A 'second Le Cam lemma' for the Ibragimov-Khasminskii limit experiment

Reinhard Höpfner

Johannes Gutenberg Universität Mainz
www.mathematik.uni-mainz.de/∼hoepfner

SAPS IX, Le Mans, March 2013

introduction: some statistical problems in periodic diffusions

parametric model: unknown $\vartheta \in \Theta$, $\Theta \subset \mathbb{R}^d$, observe solution ξ of

$$d\xi_t = [S(\vartheta, t) + b(\xi_t)]\, dt + \sigma(\xi_t) dW_t \quad , \quad t \geq 0$$

continuously in time, over a long time interval (asymptotics $n \to \infty$)

T-periodicity: in the drift, deterministic periodic input gouverned by $\vartheta \in \Theta$

$$t \longrightarrow S(\vartheta, t) = S(\vartheta, i_T(t)) \ , \quad i_T(t) := t \text{ modulo } T$$

aim: asymptotic inference about ϑ

- local models at ϑ, local scale, convergence of local models
- convergence of estimators

under assumptions which provide periodic ergodicity

- positive Harris recurrence of the grid chain $X := (\xi_{kT})_{k \in \mathbb{N}_0}$
- pos. Harris recurrence of the segment chain $\mathbb{X} := ((\xi_{kT+s})_{0 \leq s \leq T})_{k \in \mathbb{N}_0}$

examples A+B: with some 1-periodic function $S_0(\cdot)$, fixed and known,

$$S(\vartheta, t) = S_0\left(\frac{1}{\vartheta}\, t\right) \quad : \quad \text{period } T = \vartheta,\ \Theta = (0, \infty)$$

A: S_0 is continuous (H-K 2009): LAN with local scale $\delta_n(\vartheta) = n^{-3/2}$
in the sense of Le Cam (1969), Hájek (1970), Le Cam and Yang (1990), ...

B: S_0 has discontinuities (H-K 2011): obtain with local scale $\delta_n(\vartheta) = n^{-2}$
convergence of local models at ϑ to Ibragimov-Khasminskii limit experiment

$$\widetilde{\mathcal{E}} := \left\{ \widetilde{P}_u \,:\, u \in \mathbb{R} \right\}$$

(Golubev (1979), Ibragimov and Khasminskii (1981), Rubin and Song (1995), Dachian (2010), ...) with likelihoods of type

$$\widetilde{L}^{u/0} := \frac{d\widetilde{P}_u}{d\widetilde{P}_0} = \exp\left\{ \widetilde{W}_u - \frac{1}{2}|u| \right\} , \; u \in \mathbb{R}$$

where $\widetilde{W} = (\widetilde{W}_u)_{u \in \mathbb{R}}$ is double sided standard Brownian motion experiment $\widetilde{\mathcal{E}}$: unit drift switched off after an unknown amount of time, with separate branches for 'past' and 'future':

on a probability space carrying independent standard BM's $\widetilde{W}^+$ and $\widetilde{W}^-$, create double sided BM from $\widetilde{W}_v := \widetilde{W}^+_v$ if $v \geq 0$, $\widetilde{W}_v := \widetilde{W}^-_{|v|}$ if $v \leq 0$, put

$$\widetilde{P}_0 := \mathcal{L}\left(\left(\widetilde{W}^+_v \,,\, \widetilde{W}^-_v \right)_{v \geq 0} \right) \quad \text{on } C([0,\infty), \mathbb{R}^2)$$

and

$$\widetilde{P}_u \;:=\; \begin{cases} \mathcal{L}\left(\left(\widetilde{W}^+_v + v \wedge u \,,\, \widetilde{W}^-_v \right)_{v \geq 0} \mid \widetilde{P}_0 \right) & \text{for } \; u > 0 \\ \mathcal{L}\left(\left(\widetilde{W}^+_v \,,\, \widetilde{W}^-_v + v \wedge |u| \right)_{v \geq 0} \mid \widetilde{P}_0 \right) & \text{for } \; u < 0 \end{cases}$$

then observation of the bivariate canonical process on $C([0,\infty), \mathbb{R}^2)$ with infinite time horizon $v = +\infty$ leads to likelihood ratios

$$\widetilde{L}^{u/0} \;=\; \frac{d\widetilde{P}_u}{d\widetilde{P}_0} \;=\; \exp\left\{ \widetilde{W}_u - \frac{1}{2}|u| \right\} \,, \quad u \in \mathbb{R}$$

of experiment $\widetilde{\mathcal{E}}$ (use Liptser and Shiryaev 1981, Jacod and Shiryaev 1987) examples C+D: periodicity T fixed and known, not depending on $\vartheta \in \Theta$

C: parametric families $\{S(\vartheta, \cdot) : \vartheta \in \Theta\} \subset C_{\rm per}([0, T])$ [S&D '10]:

we obtain LAN at ϑ with local scale $\delta_n(\vartheta) = n^{-1/2}$ provided parametrization is locally at ϑ smooth in $L^2([0, T], \lambda^{(\vartheta)})$ sense, for a measure $\lambda^{(\vartheta)}$ on $[0, T]$ which involves $\sigma^{-2}(\cdot)$ and the invariant law of the T-segment chain $\mathbb{X}$ under ϑ

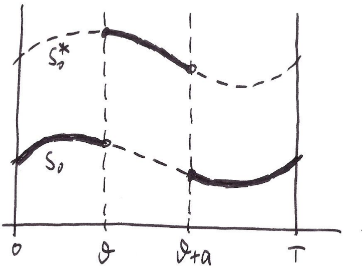
D: $\vartheta \in \Theta$ parametrizing points of discontinuity of $S(\vartheta, \cdot)$ [SISP '10]:

known functions $S_0 < S_0^* \in C_{\rm per}([0, T])$

known $0 < a < T$, put $\Theta := (0, T - a)$

the function $S(\vartheta, \cdot)$ sketched switches from S_0 at time ϑ to S_0^*, and back at time $\vartheta + a$

when $\sigma^2(\cdot)$ is bounded away from zero, with local scale $\delta_n(\vartheta) = n^{-1}$, we obtain convergence of local models at ϑ to an IK limit experiment $\widetilde{\mathcal{E}}$



linking Ibragimov-Khasminskii and Gaussian shift limit experiment

Lévy process $(X_u)_{u \geq 0}$ on $(\Omega, \mathcal{A}, P_0)$, $X_0 = 0$, with LT existing on $\Theta := \mathbb{R}$

$$\vartheta \longrightarrow E_{P_0}(e^{\vartheta X_u}) = e^{u\,\phi_0(\vartheta)} \;, \quad \vartheta \in \Theta \;, \quad u > 0$$

induces two types of one-parametric statistical models:

for $u = 1$ fixed, an exponential family $\{Q^{(1,\vartheta)} : \vartheta \in \mathbb{R}\}$ parametrized by ϑ

$$dQ^{(1,\vartheta)} := e^{\vartheta X_1 - \phi_0(\vartheta)}\, dP_0 \;, \quad \vartheta \in \mathbb{R}$$

for $\vartheta = 1$ fixed, an experiment $\{Q^{(u,1)} : u > 0\}$ parametrized by u

$$dQ^{(u,1)} := e^{X_u - u\,\phi_0(1)}\, dP_0 \;, \quad u > 0$$

of Ibragimov-Khasminkii type; double-sided version parametrized by $u \in \mathbb{R}$ can be defined using an independent copy X^* of the Lévy process X

X BM: have $\phi_0(\vartheta) = +\frac{1}{2}\,\vartheta^2$ and get a link between Gaussian shift (in $\vartheta \in \mathbb{R}$) and Ibragimov-Khasminkii (in $u \in \mathbb{R}$) limit experiment

X PP: have $\phi_0(\vartheta) = \lambda(e^{\vartheta} - 1)$ and get (in $u \in \mathbb{R}$) Poisson approximations to the Ibragimov-Khasminkii limit experiment in the sense of Dachian (2010)

limit theorems in Markov processes with T-periodic semigroup

general setting: $(\xi_t)_{t\geq 0}$ time inhomogeneous strong Markov, càdlàg, Polish state space, $(P_{s,t})_{s<t}$ semigroup of transition probabilities; we assume

$$\underline{T\text{-periodicity}} : \quad P_{s,t}(x,dy) \;=\; P_{s+kT,t+kT}(x,dy) \quad \text{for arbitrary } k\ .$$

need limit theorems as $t \to \infty$ for functionals $A = (A_t)_{t\geq 0}$ of $(\xi_t)_{t\geq 0}$

$$A_t \;=\; \int_0^t f(\xi_s)\,\Lambda_T(ds) \quad \text{or} \quad A_t \;=\; \int_0^t H(s)\,f(s,\xi_s)\,\Lambda_T(ds)$$

with

- σ-finite measures $\Lambda_T(ds)$ which are T-periodic:

$$\Lambda_T(B) \;=\; \Lambda_T(B+kT) \quad \text{for all } k\ ,$$

- measurable nonnegative functions, either $x \to f(x)$ state dependent or $(t,x) \to f(t,x)$ T-periodic in the time variable
- càdlàg nondecreasing fcts $H \in \mathcal{RV}_\rho$ (regular var. at ∞ with index $\rho > 0$)

this class is by far larger than the class of all additive functionals of $(\xi_t)_{t\geq 0}$
allows to filter out periodic phenomena

example: fix r or $r_1 < r_2$ in $[0,T)$, take $H(s) = s^2$, consider

$$\Lambda_T(ds) := 1_{(r_1,r_2)}(i_T(s))\, ds \quad , \quad A_t := \int_0^t s^2\, f(s,\xi_s)\, 1_{(r_1,r_2)}(i_T(s))\, ds$$

$$\Lambda_T(ds) := \sum_{k\in\mathbb{Z}} \epsilon_{(kT+r)}(ds) \quad , \quad A_t := \sum_{k:kT+r\leq t} (kT+r)^2\, f\,(kT+r, \xi_{kT+r})$$

we will obtain limit theorems for $A = (A_t)_{t\geq 0}$ as $t\to\infty$ under a periodic ergodicity condition below, working with the T-segment chain

$$\mathbb{X} = (\mathbb{X}_k)_k \quad : \quad \mathbb{X}_k := (\xi_{kT+s})_{0\leq s\leq T} \ , \ k\in\mathbb{N}_0 \ , \quad D([0,T])\text{-valued}$$

and with the T-grid chain

$$X = (X_k)_k \quad : \quad X_k := \xi_{kT} \ , \ k\in\mathbb{N}_0$$

which both are time-homogeneous (by T-periodic semigroup of $(\xi_t)_{t\geq 0}$)

assumption : (periodic ergodicity condition)

(H) the T-grid chain X is positive Harris recurrent with invariant prob. μ

(H-Löcherbach 2010): a sufficient condition for (H) is

- $x\,b(x) + \frac{1}{2}\sigma^2(x) + (x^- \max S^- + x^+ \max S^+) \;<\; -\varepsilon$ outside K_1
- $\sigma^2(x) > 0$ on some K_2 which is large in comparison to K_1

which allows for Nummelin splitting in the grid chain

theorem 1: (H-K 2010) condition (H) is equivalent to

the segment chain $\mathbb{X}$ is positive Harris recurrent

then for $\mathbb{X}$, the invariant probability m on $D([0,T])$ is specified by

$$\begin{cases} \text{for arbitrary } 0 = t_0 < t_1 < \ldots < t_\ell < t_{\ell+1} = T \text{ and arbitrary } A_i \;, \\ m\left(\{\pi_{t_i} \in A_i \,,\, 0 \le i \le \ell{+}1\}\right) \quad \text{is given by} \\ \int \ldots \int \mu(dx_0)\, 1_{A_0}(x_0) \prod_{i=0}^{\ell} P_{t_i,t_{i+1}}(x_i, dx_{i+1})\, 1_{A_{i+1}}(x_{i+1}) \;. \end{cases}$$

theorem 2 : under (H), for increasing processes $A = (A_t)_{t\geq 0}$ such that

$$\begin{cases} \text{there is an } m\text{-integrable function } F : D([0,T]) \to [0,\infty) \\ \text{of form} \quad F(\alpha) \;=\; \int_0^T \Lambda_T(ds)\, f(s,\alpha(s)) \quad \text{as above} \\ \text{such that} \quad A_{kT} \;=\; \sum_{j=1}^k F(\mathbb{X}_j) \quad \text{for all } k=1,2,\ldots \end{cases}$$

we have
i) almost sure convergence as $t \to \infty$

$$\frac{1}{t} A_t \;\longrightarrow\; \frac{1}{T} m(F) \;=\; \frac{1}{T} \int_0^T \Lambda_T(ds) \int [\mu P_{0,s}](dy)\, f(s,y)$$

ii) for every $H \in \mathcal{RV}_\rho$, $\rho > 0$, almost sure convergence as $t \to \infty$

$$\frac{1+\rho}{t\,H(t)} \int_0^t H(s)\, dA_s \;\longrightarrow\; \frac{1}{T} m(F)$$

(H-K 2010, 2011), for regular variation: Bingham, Goldie and Teugels (1987)

convergence of local models: $S(\vartheta, t) = S_0\left(\frac{1}{\vartheta}\, t\right)$, example B

assumptions: • $S_0(\cdot)$ fixed and known, 1-periodic, piecewise continuous, jumps of height ρ_j at times r_j, $0 < r_1 < \ldots < r_\ell < 1$

• $\sigma(\cdot) > 0$ Lipschitz, bounded, bounded away from 0; $b(\cdot)$ Lipschitz

• periodic ergodicity condition (H) holds for solution of

$$(*) \qquad d\xi_t = \left[S_0\left(\frac{t}{\vartheta}\right) + b(\xi_t)\right] dt + \sigma(\xi_t) dW_t \quad , \quad t \geq 0$$

for every $\vartheta \in \Theta$, $\Theta = (0, \infty)$: unknown parameter ϑ is length of the period

notations: • P^ϑ law of solution to $(*)$ on canonical path space $(C, \mathcal{C}, \mathbb{F})$

• $\eta = (\eta_t)_{t \geq 0}$ canonical process on $(C, \mathcal{C}, \mathbb{F})$, $m^{(\vartheta)}$ its P^ϑ-martingale part

• $L^{\zeta/\vartheta}$ likelihood ratio process of P^ζ to P^ϑ relative to $\mathbb{F}$:

$$L^{\zeta/\vartheta} = \mathcal{E}_\vartheta\left(\int_0^{\cdot} \frac{S_0(\frac{t}{\zeta}) - S_0(\frac{t}{\vartheta})}{\sigma^2(\eta_t)}\, dm_t^{(\vartheta)}\right) = \mathcal{E}_\vartheta\left(\int_0^{\cdot} \frac{S_0(\frac{t}{\zeta}) - S_0(\frac{t}{\vartheta})}{\sigma(\eta_t)}\, dB_t^{(\vartheta)}\right)$$

with P^ϑ-Brownian motion $B^{(\vartheta)}$

localization at ϑ: local scale $\delta_n(\vartheta) = n^{-2}$, observe $(\eta_t)_t$ on $[0, n]$, local model:

$$\left\{ P^{\vartheta + h/n^2} \mid \mathcal{F}_n : h \in \Theta_{\vartheta,n} \right\} \quad \text{parametrized by the local parameter } h$$

write $L_n^{(\vartheta+h/n^2)/\vartheta}$ likelihood of local parameter h to local parameter 0

theorem 3: (H-K 2011) as $n \to \infty$, we have convergence of local models at ϑ

$$\left(L_n^{(\vartheta+h/n^2)\,/\,\vartheta} \right)_{h \in \Theta_{\vartheta,n}} \quad , \quad \Theta_{\vartheta,n} := \{ h : \vartheta + h/n^2 > 0 \} \; ,$$

$$\log L_{\bullet n}^{(\vartheta+h/n^2)\,/\,\vartheta} \; = \; \mathcal{E}_\vartheta \left(\int_0^{\bullet n} \left[S_0\Big(\frac{s}{\vartheta + h/n^2}\Big) - S_0\Big(\frac{s}{\vartheta}\Big) \right] \frac{1}{\sigma(\eta_s)} dB_s^\vartheta \right)$$

in the sense of finite dimensional distributions to

$$\left(\widetilde{L}^{h/0} \right)_{h \in \mathbb{R}} \quad , \quad \widetilde{L}^{h/0} = \exp \left\{ \widetilde{W}(hJ) - \frac{1}{2} |hJ| \right\} \; , \; h \in \mathbb{R}$$

with some scaling factor $J := \frac{1}{2\vartheta^2} \sum_{j=1}^{\ell} \rho_j^2 \, [\, \mu^{(\vartheta)} \, P_{0, r_j \vartheta}^{(\vartheta)} \,](\frac{1}{\sigma^2})$ which depends on $\sigma(\cdot)$, the sequence of jump times/heights of S_0, and on the semigroup under ϑ

we shall deduce this theorem from a stronger assertion which has the structure of a '2nd Le Cam lemma' for the Ibragimov-Khasminskii limit experiment

notations: $S_0(\cdot)$ has jumps of height ρ_j at times r_j, $0 < r_1 < \ldots < r_\ell < 1$, and period 1; we define deterministic counting functions

$$N_j(s) \ := \sum_{k=1}^{\infty} 1_{[k+r_j,\infty)}(s) \quad , \quad s \geq 0 \, , \; 1 \leq j \leq \ell$$

fix $\vartheta > 0$ and put (written here for case $h < 0$, notation similiar for $h > 0$)

$$\Phi_\vartheta^{h,n}(s) \ := \sum_{j=1}^{\ell} \rho_j \left[N_j\left(\frac{s}{\vartheta + h/n^2}\right) - N_j\left(\frac{s}{\vartheta}\right) \right] \quad , \quad 0 \leq s \leq n$$

structure in $0 \leq s \leq n$ provided n is large enough: for every j, $[\ldots j \ldots]$ is a sum of indicator functions for <u>disjoint intervals of increasing length</u>, and

$$\left[S_0\left(\frac{s}{\vartheta + h/n^2}\right) - S_0\left(\frac{s}{\vartheta}\right) \right] \ \rightleftharpoons \ \Phi_\vartheta^{h,n}(s) \quad , \quad 0 \leq s \leq n$$

rough structural approximation

closer look to summands contributing to $\Phi_{\vartheta}^{h,n}(s)$, $0 \le s \le n$:

$$\left[N_j\left(\frac{s}{\vartheta + h/n^2}\right) - N_j\left(\frac{s}{\vartheta}\right) \right] = \sum_{k=0}^{\infty} 1_{\underbrace{((\vartheta + h/n^2)(k + r_j), \vartheta(k + r_j)]}_{\text{length } \approx |h| \frac{k}{n^2}}}(s)$$

in notation for $h < 0$ (analogous for $h > 0$)

in restriction to $s \in [0, n]$ when n large enough: intervals above are

- 'very small' and always located <u>left of $\vartheta(k+r_j)$</u>
- non-overlapping in $1 \le k \ne k' \le O(\frac{n}{\vartheta})$ in every $[\ldots j, h \ldots]$
- non-overlapping between $[\ldots j, h \ldots]$ and $[\ldots j', h \ldots]$ when $j \ne j'$
- when $-\infty < h_2 < h_1 < 0$, intervals in $[\ldots j, h_1 \ldots]$ are contained in the corresponding intervals in $[\ldots j, h_2 \ldots]$

case $h > 0$: analogous with intervals located <u>right of $\vartheta(k+r_j)$</u>

asymptotically as $n \to \infty$, the local parameter $h \in \mathbb{R}$ thus makes appear the covariance structure of two-sided Brownian motion $(W_h)_{h \in \mathbb{R}}$

$$\begin{aligned} K(h_1, h_2) &= |h_1| \wedge |h_2| \quad \text{if } \operatorname{sgn}(h_1) = \operatorname{sgn}(h_2) \\ K(h_1, h_2) &= 0 \qquad \text{if } \operatorname{sgn}(h_1) \neq \operatorname{sgn}(h_2) \end{aligned}$$

lemma: we have convergence in P^{ϑ}-probability as $n \to \infty$

$$\int_0^{t \cdot n} \Phi_{\vartheta}^{h_1, n}(s)\, \Phi_{\vartheta}^{h_2, n}(s)\, \frac{1}{\sigma^2(\eta_s)}\, ds \quad \longrightarrow \quad K(h_1, h_2)\, t^2\, J$$

for every $0 \leq t \leq 1$ fixed, for all h_1, h_2 in $\mathbb{R}$

t^2 arises when summing up disjoint support intervals of successive length $O(|h| \frac{k}{n^2})$, $k = 1, 2, \ldots, O(\frac{tn}{\vartheta})$ (which explains choice of local scale !)

J is the constant of theorem 3, involving the ρ_j and r_j and $\sigma(\cdot)$

consider now any finite collection of points $h_1, \ldots, h_r$ in $\mathbb{R}$

convergence in P^ϑ-probability as $n \to \infty$

$$\int_0^{t \cdot n} \Phi_\vartheta^{h_1,n}(s)\, \Phi_\vartheta^{h_2,n}(s)\, \frac{1}{\sigma^2(\eta_s)}\, ds \quad \longrightarrow \quad K(h_1, h_2)\, t^2\, J$$

allows to apply the martingale convergence theorem (martingales in time $0 \leq t \leq 1$) to prove weak convergence

$$\begin{pmatrix} \int_0^{\bullet n} \Phi_\vartheta^{h_1,n}(s)\, \frac{1}{\sigma^2(\eta_s)}\, dB_s^{(\vartheta)} \\ \ldots \\ \int_0^{\bullet n} \Phi_\vartheta^{h_r,n}(s)\, \frac{1}{\sigma^2(\eta_s)}\, dB_s^{(\vartheta)} \end{pmatrix} \longrightarrow \begin{pmatrix} \widehat{B}^{h_1} \\ \ldots \\ \widehat{B}^{h_r} \end{pmatrix}$$

to a high-dimensional limit process $\widehat{B}$, Gaussian in time t, independent of $B^{(\vartheta)}$ and thus of W, which in $h_1, \ldots, h_r$ has the covariance structure

$$E\left(\widehat{B}_t^{h_i}, \widehat{B}_t^{h_j}\right) = K(h_1, h_2)\, t^2\, J$$

of two-sided Brownian motion $(\widetilde{W}(hJ))_{h \in \mathbb{R}}$ with scaling constant J

the following is our 'second Le Cam lemma' for convergence to the Ibragimov Khasminskii limit experiment:

theorem 4: (H-K 2011) in local models at ϑ we have as $n \to \infty$

i) a stochastic approximation valid for bounded sequences $(h_n)_n$ in $\mathbb{R}$:

$$\log L_n^{(\vartheta+h_n/n^2)\,/\,\vartheta} = \int_0^n \Phi_\vartheta^{h_n,n}(s)\,\frac{1}{\sigma(\eta_s)}\,dB_s^{(\vartheta)} - \frac{1}{2}\int_0^n [\Phi_\vartheta^{h_n,n}(s)]^2\,\frac{1}{\sigma^2(\eta_s)}\,ds + o_{P^{(\vartheta)}}(1)$$

ii) finite dimensional convergence of

$$\left(\int_0^n \Phi_\vartheta^{h,n}(s)\,\frac{1}{\sigma(\eta_s)}\,dB_s^{(\vartheta)} - \frac{1}{2}\int_0^n [\Phi_\vartheta^{h,n}(s)]^2\,\frac{1}{\sigma^2(\eta_s)}\,ds\right)_{h\in\Theta_{\vartheta,n}} \quad \text{under } P^{(\vartheta)}$$

as $n \to \infty$ to the likelihood function

$$\left(\widetilde{W}(hJ) - \frac{1}{2}\,|hJ|\right)_{h\in\mathbb{R}}$$

of the IK limit experiment, with scaling constant J from theorem 3

selected references

Dachian, S.: On limiting likelihood ratio processes of some change-point type statistical models. JSPI 140, 2010.

Golubev, G.: Computation of efficiency of maximum-likelihood estimate when observing a discontinuous signal in white noise. PIT 15, 1979.

Höpfner, R., Kutoyants, Y.: Estimating a periodicity parameter in the drift of a time inhomogeneous diffusion. MMS 13, 58–74 (2011)

Höpfner, R., Kutoyants, Y.: Estimating discontinuous periodic signals in a time inhomogeneous diffusion. SISP 20, 193–230 (2010)

Höpfner, R., Kutoyants, Y.: On frequency estimation for a periodic ergodic diffusion process. PIT 48, 2012.

Höpfner, R., Löcherbach, E.: On ergodicity properties for time inhomogeneous Markov processes with T-periodic semigroup. `arXiv:1012.4916`, subm. TPA

Ibragimov, I., Has'minskii, R.: Statistical estimation. Springer 1981.

Rubin, H., Song, K.: Exact computation of the asymptotic efficiency of maximum likelihood estimators of a discontinuous signal in a Gaussian white noise. AS 23, 1995.