# Random Graphs

Catherine Li

Euler Circle

July 2021

# Random Graphs

**Definition**

A *random graph* $G(n, p)$ is the probability space of graphs with $n$ vertices, such that each edge has probability $p$ of being in the graph.
Alternatively, $G(n, m)$ is the probability space of graphs with $n$ vertices and $m$ edges, with each edge having equal probability of being chosen.

# Random Graphs

## Definition

A *random graph* $G(n, p)$ is the probability space of graphs with $n$ vertices, such that each edge has probability $p$ of being in the graph.
Alternatively, $G(n, m)$ is the probability space of graphs with $n$ vertices and $m$ edges, with each edge having equal probability of being chosen.

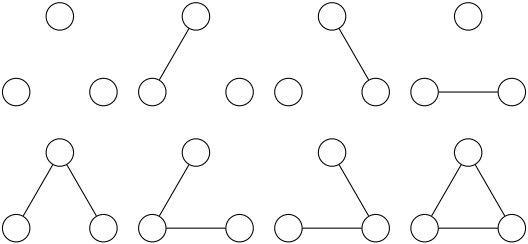
$$G\left(3, \frac{1}{2}\right):$$

# Random Graphs

## Definition

A *random graph* $G(n, p)$ is the probability space of graphs with $n$ vertices, such that each edge has probability $p$ of being in the graph.
Alternatively, $G(n, m)$ is the probability space of graphs with $n$ vertices and $m$ edges, with each edge having equal probability of being chosen.

$G\left(3, \frac{1}{2}\right)$ :



# Thresholds

## Definition

A threshold $r(n)$ for a property $A$ is a function such that, as $n \to \infty$,

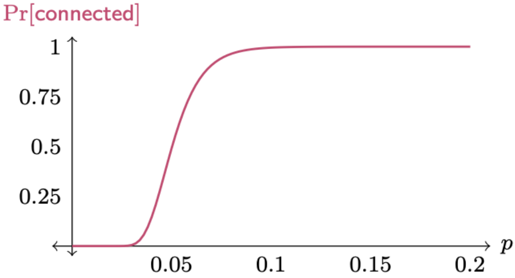
$$\Pr[G(n,p) \in A] \to \begin{cases} 0 & \text{if } p \ll r(n), \\ 1 & \text{if } p \gg r(n), \end{cases}$$

# Thresholds

## Definition

A threshold $r(n)$ for a property $A$ is a function such that, as $n \to \infty$,

$$\Pr[G(n,p) \in A] \to \begin{cases} 0 & \text{if } p \ll r(n), \\ 1 & \text{if } p \gg r(n), \end{cases}$$



# History

- 1959: Erdős and Rényi published *On random graphs*, introducing the model $G(n, m)$
- 1959: Gilbert proposed $G(n, p)$ model
- 1969: Travers and Milgram studied the small world problem
- 1984: Béla Bollobás published *Random Graphs*
- 2006: Kahn-Kalai conjectured expectation threshold $\approx$ phase transition
- 2022: Park and Pham proved the Kahn-Kalai conjecture

# First and Second Moment Methods

## Theorem (Markov's Inequality)

*For nonnegative $X$ and positive $a$,* $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

# First and Second Moment Methods

## Theorem (Markov's Inequality)

*For nonnegative $X$ and positive $a$,* $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

## Theorem (Chebyshev's Inequality)

*Let $\mu$ and $\sigma$ denote the mean and standard deviation of nonnegative $X$. For positive $\lambda$,*

$$\Pr[|X - \mu| \geq \lambda\sigma] \leq \frac{1}{\lambda^2}.$$

# First and Second Moment Methods

**Theorem (Markov's Inequality)**

*For nonnegative $X$ and positive $a$,* $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Theorem (Chebyshev's Inequality)**

*Let $\mu$ and $\sigma$ denote the mean and standard deviation of nonnegative $X$. For positive $\lambda$,*

$$\Pr[|X - \mu| \geq \lambda\sigma] \leq \frac{1}{\lambda^2}.$$

**Corollary**

*For nonnegative $X$,*

$$\Pr[X = 0] \leq \frac{\mathrm{Var}[X]}{\mathbb{E}[X]^2} = \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2}{\mathbb{E}[X]^2}.$$

# Containing Triangles

**Theorem**

*The threshold for containing triangles is* $p = \frac{1}{n}$.

# Containing Triangles

**Theorem**

*The threshold for containing triangles is* $p = \frac{1}{n}$.

- $X = \#$ of triangles

# Containing Triangles

**Theorem**

*The threshold for containing triangles is* $p = \frac{1}{n}$.

- $X = \#$ of triangles
- $\mathbb{E}[X] = \binom{n}{3} p^{\binom{3}{2}} = O(n^3) p^3$

# Containing Triangles

## Theorem

*The threshold for containing triangles is $p = \frac{1}{n}$.*

- $X = \#$ of triangles
- $\mathbb{E}[X] = \binom{n}{3} p^{\binom{3}{2}} = O(n^3)p^3$
- $p \ll \frac{1}{n}$ : Markov's Inequality

$$\Pr[X \geq 1] \leq \mathbb{E}[X]/1 = O(n^3)p^3 \ll 1$$

# Containing Triangles

**Theorem**

*The threshold for containing triangles is $p = \frac{1}{n}$.*

- $X = \#$ of triangles
- $\mathbb{E}[X] = \binom{n}{3} p^{\binom{3}{2}} = O(n^3)p^3$
- $p \ll \frac{1}{n}$ : Markov's Inequality

$$\Pr[X \geq 1] \leq \mathbb{E}[X]/1 = O(n^3)p^3 \ll 1$$

- $p \gg \frac{1}{n}$ : $\mathbb{E}[X^2]$, expected number of pairs of triangles

# Containing Triangles

**Theorem**

*The threshold for containing triangles is $p = \frac{1}{n}$.*

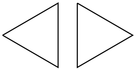
- $X = \#$ of triangles
- $\mathbb{E}[X] = \binom{n}{3} p^{\binom{3}{2}} = O(n^3)p^3$
- $p \ll \frac{1}{n}$ : Markov's Inequality

$$\Pr[X \geq 1] \leq \mathbb{E}[X]/1 = O(n^3)p^3 \ll 1$$

- $p \gg \frac{1}{n}$ : $\mathbb{E}[X^2]$, expected number of pairs of triangles
  - 4 cases: share 0,1,2,3 vertices

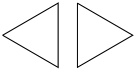
# Containing Triangles: Casework

- 0 shared vertices: $O(n^6)p^6$

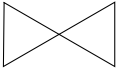


# Containing Triangles: Casework

- 0 shared vertices: $O(n^6)p^6$

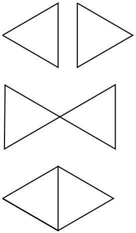


- 1 shared vertex: $O(n^5)p^6$



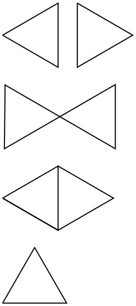
# Containing Triangles: Casework

- 0 shared vertices: $O(n^6)p^6$
- 1 shared vertex: $O(n^5)p^6$
- 2 shared vertices: $O(n^4)p^5$



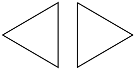
# Containing Triangles: Casework

- 0 shared vertices: $O(n^6)p^6$
- 1 shared vertex: $O(n^5)p^6$
- 2 shared vertices: $O(n^4)p^5$
- 3 shared vertices: $O(n^3)p^3$

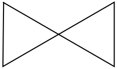


# Containing Triangles: Casework

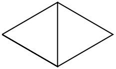
- 0 shared vertices: $O(n^6)p^6$



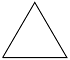
- 1 shared vertex: $O(n^5)p^6$



- 2 shared vertices: $O(n^4)p^5$



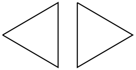
- 3 shared vertices: $O(n^3)p^3$



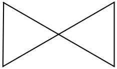
$$
\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&\sim O(n^6)p^6 + O(n^5)p^6 + O(n^4)p^5 + O(n^3)p^3 - O(n^6)p^6 \\
&= O(n^6)p^6 \left( \frac{1}{n} + \frac{1}{n^2 p} + \frac{1}{n^3 p^3} \right)
\end{aligned}
$$

# Containing Triangles: Casework

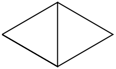
- 0 shared vertices: $O(n^6)p^6$



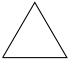
- 1 shared vertex: $O(n^5)p^6$



- 2 shared vertices: $O(n^4)p^5$



- 3 shared vertices: $O(n^3)p^3$



$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &\sim O(n^6)p^6 + O(n^5)p^6 + O(n^4)p^5 + O(n^3)p^3 - O(n^6)p^6 \\ &= O(n^6)p^6\left(\frac{1}{n} + \frac{1}{n^2p} + \frac{1}{n^3p^3}\right)\end{aligned}$$

- $\Pr[X = 0] \leq \frac{\mathrm{Var}[X]}{\mathbb{E}[X]^2} \ll 1$

# Containing Cliques

**Theorem**

*The threshold for containing* $K_4$ *is* $p = n^{-\frac{2}{3}}$.

# Containing Cliques

**Theorem**

*The threshold for containing $K_4$ is $p = n^{-\frac{2}{3}}$.*

**Theorem**

*The threshold for containing $K_5$ is $p = n^{-\frac{1}{2}}$.*

**Theorem**

*The threshold for containing* $K_4$ *is* $p = n^{-\frac{2}{3}}$.

**Theorem**

*The threshold for containing* $K_5$ *is* $p = n^{-\frac{1}{2}}$.

$$\vdots$$

**Theorem**

*The threshold for containing the complete graph on* $r$ *vertices* ($K_r$) *is* $p = n^{-\frac{2}{r-1}}$.

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

- $X_k = \#$ of components of order $k$

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

- $X_k = \#$ of components of order $k$
- $p \ll \frac{\log n}{n}$

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

- $X_k = \#$ of components of order $k$
- $p \ll \frac{\log n}{n}$
  - isolated vertices: $\mathbb{E}[X_1] = n(1-p)^{n-1}, \mathbb{E}[X_1^2] \leq \mathbb{E}[X_1]^2 + 1$

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

- $X_k = \#$ of components of order $k$
- $p \ll \frac{\log n}{n}$
  - isolated vertices: $\mathbb{E}[X_1] = n(1-p)^{n-1}, \mathbb{E}[X_1^2] \leq \mathbb{E}[X_1]^2 + 1$
  - Chebyshev's Inequality

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

- $X_k = \#$ of components of order $k$
- $p \ll \frac{\log n}{n}$
  - isolated vertices: $\mathbb{E}[X_1] = n(1-p)^{n-1}, \mathbb{E}[X_1^2] \leq \mathbb{E}[X_1]^2 + 1$
  - Chebyshev's Inequality
- $p \gg \frac{\log n}{n}$

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

- $X_k = \#$ of components of order $k$
- $p \ll \frac{\log n}{n}$
  - isolated vertices: $\mathbb{E}[X_1] = n(1-p)^{n-1}, \mathbb{E}[X_1^2] \leq \mathbb{E}[X_1]^2 + 1$
  - Chebyshev's Inequality
- $p \gg \frac{\log n}{n}$
  - $\mathbb{E}[X_k] \leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)}$

# Connectivity

**Theorem**

*The threshold for connectivity is* $p = \frac{\log n}{n}$.

- $X_k = \#$ of components of order $k$
- $p \ll \frac{\log n}{n}$
  - isolated vertices: $\mathbb{E}[X_1] = n(1-p)^{n-1}, \mathbb{E}[X_1^2] \leq \mathbb{E}[X_1]^2 + 1$
  - Chebyshev's Inequality
- $p \gg \frac{\log n}{n}$
  - $\mathbb{E}[X_k] \leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)}$
  - Markov's Inequality: $\sum_{k=1}^{\lfloor n/2 \rfloor} \mathbb{E}[X_k] \to 0$ suffices

# The Galton-Watson Branching Process

- Consider an organism that reproduces by having a random number of children, $X$

# The Galton-Watson Branching Process

- Consider an organism that reproduces by having a random number of children, $X$
- Does the population eventually go extinct?

# The Galton-Watson Branching Process

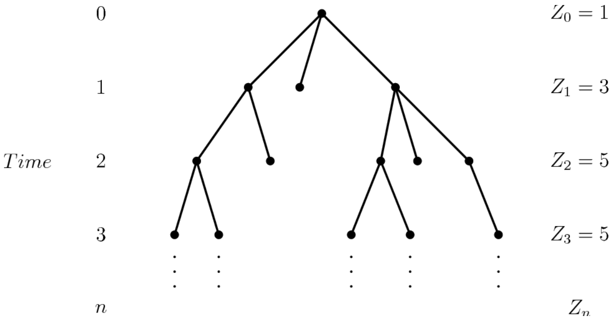
- Consider an organism that reproduces by having a random number of children, $X$
- Does the population eventually go extinct?
  - if $\mathbb{E}[X] < 1$, yes

# The Galton-Watson Branching Process

- Consider an organism that reproduces by having a random number of children, $X$
- Does the population eventually go extinct?
  - if $\mathbb{E}[X] < 1$, yes
  - if $\mathbb{E}[X] > 1$, possibly

# The Galton-Watson Branching Process

- Consider an organism that reproduces by having a random number of children, $X$
- Does the population eventually go extinct?
  - if $\mathbb{E}[X] < 1$, yes
  - if $\mathbb{E}[X] > 1$, possibly



# The Giant Component

- introduced by Erdős and Rényi

# The Giant Component

- introduced by Erdős and Rényi
- traversing components: modeled by the branching process

# The Giant Component

- introduced by Erdős and Rényi
- traversing components: modeled by the branching process
- small components: $O(\log n)$

# The Giant Component

- introduced by Erdős and Rényi
- traversing components: modeled by the branching process
- small components: $O(\log n)$
- large components: $O(n)$

# The Giant Component

- introduced by Erdős and Rényi
- traversing components: modeled by the branching process
- small components: $O(\log n)$
- large components: $O(n)$
- $p = \frac{c}{n}$ :

# The Giant Component

- introduced by Erdős and Rényi
- traversing components: modeled by the branching process
- small components: $O(\log n)$
- large components: $O(n)$
- $p = \frac{c}{n}$ :
  - $c < 1$ : many small components

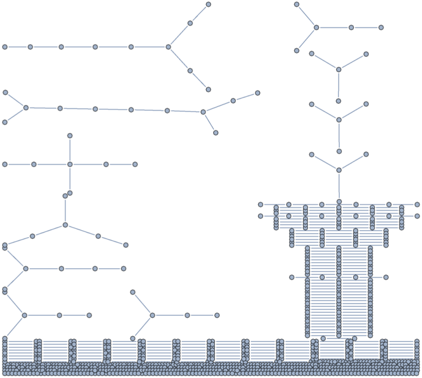
# The Giant Component

- introduced by Erdős and Rényi
- traversing components: modeled by the branching process

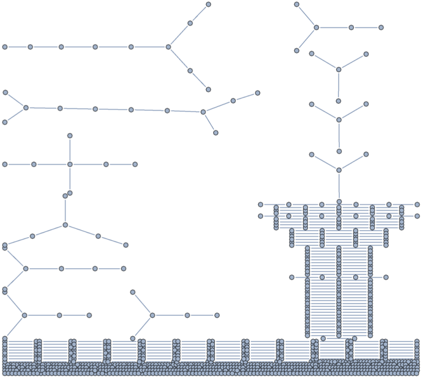
- small components: $O(\log n)$
- large components: $O(n)$

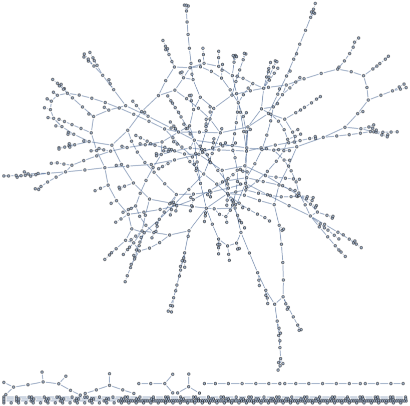
- $p = \frac{c}{n}$ :
  - $c < 1$ : many small components
  - $c > 1$ : giant component and small components

# The Giant Component



# The Giant Component





# Thank you

Thanks for listening!