Normal Approximation in Strategic Network Formation

Michael P. Leung   Hyungsik Roger Moon

University of Southern California

ASSA 2019

Goal

- We develop a CLT for network statistics,

$$\frac{1}{n}\sum_{i=1}^{n}\psi(X_i, \mathcal{X}_n, W, A),$$

  where

  - X_i is a vector of homophilous attributes of node i,
  - $\mathcal{X}_n := \{X_1, ..., X_n\}$,
  - W is the set of all other node attributes,
  - $A = [A_{ij}]$ is the observed network on n nodes.

- A simple example is

$$\psi(X_i, \mathcal{X}_n, W, A) = \sum_{j \neq i} A_{ij}.$$

Contributions

- We derive conditions under which

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\psi_i - \mathbf{E}[\psi_i]) \xrightarrow{d} \mathcal{N}(0,\sigma^2).$$

- A key high level condition for the large-network CLT for dynamic network moments is (i) "stabilization" condition (e.g. Penrose (2007), Penrose and Yukich (2005), and Leung (2018)) and (ii) bounded moments of ψ_i.
- We apply our results to
  - nonparametric bounds on the average structural function of dynamic network formation,
  - network regressions (not today),
  - treatment effects with network spillovers (not today).
- We provide lower level conditions for "stabilization" in each application.
- We also propose inference procedures for $\mathbf{E}[\psi_i]$ (not today).

Related Literature

- Leung (2018), Menzel (2016): Law of large numbers for static models with strategic interactions.
- Estimation of subgraph and exponential random graph models: Boucher and Mourifié (2015), Chandrasekhar and Jackson (2015).
- Estimation of dyadic link formation without strategic network formation: Dzemski (2014), Graham (2017).
- Estimation static models of strategic network formation: Leung (2015), Ridder and Sheng (2016)
- Bayesian approaches: Christakis et al. (2010), Mele (2017).
- Large matching models: Agarwal and Diamond (2017), Fox (2017), Menzel (2016).
- Dynamic network models: Kuersteiner and Prucha (2018, dynamic spatial panels), Graham (2016, point identification, parametric)
- Proof of CLT draws heavily from techniques in Penrose (2003) and Penrose and Yukich (2001, 2003).

Outline

Application to Dynamic Model
- Setup and Object of Interest
- Weak Dependence

Main Result

Dynamic Network Formation Model I

Notations:

- $\mathcal{N}_n := \{1, ..., n\}$ is the set of nodes.
- Each node i is endowed with a type (X_i, Z_i).
- $X_i \in \mathbb{R}^d$: "position" of node i. It is a latent, continuously distributed, time invariant characteristic.
- $Z_i := (Z_{i0}, ..., Z_{iT})$: observed, potentially time varying attributes.
- Each node pair (i, j) are endowed with a random shock ζ_{ij}.

Network on n nodes evolves from period $t-1$ to t according to myopic best-response dynamics: for every $i, j \in \mathcal{N}_n$

$$A_{ij,t} = \mathbf{1}\{V(r_n^{-1}||X_i - X_j||, \underbrace{(A_{ij,t-1}, \max_k A_{ik,t-1}A_{jk,t-1})}_{S_{ij,t}}, Z_{it}, Z_{jt}, \zeta_{ij,t}) > 0\}.$$

Here,

Dynamic Network Formation Model II

- $V(\cdot)$:
  - $V(\cdot)$: is unknown.
  - $V(\cdot)$ is strictly monotonic in $\zeta_{ij,t}$.
  - homophily: V is decreasing in $r_n^{-1}\|X_i - X_j\|$.
- Homophily:
  - Nodes homophilous $(-r_n^{-1}\|X_i - X_j\|)$ in position.
  - sparsity: $r_n \to 0$ at a certain rate.
  - Examples: income, geographic location.
  - Can also interpret more abstractly as positions in latent social space, following latent space models (Hoff et al., 2002).
- $A_{ij,t-1}$: captures state dependence.
- $\max_k A_{ik,t-1}A_{jk,t-1}$: generates network clustering.
- Examples in literature include

Dynamic Network Formation Model III

  - risk-sharing networks in the rural Phillippines (Fafchangps and Gubert, 2007)
  - research partnerships in the biotechnology industry (Powell et al, 2005)
- Graham (2016) discusses the policy implications of distinguishing between incentives based on assortative matching (homophily) and those based on strategic interactions.
- Network formation models are also useful for forecasting the effects of counterfactual interventions (Mele, 2017) and as selection models for social interactions (Badev, 2013).

Initial Network

- Simple example: A_0 follows dyadic regression model,

$$A_{ij,0} = \mathbf{1}\left\{ V_0(r_n^{-1}||X_i - X_j||, Z_{i0}, Z_{j0}, \zeta_{ij,0}) > 0 \right\}. \qquad (1)$$

- Interpret as random meeting process prior to creatation of social connections.
- More generally, we can allow for strategic interactions similar to dynamic model, except $S_{ij,0}$ depends on A_0, not a lagged network. (Not today)

Application: ATE of Network Formation - I

- Goal: inference on ASF $\mu(\mathbf{s}, z, z')$, where

$$\mu(\mathbf{s}, z, z') = \int \mathbf{1}\{V(\delta, \mathbf{s}, z, z', \zeta) > 0\}\, \mathrm{d}F(\delta, \zeta),$$

  where F is the joint distribution of $(r_n^{-1}\|X_i - X_j\|, W_{ij,t})$.
- For notational simplicity, assume that V does not depend on (Z_{it}, Z_{jt}).
- Recall $S_{ij,t} = (A_{ij,t-1}, \max_k A_{ik,t-1}A_{jk,t-1})$.
- Examples of parameters of interests:
  - (i) $(\mu(1,0)) - \mu(0,0))/\mu(0,0)$: nonparametric measure of state dependence.
  - (ii) $(\mu(0,1)) - \mu(0,0))/\mu(0,0)$: nonparametric measure of transitivity.
- In general, these objects are not point-identified (e.g., Chernozhukov et al. (2013)).

Application: ATE of Network Formation - II

- We follow the idea in Chernozhukov et al. (2013).
- $\mathcal{S}_t(\mathbf{s})$ is the set of values of $\mathbf{S}_{ij} = (\mathbf{S}_{ij,t_0}, ..., \mathbf{S}_{ij,T})$ for which the tth component first equals $\mathbf{s}$ at time t.
- $\bar{\mathcal{S}}(\mathbf{s})$ is the set of values of $\mathbf{S}_{ij}$ for which $\mathbf{s}$ is never reached between t_0 and T.
- Define

$$\hat{A}_{ij}(\mathbf{s}) = \sum_{t=t_0+1}^{T} \mathbf{1}\{\mathbf{S}_{ij} \in \mathcal{S}_t(\mathbf{s})\} A_{ij,t}, \tag{2}$$

$$P_{ij}(\mathbf{s}) = \mathbf{1}\{\mathbf{S}_{ij} \in \bar{\mathcal{S}}(\mathbf{s})\}. \tag{3}$$

- Chernozhukov et al. (2013) showed that

$$\mu_\ell(\mathbf{s}) \leqslant \mu(\mathbf{s}) \leqslant \mu_u(\mathbf{s}), \tag{4}$$

for $\mu_\ell(\mathbf{s}) = \mathbf{E}[\hat{A}_{ij}(\mathbf{s})]$ and $\mu_u(\mathbf{s}) = \mu_\ell(\mathbf{s}) + \mathbf{E}[P_{ij}(\mathbf{s})]$.

Application: ATE of Network Formation - III

- Using (4), we obtain the following upper and lower bounds on percentage marginal effects:

$$\frac{\mu_\ell(\mathbf{s}') - \mu_u(\mathbf{s})}{\mu_u(\mathbf{s})} \leqslant \frac{\mu(\mathbf{s}') - \mu(\mathbf{s})}{\mu(\mathbf{s})} \leqslant \frac{\mu_u(\mathbf{s}') - \mu_\ell(\mathbf{s})}{\mu_\ell(\mathbf{s})}.$$

- We estimate the lower and upper bounds on the ASF using their scaled sample analogs

$$\hat{\mu}_\ell(\mathbf{s}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}\hat{A}_{ij}(\mathbf{s}) \quad \text{and} \quad \hat{\mu}_u(\mathbf{s}) = \hat{\mu}_\ell(\mathbf{s}) + \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}P_{ij}(\mathbf{s}). \qquad (5)$$

- Both $\hat{\mu}_\ell(\mathbf{s})$ and $\hat{\mu}_u(\mathbf{s})$ can be written as averages

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

For example, for $\hat{\mu}_\ell(\mathbf{s})$, $\psi_i = \sum_j \hat{A}_{ij}(\mathbf{s})$, a weighted degree of node i.

Application: ATE of Network Formation - IV

- We prove a CLT for general averages of node statistics $\{\psi_i\}_{i=1}^n$ of this type under new restrictions on the model primitives that ensure weak dependence.
- ATE of network formation: Fernandez-val and Weidner (2016) and Chen, Fernandez-val and Weidner (2018) - parametric dense network formation without network externality.

Outline

Application to Dynamic Model
Setup and Object of Interest
Weak Dependence

Main Result

‘

Dependence Structure I

- For simplicity, we consider a two-period case, where $t_0 = 0$ and $T = 1$, and the node statistic is the degree in period 1, $\psi_i = \sum_j A_{ij,1}$.
- We consider asymptotics where T is fixed and network size $n \to \infty$.
- The key component in establishing asymptotics is understanding and handling "dependence" between ψ_i and ψ_j.
- An example:

Dependence Structure II

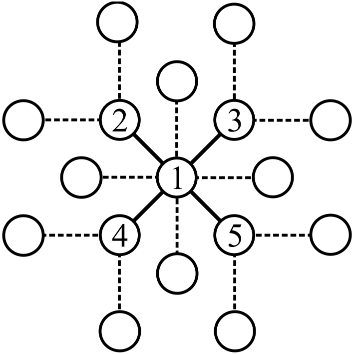

Figure: Dashed lines depict A_0, solid lines A_1.

- Denote $\mathcal{N}_{A_t}(i, K)$ is the K-neighborhood of node i in the network A_t.

Dependence Structure III

- Evidently, ψ_1 depends directly on $\mathcal{N}_{A_1}(1,1) = \{2,3,4,5\}$, but its actual "dependency neighborhood" is larger due to strategic interactions.
- For example, consider the link $A_{12,1}$. By the model specification, its realization depends only on (X_1, Z_1), (X_2, Z_2), $\zeta_{12,1}$, and A_0, the latter only through $\mathcal{N}_{A_0}(1,1) \cup \mathcal{N}_{A_0}(2,1)$, which are those in the figure connected by dotted lines to either node 1 or 2.
- Furthermore, if we were to remove all nodes from the network other than those in $\mathcal{N}_{A_0}(1,1) \cup \mathcal{N}_{A_0}(2,1)$, this would not change the set of links formed by nodes 1 and 2 in A_0.
- It follows that $A_{12,1}$ is invariant to the removal of $\mathcal{N} \backslash (\mathcal{N}_{A_0}(1,1) \cup \mathcal{N}_{A_0}(2,1))$ from the network. The same reasoning applies to $A_{1k,1}$ for $k = 3,4,5$.

Dependence Structure IV

- The realization of ψ_i is invariant to the removal of all nodes from the model, other than members of the set

$$J_i \equiv \mathcal{N}_{A_1}(i,1) \cup \bigcup_{j \in \mathcal{N}_{A_1}(i,1)} \mathcal{N}_{A_0}(j,1), \tag{6}$$

  which are all the nodes depicted in the figure.
- We call J_i the *relevant set* for ψ_i.
- Under a sparsity condition, the sizes of 1-neighborhoods are asymptotically bounded.
  → Hence, for any i, $|J_i| = O_p(1)$.
- Since J_i effectively represents a dependency neighborhood, $\{\psi_i\}_{i=1}^n$ ought to be "weakly dependent" (like a moving average in time series).
- We will show that ψ_i does satisfy one such notion, known as "stabilization," for which we can prove a CLT.

Dependence Structure V

- Two crucial properties used in this argument are that $T < \infty$ and the initial conditions model has no strategic interactions.
- Finiteness of T is important.
  - If $T = \infty$, then even under sparsity, $|J_i| = \infty$ a.s.
  - In the general model, we accommodate the "long-run" $T = \infty$ case by modeling the initial condition as a draw from a static strategic network-formation model, which informally represents a draw from the stationary distribution.
  - Also, $T < \infty$ is important because it justifies the claim used above that potential links in A_0 do not depend on the states of other potential links in that network.
  - With strategic interactions, potential links in A_0 are now dependent, so we require an additional weak-dependence condition that controls the strength of strategic interactions.

Outline

Application to Dynamic Model
Setup and Object of Interest
Weak Dependence

Main Result

General Set up I

- Recall the general network formation model

$$A_{ij,t} = \mathbf{1}\left\{V\left(r_n^{-1}||X_i - X_j||, S_{ij,t}, Z_{it}, Z_{jt}, \zeta_{ij,t}\right) > 0\right\}.$$

- $\zeta_{ij} := (\zeta_{ij,0}, \ldots, \zeta_{ij,T})$.
- $\{(X_i, Z_i)\}_{i\in\mathbb{N}}$ and $\{\zeta_{ij}\}_{i,j\in\mathbb{N}}$ are $\sim$ i.i.d. and mutually independent.
- $\mathcal{X}_n := \{X_i\}_{i=1}^n$ and $W := \{(Z_i, Z_j, \zeta_{ij}) : i, j \in \mathcal{N}_n\}$.
- For a universal constant $\kappa > 0$, define the sparsity parameter $r_n := (\kappa/n)^{1/d}$, where d is the dimension of X_i.
- X_i are continuously distributed with density f.

Main Goal I

- We prove a CLT for statistics of the form

$$\Lambda(r_n^{-1}\mathcal{X}_n, W) \equiv \sum_{X \in \mathcal{X}_n} \xi(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W),$$

  where the *node statistic* ξ has range $\mathbb{R}^m$, specifically

$$n^{-1/2}\big(\Lambda(r_n^{-1}\mathcal{X}_n, W) - \mathbf{E}[\Lambda(r_n^{-1}\mathcal{X}_n, W)]\big) \xrightarrow{d} \mathcal{N}(0, \Sigma)$$

  as $n \to \infty$.
- In the dynamic model,
  - $Z_i = (Z_{i0}, \ldots, Z_{iT})$ and $\zeta_{ij} = (\zeta_{ij,0}, \ldots, \zeta_{ij,T})$.
  - Recall that the sample analogs of the ASF bounds are determined by $n^{-1}\sum_{i=1}^n \psi_i$, where $\psi_i \equiv (\sum_{j \neq i} \hat{A}_{ij}(\mathbf{s}), \sum_{j \neq i} P_{ij}(\mathbf{s}))$.

Main Goal II

- Write as $\psi(r_n^{-1}X_i, r_n^{-1}\mathcal{X}_n, W, A)$, where $A = (A_0, \ldots, A_T)$ is the full history of the network time series.
  The first argument of ψ functions as the label i, since $r_n^{-1}X_i$ is a.s. unique to i given that positions are continuously distributed.
- Observe that A is a deterministic functional of $r_n^{-1}\mathcal{X}_n$ and W, since positions only enter the model either directly through the differences $r_n^{-1}||X_i - X_j||$ or indirectly through attributes W.
- We can then define $\xi(r_n^{-1}X_i, r_n^{-1}\mathcal{X}_n, W) \equiv \psi(r_n^{-1}X_i, r_n^{-1}\mathcal{X}_n, W, A)$ and

$$\Lambda(r_n^{-1}\mathcal{X}_n, W) \equiv \sum_{X \in \mathcal{X}_n} \psi(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W, A).$$

Stabilization Conditions I

- Let X denote a generic element of $\mathcal{X}_n$.
- Let $Q(x, r)$ be the cube in $\mathbb{R}^d$ centered at x with side length r.
- Also for any $H \subseteq \mathbb{R}^d$, define $W_H = \{(Z_i, Z_j, \zeta_{ij}) : i, j \in \mathcal{N}_n, r_n^{-1}X_i, r_n^{-1}X_j \in H\}$.
- We define $\mathbf{R}_\xi^*(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W) \in \mathbb{R}_+$ is a radius of stabilization for the node statistic $\xi(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W)$ if for any $H \supseteq Q(r_n^{-1}X, \mathbf{R}_\xi^*(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W))$,

$$\xi(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W) = \xi(r_n^{-1}X, r_n^{-1}\mathcal{X}_n \cap H, W_H) \quad \text{a.s.}$$

- The radius of stabilization defines a "relevant set" of nodes $r_n^{-1}\mathcal{X}_n \cap Q(r_n^{-1}X, \mathbf{R}_\xi^*(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W))$ (or more precisely, their positions) such that the removal of nodes outside of this set does not affect the value of the statistic ξ.

Stabilization Conditions II

- Given a radius of stabilization $\mathbf{R}_\xi^*$, we say ξ is $\mathbf{R}_\xi^*$-exponentially stabilizing if for some $\tilde{n}, c, \epsilon > 0$,

$$\sup_{n > \tilde{n}} \mathbf{P}\left(\mathbf{R}_\xi^*(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W) \geqslant r\right) \leqslant c \exp\left\{-cr^\epsilon\right\}.$$

  This implies $\mathbf{R}_\xi^*(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W) = O_p(1)$.

- We say ξ is $\mathbf{R}_\xi^*$-externally stabilizing for radius of stabilization $\mathbf{R}_\xi^*(\cdot)$ if for all n, there exists $\mathbf{R}_n(X) \geqslant 0$ such that (a) $\mathbf{R}_n(X) = O_p(1)$, and (b) for n sufficiently large,

$$Q\left(r_n^{-1}X', \mathbf{R}_\xi^*(r_n^{-1}X', r_n^{-1}\mathcal{X}_n, W)\right) \subseteq Q(r_n^{-1}X, \mathbf{R}_n(X))$$

  for all $X' \in \mathcal{X}_n$ such that $r_n^{-1}X \in Q\left(r_n^{-1}X', \mathbf{R}_\xi^*(r_n^{-1}X', r_n^{-1}\mathcal{X}_n, W)\right)$ and

$$\begin{aligned} &\xi\left(r_n^{-1}X', r_n^{-1}\mathcal{X}_n \cap Q(r_n^{-1}X', \mathbf{R}_{X'}), W_{Q(r_n^{-1}X', \mathbf{R}_{X'})}\right) \\ &\neq \xi\left(r_n^{-1}X', r_n^{-1}(\mathcal{X}_n \backslash \{X\}) \cap Q(r_n^{-1}X', \mathbf{R}_{X'}), W_{Q(r_n^{-1}X', \mathbf{R}_{X'}) \backslash \{r_n^{-1}X\}}\right) \quad (7) \end{aligned}$$

Stabilization Conditions III

a.s., where $\mathbf{R}_{X'} = \mathbf{R}^*_\xi(X', r_n^{-1}\mathcal{X}_n, W)$.

- It states that removing the node positioned at X only affects an asymptotically bounded number of other nodes' statistics.
- The "affected" nodes are those positioned at X' in part (b);
- the requirement $r_n^{-1}X \in Q\left(r_n^{-1}X', \mathbf{R}^*_\xi(X', r_n^{-1}\mathcal{X}_n, W)\right)$ states that X lies in the relevant set of X',
- (7) states that the node statistic of X' is affected by the removal of X.

- Whereas exponential stabilization limits the degree to which alters affect the ego's statistic, external stabilization limits the degree to which the ego affects alters' statistics.
- A radius of stabilization $\mathbf{R}^*_\xi$ is increasing if for any n sufficiently large and $H \subseteq \mathbb{R}^d$,

$$\mathbf{R}^*_\xi(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W) \geqslant \mathbf{R}^*_\xi(r_n^{-1}X, r_n^{-1}\mathcal{X}_n \cap H, W_H) \quad \text{a.s.}$$

Stabilization Conditions IV

- This says that removing nodes can only shrink the radius of stabilization, which will be trivially satisfied in our applications.

High Level Conditions I

The first is our main weak dependence condition.

- Stabilization - There exists an increasing radius of stabilization $\mathbf{R}^*_\xi$ such that ξ is $\mathbf{R}^*_\xi$-exponentially and -externally stabilizing.
- This implies that $\xi(r_n^{-1}X_i, r_n^{-1}\mathcal{X}_n, W)$ will only depend on its arguments through a "relevant set" of nodes $J_i \subseteq \mathcal{N}_n$ whose size has exponential tails, uniformly in n.
- Relevant sets in our applications will consist of unions of the network components of nodes in the K-neighborhood of i with respect ot a certain latent network.

The remaining two assumptions are regularity conditions.

- Bounded Moments - $\sup_n \mathbf{E}[\xi\left(r_n^{-1}X, r_n^{-1}\mathcal{X}_n, W\right)^8] < \infty$.
- Polynomial Bound - There exists $c > 0$ such that for any n, $|\xi(X, r_n^{-1}\mathcal{X}_n, W)| \leqslant cn^c$ a.s.

Limit Variance I

- Let $\mathcal{P}_{\kappa f(x)}$ be a homogeneous Poisson point process on $\mathbb{R}^d$ with intensity $\kappa f(x)$.
- Let $G \in \{\{x, y\}, \{x\}, \varnothing\}$ for $x, y \in \mathbb{R}^d$.
- Let $\mathcal{X}$ denote a random, at-most countable subset of $\mathbb{R}^d$.
- Conditional on $\mathcal{X}$, we draw i.i.d. node-level attributes $\{Z(x') : x' \in \mathcal{X}\}$ and i.i.d. pair-level shocks $\{\zeta(x', y') : x', y' \in \mathcal{X}\}$ independently of the attributes.
- Let $W^\infty(\mathcal{X}) = \{(Z(x'), Z(y'), \zeta(x', y')) : x', y' \in \mathcal{X}\}$.
- The asymptotic variance will depend on node statistics of the form $\xi(x, \mathcal{P}_{\kappa f(x)} \cup G, W^\infty)$.
- Define the "add-one cost"

$$\Xi_x = \Lambda\left(\mathcal{P}_{\kappa f(x)} \cup \{x\}, W^\infty\right) - \Lambda\left(\mathcal{P}_{\kappa f(x)}, W^\infty\right). \tag{8}$$

This measures the change in the network moments due to the addition of a single node positioned at x.

Limit Variance II

- Let
$$\alpha = \int \mathbf{E}[\Xi_x] f(x)\,\mathrm{d}x,$$
- Define

$$\begin{aligned}
&\sigma^2 \\
&= \int_{\mathbb{R}^d} \mathbf{E}\left[\xi(x, \mathcal{P}_{\kappa f(x)}, W^\infty)^2\right] f(x)\,\mathrm{d}x \\
&+ \kappa \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \Big(\mathbf{E}\left[\xi(x, \mathcal{P}_{\kappa f(x)} \cup \{x, y\}, W^\infty) \xi(y, \mathcal{P}_{\kappa f(x)} \cup \{x, y\}, W^\infty)\right] \\
&- \mathbf{E}\left[\xi(x, \mathcal{P}_{\kappa f(x)} \cup \{x\}, W^\infty)\right] \mathbf{E}\left[\xi(y, \mathcal{P}_{\kappa f(x)} \cup \{y\}, W^\infty)\right] \Big) f(x)^2\,\mathrm{d}x\mathrm{d}y.
\end{aligned}$$

Main Theorem

Theorem (CLT)

$$n^{-1/2}\left(\Lambda(r_n^{-1}\mathcal{X}_n, W) - \mathbf{E}\big[\Lambda(r_n^{-1}\mathcal{X}_n, W)\big]\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2 - \alpha^2\right).$$

Moreover, if Ξ_x has a non-degenerate distribution for any $x \in supp(f)$, then $\sigma^2 > \alpha^2$.

Proof Sketch

- Proof draws on several techniques for proving limit theorems for geometric graphs (Penrose 2003, 2005, 2007; Penrose and Yukich, 2003).
- First prove CLT for "Poissonized" model, where we replace n with $N_n \sim \text{Poisson}(n)$, independent of all other primitives.
  - Poissonized model easier to work with because it possesses a spatial independence property.
  - Construct a martingale via spatial projections.
  - Under the "stabilization" condition, we verify the regularity conditions of the martingale CLT and the convergence to the limit variance.
  - In econometrics, Poissonization used to show convergence of bootstrap empirical processes (der Vaart and Wellner, 1996).

Proof Sketch

- Then "de-Poissonize" and prove result for original process.
  - Poissonized model "close to" original model since $N_n/n \xrightarrow{p} 1$.
  - However, Poissonization increases variance since $\text{Var}(n^{-1/2}N_n) = 1$.
  - Need to subtract off appropriate term to obtain correct asymptotic variance.
- Key high-level conditions for both steps are uniform moment conditions and stabilization conditions, which formalize weak dependence in this setting.

CLT for "Poissonized" Model I

Poisson Point Process: $\mathcal{X}_{N_n}$.

Step 1: Writing the moments as a martingale difference sequence using the following spatial projection.

- Let $Q(x, r)$ be the cube centered at x with side length r.
- Suppose that we partition the support of f into $Q(x_1, r_n), \ldots, Q(x_{k_n}, r_n)$.
- Observe that the number of the cubes, k_n, is proportional to n.
- Let $\mathcal{F}_l$ be the sigma field generated by elements of $\mathcal{X}_{N_n}$ that belong in the set $Q(x_1, r_n) \cup \cdots \cup Q(x_l, r_n)$, where $l = 1, \ldots, k_n$.
- Let $\mathcal{F}_0$ be the trivial sigma field.

CLT for "Poissonized" Model II

- Then by construction, we can represent the centered moments as a telescoping sum

$$\Lambda(r_n^{-1}\mathcal{X}_{N_n}, W(\mathcal{X}_{N_n})) - \mathbf{E}(\Lambda(r_n^{-1}\mathcal{X}_{N_n}, W(\mathcal{X}_{N_n}))) = \sum_{l=1}^{k_n} \delta_l,$$

  where

$$\delta_l := \mathbf{E}(\Lambda(r_n^{-1}\mathcal{X}_{N_n}, W(\mathcal{X}_{N_n})) \,|\, \mathcal{F}_l) - \mathbf{E}(\Lambda(r_n^{-1}\mathcal{X}_{N_n}, W(\mathcal{X}_{N_n})) \,|\, \mathcal{F}_{l-1}).$$

  By definition of δ_l, $\mathbf{E}(\delta_l|\mathcal{F}_{l-1}) = 0$ for all $l = 1, ..., k_n$.
- Therefore, $(\delta_l, \mathcal{F}_l)_{l=1,\dots,k_l}$ is a martingale difference sequence with respect to the filtration $\{\mathcal{F}_l\}_{l=0}^{k_n}$.

CLT for "Poissonized" Model III

Step 2: To establish the CLT of the Poissonized Model, we apply the martingale difference CLT by verifying the following conditions:

$$\sup_n \mathbf{E}\left(\frac{1}{n}\max_{1\leqslant l\leqslant k_n}\delta_l^2\right) < \infty \tag{9}$$

$$n^{-1/2}\max_{1\leqslant l\leqslant k_n}|\delta_l| = o_p(1), \tag{10}$$

$$\frac{1}{n}\sum_{l=1}^{k_n}\delta_l^2 \xrightarrow{p} \sigma^2, \tag{11}$$

- Let $\mathcal{X}'_{N_n}$ be an independent copy of $\mathcal{X}_{N_n}$ and $Q_l = Q(x_l, r_n)$.

CLT for "Poissonized" Model IV

- Define the resampling cost

$$\Delta_{x_l} = \Lambda(r_n^{-1}\mathcal{X}_{N_n}, W(\mathcal{X}_{N_n})) - \Lambda\left(r_n^{-1}\left((\mathcal{X}_{N_n}\backslash Q_l) \cap (\mathcal{X}'_{N_n} \cap Q_l)\right), W'(\mathcal{X}_{N_n})\right), \tag{12}$$

where
$W'(\mathcal{X}_{N_n}) = \{(Z(x), Z(y), \zeta(x,y)) : x, y \in (\mathcal{X}_{N_n}\backslash Q_l) \cap (\mathcal{X}'_{N_n} \cap Q_l)\}$.
- This is the change in network moments from redrawing the positions of nodes in the cube Q_l. It is quite similar to the add-one cost Ξ_x defined in (8).
- Since $\mathcal{X}'_{N_n}$ is an independent copy of $\mathcal{X}_{N_n}$, we have

$$\delta_l = \mathbf{E}(\Delta_{x_l} \mid \mathcal{F}_l) \quad l = 1, \ldots, k_n.$$

CLT for "Poissonized" Model V

- To prove (9) and (10),

$$\sup_n \mathbf{E}\left(\frac{1}{n}\max_{1\leqslant l\leqslant k_n}\delta_l^2\right) \leqslant \sup_n \frac{1}{n}\sum_{l=1}^{k_n}\mathbf{E}(\delta_l^2) \leqslant \sup_n \frac{k_n}{n}\max_{1\leqslant l\leqslant k_n}\mathbf{E}[\Delta_{x_l}^2],$$

$$\mathbf{P}\left(n^{-1/2}\max_{1\leqslant l\leqslant k_n}|\delta_l| \geqslant \varepsilon\right) \leqslant \sum_{l=1}^{k_n}\frac{1}{n^2\epsilon^4}\mathbf{E}[\delta_l^4] \leqslant \frac{k_n}{n^2\epsilon^4}\max_{1\leqslant l\leqslant k_n}\mathbf{E}[\Delta_{x_l}^4].$$

  We use the stabilization and boundedness assumptions to establish uniform bounds on $\mathbf{E}[\Delta_{x_l}^2]$ and $\mathbf{E}[\Delta_{x_l}^4]$.
- To verify (11), we first approximate δ_l by $\delta_{l,R} = \mathbf{E}[\Delta_{x_l,R} \mid \mathcal{F}_l]$, where

$$\begin{aligned}\Delta_{x_l,R} = {} & \Lambda(r_n^{-1}(\mathcal{X}_{N_n} \cap Q_{l,R}), W_{Q_{l,R}}) \\ & - \Lambda\left(r_n^{-1}\left((\mathcal{X}_{N_n} \cap Q_{l,R}\backslash Q_l) \cap (\mathcal{X}'_{N_n} \cap Q_l)\right), W'_{Q_{l,R}}\right),\end{aligned}$$

  where $Q_{l,R} := Q(x_l, Rr_n)$ and
  $W'_{Q_{l,R}} = \{(Z(x), Z(y), \zeta(x,y)) : x, y \in (\mathcal{X}'_{N_n} \cap Q_l) \cup (\mathcal{X}_{N_n} \cap Q_{l,R}\backslash Q_l)\}$.

CLT for "Poissonized" Model VI

- Note that $\Delta_{x_l,R}$ is the resampling cost under the locally restricted Poisson process, $\mathcal{X}_{N_n} \cap Q_{l,R}$. Since

$$\frac{1}{n}\sum_{l=1}^{k_n} \delta_l^2 - \sigma^2 = \frac{1}{n}\sum_{l=1}^{k_n}(\delta_l^2 - \delta_{l,R}^2) + \frac{1}{n}\sum_{l=1}^{k_n} \delta_{l,R}^2 - \sigma^2,$$

the required result (11) by showing

$$\lim_{R\to\infty} \lim_{n\to\infty} \mathbf{E}\frac{1}{n}\sum_{l=1}^{k_n} \left|\delta_l^2 - \delta_{l,R}^2\right| = 0, \tag{13}$$

$$\frac{1}{n}\sum_{l=1}^{k_n} \delta_{l,R}^2 - \sigma^2 \xrightarrow{p} 0. \tag{14}$$

- For (13) these we use the stabilization and boundedness assumptions.

CLT for “Poissonized” Model VII

- For (14) we show

$$\lim_{n\to\infty} \operatorname{Var}\left(\frac{1}{n}\sum_{l=1}^{k_n}\delta_{l,R}^2\right) = 0 \quad \text{for any } R, \tag{15}$$

$$\lim_{n\to\infty}\lim_{R\to\infty}\frac{1}{n}\sum_{l=1}^{k_n}\mathbf{E}[\delta_{l,R}^2] \to \sigma^2. \tag{16}$$

- For (15), we use the spatial independence property that disjoint subsets of a Poisson process are independent; that is, for $A, B \subseteq \mathbb{R}^d$ with $A \cap B = \emptyset$,

$$\mathbf{E}\left[\sum_{X\in\mathcal{X}_{N_n}} \mathbf{1}\{X \in A \cup B\}\right]$$
$$= \mathbf{E}\left[\sum_{X\in\mathcal{X}_{N_n}} \mathbf{1}\{X \in A\}\right]\mathbf{E}\left[\sum_{X\in\mathcal{X}_{N_n}} \mathbf{1}\{X \in B\}\right].$$

De-Poissonization I

- It can be shown that under an appropriate coupling that

$$
\begin{aligned}
& n^{-1/2}\big(\Lambda(r_n^{-1}\mathcal{X}_{N_n}, W(\mathcal{X}_{N_n})) - \mathbf{E}[\Lambda(r_n^{-1}\mathcal{X}_{N_n}, W(\mathcal{X}_{N_n}))]\big) \\
&= n^{-1/2}\big(\Lambda(r_n^{-1}\mathcal{X}_n, W) - \mathbf{E}[\Lambda(r_n^{-1}\mathcal{X}_n, W)]\big) + n^{-1/2}(N_n - n)\alpha + o_p(1).
\end{aligned} \tag{17}
$$

- The left-hand side is asymptotically $\mathcal{N}(0, \sigma^2)$, as previously discussed.
- The second term of the right-hand side is asymptotically $\mathcal{N}(0, \alpha^2)$ by the well-known normal approximation of a Poisson random variable.
- Since $N_n \perp \mathcal{X}_n$ (under the right coupling), we have the required result.

CLT of ATE of Dynamic Network Formation I

- Since $\mathbf{S}_{ij,t}$ is finitely supported, there exist $\bar{s}, \bar{z}, \bar{z}'$ such that $V(\delta, \bar{s}, \bar{z}, \bar{z}', \zeta) \geqslant V(\delta, S_{ij,t}, Z_{it}, Z_{jt}, \zeta)$ a.s. for any δ, ζ.
- Moreover, since V is strictly increasing in its last component, we can define $\tilde{V}^{-1}(\delta, \cdot)$ as the inverse of $V(\delta, \bar{s}, \bar{z}, \bar{z}', \cdot)$.
- Let $\tilde{\Phi}_\zeta$ denote the complementary CDF of $\zeta_{ij,t}$.

The key assumptions are

- Tail Condition: There exists a constant $c > 0$ such that for δ sufficiently large,
$$\tilde{\Phi}_\zeta\left(\tilde{V}^{-1}(\delta, 0)\right) \leqslant e^{-c\delta}.$$
- For the other regularity conditions, let $\Phi(\cdot \,|\, x)$ be the conditional distribution of Z_i given $X_i = x$ and $\Phi_t(\cdot \,|\, x)$ the conditional distribution of Z_{it} given $X_i = x$.

  Assume that

CLT of ATE of Dynamic Network Formation II

(a) $\Phi(z \mid x)$ is continuous in x for any z.
(b) For all t, there exists a distribution Φ_t^* that stochastically dominates $\Phi_t(\cdot \mid x)$ for all x.
(c) The density f of X_1 is continuous and bounded away from zero and infinity.
(d) V is continuous in its arguments, and $\zeta_{ij,t}$ is continuously distributed.

Conclusion

- We develop general CLT for network moments.
- Primitive conditions for stabilization in dynamic model: sparsity and weakly dependent initial network.
- Other applications: network regression.
- Work in progress: inference procedures.

Sparsity - I

$r_n = \left(\frac{\kappa}{n}\right)^{1/d}$ and the tail condition of the distribution of $\zeta_{ij,t}$ imply G_t is "sparse"" for any t.

- Due to finite support of $S_{ij,t}$ and ρ_{ij}, we can define $\bar{s}$ and $\bar{\rho}$ such that $V(\bar{s}, \bar{\rho}, \delta_{ij}, \zeta_{ij,t}) \geqslant V(S_{ij,t}, \rho_{ij}, \delta_{ij}, \zeta_{ij,t})$ a.s.
- Since V is strictly increasing in its last component, we can define $\tilde{V}^{-1}(\delta_{ij}, \cdot)$ as the inverse of $V(\bar{s}, \bar{\rho}, \delta_{ij}, \cdot)$, that is, $v = V(\bar{s}, \bar{\rho}, \delta_{ij}, \tilde{V}^{-1}(\delta_{ij}, v))$.
- Let $\bar{\Phi}_\zeta$ denote the complementary CDF of $\zeta_{ij,t}$.

Sparsity - II

Assumption (Tail Condition)

There exist constants $c_1, c_2, \epsilon > 0$ *such that*

$$\bar{\Phi}_\zeta \left(\tilde{V}^{-1}(\delta, 0)\right) \leqslant c_1 e^{-c_2 \delta^\epsilon}.$$

Then,

$$\frac{1}{n}\sum_{i,j} \mathbf{E}[G_{ij,t}] \leqslant \frac{1}{n}\sum_{i,j} \mathbf{P}(\zeta_{ij,t} > \tilde{V}^{-1}(\delta_{ij}, 0))$$
$$\to \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \bar{\Phi}_\zeta(\tilde{V}^{-1}(\|x - x'\|, 0))\, \kappa f(x)^2 \, \mathrm{d}x' \, \mathrm{d}x$$

Back