[1]

ON INVARIANT DISTRIBUTION ESTIM
FOR A CLASS OF CONTINUOUS-TI
STATIONARY PROCESSES

Dominique Dehay

dominique.dehay@uhb.fr

Rennes 2 Haute-Bretagne

[1]SASP 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, $\forall t \geq 0$

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

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

Example : diffusion process

$$\mathrm{d}X_t = S(X_t)\mathrm{d}t + \mathrm{d}W_t$$

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

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

Then

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

$$\sigma^2(x) = 2\int_0^\infty \Big(F_t(x,x) - F(x)^2\Big)\,\mathrm{d}t$$

$$= 4f(x)^2 \mathrm{E}\left[\left(\frac{F(\xi \wedge x) - F(\xi)F(x)}{f(\xi)}\right)^2\right]$$

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

– Estimation of the asymptotic variance

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

Then (D. & Kutoyants 2004)

$$\widehat{\sigma}_x^2 \xrightarrow{\mathrm{P}} \sigma_x^2$$

– TLC fonctionnel

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

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

Discrete-time process

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

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

Then

$$\begin{aligned}
&\sqrt{n}\Big(\widehat{F}_n(x) - F(x)\Big) \xrightarrow{loi} \mathcal{N}\big(0, \sigma^2(x)\big) \\
&\sigma^2(x) = F(x)\big(1 - F(x)\big) \\
&\widehat{\sigma}_n^2(x) := \widehat{F}_n(x)\big(1 - \widehat{F}_n(x)\big) \longrightarrow \sigma^2(x) \qquad \text{a.s.}
\end{aligned}$$

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

G Gaussian process (voir Dudley 1984).

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

$\sum_n \alpha_n < \infty$

Then

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

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

Futhermore

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

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

G Gaussian process.

II - Asymptotic independence

(CL*) *Castellana Leadbetter* (1986)

$$\Gamma_T(A) \longrightarrow \Gamma(A), \qquad \forall A \in \mathcal{B}(\mathbb{R}^2) \qquad \text{where}$$

$$\Gamma_T(A) := \int_0^T \Big(\mu_t^{(2)}(A) - \mu \otimes \mu(A) \Big) \, \mathrm{d}t.$$

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

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

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

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

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

Examples

1) Diffusion process (Veretennikov 1998)

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

2) Exponentially ergodic Markov process

Markov process X with

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

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

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

III- Convergence of the empirical distribution

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

$$\operatorname{var} \widehat{F}_T(x) = \frac{2}{T}\int_0^T \Big(1-\frac{t}{T}\Big)\Big(F_t(x,x) - F(x)^2\Big)\,\mathrm{d}$$

Proposition 1 *(Asymptotic covariance)*
Under (CL*),

$$\begin{aligned} C(x,y) &:= \lim_{T\to\infty} T \operatorname{cov}[\widehat{F}_T(x), \widehat{F}_T(y)] \\ &= \int_0^\infty \Big(F_t(x,y) + F_t(y,x) - 2F(x)F(y)\Big)\,\mathrm{d}t. \end{aligned}$$

(Asymptotic variance)

$$\begin{aligned} \sigma^2(x) &:= \lim_{T\to\infty} T \operatorname{var}\big[\widehat{F}_t(x)\big] \\ &= 2\int_0^\infty \Big(F_t(x,x) - F(x)^2\Big)\,\mathrm{d}t = 2\,\Gamma\Big((-\infty,x] \end{aligned}$$

Proposition 2 (Almost-sure convergence)

Assume that one of the two following conditions is satisfi

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

(ii) (CL*).

Then

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

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

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

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

IV - Asymptotic variance estimation

$$\sigma^2(x) = 2\int_0^\infty \Big(F_t(x,x) - F(x)^2\Big)\,\mathrm{d}t.$$

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

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

$$\widehat{\sigma}^2_T(x) := 2\int_0^{T^\eta} \Big(\widehat{F}_{t,T}(x) - \widehat{F}_T(x)^2\Big)\,\mathrm{d}t.$$

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

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

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

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

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

Let

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

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

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

then

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

V - Asymptotic normality

Proposition 5

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

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

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

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

Sketch for the proof of Proposition 5

– x fixed

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

– $Y_n(x) = \int_{n-1}^{n} \mathbb{I}_{\{X_t \leq x\}}\,\mathrm{d}t - F(x)$

– strongly-mixing stationary sequence of bounded zer

– CLT (Ibraginov Linnik 1971)

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

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 $\gamma > 1$. The

$$\widehat{G}_T = \Big\{\sqrt{T}\big(\widehat{F}_T(x) - F(x)\big) : x \in \mathbb{R}\Big\} \xrightarrow{law} G, \quad \text{in} \quad \big(\mathbb{E}$$

– $G := \{G(x) : x \in \mathbb{R}\}$ Gaussian process

– $\operatorname{cov}\big[G(x), G(y)\big] = C(x,y)$

– P-almost all the paths of G are uniformly continuous

Corollary 7

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

Corollaire 8

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

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

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

Sketch for the proof of Theorem 6

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

ii) (Rio 2000)

Theorem (Pollard 1990)

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

– $\left(\widehat{G}_n(x_1),\ldots,\widehat{G}_n(x_k)\right)$ converges in law $\forall k, x_1,\ldots,$

– $\forall\epsilon>0,\forall\eta>0,\exists\delta>0,$

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

Then $\left(\widehat{G}_n\right)_n$ converges in law to a probability measure o trated on the subspace $\mathbb{U}(\mathbb{X})\subset\mathbb{B}(\mathbb{X})$ of uniformly conti functions from $\mathbb{X}$ into $\mathbb{R}$.

- $G_N(A) := \sqrt{N}\Big\{\frac{1}{N}\int_0^N \mathbb{I}_{\{X_s\in A\}}\,\mathrm{d}s - \mu(A)\Big\} = \frac{1}{\sqrt{N}}$

- $\pi_k(x) := F^{-1}\Big(2^{-k}\Big[2^k F(x)\Big]\Big)$

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

- $\sup\limits_{\rho(x,y)\le\delta}\big|\widehat{G}_N\big((y,x]\big)\big| \le 2\,\sup\limits_{x\in\mathbb{R}}\Big|\widehat{G}_N\big(\pi_k(x),x\big]\Big| + \sup\limits_{\rho(x,y)\le}$

- $\sup\limits_{x\in\mathbb{R}}\Big|\widehat{G}_N\big(\pi_k(x),x\big]\Big| \le \sum\limits_{l=k+1}^{n}\Delta_l + \Delta_k^*$ where

- $\Delta_l = \max\limits_{j=1,\dots,2^l}\Big|\widehat{G}_N(I_{l,j})\Big|$ where $I_{l,j} := \Big(F^{-1}\big($

- $\mathrm{E}\Delta_l = \int_0^\infty \mathrm{P}\big\{\Delta_l > z\big\}\,\mathrm{d}z \le c\,2^{-l/r^2}$ (by symmetri

  $r = 4\gamma/(\gamma-1)$

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. Pla*

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

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.*