Asymptotic Theory for Statistics of Geometric Structures

Joe Yukich

Universidad Carlos III, October 11, 2019

Introduction

- $\mathcal{X} \subset \mathbb{R}^d$ random finite point set.

- Convex geometry. How many vertices in convex hull of $\mathcal{X}$?

Introduction

· $\mathcal{X} \subset \mathbb{R}^d$ random finite point set.

· Convex geometry. How many vertices in convex hull of $\mathcal{X}$?

· Stochastic geometry. Fix $\rho > 0$. At each point of $\mathcal{X}$ place a ball of radius ρ. What is volume of the union of such balls? Number of components?

Introduction

· $\mathcal{X} \subset \mathbb{R}^d$ random finite point set.

· Convex geometry. How many vertices in convex hull of $\mathcal{X}$?

· Stochastic geometry. Fix $\rho > 0$. At each point of $\mathcal{X}$ place a ball of radius ρ. What is volume of the union of such balls? Number of components?

· Statistical physics. RSA packing.

Introduction

· $\mathcal{X} \subset \mathbb{R}^d$ random finite point set.

· Convex geometry. How many vertices in convex hull of $\mathcal{X}$?

· Stochastic geometry. Fix $\rho > 0$. At each point of $\mathcal{X}$ place a ball of radius ρ. What is volume of the union of such balls? Number of components?

· Statistical physics. RSA packing.

· Graph and networks. $L_G(\mathcal{X}) :=$ length of graph G on $\mathcal{X}$. What is the behavior of $L_G(\mathcal{X})$ for large $\mathcal{X}$?

· The random variable X has density $\kappa(x)$ if

$$P(X \in A) = \int_A \kappa(x) dx.$$

· Theorem (Beardwood, Halton, Hammersley (1959)): $X_i, 1 \le i \le n$, i.i.d. with density $\kappa(x)$ on $[0,1]^d$. Then

$$\lim_{n \to \infty} \frac{L_{MST}(\{X_1, ..., X_n\})}{n^{(d-1)/d}} \stackrel{P}{=} \gamma_{MST}(d) \int_{[0,1]^d} \kappa(x)^{(d-1)/d} dx.$$

Introduction

Questions pertaining to statistics of 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 statistics of 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 $\xi(x, \mathcal{X}),\ x \in \mathcal{X}$.

Introduction

Questions pertaining to statistics of 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 $\xi(x, \mathcal{X}),\ x \in \mathcal{X}$.

We give some examples.

Random graphs

$\mathcal{X} \subset \mathbb{R}^d$ finite; let $G(\mathcal{X})$ be a graph on $\mathcal{X}$.

(a) For $x \in \mathcal{X}$, put

$$\xi(x, \mathcal{X}) := \frac{1}{2}(\text{sum of lengths of edges in graph incident to x}).$$

Then $\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X})$ gives the total edge length of $G(\mathcal{X})$.

Random graphs

$\mathcal{X} \subset \mathbb{R}^d$ finite; let $G(\mathcal{X})$ be a graph on $\mathcal{X}$.

(a) For $x \in \mathcal{X}$, put

$$\xi(x, \mathcal{X}) := \frac{1}{2}(\text{sum of lengths of edges in graph incident to x}).$$

Then $\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X})$ gives the total edge length of $G(\mathcal{X})$.

(b) $k \in \mathbb{N}$; $\xi_k(x, \mathcal{X}) = \frac{1}{k+1}(\text{number of k-simplices containing x})$.

Then

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

gives the number of k-simplices in $G(\mathcal{X})$.

Random convex hulls

· $\mathcal{X} \subset \mathbb{R}^d$ finite. Let $\mathrm{co}(\mathcal{X})$ denote the convex hull of $\mathcal{X}$.

· For $x \in \mathcal{X}$, $k \in \{0, 1, ..., d-1\}$, we put

$f_k(x, \mathcal{X}) := \frac{1}{k+1}(\text{number of k} - \text{dimensional faces containing } x).$

Random convex hulls

· $\mathcal{X} \subset \mathbb{R}^d$ finite. Let $\mathrm{co}(\mathcal{X})$ denote the convex hull of $\mathcal{X}$.

· For $x \in \mathcal{X}$, $k \in \{0, 1, ..., d-1\}$, we put

$f_k(x, \mathcal{X}) := \frac{1}{k+1}(\text{number of k} - \text{dimensional faces containing } x).$

· Total number of k-dimensional faces of $\mathrm{co}(\mathcal{X})$: $\quad \sum_{x \in \mathcal{X}} f_k(x, \mathcal{X})$.

· Rényi, Sulanke

Continuum percolation

$\mathcal{X} \subset \mathbb{R}^d$; join two points with an edge iff they are distant at most one.

$\xi_{\text{comp}}(x, \mathcal{X}) := (\text{size of component containing x})^{-1}$.

Continuum percolation

$\mathcal{X} \subset \mathbb{R}^d$; join two points with an edge iff they are distant at most one.

$\xi_{\mathrm{comp}}(x, \mathcal{X}) := (\text{size of component containing x})^{-1}$.

Component count in continuum percolation model on $\mathcal{X}$:

$$\sum_{x \in \mathcal{X}} \xi_{\mathrm{comp}}(x, \mathcal{X}).$$

Random sequential adsorption

· Unit volume balls $B_{1,n}, B_{2,n} ...,$ arrive sequentially and uniformly at random on the cube $[-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

Random sequential adsorption

· Unit volume balls $B_{1,n}, B_{2,n}...,$ arrive sequentially and uniformly at random on the cube $[-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

· The first ball $B_{1,n}$ is *packed*, and recursively for $i = 2, 3, \ldots$, the i-th ball $B_{i,n}$ is packed iff $B_{i,n}$ does not overlap any ball in $B_{1,n}, ..., B_{i-1,n}$ which has already been packed. If not packed, the i-th ball is discarded.

Random sequential adsorption

· Unit volume balls $B_{1,n}, B_{2,n}...,$ arrive sequentially and uniformly at random on the cube $[-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

· The first ball $B_{1,n}$ is *packed*, and recursively for $i = 2, 3, \ldots$, the i-th ball $B_{i,n}$ is packed iff $B_{i,n}$ does not overlap any ball in $B_{1,n}, ..., B_{i-1,n}$ which has already been packed. If not packed, the i-th ball is discarded.

· $\mathcal{X} \subset \mathbb{R}^d$ a temporally marked point set. Define the 'score' at $(x, \tau_x) \in \mathcal{X}$:

$$\xi((x,\tau_x),\mathcal{X}) := \begin{cases} 1 & \text{if ball centered at x with arrival time } \tau_x \text{ is accepted} \\ 0 & \text{otherwise.} \end{cases}$$

Random sequential adsorption

· Unit volume balls $B_{1,n}, B_{2,n}...,$ arrive sequentially and uniformly at random on the cube $[-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

· The first ball $B_{1,n}$ is *packed*, and recursively for $i = 2, 3, \ldots$, the i-th ball $B_{i,n}$ is packed iff $B_{i,n}$ does not overlap any ball in $B_{1,n}, ..., B_{i-1,n}$ which has already been packed. If not packed, the i-th ball is discarded.

· $\mathcal{X} \subset \mathbb{R}^d$ a temporally marked point set. Define the 'score' at $(x, \tau_x) \in \mathcal{X}$:

$$\xi((x, \tau_x), \mathcal{X}) := \begin{cases} 1 & \text{if ball centered at x with arrival time } \tau_x \text{ is accepted} \\ 0 & \text{otherwise.} \end{cases}$$

Total number of balls accepted: $\sum_{x \in \mathcal{X}} \xi((x, \tau_x), \mathcal{X})$.

Poisson input

· For purposes of exposition, we consider Poisson input on $\mathbb{R}^d$.

· By Poisson input, we mean a Poisson point process in $\mathbb{R}^d$. The Poisson point process (PPP) on $\mathbb{R}^d$ is the probabilist's way of placing points more or less uniformly at random in space. The PPP with rate (intensity) τ is denoted by $\mathcal{P}_\tau$ and has these properties:

(i) the number of points that $\mathcal{P}_\tau$ puts in disjoint sets are independent r.v.

Poisson input

· For purposes of exposition, we consider Poisson input on $\mathbb{R}^d$.

· By Poisson input, we mean a Poisson point process in $\mathbb{R}^d$. The Poisson point process (PPP) on $\mathbb{R}^d$ is the probabilist's way of placing points more or less uniformly at random in space. The PPP with rate (intensity) τ is denoted by $\mathcal{P}_\tau$ and has these properties:

(i) the number of points that $\mathcal{P}_\tau$ puts in disjoint sets are independent r.v.

(ii) the number of points of $\mathcal{P}_\tau$ in the set B is a Poisson r.v. with parameter equal to the product of τ and Lebesque measure of B.

Dimension estimators

$\mathcal{P} :=$ homogeneous rate one Poisson pt process on $\mathbb{R}^d$, $x \in \mathbb{R}^d$, $k \geq 3$.

$D_j := D_j(x, \mathcal{P}) :=$ dist. between x and its jth nearest neighbor in $\mathcal{P}$.

Dimension estimators

$\mathcal{P}$:= homogeneous rate one Poisson pt process on $\mathbb{R}^d$, $x \in \mathbb{R}^d$, $k \geq 3$.

$D_j := D_j(x, \mathcal{P})$:= dist. between x and its jth nearest neighbor in $\mathcal{P}$.
We have:

$$(k-2)\left(\sum_{j=1}^{k-1} \log \frac{D_k}{D_j}\right)^{-1} \stackrel{\mathcal{D}}{=} d(k-2)(\Gamma_{k-1,1})^{-1}.$$

Expectation of LHS is d.

In other words the LHS is an unbiased estimator of dimension for any $k \geq 3$ (Bickel + Levina).

$D_j := D_j(x, \mathcal{P}) :=$ dist. between x and its jth nearest neighbor in $\mathcal{P}$. We have

$$\mathbb{E}\left[(k-2)\left(\sum_{j=1}^{k-1}\log\frac{D_k}{D_j}\right)^{-1}\right] = d.$$

$D_j := D_j(x, \mathcal{P})$:= dist. between x and its jth nearest neighbor in $\mathcal{P}$. We have

$$\mathbb{E}\left[(k-2)\left(\sum_{j=1}^{k-1}\log\frac{D_k}{D_j}\right)^{-1}\right] = d.$$

Let $\{X_i\}_{i=1}^n$ be i.i.d. on manifold $\mathcal{M} \subset \mathbb{R}^m$; $d := \dim(\mathcal{M}) \leq m$ **unknown**.

Problem. Estimate intrinsic dimension d.

$D_j := D_j(x, \mathcal{P})$:= dist. between x and its jth nearest neighbor in $\mathcal{P}$. We have

$$\mathbb{E}\left[(k-2)\left(\sum_{j=1}^{k-1}\log\frac{D_k}{D_j}\right)^{-1}\right] = d.$$

Let $\{X_i\}_{i=1}^n$ be i.i.d. on manifold $\mathcal{M} \subset \mathbb{R}^m$; $d := \dim(\mathcal{M}) \leq m$ **unknown**.

Problem. Estimate intrinsic dimension d. Fix $k \geq 3$. Put $\mathcal{X}_n := \{X_i\}_{i=1}^n$. Define (spatially correlated) 'score' at X_1 wrt $\mathcal{X}_n$, by

$$\xi_k(X_1, \mathcal{X}_n) := (k-2)\left(\sum_{j=1}^{k-1}\log\frac{D_k(X_1, \mathcal{X}_n)}{D_j(X_1, \mathcal{X}_n)}\right)^{-1}.$$

$D_j := D_j(x, \mathcal{P}) :=$ dist. between x and its jth nearest neighbor in $\mathcal{P}$. We have

$$\mathbb{E}\left[(k-2)\left(\sum_{j=1}^{k-1}\log\frac{D_k}{D_j}\right)^{-1}\right] = d.$$

Let $\{X_i\}_{i=1}^n$ be i.i.d. on manifold $\mathcal{M} \subset \mathbb{R}^m$; $d := \dim(\mathcal{M}) \leq m$ **unknown**.

Problem. Estimate intrinsic dimension d. Fix $k \geq 3$. Put $\mathcal{X}_n := \{X_i\}_{i=1}^n$. Define (spatially correlated) 'score' at X_1 wrt $\mathcal{X}_n$, by

$$\xi_k(X_1, \mathcal{X}_n) := (k-2)\left(\sum_{j=1}^{k-1}\log\frac{D_k(X_1, \mathcal{X}_n)}{D_j(X_1, \mathcal{X}_n)}\right)^{-1}.$$

Questions (i) Fix $k \geq 3$. What conditions on $\mathcal{M}$ insure

$$\lim_{n\to\infty}\mathbb{E}\left[\xi_k(X_1, \mathcal{X}_n)\right] = \dim\mathcal{M}?$$

(ii) Are the sums $\sum_{i\leq n}\xi_k(X_i, \mathcal{X}_n)$ asymptotically normal?

General questions

· When $\mathcal{X} \subset \mathbb{R}^d$ is a random pt configuration, we have seen that the sums

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

describe a global feature of some random structure.

General questions

· When $\mathcal{X} \subset \mathbb{R}^d$ is a random pt configuration, we have seen that the sums

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

describe a global feature of some random structure.

· What is the distribution of these sums for large random pt configurations $\mathcal{X}$?

· Laws of large numbers?

· Central limit theorems?

Goals

$\mathcal{P}$: a rate one Poisson point process on $\mathbb{R}^d$.

Restrict $\mathcal{P}$ to windows: $W_n := [-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

Goals

$\mathcal{P}$: a rate one Poisson point process on $\mathbb{R}^d$.

Restrict $\mathcal{P}$ to windows: $W_n := [-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

Goal. Given a score function $\xi(\cdot, \cdot)$ defined on pairs $(x, \mathcal{X})$, given a pt process $\mathcal{P}$, we seek the limit theory (LLN, CLT, variance asymptotics) for the total score

$$H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n)$$

and total measure

$$\mu_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n) \delta_{n^{-1/d} x}.$$

Goals

$\mathcal{P}$: a rate one Poisson point process on $\mathbb{R}^d$.

Restrict $\mathcal{P}$ to windows: $W_n := [-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

Goal. Given a score function $\xi(\cdot, \cdot)$ defined on pairs $(x, \mathcal{X})$, given a pt process $\mathcal{P}$, we seek the limit theory (LLN, CLT, variance asymptotics) for the total score

$$H_n^{\xi} := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n)$$

and total measure

$$\mu_n^{\xi} := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n) \delta_{n^{-1/d}x}.$$

Tractable problems must be *local* in the sense that points far away from x should not play a role in the evaluation of the score $\xi(x, \mathcal{P} \cap W_n)$.

Stabilization

We assume translation invariant scores: $\xi(x, \mathcal{X}) = \xi(\mathbf{0}, \mathcal{X} - x)$.

Stabilization

We assume translation invariant scores: $\xi(x, \mathcal{X}) = \xi(\mathbf{0}, \mathcal{X} - x)$.

Key Definition. ξ is *stabilizing* wrt Poisson pt process $\mathcal{P}$ on $\mathbb{R}^d$ if there is $R := R^{\xi}(\mathcal{P}) < \infty$ a.s. (a 'radius of stabilization') such that

$$\xi(\mathbf{0}, \mathcal{P} \cap B_R(\mathbf{0})) = \xi(\mathbf{0}, (\mathcal{P} \cap B_R(\mathbf{0})) \cup \mathcal{A}).$$

for any locally finite $\mathcal{A} \subset B_R^c(\mathbf{0})$.

Stabilization

We assume translation invariant scores: $\xi(x, \mathcal{X}) = \xi(\mathbf{0}, \mathcal{X} - x)$.

Key Definition. ξ is *stabilizing* wrt Poisson pt process $\mathcal{P}$ on $\mathbb{R}^d$ if there is $R := R^\xi(\mathcal{P}) < \infty$ a.s. (a 'radius of stabilization') such that

$$\xi(\mathbf{0}, \mathcal{P} \cap B_R(\mathbf{0})) = \xi(\mathbf{0}, (\mathcal{P} \cap B_R(\mathbf{0})) \cup \mathcal{A}).$$

for any locally finite $\mathcal{A} \subset B_R^c(\mathbf{0})$.

ξ is *exponentially stabilizing* wrt $\mathcal{P}$ if there is a constant $c \in (0, \infty)$ such that

$$\mathbb{P}[R^\xi(\mathbf{0}, \mathcal{P}) \geq r] \leq c \exp(-\frac{r}{c}), \quad r \in [1, \infty).$$

Stabilization

Main idea: under stabilization conditions on ξ, the sums

$$\sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n)$$

should behave like a sum of weakly dependent random variables

Stabilization

$\mathcal{P}$: rate one Poisson pt process on $\mathbb{R}^d$; consider total edge length of the nearest neighbor graph on $\mathcal{P}$.

For $x \in \mathcal{P}$, put

$$\xi(x, \mathcal{P}) := \begin{cases} \frac{1}{2}|x - x_{NN}| & \text{if } x \text{and } x_{NN} \text{are mutual nearest neighbors} \\ |x - x_{NN}| & \text{otherwise.} \end{cases}$$

Stabilization

$\mathcal{P}$: rate one Poisson pt process on $\mathbb{R}^d$; consider total edge length of the nearest neighbor graph on $\mathcal{P}$.

For $x \in \mathcal{P}$, put

$$\xi(x, \mathcal{P}) := \begin{cases} \frac{1}{2}|x - x_{NN}| & \text{if } x \text{and } x_{NN} \text{are mutual nearest neighbors} \\ |x - x_{NN}| & \text{otherwise.} \end{cases}$$

Then $\sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n)$ gives the total edge length of nearest neighbors graph on the window W_n.

The radius of stabilization is

$$R^{\xi}(x, \mathcal{P}) := 2|x - x_{NN}|.$$

Moment condition

$\mathcal{P}$: Poisson pt process on $\mathbb{R}^d$.

Definition. ξ satisfies the p moment condition wrt $\mathcal{P}$ if

$$\sup_n \sup_{x,y \in \mathbb{R}^d} \mathbb{E}\, |\xi(x, \mathcal{P} \cup \{y\})|^p < \infty.$$

Weak law of large numbers for Poisson input $\mathcal{P}$

Let $\mathcal{P}$ be a rate 1 Poisson pt process on $\mathbb{R}^d$; $W_n := [\frac{-n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

$$H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n).$$

Weak law of large numbers for Poisson input $\mathcal{P}$

Let $\mathcal{P}$ be a rate 1 Poisson pt process on $\mathbb{R}^d$; $W_n := [\frac{-n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

$$H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n).$$

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (1, \infty)$, then

$$|n^{-1} \mathbb{E}\, H_n^\xi - \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P} \cup \{\mathbf{0}\})| \leq \epsilon_n.$$

Weak law of large numbers for Poisson input $\mathcal{P}$

Let $\mathcal{P}$ be a rate 1 Poisson pt process on $\mathbb{R}^d$; $W_n := [\frac{-n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

$$H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n).$$

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (1, \infty)$, then

$$|n^{-1} \mathbb{E}\, H_n^\xi - \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P} \cup \{\mathbf{0}\})| \leq \epsilon_n.$$

$\epsilon_n = O(n^{-1/d})$ if ξ is exponentially stabilizing wrt $\mathcal{P}$.

Weak law of large numbers for Poisson input $\mathcal{P}$

Let $\mathcal{P}$ be a rate 1 Poisson pt process on $\mathbb{R}^d$; $W_n := [\frac{-n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$.

$$H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n).$$

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (1, \infty)$, then

$$|\frac{1}{n}\mathbb{E}\, H_n^\xi - \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P} \cup \{\mathbf{0}\})| \le \epsilon_n.$$

- We may replace $\mathcal{P}_n$ with n i.i.d. uniform r.v. $\{X_i\}_{i=1}^n$ on $[-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$:

$$\lim_{n \to \infty} n^{-1} \mathbb{E} \sum_{i=1}^n \xi(X_i, \{X_i\}_{i=1}^n) = \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P} \cup \{\mathbf{0}\}).$$

Weak law of large numbers

What about laws of large numbers on non-uniform input?

Again, we first consider Poisson input with a non-uniform intensity.

Weak law of large numbers

What about laws of large numbers on non-uniform input?

Again, we first consider Poisson input with a non-uniform intensity.

Let $\mathcal{P}_{ng}$ be a Poisson pt process with intensity ng, i.e. the number of points of $\mathcal{P}_{ng}$ in a Borel set B is Poisson r.v. with parameter $n \int_B g(x)dx$ and the number of points in disjoint sets are independent r.v.

It is the case that for stabilizing, trans. invariant ξ we have as $n \to \infty$

$$\xi(n^{1/d}x, n^{1/d}\mathcal{P}_{ng}) = \xi(\mathbf{0}, n^{1/d}(\mathcal{P}_{ng} - x)) \xrightarrow{\mathcal{D}} \xi(\mathbf{0}, \mathcal{P}_{g(x)}).$$

Stabilization is a surrogate for continuity.

Weak law of large numbers for binomial input

Let $\{X_i\}_{i=1}^n$ be i.i.d. r.v. with density g on $[-\frac{1}{2}, \frac{1}{2}]^d$.

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (1, \infty)$, then

$$\lim_{n\to\infty} n^{-1}\mathbb{E} \sum_{i=1}^{n} \xi(n^{1/d}X_i, n^{1/d}\{X_i\}_{i=1}^n)$$

$$= \int_{[-\frac{1}{2},\frac{1}{2}]^d} \mathbb{E}\,[\xi(\mathbf{0}, \mathcal{P}_{g(x)} \cup \{\mathbf{0}\})]g(x)dx.$$

It is possible to simplify the right-hand side....

Weak law of large numbers for binomial input

For any Poisson point process $\mathcal{P}_\tau$ of intensity τ we have $\mathcal{P}_\tau \stackrel{\mathcal{D}}{=} \tau^{-1/d}\mathcal{P}_1$.

If the score function ξ measures edge length, then $\xi(ax, a\mathcal{X}) = a\xi(x, \mathcal{X})$.

Thus $\xi(\mathbf{0}, \mathcal{P}_\tau) \stackrel{\mathcal{D}}{=} \xi(0, \tau^{-1/d}\mathcal{P}_1) = \tau^{-1/d}\xi(\mathbf{0}, \mathcal{P}_1)$. Thus

Weak law of large numbers for binomial input

For any Poisson point process $\mathcal{P}_\tau$ of intensity τ we have $\mathcal{P}_\tau \stackrel{\mathcal{D}}{=} \tau^{-1/d}\mathcal{P}_1$.

If the score function ξ measures edge length, then $\xi(ax, a\mathcal{X}) = a\xi(x, \mathcal{X})$.

Thus $\xi(\mathbf{0}, \mathcal{P}_\tau) \stackrel{\mathcal{D}}{=} \xi(0, \tau^{-1/d}\mathcal{P}_1) = \tau^{-1/d}\xi(\mathbf{0}, \mathcal{P}_1)$. Thus

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (1, \infty)$, then

$$\begin{aligned}\mathbb{E} \sum_{i=1}^{n} \xi(n^{1/d}X_i, n^{1/d}\{X_i\}_{i=1}^{n}) &\to \int_{[-\frac{1}{2},\frac{1}{2}]^d} \mathbb{E}\,[\xi(\mathbf{0}, \mathcal{P}_{g(x)} \cup \{\mathbf{0}\})]g(x)dx \\ &= \int_{[-\frac{1}{2},\frac{1}{2}]^d} \mathbb{E}\,[\xi(\mathbf{0}, \mathcal{P}_1 \cup \{\mathbf{0}\})]g(x)^{(d-1)/d}dx \\ &= \mathbb{E}\,[\xi(\mathbf{0}, \mathcal{P}_1 \cup \{\mathbf{0}\})] \int_{[-\frac{1}{2},\frac{1}{2}]^d} g(x)^{(d-1)/d}dx\end{aligned}$$

Gaussian fluctuations for Poisson input $\mathcal{P}$ on $\mathbb{R}^d$

Recall $H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n)$.

Thm (CLT): Assume ξ is exponentially stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (5, \infty)$. Then

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left[\frac{H_n^\xi - \mathbb{E}\, H_n^\xi}{\sqrt{\mathrm{Var} H_n^\xi}} \le t \right] - \mathbb{P}[N(0,1) \le t] \right| \le \epsilon_n.$$

Gaussian fluctuations for Poisson input $\mathcal{P}$ on $\mathbb{R}^d$

Recall $H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(x, \mathcal{P} \cap W_n)$.

Thm (CLT): Assume ξ is exponentially stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (5, \infty)$. Then

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left[\frac{H_n^\xi - \mathbb{E}\, H_n^\xi}{\sqrt{\mathrm{Var} H_n^\xi}} \le t \right] - \mathbb{P}[N(0,1) \le t] \right| \le \epsilon_n.$$

Penrose + Y (2005), Penrose (2007): $\epsilon_n = O(\frac{(\log n)^{3d}}{\sqrt{n}})$.

Lachièze-Rey, Schulte, + Y (2019): $\epsilon_n = O(\frac{1}{\sqrt{\mathrm{Var} H_n^\xi}})$.

Variance asymptotics for Poisson input; volume order fluctuations

Given homogenous rate 1 Poisson input $\mathcal{P}$ on $\mathbb{R}^d$, and a score ξ, put

$$\sigma^2(\xi) := \mathbb{E}\,\xi^2(\mathbf{0},\mathcal{P}) + \\ + \int_{\mathbb{R}^d} [\mathbb{E}\,\xi(\mathbf{0},\mathcal{P}\cup\{x\})\xi(x,\mathcal{P}\cup\{\mathbf{0}\}) - \mathbb{E}\,\xi(\mathbf{0},\mathcal{P})\mathbb{E}\,\xi(x,\mathcal{P})]dx.$$

Variance asymptotics for Poisson input; volume order fluctuations

Given homogenous rate 1 Poisson input $\mathcal{P}$ on $\mathbb{R}^d$, and a score ξ, put

$$\sigma^2(\xi) := \mathbb{E}\,\xi^2(\mathbf{0},\mathcal{P}) + \\ + \int_{\mathbb{R}^d} [\mathbb{E}\,\xi(\mathbf{0},\mathcal{P}\cup\{x\})\xi(x,\mathcal{P}\cup\{\mathbf{0}\}) - \mathbb{E}\,\xi(\mathbf{0},\mathcal{P})\mathbb{E}\,\xi(x,\mathcal{P})]dx.$$

Thm (variance asymptotics): If ξ is exponentially stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (2,\infty)$, then

$$\lim_{n\to\infty} n^{-1}\mathrm{Var} H_n^{\xi} = \sigma^2(\xi) \in [0,\infty).$$

Extensions:

(i) Input can have fast decay of correlations

Extensions:

(i) Input can have fast decay of correlations

(ii) Multivariate CLT with rates of convergence

Extensions:

(i) Input can have fast decay of correlations

(ii) Multivariate CLT with rates of convergence

(iii) Input on manifolds

(iv) Our approach gives limit theory for the measures:

$$\mu_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n) \delta_{n^{-1/d}x}.$$

THANK YOU