

Random Graphs (Erdos-Renyi)

A probabilistic model of graphs.

$G(n, p)$ random graph model: Given a n -vertex labeled graph, an (undirected) edge occurs between a pair of vertices with probability p , independently of other pairs.

The probability of a graph G of k edges is given by

$$p^k (1 - p)^{\binom{n}{2} - k}$$

The probability that a graph $G \in G(n, p)$ contains k edges is binomially distributed:

$$\binom{\binom{n}{2}}{k} p^k (1 - p)^{\binom{n}{2} - k}$$

Bounding Deviation from Expectation

Theorem 1. [Markov Inequality] For any non-negative random variable and for any $a > 0$

$$\Pr(X \geq a) \leq \frac{E[X]}{a}.$$

Proof. For any $a > 0$, let

I be an indicator r.v. for the event $X \geq a$.

$$I \leq X/a. \quad E[I] \leq E[X]/a \quad \square$$

Example: What is the probability of getting more than $\frac{3N}{4}$ heads in N coin flips?

$$\leq \frac{N/2}{3N/4} \leq \frac{2}{3}.$$

Variance

Definition 1. *The variance of a random variable X is*

$$\text{Var}[X] = E[(X - E[X])^2].$$

Definition 2. *The standard deviation of a random variable X is*

$$\sigma(X) = \sqrt{\text{Var}[X]}.$$

Example: Let X be a 0-1 random variable with $Pr(X = 0) = Pr(X = 1) = 1/2$.

$$E[X] = 1/2.$$

$$Var[X] = \frac{1}{2}\left(1 - \frac{1}{2}\right)^2 + \frac{1}{2}\left(0 - \frac{1}{2}\right)^2 = \frac{1}{4}.$$

Chebyshev's Inequality

Theorem 2. *For any random variable*

$$Pr(|X - E[X]| \geq a) \leq \frac{Var[X]}{a^2}.$$

Proof.

$$Pr(|X - E[X]| \geq a) = Pr((X - E[X])^2 \geq a^2)$$

By Markov inequality

$$\begin{aligned} Pr((X - E[X])^2 \geq a^2) &\leq \frac{E[(X - E[X])^2]}{a^2} \\ &= \frac{Var[X]}{a^2} \end{aligned}$$

□

Theorem 3. *For any random variable*

$$\Pr(|X - E[X]| \geq a\sigma[X]) \leq \frac{1}{a^2}.$$

Theorem 4. *For any random variable*

$$\Pr(|X - E[X]| \geq \epsilon E[X]) \leq \frac{\text{Var}[X]}{\epsilon^2 (E[X])^2}.$$

First and Second Moment Method

Theorem 5. *Let X be a non-negative integer-valued r.v. Then*

$$\Pr(X \geq 1) \leq E(X)$$

$$\Pr(X = 0) \leq \frac{\text{Var}(X)}{(E(X))^2} = \frac{E(X^2)}{(E(X))^2} - 1$$

Proof:

$$\Pr(X = 0) \leq \Pr(|X - E(X)| \geq E(X)) \leq \frac{\text{Var}(X)}{(E(X))^2}$$

These help in showing $\Pr(X = 0) \rightarrow 1$ and $\Pr(X = 0) \rightarrow 0$.

Thresholds in Random Graphs

Theorem 6. For the $G(n, p)$ random graph model, let $p = c \frac{\log n}{n}$. If $c > 1$ then almost all graphs have no isolated vertices and if $c < 1$ almost all graphs have at least one isolated vertex.

Proof:

Upper Threshold: Let X denote the number of isolated vertices in a random $G \in G(n, p)$.

Let X_i be the indicator r.v. for a vertex to be isolated.

$$E[X] = \sum_{i=1}^n X_i = n(1 - p)^{n-1}$$

$$(1 - p)^n = e^{n \log(1-p)} = e^{n(-p - p^2/2 - p^3/3 \dots)}$$

$$= e^{-np} e^{-np^2(1/2 + p/3 + \dots)}$$

$$\sim e^{-np} \text{ provided } np^2 \rightarrow 0.$$

$$E[X] = n(1 - p)^{n-1} \sim ne^{-np} \sim n^{1-c}$$

Thus, $E(X) \rightarrow 0$ if $c > 1$ and

$E(X) \rightarrow \infty$ if $c < 1$.

Lower Threshold:

$$\begin{aligned} E[X^2] &= \sum_{i=1}^n E(X^2) + \sum_{i \neq j} X_i X_j \\ &= E(X) + n(n-1)E(X_1 X_2) \\ &= E(X) + n(n-1)(1-p)^{2(n-2)+1} \end{aligned}$$

$$\begin{aligned} \Pr(X = 0) &\leq \frac{E(X^2)}{(E(X))^2} - 1 \\ &= \frac{1}{E(X)} + \frac{n(n-1)(1-p)^{2(n-2)+1}}{n^2(1-p)^{2(n-1)}} - 1 \\ &\sim \frac{1}{E(X)} \rightarrow 0 \text{ if } c < 1. \end{aligned}$$

Random Geometric Graphs

Assume that n points are thrown randomly in a unit square.

The random geometric graph $G(n, r)$ has n nodes corresponding to the n points and two nodes are connected by an edge if they are within a distance of r of each other. We assume the L_∞ norm, i.e., for two points $u = (x_1, y_1)$, $v = (x_2, y_2)$, $d(u, v) = \max(|x_1 - x_2|, |y_1 - y_2|)$.

A popular model for Sensor Networks.

Connectivity of $G(n, r)$

Theorem 7. *If $r \geq \sqrt{\frac{c \log n}{n}}$, where c is a constant > 4 , then $G(n, r)$ is connected asymptotically almost surely, i.e., $\Pr(G(n, r) \text{ is connected}) \rightarrow 1$ as $n \rightarrow \infty$.*

Proof. Divide the unit square into bins of size $r/2 \times r/2$.

Number of bins is $4/r^2$.

$G(n, r)$ is connected if no bin is empty.

Let $r = \sqrt{\frac{c \log n}{n}}$.

The probability that a bin is empty is

$$(1 - r^2/4)^n \leq e^{-nr^2/4} = n^{-c/4}$$

Let X be the number of empty bins.

$E[X] = \frac{4}{r^2} n^{-c/4} = \frac{4n}{c \log n} n^{-c/4} \leq n^{1-c/4} \rightarrow 0$ if $c > 4$.

$$\Pr(X > 0) \leq E[X] \rightarrow 0. \quad \square$$

Power-Law graphs

The degree distribution of the network follows a power law.

The number of vertices of degree k is $ck^{-\beta}$.

β is the power-law exponent.

Power-law degree distribution has been observed in:

Internet ($\beta = 2.1$)

World Wide Web ($\beta = 2.1$)

Social networks (movie actors graph with $\beta = 2.3$, citation graph with $\beta = 3$)

Biological networks (protein domains with $\beta = 1.6$

Protein-protein interaction graphs with $\beta = 2.5$).

In most real-world graphs, β ranges between 1 and 4.

Random power-law graphs

Generalization of the Erdos-Renyi ($G(n, p)$) random graphs.

Given a degree sequence $\langle d_1, \dots, d_n \rangle$, the *expected degree* model of the random graph is:

Probability of an edge between vertices i and j is $\frac{d_i d_j}{2m}$.

Other interesting graph classes

1. Random Regular graphs, Expanders (P2P networks).
2. Planar graphs and graphs excluding a fixed minor. (Planar graphs exclude K_5 and $K_{3,3}$.)
3. Networks with low doubling dimension.

The doubling dimension of a metric space is the smallest $\alpha > 0$ such that every ball can be covered by at most 2^α balls of radius r .

For (undirected, weighted) graphs: a graph has doubling dimension α if the metric space induced by the shortest path distance on the graph has doubling dimension α .

Unit disk graph has a constant doubling dimension.