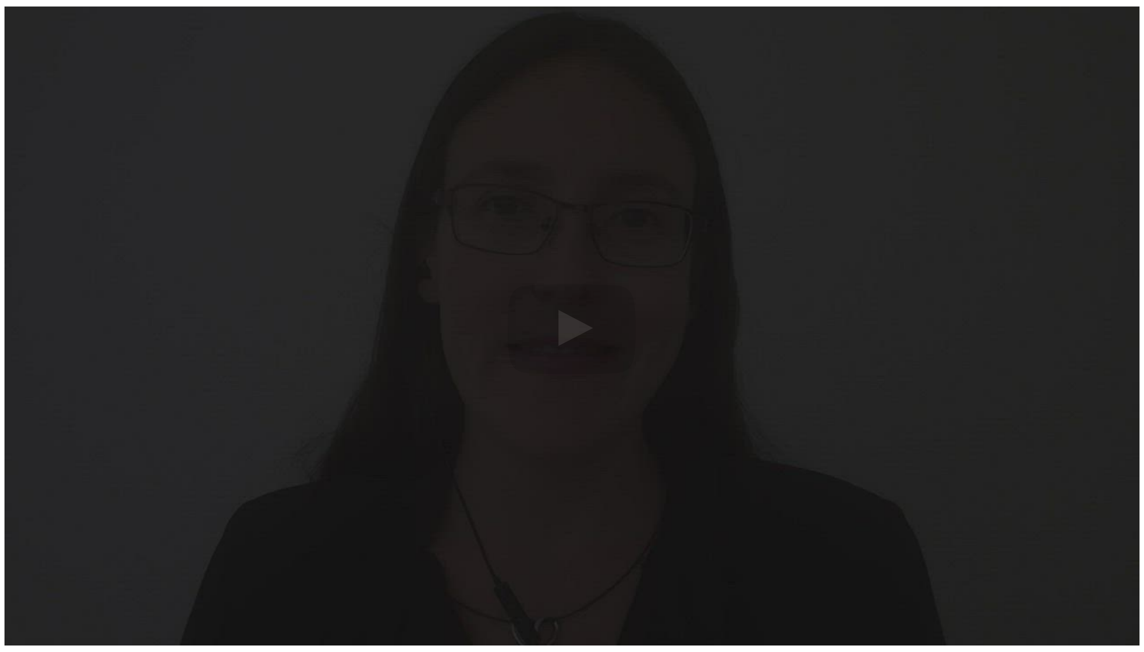
3. Erdos-Renyi model

Bookmark this page

Exercises due Oct 27, 2021 17:29 IST Completed

The Erdos-Renyi model – definition and structure



0:00 / 0:00 2.0x

Start of transcript. Skip to the end.

Prof Uhler: OK, so welcome back in this third lecture

in the networks module, where we're

talking about network models.

So we already discussed a bit about the motivation

for actually looking at network

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Erdos-Renyi model

2/2 points (graded)

Random graph models are a specification for a graph where the properties of the graph (such as the edges) are generated randomly according to some parameterized probability distribution.

The Erdos-Renyi model is a common model for a random graph. It is parameterized by

- n: The number of nodes in the graph.
- p: The probability of that any two nodes are connected by an edge.

and the Erdos-Renyi model is denoted by $G(n, p)$.

The number of nodes in the graph is a fixed parameter, and is not random. The edges, on the other hand, are drawn randomly from a binomial distribution.

For any pair of nodes in the graph, the probability to connect this pair is independent of all other pairs of nodes. The probability to connect the pair is p, and so the edge can be modelled as a binomial distribution: Binomial($1, p$).

Suppose we have an Erdos-Renyi model $G(n, 1)$. How many edges will a random realization of the graph have?

- ( ) $\binom{n}{3}$
- (•) $\binom{n}{2}$
- ( ) We can't say in advance, the number is random.
- ( ) n

Will this random realization be a complete graph?

- [x] Yes
- [ ] No

Solution:

For an Erdos-Renyi model $G(n, 1)$, we have $p = 1$, and so the probability of any pair of nodes being connected is 1. Therefore all nodes are connected, and so the graph is complete and it has $\binom{n}{2}$ edges.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

The edge probability

2/2 points (graded)

Suppose we have an Erdos-Renyi model $G(n, p)$. Let m be the number of edges of the graph.

What is the expected number of edges in terms of n and p?

(For the problems in this course, if you arrive at any expressions in terms of binomial coefficients such as \binom {b,k}, enter a simplified algebraic expression without binomial coefficients.)

$\mathbb{E}[m] =$ (n*(n-1)*p)/2 **Answer:** p*n*(n-1)/2

If we observe a random realization of this graph model to have m edges, then what is the maximum likelihood estimate for p in terms of n and m?

$\hat{p} =$ (2*m)/(n*(n-1)) **Answer:** 2*m/(n*(n-1))

STANDARD NOTATION

Solution:

There are $\binom{n}{2}$ possible edges, and each edge, μ_i, is selected according to $\mu_i \sim \text{Binomial}(1, p)$. So the total number of edges is

$$\begin{aligned}
\mathbb{E}[m] &= \mathbb{E}\left[\sum_i \mu_i\right] \\
&= \sum_i \mathbb{E}[\mu_i] \\
&= \left(\sum_i 1\right) \mathbb{E}[\mu] \\
&= \binom{n}{2} p \\
&= p\frac{n(n-1)}{2}
\end{aligned}$$

For the edge observations μ_i, the likelihood is

$$\ldots = \prod p^{\mu_i}(1-p)^{1-\mu_i}$$

$$\mathcal{L} = \prod_i p^{\mu_i} (1-p)^{1-\mu_i}$$

$$\ln \mathcal{L} = \sum_i \mu_i \ln p + (1-\mu_i) \ln (1-p)$$

Taking the derivative with respect to p gives the MLE $\hat{p}$:

$$\begin{aligned}
\frac{d \ln \mathcal{L}}{dp} &= \sum_i \frac{\mu_i}{p} - \frac{1-\mu_i}{1-p} \\
0 &= \sum_i \frac{\mu_i}{\hat{p}} - \frac{1-\mu_i}{1-\hat{p}} \\
&= \frac{m}{\hat{p}} - \frac{\left(\sum_i 1\right) - m}{1-\hat{p}} \\
\frac{\binom{n}{2} - m}{1-\hat{p}} &= \frac{m}{\hat{p}} \\
\hat{p}\binom{n}{2} - \hat{p}m &= m - \hat{p}m \\
\hat{p} &= \frac{m}{\binom{n}{2}} \\
&= \frac{2m}{n(n-1)}
\end{aligned}$$

Submit You have used 1 of 4 attempts

Answers are displayed within the problem

Degree distribution

2/2 points (graded)

Consider node j in an Erdos-Renyi model $G(n, p)$. Potentially, it could be connected through edges to up to $n-1$ other nodes, where the presence of each edge is distributed according to Binomial$(1, p)$. Thus, the degree, k_j, of node j is the sum of $n-1$ random variables distributed according to Binomial$(1, p)$. Therefore, $k_j \sim$ Binomial$(n-1, p)$.

It follows that the expected value of the degree is

$$\mathbb{E}[k_j] = p(n-1).$$

Let $K = \sum_{j=1}^{n} k_j$ be the sum of all node degrees. What is the MLE of p in terms of n and K?

$\hat{p} =$

K/(n*(n-1)) ✔ **Answer:** K/(n*(n-1))

Do the node degrees follow a power-law distribution?

- ○ Yes
- ◉ No

✔

Solution:

We could go through the full MLE derivation for $\hat{p}$; or, we can note that the sum of all the node degrees is equal to twice the number of edges:

$$\sum_{j=1}^{n} k_j = 2m$$

Substituting this into our result for the MLE in terms of m in the previous problem gives:

$$\begin{aligned}\hat{p} &= \frac{2m}{n\,(n-1)} \\ &= \frac{\sum_{j=1}^{n} k_j}{n\,(n-1)} \\ &= \frac{K}{n\,(n-1)}\end{aligned}$$

The node degrees are binomially distributed, which is not a power-law. So the node degrees do not follow a power-law distribution.

Submit

You have used 1 of 3 attempts

Answers are displayed within the problem

Graph structure

2/2 points (graded)

An Erdos-Renyi model, $G\,(n, p)$, displays, in the large node limit of $n \to \infty$, a phase transition in the global graph structure at two points.

For $p < \frac{1}{n}$, there are many small components. The size of these components tends to be bounded above:

$$P\,(S_{\text{max}} > c \ln n) \;\to 0 \quad \text{as } n \to \infty$$

where S_{max} is the size of the largest component in the graph, and $c = 2m/n$ is the average node degree. Note that this is strictly only a bound as $n \to \infty$, for any finite n there is a non-zero probability of all nodes being connected.

Between $\frac{1}{n} < p < \frac{\ln n}{n}$, a **giant** component emerges. This giant component has a size that is a significant fraction of n, around $(c-1)\,n$ for $c \approx 1$ (the relation is not linear for larger c). There will only be one of these giant components; other components will exist, but they will observe the $S < c \ln n$ bound discussed above.

Above $p > \frac{\ln n}{n}$, the giant component includes all nodes in the graph and the graph becomes connected.

Although these phase transitions are only defined in the $n \to \infty$ limit, you can observe them even for n as small as 100. When n is small, the phase transition points become “fuzzy" in that the transition may not occur at exactly the points described above, but close to them.

Suppose that you are given a graph with $n = 1 \times 10^6$ and $m = 200 \times 10^3$.

What would you expect to see if the process that created this graph can be described by an Erdos-Renyi model?

- (•) Many small connected components.
- ( ) Some small components along with a giant connected component.
- ( ) A single connected component.

✔

If you are given a graph with $n = 1 \times 10^6$ and $m = 20 \times 10^6$, what would you expect to see if the graph came from an Erdos-Renyi model?

- ◯ Many small connected components.
- ◯ Some small components along with a giant connected component.
- ◉ A single connected component.

✔

Solution:

We can use the MLE of p to estimate this parameter.

For $n = 1 \times 10^6$ and $m = 200 \times 10^3$, we have

$$\begin{aligned} \hat{p} &\approx \frac{2 \times 200 \times 10^3}{10^{12}} \\ &= 400 \times 10^{-9} \\ &< \frac{1}{n} = 10^{-6} \end{aligned}$$

So $\hat{p}$ is below the first phase transition and we expect many small connected components.

For $n = 1 \times 10^6$ and $m = 20 \times 10^6$, we have

$$\begin{aligned} \hat{p} &\approx \frac{2 \times 20 \times 10^6}{10^{12}} \\ &= 40 \times 10^{-6} \\ &> \frac{\ln n}{n} \approx 15 \times 10^{-6} \end{aligned}$$

So $\hat{p}$ is above the second phase transition and we expect a single connected component.

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Module 3: Network Analysis:Graphical models / 3. Erdos-Renyi model

Add a Post

< All Posts

Tutorial/guidance time (supplementary resources)

discussion posted 2 months ago by **lam_trinh** (Community TA)

Hi all,

I believe these 2 resources will be helpful for this tab:

1. Intro to Erdos-Renyi graph along with some primer on binomial distribution at the end (~12 mins on main topic + ~3 mins on binomial distribution): https://www.youtube.com/watch?v=XwVZ5VttrW0
2. For those brave souls who skip probability and/or statistics classes (like me) in this Micromasters and go straight for this class, this refresher/primer on MLE may help you: https://online.stat.psu.edu/stat504/lesson/1/1.5

As usual, if these resources are helpful, please give a thumbs up (or plus up) so I know these guides are good! Happy random graphing! :)

This post is visible to everyone.

Add a Response

3 responses

MPaulaWilde
2 months ago

Thanks

Add a comment

destinyZ
2 months ago

Thanks!

Add a comment

gastorgaa
2 months ago

Thanks

Add a comment

Showing all responses

Add a response:

Preview

Submit

‹ Previous Next ›

< Previous Next >

edX

About
Affiliates
edX for Business
Open edX
Careers
News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

Connect

Blog
Contact Us
Help Center
Media Kit
Donate

GET IT ON Google play
Download on the App Store