Uniform limit theorems for point processes

Giacomo Francisci [a]

joint work with Anand N. Vidyashankar [b]

[a]Institute of Mathematical Finance, Ulm University
[b]Department of Statistics, George Mason University

Fall School Time Series, Random Fields and Beyond
September 26, 2024

Outline

1. Uniform limit theorems
   - Point processes
   - Measurability conditions
   - Main results
2. Applications
   - Tree-indexed random elements
   - Depth functions for point processes

Outline

Empirical measure

Let $\{Y_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. point processes, where

- $Y_i = \sum_{j=1}^{L_i} \delta_{X_{i,j}}$,
- δ_x is the Dirac measure at x,
- $X_{i,j}$ are random elements in a Polish space, and
- L_i is a random variable in $\mathbb{N} = \{1, 2, \dots\}$.

The empirical measure is given by

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} Y_i.$$

The case $L_i \equiv 1$ is well-known (Giné and Nickl; 2016; van der Vaart and Wellner; 1996).

Intensity measure

The intensity measure μ of the point process Y_1 is given for all Borel sets B by

$$\mu(B) = \mathbf{E}[Y_1(B)].$$

Let $\mathcal{F}$ be a uniformly bounded class of functions. Then, for all $f \in \mathcal{F}$

$$\mu(f) = \mathbf{E}[Y_1(f)],$$

where for any finite measure ν

$$\nu(f) := \int f \, d\nu.$$

We study convergence of the empirical process

$$\mu_n - \mu = \{\mu_n(f) - \mu(f)\}_{f \in \mathcal{F}}.$$

The space $\ell_\infty(\mathcal{F})$

- Since $\mathcal{F}$ is uniformly bounded, the empirical process $\mu_n - \mu$ takes values on the space $\ell_\infty(\mathcal{F})$ of bounded functionals on $\mathcal{F}$ with cylindrical σ-algebra.
- The space $\ell_\infty(\mathcal{F})$ is endowed with the norm

$$\|H\| := \|H\|_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |H(f)|, \text{ where } H \in \ell_\infty(\mathcal{F}).$$

- The (uncountable) supremum of random variables is not measurable in general.
- $\ell_\infty(\mathcal{F})$ is not separable unless $\mathcal{F}$ is finite.

Outline

1. Uniform limit theorems
   - Point processes
   - Measurability conditions
   - Main results
2. Applications
   - Tree-indexed random elements
   - Depth functions for point processes

Convergence in distribution

The outer expectation of any function T from a probability space $(\Omega, \Sigma, \mathbf{P})$ to the extended real numbers line $\bar{\mathbb{R}}$ is

$$\mathbf{E}^*[T] = \inf\{\mathbf{E}[U] \,:\, U \geq T,\, U : \Omega \to \bar{\mathbb{R}} \text{ measurable and } \mathbf{E}[U] \text{ exists}\}.$$

The infimum is achieved in the sense that there exists a measurable function $T^* \geq T$ such that $\mathbf{E}^*[T] = \mathbf{E}[T^*]$.

Definition (Hoffmann-Jørgensen (1991))

Let E be a metric space and $X_n : \Omega \to E$ (not necessarily measurable). We say that X_n converges in distribution to X with Borel law ν on E, that is, $X_n \xrightarrow{d^*} X$, if $\lim_{n\to\infty} \mathbf{E}^*[g(X_n)] = \int g \, d\nu$ for all bounded, continuous real functions g.

Measurability conditions

Definition (Giné and Zinn (1984))

A class of functions $\mathcal{F}$ is measurable if for each $a_1, \ldots, a_n, b \in \mathbb{R}$ and $n \in \mathbb{N}$, the quantity $\|\sum_{i=1}^{n} a_i Y_i + b\mu\|$ is measurable on $(\Omega, \Sigma, \mathbf{P})$.

We make the following assumptions:

(H1) $\mathbf{E}[L_1^2] < \infty$, and

(H2) $\mathcal{F}$ is a uniformly bounded non-empty measurable class of real functions.

Entropy conditions

- For any $\epsilon > 0$ the covering number of a pseudo-metric space (T, e) is

$$N(T, e, \epsilon) := \inf\{N : \exists\ t_1, \ldots, t_N \in T : \min_{i=1,\ldots,N} e(t_i, t) \leq \epsilon\ \forall t \in T\}.$$

- Sufficient conditions are given in terms of random metric entropy, that is, logarithms of the covering numbers $N(\mathcal{F}, e_{n,p}, \epsilon)$ of $\mathcal{F}$ w.r.t. the L^p empirical pseudo-distance $e_{n,p}$ given by

$$e_{n,p}^p(f, g) := \frac{1}{n} \sum_{i=1}^n Y_i(|f - g|^p), \quad f, g \in \mathcal{F}.$$

- These conditions hold if $\mathcal{F}$ is a VC-subgraph class.

Outline

Uniform LLN

Theorem

Assume **(H1)-(H2)**. *Then,* $\|\mu_n - \mu\| \xrightarrow{a.s.} 0$ *if one of the following conditions hold:*
(i) for all $\epsilon > 0$ *and some* $p \geq 1$ $\frac{1}{n}\log(N^*(\mathcal{F}, e_{n,p}, \epsilon)) \xrightarrow{p} 0$*, or*
(ii) for all $\delta > 0$

$$\lim_{n\to\infty} \mathbf{E}[\min(1, \frac{1}{\sqrt{n}} \int_0^\delta \sqrt{\log(N^*(\mathcal{F}, e_{n,2}, \epsilon))}\, d\epsilon)] = 0.$$

Proof idea

- Convergence in probability and in L^1 of $\|\mu_n - \mu\|$ is equivalent to convergence of $\|\mu_{\xi,n}\|$, where $\mu_{\xi,n} = \frac{1}{n}\sum_{i=1}^n \xi_i Y_i$ and $\{\xi_i\}_{i=1}^\infty$ is a sequence of independent Rademacher random variables.
- Conditionally on $\{Y_i\}_{i=1}^\infty$, the process $\{\sqrt{n}\mu_{\xi,n}(f)\}_{f\in\mathcal{F}}$ is subgaussian w.r.t. the distance $e_{n,2}$, that is, for all $\lambda \in \mathbb{R}$

$$\mathbf{E}_\xi[\exp(\lambda\sqrt{n}(\mu_{\xi,n}(f) - \mu_{\xi,n}(g)))] \leq \exp(\lambda^2 e_{n,2}^2(f,g)/2).$$

- Inequalities for subgaussian processes in terms of metric entropy and condition (i) or (ii) yield convergence in probability.
- General results on convergence of averages of random elements in a Banach space (extended to cylindrical σ-algebra) yield equivalence between convergence in probability and almost sure convergence (Kuelbs and Zinn; 1979; de Acosta; 1981).

Uniform CLT

Let $\mathcal{F}'_{\delta,p} := \{(f-g)^p : f,g \in \mathcal{F} \text{ and } \|f-g\|_{L^2(\mu)} \le \delta\}$.

Theorem

Assume **(H1)-(H2)** *and that the classes of functions* $\mathcal{F}'_{\infty,2}$ *and* $\{\mathcal{F}'_{\delta,1}\}_{\delta>0}$ *are measurable. If*

$$\lim_{\delta\to 0^+} \limsup_{n\to\infty} \mathbf{E}[\min(1, \int_0^\delta \sqrt{\log(N^*(\mathcal{F}, e_{n,2}, \epsilon))}\, d\epsilon)] = 0,$$

then

$$\sqrt{n}(\mu_n - \mu) \xrightarrow{d^*} W \text{ in } \ell_\infty(\mathcal{F}),$$

where W *is a Gaussian process with covariance function*

$$\mathbf{Cov}[W(f), W(g)] = \gamma(f,g) := \mathbf{E}[(Y_1(f) - \mu(f))(Y_1(g) - \mu(g))].$$

Proof idea

By Theorem 3.7.23 of Giné and Nickl (2016) it is enough to show that

1. the finite dimensional distributions of the process $W_n := \sqrt{n}(\mu_n - \mu)$ converge in law,
2. the space $(\mathcal{F}, \|\cdot\|_{L^2(\mu)})$ is totally bounded, and
3. the process W_n is asymptotically equicontinuous, that is, for all $\epsilon > 0$

$$\lim_{\delta\to 0^+} \limsup_{n\to\infty} \mathbf{P}^*\Big(\sup_{f,g\in\mathcal{F}\,:\,\|f-g\|_{L^2(\mu)}\le\delta} |W_n(f) - W_n(g)| \ge \epsilon\Big) = 0.$$

The proof of 2 uses the random metric condition and the uniform LLN for the class of functions $\mathcal{F}'_{\infty,2}$. The proof of 3 uses inequalities for subgaussian processes and the random metric condition.

Uniform rates of convergence

- Let $\mathcal{F} = \{\mathbf{I}_D \, : \, D \in \mathcal{D}\}$ be a class of indicators of a VC-class of sets $\mathcal{D}$ with VC-index $\mathbf{v}$.
- Let $S_n := \sum_{i=1}^{n} L_i$ and $S_{n,2} := \sum_{i=1}^{n} L_i^2$.

Theorem

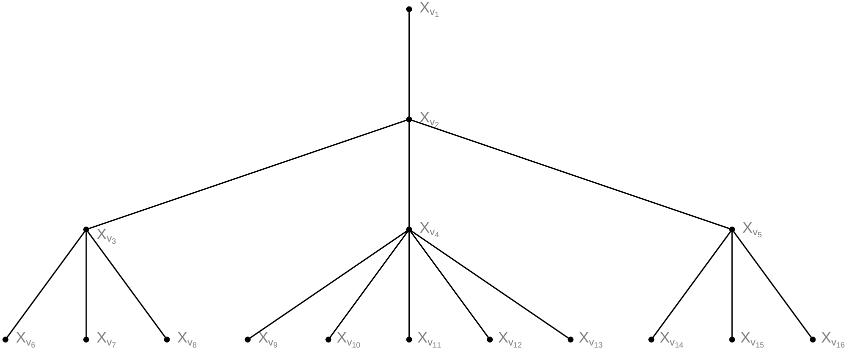
Assume **(H1)-(H2)**. *For all* $\alpha, \beta, \epsilon > 0$ *and* $n \geq 8 \cdot \mathbf{E}[L_1^2]/\epsilon^2$ *it holds that*

$$\mathbf{P}(\|\mu_n - \mu\| \geq \epsilon) \leq 16 \cdot (\alpha n)^{\mathbf{v}-1} \cdot \exp\left(-\frac{\epsilon^2}{2^5} \cdot \frac{n}{\beta}\right) + \mathbf{P}(S_n > \alpha n) + \mathbf{P}(S_{n,2} > \beta n).$$

Outline

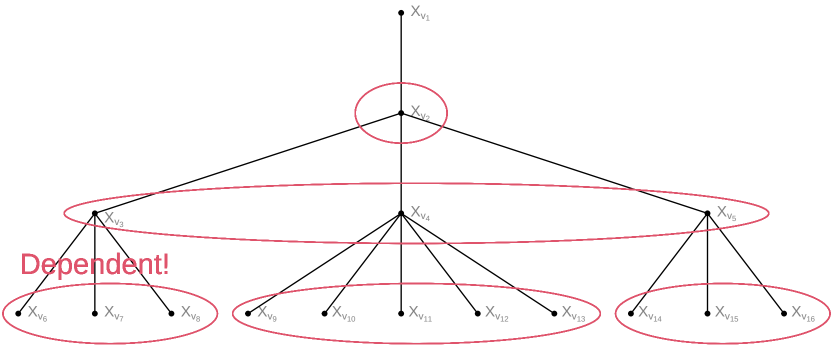
Tree-indexed random elements I

- Random elements $\{X_{v_i}\}_{i=1}^{n}$ indexed by vertices $v_1, \ldots, v_n$ of a random tree starting from the vertex $v_1 = \emptyset$.



Tree-indexed random elements I

- Random elements $\{X_{v_i}\}_{i=1}^{n}$ indexed by vertices $v_1, \dots, v_n$ of a random tree starting from the vertex $v_1 = \emptyset$.
- Random elements coming from same ancestor in the tree may be dependent.



Tree-indexed random elements II

- Using Ulam-Harris notation we write

$$\mathbb{V} = \{\emptyset\} \cup \cup_{k=1}^{\infty} \mathbb{N}^k$$

  for the set of all potential vertices.
- The random elements associated with the direct descendants of vertex v are given by the point process $Y_v = \sum_w \delta_{X_w}$.
- $\{Y_v\}_{v \in \mathbb{V}}$ are independent and identically distributed (i.i.d.).
- The set of actual vertices is denoted by $V \subset \mathbb{V}$.
- The vertices are ordered according to the breadth-first order induced by Ulam-Harris notation so that $V = \{v_1, v_2, \dots\}$.

Tree-indexed random elements III

- Let V_j be the vertex set at time j.
- $|V_j|$ is the cardinality of V_j.
- We obtain a Galton-Watson process $\{|V_j|\}_{j=0}^{\infty}$ with random elements attached to each vertex.
- By setting $Y_i = Y_{v_i}$ we obtain a sequence of i.i.d. point processes.

Lotka-Nagaev and Harris-type estimators

- The Lotka-Nagaev estimator $\hat{\mu}_j$ of the intensity measure μ is given by

$$\hat{\mu}_j(f) = \frac{1}{|V_j|} \sum_{v \in V_j} Y_v(f).$$

- The Harris-type estimator $\tilde{\mu}_j$ of μ is given by

$$\tilde{\mu}_j(f) = \frac{1}{\sum_{l=0}^{j} |V_l|} \sum_{i=1}^{\sum_{l=0}^{j} |V_l|} Y_{v_i}(f).$$

- When $f \equiv 1$ one obtains $Y_{v_i}(1) = L_i$ and the estimators reduce to the classical Lotka-Nagaev and Harris estimators of the mean of a supercritical Galton-Watson process, that is,

$$\hat{\mu}_j(1) = \frac{|V_{j+1}|}{|V_j|} \text{ and } \tilde{\mu}_j(1) = \frac{\sum_{l=1}^{j+1} |V_l|}{\sum_{l=0}^{j} |V_l|}.$$

Uniform converge for the Lotka-Nagaev estimator

Proposition

Assume **(H1)-(H2)**, $\mathbf{E}[L_1] > 1$, *and that* $\mathcal{F}$ *is a VC-subgraph class of functions. The following holds:*
(i) $\|\hat{\mu}_j - \mu\| \xrightarrow{a.s.} 0$ *and*
(ii) if $\mathcal{F}'_{\infty,2}$ *and* $\{\mathcal{F}'_{\delta,1}\}_{\delta>0}$ *are measurable, then* $|V_j|^{1/2}(\hat{\mu}_j - \mu) \xrightarrow{d^*} W$, *where* W *is the Gaussian process in the uniform CLT.*

Proof idea: We condition on $|V_j| = k$. Using the uniform CLT we see that $|V_j|^{1/2}(\hat{\mu}_j - \mu)$ is close to W for every large k. On the other hand, $\mathbf{P}(|V_j| = k) \to 0$ as $j \to \infty$ for every fixed k.

Uniform converge for the Harris-type estimator

Proposition

Assume **(H1)-(H2)** *and that $\mathcal{F}$ is a VC-subgraph class of functions. The following holds:*
(i) $\|\tilde{\mu}_j - \mu\| \xrightarrow{a.s.} 0$ *and*
(ii) if $\mathcal{F}'_{\infty,2}$ *and* $\{\mathcal{F}'_{\delta,1}\}_{\delta>0}$ *are measurable, then*

$$(\textstyle\sum_{l=0}^{j}|V_l|)^{1/2}(\tilde{\mu}_j - \mu) \xrightarrow{d^*} W.$$

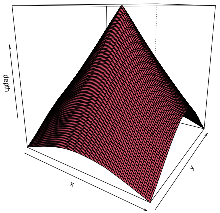
Part (ii) provides a uniform version of Theorem 3 of Kuelbs and Vidyashankar (2011) on convergence in law of Harris estimator. The proof requires an extension of the uniform CLT allowing for a random number of terms.

Outline

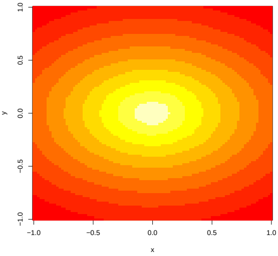
1. Uniform limit theorems
   - Point processes
   - Measurability conditions
   - Main results
2. Applications
   - Tree-indexed random elements
   - Depth functions for point processes

Depth functions

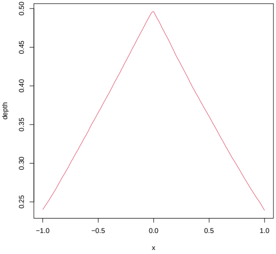
- Depth functions specify a center-outward order with respect to a finite measure ν on $\mathbb{R}^d$ (usually a probability measure).
- Depth functions contours yield multivariate quantiles.



Bivariate depth



Heat map



Section along x-axis

Half-space depth

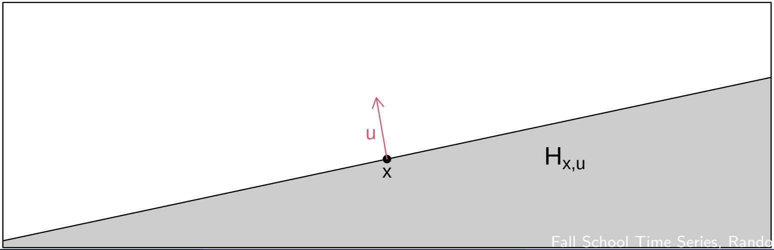
The half-space depth (Zuo and Serfling; 2000) of $x \in \mathbb{R}^d$ with respect to a finite measure ν is

$$D(x, \nu) = \inf_{u \in S^{d-1}} \nu(H_{x,u}),$$

where S^{d-1} is the unit sphere and

$$H_{x,u} = \{y \in \mathbb{R}^d \,:\, \langle y, u \rangle \leq \langle x, u \rangle\}$$

is the closed half-space with outer normal u and x on the boundary.



Properties

1. Affine-invariance: for any affine transformation T, that is, $T : \mathbb{R}^d \to \mathbb{R}^d$ given by $T(x) = Ax + b$ for a non-singular $d \times d$ matrix A and $b \in \mathbb{R}^d$,

$$D(T(x), \nu_T) = D(x, \nu), \quad x \in \mathbb{R}^d,$$

   where ν_T is the pushforward of ν.

2. If ν is half-space symmetric about $y \in \mathbb{R}^d$, that is,

$$\nu(H_{y,u}) \geq \nu(\mathbb{R}^d)/2, \quad u \in S^{d-1},$$

   then

$$D(y, \nu) \geq D(x, \nu), \quad x \in \mathbb{R}^d.$$

Properties

3. $D(\cdot, \nu)$ is monotonically non-increasing along rays from the point of maximum depth: if

$$D(y, \nu) \geq D(x, \nu), \quad x \in \mathbb{R}^d,$$

then

$$D(y + \alpha(x - y), \nu) \geq D(x, \nu), \quad \alpha \in [0, 1].$$

4. Vanishing at infinity:

$$\sup_{x \in \mathbb{R}^d : \|x\| \geq r} D(x, \nu) \to 0 \text{ as } r \to \infty.$$

Half-space depth for point processes

- The half-space depth of $x \in \mathbb{R}$ w.r.t. the intensity measure μ is given by
$$D(x, \mu) = \inf_{u \in S^{d-1}} \mu(H_{x,u})$$
- Similarly, the empirical half-space depth is
$$D(x, \mu_n) = \inf_{u \in S^{d-1}} \mu_n(H_{x,u}).$$

The following assumption ensures that the infimum is obtained:

(H3) $\mu(\partial H) = 0$ for all half-spaces $H \subset \mathbb{R}^d$.

We let $\mathcal{R}(\mu)$ be the set of points $x \in \mathbb{R}^d$ that have a unique minimizing direction $u_x \in S^{d-1}$. In particular, $D(x, \mu) = \mu(H_{x,u_x})$ (Massé; 2004).

Asymptotics for half-space depth

Proposition

Assume **(H1)**. *The following holds:*

(i) $\sup_{x\in\mathbb{R}^d} |D(x,\mu) - D(x,\mu_n)| \xrightarrow{a.s.} 0,$

(ii) If **(H3)** *holds true and* $A \neq \emptyset$ *is a closed subset of* $\mathcal{R}(\mu)$, *then*

$$\sqrt{n}(D(\cdot,\mu) - D(\cdot,\mu_n)) \xrightarrow{d^*} W \text{ in } \ell_\infty(A),$$

where W *is a Gaussian process.*
(iii) For all $\alpha, \beta, \epsilon > 0$ *and* $n \geq 8 \cdot \mathbf{E}[L_1^2]/\epsilon^2$

$$\mathbf{P}\big(\sup_{x\in\mathbb{R}^d} |D(x,\mu) - D(x,\mu_n)| \geq \epsilon\big) \leq 16 \cdot (\alpha n)^d \cdot \exp\left(-\frac{\epsilon^2}{2^5} \cdot \frac{n}{\beta}\right) + \mathbf{P}(S_n > \alpha n) + \mathbf{P}(S_{n,2} > \beta n).$$

Summary

1. Motivated by applications to tree-indexed random elements and depth functions, we study empirical point processes indexed by a class $\mathcal{F}$.
2. We provide sufficient conditions for the uniform LLN and CLT in terms of random metric entropy, which hold if $\mathcal{F}$ is VC-subgraph.
3. We derive uniform LLN and CLT for Lotka-Nagaev and Harris-type estimators.
4. We establish uniform consistency and asymptotic normality of the half-space depth based on the intensity measure of the point processes.

References

de Acosta, A. (1981). Inequalities for B-valued random vectors with applications to the strong law of large numbers, *The Annals of Probability* **9**: 157–161.

Francisci, G. and Vidyashankar, A. N. (2024). Functional limit laws for the intensity measure of point processes and applications, *arXiv preprint 2402.05087* .

Giné, E. and Nickl, R. (2016). *Mathematical foundations of infinite-dimensional statistical models*, Cambridge University Press.

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes, *The Annals of Probability* **12**: 929–998.

Hoffmann-Jørgensen, J. (1991). *Stochastic processes on Polish spaces*, Matematisk Institut Århus: Various publications series, Inst., Univ.

Kuelbs, J. and Vidyashankar, A. N. (2011). Weak convergence results for multiple generations of a branching process, *Journal of Theoretical Probability* **24**: 376–396.

Kuelbs, J. and Zinn, J. (1979). Some stability results for vector valued random variables, *The Annals of Probability* **7**: 75–84.

Massé, J.-C. (2004). Asymptotics for the Tukey depth process, with an application to a multivariate trimmed mean, *Bernoulli* **10**: 397–419.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes*, Springer.

Zuo, Y. and Serfling, R. (2000). General notions of statistical depth function, *The Annals of statistics* **28**: 461–482.