Extremes of Gaussian non-stationary processes and honest confidence bands for densities

Vladimir Panov

Joint work with Valentin Konakov and Vladimir Piterbarg

Higher School of Economics (Moscow, Russia)

LSA Winter Meeting
Snegiri, 4. December 2019

Honest confidence sets

Given: $X_1, ...X_n \sim f$,
f is an infinitely dimensional object (distribution function, density, ...).

Definition. (*Li, 1989*) Given any $\alpha \in (0,1)$, we aim to construct $(1-\alpha)$-confidence set $\mathcal{C}_n(x)$ for f that are *honest* to a given class $\mathcal{F}$ of functions in the sense

$$\sup_{f \in \mathcal{F}} \mathbb{P}\Big\{ f(x) \notin \mathcal{C}_n(x),\ x \in \mathbb{R} \Big\} \leq \alpha + e_n,$$

where $e_n \to 0$ as $n \to \infty$.

Honest / dishonest

$$\limsup_{n\to\infty} \sup_{f\in\mathcal{F}} \mathbb{P}\left\{ f \notin \mathcal{C}_n \right\} \leq \alpha \quad - \quad \text{honest},$$

$$\sup_{f\in\mathcal{F}} \limsup_{n\to\infty} \mathbb{P}\left\{ f \notin \mathcal{C}_n \right\} \leq \alpha \quad - \quad \text{dishonest}.$$

Dvoretsky-Kiefer-Wolfowitz inequality (in strong form - *Massart, 1990*)

$$\mathbb{P}\Big\{ \sqrt{n} \sup_{u \in \mathbb{R}} \Big| \hat{F}_n(u) - F(u) \Big| > x \Big\} \leq 2e^{-2x^2}.$$

Confidence bands for densities

Maximal deviation: for an estimate $\hat{p}_n$ of the density function p, denote

$$\mathcal{D}_n := \sup_{u \in \mathbb{R}} \frac{\left|\hat{p}_n(u) - p(u)\right|}{\sqrt{p(u)}}.$$

SBR-type ("Smirnov-Bickel-Rosenblatt") limit theorems

$$\boxed{\sup_{p \in \mathcal{F}} \left| \mathbb{P}\left\{ \mathcal{D}_n \leq \frac{x}{a_n} + b_n \right\} - e^{-e^{-x}} \right| \to 0}$$

for some sequences a_n and b_n tending to infinity as $n \to \infty$.

Smirnov (1950), Bickel and Rosenblatt (1973), Konakov and Piterbarg (1984), Giné, Koltchinskii, Sakhanenko (2004), Giné and Nickl (2010), Bull (2012).

Giné and Nickl (2016). Mathematical foundations of of infinitely-dimensional statistical models.

Challenges

1. SBR-type theorems are known only for kernel density estimate and *certain* wavelet projection density estimates like Haar wavelets or Battle-Lemarie wavelets.

   Chernozhukov, Chetverikov and Kato (Annals of Statistics, 2014):
   ... the SBR condition has not been obtained for other density estimators such as nonwavelet projection kernel estimators based, for example, on Legendre polynomials or Fourier series.

2. The rates of convergence.

   Giné and Nickl (Annals of Statistics, 2010):
   ... we finally remark that the results in this article are clearly of an asymptotic (and hence "theoretical" nature).

Key ingredient

Most estimates can be represented as

$$\hat{p}_n(x) = \int_{\mathbb{R}} K(x,y) d\mathbb{P}_n(y).$$

Komlós - Major -Tusnady construction: the analysis of $\mathcal{D}_n$ leads to the study of asymptotic behaviour of the Gaussian process

$$\Upsilon(x) = \int_{\mathbb{R}} K(x,y) dW(y).$$

Examples:

1. Kernel density estimates: $K(x,y) = \mathcal{K}((x-y)h^{-1})/h^{-1}$.
   $\Upsilon(x)$ is a stationary process.
2. Wavelets: $K(x,y) = 2^j \sum_k \phi(2^j x - k)\phi(2^j y - k)$, $j \in \mathbb{N}$.
   Then $\Upsilon(x)$ is a nonstationary sprocess of some special type: $r(x, x+u)$ is periodic in x with the same period for any u.
   Cyclostationary processes: *Konstant and Piterbarg (1993), Hüsler, Piterbarg, and Seleznjev (2003)*

Two types of projection estimates

Let $\Psi := \{\psi_0, \psi_1, \psi_2, ...\}$ be an orthonormal basis of $L^2([A, B])$.

1. Consider with $J \to \infty$

$$\hat{p}_n(x) = \sum_{j=0}^{J} \left[\int \psi_j(y) d\mathbb{P}_n(y) \right] \psi_j(x).$$

2. Let us divide $I := [A, B]$ on M subintervals, and on each subinterval $I_m = [a_m, b_m] := [A + \delta(m-1), A + \delta m]$, $m = 1..M$, we reproduce Ψ:

$$\psi_j^{(m)}(x) = \sqrt{M} \cdot \psi_j \left(M(x - a_m) + A \right), \quad m = 1..M \;\; j = 0, 1, ...$$

Consider with $M \to \infty$

$$\hat{p}_n(x) = \sum_{m=1}^{M} \sum_{j=0}^{J} \left[\int \psi_j^{(m)}(y) d\mathbb{P}_n(y) \right] \psi_j^{(m)}(x).$$

In both cases,

$$\Upsilon(x) = \int_I \left(\sum_{j=0}^{J} \psi_j(x) \psi_j(y) \right) dW(y).$$

Building bridge to the Gaussian process $\Upsilon(x)$

$$\mathcal{P}_{q,H,\beta} := \Big\{ p - \text{ p.d.f. }, \quad p \in L^2([A,B]), \quad \inf_{x\in[A,B]} p(x) \geq q, \quad \big|p\big|_\beta \leq H \Big\},$$

where $\big|p\big|_\beta := \sup_{x \neq y} \frac{|p(x)-p(y)|}{|x-y|^\beta}$ is the Hölder coefficient of the function p.

Theorem

There exists a positive constant κ such that for any $p \in \mathcal{P}_{q,H,\beta}$ and any $u \in \mathbb{R}$ it holds

$$\mathbb{P}\Big\{ \sqrt{\frac{n}{M_n}} \mathcal{R}_n \leq u \Big\} \quad \leq \quad \Big[\mathbb{P}\Big\{ \zeta \leq u + \gamma_{n,M} \Big\} \Big]^{M_n} + \mathcal{C}_1 n^{-\kappa},$$

$$\mathbb{P}\Big\{ \sqrt{\frac{n}{M_n}} \mathcal{R}_n \leq u \Big\} \quad \geq \quad \Big[\mathbb{P}\Big\{ \zeta \leq u - \gamma_{n,M} \Big\} \Big]^{M_n} - \mathcal{C}_1 n^{-\kappa},$$

where $\zeta := \sup_{x\in\mathbb{R}} |\Upsilon(x)|$, $\gamma_{n,M} = \mathcal{C}_2 \frac{\log(n)}{\sqrt{n/M_n}} + \mathcal{C}_3 \frac{\sqrt{\log(n)}}{\sqrt{M_n}}$, and $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3 > 0$ depend on q, H, β.

Supremum of non-stationary Gaussian process

Pickands and Berman methods
Hashorva, Hüsler (2000), Hüsler, Piterbarg, Rumyantseva (2011), Piterbarg, Popivoda, Stamatović (2017), Bai, Dębicki, Hashorva (2018)

Drawback: no results on the asymptotics of the remainder term.

Standard result: let the variance $\sigma^2(t) = \mathrm{Var}(X(t))$ attain its maximum S in a finite number of points $t_\circ^{(1)}, ..., t_\circ^{(K)}$. Assume that for any $k = 1..K$,

$$\rho(t,s) = 1 - A_k |t-s|^{\alpha_k} (1 + o(1)), \qquad \text{as} \quad t \to t_\circ^{(k)},\ s \to t_\circ^{(k)},$$
$$\sigma(t) = 1 - B_k \left| t - t_\circ^{(k)} \right|^{\beta_k} (1 + o(1)), \qquad \text{as} \quad t \to t_\circ^{(k)}.$$

Then

$$\mathbb{P}\left\{ \sup_t X(t) \geq u \right\} = C u^{b-1} \exp\Big(-\frac{u^2}{2S} \Big) \Big(1 + o(1) \Big),$$

where $b = \min_{k=1..K} (2/\beta_k - 2/\alpha_k)_-$.

Rice's method of moments. No papers in mathematical journals.

Main result

Theorem

Let $X(t)$ be a centered Gaussian process with a.s. twice differentiable trajectories. Assume that the variance $\sigma^2(t) = \mathrm{Var}(X(t))$ reaches its maximum S at only one point $t_\circ \in [A, B]$. For $\delta > 0$, introduce the informative set

$$\mathcal{M}(\delta) := \left\{ t \in [A, B] : \sigma^2(t) \geq \frac{S}{1+\delta} \right\}.$$

Then there exists some $\chi > 0$ (depending on δ and the process X_t) such that

$$\mathbb{P}\left\{ \max_{t\in[A,B]} |X(t)| \geq u \right\} = \mathscr{P}(u) + O\left(e^{-u^2(1+\chi)/(2S)} \right), \qquad u \to \infty,$$

where

$$\mathscr{P}(u) = \begin{cases} 2\mathbb{E}\left[N_u^+(\mathcal{M}(\delta)) \right], & \text{if } t_0 \in (A, B], \sigma'(t_0) = 0, \\ 2\mathbb{P}\left\{ X(t_0) \geq u \right\} + 2\mathbb{E}\left[N_u^+(\mathcal{M}(\delta)) \right], & \text{if } t_0 = A, \sigma'(t_0) = 0, \\ 2\mathbb{P}\left\{ X(t_0) \geq u \right\}, & \text{if } t_0 = A \text{ or } t_0 = B \text{ and } \sigma'(t_0) \neq 0. \end{cases}$$

Some remarks to the main result

Standard setup: $\operatorname{argmax} \sigma^2(t) = \{t_\circ^{(1)}, \ldots, t_\circ^{(K)}\}$. Let us choose disjoint intervals $\mathcal{M}_i, \; t_\circ^{(i)} \in \mathcal{M}_i, \; i = 1, \ldots, K$:

$$\max_{(s,t) \in \mathcal{M}_i \times \mathcal{M}_j} \rho(s,t) < 1, \qquad \forall i,j = 1..k, \; i \neq j.$$

Then there exists some $\chi > 0$ such that

$$\mathbb{P}\Big\{ \max_{t \in [A,B]} \big|X(t)\big| \geq u \Big\} = \sum_{i=1}^{k} \mathscr{P}_i(u) + O\Big(e^{-u^2(1+\chi)/(2S)} \Big),$$

where $\mathscr{P}_i(u)$ are defined above for $\mathbb{P}\Big\{ \max_{t \in \mathcal{M}_i} \big|X(t)\big| \geq u \Big\}$.

Example: for the process $\Upsilon(x) = \int_I \Big(\sum_{j=0}^{J} \psi_j(x)\psi_j(y) \Big) \, dW(y)$, there exists some $\chi > 0$ such that

$$\mathbb{P}\Big\{ \max_{t \in [-1,1]} \big|\Upsilon(t)\big| \geq u \Big\} = 4\Big(1 - \Phi\big(u/\sqrt{S}\big)\Big) + O\Big(e^{-u^2(1+\chi)/(2S)} \Big).$$

Dependence between χ and δ

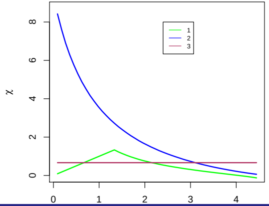
Let δ be a small number such that $\mathcal{M}(\delta) \cap [-1, 0] = [-1, b]$.
Then $\chi < \min\left(\chi_1(\delta), \chi_2(\delta), \chi_3(\delta)\right)$, where

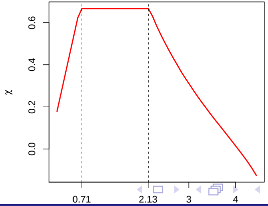
$$\chi_1(\delta) := \min\left\{\delta, \frac{4}{(b-A)J(J+2)} - 1\right\},$$

$$\chi_2(\delta) := \min_{t\in[-1,b]} \left(r_{10}^2(t,t)/r_{11}(t,t)\right),$$

$$\chi_3(\delta) := \begin{cases} J/(J+2), & J \text{ is even}, \\ (J+2)/J, & J \text{ is odd}. \end{cases}$$

Empirical result for $J = 4$: $\chi_{max} = 2/3$, for any $\delta \in (0.71, 2.13)$.





Return to statistical problem

Recall that we consider the estimate

$$\hat{p}_n(x) = \sum_{m=1}^{M_n} \sum_{j=0}^{J} \left[\int \psi_j^{(m)}(y) d\mathbb{P}_n(y) \right] \psi_j^{(m)}(x),$$

where $\psi_j^{(m)}(x) = \sqrt{M_n} \cdot \psi_j \left(M_n(x - a_m) + A \right), \quad m = 1..M_n \;\; j = 0, 1, \ldots.$

Theorem

Let $M_n = \lfloor n^\lambda \rfloor$ *with* $\lambda \in (1/3, 1)$. *For any* $x \in \mathbb{R}$, *it holds*

$$\mathbb{P}\left\{ \sqrt{\frac{n}{M_n}} \mathcal{D}_n \leq u_M(x) \right\} = e^{-e^{-x}} \left(1 + e^{-x} \Lambda_M (1 + o(1)) \right),$$

as $n \to \infty$, *where*

$$\Lambda_M = \frac{\left(\log\log(M_n)\right)^2}{16 \log(M_n)}.$$

$$u_M(x) = a_M + \frac{xS}{a_M}, \qquad a_M \asymp \left(2S \log(M_n)\right)^{1/2}.$$

Improvement of convergence rates

Theorem

Assume that $p \in \mathcal{P}_{q,H,\beta}$ *with some* $q, H > 0$, $\beta \in (0,1]$. *Denote the sequence of distribution functions*

$$A_M(x) := \begin{cases} \exp\left\{-M_n \sum_{i=1}^{k} \mathscr{P}_i(x)\right\}, & \text{if } x \geq c_M, \\ 0, & \text{if } x < c_M, \end{cases}$$

where $c_M = (2S \log M_n)^{1/2} - S$. *If* $M = \lfloor n^\lambda \rfloor$ *with* $\lambda \in ((2\beta+1)^{-1}, 1)$, *then*

$$\boxed{\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left\{ \sqrt{\frac{n}{M_n}} \mathcal{D}_n \leq x \right\} - A_M(x) \right| \leq \bar{c} n^{-\gamma}.}$$

for some positive constants $\bar{c}$ *and* γ.

Honest confidence bands

Denote

$$k_{\alpha,M} := \sqrt{M_n/n} \cdot q_{\alpha,M},$$

where $q_{\alpha,M}$ is the $(1-\alpha)$- quantile of the distribution function $A_M(\cdot)$.

Then

$$\mathbb{P}\left\{ \frac{|\hat{p}_n(x) - p(x)|}{\sqrt{p(x)}} \le k_{\alpha,M}, \quad \forall x \in I \right\} = 1 - \alpha + e_{n,M},$$

where $e_{n,M}$ converges to zero at polynomial level in both n and M.

Outcome:

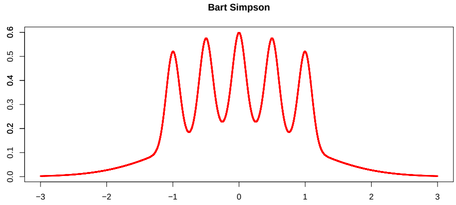
$$\mathcal{C}_n(x) := \Big(\hat{p}_n(x) + (k^2_{\alpha,M}/2) - \left[\hat{p}_n(x) k^2_{\alpha,M} + (k^4_{\alpha,M}/4)\right]^{1/2}, \\ \hat{p}_n(x) + (k^2_{\alpha,M}/2) + \left[\hat{p}_n(x) k^2_{\alpha,M} + (k^4_{\alpha,M}/4)\right]^{1/2} \Big)$$

is $(1-\alpha)-$**confidence set, which is honest to a class** $\mathcal{P}_{q,H,\beta}$ **at polynomial rate**.

Numerical example

Consider the density (Bart Simpson density; the claw)

$$p(x) = \frac{1}{2}\phi_{(0,1)}(x) + \frac{1}{10}\sum_{j=0}^{4}\phi_{((j/2)-1,1/100)}(x),$$



Approximate the distribution of $\mathcal{D}_n = \sup_{u\in\mathbb{R}} |\hat{p}_n(u) - p(u)|/\sqrt{p(u)}$ via

$$A_M(x) := \exp\left\{-4M\left(1-\Phi\left(\sqrt{6}x/(J+1)\right)\right)\right\} \cdot I\{x \geq c_M\}$$

where $c_M = \frac{(J+1)}{\sqrt{3}}\sqrt{\log(M)} - \frac{(J+1)^2}{6}$.

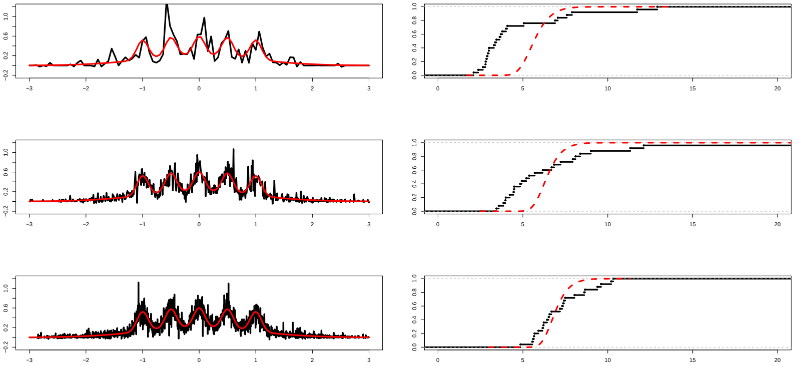


Рис.: First raw: projection density estimates (black solid lines) in comparison withth the true densities (red lines) based on $n = 500, 3000, 10000$ simulations. In this example, we take $M_n = \lfloor n^{2/3} \rfloor$. Second raw: empirical distribution functions of $\sqrt{n/M_n} \cdot \mathcal{D}_n$ (black solid lines) based on 25 simulation runs in comparison with the distribution function $A_M(x)$.

Summary

1. Extremes of Gaussian non-stationary processes: a new theoretical result revealing the asymptotic behaviour of Gaussian processes, **which are neither stationary nor cyclostationary**.
2. Sequence of accompanying laws, which approximates the distribution of maximal deviation at polynomial rate. Confidence sets, which are honest to some classes of densities **at polynomial rate**.
3. Several references:

- Piterbarg, V. *Twenty lectures about Gaussian processes.* Atlantic Financial Press. 2015.
- Konakov, V. and Panov, V. *Sup-norm convergence rates for Levy density estimation.* Extremes. 2016. Vol. 19. No. 3. P. 371-403
- Konakov, V., Panov, V. and Piterbarg, V. *Extremes of Gaussian non-stationary processes and honest confidence bands for densities.* To appear soon on ArXiv.

Thank you for your attention.