Epidemics & Random Graphs

SIR (susceptible, infective, removed) epidemic process:
- A population initially consisting of healthy individuals and a small number of infected individuals.
- Infected individuals have an infectious period during which they can infect the healthy individual encountered with. After the infectious period, the infected individual ceases spreading (e.g., due to death or immunization) becomes removed individual.

Reed–Frost epidemic
- A discrete time version of SIR model.
- infectious period lasts one unit of time. (or iid exponential list in a continuous version)
- A set of n individuals.
- At time 0, a single individual is infected.
- Each infected node can infect a healthy node with probability \( p \). (independently over all healthy nodes and infectious nodes)
- Each infected individual is removed after one time slot.

The evolution can be described by a Discrete-time Markov chain \( \{ Z(t) \}_{t=0,1,2,...} \) where \( Z(t) = (\text{Set}(t), \text{Inf}(t)) \)

set of susceptible nodes

set of infected nodes.
Another representation of the model can be provided based on Erdős–Rényi random graph (also called E-R graph, \( G(n, p) \)).

**Def:** E-R random graph \( G(n, p) \): A graph of \( n \) nodes \( \{1, \ldots, n\} \) where each pair \((u, v)\) of nodes is connected (i.e., there is an edge \((u, v)\)) with probability \( p \), independently of other edges. Let \( s_{uv} = 1 \) if edge \((u, v)\) exists and \( s_{uv} = 0 \) otherwise.

Then the Reed–Frost Epidemic can be described based on \( G(n, p) \) as follows:

- Suppose node \( u \) is infected at time \( t \).

- Then \( u \) infects all its susceptible neighbors (i.e., all \( v \) s.t. \( s_{uv} = 1 \) and \( v \) susceptible).

- \( u \) is "removed" at time \( t+1 \).

Let \( d_G(u, v) \) be the length of the shortest path from \( u \) to \( v \).

Define the \( i \)-hop neighborhood of \( u \) as \( T_i(u) = \{ v \text{ s.t. } d_G(u, v) = i \} \).

Then the evolution of the Reed–Frost epidemic started at \( u \), at time \( t \), is given by:

\[
Z_v(t) = \begin{cases} 
R & \text{if } d_G(u, v) < t \\
I & \text{if } d_G(u, v) = t \\
S & \text{if } d_G(u, v) > t 
\end{cases}
\]
R: Removed  I: Infected  S: Susceptible.

Let $C(u)$ be the connected component of $G(n,p)$ containing $u$, i.e., a connected subgraph of $G(n,p)$ containing $u$. Hence, if epidemic starts at $u$, the set of nodes in $C(u)$ will be eventually removed. Thus, we need to study the size of the connected components in ER graph.

E. Emergence of giant component in ER graphs.

We study the size of the largest component of the graph $G(n,p)$ as $n \to \infty$ and $pn = \lambda > 0$ is fixed. Let $C_1$ be the largest component, $C_2$ be the second largest and so on. $|C_i|$ is the size (number of nodes) of $C_i$.

Theorem: (i) $\lambda < 1$ (subcritical regime): There is a constant $a = a(\lambda)$ s.t. \[ \lim_{n \to \infty} \Pr(|C_1| < a \log n) = 0. \]

(ii) $\lambda > 1$ (supercritical regime): Let $p_{ext}(\lambda)$ be the extinction probability of Galton-Watson branching process with Poisson law $\lambda$ offspring (i.e., the root of $x = \exp(-\lambda(1-x))$ in $[0,1]$).
Then there is a $a' = a'(\lambda)$ s.t. for all $\delta > 0$:

$$\lim_{n \to \infty} \mathbb{P}\left( \left| \frac{\langle c \rangle}{n} - (1 - P_{\text{ext}}(\lambda)) \right| \leq \delta, \quad |c| \leq a' \log n \right) = 1$$

(iii) $\lambda = 1$ (critical regime): There is a constant $k > 0$ s.t.

for all $a > 0$:

$$\mathbb{P}(\langle c \rangle \geq a n^{2/3}) \leq \frac{k}{a^2}.$$

In terms of outcome of Red-First epidemic: In subcritical case, a "small outbreak" happens, while in supercritical case, a large outbreak happens with size $(1 - P_{\text{ext}}(\lambda))$ fraction of nodes.

Proof idea is to approximate the connected component by a Galton-Watson Process with offspring dist. $\text{Binomial}(n - 1, p)$, thus the mean is $\mu = (n - 1)p$.

Proof of (ii) subcritical regime

Consider a node $v \in \{1, \ldots, n\}$ and the connected component containing $v$ denoted by $C(v)$. We construct $C(v)$ using a one-by-one exploration as follows. At each step $k$;
At each step $k$, an arbitrary node from $A_{k-1}$ is chosen and all its neighbors in $\{1, \ldots, n\} \setminus \{A_{k-1} \cup B_{k-1}\}$ are turned to active nodes. We use $D_k$ to denote this set of newly activated nodes. Then,

$$A_k = A_{k-1} \cup D_k \setminus \{v_{k-1}\}$$

$$B_k = B_{k-1} \cup \{v_{k-1}\}$$

Let $s_k = |D_k|$, then in terms of set sizes:

$$|A_0| = 1$$

$$|A_k| = |A_{k-1}| - 1 + s_k, \quad k > 0$$

Conditional on $s_1, s_2, \ldots, s_{k-1}$ (history),

$$s_k \sim \text{Binomial}(n - k - 1 - |A_{k-1}|, p)$$

The process stops when $A_k = \emptyset$, i.e., all nodes in $\mathcal{C}(v)$ have been found and deactivate:

$$|\mathcal{C}(v)| = \inf \{k > 0 : A_k = \emptyset\}$$

\textbf{Lemma:} $|A_k + k - 1 \sim \text{Binomial}(n - 1, 1 - (1 - p)^k)$

\textbf{proof:} Let $X_k$ be the number of nodes that have not been explored after $k$ steps.
It suffices to show that $X_k \sim \text{Binomial} \left( n-1, (1-p)^k \right)$ because

$$X_k = n - k - |A_k|$$

Suppose a node $w \not= v$ has not been explored after $k$ steps, this means that $w \not\in D_1, w \not\in D_2, \ldots, w \not\in D_k$, i.e., there are no edges between $(w, v), (w, v_1), \ldots, (w, v_{k-1})$, which happens with probability $(1-p)^k$, i.i.d. for each $w \not= v$.

$$X_k = \sum_{w \not= v} \mathbb{1}(w \not\in D_1, \ldots, w \not\in D_k) \sim \text{Binomial} \left( n-1, (1-p)^k \right).$$

Consider $\theta > 0$. Then

$$\mathbb{P}(|C(w)| \geq k) = \mathbb{P}(|A_1| > 0, \ldots, |A_k| < k)$$

$$\leq \mathbb{P}(|A_k| > 0)$$

$$= \mathbb{P}(|A_k| + k - 1 \geq k)$$

$$= \mathbb{P}(\text{Binomial} \left( n-1, 1-(1-p)^k \right) \geq k)$$

$$\leq \mathbb{P}(\text{Binomial} \left( n, kp \right) \geq k)$$

$$\leq \frac{\binom{n}{k} (1-kp)^{n-k} (kp e^{\theta})^{k}}{e^{\theta k}}$$

Chernoff bound

$$\leq \frac{(1-kp)^{n-k} (kp e^{\theta})^{k}}{e^{\theta k}} = \exp \left( -k(\theta + \lambda (1-e^\theta)) \right)$$

$$1 - 2 \leq e^{-x}$$
For $\lambda = np < 1$, $\max_{\theta > 0} \{ \theta + \lambda (1-e^\theta) \} = -\log \lambda - 1 + \lambda = \beta > 0$

For any $\delta > 0$:

$$P \left( \| c \| > \beta \| 1 + \delta \| \log n \right) \leq P \left( \max_{i \leq n} \| c_i \| > \beta \| 1 + \delta \| \log n \right)$$

$$\leq \sum_{i=1}^n P \left( \| c_i \| > \beta \| 1 + \delta \| \log n \right)$$

$$\leq n P \left( \| c \| > \beta \| 1 + \delta \| \log n \right)$$

$$\leq n \cdot n^{-1 - \delta} = n^{-\delta} \to 0 \text{ as } n \to \infty$$

So all connected components are of size at most $O(\log n)$.

(vii) proof of supercritical regime.

- Chernoff bound for sum of independent Bernoulli random variables:

Let $X = \sum_{i=1}^n X_i$, where $X_i$'s are independent Bernoulli but not necessarily identically distributed. Let $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, then

$$P \left( X - \bar{X} \geq \varepsilon \bar{X} \right) \leq e^{-\bar{X} h(\varepsilon)}$$

$$P \left( X - \bar{X} \leq -\varepsilon \bar{X} \right) \leq e^{-\bar{X} h(\varepsilon)}$$

- Proof very similar to regular Chernoff bound:

$$P \left( X - \bar{X} \geq \varepsilon \bar{X} \right) \leq E \left[ e^{\varepsilon X / \bar{X}} \right] e^{-\bar{X} (1+\varepsilon)}$$

$$= \prod_{i=1}^n E \left[ e^{\varepsilon X_i / \bar{X}} \right] e^{-\bar{X} (1+\varepsilon)}$$
\[ n \prod_{i=1}^{n} \left( 1 - \bar{X}_i \right) + \bar{X}_i e^\theta \right)^{-\theta (1+\varepsilon)} \bar{X}_i \]

\[ = \exp \left( -\bar{X}_i \left( 1 - e^\theta + \varepsilon (1+\varepsilon) \right) \right) \]

Then optimize over \( \theta \).

We also need the following:

**Lemma:** Let \( \lambda = np > 1 \), then for any \( \varepsilon > 0 \), \( \exists K > 0 \) s.t.

\[ P_{\text{ext}}(\lambda) - \varepsilon \leq P(\left| C(v) \right| < K) \leq P_{\text{ext}}(\lambda) + \varepsilon \]

**Proof:** Recall the previous construction based on exploration. Conditional on \( \xi_1, \ldots, \xi_{k-1} \), \( \xi_k \) is Binomial \( \left( n-1 - \xi_1 - \ldots - \xi_{k-1}, p \right) \).

Then for any fixed \( k \), \( x = (x_1, \ldots, x_k) \in \mathbb{N}^k \),

\[ P(\xi_1 = x_1, \ldots, \xi_k = x_k) = P(\xi_1 = x_1) P(\xi_2 = x_2 | \xi_1 = x_1) \ldots P(\xi_k = x_k | \xi_1 = x_1, \ldots, \xi_{k-1} = x_{k-1}) \]

\[ = \left( 1 + O(1) \right) \prod_{i=1}^{k} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \]

where we have used the fact that Binomial \( (n, p) \rightarrow \text{Poisson}(\lambda) \) as \( n \to \infty \)

This shows that up to \( k \) steps, the vector \( (\xi_1, \ldots, \xi_k) \) can be approximated by a Galton-Watson branching process with Poisson \( (\lambda) \) offspring distribution.
\[ \mathbb{P}(|C(u)| \leq k_0) = \mathbb{P}(\text{number of offspring} \leq k_0) \quad \text{as} \quad n \to \infty. \]

\[ \implies p_{\text{ext}}(\lambda) \quad \text{as} \quad k_0 \to \infty. \]

Next, we show the following:

**Lemma.** \( \forall \lambda, \delta > 0, \exists k_0 \text{ s.t. for all } n \text{ large enough } (\lambda > np > 1) \)

\[ \mathbb{P}(\left| |C(u)| - (1 - p_{\text{ext}}(\lambda))n \right| > \delta n) \leq \varepsilon \]

**Proof.** Let \( X_k \) be the \( k \)th node explored after \( k \) steps, then as we saw before: \( X_k \sim \text{Binomial} \left( n - 1, 1 - (1 - p)^k \right) \). Thus

\[ \mathbb{P}(|C(u)| = k) \leq \mathbb{P}(A_k = 0) = \mathbb{P}(X_k = k - 1) \]

To apply Chernoff bound, need to compare \( k - 1 \) and \( \mathbb{E}[X_k] \).

\[ \frac{\mathbb{E}[X_k]}{k-1} = \frac{(n-1)(1-(1-p)^k)}{k-1} = \left(1 + o(1)\right) \frac{n}{k} \left(1 - e^{-kp}\right) \]

\[ = \left(1 + o(1)\right) g \left( \frac{k}{n} \right) \]

where \( g(x) = x - (1 - e^{-x}) \).

Check that \( g(x) \) is decreasing and
g(0) = 1

g(1) = 1 - e < 1

\( \Rightarrow \exists \text{ unique } x^* \text{ s.t. } g(x^*) = 1 \)

\( \in (0, 1) \)

\( x^* = 1 - e^{-x^*} \)

We know for a G-W process that \( p_{\text{ext}}(\lambda) \) is the solution in \( (0, 1) \) to
\[ x = e^{-\lambda (1-\delta)} \] This shows that \( x^* = \frac{1}{1 - P_{\text{Ext}}(\Lambda)} \).

Hence,

- If \( k < (1 - P_{\text{Ext}}(\Lambda) - \delta) n \), \( \frac{\mathbb{E}X_k}{k - 1} > 1 \) or \( \frac{k}{\mathbb{E}[X_k]} < 1 - \eta \) for some \( \eta \in (0, 1) \)

\[ P(\mid \mathbf{c(w)} \mid = k) \leq P(X_k \leq k) \]

\[ \leq P(X_k - \mathbb{E}[X_k] \leq -\eta \mathbb{E}[X_k]) \]

\[ \leq \exp \left( -\mathbb{E}[X_k] h(-\eta) \right) \]

\[ < \exp \left( -k h(-\eta) \right) \]

- If \( k > (1 - P_{\text{Ext}}(\Lambda) + \delta) n \), then \( \frac{k - 1}{\mathbb{E}[X_k]} > 1 + \eta \), so

\[ P(\mid \mathbf{c(w)} \mid = k) \leq P(X_k > k - 1) \]

\[ \leq P(X_k - \mathbb{E}[X_k] > \eta \mathbb{E}[X_k]) \]

\[ \leq \exp \left( -\mathbb{E}[X_k] h(\eta') \right) \]

\[ < \exp \left( -k \frac{h(\eta')}{1 - \eta'} \right) \]

So, in both cases, for large \( n \)

\[ P(\mid \mathbf{c(w)} \mid = k) < \exp(-k \eta') \]

where \( \eta' = \min(h(-\eta), \frac{h(\eta)}{1 - \eta}) > 0 \).

Finally,

\[ P\left( \mid \mathbf{c(w)} \mid > k_0, \mid \mathbf{c(w)} - (1 - P_{\text{Ext}}(\Lambda)) n \mid > \delta n \right) < \sum_{n \geq k > k_0} e^{-k \eta} \]
\[
\frac{e^{-\eta}}{1 - e^{-\eta}} \to 0 \quad \text{as} \quad k_0 \to \infty.
\]

Now, we conclude the proof of part (ii),

- Consider the following procedure for extracting the component of \( G(n, p) \)
  starting from node \( i = 1 \)

  1. Extract the connected component of \( i : C(i) \)
  2. If \( |C(i)| \leq k_0 \), pick another node not in the previously selected
      components
  3. If \( |C(i)| - (1 - \text{Pext}(k))n \leq \delta n \), STOP; declare success!
  4. If neither of 2 or 3, STOP; declare failure!

After \( i \) iterations (which have been neither failure nor success),
we are left with a \( G'(n', p') \) where

\[ n' \in (n - i k_0, n - i) \]

Since in each iteration, we remove at most \( k_0 \) nodes and at least one,

Hence \( n' p' \approx n p = \lambda \cdot (i, k_0 \text{ fixed}) \)

So we can apply the previous two lemmas to the remaining \( G' \) graph

at each iteration. At each iteration:

- the probability that \( |C(i)| \leq k_0 \) is less than \( \text{Pext}(N) + \varepsilon \).
the probability of failure is less than $\epsilon$.

so probability of success in greater than $1 - p_{ext}(x) - 2\epsilon$.

$\Rightarrow \Pr(\text{success in at most } k \text{ steps})$

$$\geq \sum_{i=1}^{k} \left( p_{ext} \right)^{i-1} \left( 1 - p_{ext} - 2\epsilon \right) = \left( 1 - p_{ext} - 2\epsilon \right) \left( \frac{1 - p_{ext}^k}{1 - p_{ext}} \right)$$

For any $\epsilon > 0$, we can choose $k$ large enough so that this probability is more than $1 - O(\epsilon)$.

So upon success, we have one giant component of the right size, up to $k$ component of size $\leq k_0$, and the remaining graph is a

$G(n'', p)$ with $n'' \in \left[ n\left( p_{ext}(x) - 8 \right) - k_0, n\left( p_{ext}(x) + 8 \right) \right]$.

Note that $n''p = \lambda p_{ext}(x) < 1$, i.e., $G(n'', p)$ is subcritical so its largest component has to be of logarithmic size.
Other properties of random graphs:

Connectivity

Q: Under what condition the graph is connected?
A: \( np = \lambda = c \log n \) for some \( c > 1 \).

Why? Suppose \( I_j = 1 \) (node \( j \) is isolated), then
\[
P(I_j = 1) = (1-p)^{n-1} = \left(1 - \frac{\lambda}{n} \right)^{n-1} \rightarrow e^{-\lambda} = n^{-c}, \text{ as } n \rightarrow \infty
\]

\[
P(\text{there is an isolated node}) = P\left( \bigcup_{j=1}^{n} (I_j = 1) \right) \leq n \cdot n^{-c} = n^{1-c}
\]
\[
\rightarrow 0 \text{ as } n \rightarrow \infty \text{ and } c > 1.
\]

So with high probability there are no isolated nodes, of course this is just a necessary condition for \( G(n, p) \) to be connected. Is \( c > 1 \) right? i.e., is there an isolated node with high probability if \( c \leq 1 \)? The answer is yes!

A nice way to see this is by using the Stein-Chen method. Let \( W = \sum_{j=1}^{n} I_j \). \( I_j \sim \text{Bernoulli}(n^{-c}) \)

If \( I_j \)'s are independent, then \( W \sim \text{Poisson}(n \cdot n^{-c}) \) and
\[
P(W = 0) = e^{-n^{1-c}} \rightarrow 0 \text{ if } c \leq 1.
\]
The problem is that $\overline{I}_j$'s are \underline{not} independent. The Stein-Chen method says that if the influence of each $I_j$ individually on $W$ is small ($\approx \partial W$), then $W \approx \text{Poisson} \left( \sum_{i=1}^{n} I_i \right)$.

Let $\overline{I}_j = \pi_j = P( I_j = 1)$.

\[
\partial W = \sum_{i=1}^{n} \pi_j \partial W_j
\]

\[
\partial_j W = E \left[ \left| \sum_{i=1}^{n} I_i - \sum_{i=1 \neq j}^{n} I_i \right| \right]
\]

\[
\text{if } i \neq j, \left( \prod_{N \neq i, j} (1 - \xi_{vi}) \right) = \text{"i" being isolated, given "j" is isolated.}
\]

\[
\xi_{vi} = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{o.w.}
\end{cases}
\sim \text{Bernoulli} \left( (1-p)^{n-2} \right)
\]

\[
\partial_j W = \sum_{i \neq j} E \left[ \left| I_i - \xi_{vi} \right| \right]
\]

\[
= \sum_{i \neq j} E \left[ \xi_{ij} \prod_{N \neq i, j} (1 - \xi_{vi}) \right]
\]

\[
= (n-1) p (1-p)^{n-2} = \frac{p(n-1)}{1-p} (1-p)^{-1} \sim \text{order } n^{-c}
\]

Definition: variation distance between two probability distributions $P$ and $Q$:

\[
d_{\text{var}}(P, Q) = \frac{1}{2} \sum_{i=1}^{n} |P(i) - Q(i)|
\]

Stein-Chen lemma:

\[
d_{\text{var}}(P_W, \text{Poisson}(n \pi_j)) = O(\partial W) \rightarrow 0
\]
It remains to show that the graph has no connected components of size $2, 3, \ldots, \frac{n}{2}$.

(Retrieve Chapter 3 of the book.)

In the context of Reed–Frost epidemic, the connectivity corresponds to all nodes getting ultimately infected.

**Diameter:**

Recall $d_G(u, v)$: minimum distance between nodes $u, v$ in $G$.

The diameter $D(G) = \sup_{u,v} d_G(u, v)$.

In the context of Reed–Frost epidemic, diameter provides an upper bound on the time it takes for the epidemic to reach all nodes.

**Lemma:** Given a graph $G$ on $n$ nodes, with maximum degree $\Delta > 2$,

its diameter $D$ satisfies:

\[
\frac{n}{2} < 1 + \frac{(\Delta-1)\Delta}{\Delta-2} - 1
\]

or

\[
D \geq \frac{\log(n) + 2\left(1 - \frac{2}{\Delta} + \frac{1}{\Delta}\right)}{\log(\Delta-1)}
\]

**Proof:**

$\delta_i(u) = \# of nodes in the $i$-hop distance of $u$.

$\delta_1(u) < \Delta$, \ldots, $\delta_i(u) < \Delta(\Delta-1)^{i-1}$.
So
\[ n = 1 + d_1(w) + \ldots + d_D(w) \leq 1 + D(1 + (1-\Delta) + \ldots + (1-\Delta)^{D-1}) \]

Graphs with small diameter have many applications, e.g., efficient routing.

\( E_R \) graphs have diameter close to optimal (lower bound in the lemma). Precisely, suppose \( p = \frac{c \log n}{n}, \ c > 1 \), i.e., average degree is \( \lambda = \frac{c \log n}{(c > 1)} \). Then the diameter concentrates on only a finite number of values around \( \log n \frac{2}{\log \lambda} \).

(we ignore the part here, It can be found in Ch. 4 of the book)