Math 554

Hemanshu Kaul

kaul@iit.edu

Random Graph Models & Properties

Defn Given a sequence Ω_n of probability spaces, let q_n be the probability that property Q holds in Ω_n.
We say Q holds almost always (asymptotatically almost always) if $\lim_{n\to\infty} q_n = 1$.

}

Random Graph Models & Properties

<u>Defn</u> Given a sequence Ω_n of probability spaces, let q_n be the probability that property Q holds in Ω_n.
We say Q holds <u>almost always</u> (asymptotatically almost always) if $\lim_{n\to\infty} q_n = 1$.

Standard or Binomial or Erdös-Renyi Random Graph Model

Let p be a function of n, $p(n)$.
The probability space $G(n,p)$ is defined on a vertex set $[n]=\{1,2,..,n\}$ as follows: let any pair of vertices form an edge, independently with probab. p.

So, each graph with m edges occurs with probability $p^m(1-p)^{\binom{n}{2}-m}$

Uniform Random Graph Model

Let m be a function of n, $m(n)$.

The probability space $G'(n,m)$ assigns to each graph on $[n]$ with m edges the probability $1/\binom{N}{m}$ where $N=\binom{n}{2}$

Uniform Random Graph Model

Let m be a function of n, $m(n)$.
The probability space $G'(n,m)$ assigns to each graph on $[n]$ with m edges the probability $1/\binom{N}{m}$ where $N=\binom{n}{2}$

Given a random graph model ($G(n,p)$, $G'(n,m)$, Geometric random graphs, "Scale-free" random graphs, "small world" random graphs, ...)
we wish to understand the probability of satisfying a graph property.

Uniform Random Graph Model

Let m be a function of n, $m(n)$.

The probability space $G'(n,m)$ assigns to each graph on $[n]$ with m edges the probability $1/\binom{N}{m}$ where $N=\binom{n}{2}$

Given a random graph model ($G(n,p)$, $G'(n,m)$, Geometric random graphs, "scale-free" random graphs, "small world" random graphs, ...) we wish to understand the probability of satisfying a graph property.

Defn A graph property $\mathcal{P}$ is a subset of all graphs.

We say $\mathcal{P}$ is convex if for any graphs $G \subseteq H \subseteq K$, whenever G, K satisfy $\mathcal{P}$, then H must also satisfy $\mathcal{P}$.

$\mathcal{P}$ is monotone (increasing) if whenever $G \in \mathcal{P}$ then any graph obtained by adding edges to G also satisfies $\mathcal{P}$.

examples of graph properties:

- "contains K_4", i.e., $\mathcal{P} = \{G : K_4 \subseteq G\}$,
- Connected,
- Hamiltonian,
- k-colorable
- Planar
- contains a vertex of degree 1
- $\vdots$
-

examples of graph properties:

- "contains K_4", i.e., $P = \{G : K_4 \subseteq G\}$, convex & monotone
- Connected, convex & monotone
- Hamiltonian, monotone
- k-colorable not monotone, but monotone dec.
- Planar not monotone, but monotone dec.
- contains a vertex of degree 1 neither mono. inc. nor mono. dec.

$G(n,p)$ & $G'(n,m)$ behave somewhat similarly for convex properties.

<u>**Theorem**</u> If Q is a convex property and $p(1-p)\binom{n}{2} \to \infty$ then almost every G^p from $G(n,p)$ satisfies Q if and only if for every fixed x, almost every G^m from $G'(n,m)$ satisfies Q where $m = \lfloor p\binom{n}{2} + x\left[p(1-p)\binom{n}{2}\right]^{1/2} \rfloor$

$G(n,p)$ & $G'(n,m)$ behave somewhat similarly for convex properties.

<u>Theorem</u> If Q is a convex property and $p(1-p)\binom{n}{2} \to \infty$ then almost every G^p from $G(n,p)$ satisfies Q if and only if for every fixed x, almost every G^m from $G'(n,m)$ satisfies Q where $m = \lfloor p\binom{n}{2} + x[p(1-p)\binom{n}{2}]^{1/2} \rfloor$

We will focus on $G(n,p)$, easier to work with.

Note, when p is fixed $G(n,p)$ gives "dense" graphs.

we must let p get smaller with n to get "sparse" graphs.

Thm [Gilbert 1959] For a fixed $p \in (0,1)$, almost every G^p from $G(n,p)$ is connected.

Thm [Gilbert 1959] For a fixed $p \in (0,1)$, almost every G^p from $G(n,p)$ is connected.

<u>Proof</u> Let q_n = probability G^p is disconnected

Consider $V(G) = S \cup \bar{S}$ with $|S| = k$

then $\mathbb{P}[\,[S,\bar{S}] = \emptyset\,] = (1-p)^{(n-k)k}$

Hence $q_n \leq$

Thm [Gilbert 1959] For a fixed $p \in (0,1)$, almost every G^p from $G(n,p)$ is connected.

Proof Let q_n = probability G^p is disconnected

Consider $V(G) = S \cup \bar{S}$ with $|S| = k$

then $\mathbb{P}[\,[S,\bar{S}] = \emptyset\,] = (1-p)^{(n-k)k}$

Hence $q_n \le \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} (1-p)^{(n-k)k}$

$\le \sum_{k=1}^{\lfloor n/2 \rfloor} \left(n(1-p)^{n/2}\right)^k$, using $\binom{n}{k} \le n^k$, $(1-p)^{n-k} \le (1-p)^{n/2}$

why ?

$\le \frac{\alpha}{1-\alpha}$, where $\alpha = n(1-p)^{n/2}$

Thm [Gilbert 1959] For a fixed $p \in (0,1)$, almost every G^p from $G(n,p)$ is connected.

Proof Let q_n = probability G^p is disconnected

Consider $V(G) = S \cup \bar{S}$ with $|S| = k$

then $\mathbb{P}[\,[S,\bar{S}] = \emptyset\,] = (1-p)^{(n-k)k}$

Hence $q_n \le \displaystyle\sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} (1-p)^{(n-k)k}$

$\le \displaystyle\sum_{k=1}^{\lfloor n/2 \rfloor} \left(n(1-p)^{n/2}\right)^k$, using $\binom{n}{k} \le n^k$, $(1-p)^{n-k} \le (1-p)^{n/2}$

Why?

For large n, $n(1-p)^{n/2} < 1$ & $\sum_k (n(1-p)^{n/2})^k$ is initial portion of a cgt. geom series

$\le \dfrac{\alpha}{1-\alpha}$, where $\alpha = n(1-p)^{n/2}$

$\to 0$ as $n \to \infty$ since $\alpha \to 0$ as $n \to \infty$.

$\blacksquare$

p as a function of n

When does $G(n,p)$ contain a Triangle, K_3?

Evolution of random graph as p increases.....

When exactly do K_3 start appearing in this evolution?

p as a function of n

When does $G(n,p)$ contain a Triangle, K_3?

Evolution of random graph as p increases.....

when exactly do K_3 start appearing in this evolution?

Let $T = \#$ of K_3 in G^p sampled from $G(n,p)$

$$\mathbb{E}[T] = \binom{n}{3} p^3$$

If $p(n) << 1/n$ then $\mathbb{E}[T] \to 0$ as $n \to \infty$

$G(n,p)$ almost always has no K_3 when $p(n) << 1/n$

p as a function of n

When does $G(n,p)$ contain a Triangle, K_3?

Evolution of random graph as p increases.....

when exactly do K_3 start appearing in this evolution?

Let $T = \#$ of K_3 in G^p sampled from $G(n,p)$

$$\mathbb{E}[T] = \binom{n}{3} p^3$$

If $p(n) << 1/n$ then $\mathbb{E}[T] \to 0$ as $n \to \infty$

By Markov inequality, $\mathbb{P}[T=0] \to 1$ as $n \to \infty$

i.e., $\mathbb{P}[G(n,p) \text{ has a } K_3] \to 0$ as $n \to \infty$

$G(n,p)$ almost always has no K_3 when $p(n) << 1/n$

$G(n,p)$ almost always has a K_3 when $p(n)$??

p as a function of n

When does $G(n,p)$ contain a Triangle, K_3?

Evolution of random graph as p increases.....

when exactly do K_3 start appearing in this evolution?

Let $T = \#$ of K_3 in G^p sampled from $G(n,p)$

$$\mathbb{E}[T] = \binom{n}{3} p^3$$

If $p(n) << 1/n$ then $\mathbb{E}[T] \to 0$ as $n \to \infty$

By Markov inequality, $\mathbb{P}[T=0] \to 1$ as $n \to \infty$

i.e., $\mathbb{P}[G(n,p) \text{ has a } K_3] \to 0$ as $n \to \infty$

$G(n,p)$ almost always has no K_3 when $p(n) << 1/n$

$G(n,p)$ almost always has a K_3 when $p(n) >> 1/n$?

$p(n) >> 1/n \Rightarrow \mathbb{E}[T] \longrightarrow \infty$

Why can we not conclude $\mathbb{P}[G(n,p) \text{ has a } K_3] \to 1$ as $n \to \infty$?

<u>Defn</u> We say $t(n)$ is a <u>threshold function</u> for a monotone (inc.) property $\mathcal{P}$ in $G(n,p)$ if

$$\mathbb{P}[G^p \text{ from } G(n,p) \text{ satisfies } \mathcal{P}] \longrightarrow \begin{cases} 0 & \text{if } \frac{p(n)}{t(n)} \to 0 \\ 1 & \text{if } \frac{p(n)}{t(n)} \to \infty \end{cases}$$

Defn We say $t(n)$ is a threshold function for a monotone (inc.) property $\mathcal{P}$ in $G(n,p)$ if

$$\mathbb{P}[G^p \text{ from } G(n,p) \text{ satisfies } \mathcal{P}] \longrightarrow \begin{cases} 0 & \text{if } \frac{p(n)}{t(n)} \to 0 \\ 1 & \text{if } \frac{p(n)}{t(n)} \to \infty \end{cases}$$

$p(n) << t(n)$
$p(n) = o(t(n))$

$p(n) >> t(n)$
$t(n) = o(p(n))$

That is,

when $p(n)$ is sufficiently smaller than $t(n)$,
almost no G^p in $G(n,p)$ satisfies $\mathcal{P}$.

when $p(n)$ is sufficiently larger than $t(n)$
almost all G^p in $G(n,p)$ satisfy $\mathcal{P}$.

Comment
- "Threshold function" is not unique, but all are within constant multiple of others.
- **Theorem** Threshold function exists for every monotone property.

Second Moment Method Show a random variable is "concentrated" near its mean by bounding its variance.

Second Moment Method Show a random variable is "concentrated" near its mean by bounding its variance.

Review

$$\mathrm{Var}[X] = \mathbb{E}[(X-\mathbb{E}(X))^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$$\mathrm{Cov}[X,Y] = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$$

$$\mathrm{Var}\Big[\sum_i X_i\Big] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + \sum_{i\neq j} \mathrm{Cov}[X_i, X_j]$$

$$= \sum_{i,j=1}^{n} \mathrm{Cov}[X_i, X_j]$$

Second Moment Method Show a random variable is "concentrated" near its mean by bounding its variance.

Review

$$\mathrm{Var}[X] = \mathbb{E}[(X-\mathbb{E}(X))^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$$\mathrm{Cov}[X,Y] = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\mathrm{Var}\Big[\sum_i X_i\Big] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + \sum_{i \neq j} \mathrm{Cov}[X_i, X_j]$$

$$= \sum_{i,j=1}^{n} \mathrm{Cov}[X_i, X_j]$$

Chebyshev Inequality Let X be a random variable with finite variance. For all $t > 0$, $\mathbb{P}[|X-\mathbb{E}[X]| > t] \leq \mathrm{Var}[X]/t^2$

Lemma [Second Moment Method] Let $X_1, \ldots, X_n$ be non-neg. r.v.s s.t. $\lim_{n\to\infty} \frac{\mathrm{Var}[X_n]}{(\mathbb{E}[X_n])^2} = 0$ then $\mathbb{P}[X_n > 0] \to 1$ as $n \to \infty$.

Second Moment Method Show a random variable is "concentrated" near its mean by bounding its variance.

Chebyshev Inequality Let X be a random variable with finite variance. For all $t > 0$, $\mathbb{P}[|X - \mathbb{E}[X]| \geq t] \leq \mathrm{Var}[X]/t^2$

Lemma [Second Moment Method] Let $X_1, X_2, \ldots, X_n$ be non-neg. r.v.s s.t. $\lim_{n\to\infty} \frac{\mathrm{Var}[X_n]}{(\mathbb{E}[X_n])^2} = 0$, then $\mathbb{P}[X_n > 0] \to 1$ as $n \to \infty$.

Proof Markov's ineq applied to r.v. $(X-\mathbb{E}[X])^2$ and value t^2 gives $\mathbb{P}[(X-\mathbb{E}[X])^2 \geq t^2] \leq \mathbb{E}[(X-\mathbb{E}[X])^2]/t^2$

i.e., Chebyshev Inequality: $\mathbb{P}[|X - \mathbb{E}[X]| \geq t] \leq \mathrm{Var}[X]/t^2$

Choose $t = \mathbb{E}[X_n]$ & plug into Chebyshev —

$$\mathbb{P}[|X_n - \mathbb{E}[X_n]| \geq \mathbb{E}[X_n]] \leq \mathrm{Var}[X_n]/(\mathbb{E}[X_n])^2 \to 0 \quad \text{as } n \to \infty$$

i.e. $\mathbb{P}[X_n = 0] = \mathbb{P}[X_n \leq 0] \to 0$ as $n \to \infty$ $\blacksquare$

p as a function of n

When does $G(n,p)$ contain a Triangle, K_3?

Evolution of random graph as p increases.....

when exactly do K_3 start appearing in this evolution?

Let $T = \#$ of K_3 in G^p sampled from $G(n,p)$

$$\mathbb{E}[T] = \binom{n}{3} p^3$$

If $p(n) << 1/n$ then $\mathbb{E}[T] \to 0$ as $n \to \infty$

By Markov inequality, $\mathbb{P}[T=0] \to 1$ as $n \to \infty$

i.e., $\mathbb{P}[G(n,p)$ has a $K_3] \to 0$ as $n \to \infty$

$G(n,p)$ almost always has no K_3 when $p(n) << 1/n$

$G(n,p)$ almost always has a K_3 when $p(n) >> 1/n$?

$p(n) >> 1/n \Rightarrow \mathbb{E}[T] \longrightarrow \infty$

Why can we not conclude $\mathbb{P}[G(n,p)$ has a $K_3] \to 1$ as $n \to \infty$?

p as a function of n

When does $G(n,p)$ contain a Triangle, K_3?

Evolution of random graph as p increases.....

when exactly do K_3 start appearing in this evolution?

Let $T = \#$ of K_3 in G^p sampled from $G(n,p)$

$$\mathbb{E}[T] = \binom{n}{3} p^3$$

If $p(n) << 1/n$ then $\mathbb{E}[T] \to 0$ as $n \to \infty$

By Markov inequality, $\mathbb{P}[T=0] \to 1$ as $n \to \infty$

i.e., $\mathbb{P}[G(n,p) \text{ has a } K_3] \to 0$ as $n \to \infty$

$G(n,p)$ almost always has no K_3 when $p(n) << 1/n$

$G(n,p)$ almost always has a K_3 when $p(n) >> 1/n$

i.e. $\mathbb{P}[G(n,p) \text{ has a } K_3] \to 1$ when $p(n) >> 1/n$ ← Use Second Moment Method

Let $T = \# K_3 = \sum_{i=1}^{\binom{n}{3}} T_i$, where T_i is an indicator variable for a triple of vertices to form a K_3

note T is implicitly a function of $n = \#$ vertices & our asymptotics are based on $n \to \infty$

$$\begin{aligned} Var[T] &= \sum_{i,j} Cov[T_i, T_j] \\ &= \sum_i Var[T_i] + \sum_{i \neq j} Cov[T_i, T_j] \end{aligned}$$

Let $T = \# K_3 = \sum_{i=1}^{\binom{n}{3}} T_i$, where T_i is an indicator variable for a triple of vertices to form a K_3

note T is implicitly a function of $n = \#$ vertices & our asymptotics are based on $n \to \infty$

$$\begin{aligned} Var[T] &= \sum_{i,j} Cov[T_i, T_j] \\ &= \sum_{i} Var[T_i] + \sum_{i \neq j} Cov[T_i, T_j] \end{aligned}$$

$$Var[T_i] = \mathbb{E}[T_i^2] - (\mathbb{E}[T_i])^2 \leq \mathbb{E}[T_i^2] = \mathbb{E}[T_i] = \mathbb{P}[T_i = 1] = ?$$

$$\begin{aligned} Cov[T_i, T_j] &= \mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j] \\ &\leq \mathbb{E}[T_i T_j] = \mathbb{P}[T_i T_j = 1] \\ &= \mathbb{P}[T_i = 1 \wedge T_j = 1] \leq ? \end{aligned}$$

Let $T = \# K_3 = \sum_{i=1}^{\binom{n}{3}} T_i$, where T_i is an indicator variable for a triple of vertices to form a K_3

note T is implicitly a function of $n = \#$ vertices & our asymptotics are based on $n \to \infty$

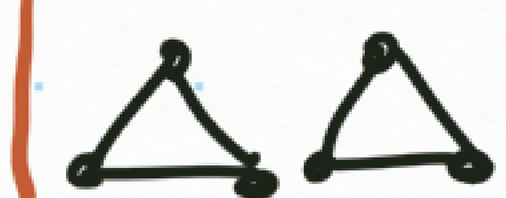
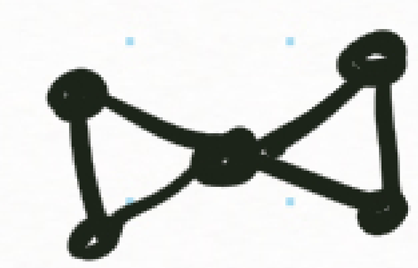
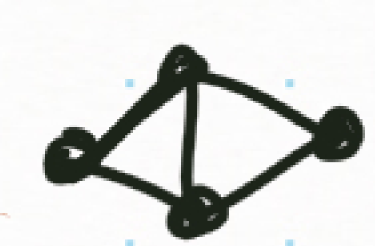
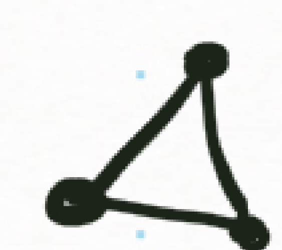
$$\mathrm{Var}[T] = \sum_{i,j} \mathrm{Cov}[T_i, T_j]$$
$$= \sum_i \mathrm{Var}[T_i] + \sum_{i \neq j} \mathrm{Cov}[T_i, T_j]$$

$$\mathrm{Var}[T_i] = \mathbb{E}[T_i^2] - (\mathbb{E}[T_i])^2 \leq \mathbb{E}[T_i^2] = \mathbb{E}[T_i] = \mathbb{P}[T_i = 1] = p^3$$

$$\mathrm{Cov}[T_i, T_j] = \mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$$
$$\leq \mathbb{E}[T_i T_j] = \mathbb{P}[T_i T_j = 1]$$
$$= \mathbb{P}[T_i = 1 \wedge T_j = 1] \leq p^5$$

Triples corresponding to i & j

How much do they overlap?

	$\mathbb{E}[T_i T_j]$	$\mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$
	p^6	$p^6 - p^6 = 0$
	p^6	$p^6 - p^6 = 0$
	p^5	$p^5 - p^6$
	p^3	$p^3 - p^6$

Let $T = \# K_3 = \sum_{i=1}^{\binom{n}{3}} T_i$, where T_i is an indicator variable for a triple of vertices to form a K_3

note T is implicitly a function of n = # vertices & our asymptotics are based on $n \to \infty$

$$\mathrm{Var}[T] = \sum_{i,j} \mathrm{Cov}[T_i, T_j]$$
$$= \sum_i \mathrm{Var}[T_i] + \sum_{i \neq j} \mathrm{Cov}[T_i, T_j]$$

$$\mathrm{Var}[T_i] = \mathbb{E}[T_i^2] - (\mathbb{E}[T_i])^2 \leq \mathbb{E}[T_i^2] = \mathbb{E}[T_i] = \mathbb{P}[T_i = 1] = p^3$$

$$\mathrm{Cov}[T_i, T_j] = \mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$$
$$\leq \mathbb{E}[T_i T_j] = \mathbb{P}[T_i T_j = 1]$$
$$= \mathbb{P}[T_i = 1 \wedge T_j = 1] \leq p^5$$

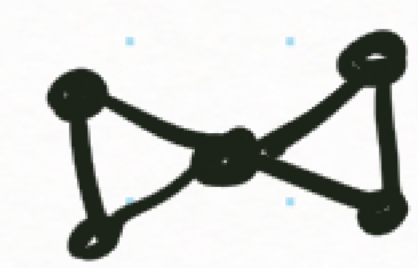
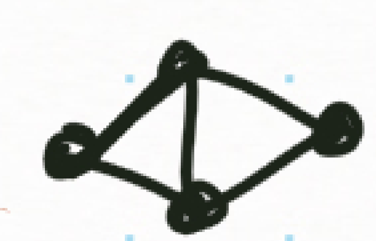
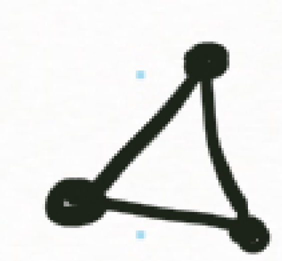
$$\therefore \mathrm{Var}[T] \leq \binom{n}{3} p^3 + 2 \binom{n}{4} \binom{4}{2} p^5$$

? ? ?

$$\leq n^3 p^3 + n^4 p^5$$

Triples corresponding to i & j

How much do they overlap?

	$\mathbb{E}[T_i T_j]$	$\mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$
	p^6	$p^6 - p^6 = 0$
	p^6	$p^6 - p^6 = 0$
	p^5	$p^5 - p^6$
	p^3	$p^3 - p^6$

Let $T = \# K_3 = \sum_{i=1}^{\binom{n}{3}} T_i$, where T_i is an indicator variable for a triple of vertices to form a K_3

note T is implicitly a function of $n = \#$ vertices & our asymptotics are based on $n \to \infty$

$$\mathrm{Var}[T] = \sum_{i,j} \mathrm{Cov}[T_i, T_j]$$
$$= \sum_i \mathrm{Var}[T_i] + \sum_{i \neq j} \mathrm{Cov}[T_i, T_j]$$

$$\mathrm{Var}[T_i] = \mathbb{E}[T_i^2] - (\mathbb{E}[T_i])^2 \leq \mathbb{E}[T_i^2] = \mathbb{E}[T_i] = \mathbb{P}[T_i = 1] = p^3$$

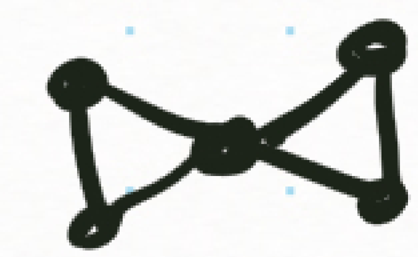
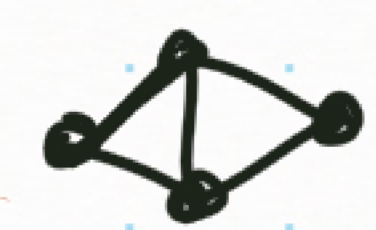
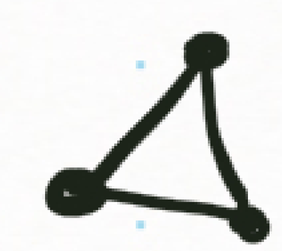
$$\mathrm{Cov}[T_i, T_j] = \mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$$
$$\leq \mathbb{E}[T_i T_j] = \mathbb{P}[T_i T_j = 1]$$
$$= \mathbb{P}[T_i = 1 \wedge T_j = 1] \leq p^5$$

$$\therefore \mathrm{Var}[T] \leq \binom{n}{3} p^3 + \underset{i \neq j}{2} \underset{4 \text{ vertices}}{\binom{n}{4}} \underset{\text{2 vertices in common}}{\binom{4}{2}} p^5$$

$$\leq n^3 p^3 + n^4 p^5$$

Triples corresponding to i & j

How much do they overlap?

	$\mathbb{E}[T_i T_j]$	$\mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$
	p^6	$p^6 - p^6 = 0$
	p^6	$p^6 - p^6 = 0$
	p^5	$p^5 - p^6$
	p^3	$p^3 - p^6$

Let $T = \# K_3 = \sum_{i=1}^{\binom{n}{3}} T_i$, where T_i is an indicator variable for a triple of vertices to form a K_3

note T is implicitly a function of $n = \#$ vertices & our asymptotics are based on $n \to \infty$

$$\mathrm{Var}[T] = \sum_{i,j} \mathrm{Cov}[T_i, T_j]$$
$$= \sum_i \mathrm{Var}[T_i] + \sum_{i \neq j} \mathrm{Cov}[T_i, T_j]$$

$$\mathrm{Var}[T_i] = \mathbb{E}[T_i^2] - (\mathbb{E}[T_i])^2 \leq \mathbb{E}[T_i^2] = \mathbb{E}[T_i] = \mathbb{P}[T_i = 1] = p^3$$

$$\mathrm{Cov}[T_i, T_j] = \mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$$
$$\leq \mathbb{E}[T_i T_j] = \mathbb{P}[T_i T_j = 1]$$
$$= \mathbb{P}[T_i = 1 \wedge T_j = 1] \leq p^5$$

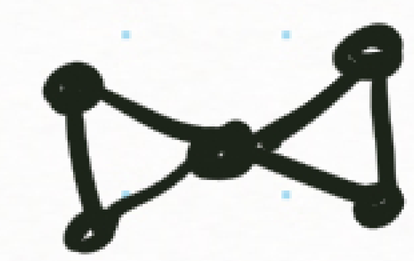
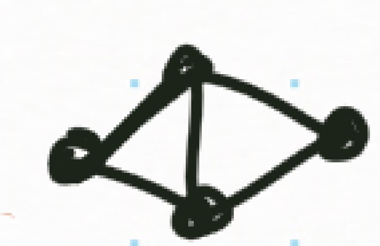
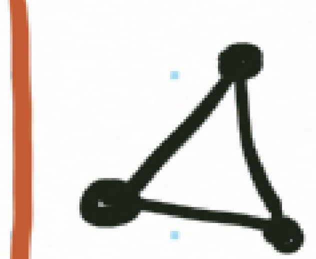
$$\therefore \mathrm{Var}[T] \leq \binom{n}{3} p^3 + 2 \binom{n}{4} \binom{4}{2} p^5$$

? ? ?

$$\leq n^3 p^3 + n^4 p^5$$

Triples corresponding to i & j

How much do they overlap?

	$\mathbb{E}[T_i T_j]$	$\mathbb{E}[T_i T_j] - \mathbb{E}[T_i]\mathbb{E}[T_j]$
	p^6	$p^6 - p^6 = 0$
	p^6	$p^6 - p^6 = 0$
	p^5	$p^5 - p^6$
	p^3	$p^3 - p^6$

Let $T = \# K_3 = \sum_{i=1}^{\binom{n}{3}} T_i$, where T_i is an indicator variable for a triple of vertices to form a K_3

note T is implicitly a function of $n = \#$ vertices & our asymptotics are based on $n \to \infty$

$$\begin{aligned} Var[T] &= \sum_{i,j} Cov[T_i, T_j] \\ &= \sum_i Var[T_i] + \sum_{i \neq j} Cov[T_i, T_j] \\ &\leq n^3 p^3 + n^4 p^5 \end{aligned}$$

So, $$\frac{Var[T]}{(E[T])^2} \leq \frac{n^3p^3 + n^4p^5}{(\binom{n}{3} p^3)^2} \leq O\left(\frac{n^3p^3 + n^4p^5}{(n^3p^3)^2}\right)$$

$\left[\text{Recall } \binom{n}{3} \geq \left(\frac{n}{3}\right)^3\right]$

$$= O\left(\frac{1}{n^3p^3} + \frac{n^4p^5}{n^6p^6}\right)$$

$$= O\left(\frac{1}{n^3p^3} + \frac{1}{n^2p}\right) \to 0 \text{ as } n \to \infty$$

when $p(n) >> \frac{1}{n}$

i.e., $pn >> 1$

i.e., $pn \to \infty$

$\therefore$ By Second Moment Method, ...

Theorem $t(n) = \frac{1}{n}$ is a threshold function for the event that K_3 is a subgraph of $G(n,p)$

Comments

- It is known that $T = \# K_3$ in $G(n,p)$ satisfies an asymptotic central limit Thm., $(T - \mathbb{E}[T])/\sigma_T \to N(0,1)$ in distribution when $p(n) \gg 1/n$

When $p(n) \ll 1/n$, T is asymtotically Poisson distribution

What is the threshold function for K_4 as subgraph of $G(n,p)$?
Try it by mimicing our proof for K_3.

Defn Let H be a graph with v vertices and e edges
we define density of H as $\rho(H) = e/v$ (half of average degree)

we call H balanced if no subgraph of H has density strictly greater than $\rho(H)$.

e.g. Trees, cliques, k-trees, k-reg connected, hypercubes, etc.

<u>Defn</u> Let H be a graph with v vertices and e edges
we define <u>density of H</u> as $\rho(H) = e/v$ (half of average degree)
we call H <u>balanced</u> if no subgraph of H has density
strictly greater than $\rho(H)$.
e.g. Trees, cliques, k-trees, k-reg connected, hypercubes, etc.

<u>Theorem</u> Let H be a balanced graph with density ρ.
Then, $t(n) = n^{-1/\rho}$ is a threshold function for the
event that H is a subgraph of $G(n,p)$.

What does this thm tell us about $H = K_3$?

$\rho(K_3) = \frac{3}{3} = 1$, so $t(n) = n^{-1/1} = \frac{1}{n}$

Theorem Let H be a balanced graph with density ρ.
Then $t(n) = n^{-1/\rho}$ is a threshold function for "H is a subgraph of $G(n,p)$".

Proof Let H have v vertices and e edges, so $\rho = e/v$
Denote $V(H)$ as $\{a_1, a_2, \dots, a_v\}$
For any ordered v-tuple, $\beta = (b_1, b_2, \dots, b_v)$ of distinct vertices $b_1, b_2, \dots, b_v$ in $V(G(n,p))$,
let A_β be the event that "$G(n,p)$ contains an appropriately ordered copy of H on $(b_1, b_2, \dots, b_v)$"
that is, A_β occurs if $b_i b_j \in E(G(n,p))$ whenever $a_i a_j \in E(H)$.

Theorem Let H be a balanced graph with density ρ.
Then $t(n) = n^{-1/\rho}$ is a threshold function for "H is a subgraph of $G(n,p)$".

Proof Let H have v vertices and e edges, so $\rho = e/v$

Denote $V(H)$ as $\{a_1, a_2, \dots, a_v\}$

For any ordered v-tuple, $\beta = (b_1, b_2, \dots, b_v)$ of distinct vertices $b_1, b_2, \dots, b_v$ in $V(G(n,p))$,

let A_β be the event that "$G(n,p)$ contains an appropriately ordered copy of H on $(b_1, b_2, \dots, b_v)$"

that is, A_β occurs if $b_i b_j \in E(G(n,p))$ whenever $a_i a_j \in E(H)$.

Let X_β = indicator r.v. for event A_β

Let $X = \sum_\beta X_\beta$

\# of ordered copies of H, which could be larger than \# copies of H

However, $X = 0$ & $X > 0$ are equivalent to absence & appearance of H in $G(n,p)$

$$
\begin{aligned}
\mathbb{E}[X] &= \sum \mathbb{E}[X_\beta] \\
&= \sum \mathbb{P}[A_\beta] \\
&= \sum p^e = \Theta(n^v p^e)
\end{aligned}
$$

So, $\mathbb{E}[X] = \Theta(n^v p^e)$

If $\underline{p(n) << n^{-1/\rho}} = n^{-v/e}$ then $p^e n^v << 1$, ie., $p^e n^v \to 0$

so, $\lim_{n\to\infty} \mathbb{E}[X] = 0$

& by Markov, $\mathbb{P}[X=0] \to 1$ as $n \to \infty$.

So, $\mathbb{E}[X] = \Theta(n^v p^e)$

If $\underline{p(n) << n^{-1/g}} = n^{-v/e}$ then $p^e n^v << 1$, ie., $p^e n^v \to 0$

so, $\lim_{n\to\infty} \mathbb{E}[X] = 0$

& by Markov, $\mathbb{P}[X=0] \to 1$ as $n \to \infty$.

Now, we assume $\underline{p(n) >> n^{-1/g}}$

$$\mathrm{Var}[X] = \sum_{\beta} \mathrm{Var}[X_\beta] + \sum_{\alpha \neq \beta} \mathrm{Cov}[X_\alpha, X_\beta]$$

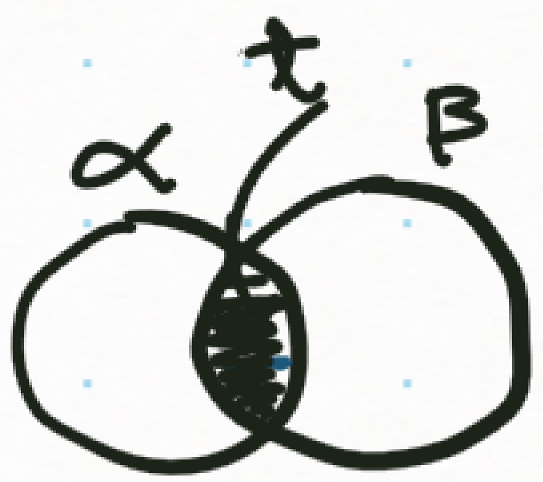
So, $\mathbb{E}[X] = \Theta(n^v p^e)$

If $\underline{p(n) << n^{-1/\rho}} = n^{-v/e}$ then $p^e n^v << 1$, ie., $p^e n^v \to 0$

so, $\lim_{n\to\infty} \mathbb{E}[X] = 0$

& by Markov, $\mathbb{P}[X=0] \to 1$ as $n \to \infty$.

Now, we assume $\underline{p(n) >> n^{-1/\rho}}$



$$Var[X] = \sum_{\beta} Var[X_\beta] + \sum_{\alpha \neq \beta} Cov[X_\alpha, X_\beta]$$

Let α and β share $t \geq 2$ vertices, then the two copies of H on α & β have at most ? edges in common

$\therefore Cov[X_\alpha, X_\beta] \leq \mathbb{E}[X_\alpha X_\beta] \leq p^{?}$, since their union has at least ? edges

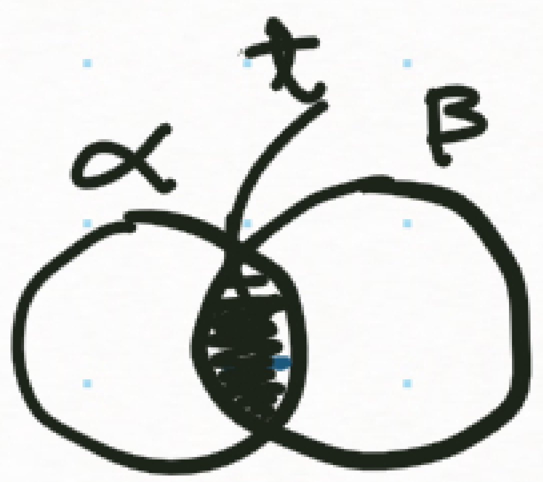
So, $\mathbb{E}[X] = \Theta(n^v p^e)$

If $\underline{p(n) << n^{-1/\varrho} = n^{-v/e}}$ then $p^e n^v << 1$, ie., $p^e n^v \to 0$

so, $\lim_{n\to\infty} \mathbb{E}[X] = 0$

& by Markov, $\mathbb{P}[X=0] \to 1$ as $n \to \infty$.

Now, we assume $\underline{p(n) >> n^{-1/\varrho}}$



$$\mathrm{Var}[X] = \sum_{\beta} \mathrm{Var}[X_\beta] + \sum_{\alpha \neq \beta} \mathrm{Cov}[X_\alpha, X_\beta]$$

Let α and β share $t \geq 2$ vertices, then the two copies of H on α & β have at most $\underline{t\varrho}$ edges in common

$\therefore \mathrm{Cov}[X_\alpha, X_\beta] \leq \mathbb{E}[X_\alpha X_\beta] \leq p^{2e - t\varrho}$, since their union has at least $\underline{2e - t\varrho}$ edges

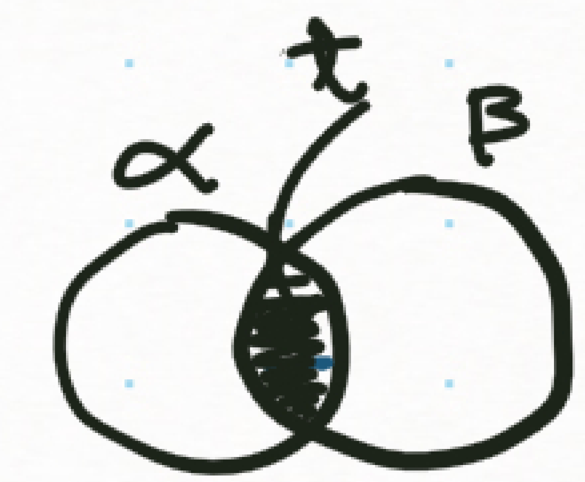
So, $\mathbb{E}[X] = \Theta(n^{v} p^{e})$

If $\underline{p(n) << n^{-1/\rho}} = n^{-v/e}$ then $p^{e} n^{v} << 1$, ie., $p^{e} n^{v} \to 0$

so, $\lim\limits_{n\to\infty} \mathbb{E}[X] = 0$

& by Markov, $\mathbb{P}[X=0] \to 1$ as $n \to \infty$.

Now, we assume $\underline{p(n) >> n^{-1/\rho}}$



$$\mathrm{Var}[X] = \sum_{\beta} \mathrm{Var}[X_\beta] + \sum_{\alpha \neq \beta} \mathrm{Cov}[X_\alpha, X_\beta]$$

Let α and β share $t \geq 2$ vertices, then the two copies of H on α & β have at most $t\rho$ edges in common

$\therefore \mathrm{Cov}[X_\alpha, X_\beta] \leq \mathbb{E}[X_\alpha X_\beta] \leq p^{2e - t\rho}$, since their union has at least $2e - t\rho$ edges

The number of pairs α, β sharing t vertices is $\Theta(n^{2v-t})$

(we can choose $2v-t$ vertices in $\binom{n}{2v-t}$ ways and there are only constant many ways to choose α & β from them since H is fixed so $v = |\alpha| = |\beta|$ is fixed)

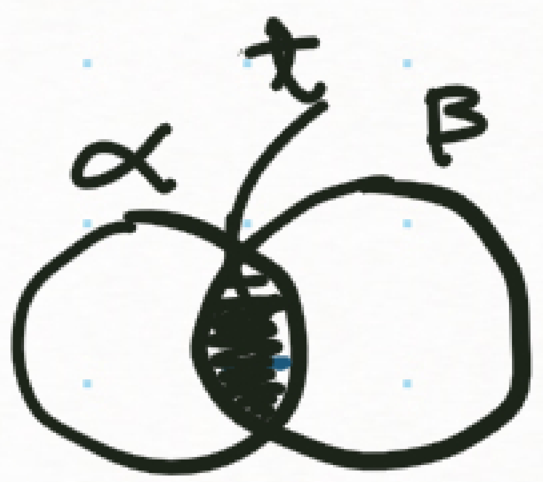
So, $\mathbb{E}[X] = \Theta(n^v p^e)$

If $\underline{p(n) << n^{-1/\rho}} = n^{-v/e}$ then $p^e n^v << 1$, ie, $p^e n^v \to 0$

so, $\lim_{n\to\infty} \mathbb{E}[X] = 0$

& by Markov, $\mathbb{P}[X=0] \to 1$ as $n \to \infty$.

Now, we assume $\underline{p(n) >> n^{-1/\rho}}$



$$\mathrm{Var}[X] = \sum_\beta \mathrm{Var}[X_\beta] + \sum_{\alpha \neq \beta} \mathrm{Cov}[X_\alpha, X_\beta]$$

Let α and β share $t \geq 2$ vertices, then the two copies of H on α & β have at most $t\rho$ edges in common

$\therefore \mathrm{Cov}[X_\alpha, X_\beta] \leq \mathbb{E}[X_\alpha X_\beta] \leq p^{2e-t\rho}$, since their union has at least $2e-t\rho$ edges

The number of pairs α, β sharing t vertices is $\Theta(n^{2v-t})$

$$\therefore \sum_{|\alpha\cap\beta|=t} \mathrm{Cov}[X_\alpha, X_\beta] \leq \Theta(n^{2v-t} p^{2e-t\rho}) = \Theta((n^v p^e)^{2-t/v})$$

So $\mathrm{Var}[X] \leq O\left(\sum_{t=2}^{v} (n^v p^e)^{2-t/v}\right)$

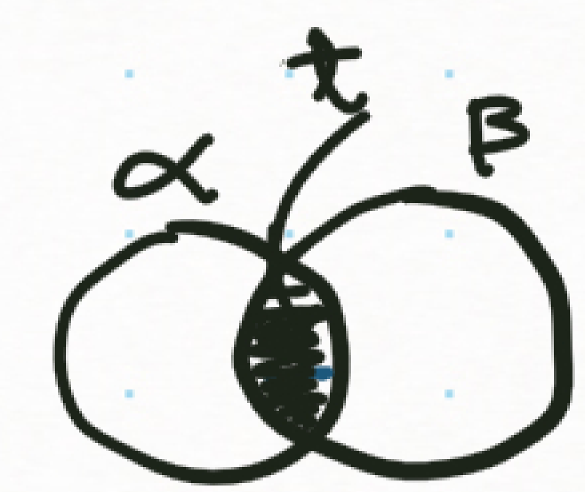
So, $\mathbb{E}[X] = \Theta(n^{v} p^{e})$

If $\underline{p(n) << n^{-1/\varrho}} = n^{-v/e}$ then $p^{e} n^{v} << 1$, ie., $p^{e} n^{v} \to 0$

so, $\lim_{n\to\infty} \mathbb{E}[X] = 0$

& by Markov, $\mathbb{P}[X=0] \to 1$ as $n \to \infty$.

Now, we assume $\underline{p(n) >> n^{-1/\varrho}}$



$$\mathrm{Var}[X] = \sum_{\beta} \mathrm{Var}[X_\beta] + \sum_{\alpha \neq \beta} \mathrm{Cov}[X_\alpha, X_\beta]$$

Let α and β share $t \geq 2$ vertices, then the two copies of H on α & β have at most $t\varrho$ edges in common

$\therefore \mathrm{Cov}[X_\alpha, X_\beta] \leq \mathbb{E}[X_\alpha X_\beta] \leq p^{2e - t\varrho}$, since their union has at least $2e - t\varrho$ edges

The number of pairs α, β sharing t vertices is $\Theta(n^{2v-t})$

$$\therefore \sum_{|\alpha \cap \beta| = t} \mathrm{Cov}[X_\alpha, X_\beta] \leq \Theta(n^{2v-t} p^{2e - t\varrho}) = \Theta((n^{v} p^{e})^{2 - t/v})$$

So $\mathrm{Var}[X] \leq \Theta\left(\sum_{t=2}^{v} (n^{v} p^{e})^{2-t/v}\right)$

$$\therefore \lim_{n\to\infty} \frac{\mathrm{Var}[X]}{(\mathbb{E}[X])^2} = \lim_{n\to\infty} \Theta\left(\sum_{t=2}^{v} (n^{v} p^{e})^{-t/v}\right) = 0, \text{ since } \lim_{n\to\infty} n^{v} p^{e} = \infty \quad \blacksquare$$

<u>Proposition</u> If H is not balanced then $t(n) = n^{-v/e}$ is not a threshold function for appearance of H.

Proposition If H is not balanced then $t(n) = n^{-v/e}$ is not a threshold function for appearance of H.

Proof Pick a subgraph H' of H s.t. $\frac{e(H')}{v(H')} > \frac{e}{v}$

Let $v/e > \gamma > v'/e'$ & $p(n) = n^{-\gamma}$

Since $p(n) = n^{-\gamma} >> t(n) = n^{-v/e}$, there must be a copy of H w.h.p.

Since $p(n) = n^{-\gamma} << n^{-v'/e'}$, the expected number of copies of H' must tend to 0 as $n \to \infty$

which implies w.h.p. there are no copies of H'

which means there are no copies of H.

contradiction ☒

So, what should be the threshold function for the appearance of any graph H?

Proposition If H is not balanced then $t(n) = n^{-v/e}$ is not a threshold function for appearance of H.

Proof Pick a subgraph H' of H s.t. $\frac{e(H')}{v(H')} > \frac{e}{v}$ [Let H' be a subgraph of max density]

Let $v/e > \gamma > v'/e'$ & $p(n) = n^{-\gamma}$

Since $p(n) = n^{-\gamma} >> t(n) = n^{-v/e}$, there must be a copy of H w.h.p.

Since $p(n) = n^{-\gamma} << n^{-v'/e'}$, the expected number of copies of H' must tend to 0 as $n \to \infty$

which implies w.h.p. there are no copies of H'

which means there are no copies of H.

contradiction $\boxtimes$

Theorem For any graph H, let max density of H be $\rho_{max} = \max\limits_{H' \subseteq H} \left(\frac{e(H')}{v(H')}\right)$

Then, $t(n) = n^{-1/\rho_{max}}$ is a threshold function for appearance of H.

Much more is known about Threshold functions in $G(n,p)$.

Theorem [Bollobás & Thomason 1987] Every non-trivial monotone graph property has a threshold function.

Sharp Thresholds

e.g. ① $\mathbb{P}[G(n, \frac{c_n}{n})$ contains a $K_3] \to \begin{cases} 0 & \text{if } c_n \to -\infty \\ 1-e^{-c^3/6} & \text{if } c_n \to c, \text{ constant} \\ 1 & \text{if } c_n \to \infty \end{cases}$

② $\mathbb{P}[G(n, \frac{\log n + c_n}{n})$ is connected$] \to \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-e^{-c}} & \text{if } c_n \to c \\ 1 & \text{if } c_n \to \infty \end{cases}$

③ $\mathbb{P}[G(n, \frac{\log n + \log\log n + c_n}{n})] \to \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-e^{-c}} & \text{if } c_n \to c \\ 1 & \text{if } c_n \to \infty \end{cases}$