Multivariate Second Order Poincaré Inequalities for Statistics in Geometric Probability

Joe Yukich (joint with Matthias Schulte)

Lehigh University

BIRS, Banff, April 13, 2022

Introduction, Main Results

· First order Poincaré inequality: measures the closeness of a random variable F to its mean.

· Second order Poincaré inequality: measures the closeness of F to a Gaussian r.v., with distance given by some metric on the space of distribution functions.

· Last, Peccati, Schulte (2016): Normal approximation on the Poisson space: Mehler's formula, second order Poincaré inequalities and stabilization, PTRF 165.

Introduction: Poisson functionals

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

· $\mathbf{N}$: space of all σ-finite counting measures on $\mathbb{X}$, equipped with σ-field generated by mappings $\nu \mapsto \nu(A), A \in \mathcal{F}$.

· F is a Poisson functional if there is measurable map $f : \mathbf{N} \to \mathbb{R}$ such that $F = f(\eta)$ a.s.

· Difference operators: $D_x F := f(\eta + \delta_x) - f(\eta)$.

· F belongs to the domain of difference operator, written $F \in \mathrm{DomD}$, if $\mathbb{E}\, F^2 < \infty$ and $\int_{\mathbb{X}} \mathbb{E}\,(D_x F)^2 \lambda(dx) < \infty$.

· $D^2_{x,y} F := D_x(D_y F) := f(\eta + \delta_x + \delta_x) - f(\eta + \delta_x) - f(\eta + \delta_y) + f(\eta)$.

Introduction

· First order Poincaré inequality:

$$\mathrm{Var}F \leq \int_{\mathbb{X}} \mathbb{E}\,(D_x F)^2 \lambda(dx).$$

· First and second order difference operators control the closeness to Gaussianity.

· We will be interested in the behavior of a vector

$$F = (F_1, ..., F_m), \quad m \in \mathbb{N},$$

of Poisson functionals. We want to compare F with an m-dimensional centered Gaussian vector N_Σ with covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$.

· We are not only interested in the weak convergence of F to N_Σ, but in quantitative bounds for the proximity between F and N_Σ.

Three distances between m-dimensional vectors

(i) $\mathcal{H}_m^{(2)}$: 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) $\mathcal{H}_m^{(3)}$: 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)

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

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

Main Results

· Let $F = (F_1, \dots, F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals $F_1, \dots, F_m \ \in \mathrm{DomD}$ with $\mathbb{E}\, F_i = 0$, $i \in \{1, \dots, m\}$. Define

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

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

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

Main results

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

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

and

$$d_2(F, N_\Sigma) \leq \|\Sigma^{-1}\|_{op} \|\Sigma\|_{op}^{1/2} \sum_{i,j=1}^{m} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)| + 2\|\Sigma^{-1}\|_{op} \|\Sigma\|_{op}^{1/2} \gamma_1$$
$$+ \|\Sigma^{-1}\|_{op} \|\Sigma\|_{op}^{1/2} \gamma_2 + \frac{\sqrt{2\pi} m^2}{8} \|\Sigma^{-1}\|_{op}^{3/2} \|\Sigma\|_{op} \gamma_3.$$

Literature

· Multivariate CLTs for vectors with dependency structures: Rai$\breve{c}$ (2004), Goldstein and Rinott (2005), Chen, Goldstein, and Shao (2011), Fang and Röllin (2015), Fang (2016), Reinert and Röllin (2009), Fang (2011), Fang and Koike (2021).

· Peccati and Zheng (2010): bounds in terms of inverse O-U operator and difference operator D.

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

Main results

For a vector $F = (F_1, \dots, F_m)$, $m \in \mathbb{N}$, of Poisson functionals with $\mathbb{E}\, F_i = 0$, $i \in \{1, \dots, m\}$, we use the abbreviations

$$D_x F := (D_x F_1, \dots, D_x F_m) \qquad \text{for } x \in \mathbb{X},$$
$$D^2_{x,y} F := (D^2_{x,y} F_1, \dots, D^2_{x,y} F_m) \text{ for } x, y \in \mathbb{X}.$$

$$\gamma_4 := \Bigg(\sum_{i,j=1}^{m} \int_{\mathbb{X}} \mathbb{E}\,(D_x F_i)^4 \,\lambda(\mathrm{d}x) + 6 \int_{\mathbb{X}^2} \big(\mathbb{E}\,(D^2_{x,y} F_i)^4\big)^{1/2} \big(\mathbb{E}\,(D_x F_j)^4\big)^{1/2} \lambda^2(\mathrm{d}(x,y)) + 3 \int_{\mathbb{X}^2} \big(\mathbb{E}\,(D^2_{x,y} F_i)^4\big)^{1/2} \big(\mathbb{E}\,(D^2_{x,y} F_j)^4\big)^{1/2} \lambda^2(\mathrm{d}(x,y)) \Bigg)^{1/2}.$$

Main results: CLTs for Poisson functionals

$$
\begin{aligned}
\gamma_5 := \Bigg(& 3 \sum_{i,j=1}^{m} \int_{\mathbb{X}^3} \Big(\mathbb{E}\, \mathbf{1}\{D^2_{x_1,y}F \neq \mathbf{0}, D^2_{x_2,y}F \neq \mathbf{0}\} \big(\|D_{x_1}F\| + \|D^2_{x_1,y}F\|\big)^{3/4} \\
& \times \big(\|D_{x_2}F\| + \|D^2_{x_2,y}F\|\big)^{3/4} |D_{x_1}F_i|^{3/2}\, |D_{x_2}F_i|^{3/2}\Big)^{2/3} \\
& \times \big(\mathbb{E}\, |D_{x_1}F_j|^3 |D_{x_2}F_j|^3\big)^{1/3}\, \lambda^3(\mathsf{d}(x_1,x_2,y)) \\
& + \sum_{i,j=1}^{m} \int_{\mathbb{X}^3} \Big(\mathbb{E}\, \big(\|D_{x_1}F\| + \|D^2_{x_1,y}F\|\big)^{3/2} \big(\|D_{x_2}F\| + \|D^2_{x_2,y}F\|\big)^{3/2}\Big)^{1/3} \\
& \times \Bigg(\frac{45}{2} \big(\mathbb{E}\, |D^2_{x_1,y}F_i|^3 |D^2_{x_2,y}F_i|^3\big)^{1/3} \big(\mathbb{E}\, |D_{x_1}F_j|^3 |D_{x_2}F_j)^3\big)^{1/3} \\
& + \frac{9}{2} \big(\mathbb{E}\, |D^2_{x_1,y}F_i|^3\, |D^2_{x_2,y}F_i|^3\big)^{1/3} \big(\mathbb{E}\, |D^2_{x_1,y}F_j|^3 |D^2_{x_2,y}F_j|^3\big)^{1/3} \Bigg) \\
& \lambda^3(\mathsf{d}(x_1,x_2,y)) \Bigg)^{1/3}.
\end{aligned}
$$

Main results: CLTs for Poisson functionals

$$\gamma_6^4 = 3 \sum_{i,j=1}^{m} \int_{\mathbb{X}^3} \Big(\mathbb{E}\, \mathbf{1}\{D^2_{x_1,y}F \neq \mathbf{0}, D^2_{x_2,y}F \neq \mathbf{0}\} \big(\|D_{x_1}F\|^2 + \|D^2_{x_1,y}F\|^2\big)^{3/4}$$

$$\times \big(\|D_{x_2}F\|^2 + \|D^2_{x_2,y}F\|^2\big)^{3/4} |D_{x_1}F_i|^{3/2}\, |D_{x_2}F_i|^{3/2}\Big)^{2/3}$$

$$\times \big(\mathbb{E}\, |D_{x_1}F_j|^3 |D_{x_2}F_j|^3\big)^{1/3}\, \lambda^3(\mathrm{d}(x_1,x_2,y))$$

$$+ \sum_{i,j=1}^{m} \int_{\mathbb{X}^3} \Big(\mathbb{E}\, \big(\|D_{x_1}F\|^2 + \|D^2_{x_1,y}F\|^2\big)^{3/2} \big(\|D_{x_2}F\|^2 + \|D^2_{x_2,y}F\|^2\big)^{3/2}\Big)^{1/3}$$

$$\times \Bigg(\frac{135}{8} \big(\mathbb{E}\, |D^2_{x_1,y}F_i|^3 |D^2_{x_2,y}F_i|^3\big)^{1/3} \big(\mathbb{E}\, |D_{x_1}F_j|^3 |D_{x_2}F_j|^3\big)^{1/3}$$

$$+ \frac{27}{8} \big(\mathbb{E}\, |D^2_{x_1,y}F_i|^3\, |D^2_{x_2,y}F_i|^3\big)^{1/3} \big(\mathbb{E}\, |D^2_{x_1,y}F_j|^3 |D^2_{x_2,y}F_j|^3\big)^{1/3} \Bigg)$$

$$\lambda^3(\mathrm{d}(x_1,x_2,y)).$$

Main results: CLTs for Poisson functionals

For a positive definite matrix $\Sigma \in \mathbb{R}^{m\times m}$ let $\Sigma^{1/2}$ be the positive definite matrix in $\mathbb{R}^{m\times m}$ such that $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$ and let $\Sigma^{-1/2} := (\Sigma^{1/2})^{-1}$.

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

$$d_{convex}(F, N_\Sigma) \leq 941 m^5 \max\{\|\Sigma^{-1/2}\|_{op}, \|\Sigma^{-1/2}\|_{op}^3\} \times \max\left\{\sum_{i,j=1}^{m} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)|, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6\right\}.$$

Main results ... remarks

(i) For each distance d_2, d_3, d_{convex} the bounds are of the same optimal order. It is more delicate to deal with non-smooth test functions when using Stein's method for normal approximation.

(ii) No logarithmic terms.

(iii) Existing results often require a.s. boundedness assumptions. In our set-up this would require

$$\sup_{x \in \mathbb{X}} |D_x(F_i)| \leq C \quad \text{a.s.}, \ i \in \{1, ..., m\}.$$

Ingredients of proof 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

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

where $f_h : \mathbb{R}^m \to \mathbb{R}$ solves the Stein equation:

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

· For h smooth one can give a formula for f_h, but for non-smooth h (e.g. indicators) it is unclear how to proceed.

2. Smoothing: Given $t \in (0,1)$, and a 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.

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

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

· This lemma actually holds for any m-dimensional random vector F.

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

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

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

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

· Here $f_{t,h,\Sigma} : \mathbb{R}^m \to \mathbb{R}$ solves the Stein equation for N_Σ:

$$f_{t,h,\Sigma}(y) := \frac{1}{2} \int_t^1 \frac{1}{1-s} \int_{\mathbb{R}^m} (h(\sqrt{s}z + \sqrt{1-s}y) - h(z))\, \varphi_\Sigma(z) \mathrm{d}z \mathrm{d}s, y \in \mathbb{R}^m.$$

· Now show that an upper bound for rhs of (*) involves terms $\gamma_1, ..., \gamma_6$ and factors such as $|\log t| \sqrt{d_{convex}(F, N_\Sigma)}$ and then choose t appropriately. This is done as follows....

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

· Some of the bounding terms for $d_{convex}(F, N_\Sigma)$ involve

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

· However,

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

$$\leq ||\Sigma^{-1}||_{\mathsf{op}} \left(m |\log t| \sqrt{d_{convex}(F, N_\Sigma)} + 24 m^{17/12} \right).$$

· This gives recursive inequality for $d_{convex}(F, N_\Sigma)$.

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

Summary so far..

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

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

· We bounded $d_2(F, N_\Sigma), d_3(F, N_\Sigma), d_{convex}(F, N_\Sigma)$ in terms of integrated difference operators and $\sum_{i,j=1}^m |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)|$.

· Are the integrated difference operators easy to evaluate? Same question for $\sum_{i,j=1}^m |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)|$.

· We show that our general results apply to a large class of Poisson functionals known as stabilizing Poisson functionals.

Poisson statistics in geometric probability

· Put $\mathbb{X} := W \subset \mathbb{R}^d$, $d \geq 2$, a fixed measurable set (W is a'window').

· Put $\lambda(dx) := sg(x)dx$, $g : \mathrm{Lip}(W) \to \mathbb{R}^+$.
· η_{sg}, a Poisson point process on W with intensity sg. Thus, for $A \subseteq W$, $|\eta_{sg} \cap A|$ is Poisson distributed with parameter $s \int_A g(x)dx$.

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

· **Poisson statistics**: $H_s^{(i)} := H_s^{(i)}(\eta_{sg}) := \sum_{x \in \eta_{sg}} \xi_s^{(i)}(x, \eta_{sg}), 1 \leq i \leq m$.

· Typically the $H_s^{(i)}$ describe a global feature of a random structure in terms of local contributions $\xi_s^{(i)}(x, \eta_{sg}),\ x \in \eta_{sg}$.

· **Poisson statistics**: $H_s^{(i)} := H_s^{(i)}(\eta_{sg}) := \sum_{x \in \eta_{sg}} \xi_s^{(i)}(x, \eta_{sg}), 1 \leq i \leq m.$

· **Goal**. Use the announced general results to find rates of multivariate normal convergence for the m-vector of Poisson functionals:

$$\left(\frac{H_s^{(1)} - \mathbb{E}\, H_s^{(1)}}{\sqrt{s}}, \ldots, \frac{H_s^{(m)} - \mathbb{E}\, H_s^{(m)}}{\sqrt{s}} \right)$$

as intensity $s \to \infty$.

Stabilization of scores

· The ith score $\xi_s^{(i)}$ generates the Poisson statistic

$$H_s^{(i)} := H_s^{(i)}(\eta_{sg}) := \sum_{x \in \eta_{sg}} \xi_s^{(i)}(x, \eta_{sg}).$$

· 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 η_{sg} if there are constants C_{stab} and $c_{stab} \in (0,\infty)$ such that

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

· 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.

· Stabilization often holds for scores which are locally defined.

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

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

Main results: rates of multivariate normal convergence

· $H_s^{(i)} := H_s^{(i)}(\eta_{sg}) := \sum_{x \in \eta_{sg}} \xi_s^{(i)}(x, \eta_{sg}), \; s \geq 1$

· Σ_s: covariance matrix of $s^{-1/2}(H_s^{(1)}, ..., H_s^{(m)})$. Assume Σ_s is positive definite for $s \geq 1$.

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

(i) are exponentially stabilizing, and

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

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

$$\tilde{d}\left(\left(\frac{H_s^{(1)} - \mathbb{E}\, H_s^{(1)}}{s^{1/2}}, ..., \frac{H_s^{(m)} - \mathbb{E}\, H_s^{(m)}}{s^{1/2}} \right), N_{\Sigma_s} \right) \leq \frac{C}{s^{1/2}}, \; s \geq 1. \; (*)$$

· The rate (*) is of correct order for d_{convex} if at least one of the scores $(\xi_s^{(i)})_{s \geq 1}, i \in \{1, ..., m\}$ is integer valued and Σ_s converges to a positive definite matrix.

Main results: rates of multivariate normal convergence

· $H_s^{(i)} := \sum_{x \in \eta_{sg}} \xi_s^{(i)}(x, \eta_{sg}),\ s \geq 1.$ $\bar{H}_s^{(i)} := H_s^{(i)} - \mathbb{E}\, H_s^{(i)}$. Given:

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

· Assume $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq m}$ is positive definite, η_{sg} a PPP on $W \subset \mathbb{R}^d$.

·**Theorem (Schulte + Y.)** Assume $(\xi_s^{(i)})_{s \geq 1}, 1 \leq i \leq m$,
(i) are exponentially stabilizing, and
(ii) satisfy the p-moment condition for some $p > 6$.
Then for $\tilde{d} \in \{d_2, d_3, d_{convex}\}$ we have the sharp bound

$$\tilde{d}(s^{-1/2}(\bar{H}_s^{(1)}, ..., \bar{H}_s^{(m)}), N_\Sigma) \leq C s^{-1/2} + C \sum_{i,j=1}^{m} \left| \sigma_{ij} - \frac{\mathrm{Cov}(H_s^{(i)}, H_s^{(j)})}{s} \right|$$

$$\leq C s^{-1/d},\ s \geq 1.$$

Extensions

· (i) replace Poisson functionals by Poisson measures:

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

· (ii) replace η_{sg} with a marked Poisson point process, where each Poisson pt in η_{sg} carries an independent mark.

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 η_{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

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

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

Applications

(ii) **Multivariate statistics 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 $\mathcal{X}$. With probability $\pi_i, 1 \leq i \leq m$, independently color the nodes of $\mathcal{X}$ with color i. These are 'marks'.

· 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 η_{sg} be the above marked Poisson point process on $[0,1]^d$ with intensity sg, g bounded away from 0 and ∞. We have

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

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

· This vector features in tests for equality of distributions.

Applications

(iii) **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 η_{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 $i_j \in \mathbb{N}$, $1 \leq j \leq m$

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

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

References

Last, G., Peccati, G. and M. Schulte (2016) Normal approximation on the Poisson space: Mehler's formula, second order Poincaré inequalities and stabilization, PTRF, 165, 667-723.

Schulte, M. and J. E. Yukich (2019) Multivariate second order Poincaré inequalities for Poisson functionals, EJP, 24, paper 130.

Schulte, M. and J. E. Yukich (2022) Rates of multivariate normal approximation for statistics in geometric probability, Ann. Appl. Prob., to appear.

THANK YOU