ulm university universität ulm

Nonparametric estimation of entropy in Poisson marked point processes

Patricia Alonso Ruiz
joint work with E. Spodarev | 02.03.2016
Institut of Stochastics, Ulm University

GPSD Bochum 2016

Motivation

Reinforced plastic:

- short fibers;
- no intersections;
- strongly anisotropic;
- locally stationary.

Motivation

Reinforced plastic:

- short fibers;
- no intersections;
- strongly anisotropic;
- locally stationary.

Motivation

Reinforced plastic:

- short fibers;
- no intersections;
- strongly anisotropic;
- locally stationary.

$\rightsquigarrow$ Detection of clumps, deformations, production damages.

Measuring homogeneity - Kolmogorov entropy

Let (M, g) be a complete Riemannian manifold of dimension $p > 0$ without boundary. The entropy of a r.v. $\xi : \Omega \to M$ with density distribution f_ξ is given by

$$\mathcal{E}_{f_\xi} := -\mathbb{E}[\log f_\xi(\xi)] = -\int_M \log f_\xi(s) f_\xi(s) \upsilon_g(ds).$$

Measuring homogeneity - Kolmogorov entropy

Let (M, g) be a complete Riemannian manifold of dimension $p > 0$ without boundary. The entropy of a r.v. $\xi : \Omega \to M$ with density distribution f_ξ is given by

$$\mathcal{E}_{f_\xi} := -\mathbb{E}[\log f_\xi(\xi)] = -\int_M \log f_\xi(s) f_\xi(s) \upsilon_g(ds).$$

lower entropy $\rightsquigarrow$ higher homogeneity $\rightsquigarrow$ possible clump

Geometric model – assumptions

Marked Poisson point process (MPPP)

$$\Psi := \{(Y_i, \xi_i)\}_{i \geq 1} \subset \mathbb{R}^d \times M.$$

- $\Pi := \{Y_i\}_{i \geq 1}$ homogeneous PPP of intensity $\lambda > 0$;
- the mark process $\{\xi_i\}_{i \geq 1}$ is independent of Π;
- ξ_i are i.i.d. r.v.'s;
- the distribution function of ξ_i has a density $f \colon M \to \mathbb{R}$;
- Ψ is observed in a window $W \subset \mathbb{R}^d$.

Example: fiber system

Boolean model: For each $Y \in \Pi$, $F_Y \subset \mathbb{R}$ is an independent copy of a segment of finite length $\ell > 0$,

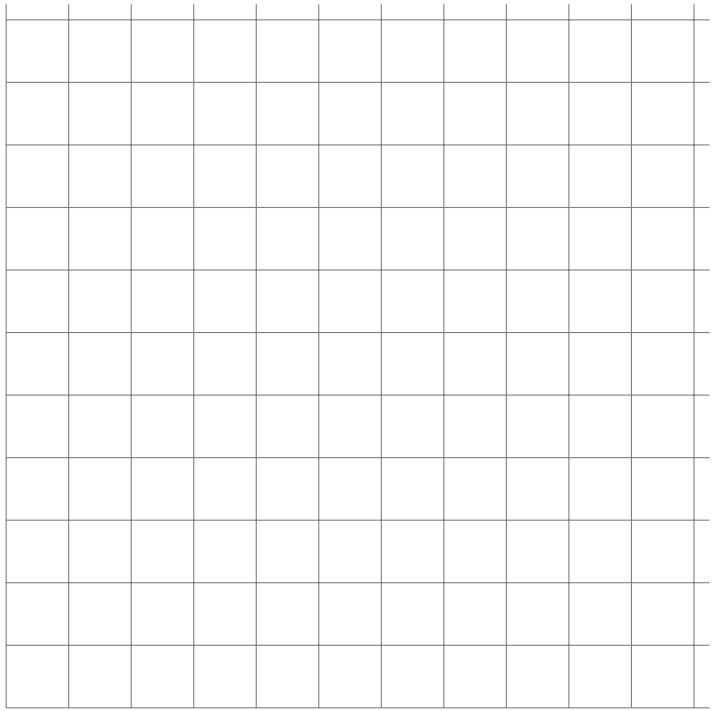
$$\Phi := \bigcup_{Y \in \Pi} Y \oplus F_Y.$$

Associated MPPP: The direction of a fiber F_Y is given by $\xi_Y : \Omega \to S^{d-1}$. Define

$$\{(Y, \xi_Y)\}_{Y \in \Pi} \subset \mathbb{R}^d \times S^{d-1}.$$

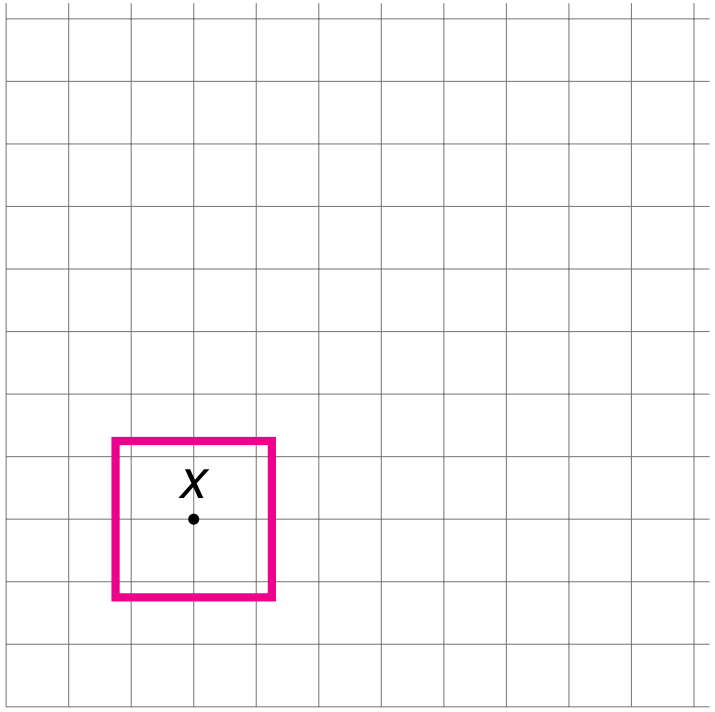
$A \oplus B := \{x \in \mathbb{R}^d \mid x = a + b,\ a \in A, b \in B\}$.

Clump detection



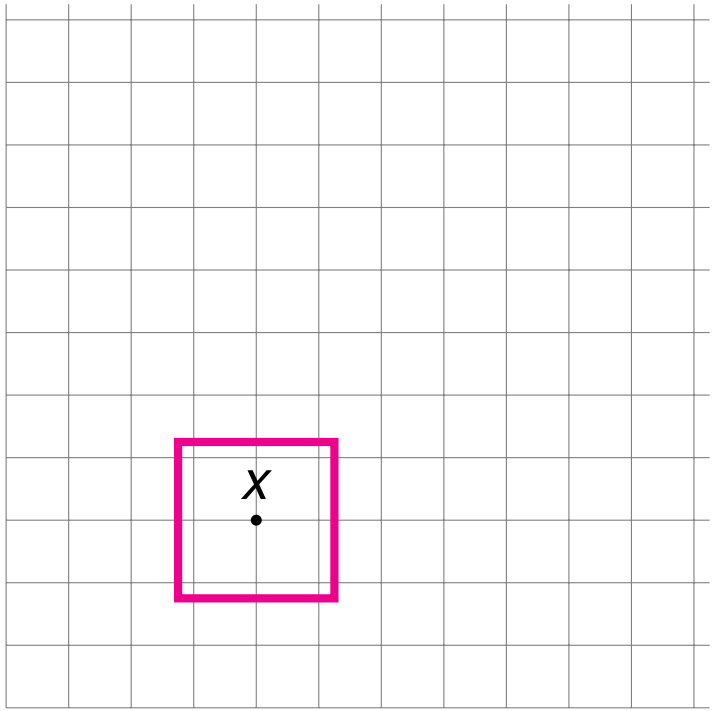
Clump detection

- Estimate the entropy on $B + x, x \in W$.



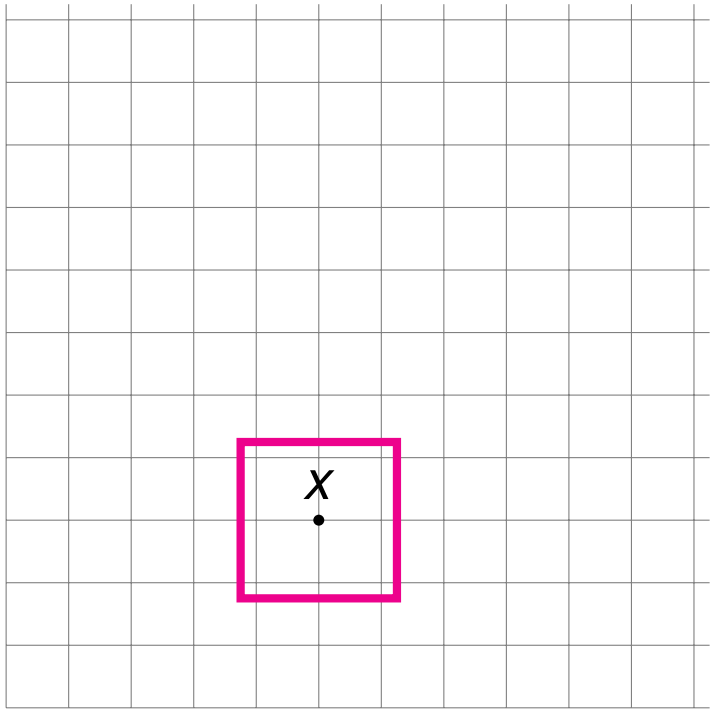
Clump detection

- Estimate the entropy on $B + x, x \in W$.



Clump detection

- Estimate the entropy on $B + x, x \in W$.

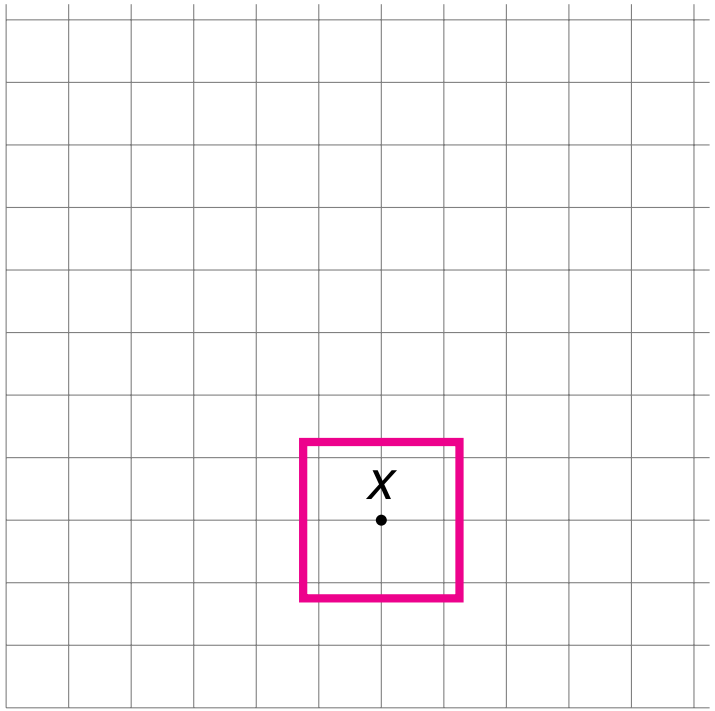


Clump detection

- Estimate the entropy on $B + x, x \in W$.

Lowest entropy region

clump candidate



Plan

Density f a priori unknown $\rightsquigarrow$ nonparametric estimation

1. Estimate f in $B \rightsquigarrow \widehat{f}_B$
2. Estimate entropy $\mathcal{E}_f$ in $B \rightsquigarrow \widehat{\mathcal{E}}_f(B)$
3. Study asymptotic properties of $\widehat{\mathcal{E}}_f^*(B_n)$, $B_n \uparrow \mathbb{R}^d$.

Estimator of the density $f\colon M \to \mathbb{R}$

$$\hat{f}_{B_n}(\eta) := \frac{1}{\lambda|B_n|} \sum_{i\geq 1} \frac{\mathbb{1}_{\{Y_i \in B_n\}}}{b_n^p \theta_\eta(\xi_i)} K\left(\frac{d_g(\eta,\xi_i)}{b_n}\right) \qquad \eta \in M.$$

- B_n - observation window $\rightsquigarrow$ $\lambda|B_n|$ estimator of $\Pi(B_n)$;
- b_n - bandwidth;
- $\theta_\eta\colon M \to \mathbb{R}$ - volume density function;
- $K\colon \mathbb{R}_+ \to \mathbb{R}$ - kernel function;
- $d_g(\eta,\xi)$ - geodesic distance.

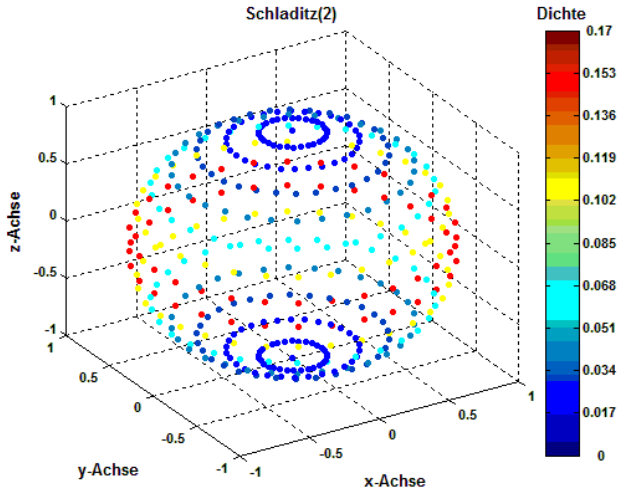
Properties

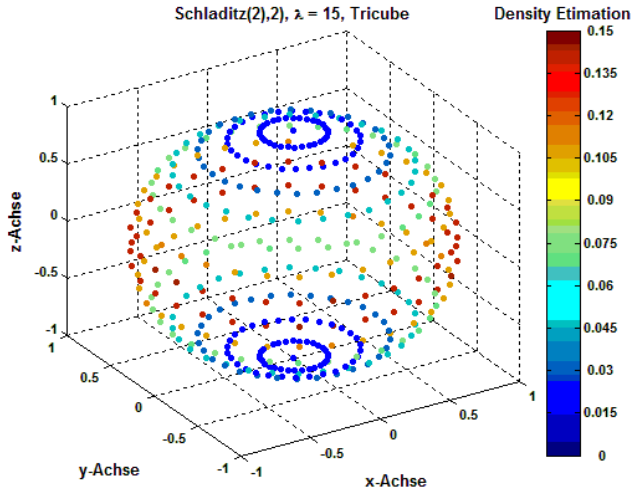
- Under assumptions (b)(K)(f), $\hat{f}_{B_n}$ is L^2-consistent ($b_n^p|B_n| \to \infty$).
- Estimation error:

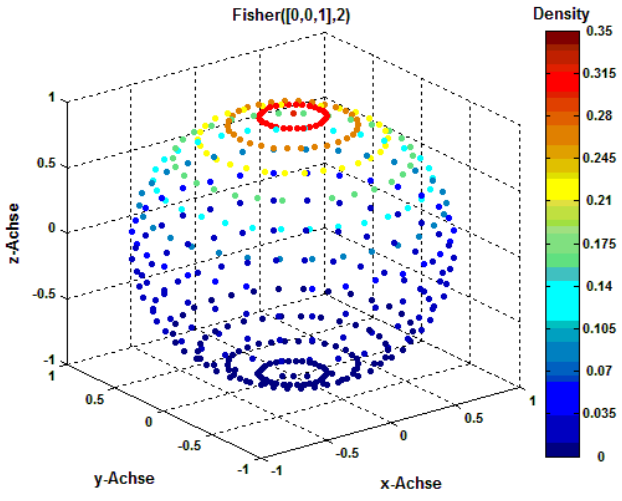
$$\mathbb{E}[\|\hat{f}_n - f\|_2^2] \leq \frac{C_\theta \omega_p K_0^2}{\lambda |B_n'| b_n^p} + b_n^4 C_2^2 K_2^2 \upsilon_g(M).$$

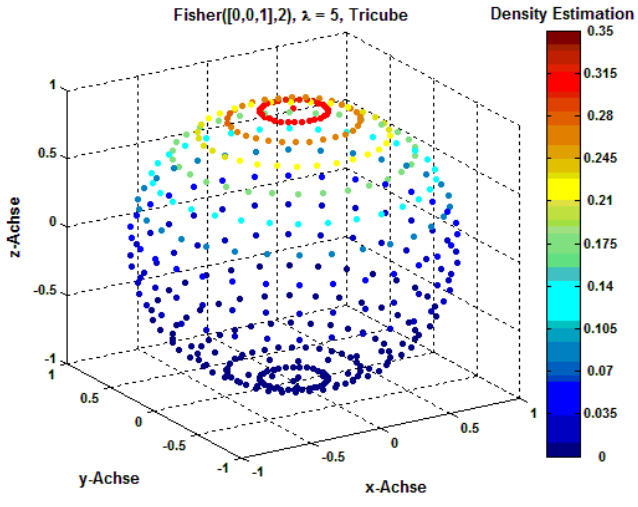
- Optimal bandwidth: $b_{opt} = \left(\frac{C_\theta \omega_p K_0^2}{2 C_2^2 K_2^2 \upsilon_g(M) \lambda |B_n'|} \right)^{\frac{1}{p+4}}$.
- Stronger assumptions on $|B_n| \rightsquigarrow$ a.s. consistency of $\hat{f}_{B_n}$.

Estimation error in simulations







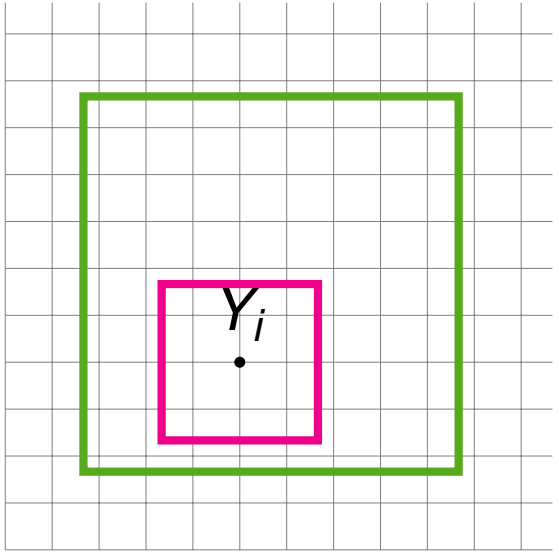


2. Entropy estimator

Observation window: $B_n \subset \mathbb{R}^d$.

$$\widehat{\mathcal{E}}_f(B_n) := -\frac{1}{\lambda|B_n|} \sum_{i \geq 1} \mathbb{1}_{\{Y_i \in B_n\}} \log \hat{f}_{B'_n + Y_i}(\xi_i),$$

$$B_n = [0, p_n]^d, \qquad B'_n = [0, m_n]^d.$$



Properties

- Under assumptions (b)(K)(f)(E), $\widehat{\mathcal{E}}_f(B_n)$ is L^2-consistent ($b_n^p|B_n'| \to \infty$).
- Estimation error:

$$\mathbb{E}[|\widehat{\mathcal{E}}_f(B_n) - \mathcal{E}_f|^2]^{1/2} \leq \frac{2K_0 C_\theta \upsilon_g(M)}{\lambda^2 c_1^2 |B_n||B_n'|b_n^p} + \frac{1}{\lambda^2|B_n'|} + \frac{8b_n^2 L_2}{c_2^2} + \frac{L_1}{\lambda|B_n|}.$$

- Optimal bandwidth: $b_{opt} = \left(\frac{c_1^2 K_0 C_\theta \upsilon_g(M)}{4L_2\lambda^2 c_2^2|B_n||B_n'|}\right)^{\frac{1}{p+2}}$.

Simulations

$M = S^2$, intensity $\lambda = 15$, tricube kernel.

Distribution	$\mathcal{E}_f$	$\overline{\widehat{\mathcal{E}}_f(B_n)}$	Var $\widehat{\mathcal{E}}_f(B_n)$	$\|Err\|_\infty$	MSQE
Uniform	2.5310	2.5165	8.7105e-07	0.0157	0.0459
Schladitz(2)	2.3554	2.3525	9.0843e-07	0.0039	0.0067
Fisher(2)	1.7239	1.8930	2.3662e-06	0.1715	0.3782
Watson(2)	1.8646	1.6849	2.8397e-07	0.1804	0.5682

Asymptotic distribution

$$\widehat{\mathcal{E}}_f(B_n) := -\frac{1}{\lambda|B_n|}\sum_{i\geq 1}\mathbb{1}_{\{Y_i\in B_n\}}\log\hat{f}_{B_n'+Y_i}(\xi_i).$$

- $B_n = [0, p_n]^d$, $B_n' = [0, m_n]^d$;
- $\{\log\widehat{f}_{y\oplus B_n'}(\xi_y)\}_{y\in\mathbb{R}^d}$ is an m_n-dependent random field;
- $\widehat{\mathcal{E}}_f(B_n)$ is a partial random sum of an m_n-dependent random field.

$\log\widehat{f}_{y_1\oplus B_n'}(\xi_{y_1}), \log\widehat{f}_{y_2\oplus B_n'}(\xi_{y_2})$ independent $\Leftrightarrow \|y_1 - y_2\|_\infty > m_n$.

General theorem

(Corollary to Wang, Woodroofe'14)

Let $\{X_{n,y}, y \in \mathbb{R}_+^d\}$, $n \in \mathbb{N}$, be a stationary m_n-dependent zero mean random field observed in $B_n = [0, p_n]^d$, $m_n < p_n \xrightarrow{n\to\infty} \infty$. Further, let Π be a stationary PPP on $\mathbb{R}_+^d$ of intensity $\lambda > 0$. Under assumptions (A1)–(A4),

$$\frac{\sum_{y\in\Pi\cap B_n} X_{n,y}}{\sqrt{|B_n|\sigma_n^2}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1).$$

$\rightsquigarrow$ Independence between $\{X_{n,y}, y \in \mathbb{R}_+^d\}_{n\in\mathbb{N}}$ and Π is not required!

Modifications

- No independence $\rightsquigarrow$ intractable assumptions.

Modifications

- No independence $\rightsquigarrow$ intractable assumptions.
- Assuming independence between the random field and the PPP:

Modifications

- No independence $\rightsquigarrow$ intractable assumptions.
- Assuming independence between the random field and the PPP:

(A1') $\sup_{n\in\mathbb{N}} \left(\mathbb{E}[X_{n,o}^2] + \int_{\mathbb{R}_+^d} |\operatorname{Cov}(X_{n,o}, X_{n,y})| \, dy \right) < \infty.$

(A3') $\liminf_{n\to\infty} \left(\mathbb{E}[X_{n,o}^2] + \lambda \int_{\mathbb{R}_+^d} \frac{|[0,q_n]^d \cap ([0,q_n]^d - y)|}{q_n^d} \operatorname{Cov}(X_{n,o}, X_{n,y}) \, dy \right) > 0,$

where $\frac{m_n}{q_n} \xrightarrow{n\to\infty} 0$ and $\frac{q_n}{p_n} \xrightarrow{n\to\infty} 0$.

Application to entropy

Assuming independence between the random field and the PPP:

$$\widehat{\mathcal{E}}_f^*(B_n) := -\frac{1}{\lambda|B_n|}\sum_{i\geq 1}\mathbb{1}_{\{Y_i^*\in B_n\}}\log\hat{f}_{B_n'+Y_i^*}(\xi_i^*),$$

where $\Psi^* := \{(Y_i^*, \xi_i^*)\}_{i\geq 1}$ is an independent copy of Ψ.

$\rightsquigarrow$ Extra assumption (f3): $\inf_{\eta\in\operatorname{supp} f} f(\eta) =: c_0 > 0$.

Theorem: Let $B_n = [0, p_n]^d$, $B'_n = [0, m_n]^d$ such that $p_n = m_n^{4+\delta}$ $\delta > 0$, and $m_n \xrightarrow{n\to\infty} \infty$. Under assumptions (b)(K)(f)(E)(f3) it holds that

$$\sqrt{|B_n|}\frac{\widehat{\mathcal{E}}^*_f(B_n) - \widehat{\mu}_{B_n}}{\sigma_n} \xrightarrow{d} \mathcal{N}(0,1),$$

where

$$\widehat{\mu}_{B_n} = \frac{\Pi(B_n)}{\lambda|B_n|}\mathbb{E}[-\log \hat{f}_{B'_n}(\xi^*_0)],$$

$$\sigma_n^2 = \lambda\mathbb{E}[\log^2 \hat{f}_{B'_n}(\xi^*_0)] + \lambda^2 \int_{B'_n} \frac{|[0,q_n]^d \cap ([0,q_n]^d - y)|}{q_n^d} \operatorname{Cov}(\log \hat{f}_{B'_n}(\xi^*_0), \log \hat{f}_{B'_n}(\xi'_y))\, dy,$$

and $\{\xi'_y, y \in \mathbb{R}^d\}$ are i.i.d. copies of ξ_0, $q_n = m_n^{1+\delta'}$, $4\delta' < \delta$.

References


I.A. Ahmad and P.E. Lin, *A nonparametric estimation of the entropy for absolutely continuous distributions*, IEEE Trans. Information Theory **IT-22** (1976), no. 3, 372–375.


P. Alonso-Ruiz and E. Spodarev, *Estimation of entropy for poisson marked point processes*, ArXiv e-prints (2015), arXiv:1511.03830.


J. Beirlant, E. J. Dudewicz, L. Györfi, and E. C. van der Meulen, *Nonparametric entropy estimation: an overview*, Int. J. Math. Stat. Sci. **6** (1997), no. 1, 17–39.


L. Heinrich, *Asymptotic behaviour of an empirical nearest-neighbour distance function for stationary Poisson cluster processes*, Math. Nachr. **136** (1988), 131–148.


M. Kiderlen, *Non-parametric estimation of the directional distribution of stationary line and fibre processes*, Adv. in Appl. Probab. **33** (2001), no. 1, 6–24.


B. Pelletier, *Kernel density estimation on Riemannian manifolds*, Statist. Probab. Lett. **73** (2005), no. 3, 297–304.


Y. Wang and M. Woodroofe, *On the asymptotic normality of kernel density estimators for causal linear random fields*, J. Multivariate Anal. **123** (2014), 201–213.

Assumptions type (b)

(b1) $0 \leq b_n < r_0 < \text{inj}_g M := \inf_{\eta \in M} \{r > 0, B_r(\eta) \text{ normal nbhd}\}$;

(b2) $b_n \downarrow 0$;

(b3) $\lim_{n \to \infty} b_n |B_n| = \infty$.

Assumptions type (K)

(K1) $\int_{\mathbb{R}^d} K(\|x\|)dx = 1$;

(K2) $0 < \int_{\mathbb{R}^d} K(\|x\|)\|x\|^2 dx =: K_2 < \infty$;

(K3) $\operatorname{supp} K = [0, 1]$;

(K4) $\sup_{r \geq 0} K(r) =: K_0 > 0$;

(K5) $\int_{\mathbb{R}^d} K(\|x\|)x\, dx = 0$ (symmetry property).

Assumptions type (f) and (E)

(f1) $f \in L^2(M)$, eventually f is continuous;

(f2) f has bounded second covariant derivative, $\|D^2 f\| < C_2$.

(E1) $\mathbb{E}[\log^2 \hat{f}_{B_n'}(\xi_o)] =: L_1 < \infty$;

(E2) $\mathbb{E}\left[\left(\frac{\|\nabla f(\xi_o)\|}{f(\xi_o)}\right)^2\right] =: L_2 < \infty$.

Theorem: (Corollary to Wang and Woodrofe 2014) Suppose that there exist $C > 0$ and $\{p_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ such that

(A1) $\mathbb{E}\Big[\Big(\sum_{i\geq 1} \mathbb{1}_{\{Y_i\in[[0,j]]\}} X_{n,i}\Big)^2\Big] \leq C\sqrt{j^1\cdots j^d} \quad \forall\, n\in\mathbb{N},\ j\in\mathbb{N}^d,$

(A2) $\frac{m'_n}{p_n} \xrightarrow{n\to\infty} 0 \qquad \text{and} \qquad \frac{p_n}{m_n} \xrightarrow{n\to\infty} 0,$

(A3) $\lim_{n\to\infty} \frac{1}{p_n^d}\mathbb{E}\Big[\Big(\sum_{i\geq 1} \mathbb{1}_{\{Y_i\in[[0,j]]\}} X_{n,i}\Big)^2\Big] = \sigma^2 > 0,$

(A4) $\forall\,\varepsilon>0,\ \lim_{n\to\infty} \frac{1}{p_n^d}\mathbb{E}\Big[\Big(\sum_{i\geq 1} \mathbb{1}_{\{Y_i\in[[0,j]]\}} X_{n,i}\Big)^2 \mathbf{1}\Big(\big|\sum_{i\geq 1} \mathbb{1}_{\{Y_i\in[[0,j]]\}} X_{n,i}\big| > \sqrt{j^1\cdots j^d}\varepsilon\Big)\Big] = 0.$

Then

$$\frac{\sum_{i\geq 1} \mathbb{1}_{\{Y_i\in[[0,p_n]]\}} X_{n,i}}{\sqrt{p_n^d\sigma^2}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1).$$