

Stable convergence of CLS estimators for supercritical CBI processes

Mátyás Barczy

MTA-SZTE Analysis and Applications Research Group
University of Szeged, Hungary

Branching Processes and Applications

University of Angers, Angers, France

22 – 26 May, 2023

Supported by the project TKP2021-NVA-09.

Outline of my talk

- Stable and mixing convergences.
- Motivation of research: Häusler and Luschgy (2015) proved stable convergence of Conditional Least Squares Estimator (CLSE) of the offspring mean for supercritical Galton-Watson processes under non-extinction.
- Continuous state and continuous time branching processes with immigration (CBI processes): definition, classification and asymptotic behaviour in the supercritical case.
- CLSE of drift parameters for CBI processes based on discrete time observations.
- Some references on asymptotic behaviour of CLSE (all about convergence in distribution).
- A new result: stable convergence of CLSE of drift parameters for supercritical CBI processes based on discrete time observations.
- Proofs are based on a general multidimensional stable limit theorem due to Barczy and Pap (2023).

Stable and mixing convergences

The notion of stable (mixing) convergence is due to

Rényi (1950, 1958, 1963) and Rényi and Révész (1958).

Stable convergence is a type of convergence:

- in the classical central limit theorem, not only convergence in distribution, but mixing convergence holds as well.
- limit theorems with random indices.
- limit theorems for martingale difference arrays.
- asymptotic behaviour of estimators, high frequency statistics.

Stable convergence has nothing to do with stable distribution.

Stable and mixing convergences

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G} \subset \mathcal{F}$ be a sub- σ -algebra. Let $(\mathbf{X}_n)_{n \geq 1}$ and \mathbf{X} be \mathbb{R}^d -valued random variables defined on Ω .

Stable convergence

We say that \mathbf{X}_n converges \mathcal{G} -stably to \mathbf{X} as $n \rightarrow \infty$, if

$$\lim_{n \rightarrow \infty} \mathbb{E}(\xi \mathbb{E}(h(\mathbf{X}_n) \mid \mathcal{G})) = \mathbb{E}(\xi \mathbb{E}(h(\mathbf{X}) \mid \mathcal{G}))$$

for all random variables $\xi : \Omega \rightarrow \mathbb{R}$ with $\mathbb{E}(|\xi|) < \infty$ and for all bounded and continuous functions $h : \mathbb{R}^d \rightarrow \mathbb{R}$.

Mixing convergence

We say that \mathbf{X}_n converges \mathcal{G} -mixing to \mathbf{X} as $n \rightarrow \infty$, if \mathbf{X}_n converges \mathcal{G} -stably to \mathbf{X} as $n \rightarrow \infty$, and $\sigma(\mathbf{X})$ and \mathcal{G} are independent. This equivalently means that

$$\lim_{n \rightarrow \infty} \mathbb{E}(\xi \mathbb{E}(h(\mathbf{X}_n) \mid \mathcal{G})) = \mathbb{E}(\xi) \mathbb{E}(h(\mathbf{X}))$$

for all random variables $\xi : \Omega \rightarrow \mathbb{R}$ with $\mathbb{E}(|\xi|) < \infty$ and for all bounded and continuous functions $h : \mathbb{R}^d \rightarrow \mathbb{R}$.

Portmanteau theorem for stable convergence

The following assertions are equivalent:

- (i) \mathbf{X}_n converges \mathcal{G} -stably to \mathbf{X} as $n \rightarrow \infty$.
- (ii) for all $F \in \mathcal{G}$ with $\mathbb{P}(F) > 0$, we have

$\mathbb{P}_F^{\mathbf{X}_n}$ converges weakly to $\mathbb{P}_F^{\mathbf{X}}$ as $n \rightarrow \infty$,

where

- \mathbb{P}_F denotes the conditional probability measure given F :

$$\mathbb{P}_F(B) := \frac{\mathbb{P}(B \cap F)}{\mathbb{P}(F)}, \quad B \in \mathcal{F},$$

- $\mathbb{P}_F^{\mathbf{X}_n}$ and $\mathbb{P}_F^{\mathbf{X}}$ are the distributions of \mathbf{X}_n and \mathbf{X} under \mathbb{P}_F , resp.

Some further results on stable (mixing) convergences

- (i) If $\mathcal{G} = \{\emptyset, \Omega\}$, then \mathcal{G} -stable convergence is nothing else but convergence in distribution.
- (ii) \mathcal{G} -stable (mixing) convergence yields convergence in distribution. (Indeed, one can choose $\xi \equiv 1$ in the definition, and then use Portmanteau theorem for convergence in distribution.)
- (iii) If $X_n, n \geq 1$, and X are random variables and X is \mathcal{G} -measurable, then

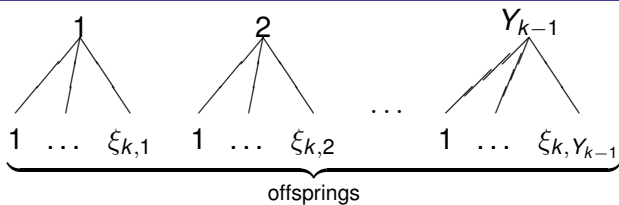
$$X_n \text{ converges to } X \text{ in probability} \iff X_n \rightarrow X \text{ } \mathcal{G}\text{-stably.}$$

By (i) and (iii), *\mathcal{G} -stable convergence is a type of convergence between convergences in probability and in distribution.*

- (iv) If $(X_n)_{n \geq 1}$, X , and Y are random variables, then

$$X_n \rightarrow X \text{ } \sigma(Y)\text{-stably} \iff (X_n, Y) \xrightarrow{\mathcal{D}} (X, Y).$$

Galton-Watson process (without immigration)



Y_k is the number of individuals in the k^{th} generation, $k \in \mathbb{Z}_+ := \{0, 1, \dots\}$.

$$Y_k = \sum_{j=1}^{Y_{k-1}} \xi_{k,j}, \quad k \in \mathbb{N} := \{1, 2, \dots\},$$

where

- $\{Y_0, \xi_{k,j} : k, j \in \mathbb{N}\}$ are independent random variables,
- Y_0 is \mathbb{N} -valued, $\xi_{k,j}$ is \mathbb{Z}_+ -valued,
- $\{\xi, \xi_{k,j} : k, j \in \mathbb{N}\}$ are identically distributed (offspring distribution).

Asymptotic behaviour of supercritical Galton-Watson processes

A classification:

$\mathbb{E}(\xi) \in (0, 1)$
subcritical

$\mathbb{E}(\xi) = 1$
critical

$\mathbb{E}(\xi) \in (1, \infty)$
supercritical

In the supercritical case (i.e., when $\mathbb{E}(\xi) > 1$), if in addition, $\mathbb{E}(Y_0^2) < \infty$ and $\mathbb{E}(\xi^2) < \infty$, then there exists a nonnegative random variable M_∞ such that $\mathbb{E}(M_\infty^2) < \infty$ and

$$\frac{Y_n}{(\mathbb{E}(\xi))^n} \rightarrow M_\infty \quad \text{as } n \rightarrow \infty \text{ in } L^2 \text{ and } \mathbb{P}\text{-a.s.}$$

Further,

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} Y_n = \infty\right) = \mathbb{P}(M_\infty > 0) > 0.$$

Hence, in the supercritical case, under the above second order moment assumptions, the conditional probability measure $\mathbb{P}_{\{M_\infty > 0\}}$ given $\{M_\infty > 0\}$ is well-defined (that will be used later on).

CLSE for supercritical Galton-Watson processes

The **CLSE** of the offspring mean $\mathbb{E}(\xi)$ based on the observations Y_0, Y_1, \dots, Y_n , $n \in \mathbb{N}$, is defined as

$$\arg \min_{\alpha \in \mathbb{R}} \sum_{i=1}^n (Y_i - \alpha Y_{i-1})^2.$$

Here note that

$$Y_i - \mathbb{E}(Y_i \mid Y_0, \dots, Y_{i-1}) = Y_i - \mathbb{E}(\xi) Y_{i-1}, \quad i = 1, \dots, n.$$

The CLSE of $\mathbb{E}(\xi)$ takes the form

$$\hat{\alpha}_n^{(\text{CLSE})} := \frac{\sum_{k=1}^n Y_{k-1} Y_k}{\sum_{k=1}^n Y_{k-1}^2},$$

where $\sum_{k=1}^n Y_{k-1}^2 \geq 1$, since $Y_0 \geq 1$ (by our assumption).

Theorem (Häusler and Luschgy (2015))

Suppose that $\mathbb{E}(Y_0^4) < \infty$, $\mathbb{E}(\xi^4) < \infty$, $\text{var}(\xi) > 0$ and $\alpha := \mathbb{E}(\xi) > 1$.
Then

$$\mathbb{P}_{\{M_\infty > 0\}}(\hat{\alpha}_n^{(\text{CLSE})} \rightarrow \alpha \text{ as } n \rightarrow \infty) = 1 \quad (\text{strong consistency}),$$

and

$$\begin{aligned} & \frac{(\alpha^3 - 1)^{1/2}}{\alpha^2 - 1} \alpha^{n/2} (\hat{\alpha}_n^{(\text{CLSE})} - \alpha) \\ & \rightarrow \sqrt{\text{var}(\xi)} \frac{N}{\sqrt{M_\infty}} \quad \mathcal{F}_\infty^Y\text{-stably under } \mathbb{P}_{\{M_\infty > 0\}} \text{ as } n \rightarrow \infty, \end{aligned}$$

where

- $\mathcal{F}_\infty^Y := \sigma(\bigcup_{n=0}^\infty \sigma(Y_0, Y_1, \dots, Y_n))$,
- M_∞ is an \mathcal{F}_∞^Y -measurable random variable satisfying $\mathbb{E}(M_\infty^2) < \infty$ and $\alpha^{-n} Y_n$ converges to M_∞ as $n \rightarrow \infty$ in L^2 and \mathbb{P} -a.s.,
- $N \sim \mathcal{N}(0, 1)$, and N is \mathbb{P} -independent of \mathcal{F}_∞^Y .

This theorem has a version with random scaling, mixing convergence and normal limit law.

Notations:

- $\mathbb{R}_+ = [0, \infty)$: non-negative real numbers,
- $\mathbb{R}_{++} = (0, \infty)$: positive real numbers,
- \mathbb{C}_- : complex numbers with non-positive real parts,
- \mathbb{C}_{--} : complex numbers with negative real parts,
- $x \wedge y := \min(x, y)$, $x, y \in \mathbb{R}$.

Set of admissible parameters

A tuple (c, a, b, ν, μ) is called a **set of admissible parameters** if

- $c, a \in \mathbb{R}_+$,
- $b \in \mathbb{R}$,
- ν is a Borel measure on \mathbb{R}_{++} satisfying $\int_0^\infty (1 \wedge r) \nu(dr) < \infty$,
- μ is a Borel measure on \mathbb{R}_{++} satisfying $\int_0^\infty (z \wedge z^2) \mu(dz) < \infty$.

Existence of CBI processes (see, e.g., Li (2011))

Let (c, a, b, ν, μ) be a set of admissible parameters. Then there exists a unique transition semigroup $(P_t)_{t \in \mathbb{R}_+}$ on \mathbb{R}_+ such that

$$\int_0^\infty e^{uy} P_t(x, dy) = \exp \left\{ x\psi(t, u) + \int_0^t F(\psi(s, u)) ds \right\}$$

for $x, t \in \mathbb{R}_+$ and $u \in \mathbb{C}_-$, where

- for any $u \in \mathbb{C}_-$, the continuously differentiable function $\mathbb{R}_+ \ni t \mapsto \psi(t, u) \in \mathbb{C}_-$ is the unique locally bounded solution to the DE

$$\partial_t \psi(t, u) = R(\psi(t, u)), \quad \psi(0, u) = u,$$

with

$$R(u) := cu^2 + bu + \int_0^\infty (e^{uz} - 1 - u(1 \wedge z)) \mu(dz), \quad u \in \mathbb{C}_-,$$

- $F(u) := au + \int_0^\infty (e^{ur} - 1) \nu(dr)$, $u \in \mathbb{C}_-$.

CBI processes

A Markov process with state space \mathbb{R}_+ and with transition semigroup $(P_t)_{t \in \mathbb{R}_+}$ given above is called a

CBI process with parameters (c, a, b, ν, μ) .

The function $\mathbb{C}_- \ni u \mapsto R(u) \in \mathbb{C}$ is called the **branching mechanism**.

The function $\mathbb{C}_- \ni u \mapsto F(u) \in \mathbb{C}_-$ is called the **immigration mechanism**.

CB process: $a = 0$ and $\nu = 0$ (i.e., $F = 0$).

SDE for CBI processes (Dawson and Li (2006))

Let (c, a, b, ν, μ) be a set of admissible parameters. If $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty r \nu(dr) < \infty$, then there exists a pathwise unique \mathbb{R}_+ -valued solution to the jump-type SDE

$$X_t = X_0 + \int_0^t (a + BX_s) ds + \int_0^t \sqrt{2c \max\{0, X_s\}} dW_s \\ + \int_0^t \int_0^\infty \int_0^\infty z \mathbb{1}_{\{u \leq X_{s-}\}} \tilde{N}(ds, dz, du) + \int_0^t \int_0^\infty r M(ds, dr)$$

for $t \in \mathbb{R}_+$, where

- $B := b + \int_1^\infty (z - 1) \mu(dz) \in \mathbb{R}$,
- $(W_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process,
- N and M are Poisson random measures on $(0, \infty)^3$ and on $(0, \infty)^2$ with intensity measures $ds \mu(dz) du$ and $ds \nu(dr)$, resp.,
- $\tilde{N}(ds, dz, du) := N(ds, dz, du) - ds \mu(dz) du$ is the compensated Poisson random measure corresponding to N ,
- $X_0, (W_t)_{t \in \mathbb{R}_+}, N$ and M are independent.

The solution is a CBI process with parameters (c, a, b, ν, μ) .

Rewriting the SDE

Under the given moment conditions $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty r \nu(dr) < \infty$, the SDE above can also be written as

$$X_t = X_0 + \int_0^t (A + BX_s) ds + \int_0^t \sqrt{2c \max\{0, X_s\}} dW_s \\ + \int_0^t \int_0^\infty \int_0^\infty z \mathbb{1}_{\{u \leq X_{s-}\}} \tilde{N}(ds, dz, du) + \int_0^t \int_0^\infty r \tilde{M}(ds, dr)$$

for $t \in \mathbb{R}_+$, where

- $A := a + \int_0^\infty r \nu(dr)$,
- $\tilde{M}(ds, dr) := M(ds, dr) - ds \nu(dr)$ is the compensated Poisson random measure corresponding to M .

Note that this SDE contains integrals with respect to compensated Poisson random measures.

Aim: to estimate the parameters A and B based on discrete time observations, supposing that c , μ and ν are known.

Interpretation of e^B and A , and classification

One can derive

$$\mathbb{E}(X_t | X_0 = x) = e^{Bt}x + A \int_0^t e^{Bu} du, \quad x \in \mathbb{R}_+, \quad t \in \mathbb{R}_+,$$

which shows that

- $e^B = \mathbb{E}(Y_1 | Y_0 = 1)$, where $(Y_t)_{t \in \mathbb{R}_+}$ is a CBI process with parameters $(c, 0, b, 0, \mu)$ (a *pure branching* process, CB process).
- $A = \mathbb{E}(Z_1 | Z_0 = 0)$, where $(Z_t)_{t \in \mathbb{R}_+}$ is a CBI process with parameters $(0, a, 0, \nu, 0)$ (a *pure immigration* process).

Hence one may call

e^B the branching mean, and A the immigration mean.

Classification: a CBI process $(X_t)_{t \in \mathbb{R}_+}$ is called

$$\begin{cases} \textit{subcritical} & \text{if } B < 0 & (\iff e^B < 1), \\ \textit{critical} & \text{if } B = 0 & (\iff e^B = 1), \\ \textit{supercritical} & \text{if } B > 0 & (\iff e^B > 1). \end{cases}$$

Some known results about supercritical CBI processes

Asymptotic behaviour of supercritical CBI processes

Let $(X_t)_{t \in \mathbb{R}_+}$ be a supercritical CBI process with parameters (c, a, b, ν, μ) such that $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty r \nu(dr) < \infty$. Then there exists a non-negative random variable w_{X_0} with $\mathbb{E}(w_{X_0}) < \infty$ such that

$$e^{-Bt} X_t \xrightarrow{\mathbb{P}\text{-a.s.}} w_{X_0} \quad \text{as } t \rightarrow \infty.$$

Further, if $\int_1^\infty z \log(z) \mu(dz) = \infty$, then $\mathbb{P}(w_{X_0} = 0) = 1$.

Corollary

Under the assumptions of the previous result, for each $\ell \in \mathbb{N}$,

$$e^{-\ell Bn} \sum_{k=1}^n X_{k-1}^\ell \xrightarrow{\mathbb{P}\text{-a.s.}} \frac{w_{X_0}^\ell}{e^{\ell B} - 1} \quad \text{as } n \rightarrow \infty,$$

$$e^{-2Bn} \sum_{k=1}^n X_{k-1} X_k \xrightarrow{\mathbb{P}\text{-a.s.}} \frac{e^B}{e^{2B} - 1} w_{X_0}^2 \quad \text{as } n \rightarrow \infty.$$

A set of sufficient conditions in order that $\mathbb{P}(w_{X_0} = 0) = 0$

Let $(X_t)_{t \in \mathbb{R}_+}$ be a supercritical CBI process with parameters (c, a, b, ν, μ) such that $\mathbb{E}(X_0) < \infty$ and

$$\int_1^\infty z^2 \mu(dz) + \int_1^\infty r^2 \nu(dr) < \infty.$$

If $A > 0$, i.e., $a > 0$ or $\nu \neq 0$ (i.e., $(X_t)_{t \in \mathbb{R}_+}$ is not a CB process), then $\mathbb{P}(w_{X_0} = 0) = 0$.

This can be found in [Barczy, Palau and Pap \(2020, 2021\)](#).

CLSEs for CBI processes

- Let $\mathcal{F}_k^X := \sigma(X_0, X_1, \dots, X_k)$, $k \in \mathbb{Z}_+$.
- Let $\mathcal{F}_\infty^X := \sigma(\bigcup_{k=0}^\infty \mathcal{F}_k^X)$.
- Since $(X_t)_{t \in \mathbb{R}_+}$ is a time-homogeneous Markov process, we get

$$\mathbb{E}(X_k | \mathcal{F}_{k-1}^X) = \mathbb{E}(X_k | X_{k-1}) = \varrho X_{k-1} + \mathcal{A}, \quad k \in \mathbb{N},$$

where

$$\varrho := e^B \in \mathbb{R}_{++}, \quad \mathcal{A} := A \int_0^1 e^{Bs} ds = \left(a + \int_0^\infty r \nu(dr) \right) \int_0^1 e^{Bs} ds \in \mathbb{R}_+.$$

- Let us introduce the sequence

$$M_k := X_k - \mathbb{E}(X_k | \mathcal{F}_{k-1}^X) = X_k - \varrho X_{k-1} - \mathcal{A}, \quad k \in \mathbb{N},$$

of martingale differences with respect to the filtration $(\mathcal{F}_k^X)_{k \in \mathbb{Z}_+}$.

- Then $X_k = \varrho X_{k-1} + \mathcal{A} + M_k$, $k \in \mathbb{N}$.

In all what follows, we suppose that c , μ and ν are known.

CLSE of (ϱ, \mathcal{A})

For each $n \in \mathbb{N}$, a CLSE $(\hat{\varrho}_n, \hat{\mathcal{A}}_n)$ of (ϱ, \mathcal{A}) based on observations X_0, X_1, \dots, X_n can be obtained by minimizing the sum of squares

$$\sum_{k=1}^n M_k^2 = \sum_{k=1}^n (X_k - \varrho X_{k-1} - \mathcal{A})^2$$

with respect to (ϱ, \mathcal{A}) over \mathbb{R}^2 , and it has the form

$$\begin{bmatrix} \hat{\varrho}_n \\ \hat{\mathcal{A}}_n \end{bmatrix} := \frac{1}{n \sum_{k=1}^n X_{k-1}^2 - \left(\sum_{k=1}^n X_{k-1} \right)^2} \begin{bmatrix} n \sum_{k=1}^n X_k X_{k-1} - \sum_{k=1}^n X_k \sum_{k=1}^n X_{k-1} \\ \sum_{k=1}^n X_k \sum_{k=1}^n X_{k-1}^2 - \sum_{k=1}^n X_k X_{k-1} \sum_{k=1}^n X_{k-1} \end{bmatrix}$$

on the set

$$H_n := \left\{ \omega \in \Omega : n \sum_{k=1}^n X_{k-1}^2(\omega) - \left(\sum_{k=1}^n X_{k-1}(\omega) \right)^2 > 0 \right\}.$$

References on CLSE for CBI processes (discrete time observations, all about convergence in distribution)

- **Huang, Ma and Zhu (2014):**

weighted CLSE of (B, a) based on observations X_0, X_1, \dots, X_n , $n \in \mathbb{N}$, can be obtained by minimizing the sum of squares

$$\sum_{k=1}^n \frac{1}{X_{k-1} + 1} (X_k - \mathbb{E}(X_k | X_{k-1}))^2 = \sum_{k=1}^n \frac{1}{X_{k-1} + 1} \left(X_k - e^B X_{k-1} - A \int_0^1 e^{Bs} ds \right)^2$$

with respect to (B, a) over \mathbb{R}^2 , where $A = a + \int_0^\infty r \nu(dr)$.

In the **supercritical case**, under 2^{nd} -order moment conditions on ν and μ , and supposing that X_0 is a nonnegative constant,

- the weighted CLSE of B is strongly consistent and is asymptotically normal using an appropriate random scaling,
- the weighted CLSE of a is not weakly consistent, but asymptotically normal using an appropriate random scaling.

They proved results in the subcritical and critical cases as well.

- Barczy, Körmendi and Pap (2016):

CLSE of (B, A) based on observations X_1, \dots, X_n , $n \in \mathbb{N}$, assuming $X_0 = 0$ and 8th-order moment conditions on ν and μ .

In the **critical case**, provided that $A > 0$, they described its asymptotic behavior, and the limit distribution is non-normal except the case when the CBI process is a pure immigration process (otherwise, it is normal).

References on CLSE for some special CBI processes

- Overbeck and Rydén (1997):

CLSE and weighted CLSE for a [Cox–Ingersoll–Ross process](#), which is a CBI process of diffusion type (without jump part), i.e., when $\mu = 0$ and $\nu = 0$.

They studied the [subcritical case](#) assuming that X_0 is distributed according to the unique stationary distribution.

- Li and Ma (2015):

CLSE and weighted CLSE of (B, a) for an [\$\alpha\$ -stable CIR process](#), which is a CBI process with $c = 0$, $\nu = 0$ and $\mu(dz) = z^{-1-\alpha} dz$, where $\alpha \in (1, 2)$.

They studied the [subcritical case](#) assuming that X_0 is distributed according to the unique stationary distribution.

Form of CLSE of (ϱ, \mathcal{A})

Recall that the CLSE $(\widehat{\varrho}_n, \widehat{\mathcal{A}}_n)$ of (ϱ, \mathcal{A}) based on observations X_0, X_1, \dots, X_n takes the form

$$\begin{bmatrix} \widehat{\varrho}_n \\ \widehat{\mathcal{A}}_n \end{bmatrix} = \frac{1}{n \sum_{k=1}^n X_{k-1}^2 - \left(\sum_{k=1}^n X_{k-1} \right)^2} \begin{bmatrix} n \sum_{k=1}^n X_k X_{k-1} - \sum_{k=1}^n X_k \sum_{k=1}^n X_{k-1} \\ \sum_{k=1}^n X_k \sum_{k=1}^n X_{k-1}^2 - \sum_{k=1}^n X_k X_{k-1} \sum_{k=1}^n X_{k-1} \end{bmatrix}$$

on the set

$$H_n = \left\{ \omega \in \Omega : n \sum_{k=1}^n X_{k-1}^2(\omega) - \left(\sum_{k=1}^n X_{k-1}(\omega) \right)^2 > 0 \right\}.$$

Existence and uniqueness of CLSE of (ϱ, \mathcal{A}) in supercritical case

Let $(X_t)_{t \in \mathbb{R}_+}$ be a supercritical CBI process with parameters (c, a, b, ν, μ) such that $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty z^2 \mu(dz) + \int_1^\infty r^2 \nu(dr) < \infty$. Suppose that $A = a + \int_0^\infty r \nu(dr) > 0$. Then

- (i) $\mathbb{P}(w_{X_0} > 0) = 1$,
- (ii) $\lim_{n \rightarrow \infty} \mathbb{P}(H_n) = 1$, i.e., the probability of the existence of a unique CLSE $(\widehat{\varrho}_n, \widehat{\mathcal{A}}_n)$ of (ϱ, \mathcal{A}) converges to 1 as $n \rightarrow \infty$,
- (iii) $(\widehat{\varrho}_n, \widehat{\mathcal{A}}_n)$ has the form given above on the event H_n .

CLSE of (B, A)

Recall that $\varrho = e^B \in \mathbb{R}_{++}$ and $\mathcal{A} = A \int_0^1 e^{Bs} ds \in \mathbb{R}_+$.

Hence $(\varrho, \mathcal{A}) = h(B, A)$, where $h: \mathbb{R}^2 \rightarrow \mathbb{R}_{++} \times \mathbb{R}$ given by

$$h(x, y) := \left(e^x, y \int_0^1 e^{xs} ds \right), \quad (x, y) \in \mathbb{R}^2.$$

Note that h is bijective having inverse $h^{-1}: \mathbb{R}_{++} \times \mathbb{R} \rightarrow \mathbb{R}^2$ given by

$$h^{-1}(u, v) = \left(\log(u), \frac{v}{\int_0^1 u^s ds} \right), \quad (u, v) \in \mathbb{R}_{++} \times \mathbb{R}.$$

Motivated by $(B, A) = h^{-1}(\varrho, \mathcal{A})$, one can introduce a natural estimator of (B, A) based on the observations X_0, X_1, \dots, X_n by applying h^{-1} to the CLSE $(\hat{\varrho}_n, \hat{\mathcal{A}}_n)$, i.e.,

$$(\hat{B}_n, \hat{A}_n) := h^{-1}(\hat{\varrho}_n, \hat{\mathcal{A}}_n) = \left(\log(\hat{\varrho}_n), \frac{\hat{\mathcal{A}}_n}{\int_0^1 (\hat{\varrho}_n)^s ds} \right), \quad n \in \mathbb{N},$$

on the set $\{\omega \in \Omega : \hat{\varrho}_n(\omega) \in \mathbb{R}_{++}\}$.

One can derive that

- the probability of the existence of the estimator $(\widehat{B}_n, \widehat{A}_n)$ converges to 1 as $n \rightarrow \infty$ (following from the strong consistency of $\widehat{\varrho}_n$, that can be proved independently of the forthcoming limit theorem),
- on the set $\{\omega \in \Omega : \widehat{\varrho}_n(\omega) \in \mathbb{R}_{++}\}$, we have

$$(\widehat{B}_n, \widehat{A}_n) = \arg \min_{(B, A) \in \mathbb{R}^2} \sum_{k=1}^n \left(X_k - e^B X_{k-1} - A \int_0^1 e^{Bs} ds \right)^2.$$

In what follows, we will simply call $(\widehat{B}_n, \widehat{A}_n)$ the CLSE of (B, A) based on the observations X_0, X_1, \dots, X_n .

Strong consistency

Let $(X_t)_{t \in \mathbb{R}_+}$ be a supercritical CBI process with parameters (c, a, b, ν, μ) such that $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty z^2 \mu(dz) + \int_1^\infty r^2 \nu(dr) < \infty$. Suppose that $A = a + \int_0^\infty r \nu(dr) > 0$. Then

- the CLSE $\hat{\varrho}_n$ of ϱ is strongly consistent, i.e., $\hat{\varrho}_n \xrightarrow{\mathbb{P}\text{-a.s.}} \varrho$ as $n \rightarrow \infty$.
- the CLSE \hat{B}_n of B is strongly consistent, i.e., $\hat{B}_n \xrightarrow{\mathbb{P}\text{-a.s.}} B$ as $n \rightarrow \infty$.

Remark. What about (weak) consistency of \hat{A}_n ?

Our forthcoming limit theorem on CLSE will imply that

the CLSE of \hat{A}_n of A is **not** (weakly) consistent if $C \neq 0$,

and

the CLSE of \hat{A}_n of A is (weakly) consistent if $C = 0$,

where

$$C := 2c + \int_0^\infty z^2 \mu(dz).$$

In case of $C = 0$, the question of strong consistency of \hat{A}_n remains open.

Main result: asymptotic behaviour of CLSE $(\widehat{B}_n, \widehat{A}_n)$

Let $(X_t)_{t \in \mathbb{R}_+}$ be a supercritical CBI process with parameters (c, a, b, ν, μ) such that $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty z^2 \mu(dz) + \int_1^\infty r^2 \nu(dr) < \infty$. Suppose that $A = a + \int_0^\infty r \nu(dr) > 0$ (not a CB process).

Then

$$\begin{bmatrix} e^{Bn/2}(\widehat{B}_n - B) \\ ne^{-Bn/2}(\widehat{A}_n - A) \end{bmatrix} \rightarrow \mathbf{S}^{1/2} \mathbf{N} \quad \mathcal{F}_\infty^X\text{-stably as } n \rightarrow \infty,$$

where

- \mathbf{N} is a 2-dimensional random vector \mathbb{P} -independent of \mathcal{F}_∞^X such that $\mathbf{N} \stackrel{\mathcal{D}}{=} \mathcal{N}_2(\mathbf{0}, C I_2)$ with $C = 2c + \int_0^\infty z^2 \mu(dz)$,
- the random matrix \mathbf{S} , defined by

$$\mathbf{S} := \begin{bmatrix} \frac{(e^B - 1)(e^{2B} - 1)^2}{Be^B(e^{3B} - 1)} W_{X_0}^{-1} & -\frac{e^B(e^B - 1)}{e^{3B} - 1} \\ -\frac{e^B(e^B - 1)}{e^{3B} - 1} & \frac{Be^{2B}}{(e^B - 1)(e^{3B} - 1)} W_{X_0} \end{bmatrix},$$

is \mathbb{P} -independent of \mathbf{N} .

Remark. In case of $C = 0$, (i.e., $c = 0$ and $\mu = 0$), we have

- the previous result implies that
 - $e^{Bn/2}(\widehat{B}_n - B) \xrightarrow{\mathcal{D}(\mathbb{P})} 0$ as $n \rightarrow \infty$,
 - $ne^{-Bn/2}(\widehat{A}_n - A) \xrightarrow{\mathcal{D}(\mathbb{P})} 0$ as $n \rightarrow \infty$.

Hence the scaling factors $e^{Bn/2}$ and $ne^{-Bn/2}$ are *not the good ones* in the sense that the limit distribution is zero.

This motivates a separate study of the case $C = 0$.

- If, in addition, $\nu \neq 0$, then the next result shows that the scaling factors e^{Bn} and $n^{1/2}$ are *the good ones* in the sense that the limit distribution is not zero.

Main result: asymptotic behaviour of CLSE $(\widehat{B}_n, \widehat{A}_n)$

If, in addition, $C = 0$ (i.e., $c = 0$ and $\mu = 0$), then

$$e^{Bn}(\widehat{B}_n - B) \rightarrow \frac{e^{2B} - 1}{e^{2B}} w_{X_0}^{-1} \sum_{j=0}^{\infty} e^{-Bj} Z_j \quad \mathcal{F}_{\infty}^X\text{-stably as } n \rightarrow \infty,$$

and

$$n^{1/2}(\widehat{A}_n - A) \rightarrow N_1 \quad \mathcal{F}_{\infty}^X\text{-mixing as } n \rightarrow \infty,$$

where

- N_1 is a random variable \mathbb{P} -independent of \mathcal{F}_{∞}^X such that

$$N_1 \stackrel{\mathcal{D}}{=} \mathcal{N}\left(0, \frac{B(e^{2B} - 1)}{2(e^B - 1)^2} \int_0^{\infty} r^2 \nu(dr)\right),$$

- $(Z_j)_{j \in \mathbb{Z}_+}$ are \mathbb{P} -i.i.d. random variables being \mathbb{P} -independent of \mathcal{F}_{∞}^X such that Z_1 has a characteristic function

$$\mathbb{E}(e^{i\theta Z_1}) = \exp\left\{\int_0^1 \int_0^{\infty} (e^{i\theta r e^{Bu}} - 1 - i\theta r e^{Bu}) du \nu(dr)\right\}, \quad \theta \in \mathbb{R}.$$

Here $Z_1 \stackrel{\mathcal{D}}{=} M_1 = X_1 - \mathbb{E}(X_1 | X_0)$, and the series above is absolutely convergent \mathbb{P} -almost surely. 30

Some remarks

- 1 The independence of \mathbf{N} and \mathbf{S} is a consequence of that \mathbf{N} is independent of \mathcal{F}_∞^X and \mathbf{S} is \mathcal{F}_∞^X -measurable following from the \mathcal{F}_∞^X -measurability of w_{X_0} .
- 2 The \mathcal{F}_∞^X -measurability of w_{X_0} also implies that the sequence of random variables $(Z_j)_{j \in \mathbb{Z}_+}$ and w_{X_0} are independent.
- 3 The limit law for $\widehat{B}_n - B$ may depend on the law of the initial value X_0 , since the law of w_{X_0} may depend on the law of X_0 .
This phenomenon usually happens for limit laws of CLSEs for *supercritical* models using *deterministic scalings*.

Sketch of proof

The proof can be traced back to describe the asymptotic behaviour of the CLSE $(\hat{\varrho}_n, \hat{\mathcal{A}}_n)$ of (ϱ, \mathcal{A}) , since



$$\begin{bmatrix} e^{Bn/2}(\hat{B}_n - B) \\ ne^{-Bn/2}(\hat{A}_n - A) \end{bmatrix} = f_n \left(\begin{bmatrix} e^{Bn/2}(\hat{\varrho}_n - \varrho) \\ ne^{-Bn/2}(\hat{\mathcal{A}}_n - \mathcal{A}) \end{bmatrix} \right), \quad n \in \mathbb{N},$$

on the set $\{\omega \in \Omega : \hat{\varrho}_n(\omega) \in \mathbb{R}_{++}\}$, where $f_n : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by

$$f_n \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) := \begin{bmatrix} e^{Bn/2} \log \left(1 + \frac{x}{e^{Bn/2} \varrho} \right) \\ \frac{y + ne^{-Bn/2} \mathcal{A}}{\int_0^1 (\varrho + \frac{x}{e^{Bn/2}})^s ds} - ne^{-Bn/2} \mathcal{A} \end{bmatrix}$$

for $(x, y) \in \mathbb{R}^2$ with $x > -e^{Bn/2} \varrho$, and $f_n(x, y) := 0$ otherwise,

- and then one can use the following continuous mapping theorem:

Continuous mapping theorem for stable convergence

Let

- $d \in \mathbb{N}$, $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space,
- $\mathcal{G} \subset \mathcal{F}$ be a sub- σ -algebra,
- $(\xi_n)_{n \in \mathbb{N}}$ and ξ be \mathbb{R}^d -valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\xi_n \rightarrow \xi$ \mathcal{G} -stably as $n \rightarrow \infty$,
- $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $f_n : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $n \in \mathbb{N}$, be Borel measurable mappings such that $\lim_{n \rightarrow \infty} f_n(s_n) = f(s)$ if $\lim_{n \rightarrow \infty} s_n = s \in \mathbb{R}^d$.

Then

$$f_n(\xi_n) \rightarrow f(\xi) \quad \mathcal{G}\text{-stably as } n \rightarrow \infty.$$

Sketch of proof

- A useful decomposition for handling the CLSE $(\hat{\varrho}_n, \hat{\mathcal{A}}_n)$ of (ϱ, \mathcal{A}) :

$$\begin{bmatrix} \hat{\varrho}_n - \varrho \\ \hat{\mathcal{A}}_n - \mathcal{A} \end{bmatrix} = \frac{1}{n \sum_{k=1}^n X_{k-1}^2 - \left(\sum_{k=1}^n X_{k-1} \right)^2} \begin{bmatrix} n \sum_{k=1}^n M_k X_{k-1} - \sum_{k=1}^n M_k \sum_{k=1}^n X_{k-1} \\ \sum_{k=1}^n M_k \sum_{k=1}^n X_{k-1}^2 - \sum_{k=1}^n M_k X_{k-1} \sum_{k=1}^n X_{k-1} \end{bmatrix},$$

where $e^{-Bn} \sum_{k=1}^n X_{k-1}$ and $e^{-2Bn} \sum_{k=1}^n X_{k-1}^2$ converges a.s.

- We need to study the asymptotic behaviour of

$$\begin{bmatrix} \sum_{k=1}^n M_k \\ \sum_{k=1}^n M_k X_{k-1} \end{bmatrix} \quad \text{as } n \rightarrow \infty,$$

with the aim of proving **stable convergence**.

Sketch of proof

Theorem

Let $(X_t)_{t \in \mathbb{R}_+}$ be a supercritical CBI process with parameters (c, a, b, ν, μ) such that $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty z^2 \mu(dz) + \int_1^\infty r^2 \nu(dr) < \infty$. Suppose that $A = a + \int_0^\infty r \nu(dr) > 0$.

Then

$$\begin{bmatrix} e^{-Bn/2} & 0 \\ 0 & e^{-3Bn/2} \end{bmatrix} \begin{bmatrix} \sum_{k=1}^n M_k \\ \sum_{k=1}^n M_k X_{k-1} \end{bmatrix} \rightarrow \mathbf{R}^{1/2} \tilde{\mathbf{N}} \quad \mathcal{F}_\infty^X\text{-stably as } n \rightarrow \infty,$$

where

- $\tilde{\mathbf{N}}$ is a 2-dimensional rand. vector \mathbb{P} -independent of \mathcal{F}_∞^X such that $\tilde{\mathbf{N}} \stackrel{\mathcal{D}}{=} \mathcal{N}_2(\mathbf{0}, V I_2)$ with $V := C \int_0^1 e^{B(1+u)} du$ and $C = 2c + \int_0^\infty z^2 \mu(dz)$,
- the random matrix

$$\mathbf{R} := \begin{bmatrix} \frac{w_{X_0}}{e^B - 1} & \frac{w_{X_0}^2}{e^{2B} - 1} \\ \frac{w_{X_0}^2}{e^{2B} - 1} & \frac{w_{X_0}^3}{e^{3B} - 1} \end{bmatrix} \quad \text{is } \mathbb{P}\text{-independent of } \tilde{\mathbf{N}}.$$

Sketch of proof

If $C = 2c + \int_0^\infty z^2 \mu(dz) = 0$ (i.e., $c = 0$ and $\mu = 0$), then the limit distribution above is $(0, 0)$, so other scaling factors should be found in order to get a non-zero limit distribution.

Theorem

Let $(X_t)_{t \in \mathbb{R}_+}$ be a supercritical CBI process with parameters (c, a, b, ν, μ) such that $\mathbb{E}(X_0) < \infty$ and $\int_1^\infty z^2 \mu(dz) + \int_1^\infty r^2 \nu(dr) < \infty$. Suppose that $A = a + \int_0^\infty r \nu(dr) > 0$ and $C = 0$.

(i) We have

$$n^{-1/2} \sum_{k=1}^n M_k \rightarrow \tilde{N}_1 \quad \mathcal{F}_\infty^X\text{-mixing as } n \rightarrow \infty,$$

where \tilde{N}_1 is a random variable \mathbb{P} -independent of \mathcal{F}_∞^X such that

$$\tilde{N}_1 \stackrel{\mathcal{D}}{=} \mathcal{N} \left(0, \frac{e^{2B} - 1}{2B} \int_0^\infty r^2 \nu(dr) \right).$$

(ii) We have

$$e^{-Bn} \sum_{k=1}^n M_k X_{k-1} \rightarrow \frac{W_{X_0}}{e^B} \sum_{j=0}^{\infty} e^{-Bj} Z_j \quad \mathcal{F}_{\infty}^X\text{-stably as } n \rightarrow \infty,$$

where

- $(Z_j)_{j \in \mathbb{Z}_+}$ are \mathbb{P} -independent and identically distributed random variables being \mathbb{P} -independent from \mathcal{F}_{∞}^X ,
- Z_1 has a characteristic function

$$\mathbb{E}(e^{i\theta Z_1}) = \exp \left\{ \int_0^1 \int_0^{\infty} \left(e^{i\theta r e^{Bu}} - 1 - i\theta r e^{Bu} \right) du \nu(dr) \right\}, \quad \theta \in \mathbb{R}.$$

In particular, we have $Z_1 \stackrel{\mathcal{D}}{=} M_1 = X_1 - \mathbb{E}(X_1 | X_0)$.

- the series $\sum_{j=0}^{\infty} e^{-Bj} Z_j$ is absolutely convergent \mathbb{P} -almost surely.

In case of $C = 0$, we could not prove joint stable convergence of $n^{-1/2} \sum_{k=1}^n M_k$ and $e^{-Bn} \sum_{k=1}^n M_k X_{k-1}$ as $n \rightarrow \infty$.

Sketch of proof

The proof of the previous theorem,

- in case of $C \neq 0$, is based on a *multidimensional stable limit theorem* due to [Barczy and Pap \(2023\)](#).

It is a generalization of the corresponding 1-dimensional result due to [Häusler and Luschgy \(2015\)](#).

- in case of $C = 0$, is based on a *one dimensional stable limit theorem* and on a *martingale central limit theorem involving mixing convergence* due to [Häusler and Luschgy \(2015\)](#).

Next, we recall these results.

Some notations:

- for an event $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$, let

$$\mathbb{P}_A(B) := \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}, \quad B \in \mathcal{F},$$

denote the conditional probability measure given A .

- for an \mathbb{R}^d -valued stochastic process $(\mathbf{U}_n)_{n \in \mathbb{Z}_+}$, its increments are defined by

$$\Delta \mathbf{U}_n := \mathbf{U}_n - \mathbf{U}_{n-1}, \quad n \in \mathbb{N}, \quad \text{and} \quad \Delta \mathbf{U}_0 := \mathbf{0}.$$

- $\varrho(\mathbf{A})$ is the spectral radius of a matrix $\mathbf{A} \in \mathbb{R}^{d \times d}$.

A multidimensional stable limit theorem

Theorem (Barczy and Pap (2023))

Let

- $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$ be a filtration,
- $(\mathbf{U}_n)_{n \in \mathbb{Z}_+}$ be an \mathbb{R}^d -valued stochastic process adapted to $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$,
- $(\mathbf{B}_n)_{n \in \mathbb{Z}_+}$ be an $\mathbb{R}^{d \times d}$ -valued stochastic process adapted to $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$ such that \mathbf{B}_n is invertible for sufficiently large $n \in \mathbb{N}$,
- $(\mathbf{Q}_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}^{d \times d}$ such that $\mathbf{Q}_n \rightarrow \mathbf{0}$ as $n \rightarrow \infty$ and \mathbf{Q}_n is invertible for sufficiently large $n \in \mathbb{N}$,
- $G \in \mathcal{F}_\infty := \sigma(\bigcup_{k=0}^\infty \mathcal{F}_k)$ with $\mathbb{P}(G) > 0$.

\mathbf{B}_n will serve as a *random scaling* for \mathbf{U}_n ,

\mathbf{Q}_n will serve as a *deterministic scaling* for \mathbf{U}_n .

A multidimensional stable limit theorem

Assume the following conditions:

- (i) there exists an $\mathbb{R}^{d \times d}$ -valued, \mathcal{F}_∞ -measurable random matrix $\eta : \Omega \rightarrow \mathbb{R}^{d \times d}$ such that $\mathbb{P}(\mathbf{G} \cap \{\exists \eta^{-1}\}) > 0$ and

$$\mathbf{Q}_n \mathbf{B}_n^{-1} \xrightarrow{\mathbb{P}_{\mathbf{G}}} \eta \quad \text{as } n \rightarrow \infty,$$

- (ii) $(\mathbf{Q}_n \mathbf{U}_n)_{n \in \mathbb{N}}$ is stochastically bounded in $\mathbb{P}_{\mathbf{G} \cap \{\exists \eta^{-1}\}}$ -probability:

$$\lim_{K \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{P}_{\mathbf{G} \cap \{\exists \eta^{-1}\}}(\|\mathbf{Q}_n \mathbf{U}_n\| > K) = 0,$$

- (iii) there exists an invertible matrix $\mathbf{P} \in \mathbb{R}^{d \times d}$ with $\varrho(\mathbf{P}) < 1$ such that

$$\mathbf{B}_n \mathbf{B}_{n-r}^{-1} \xrightarrow{\mathbb{P}_{\mathbf{G}}} \mathbf{P}^r \quad \text{as } n \rightarrow \infty \text{ for every } r \in \mathbb{N},$$

- (iv) there exists a probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $\int_{\mathbb{R}^d} \log^+(\|\mathbf{x}\|) \mu(d\mathbf{x}) < \infty$ such that for all $\boldsymbol{\theta} \in \mathbb{R}^d$, we have

$$\mathbb{E}_{\mathbb{P}}(e^{i\langle \boldsymbol{\theta}, \mathbf{B}_n \Delta \mathbf{U}_n \rangle} \mid \mathcal{F}_{n-1}) \xrightarrow{\mathbb{P}_{\mathbf{G} \cap \{\exists \eta^{-1}\}}} \int_{\mathbb{R}^d} e^{i\langle \boldsymbol{\theta}, \mathbf{x} \rangle} \mu(d\mathbf{x}) \quad \text{as } n \rightarrow \infty.$$

A multidimensional stable limit theorem

Then

- with *random* scaling:

$$\mathbf{B}_n \mathbf{U}_n \rightarrow \sum_{j=0}^{\infty} \mathbf{P}^j \mathbf{Z}_j \quad \mathcal{F}_{\infty}\text{-mixing under } \mathbb{P}_{G_n\{\exists \eta^{-1}\}} \text{ as } n \rightarrow \infty,$$


- with *deterministic* scaling:


$$\mathbf{Q}_n \mathbf{U}_n \rightarrow \eta \sum_{j=0}^{\infty} \mathbf{P}^j \mathbf{Z}_j \quad \mathcal{F}_{\infty}\text{-stably under } \mathbb{P}_{G_n\{\exists \eta^{-1}\}} \text{ as } n \rightarrow \infty,$$

where $(\mathbf{Z}_j)_{j \in \mathbb{Z}_+}$ denotes a \mathbb{P} -independent and identically distributed sequence of \mathbb{R}^d -valued random vectors being \mathbb{P} -independent of \mathcal{F}_{∞} with $\mathbb{P}(\mathbf{Z}_0 \in B) = \mu(B)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$.

Here η and $(\mathbf{Z}_j)_{j \in \mathbb{Z}_+}$ are \mathbb{P} -independent (since η is \mathcal{F}_{∞} -measurable, and $(\mathbf{Z}_j)_{j \in \mathbb{Z}_+}$ is \mathbb{P} -independent of \mathcal{F}_{∞}).

This talk is based on the following two papers:

 MÁTYÁS BARCZY AND GYULA PAP (2023).
A multidimensional stable limit theorem.
Filomat **37(11)** 3493–3512.

 MÁTYÁS BARCZY (2022+).
Stable convergence of conditional least squares estimators for
supercritical continuous state and continuous time branching
processes with immigration.
arXiv: 2207.14056

Thank you for your attention!