Vertex exchangeable and edge exchangeable random graphs

Svante Janson

Random Structures and Algorithms 2017
Gniezno, 11 August 2017.

Background: Exchangeability

A finite or infinite sequence (X_i) of random variables is *exchangeable* if any permutation of the variables yields a sequence with the same distribution.

Background: Exchangeability

A finite or infinite sequence (X_i) of random variables is *exchangeable* if any permutation of the variables yields a sequence with the same distribution.

Example An i.i.d. sequence. For example, toss a coin n times. $(n \leq \infty)$

Background: Exchangeability

A finite or infinite sequence (X_i) of random variables is *exchangeable* if any permutation of the variables yields a sequence with the same distribution.

Example An i.i.d. sequence. For example, toss a coin n times. $(n \leq \infty)$

Example You have a collection of false coins. Pick one at random, and then toss the same coin n times.

Background: Exchangeability

A finite or infinite sequence (X_i) of random variables is *exchangeable* if any permutation of the variables yields a sequence with the same distribution.

Example An i.i.d. sequence. For example, toss a coin n times. $(n \leq \infty)$

Example You have a collection of false coins. Pick one at random, and then toss the same coin n times.

Theorem (De Finetti)

Every infinite exchangeable sequence is of this type, i.e., a conditionally i.i.d. sequence with a random distribution.

Exchangeability and random graphs

- (Vertex) exchangeable random graphs and graphons
- Sparse exchangeable random graphs and graphons on $[0, \infty)$.
- Edge exchangeable random graphs.

Vertex exchangeable random graphs

A vertex exchangeable random graph ("exchangeable random graph") is a random graph on labelled vertices such that any (fixed) permutation of the labels yields a random graph with the same distribution.

Vertex exchangeable random graphs

A vertex exchangeable random graph ("exchangeable random graph") is a random graph on labelled vertices such that any (fixed) permutation of the labels yields a random graph with the same distribution.

This is natural if the labels are just labels without intrinsic significance.

Vertex exchangeable random graphs

A vertex exchangeable random graph ("exchangeable random graph") is a random graph on labelled vertices such that any (fixed) permutation of the labels yields a random graph with the same distribution.

This is natural if the labels are just labels without intrinsic significance.

Example $G(n, p)$

Vertex exchangeable random graphs

A vertex exchangeable random graph ("exchangeable random graph") is a random graph on labelled vertices such that any (fixed) permutation of the labels yields a random graph with the same distribution.

This is natural if the labels are just labels without intrinsic significance.

Example $G(n, p)$

Example $G(n, m)$

Example (Standard construction)

1. Fix a type space S, and a probability distribution μ in S, and a kernel (graphon) $W : S \times S \to [0, 1]$.
2. Give each vertex i a type $x_i \in S$ (i.i.d. random according to μ).
3. Add edge ij with probability $W(x_i, x_j)$ (independently, conditioned on the types).

Example (Standard construction)

1. Fix a type space S, and a probability distribution μ in S, and a kernel (graphon) $W : S \times S \to [0, 1]$.
2. Give each vertex i a type $x_i \in S$ (i.i.d. random according to μ).
3. Add edge ij with probability $W(x_i, x_j)$ (independently, conditioned on the types).

Rediscovered many times: statistical block design (finite type space); "inhomogeneous random graphs"; graph limits and graphons (W is the graphon); ...

Example (Standard construction)

1. Fix a type space S, and a probability distribution μ in S, and a kernel (graphon) $W : S \times S \to [0, 1]$.
2. Give each vertex i a type $x_i \in S$ (i.i.d. random according to μ).
3. Add edge ij with probability $W(x_i, x_j)$ (independently, conditioned on the types).

Rediscovered many times: statistical block design (finite type space); "inhomogeneous random graphs"; graph limits and graphons (W is the graphon); ...

We may w.l.o.g. assume that $S = [0, 1]$ and μ uniform distribution.

Example (Standard construction)

1. Fix a type space S, and a probability distribution μ in S, and a kernel (graphon) $W : S \times S \to [0,1]$.
2. Give each vertex i a type $x_i \in S$ (i.i.d. random according to μ).
3. Add edge ij with probability $W(x_i, x_j)$ (independently, conditioned on the types).

Rediscovered many times: statistical block design (finite type space); "inhomogeneous random graphs"; graph limits and graphons (W is the graphon); ...

We may w.l.o.g. assume that $S = [0,1]$ and μ uniform distribution.

But we don't have to, and sometimes we don't want to!

Graph limits, graphons and ...

Lovász et al (Lovász and Szegedy (2006); Borgs, Chayes, Lovász, Sós, Vesztergombi (2008, 2012)):

(i). If G_n is a sequence of graphs with $|G_n| \to \infty$ such that subgraph densities converge, then there exists a limit object, a *graph limit*.

Graph limits, graphons and . . .

Lovász et al (Lovász and Szegedy (2006); Borgs, Chayes, Lovász, Sós, Vesztergombi (2008, 2012)):

(i). If G_n is a sequence of graphs with $|G_n| \to \infty$ such that subgraph densities converge, then there exists a limit object, a *graph limit*.

(ii). A graph limit may be represented (non-uniquely) by a *graphon*. Conversely, every graphon defines a graph limit.

Graph limits, graphons and . . .

Lovász et al (Lovász and Szegedy (2006); Borgs, Chayes, Lovász, Sós, Vesztergombi (2008, 2012)):

(i). If G_n is a sequence of graphs with $|G_n| \to \infty$ such that subgraph densities converge, then there exists a limit object, a *graph limit*.

(ii). A graph limit may be represented (non-uniquely) by a *graphon*. Conversely, every graphon defines a graph limit.

(iii). Given a graphon W, the random graphs $G(n, W)$ defined above a.s. converge to W (in the sense of (i)).

. . . and exchangeable random graphs

Diaconis and Janson (2008), Austin (2008):

Take $n = \infty$. If W is a graphon, then $G(\infty, W)$ is an exchangeable infinite random graph.

Conversely, every exchangeable infinite random graph is $G(\infty, W)$ for some (possibly random) W. (A special case of the representation theorem by Aldous and Hoover for exchangeable arrays, applied to the array of edge indicators (I_{ij}).)

. . . and exchangeable random graphs

Diaconis and Janson (2008), Austin (2008):

Take $n = \infty$. If W is a graphon, then $G(\infty, W)$ is an exchangeable infinite random graph.

Conversely, every exchangeable infinite random graph is $G(\infty, W)$ for some (possibly random) W. (A special case of the representation theorem by Aldous and Hoover for exchangeable arrays, applied to the array of edge indicators (I_{ij}).)

Moreover, if G_n is a sequence of graphs with $|G_n| \to \infty$, (re)label each G_n (by 1,2,. . .) at random. This gives a sequence $\tilde{G}_n$ of exchangeable random labelled graphs; $\tilde{G}_n$ converge in distribution to an (exchangeable) infinite random graph iff G_n converges (in the sense of subgraph densities).

Sparse graphs

Caron and Fox(2104); Borgs, Chayes, Cohn and Holden(2016+); Veitch and Roy(2015+)

New construction:

1. Fix a type space (S, μ), where μ is a σ-finite measure, and a graphon $W : S \times S \to [0, 1]$.
   (Can take $([0, \infty), \lambda)$, but don't have to.)
2. Generate vertices $\{(t_i, x_i)\}_1^\infty$ by a Poisson point process on $[0, \infty) \times S$ with intensity $\lambda \times \mu$.
   (x_i is the type of the vertex; t_i is a (unique) label, and may also be thought of as the time the vertex is born.)
3. Add edge ij with probability $W(x_i, x_j)$ (independently, conditioned on the types).
4. Define $\tilde{G}_t$ as the induced subgraph using only vertices with $t_i \leq t$. Define G_t by deleting all isolated vertices.

If, for example, W is integrable, then G_t is a.s. a finite graph for every $t < \infty$. Typically sparse.

The formal definition of exchangeability is more technical here:

Represent the edge set of the graph G_t as a subset of $[0, \infty)^2$: an edge between t_i and t_j is represented by (t_i, t_j) and (t_j, t_i). Then the edge set of G_∞ is an exchangeable random point process in $[0, \infty)^2$.

The formal definition of exchangeability is more technical here:

Represent the edge set of the graph G_t as a subset of $[0, \infty)^2$: an edge between t_i and t_j is represented by (t_i, t_j) and (t_j, t_i). Then the edge set of G_∞ is an exchangeable random point process in $[0, \infty)^2$.

Conversely, Kallenberg (1990) showed (almost) that every such exchangeable random point process is obtained from a graphon by the construction above.

One can define convergence of such graphons, and convergence of graphs to such graphons, in several ways. (Not yet clear which is best.)

GP-convergence (Veitch and Roy) can be defined by:

$$W_n \to_{GP} W \iff G_t(W_n) \xrightarrow{\mathrm{d}} G_t(W), \quad \text{every } t < \infty.$$

One can define convergence of such graphons, and convergence of graphs to such graphons, in several ways. (Not yet clear which is best.)

GP-convergence (Veitch and Roy) can be defined by:

$$W_n \to_{GP} W \iff G_t(W_n) \xrightarrow{\mathrm{d}} G_t(W), \quad \text{every } t < \infty.$$

Theorem (Veitch & Roy (2016+), Janson (2017+))

For every graphon W, $G_s(W) \to_{GP} W$ a.s. as $s \to \infty$.

Edge exchangeable random graphs

Edge exchangeable random graphs were introduced by Crane and Dempsey (2016+). An equivalent model, using somewhat different formulations, was given by Broderick and Cai (2016+) and Campbell, Cai and Broderick (2016+).

The idea is that we have a fixed (labelled) vertex set, and add a sequence of edges (regarded as pairs of vertices). Repetions are allowed, so we construct a multigraph. The sequence of edges is supposed to be exchangeable.

By De Finetti's theorem, this is equivalent to the following:

Let V be a finite or infinite set, and let μ be a deterministic or random probability measure on the edges of the complete graph on V.

1. Given μ, take N i.i.d. edges with distribution μ.
2. Delete all isolated vertices.

Some similarities with vertex exchangeable random graphs with a discrete type space $\mathbb{N}$, but quite different.

For example, at most one vertex of each type.

Example

Let (q_i) be a probability distribution on $\mathbb{N}$. For each edge, just pick its two endpoints independently withthis distribution.

Thus $\mu(ij) = q_i q_j$.

Cf. similar "rank 1" cases of vertex exchangeable graphs, with $W(x, y) = \phi(x)\phi(y)$.

Example Pittel (2010) considered a random multigraph process with a fixed vertex set $[n]$ and N edges added one by one, with an edge ij added with probability proportional to $(d_i + \alpha)(d_j + \alpha)$, where d_i is the current degree of i. (Slightly modified for loops). Here $\alpha > 0$ is a fixed parameter.

Example Pittel (2010) considered a random multigraph process with a fixed vertex set $[n]$ and N edges added one by one, with an edge ij added with probability proportional to $(d_i + \alpha)(d_j + \alpha)$, where d_i is the current degree of i. (Slightly modified for loops). Here $\alpha > 0$ is a fixed parameter.

Equivalently: choose vertices with probability proportional to $d_i + \alpha$. Then join the first two vertices to an edge, then the next two, and so on.

Example Pittel (2010) considered a random multigraph process with a fixed vertex set $[n]$ and N edges added one by one, with an edge ij added with probability proportional to $(d_i + \alpha)(d_j + \alpha)$, where d_i is the current degree of i. (Slightly modified for loops). Here $\alpha > 0$ is a fixed parameter.

Equivalently: choose vertices with probability proportional to $d_i + \alpha$. Then join the first two vertices to an edge, then the next two, and so on.

Thus, the vertices are chosen according to a Pólya urn process, starting with α balls of each colour ($=$ vertex). The sequence of vertices is exchangeable, and thus so is the sequence of edges. Hence, this is an edge exchangeable random multigraph.

Remarks.

1. Exchangeability implies that conditioned on the final degree of each vertex, all possible edge sequences have the same probability. Hence, conditioned on the degree sequence, the random multigraph is the multigraph given by the configuration model.

Remarks.

1. Exchangeability implies that conditioned on the final degree of each vertex, all possible edge sequences have the same probability. Hence, conditioned on the degree sequence, the random multigraph is the multigraph given by the configuration model.
2. A standard result for Pólya urn processes shows that the vector $(d_i/2N)$ converges to a Dirichlet$(\alpha, \ldots, \alpha)$ distribution as $N \to \infty$.

Remarks.

1. Exchangeability implies that conditioned on the final degree of each vertex, all possible edge sequences have the same probability. Hence, conditioned on the degree sequence, the random multigraph is the multigraph given by the configuration model.
2. A standard result for Pólya urn processes shows that the vector $(d_i/2N)$ converges to a Dirichlet$(\alpha, \ldots, \alpha)$ distribution as $N \to \infty$.
3. The random sequence of vertices in the construction can be seen as a two-parameter Chinese restaurant process with parameters $(-\alpha, n\alpha)$. A Chinese restaurant process with other parameters yields a similar edge exchangeable random multigraph (on a number of vertices growing to ∞).

Chinese restaurant process with parameters (θ, α):

When there are N customers seated at k tables, with $N_i \geq 1$ customers at table i, a new customer is placed at:

$$\begin{cases} \text{table } i \ (\leq k) \text{ with probability} & (N_1 - \alpha)/(n + \theta) \\ \text{table } k+1 \text{ (new) with probability} & (\theta + k\alpha)/(n + \theta) \end{cases}$$

For $\alpha \in [0, 1]$ and $\theta > -\alpha$, the number of tables grows to ∞ a.s. The vector of proportions sitting at each table converges a.s. to a GEM distribution (given ordering of tables) and to a Poisson–Dirichlet distribution (decreasing order of frequencies).

Simple graph version

We can merge multiple edges and ignore loops, and thus obtain a random simple graph. This gives an increasing sequence of simple graphs.

Let G_m be the resulting simple graph with m edges.

Simple graph version

We can merge multiple edges and ignore loops, and thus obtain a random simple graph. This gives an increasing sequence of simple graphs.

Let G_m be the resulting simple graph with m edges.

Example If $P(ij) \sim ((i \vee j)!)^{-4}$, then a.s. $G_m = K_n$ when $m = \binom{n}{2}$ for all large n. Thus $G_n \to$ the graphon 1 a.s.

Simple graph version

We can merge multiple edges and ignore loops, and thus obtain a random simple graph. This gives an increasing sequence of simple graphs.

Let G_m be the resulting simple graph with m edges.

Example If $P(ij) \sim ((i \vee j)!)^{-4}$, then a.s. $G_m = K_n$ when $m = \binom{n}{2}$ for all large n. Thus $G_n \to$ the graphon 1 a.s.

Example There exists a distribution μ of edges on $V = \mathbb{N}$ such that a.s. the sequence G_n is dense in the space of graph limits, i.e., for every graph limit (graphon), there exists a subsequence G_{m_i} converging to it.

Simple graph version

We can merge multiple edges and ignore loops, and thus obtain a random simple graph. This gives an increasing sequence of simple graphs.

Let G_m be the resulting simple graph with m edges.

Example If $P(ij) \sim ((i \vee j)!)^{-4}$, then a.s. $G_m = K_n$ when $m = \binom{n}{2}$ for all large n. Thus $G_n \to$ the graphon 1 a.s.

Example There exists a distribution μ of edges on $V = \mathbb{N}$ such that a.s. the sequence G_n is dense in the space of graph limits, i.e., for every graph limit (graphon), there exists a subsequence G_{m_i} converging to it.

An example of everything?

Simple graph version

We can merge multiple edges and ignore loops, and thus obtain a random simple graph. This gives an increasing sequence of simple graphs.

Let G_m be the resulting simple graph with m edges.

Example If $P(ij) \sim ((i \vee j)!)^{-4}$, then a.s. $G_m = K_n$ when $m = \binom{n}{2}$ for all large n. Thus $G_n \to$ the graphon 1 a.s.

Example There exists a distribution μ of edges on $V = \mathbb{N}$ such that a.s. the sequence G_n is dense in the space of graph limits, i.e., for every graph limit (graphon), there exists a subsequence G_{m_i} converging to it.

An example of everything?

Or of nothing?