Rates of Multivariate Normal Approximation for Statistics in Geometric Probability

Joseph Yukich (joint with Matthias Schulte)

Lehigh University

Euler International Mathematical Institute

Introduction

· Questions pertaining to geometric structures on random input $\mathcal{X} \subset \mathbb{R}^d$ often involve analyzing sums of spatially correlated terms

$$\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X}),$$

where the $\mathbb{R}$-valued score function ξ, defined on pairs $(x, \mathcal{X})$, represents the interaction of x with respect to $\mathcal{X}$.

Introduction

· Questions pertaining to geometric structures on random input $\mathcal{X} \subset \mathbb{R}^d$ often involve analyzing sums of spatially correlated terms

$$\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X}),$$

where the $\mathbb{R}$-valued score function ξ, defined on pairs $(x, \mathcal{X})$, represents the interaction of x with respect to $\mathcal{X}$.

· The sums describe some global feature of the random structure in terms of local contributions $\xi(x, \mathcal{X}),\ x \in \mathcal{X}$.

Set-up

· Fix a window $W \subset \mathbb{R}^d, d \geq 2$. $\mathcal{P}_s$ a Poisson point process of intensity s on W.

· **Goal.** Establish central limit theorems for

$$\sum_{x \in \mathcal{P}_s} \xi(x, \mathcal{P}_s) \qquad (*)$$

as $s \to \infty$ and, more generally, establish rates of multivariate normal approximation for vectors whose components have the form (*).

Set-up

· Fix a window $W \subset \mathbb{R}^d, d \geq 2$. $\mathcal{P}_s$ a Poisson point process of intensity s on W.

· **Goal.** Establish central limit theorems for

$$\sum_{x \in \mathcal{P}_s} \xi(x, \mathcal{P}_s) \qquad (*)$$

as $s \to \infty$ and, more generally, establish rates of multivariate normal approximation for vectors whose components have the form (*).

· If summands are identically distributed then usually one has

$$\mathbb{E} \sum_{x \in \mathcal{P}_s} \xi(x, \mathcal{P}_s) = \Theta(s), \quad \mathrm{Var} \sum_{x \in \mathcal{P}_s} \xi(x, \mathcal{P}_s) = \Theta(s).$$

Set-up and Main Goal

$W \subset \mathbb{R}^d$, $d \geq 2$, a fixed measurable set.

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg; $g : W \to \mathbb{R}^+$. Thus, for $A \subseteq W$, $|\mathcal{P}_{sg} \cap A|$ is Poisson distributed with parameter $\int_A sg(x)dx$.

Set-up and Main Goal

$W \subset \mathbb{R}^d$, $d \geq 2$, a fixed measurable set.

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg; $g : W \to \mathbb{R}^+$. Thus, for $A \subseteq W$, $|\mathcal{P}_{sg} \cap A|$ is Poisson distributed with parameter $\int_A sg(x)dx$.

$\mathbf{N}$: the set of simple σ-finite counting measures on $\mathbb{R}^d$.

$(\xi_s^{(1)})_{s \geq 1}, ..., (\xi_s^{(m)})_{s \geq 1}$, measurable maps ('scores') from $W \times \mathbf{N} \to \mathbb{R}$.

$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \quad A_i \subseteq W.$

Set-up and Main Goal

$W \subset \mathbb{R}^d$, $d \geq 2$, a fixed measurable set.

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg; $g : W \to \mathbb{R}^+$. Thus, for $A \subseteq W$, $|\mathcal{P}_{sg} \cap A|$ is Poisson distributed with parameter $\int_A sg(x)dx$.

$\mathbf{N}$: the set of simple σ-finite counting measures on $\mathbb{R}^d$.

$(\xi_s^{(1)})_{s\geq 1}, ..., (\xi_s^{(m)})_{s\geq 1}$, measurable maps ('scores') from $W \times \mathbf{N} \to \mathbb{R}$.

$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x\in\mathcal{P}_{sg}\cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \quad A_i \subseteq W.$

Goal. Find rates of multivariate normal convergence for the m-vector

$$\left(\frac{H_s^{(1)} - \mathbb{E}\, H_s^{(1)}}{\sqrt{\mathrm{Var} H_s^{(1)}}}, ..., \frac{H_s^{(m)} - \mathbb{E}\, H_s^{(m)}}{\sqrt{\mathrm{Var} H_s^{(m)}}} \right)$$

as intensity $s \to \infty$.

Underlying assumption on scores $(\xi_s^{(i)})_{s \geq 1}$

Recall that the ith score $\xi_s^{(i)}$ generates the statistic

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \quad A_i \subseteq W.$$

Assume for all $i \in \{1, ..., m\}$ that $\xi_s^{(i)}$ is the score $\xi^{(i)}$ at x evaluated on an s-dilation of the underlying point set:

$$\xi_s^{(i)}(x, \mathcal{M}) = \xi^{(i)}(x, x + s^{1/d}(\mathcal{M} - x)), \ x \in W, \mathcal{M} \in \mathbf{N}, \ s \geq 1.$$

Underlying assumption on scores $(\xi_s^{(i)})_{s\geq 1}$

Recall that the ith score $\xi_s^{(i)}$ generates the statistic

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x\in\mathcal{P}_{sg}\cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \quad A_i \subseteq W.$$

Assume for all $i \in \{1, ..., m\}$ that $\xi_s^{(i)}$ is the score $\xi^{(i)}$ at x evaluated on an s-dilation of the underlying point set:

$$\xi_s^{(i)}(x, \mathcal{M}) = \xi^{(i)}(x, x + s^{1/d}(\mathcal{M} - x)),\ x \in W, \mathcal{M} \in \mathbf{N},\ s \geq 1.$$

If $\xi_s^{(i)}$ is translation invariant then this says

$$\xi_s^{(i)}(x, \mathcal{M}) = \xi^{(i)}(\mathbf{0}, s^{1/d}(\mathcal{M} - x)).$$

Stabilization of scores

For $s \geq 1$ we say that $R_s : W \times \mathbf{N} \to \mathbb{R}^+$ is a radius of stabilization for $(\xi_s^{(1)})_{s\geq 1}, ..., (\xi_s^{(m)})_{s\geq 1}$, if for all $x \in W$, $\mathcal{M} \in \mathbf{N}$, $s \geq 1$, $i \in \{1, ..., m\}$ we have

$$\xi_s^{(i)}(x, \mathcal{M}) = \xi_s^{(i)}(x, \mathcal{M} \cap B^d(x, R_s(x, \mathcal{M}))).$$

Stabilization of scores

For $s \geq 1$ we say that $R_s : W \times \mathbf{N} \to \mathbb{R}^+$ is a radius of stabilization for $(\xi_s^{(1)})_{s \geq 1}, ..., (\xi_s^{(m)})_{s \geq 1}$, if for all $x \in W$, $\mathcal{M} \in \mathbf{N}$, $s \geq 1$, $i \in \{1, ..., m\}$ we have

$$\xi_s^{(i)}(x, \mathcal{M}) = \xi_s^{(i)}(x, \mathcal{M} \cap B^d(x, R_s(x, \mathcal{M}))).$$

Loosely speaking, this says the scores $\xi_s^{(i)}, i \in \{1, ..., m\}$, are determined by data at distance $R_s(x, \mathcal{M})$ from x.

Exponential stabilization of scores

We say that $(\xi_s^{(1)})_{s\geq 1}, ..., (\xi_s^{(m)})_{s\geq 1}$ are exponentially stabilizing wrt $\mathcal{P}_{sg}$ if there are constants C_{stab} and $c_{stab} \in (0,\infty)$ such that

$$\mathbb{P}(R_s(x,\mathcal{P}_{sg}) \geq r) \leq C_{stab} \exp(-c_{stab} s r^d), \quad r \geq 0, x \in W, s \geq 1.$$

Exponential stabilization of scores

We say that $(\xi_s^{(1)})_{s\geq 1}, ..., (\xi_s^{(m)})_{s\geq 1}$ are exponentially stabilizing wrt $\mathcal{P}_{sg}$ if there are constants C_{stab} and $c_{stab} \in (0, \infty)$ such that

$$\mathbb{P}(R_s(x, \mathcal{P}_{sg}) \geq r) \leq C_{stab} \exp(-c_{stab} s r^d), \quad r \geq 0, x \in W, s \geq 1.$$

This says that scores $(\xi_s^{(1)})_{s\geq 1}, ..., (\xi_s^{(m)})_{s\geq 1}$ have spatial dependencies which decay exponentially fast.

Idea: Sums of exponentially stabilizing scores should behave like sums of i.i.d. random variables.

p-moment condition on scores $(\xi_s^{(i)})_{s\geq 1}$

We say that $(\xi_s^{(1)})_{s\geq 1}, ..., (\xi_s^{(m)})_{s\geq 1}$, satisfy a p-moment condition, $p \geq 1$, if there is $C_p \in (0,\infty)$ such that for all $i \in \{1, ..., m\}$, we have

$$\sup_{s\in[1,\infty)} \sup_{x,y\in W} \mathbb{E}\,|\xi_s^{(i)}(x, \mathcal{P}_{sg} \cup \{y\})|^p \leq C_p,$$

Rates of univariate normal convergence

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg; where $g : W \to \mathbb{R}^+$.

Put $H_s := H_s(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg}} \xi_s(x, \mathcal{P}_{sg})$.

Rates of univariate normal convergence

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg; where $g : W \to \mathbb{R}^+$.

Put $H_s := H_s(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg}} \xi_s(x, \mathcal{P}_{sg})$.

Theorem (Lachieze-Rey, Schulte + Y. (2019)) Assume $(\xi_s), s \geq 1$, are exponentially stabilizing and satisfy the p-moment condition for some $p \in (4, \infty)$. If $\mathrm{Var} H_s = \Omega(s)$, then

$$d_K\left(\frac{H_s - \mathbb{E}\, H_s}{\sqrt{\mathrm{Var} H_s}}, N(0,1)\right) \leq \frac{c}{\sqrt{s}}, \qquad s \geq 1.$$

Rates of univariate normal convergence

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg; where $g : W \to \mathbb{R}^+$.

Put $H_s := H_s(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg}} \xi_s(x, \mathcal{P}_{sg})$.

Theorem (Lachieze-Rey, Schulte + Y. (2019)) Assume $(\xi_s), s \geq 1$, are exponentially stabilizing and satisfy the p-moment condition for some $p \in (4, \infty)$. If $\mathrm{Var} H_s = \Omega(s)$, then

$$d_K \left(\frac{H_s - \mathbb{E}\, H_s}{\sqrt{\mathrm{Var} H_s}}, N(0,1) \right) \leq \frac{c}{\sqrt{s}}, \qquad s \geq 1.$$

Question: what are good proximity bounds for

$$d_K \left(\frac{H_s - \mathbb{E}\, H_s}{\sqrt{s}}, N(0,1) \right)?$$

Multivariate CLT without rates

Recall $H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \; s \geq 1.$
Centered version: $\bar{H}_s^{(i)} := H_s^{(i)} - \mathbb{E}\, H_s^{(i)}.$

Multivariate CLT without rates

Recall $H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}),\ s \geq 1.$

Centered version: $\bar{H}_s^{(i)} := H_s^{(i)} - \mathbb{E}\, H_s^{(i)}.$

Thm (Penrose; Baryshnikov + Y.) Assume $(\xi_s^{(1)})_{s \geq 1}, ..., (\xi_s^{(m)})_{s \geq 1}$, are

(i) exponentially stabilizing, and

(ii) satisfy the p-moment condition for some $p > 2$.

Then for all $i, j \in \{1, ..., m\}$ as $s \to \infty$ we have

$$\frac{\mathrm{Cov}(H_s^{(i)}, H_s^{(j)})}{s} \to \sigma_{ij}, \qquad s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}) \xrightarrow{\mathcal{D}} N_\Sigma,$$

where N_Σ is multivariate normal with covariance matrix

$$\Sigma = (\sigma_{ij})_{1 \leq i,j \leq m}.$$

Three distances between m-dimensional vectors

(i) $\mathcal{H}_m^{(2)}$: all C^2-functions $h : \mathbb{R}^m \to \mathbb{R}$ such that

$$|h(x) - h(y)| \leq ||x - y||, \ x, y \in \mathbb{R}^m,$$

$$\sup_{x \in \mathbb{R}^m} ||\text{Hess } h(x)||_{\text{op}} \leq 1.$$

Three distances between m-dimensional vectors

(i) $\mathcal{H}_m^{(2)}$: all C^2-functions $h : \mathbb{R}^m \to \mathbb{R}$ such that

$$|h(x) - h(y)| \leq ||x - y||, \ x, y \in \mathbb{R}^m,$$

$$\sup_{x \in \mathbb{R}^m} ||\text{Hess } h(x)||_{\text{op}} \leq 1.$$

Given m-dimensional random vectors Y, Z we put

$$d_2(Y, Z) := \sup_{h \in \mathcal{H}_m^{(2)}} |\mathbb{E}\, h(Y) - \mathbb{E}\, h(Z)|$$

if $\mathbb{E}\, ||Y||, \mathbb{E}\, ||Z|| < \infty$.

Three distances between m-dimensional vectors

(ii) $\mathcal{H}_m^{(3)}$: all C^3-functions $h : \mathbb{R}^m \to \mathbb{R}$ such that absolute values of the second and third partial derivatives are bounded by 1.

Three distances between m-dimensional vectors

(ii) $\mathcal{H}_m^{(3)}$: all C^3-functions $h : \mathbb{R}^m \to \mathbb{R}$ such that absolute values of the second and third partial derivatives are bounded by 1.

Given m-dimensional random vectors Y, Z we put

$$d_3(Y, Z) := \sup_{h \in \mathcal{H}_m^{(3)}} |\mathbb{E}\, h(Y) - \mathbb{E}\, h(Z)|$$

if $\mathbb{E}\, ||Y||^2, \mathbb{E}\, ||Z||^2 < \infty$.

Three distances between m-dimensional vectors

(ii) $\mathcal{H}_m^{(3)}$: all C^3-functions $h : \mathbb{R}^m \to \mathbb{R}$ such that absolute values of the second and third partial derivatives are bounded by 1.

Given m-dimensional random vectors Y, Z we put

$$d_3(Y,Z) := \sup_{h \in \mathcal{H}_m^{(3)}} |\mathbb{E}\, h(Y) - \mathbb{E}\, h(Z)|$$

if $\mathbb{E}\, ||Y||^2, \mathbb{E}\, ||Z||^2 < \infty$.

(iii)

$$d_{convex}(Y,Z) := \sup_{h \in \mathcal{I}} |\mathbb{E}\, h(Y) - \mathbb{E}\, h(Z)|,$$

where $\mathcal{I}$ is the set of indicators of convex sets in $\mathbb{R}^m$.

Main Theorem: Rates of Multivariate Normal Convergence

Recall $H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \ s \geq 1.$

$$\frac{\mathrm{Cov}(H_s^{(i)}, H_s^{(j)})}{s} \to \sigma_{ij}, \qquad s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}) \xrightarrow{\mathcal{D}} N_{\Sigma}.$$

· Assume $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq m}$ is positive definite.

Main Theorem: Rates of Multivariate Normal Convergence

Recall $H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \ s \geq 1.$

$$\frac{\text{Cov}(H_s^{(i)}, H_s^{(j)})}{s} \to \sigma_{ij}, \qquad s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}) \xrightarrow{\mathcal{D}} N_\Sigma.$$

· Assume $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq m}$ is positive definite.

Theorem (Schulte + Y.) Assume $(\xi_s^{(1)})_{s \geq 1}, ..., (\xi_s^{(m)})_{s \geq 1}$ are

(i) exponentially stabilizing, and

(ii) satisfy the p-moment condition for some $p > 6$.

Recall $H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}),\ s \geq 1.$

$$\frac{\mathrm{Cov}(H_s^{(i)}, H_s^{(j)})}{s} \to \sigma_{ij}, \qquad s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}) \xrightarrow{\mathcal{D}} N_\Sigma.$$

· Assume $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq m}$ is positive definite.

Theorem (Schulte + Y.) Assume $(\xi_s^{(1)})_{s \geq 1}, ..., (\xi_s^{(m)})_{s \geq 1}$ are
(i) exponentially stabilizing, and
(ii) satisfy the p-moment condition for some $p > 6$.
Then there is a constant $C \in (0, \infty)$ such that

$$\tilde{d}(s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}), N_\Sigma) \leq C s^{-1/d},\ s \geq 1,\ (*)$$

for $\tilde{d} \in \{d_2, d_3, d_{convex}\}$.

Two remarks

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \ \ s \geq 1.$$

(i) A main ingredient to the proof: For all $i, j \in \{1, ..., m\}$

$$\left| \sigma_{ij} - \frac{\mathrm{Cov}(H_s^{(i)}, H_s^{(j)})}{s} \right| \leq C s^{-1/d}, \quad s \geq 1.$$

Two remarks

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \ \ s \geq 1.$$

(i) A main ingredient to the proof: For all $i, j \in \{1, ..., m\}$

$$\left| \sigma_{ij} - \frac{\mathrm{Cov}(H_s^{(i)}, H_s^{(j)})}{s} \right| \leq C s^{-1/d}, \quad s \geq 1.$$

(ii) If we replace N_Σ by $N_{\Sigma(s)}$, where $\Sigma(s)$ is the covariance matrix of

$$s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}),$$

then the rates of multivariate normal convergence improve to

$$\tilde{d}(s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}), N_{\Sigma(s)}) \leq C s^{-1/2}, \ \ s \geq 1,$$

for $\tilde{d} \in \{d_2, d_3, d_{convex}\}$. Rates are not improvable in general.

Literature

· Multivariate clts for vectors with certain dependency structures: Rai$\breve{c}$ (2004), Goldstein and Rinott (2005), Chen, Goldstein, and Shao (2011)

· Penrose and Wade (2008): consider the special case $\xi_s^{(1)} = ... = \xi_s^{(m)}$ and all sets $A_i, i \in \{1, ..., m\}$, are disjoint. They establish rate of normal convergence $O(s^{-1/(2d+\epsilon)}),\ \epsilon > 0$, wrt Kolmogorov distance in $\mathbb{R}^d$.

Literature

· Multivariate clts for vectors with certain dependency structures: Raič (2004), Goldstein and Rinott (2005), Chen, Goldstein, and Shao (2011)

· Penrose and Wade (2008): consider the special case $\xi_s^{(1)} = ... = \xi_s^{(m)}$ and all sets $A_i, i \in \{1, ..., m\}$, are disjoint. They establish rate of normal convergence $O(s^{-1/(2d+\epsilon)}),\ \epsilon > 0$, wrt Kolmogorov distance in $\mathbb{R}^d$.

· Peccati and Zheng (2010)

Literature

· Multivariate clts for vectors with certain dependency structures: Rai$\breve{c}$ (2004), Goldstein and Rinott (2005), Chen, Goldstein, and Shao (2011)

· Penrose and Wade (2008): consider the special case $\xi_s^{(1)} = ... = \xi_s^{(m)}$ and all sets $A_i, i \in \{1, ..., m\}$, are disjoint. They establish rate of normal convergence $O(s^{-1/(2d+\epsilon)}),\ \epsilon > 0$, wrt Kolmogorov distance in $\mathbb{R}^d$.

· Peccati and Zheng (2010)

· Hug, Last, Schulte (2016): establish rates with respect to d_3 which depend on knowledge of Wiener-It$\hat{o}$ chaos expansion.

CLTs for Poisson functionals, Malliavin calculus

· **Goal:** Find rates of multivariate normal approximation for $s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)})$; each $H_s^{(i)}$ is a function of $\mathcal{P}_{sg}$ and is thus a Poisson functional.

CLTs for Poisson functionals, Malliavin calculus

· **Goal:** Find rates of multivariate normal approximation for $s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)})$; each $H_s^{(i)}$ is a function of $\mathcal{P}_{sg}$ and is thus a Poisson functional.

· We first prove rates of multivariate normal approximation for a vector of general Poisson functionals.

· η a Poisson process over $(\mathbb{X}, \mathcal{F})$ with intensity measure λ; thus $|\eta \cap A|$ is Poisson distributed with parameter $\lambda(A), A \in \mathcal{F}$.

CLTs for Poisson functionals, Malliavin calculus

· **Goal:** Find rates of multivariate normal approximation for $s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)})$; each $H_s^{(i)}$ is a function of $\mathcal{P}_{sg}$ and is thus a Poisson functional.

· We first prove rates of multivariate normal approximation for a vector of general Poisson functionals.

· η a Poisson process over $(\mathbb{X}, \mathcal{F})$ with intensity measure λ; thus $|\eta \cap A|$ is Poisson distributed with parameter $\lambda(A), A \in \mathcal{F}$.

· $F = (F_1, .., F_m)$, $m \in \mathbb{N}$, a vector of (Poisson) functionals of η if each F_i is represented as a measurable function f_i of η.

CLTs for Poisson functionals, Malliavin calculus

· **Goal:** Find rates of multivariate normal approximation for $s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)})$; each $H_s^{(i)}$ is a function of $\mathcal{P}_{sg}$ and is thus a Poisson functional.

· We first prove rates of multivariate normal approximation for a vector of general Poisson functionals.

· η a Poisson process over $(\mathbb{X}, \mathcal{F})$ with intensity measure λ; thus $|\eta \cap A|$ is Poisson distributed with parameter $\lambda(A), A \in \mathcal{F}$.

· $F = (F_1, .., F_m)$, $m \in \mathbb{N}$, a vector of (Poisson) functionals of η if each F_i is represented as a measurable function f_i of η.

· $D_x F_i := f_i(\eta \cup \{x\}) - f_i(\eta)$.

· $D^2_{x,y} F_i := f_i(\eta \cup \{x\} \cup \{y\}) - f_i(\eta \cup \{x\}) - f_i(\eta \cup \{y\}) + f_i(\eta)$.

CLTs for Poisson functionals, Malliavin calculus

· Peccati and Zheng (2010): η a Poisson process over $(\mathbb{X}, \mathcal{F})$ with intensity measure λ.

· $F = (F_1, .., F_m)$, $m \in \mathbb{N}$, a vector of (Poisson) functionals of η.

CLTs for Poisson functionals, Malliavin calculus

· Peccati and Zheng (2010): η a Poisson process over $(\mathbb{X}, \mathcal{F})$ with intensity measure λ.

· $F = (F_1, .., F_m)$, $m \in \mathbb{N}$, a vector of (Poisson) functionals of η.

· $\mathbb{E}\, F_i = 0$, $\Sigma = (\sigma_{i,j})_{i,j \in \{1,...,m\}}$ positive definite matrix. Put

$$\beta_1 := \sqrt{\sum_{i,j=1}^{m} \mathbb{E}\,(\sigma_{i,j} - \int_{\mathbb{X}} D_x F_i(-D_x L^{-1} F_j)\lambda(dx))^2}$$

$$\beta_2 := \int_{\mathbb{X}} \mathbb{E}\,(\sum_{i=1}^{m} |D_x F_i|)^2 \sum_{j=1}^{m} |D_x L^{-1} F_j| \lambda(dx).$$

· Then $d_2(F, N_\Sigma)$ and $d_3(F, N_\Sigma)$ are both bounded by $C(m, \Sigma) \cdot (\beta_1 + \beta_2)$.

CLTs for Poisson functionals, Malliavin calculus

· Peccati and Zheng (2010): η a Poisson process over $(\mathbb{X}, \mathcal{F})$ with intensity measure λ.

· $F = (F_1, .., F_m)$, $m \in \mathbb{N}$, a vector of (Poisson) functionals of η.

· $\mathbb{E}\, F_i = 0$, $\Sigma = (\sigma_{i,j})_{i,j \in \{1,\ldots,m\}}$ positive definite matrix. Put

$$\beta_1 := \sqrt{\sum_{i,j=1}^{m} \mathbb{E}\,(\sigma_{i,j} - \int_{\mathbb{X}} D_x F_i(-D_x L^{-1} F_j) \lambda(dx))^2}$$

$$\beta_2 := \int_{\mathbb{X}} \mathbb{E}\,(\sum_{i=1}^{m} |D_x F_i|)^2 \sum_{j=1}^{m} |D_x L^{-1} F_j| \lambda(dx).$$

· Then $d_2(F, N_\Sigma)$ and $d_3(F, N_\Sigma)$ are both bounded by $C(m, \Sigma) \cdot (\beta_1 + \beta_2)$.

· Last, Peccati, Schulte (2016): $\mathbb{E}\,|D_x L^{-1} F|^p$ and $\mathbb{E}\,|D^2_{x,y} L^{-1} F|^p$ bounded by moments of difference operators.

CLT for general Poisson functionals

Theorem (Schulte and Y.) Let $F = (F_1, \ldots, F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals; $\mathbb{E}\, F_i = 0$, $i \in \{1, \ldots, m\}$. Let $\Sigma = (\sigma_{ij})_{i,j\in\{1,\ldots,m\}} \in \mathbb{R}^{m\times m}$ be positive definite. Then

$$d_3(F, N_\Sigma) \leq \frac{m}{2} \sum_{i,j=1}^{m} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)| + m\gamma_1 + \frac{m}{2}\gamma_2 + \frac{m^2}{4}\gamma_3.$$

CLT for general Poisson functionals

Here

$$\gamma_1 := \Bigg(\sum_{i,j=1}^{m} \int_{\mathbb{X}^3} \big(\mathbb{E}\,(D^2_{x_1,x_3}F_i)^2(D^2_{x_2,x_3}F_i)^2\big)^{1/2} \times \big(\mathbb{E}\,(D_{x_1}F_j)^2(D_{x_2}F_j)^2\big)^{1/2}\lambda^3(\mathrm{d}(x_1,x_2,x_3))\Bigg)^{1/2}$$

$$\gamma_2 := \Bigg(\sum_{i,j=1}^{m} \int_{\mathbb{X}^3} \big(\mathbb{E}\,(D^2_{x_1,x_3}F_i)^2(D^2_{x_2,x_3}F_i)^2\big)^{1/2} \times \big(\mathbb{E}\,(D^2_{x_1,x_3}F_j)^2(D^2_{x_2,x_3}F_j)^2\big)^{1/2}\,\lambda^3(\mathrm{d}(x_1,x_2,x_3))\Bigg)^{1/2}$$

$$\gamma_3 := \sum_{i=1}^{m} \int_{\mathbb{X}} \mathbb{E}\,|D_x F_i|^3\,\lambda(\mathrm{d}x)$$

Special case: CLT for vector of stabilizing functionals

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \ \ s \geq 1.$$

· Let $F_i := s^{-1/2} \bar{H}_s^{(i)}$. F_i is a sum of stabilizing scores. Put $\mathbb{X}$ to be window W, put $\lambda(\mathrm{d}x)$ to be $sg(x)dx$. Then the integrals of moments of difference operators are $O(s^{-1/2})$.

Special case: CLT for vector of stabilizing functionals

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \;\; s \geq 1.$$

· Let $F_i := s^{-1/2} \bar{H}_s^{(i)}$. F_i is a sum of stabilizing scores. Put $\mathbb{X}$ to be window W, put $\lambda(\mathrm{d}x)$ to be $sg(x)dx$. Then the integrals of moments of difference operators are $O(s^{-1/2})$.

· Schulte + Y show $\gamma_1, \gamma_2, \gamma_3$ are all $O(s^{-1/2})$. Thus

$$d_3(F, N_\Sigma) \leq \frac{m}{2} \sum_{i,j=1}^{m} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)| + m\gamma_1 + \frac{m}{2}\gamma_2 + \frac{m^2}{4}\gamma_3$$

$$= O(s^{-1/d}) + O(s^{-1/2}).$$

Main thm for d_{convex}

Theorem (Schulte + Y.) Let $F = (F_1, \ldots, F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals $F_1, \ldots, F_m$ with $\mathbb{E}\, F_i = 0$, $i \in \{1, \ldots, m\}$.

Main thm for d_{convex}

Theorem (Schulte + Y.) Let $F = (F_1, \ldots, F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals $F_1, \ldots, F_m$ with $\mathbb{E}\, F_i = 0$, $i \in \{1, \ldots, m\}$.

If $\Sigma \in \mathbb{R}^{m \times m}$ is positive definite then there exists a constant $C \in (0, \infty)$ depending on m and Σ such that

$$d_{convex}(F, N_\Sigma) \leq C \max\left\{ \sum_{i,j \in \{1,\ldots,m\}} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)|, \gamma_1, \gamma_2, \tilde{\gamma}_3, \gamma_4, \gamma_5 \right\}.$$

Main thm for d_{convex}

Here the terms $\gamma_1, \gamma_2, \tilde{\gamma_3}, \gamma_4, \gamma_5$ are integrals of products of moments of difference operators applied to the Poisson functionals $F_1, ..., F_m$. For example,

$$\tilde{\gamma_3} := [\sum_{j,k=1}^{m} \int_{\mathbb{X}} \mathbb{E}\,(D_x F_j)^4 \lambda(dx)$$

$$+\frac{15}{4} \int_{\mathbb{X}^2} (\mathbb{E}\, D^2_{x,y} F_j)^4)^{1/2} (\mathbb{E}\,(D_x F_k)^4)^{1/2} \lambda^2(d(x,y))$$

$$+\frac{3}{4} \int_{\mathbb{X}^2} (\mathbb{E}\, D^2_{x,y} F_j)^4)^{1/2} (\mathbb{E}\,(D_{x,y} F_k)^4)^{1/2} \lambda^2(d(x,y))]^{1/2}.$$

Main thm for d_{convex}

Here the terms $\gamma_1, \gamma_2, \tilde{\gamma_3}, \gamma_4, \gamma_5$ are integrals of products of moments of difference operators applied to the Poisson functionals $F_1, ..., F_m$. For example,

$$\tilde{\gamma_3} := [\sum_{j,k=1}^{m} \int_{\mathbb{X}} \mathbb{E}\,(D_x F_j)^4 \lambda(dx)$$

$$+\frac{15}{4} \int_{\mathbb{X}^2} (\mathbb{E}\,D^2_{x,y} F_j)^4)^{1/2} (\mathbb{E}\,(D_x F_k)^4)^{1/2} \lambda^2(d(x,y))$$

$$+\frac{3}{4} \int_{\mathbb{X}^2} (\mathbb{E}\,D^2_{x,y} F_j)^4)^{1/2} (\mathbb{E}\,(D_{x,y} F_k)^4)^{1/2} \lambda^2(d(x,y))]^{1/2}.$$

Fortunately, one may control the order of growth of the terms $\gamma_1, ..., \gamma_5$ when the Poisson functionals F_i are sums of stabilizing score functions.

d_{convex} applied to stabilizing functionals $H_s^{(i)}$

Let $F_i := s^{-1/2}\bar{H}_s^{(i)}$; $H_s^{(i)} := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg})$.

Then $\max(\gamma_1, ..., \gamma_5) = O(s^{-1/2})$.

d_{convex} applied to stabilizing functionals $H_s^{(i)}$

Let $F_i := s^{-1/2}\bar{H}_s^{(i)}$; $H_s^{(i)} := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg})$.

Then $\max(\gamma_1, ..., \gamma_5) = O(s^{-1/2})$.

Theorem (Schulte + Y; 2019) Assume $(\xi_s^{(1)})_{s \geq 1}, ..., (\xi_s^{(m)})_{s \geq 1}$ are

(i) exponentially stabilizing, and

(ii) satisfy the p-moment condition for some $p > 6$.

d_{convex} applied to stabilizing functionals $H_s^{(i)}$

Let $F_i := s^{-1/2}\bar{H}_s^{(i)}$; $H_s^{(i)} := \sum_{x\in\mathcal{P}_{sg}\cap A_i} \xi_s^{(i)}(x,\mathcal{P}_{sg})$.

Then $\max(\gamma_1,...,\gamma_5) = O(s^{-1/2})$.

Theorem (Schulte + Y; 2019) Assume $(\xi_s^{(1)})_{s\geq 1},...,(\xi_s^{(m)})_{s\geq 1}$ are
(i) exponentially stabilizing, and
(ii) satisfy the p-moment condition for some $p > 6$.
Then there is a constant $C \in (0,\infty)$ such that

$$d_{convex}(s^{-1/2}(\bar{H}_s^{(1)},...,\bar{H}_s^{(m)}), N_\Sigma) \leq C s^{-1/d},\ \ s \geq 1$$

and

$$d_{convex}(s^{-1/2}(\bar{H}_s^{(1)},...,\bar{H}_s^{(m)}), N_{\Sigma(s)}) \leq C s^{-1/2},\ \ s \geq 1,$$

where $\Sigma(s)$ is the covariance matrix of $s^{-1/2}(\bar{H}_s^{(1)},...,\bar{H}_s^{(m)})$.

(i) **Multivariate statistics of kNN graph.** Let $k \in \mathbb{N}$ and $\mathcal{X} \subset \mathbb{R}^d$ a finite point set. For $x, y \in \mathcal{X}$, we put an undirected edge between x and y if x is one of the k nearest neighbors of y and/or y is a k nearest neighbor of x.

(i) **Multivariate statistics of kNN graph.** Let $k \in \mathbb{N}$ and $\mathcal{X} \subset \mathbb{R}^d$ a finite point set. For $x, y \in \mathcal{X}$, we put an undirected edge between x and y if x is one of the k nearest neighbors of y and/or y is a k nearest neighbor of x. Put

$$H^{(k)}(\mathcal{X}) := \text{sum of lengths of edges in kNN on } \mathcal{X}.$$

Applications

(i) **Multivariate statistics of kNN graph.** Let $k \in \mathbb{N}$ and $\mathcal{X} \subset \mathbb{R}^d$ a finite point set. For $x, y \in \mathcal{X}$, we put an undirected edge between x and y if x is one of the k nearest neighbors of y and/or y is a k nearest neighbor of x. Put

$$H^{(k)}(\mathcal{X}) := \text{sum of lengths of edges in kNN on } \mathcal{X}.$$

Theorem. Let $\mathcal{P}_{sg}$ be a Poisson point process on $[0,1]^d$ with intensity sg, g bounded away from 0 and ∞. Then for all $k_i \in \mathbb{N}$, $1 \leq i \leq m$, we have

$$\tilde{d}(s^{-1/2}(\bar{H}_s^{(k_1)}(\mathcal{P}_{sg}), ..., \bar{H}_s^{(k_m)}(\mathcal{P}_{sg})), N_\Sigma) \leq C s^{-1/d}, \ s \geq 1,$$

for $\tilde{d} \in \{d_2, d_3, d_{convex}\}$.

Applications

(ii) **Multivariate statistics of random geometric graph.** Fix $r > 0$. Let $\mathcal{X} \subset \mathbb{R}^d$ be a finite point set. Put $N_s^{(i)}(\mathcal{X})$ to be the number of components of random geometric graph $G(s^{1/d}\mathcal{X}, s^{1/d}r)$ of size i.

Applications

(ii) **Multivariate statistics of random geometric graph.** Fix $r > 0$. Let $\mathcal{X} \subset \mathbb{R}^d$ be a finite point set. Put $N_s^{(i)}(\mathcal{X})$ to be the number of components of random geometric graph $G(s^{1/d}\mathcal{X}, s^{1/d}r)$ of size i.

Theorem. Let $\mathcal{P}_{sg}$ be a Poisson point process on $[0,1]^d$ with intensity sg, g bounded away from 0 and ∞. When $r = \rho s^{-1/d}$ we have for all $i_j \in \mathbb{N}$, $1 \leq j \leq m$

$$\tilde{d}(s^{-1/2}(\bar{N}_s^{(i_1)}(\mathcal{P}_{sg}), ..., \bar{N}_s^{(i_m)}(\mathcal{P}_{sg})), N_\Sigma) \leq Cs^{-1/d}, \ s \geq 1,$$

for $\tilde{d} \in \{d_2, d_3, d_{convex}\}$.

Proof idea for d_{convex}

· **1. Stein:** Let $F = (F_1, ..., F_m)$ be a vector of Poisson functionals; let $\Sigma \in \mathbb{R}^{m \times m}$ be positive definite; $h : \mathbb{R}^m \to \mathbb{R}$.

Proof idea for d_{convex}

· **1. Stein:** Let $F = (F_1, ..., F_m)$ be a vector of Poisson functionals; let $\Sigma \in \mathbb{R}^{m\times m}$ be positive definite; $h: \ \mathbb{R}^m \to \mathbb{R}$.

· To assess the difference $\mathbb{E}\, h(F) - \mathbb{E}\,(h(N_\Sigma))$, where h belongs to a class of test functions, it is enough to assess the difference

$$\mathbb{E} \sum_{i=1}^{m} F_i \frac{\partial f_h}{\partial y_i}(F) - \frac{\partial^2 f_h}{\partial y_i^2}(F),$$

where $f_h : \mathbb{R}^m \to \mathbb{R}$ is a solution of the multivariate Stein equation:

$$\sum_{i=1}^{m} y_i \frac{\partial f}{\partial y_i}(y) - \frac{\partial^2 f}{\partial y_i^2}(y) = h(y) - \mathbb{E}\, h(N_\Sigma), \ y \in \mathbb{R}^m.$$

Proof idea for d_{convex}

· **1. Stein:** Let $F = (F_1, ..., F_m)$ be a vector of Poisson functionals; let $\Sigma \in \mathbb{R}^{m \times m}$ be positive definite; $h: \mathbb{R}^m \to \mathbb{R}$.

· To assess the difference $\mathbb{E}\, h(F) - \mathbb{E}\,(h(N_\Sigma))$, where h belongs to a class of test functions, it is enough to assess the difference

$$\mathbb{E} \sum_{i=1}^{m} F_i \frac{\partial f_h}{\partial y_i}(F) - \frac{\partial^2 f_h}{\partial y_i^2}(F),$$

where $f_h : \mathbb{R}^m \to \mathbb{R}$ is a solution of the multivariate Stein equation:

$$\sum_{i=1}^{m} y_i \frac{\partial f}{\partial y_i}(y) - \frac{\partial^2 f}{\partial y_i^2}(y) = h(y) - \mathbb{E}\, h(N_\Sigma),\ y \in \mathbb{R}^m.$$

· Given $t \in (0,1)$, and test function h, we introduce its smoothed version

$$h_{t,\Sigma}(y) := \int_{\mathbb{R}^m} h(\sqrt{t}z + \sqrt{1-t}y)\phi_\Sigma(z)dz,$$

where $\phi_\Sigma(z)$ is the density of N_Σ.

Proof idea for d_{convex}

· **2. Smoothing lemma:** Let $\mathcal{I}$ be collection of indicators of convex sets in $\mathbb{R}^m$.

$$d_{convex}(F, N_\Sigma) \leq \frac{4}{3} \sup_{h \in \mathcal{I}} |\mathbb{E}\, h_{t,\Sigma}(F) - \mathbb{E}\, h_{t,\Sigma}(N_\Sigma)| + \frac{20}{\sqrt{2}} m \frac{\sqrt{t}}{1-t}.$$

So it is enough to assess the difference of expectations over the smooth class of test functions $h_{t,\Sigma}$. This is accomplished with:

Proof idea for d_{convex}

· **2. Smoothing lemma:** Let $\mathcal{I}$ be collection of indicators of convex sets in $\mathbb{R}^m$.

$$d_{convex}(F, N_\Sigma) \le \frac{4}{3} \sup_{h \in \mathcal{I}} |\mathbb{E}\, h_{t,\Sigma}(F) - \mathbb{E}\, h_{t,\Sigma}(N_\Sigma)| + \frac{20}{\sqrt{2}} m \frac{\sqrt{t}}{1-t}.$$

So it is enough to assess the difference of expectations over the smooth class of test functions $h_{t,\Sigma}$. This is accomplished with:

· **3. Peccati + Zheng (Malliavin calculus on Poisson space):**

$$\mathbb{E}\, h_{t,\Sigma}(F) - \mathbb{E}\, h_{t,\Sigma}(N_\Sigma) = \sum_{i,j=1}^{m} \sigma_{ij} \mathbb{E}\, \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F)$$

$$- \sum_{k=1}^{m} \mathbb{E} \int_{\mathbb{X}} D_x \frac{\partial f_{t,h,\Sigma}}{\partial y_k}(F)(-D_x L^{-1} F_k) \lambda(dx).$$

· Here $f_{t,h,\Sigma}$ is the sol. to the MV Stein eq. associated with $h_{t,\Sigma}$.

Proof idea for d_{convex}

4. **Good sup norm and L^2 bounds on the 2nd derivatives of $f_{t,h,\Sigma}$.**

$$\mathbb{E}\, h_{t,\Sigma}(F) - \mathbb{E}\, h_{t,\Sigma}(N_\Sigma) = \sum_{i,j=1}^{m} \sigma_{ij} \mathbb{E}\, \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F)$$

$$- \sum_{k=1}^{m} \mathbb{E} \int_{\mathbb{X}} D_x \frac{\partial f_{t,h,\Sigma}}{\partial y_k}(F)(-D_x L^{-1} F_k) \lambda(dx).$$

Proof idea for d_{convex}

4. Good sup norm and L^2 bounds on the 2nd derivatives of $f_{t,h,\Sigma}$.

$$\mathbb{E}\, h_{t,\Sigma}(F) - \mathbb{E}\, h_{t,\Sigma}(N_\Sigma) = \sum_{i,j=1}^m \sigma_{ij} \mathbb{E}\, \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F)$$

$$- \sum_{k=1}^m \mathbb{E} \int_{\mathbb{X}} D_x \frac{\partial f_{t,h,\Sigma}}{\partial y_k}(F)(-D_x L^{-1} F_k)\lambda(dx).$$

However,

$$\sup_{h \in \mathcal{I}} \mathbb{E} \sum_{i,j=1}^m \left(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) \right)^2$$

$$\leq ||\Sigma^{-1}||_{\mathsf{op}}^2 \left(m^2 (\log t)^2 d_{convex}(F, N_\Sigma) + 444 m^{23/6} \right).$$

Combine steps 2, 3, 4 and choose parameter t in the right way.

Summary

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg.

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \quad A_i \subset W.$$

Summary

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg.

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \quad A_i \subset W.$$

We have found presumably optimal rates of multivariate normal convergence for the vector

$$\left(\frac{H_s^{(1)} - \mathbb{E}\, H_s^{(1)}}{\sqrt{s}}, \ldots, \frac{H_s^{(m)} - \mathbb{E}\, H_s^{(m)}}{\sqrt{s}} \right), \quad \text{as intensity } s \to \infty.$$

Summary

$\mathcal{P}_{sg}$, a Poisson point process on W with intensity sg.

$$H_s^{(i)} := H_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg}), \quad A_i \subset W.$$

We have found presumably optimal rates of multivariate normal convergence for the vector

$$\left(\frac{H_s^{(1)} - \mathbb{E}\, H_s^{(1)}}{\sqrt{s}}, ..., \frac{H_s^{(m)} - \mathbb{E}\, H_s^{(m)}}{\sqrt{s}} \right), \quad \text{as intensity } s \to \infty.$$

We needed to show that the scores satisfy two conditions:

(i) exponential stabilization

(ii) moment conditions.

Summary

Extensions:

(i) points in $\mathcal{P}_{sg}$ may carry independent marks

Summary

Extensions:
(i) points in $\mathcal{P}_{sg}$ may carry independent marks
(ii) rates of multivariate normal convergence for random measures

Summary

Extensions:

(i) points in $\mathcal{P}_{sg}$ may carry independent marks

(ii) rates of multivariate normal convergence for random measures

$$\mu_s^{(i)}(\mathcal{P}_{sg}) := \sum_{x \in \mathcal{P}_{sg} \cap A_i} \xi_s^{(i)}(x, \mathcal{P}_{sg})\delta_x, \quad A_i \subset W.$$

(iii) to be done: replace $\mathcal{P}_{sg}$ by more general input, including binomial input, Gibbsian input, and input with fast decaying correlations.

THANK YOU

Applications

(iii) **Multivariate statistics for equality of distributions.** Let $\mathcal{X} \subset \mathbb{R}^d$ be a finite point set. Consider the undirected nearest neighbors graph $NNG(\mathcal{X})$ on $\mathcal{X}$. Color the nodes of $\mathcal{X}$ with color i with probability $\pi_i, 1 \leq i \leq m$.

Let $H^{(i)}(\mathcal{X})$ be the number of edges in $NNG(\mathcal{X})$ which join nodes of color $i, 1 \leq i \leq m$.

Theorem. Let $\mathcal{P}_{sg}$ be a Poisson point process on $[0,1]^d$ with intensity sg, $g \in$ Lip$([0,1]^d)$, g bounded away from 0 and ∞. We have

$$\tilde{d}(s^{-1/2}(\bar{H}_s^{(1)}(\mathcal{P}_{sg}), ..., \bar{H}_s^{(m)}(\mathcal{P}_{sg})), N_\Sigma) \leq Cs^{-1/d}, \; s \geq 1,$$

for $\tilde{d} \in \{d_2, d_3, d_{convex}\}$.

This vector features in tests for equality of distributions.