08/09/20 : L2 : Erdös-Rényi random graphs.

ER : $G(n,p)$, $p \in [0,1]$ if $p > 1$, take $p \wedge 1$.

$$V = [n] = \{1, \dots, n\}$$

$$E = \{(i,j) : X_{ij} = 1\}$$ – unordered pairs.

where X_{ij}, $1 \le i < j \le n$ are i.i.d. Ber(p) rand. variables.

Also we take $X_{ij} = X_{ji}$, $1 \le i, j \le n$. $X_{ii} = 0$.

Undirected & simple graph.

The graph is formed by keeping edges in a complete graph K_n with prob p & each edge is kept independently of other edges.

Let G be a fixed graph on n vertices with e edges

Then

$$\mathbb{P}(G(n,p) = G) = p^e (1-p)^{\binom{n}{2} - e}.$$

If $p = \frac{1}{2}$, $\mathbb{P}(G(n,p) = G) = \left(\frac{1}{2}\right)^{\binom{n}{2}}$

i.e., $G(n, 1/2)$ is uniformly selecting a graph on n vertices. Let $\mathcal{G}_n$ – graphs on n vertices.

$\Rightarrow$ $G(n, 1/2) \stackrel{d}{=} \text{Unif}(\mathcal{G}_n)$ (one original motivation)

$\rightarrow$ $\sigma : [n] \to [n]$, permutation.

$$\sigma(G) = \left(\sigma(V), \{\{\sigma(i), \sigma(j)\} : \{i,j\} \in E\}\right)$$

Ex. (A) Check that $\sigma(G(n,p)) \overset{d}{=} G(n,p)$ i.e.,

$$\mathbb{P}(G(n,p) = G) = \mathbb{P}(\sigma(G(n,p)) = G)$$
$$\forall\, G \in \mathcal{G}_n.$$

$$\rightarrow \deg_{G(n,p)}(i) = \sum_{j=1}^{n} X_{ij} = \sum_{j \neq i} X_{ij}$$

$$\overset{d}{=} \mathrm{Bin}(n-1, p).$$

If $p = \frac{\lambda}{n}$, $\lambda > 0$ then $\deg(i) \xrightarrow{d} \mathrm{Poi}(\lambda)$ i.e.,

$[\lambda > n,\ p = (\frac{\lambda}{n} \wedge 1)]$ $\qquad \mathbb{P}(\deg(i) = k) \to e^{-\lambda} \frac{\lambda^k}{k!} \quad \forall\, k \geq 0.$

$$\boxed{\rightarrow \text{SPARSE REGIME:}}$$

$E_n = E(G(n,p))$ – Edge set of $G(n,p)$.

$$|E_n| = \frac{1}{2} \sum_{i,j=1}^{n} X_{i,j} = \sum_{1 \leq i < j \leq n} X_{ij}$$

$[p := p(n) = p_n]$

Ex(A) $\mathbb{E}[|E_n|] = \binom{n}{2} p$. (by lin. of Exp.)

1. If $n^2 p(n) \to 0 \Rightarrow \mathbb{P}(|E_n| \geq 1) \to 0$ as $n \to \infty$.
2. If $n p(n) \to 0 \Rightarrow \deg(i) \xrightarrow{d} 0 \quad \forall\, i$.
3. If $n p(n) \to \lambda \Rightarrow \deg(i) \xrightarrow{d} \mathrm{Poi}(\lambda), \ \forall\, i$
4. If $n p(n) \to \infty \Rightarrow \mathbb{P}(\deg(i) \geq m) \to 1 \quad \forall\, m \geq 1$.

$\deg(i) \xrightarrow{d} \infty$.

Ex(A) – Assignment question.

So we can get locally-finite graphs only if
$np \to \lambda \in [0, \infty)$.

Further if $\lambda = 0$, the graph becomes trivial.

So $\lambda \in (0, \infty)$ is the interesting case

This is called the sparse regime $\longrightarrow$ & $\sup_n \mathbb{E}\, \deg_n(i) < \infty$ $\forall i$)

we'll assume $p = \frac{\lambda}{n}$ for simplicity
[instead of $np \to \lambda$]

Degrees are "finite" in the sparse regime but what about the graph overall?

$G = ([n], E)$, $[n] = \{1, \cdots, n\}$.

$H = ([k], E(H))$ - Another graph, connected. $k \leq$

$$X(H; G) = \sum_{\substack{F \subseteq G \\ |V(F)| = k}} \mathbb{1}(F \cong H), \quad \boxed{\text{\# of copies of } H \text{ in } G}$$

(no. of copies of H in G)

- Sum is over all subgraphs of G on k vertices.

More useful representation

$$X(H; G) = \frac{1}{C_H} \sum_{\substack{\tau : [k] \to [n] \\ \text{injective}}} \mathbb{1}[\tau(H) \subseteq G]$$

$$\underset{\text{(Ex.)}}{=} \frac{1}{C_H} \sum_{\tau : \cdots} \prod_{(i,j) \in E(H)} \mathbb{1}[\tau(i) \overset{G}{\sim} \tau(j)]$$

C_H - # of automorphisms of H, i.e., isom : $H \to H$

Let $G_n = G(n,p)$

$$\mathbb{1}[\tau(i) \overset{G_n}{\sim} \tau(j)] = X_{\tau(i),\tau(j)}$$

So $$X(H;G_n) = \frac{1}{C_H} \sum^{\neq}_{i_1,\dots,i_k} \prod_{(j,l)\in E(H)} X_{i_j,i_l}$$

($\sum^{\neq}$ – distinct summands)

By linearity of Expectations,

$$\mathbb{E}[X(H;G_n)] = \frac{1}{C_H} \sum^{\neq}_{i_1,\dots,i_k} \mathbb{E}\Big[\prod \cdots X_{i_j,i_l}\Big]$$

$$= \frac{k!}{C_H}\binom{n}{k} p^{e_H} \qquad e_H = |E(H)|$$

[$X_{i,j}$'s are indep. & distinct]

PROPOSITION: Let $G_n = G(n,p)$, $p = \frac{\lambda}{n}$, $\lambda \in (0,\infty)$

H as above. $exc(H) = e_H - k \geq -1$

conn. on k vertices

Then

$$\mathbb{E}[X(H;G_n)] \underset{n\to\infty}{\sim} C_H^{-1} \lambda^{e_H} n^{-exc(H)}.$$

$$[a_n \sim b_n \iff \frac{a_n}{b_n} \to 1].$$

Proof:

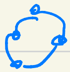
$$\mathbb{E}[X(H;G_n)] = \frac{1}{C_H} \frac{n!}{(n-k)!} \left(\frac{\lambda}{n}\right)^{e_H}$$

$$\sim \frac{1}{C_H} n^k \left(\frac{\lambda}{n}\right)^{e_H}$$

Remarks:

(1) $\mathrm{exc}(H) = -1$ iff H is a tree [Tree is a Connected acyclic graph]

$$\frac{1}{n}\mathbb{E}[X(H;G_n)] \Rightarrow \frac{\lambda^{k-1}}{c_H}$$

(2) $H = C_k$ – k-cycle  k vertices $(\mathbb{Z}/k\mathbb{Z})$

$\mathrm{exc}(C_k) = 0$

$$\mathbb{E}[X(C_k;G_n)] \longrightarrow \frac{\lambda^k}{2k}$$

(3) H has at least two cycles (i.e., $\mathrm{exc}(H) \geq 1$)

$$\Rightarrow \mathbb{E}[X(H;G_n)] \longrightarrow 0$$

Markov's ineq $\Rightarrow$ $\mathbb{P}(X(H;G_n) \geq 1) \to 0$.

$\Rightarrow$ In the sparse regime, the only subgraphs in G_n are trees & unicycles

Also no. of trees >> no. of unicycles.

What does it mean for random rooted graphs?

$G_n = G(n,p)$. Choose $O_n \in [n]$ uniformly at random & independently of G_n.

$(G_n, O_n) \in \mathcal{G}_*$ – loc. fin, rooted graphs

$$X(H;(G,o)) = \sum_{\substack{F \subseteq G \\ o \in F,\ |V(F)| = k}} \mathbb{1}[F \cong H]$$

$$= \frac{1}{C_H} \sum_{i_1,\dots,i_k}^{\neq} \mathbb{1}[o \in \{i_1,\dots,i_k\}] \prod_{(j,l) \in H} \mathbb{1}[i_j \overset{G}{\sim} i_l]$$

$$\mathbb{E} X(H; (G_n, o_n)) = \frac{1}{C_H} \sum_{i_1,\dots,i_k} \mathbb{E}\Big[\mathbb{1}[o_n \in \{i_1,\dots,i_k\}] \prod_{\dots} X_{i_j i_l}\Big]$$

[o_n indep. of $X_{i,j}$'s
& unif $[n]$]

$\Downarrow$

$$= \frac{k}{C_H\, n} \frac{n!}{(n-k)!} p^{e_H}$$

$$\overset{n\to\infty}{\sim} \frac{k}{n} \frac{1}{C_H} \lambda^{e_H} n^{-\mathrm{exc}(H)}$$

$$\Rightarrow \mathbb{E} X(H; (G_n, o_n)) \longrightarrow \frac{k \lambda^{k-1}}{C_H} \quad \text{if } H \text{ is a tree}$$

$$\longrightarrow 0 \quad \forall \text{ all conn. } H \neq \text{tree.}$$

So, Markov's inequality

$\Rightarrow \mathbb{P}(o_n \in H \text{ in } (G_n, o_n)) \to 0$ if H has a cycle.

$\Rightarrow$ "From o_n, (G_n, o_n) is a tree"!

$\Rightarrow$ From a "typical" vertex in G_n, G_n "looks locally like a tree".

Next few weeks towards showing that

$$(G_n, o_n) \xrightarrow{LW} GW(\lambda)$$

$\hookrightarrow$ Galton-Watson tree with Poisson(λ) distribn.

Cycle counts:

Defn: Total variation distance between 2 prob. measures P & Q on $(S, \mathcal{S})$ is defined as

$$d_{TV}(P, Q) := \sup_{A \in \mathcal{S}} |P(A) - Q(A)|$$

observation: $P(A^c) - Q(A^c) = Q(A) - P(A)$

$$\Rightarrow \quad d_{TV}(P, Q) = \sup_{A \in \mathcal{S}} P(A) - Q(A)$$

S is countable; Supremum is achieved for

$$A = \{ x \in S : P(x) \geq Q(x) \}$$

$$P(A) - Q(A) = \sum_{x \in A} |P(x) - Q(x)|$$

$$\parallel$$

$$Q(A^c) - P(A^c) = \sum_{x \in A^c} |P(x) - Q(x)|$$

$$\Rightarrow d_{TV}(P,Q) = \frac{1}{2} \sum_{x \in S} |P(x) - Q(x)|$$

Notn: If $X \stackrel{d}{=} P$, $Y \stackrel{d}{=} Q$ $d_{TV}(X,Y) = d_{TV}(P,Q)$

T.V. distance dep. on prob. distribn & not on r.v.'s.

If $d_{TV}(P_n, P) \to 0$ then $P_n(x) \to P(x)$ $\forall x \in S$.

[S is countable]

If $S = \mathbb{Z}$; $X_n \stackrel{d}{=} P_n$ & $X \stackrel{d}{=} P$ the above $\Rightarrow$

$$X_n \xrightarrow{d} X.$$

Def: Suppose $\{I_a\}_{a \in \Gamma}$ is a coll^n of r.v. We say $L = (\Gamma, E(L))$ is a DEPENDENCY GRAPH for $\{I_a\}$ if whenever $A, B \subseteq \Gamma$ & $\nexists$ no edges between A & B ($A \cap B = \phi$) then $\{I_a\}_{a \in A}$ & $\{I_b\}_{b \in B}$ are independent.

Eg: $\{I_a\}_{a \in \Gamma}$ are indep.; $E(L) = \phi$ works.

Eg: $\{X_i\}_{i=1}^{\infty}$ indep. r.v.'s.

$$I_i = X_i X_{i+1} \cdots X_{i+k}, \quad i \geq 1$$

$$L = \left(\mathbb{N}, \{(i,j) : |i-j| < 3k\}\right)$$

(Ex.) Then L is a dep. graph for $(I_i)_{i=1}^{\infty}$

$\rightarrow$ Complete graph is always a dep. graph for finite Γ.

THM [Holst, Barbour & Janson] (see Frieze-Karonski THM 20.12)

$X = \sum_{a \in \Gamma} I_a$, I_a are Ber(p_a) rand. variables

& has a dep. graph $L = (\Gamma, E(L))$. $\lambda = \sum_{a \in \Gamma} p_a$.

$Z_\lambda \overset{d}{=} \mathrm{Poi}(\lambda)$

Then $d_{TV}(X, Z_\lambda) \leq \min\{\lambda^{-1}, 1\} \left[\sum_{a \in \Gamma} \sum_{b \in N_a \cup \{a\}} p_a p_b + \sum_{a \in \Gamma} \sum_{b \in N_a} \mathbb{E}[I_a I_b] \right]$

$N_a = \{b : b \overset{L}{\sim} a\}$.

Eg: $H = C_3$. $\triangle$ $X(H; G_n) = \frac{1}{6} \sum_{i,j,k}^{\neq} X_{ij} X_{jk} X_{ik}$

$$= \sum_{\{i,j,k\}}^{\neq} Y_{ijk}$$

$Y_{ijk} = X_{ij} X_{jk} X_{ik}$; $\Gamma = \{\{i,j,k\} : i \neq j \neq k\}$

Y_{ijk} are Ber(p^3) r.v. & not indep.

But if $|I \cap J| \leq 1$ then Y_I & Y_J are independent,
$I \in \Gamma$, $J \in \Gamma$

because they don't share an edge.

$|\Gamma| = \binom{n}{3}$

$\Gamma = \{ I \subseteq [n] : |I| = 3 \}$ — All triangles in K_n.

$E(L) = \{ (I, J) : |I \cap J| \geq 2 \}$ — Pair of triangles that share an edge.

(Ex.) S.T. L is dep. graph for $\{Y_I\}_{I \in \Gamma}$.

$d_{TV}(X(C_3; G_n), \mathrm{Poi}(\lambda_n)) \leq T1 + T2$ (BHJ Thm)

$\lambda_n = \binom{n}{3} p^3 = \frac{n(n-1)(n-2)p^3}{6}$

$$T2 = \sum_{I \in \Gamma} \sum_{\substack{J \sim I \\ J \neq I}} E[Y_I Y_J]$$

$$= \sum_{I \in \Gamma} \sum_{|J \cap I| = 2} E[Y_I Y_J] \Leftarrow \left(\begin{matrix} J \sim I, & |J \cap I| \geq 2 \\ J \neq I, & |J \cap I| < 3 \end{matrix} \right)$$

Suppose $I = \{i, j, k\}$ & $J = \{i, j, l\}$

the $Y_I Y_J = X_{ij} X_{ik} X_{jk} X_{il} X_{jl}$

Since all are independent

$\Rightarrow E Y_I Y_J = p^5$; $\#|\{J : |J \cap I| = 2\}| = 3(n-3)$.

$\Rightarrow T2 = 3(n-3) \binom{n}{3} p^5$

$$T1 = \sum_{I \in \Gamma} \left(\sum_{\substack{J \sim I \\ J \neq I}} E[Y_I] E[Y_J] + E[Y_I]^2 \right)$$

$$= \binom{n}{3} p^6 \left[3(n-3) + 1 \right]$$

$$\Rightarrow d_{TV}\left(X(C_3, G_n), \mathrm{Poi}\left(\binom{n}{3} p^3\right) \right) \le 3(n-3)\binom{n}{3}(p^5 + p^6) + \binom{n}{3} p^6 \quad \text{--- (1)}$$

THM: Suppose $p = \frac{\lambda}{n}$, $\lambda \in (0, \infty)$.

$$d_{TV}\left(X(C_3, G_n), \mathrm{Poi}\left(\frac{\lambda^3}{6}\right) \right) \longrightarrow 0 \text{ as } n \to \infty.$$

Proof: $d_{TV}(\ \downarrow \qquad \downarrow\ )$ By (1)

(Triangle inequality)

$$\le d_{TV}\left(X(C_3; G_n), \mathrm{Poi}\left(\binom{n}{3} p^3\right) \right) \longrightarrow 0$$

$$+ \, d_{TV}\left(\mathrm{Poi}\left(\binom{n}{3} p^3\right), \mathrm{Poi}\left(\frac{\lambda^3}{6}\right) \right) \longrightarrow 0$$

$$d_{TV}(\mathrm{Poi}(a), \mathrm{Poi}(b)) \le |a - b|.$$ (Proof next).

DEFN: Given prob. measures P, Q on $(S_1, \mathcal{S}_1)$ & $(S_2, \mathcal{S}_2)$ respectively, a coupling of

P & Q is a prob. measure Π on $(S_1 \times S_2, \mathcal{S}_1 \times \mathcal{S}_2)$
$\ni \quad \Pi \circ \pi_1^{-1} = P \qquad \Pi \circ \pi_2^{-1} = Q$
$(\pi_i : S \longrightarrow S_i$ projn.$)$

[Eg. $\Pi = P \times Q$]

Probabilistically, coupling of random elements $X \stackrel{d}{=} P$ & $Y \stackrel{d}{=} Q$ is a random vector $(\hat{X}, \hat{Y}) \in S_1 \times S_2 \ni$
$\hat{X} \stackrel{d}{=} X$ & $\hat{Y} \stackrel{d}{=} Y$.

Originally $\quad X : (\Omega_1, \mathcal{F}_1, \mathbb{P}_1) \longrightarrow (S_1, \mathcal{S}_1)$
$\quad Y : (\Omega_2, \cdot, \mathbb{P}_1) \longrightarrow (\quad)$

But $(\hat{X}, \hat{Y}) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\underset{S_1 \times S_2,}{S}, \mathcal{S})$

We say $(\hat{X}, \hat{Y})$ is a coupling of X & Y or P & Q.

$S_1 = S_2$.

PROPN: $d_{TV}(P, Q) \leq \mathbb{P}(\hat{X} \neq \hat{Y})$ for any coupling $(\hat{X}, \hat{Y})$

Proof: $\quad P(A) - Q(A) = \mathbb{P}(\hat{X} \in A) - \mathbb{P}(\hat{Y} \in A)$

$= \mathbb{E}[\mathbb{1}[\hat{X} \in A] - \mathbb{1}[\hat{Y} \in A]]$ [by linearity & coupling]

$$= \mathbb{E}\left[\left(\mathbb{1}[\hat{X} \in A] - \mathbb{1}[\hat{Y} \in A]\right) \mathbb{1}[\hat{X} \neq \hat{Y}]\right]$$

$$\leq \mathbb{P}(\hat{X} \neq \hat{Y}) .$$

$a<b$.

Let Z_a be $\mathrm{Poi}(a)$ r.v., Z_b be $\mathrm{Poi}(b)$ r.v. $a<b$.

& let Y be $\mathrm{Poi}(b-a)$ r.v. & independent of Z_a.

$\Rightarrow$ $(Z_a, Y) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \cdot)$ exists

Then $\hat{Z}_b = Z_a + Y \overset{d}{=} Z_b$; $\hat{Z}_a = Z_a$.

Prop $\Rightarrow$ $d_{TV}(\mathrm{Poi}(a), \mathrm{Poi}(b)) \leq \mathbb{P}(\hat{Z}_b \neq \hat{Z}_a)$

[$(\hat{Z}_b, \hat{Z}_a)$ is a coupling of Z_a & Z_b]

$$= \mathbb{P}(Y \geq 1) = 1 - e^{-(b-a)}$$

$$\leq (b-a)$$

$\Rightarrow$ $d_{TV}(\mathrm{Poi}(a), \mathrm{Poi}(b)) \leq |b-a| \quad \forall a, b$.

$\to$ Completes proof of Poisson approximation THM for $X(C_3; G_n)$.

Remarks:

(1) $\forall (P,Q)$ $\exists$ a coupling $(\hat{X},\hat{Y})$ $\ni$

$$d_{TV}(P,Q) = \mathbb{P}(\hat{X} \neq \hat{Y})$$

[vdH-1 (Ch. 2) ; Bordenave (Ch. 2)].

(2) Poisson Convergence: $X_n \xrightarrow{d} \mathrm{Poi}(\lambda)$

(3) Poisson Approximation: $d_*(X_n, \mathrm{Poi}(\lambda)) \leq \ldots$
$\longrightarrow 0$.

d_* is a metric on the space of prob. measures.

(4) Suppose $\{X_n\}_{n\geq 1}$ are $\mathbb{N}\cup\{0\}$- r.v. & $X \overset{d}{=} \mathrm{Poi}(\lambda)$.

If $\mathbb{E}[X_n(X_n-1)\cdots(X_n-k+1)] \longrightarrow \lambda^k \quad \forall k\geq 1$

(Factorial moments) — λ^k: k^{th} Factorial mom. of $\mathrm{Poi}(\lambda)$

then $X_n \xrightarrow{d} \mathrm{Poi}(\lambda)$.

[vdH-1, Ch. 2]

Earlier proofs of $X(C_k; G_n) \xrightarrow{d} \mathrm{Poi}\left(\frac{\lambda^k}{2k}\right)$

follows by factorial moment / moment method.

Eg: $X = \sum_{a \in \Gamma} Y_a$.

$$X(X-1)\cdots(X-k+1) = \sum^{\neq}_{a_1,\dots,a_k \in \Gamma} Y_{a_1}\cdots Y_{a_k}.$$

[Poisson Convergence: Alon & Spencer - Probabilistic Method].

(5) We have rate of convergence of
$d_{TV}(X(C_3; G_n), Z_{\frac{\lambda^3}{6}})$

H Conn. subgraph

$exc(H) \geq 1 \Rightarrow X(H; G_n) \to 0$

$H = C_k \Rightarrow X(C_k; G_n) \xrightarrow{d} Poi\left(\frac{\lambda^k}{2k}\right)$ [Ex (A)]

$exc(H) = -1$ i.e, H is a tree

$n^{-1} X(H; G_n) \xrightarrow{d} ?$

$n^{-1} \mathbb{E} X(H; G_n) \to$?