

ON INVARIANT DISTRIBUTION ESTIMATION FOR A CLASS OF CONTINUOUS-TIME STATIONARY PROCESSES

Dominique Dehay

dominique.dehay@uhb.fr

Rennes 2 Haute-Bretagne

¹SASP V, Le Mans, January 7th

- I - Introduction
- II - Asymptotic independence
- III - Convergence of the empirical distribution function
- IV - Asymptotic normality
- V - Asymptotic variance estimation
- VI - Weak convergence

I - Introduction

$$X := \{X_t, t \geq 0\}$$

- continuous-time,
- real-valued
- stationary,
- μ invariant marginal distribution, law of $X_t, \forall t \geq 0$

$$F(x) := \mathbb{P}[X_0 \leq x]$$

$$\widehat{F}_T(x) := \frac{1}{T} \int_0^T \mathbb{I}_{\{X_s \leq x\}} ds$$

Example : diffusion process

$$dX_t = S(X_t)dt + dW_t$$

where $S : \mathbb{R} \rightarrow \mathbb{R}$ Borelian with a polynomial majorant

$$\limsup_{|x| \rightarrow \infty} \operatorname{sgn}(x)S(x) < 0.$$

Then

- existence and unicity of the invariant law μ with de
- $\sqrt{T} \left(\widehat{F}_T(x) - F(x) \right) \xrightarrow{loi} \mathcal{N}(0, \sigma^2(x))$

$$\begin{aligned} \sigma^2(x) &= 2 \int_0^\infty \left(F_t(x, x) - F(x)^2 \right) dt \\ &= 4f(x)^2 \mathbb{E} \left[\left(\frac{F(\xi \wedge x) - F(\xi)F(x)}{f(\xi)} \right)^2 \right] \end{aligned}$$

$$\mathcal{L}(\xi) = \mu$$

- Estimation of the asymptotic variance

$$\hat{\sigma}_x^2 := 4\hat{f}(x)^2 \frac{1}{T} \int_0^T \left(\frac{\hat{F}_T(X_t \wedge x) - \hat{F}_T(X_t)\hat{F}_T(x)}{\hat{f}_T(X_t) + T^{-1/4}} \right)^2 dt$$

Then (D. & Kutoyants 2004)

$$\hat{\sigma}_x^2 \xrightarrow{\text{P}} \sigma_x^2$$

- TLC fonctionnel

$$\left\{ \sqrt{T}(\hat{F}_T(x) - F(x)) : x \in \mathbb{R} \right\} \xrightarrow{\text{law}} G, \quad (\mathcal{C}_0(\mathbb{R})),$$

$$G = \left\{ G(x) : x \in \mathbb{R} \right\} \text{ Gaussian process (Negri 1998).}$$

Discrete-time process

1) $(Y_n)_n$ i.i.d. r.v. distribution $F(\cdot)$.

$$\widehat{F}_n(x) := \frac{1}{n} \sum_{k=1}^n \mathbb{I}_{\{Y_k \leq x\}}.$$

Then

$$\sqrt{n} \left(\widehat{F}_n(x) - F(x) \right) \xrightarrow{loi} \mathcal{N}(0, \sigma^2(x))$$

$$\sigma^2(x) = F(x)(1 - F(x))$$

$$\widehat{\sigma}_n^2(x) := \widehat{F}_n(x)(1 - \widehat{F}_n(x)) \longrightarrow \sigma^2(x) \quad \text{a.s.}$$

$$\left\{ \sqrt{n}(\widehat{F}_n(x) - F(x)) : x \in \mathbb{R} \right\} \xrightarrow{law} G, \quad (\mathbb{D}(\mathbb{R})),$$

G Gaussian process (voir Dudley 1984).

2) $(Y_n)_n$ strongly mixing stationary process with

$$\sum_n \alpha_n < \infty$$

Then

$$\sqrt{n} \left(\widehat{F}_n(x) - F(x) \right) \xrightarrow{law} \mathcal{N}(0, \sigma^2(x))$$

$$\sigma^2(x) = F(x)(1 - F(x)) + 2 \sum_{k=1}^{\infty} (F_k(x, x) - F(x)^2)$$

Furthermore

If $\sum_n n^2 \varphi_n^{1/2} < \infty$ (Billingsley 1968) or

if $F(\cdot)$ is continuous and $\alpha_n = \mathcal{O}(n^{-a})$, $a > 1$ (Rio 2000)

then

$$\left\{ \sqrt{n} (\widehat{F}_n(x) - F(x)) : x \in \mathbb{R} \right\} \xrightarrow{law} G, \quad (\mathbb{D}(\mathbb{R})),$$

G Gaussian process.

II - Asymptotic independence

(CL*) *Castellana Leadbetter*(1986)

$$T(A) \longrightarrow (A), \quad \forall A \in \mathcal{B}(\mathbb{R}^2) \quad \text{where}$$

$$T(A) := \int_0^T \left(\mu_t^{(2)}(A) - \mu \otimes \mu(A) \right) dt.$$

(SM) *Strong mixing α -mixing*(Rosenblatt 1956)

$$\sigma^s(X) := \sigma\{X(u) : u \leq s\} \quad \text{et} \quad \sigma_s(X) := \sigma\{X(u) :$$

$$\alpha(t) := \sup \left\{ |P[A \cap B] - P[A]P[B]| : A \in \sigma^s(X), B \in \sigma_{s+t} \right.$$

X is *strongly mixing* when $\lim_{t \rightarrow \infty} \alpha(t) = 0$.

Remark : $\alpha(\cdot) \in L^1([0, \infty)) \implies (CL^*)$.

Examples

1) Diffusion process (Veretennikov 1998)

$$\alpha(t) \leq \exp(-\lambda t)$$

2) Exponentially ergodic Markov process

Markov process X with

- $(P^t)_{t \geq 0}$ transition semi-group
- $\|P^t(x, \cdot) - \mu\|_{\text{var}} \leq M(x)\rho^t$
- $0 < \rho < 1, M \in L^1(\mu)$.

Then X is strongly mixing and $\alpha(\cdot) \in L^1([0, \infty))$ with

$$\alpha(t) \leq \int_{\mathbb{R}} M(x) \mu(dx) \rho^t.$$

III- Convergence of the empirical distribution

X stationary $\implies \mathbb{E}[\widehat{F}_t(x)] = F(x)$ and

$$\text{var } \widehat{F}_T(x) = \frac{2}{T} \int_0^T \left(1 - \frac{t}{T}\right) \left(F_t(x, x) - F(x)^2\right) dt$$

Proposition 1 (Asymptotic covariance)

Under (CL*),

$$\begin{aligned} C(x, y) &:= \lim_{T \rightarrow \infty} T \text{cov}[\widehat{F}_T(x), \widehat{F}_T(y)] \\ &= \int_0^\infty \left(F_t(x, y) + F_t(y, x) - 2F(x)F(y)\right) dt. \end{aligned}$$

Asymptotic variance)

$$\begin{aligned} \sigma^2(x) &:= \lim_{T \rightarrow \infty} T \text{var}[\widehat{F}_T(x)] \\ &= 2 \int_0^\infty \left(F_t(x, x) - F(x)^2\right) dt = 2 \int_{(-\infty, x]} \dots \end{aligned}$$

Proposition 2 Almost-sure convergence)

Assume that one of the two following conditions is satisfied

(i) $\alpha(t) = \mathcal{O}(t^{-\gamma})$ for some $0 < \gamma < 1$,

(ii) (CL*).

Then

$$\lim_{T \rightarrow \infty} T^\delta \left| \widehat{F}_T(x) - F(x) \right| = 0 \quad \text{a.s.}$$

for each $0 \leq \delta < \gamma/3$, in the case (ii) we take $\gamma = 1$.

The process is ergodic and satisfies *Glivenko-Cantelli* property

$$\mathbb{P} \left[\lim_{T \rightarrow \infty} \sup_x \left| \widehat{F}_T(x) - F(x) \right| = 0 \right] = 1.$$

IV - Asymptotic variance estimation

$$\sigma^2(x) = 2 \int_0^\infty \left(F_t(x, x) - F(x)^2 \right) dt.$$

Let $0 < \eta < 1/3$ fixed

$$\widehat{F}_{t,T}(x) := \frac{1}{T} \int_0^T \mathbb{I}_{\{X_s \leq x\}} \mathbb{I}_{\{X_{s+t} \leq x\}} ds$$

$$\widehat{\sigma}_T^2(x) := 2 \int_0^{T^\eta} \left(\widehat{F}_{t,T}(x) - \widehat{F}_T(x)^2 \right) dt.$$

Proposition 3 If $\alpha(\cdot) \in L^1([0, \infty))$ then $\lim_{T \rightarrow \infty} \widehat{\sigma}_T^2(x) = \sigma^2(x)$

Proposition 4 Assume that $\alpha(t) = \mathcal{O}(t^{-\gamma})$ for some $\gamma > 0$

Then $\lim_{T \rightarrow \infty} T^\delta \{ \widehat{\sigma}_T^2(x) - \sigma^2(x) \} = 0$

in quadratic mean for $0 < \eta < \frac{1}{3}$ and $\delta < \min\left\{ \frac{1}{2}(1 - 3\eta) \right\}$

almost-surely for $0 < \eta < \frac{1}{4}$, $\gamma > \frac{3}{2}$ and $\delta < \min\left\{ \frac{1}{3}(1 - 4\eta) \right\}$

Let

$$\widehat{C}_T(x, y) := \int_0^{T^\eta} \left(\widehat{F}_{t,T}(x, y) + \widehat{F}_{t,T}(y, x) - 2 \widehat{F}_T(x) \widehat{F}_T(y) \right) dt$$

$$\widehat{F}_{t,T}(x, y) := \frac{1}{T} \int_0^T \mathbb{I}_{\{X_s \leq x\}} \mathbb{I}_{\{X_{s+t} \leq y\}} ds$$

$$0 < \eta < 1/3,$$

then

$$\lim_{T \rightarrow \infty} \widehat{C}_T(x, y) = C(x, y) \quad \text{q.m.}$$

V - Asymptotic normality

Proposition 5

If $\alpha(\cdot) \in L^1([0, \infty))$ then

$$\widehat{G}_T(x) := \sqrt{T} \left(\widehat{F}_T(x) - F(x) \right) \xrightarrow{\text{law}} \mathcal{N}(0, \sigma^2)$$

If in addition $\sigma(x) \neq 0$ we deduce that

$$\frac{\widehat{G}_T(x)}{\widehat{\sigma}_T^2(x)} := \frac{\sqrt{T}}{\widehat{\sigma}_T^2(x)} \left(\widehat{F}_T(x) - F(x) \right) \xrightarrow{\text{law}} \mathcal{N}(0, \sigma^2)$$

Sketch for the proof of Proposition 5

– x fixed

$$- \left| \sqrt{T} \{ \hat{F}_T(x) - F(x) \} - \sqrt{[T]} \{ \hat{F}_{[T]}(x) - F(x) \} \right| \leq$$

$$- Y_n(x) = \int_{n-1}^n \mathbb{I}_{\{X_t \leq x\}} dt - F(x)$$

– strongly-mixing stationary sequence of bounded zero

– CLT (Ibragimov Linnik 1971)

$$- \sqrt{N} \{ \hat{F}_N(x) - F(x) \} = \frac{1}{\sqrt{N}} \sum_{k=1}^N Y_k(x) \longrightarrow \mathcal{N}(0, \sigma)$$

VI- Weak convergence

Notations : (Pollard 1990, Rio 2000)

- $\mathbb{X} = \mathbb{R}$ with pseudo-metric

$$\rho(x, y) := |F(x) - F(y)|$$

- $\mathbb{B}(\mathbb{X})$ space of bounded measurable functions $f : \mathbb{X}$ with the uniform norm $\|f(\cdot)\|_\infty = \sup_x |f(x)|$.

Theorem 6 Assume that $\mathcal{O}(t^{-\gamma})$ for some $\gamma > 1$. Then

$$\widehat{G}_T = \left\{ \sqrt{T}(\widehat{F}_T(x) - F(x)) : x \in \mathbb{R} \right\} \xrightarrow{law} G, \quad \text{in } (\mathbb{B}(\mathbb{X}))$$

- $G := \{G(x) : x \in \mathbb{R}\}$ Gaussian process
- $\text{cov}[G(x), G(y)] = C(x, y)$
- P-almost all the paths of G are uniformly continuous

Corollary 7

$$\widehat{G}_T = \left\{ \sqrt{T}(\widehat{F}_T(x) - F(x)) : x \in \mathbb{R} \right\} \xrightarrow{law} G, \quad (\mathbb{D}(\mathbb{R}))$$

Corollaire 8

If in addition $F(\cdot)$ is continuous, then

$$\widehat{G}_T = \left\{ \sqrt{T}(\widehat{F}_T(x) - F(x)) : x \in \mathbb{R} \right\} \xrightarrow{law} G, \quad (\mathcal{C}_b(\mathbb{R}))$$

\mathbb{P} -almost all paths of G are uniformly continuous in \mathbb{R}

Sketch for the proof of Theorem 6

$$i) \quad \left| \sqrt{T} \{ \widehat{F}_T(x) - F(x) \} - \sqrt{[T]} \{ \widehat{F}_{[T]}(x) - F(x) \} \right|$$

ii) (Rio 2000)

Theorem (Pollard 1990)

Let (\mathbb{X}, ρ) be a totally-bounded pseudo-metric space, $\{\widehat{G}_n(x) : x \in \mathbb{X}\}$, $n \geq 0$ a sequence of measurable processes

- $(\widehat{G}_n(x_1), \dots, \widehat{G}_n(x_k))$ converges in law $\forall k, x_1, \dots,$
- $\forall \epsilon > 0, \forall \eta > 0, \exists \delta > 0,$

$$\limsup_{n \rightarrow \infty} P^* \left[\sup_{\rho(x,y) < \delta} |\widehat{G}_n(x) - \widehat{G}_n(y)| > \eta \right] < \epsilon$$

Then $(\widehat{G}_n)_n$ converges in law to a probability measure on $\mathcal{U}(\mathbb{X}) \subset \mathcal{B}(\mathbb{X})$ of uniformly continuous functions from \mathbb{X} into \mathbb{R} .

$$- \quad G_N(A) := \sqrt{N} \left\{ \frac{1}{N} \int_0^N \mathbb{I}_{\{X_s \in A\}} ds - \mu(A) \right\} = \frac{1}{\sqrt{N}}$$

$$- \quad \pi_k(x) := F^{-1} \left(2^{-k} \left[2^k F(x) \right] \right) \quad \gamma/(\gamma - 1)$$

$$F(x) - 2^{-k} \leq F\pi_k(x) \leq F\pi_{k+1}(x) \leq F(x)$$

$$- \quad \sup_{\rho(x,y) \leq \delta} |\widehat{G}_N((y, x])| \leq 2 \sup_{x \in \mathbb{R}} |\widehat{G}_N(\pi_k(x), x)]| + \sup_{\rho(x,y) \leq \delta}$$

$$- \quad \sup_{x \in \mathbb{R}} |\widehat{G}_N(\pi_k(x), x)]| \leq \sum_{l=k+1}^n l + \binom{*}{k} \quad \text{where}$$

$$- \quad l = \max_{j=1, \dots, 2^l} |\widehat{G}_N(I_{l,j})| \quad \text{where} \quad I_{l,j} := \left(F^{-1} \left(\frac{j-1}{2^l} \right), F^{-1} \left(\frac{j}{2^l} \right) \right)$$

$$- \quad \mathbb{E} \, l = \int_0^\infty \mathbb{P}\{l > z\} dz \leq c 2^{-l/r} = 4$$

BIBLIOGRAPHY

Billingsley, P. (1968,1995), Wiley.

Castellana, J.V. and Leadbetter, M.R. (1986), *Stochastic*

Davydov, Yu (2001), *Stat. Inference. Stoch. Process.*

Dehay, D. and Kutoyants, Yu.A. (2004), *J. Statist. Plan*

Dehling, H., Mikosch, T. and Sørensen, M (2002), Birkhå

Dudley, R.M. (1984), L.N.M. 1097, Springer.

Ibragimov I.A. and Linnik Yu.V. (1971), Wolters-Noordh

Kutoyants, Yu.A. (2003), Springer.

Meyn, S.P. and Tweedie, R.L. (1993), *Adv. Appl. Prob.*

Negri, I. (1998), *Inference Stoch. Process.*

Pollard, D. (1990), I.M.S. and A.S.A.

Rio, E. (2000), Springer.

Shorack, G.R. and Wellner, J.A. (1986), Wiley.

Veretennikov, A.Yu. (1988), *Theory Probab. Appl.*

Veretennikov, A.Yu. (1999), *Stat. Inf. Stoch. Proc.*