(\xi_j; j \geq n)$ (future).

Note that D_k is adapted to $\mathcal{F}_{(k+1)m} = \mathcal{G}_k$ and $E(D_k|\mathcal{G}_{k-1}) = 0$ a.s.

So $(D_k, \mathcal{G}_k)_{k\geq 0}$ is a **stationary and ergodic** sequence of square integrable **martingale differences.**

CLT for the martingale difference array

By the classical central limit theorem for ergodic martingales, for every m, a fixed positive integer, we have

$$\frac{1}{\sqrt{u}} M_u(m) := \frac{1}{\sqrt{u}} \sum_{k=0}^{u-1} D_k(m) \Rightarrow N_m \text{ as } n \to \infty,$$

where N_m is a normally distributed random variable with mean 0 and variance

$$m^{-1}||S_m - E(S_m|\xi_0, \xi_m)||^2.$$

Denote **the compensator** by $Z_k = m^{-1/2}E(Y_k|\xi_{km}, \xi_{(k+1)m})$ and let $R_u(m) = \sum_{k=0}^{u-1} Z_k$. So

$$\frac{1}{\sqrt{n}} S_n \approx \frac{1}{\sqrt{u}} M_u(m) + \frac{1}{\sqrt{u}} R_u(m).$$

We can show that $M_n(m)$ and $R_n(m)$ are orthogonal but $R_u(m)/\sqrt{u}$ **might not be negligible**.

Martingale approximation

With the **notation** $T_n = S_n - E(S_n|\xi_0, \xi_n)$, for m fixed

$$||\frac{1}{\sqrt{n}}T_n - \frac{1}{\sqrt{u}}M_u(m)||^2 \approx \left(\frac{1}{n}||T_n||^2 - \frac{1}{m}||T_m||^2\right) \text{ as } n \to \infty.$$

We do not know (yet) whether the limit of $||T_n||^2/n$ **exists**. But clearly

$$\liminf_{m} \limsup_{n\to\infty} ||\frac{1}{\sqrt{n}}T_n - \frac{1}{\sqrt{u}}M_u(m)||^2 = 0.$$

Since the martingale satisfies the CLT, we obtain

$$\frac{1}{\sqrt{n}}T_n \Rightarrow cN(0,1).$$

where $c^2 = \limsup ||T_n||^2/n$. Finally, by Skorokhod's theorem and Fatou's lemma, we identify

$$c^2 = \lim ||T_n||^2/n.$$

The advantage of conditioning with respect to past and future

This new idea of **conditioning with respect to both the past and the future** of the chain has the **advantage** that $R_{n,m} = S_n - M_{n,m}$ is **orthogonal** to $M_{n,m}$, which allows us to study the behavior of $R_{n,m}$.

If we had conditioned **only** with respect to **the past** in the construction of D_k this property would not hold, and the conditions to make $R_{n,m}/\sqrt{n}$ negligible would involve **rates** of convergence to 0 of some coefficients of ergodicity. By using this new method **the rates can be avoided**, a very useful feature in applications.

Idea of proof of Lemma for removing random centering

How to show that $\sup E(S_n^2)/n < \infty$ **plus trivial tail sigma field implies**

$$E\left|E(S_n/\sqrt{n}|\mathcal{F}_0 \vee \mathcal{F}^n)\right| \to 0.$$

By reasoning on subsequences, without restricting the generality we assume that

$$\frac{S_n}{\sqrt{n}} \Rightarrow L.$$

Because $E(X_1) = 0$, by the convergence of moments in the weak laws we have that

$$E(L) = 0.$$

We prove it first under additional condition that

$$E\left|\frac{S_n}{\sqrt{n}} \to L\right| \to 0$$

and after we remove this restriction.

By stationarity and the triangle inequality, note that for every $m \in N$, $m \leq n$,

$$E\left|E\left(\frac{S_n}{\sqrt{n}}|\mathcal{F}_0 \vee \mathcal{F}^n\right)\right| \leq E\left|E\left(\frac{S_n - S_m}{\sqrt{n}}|\mathcal{F}_0 \vee \mathcal{F}^n\right)\right| +$$

$$E\left|E\left(\frac{S_n - S_m}{\sqrt{n}}|\mathcal{F}_0 \vee \mathcal{F}^n\right) - E\left(\frac{S_n}{\sqrt{n}}|\mathcal{F}_0 \vee \mathcal{F}^n\right)\right|$$

$$\leq E\left|E\left(\frac{S_{n-m}}{\sqrt{n}} - L|\mathcal{F}_{-m} \vee \mathcal{F}^{n-m}\right)\right| + \frac{E|S_m|}{\sqrt{n}} + |E(L|\mathcal{F}_{-m} \vee \mathcal{F}^{n-m})|.$$

But, by the properties of conditional expectation,

$$E\left|E\left(\frac{S_{n-m}}{\sqrt{n}}-L|\mathcal{F}_{-m}\vee\mathcal{F}^{n-m}\right)\right|\leq E\left|\frac{S_{n-m}}{\sqrt{n}}-L\right|.$$

Therefore, for $m\in N$ fixed, by letting $n\longrightarrow\infty$, and the fact that $\mathcal{F}_{-m}\vee\mathcal{F}^{n-m}$ is decreasing in n

$$\limsup_{n} E\left|E\left(\frac{S_n}{\sqrt{n}}|\mathcal{F}_0\vee\mathcal{F}^n\right)\right|\leq E\left|E(L|\cap_{n\geq 1}\left(\mathcal{F}_{-m}\vee\mathcal{F}^{n-m}\right))\right|.$$

Now, by letting $m\rightarrow\infty$, and using the fact that $\cap_{n\geq 1}\left(\mathcal{F}_{-m}\vee\mathcal{F}^{n}\right)$ is decreasing in m, we obtain that

$$\begin{aligned}\limsup_{n} E\left|E\left(\frac{S_n}{\sqrt{n}}|\mathcal{F}_0\vee\mathcal{F}^n\right)\right| &\leq E\left|E(L|\cap_{m\geq 1}\cap_{n\geq 1}\left(\mathcal{F}_{-m}\vee\mathcal{F}^{n}\right))\right| \\ &\leq E\left|E(L|\mathcal{T}_d)\right|=|E(L)|=0.\end{aligned}$$

Set $W_n = (S_n, \xi)$ where $\xi = (\xi_n)_n$. Then $W_n \Rightarrow (L, \xi)$.
By the Skorohod representation theorem, we can expand the probability space to $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and construct on this expanded probability space, vectors $\tilde{W}_n = (\tilde{S}_n, \tilde{\xi}^n)$ and $\tilde{W} = (\tilde{L}, \tilde{\xi}')$ such that for each n, $\tilde{W}_n$ is distributed as W_n, $\tilde{W}$ is distributed as W, and $\tilde{W}_n \to \tilde{W}$ a.s.
Clearly,

$$\frac{\tilde{S}_n}{\sqrt{n}} \to \tilde{L} \text{ a.s. as } n \to \infty \text{ and } \tilde{\xi}^n = \tilde{\xi}' \text{ a.s.}$$

Now since $(\tilde{S}_n/\sqrt{n})$ is uniformly integrable, we also have

$$\tilde{E}\left|\frac{\tilde{S}_n}{\sqrt{n}} - \tilde{L}\right| \to 0 \text{ as } n \to \infty,$$

and

$$\tilde{E}(\tilde{L}) = 0.$$

Then the Skorohod representation has to have $\tilde{\mathcal{T}}_d$ also trivial.

References

1. A new CLT for additive functionals of Markov chains. (2020). Stochastic Processes and their Applications. 130 (9) 5695-5709.
2. On the CLT for additive functionals of Markov chains. (2020). Electronic Communications in Probability. 25. article number 40, 1-10.
3. On the CLT for stationary Markov chains with trivial tail sigma field. (2023) Electron. Commun. Probab. 28: 1-7.
4. On the quenched CLT for stationary Markov chains. (2023). Journal of Theoretical Probabilities. online 4 May 2023.

Thank you for your attention!