Random Graphs
Lecture IV: May 24

1 Random Graphs \( G(n, p) \)

Started in 1960 with "The Evolution of the Random Graph" by Paul Erdős and Alfred Rényi. \( n \) vertices, each pair adjacent with independent probability \( p \). Critically, we generally have \( p \) as a function of \( n \).

1.1 Property A: \( \exists K_4 \)

For \( |s| = 4 \) let \( A_s := S \) is a \( K_4 \). Let \( X_s := \) indicator r.v. for \( A_s \) with values \( \{0,1\} \).

\[
X = \sum_{|s|=4} X_s
\]

\[
E[X_s] = Pr[A_s] = p^6 \Rightarrow E[X] = \left(\frac{n}{4}\right) p^6 \sim \frac{n^4 p^6}{24}
\]

We are lead to consider \( p \approx o(n^{-\frac{2}{3}}) \)

**Theorem 1.1**

(a) If \( p(n) = o(n^{-\frac{2}{3}}) \) then \( \lim_{n\to\infty} Pr[A] = 0 \)

(b) If \( p(n) \gg n^{-\frac{2}{3}} \) then \( \lim_{n\to\infty} Pr[A] = 1 \)

We say that \( p_0(n) \) is a threshold function for event \( A \), if:

(i) \( p \ll p_0(n) \Rightarrow Pr[A] \to 0 \)

(ii) \( p \gg p_0(n) \Rightarrow Pr[A] \to 1 \)

**Remark:** If \( p_0(n) \) is a threshold function then \( c p_0(n) \) is a threshold function as well.

1.2 Example:

Planarity has \( \frac{\log n}{n} = p_0(n) \). "Sharp threshold" = \( \frac{(1+\epsilon)\log n}{n} \) is not planar, but \( \frac{(1-\epsilon)\log n}{n} \) is planar.
1.3 FACT:
If $E[X] \to 0$ then $Pr[X = 0] \to 1$. $X$ is a counting r.v. (i.e. 0,1,2,3,...)
Proof:
(i) $p \ll n^{-\frac{2}{3}} \Rightarrow E[X] \to 0 \Rightarrow Pr[A] \to 0$
(ii) $p \gg n^{-\frac{2}{3}} \Rightarrow E[X] \to \infty \not\Rightarrow Pr[X \neq 0] \to 1$
End of Proof

2 Example
Consider a $K_4$ with an extra vertex joined to one of the four vertices. So it has 5 vertices and 7 edges. What is the threshold for its appearance. Let $X$ be the number of these “fish”. $E[X] = \Theta(n^5 p^7)$. If $p \gg n^{-\frac{2}{3}}$ then we don’t have $Pr[B] \to 1$. If $n^{-\frac{2}{3}} \ll p \ll n^{-\frac{2}{3}}$ we only have “fish” without the tail, as we don’t have the $K_4$ with probability 1.

2.1 FACT:
If $E[X] \to \infty$ and $Var[X] = o(E[X]^2)$ then:
(i) $Pr[X = 0] \to 0$
(ii) $X \sim E[X]$ (asymptotic, almost surely)
Proof: Recall Chebychev: $Pr[|X - \mu| \geq \lambda \sigma] \leq \lambda^2$
(i) $Pr[X = 0] \leq Pr[|X - \mu| \geq (\frac{\mu}{\sigma}) \sigma] \leq \frac{\sigma^2}{\mu^2}$
but $\frac{\sigma^2}{\mu^2} \to 0$ so $Pr[X = 0] \to 0$.
(ii) $Pr[|X - \mu| \geq \epsilon \mu] \leq \frac{\sigma^2}{\epsilon^2 \mu^2}$
for fixed $\epsilon$ we get that $Pr[|X - \mu| \geq \epsilon \mu] \to 0$. So $E[X]$ is close to $X$. End of Proof. Let
$$X = \sum X_S$$
and
$$Var[X] = \sum Var[X_S] + \sum_{S \neq T} Cov[X_S, X_T] \leq E[X] + \Delta$$
where
$$\Delta := \sum Pr[A_S \cap A_T]$$
and the sum is over $S \sim T$ but $S \neq T$. Suppose we have a relation $\sim$ on an index set such that $\neg(S \sim T) \Rightarrow A_S, A_T$ independent. In our case: $S \sim T := |S \cap T| \geq 2.$
If $E[X] \to \infty$ and $\Delta = o(E[X]^2)$ then $X > 0$ almost surely.
In our case:

\[ \Delta = \sum \Pr[A_S \land A_T] = \sum_{i=2}^{3} \sum_{|S \cap T|=i} \Pr[A_S \land A_T] \]

For \( i = 2 \)

\[ O(n^6p^{11}) = O\left(\frac{(n^4p^6)^2}{n^2p}\right) = O(E[X]^2) \]

For \( i = 3 \)

\[ O(n^5p^9) = O\left(\frac{(n^4p^6)^2}{n^3p^3}\right) = O(E[X]^2) \]

**Definition:** Let \( H \) have \( v \) vertices and \( e \) edges. We call \( H \) strictly balanced if every subgraph \( H' \) with \( v' \) vertices, \( e' \) edges is such that \( \frac{e'}{v'} < \frac{e}{v} \). It is balanced when \( \frac{e'}{v'} \leq \frac{e}{v} \).

The “fish” is not balanced, because \( \frac{6}{4} > \frac{7}{5} \).

**Theorem 2.1** Let \( H \) be balanced. Then \( p = n^{-\frac{\xi}{2}} \) is a threshold function for \( A := \exists H \) (this is actually an iff statement)

Proof: For each \( |S| = v \) let \( A_S := S \) contains \( H \) and \( X_S := \) indicator r.v. of \( A_S \). Then

\[ E[X_S] = \Theta(p^v) \Rightarrow E[X] = \Theta(n^v p^v) \]

\[ \Delta = \sum_{i=2}^{v-1} \sum_{|S \cap T|=i} \Pr[A_S \land A_T] \]

We need for each \( i \) that \( \sum_i = o(\mu^2) \).

Given \( i \) the number of \( (S, T) = O(n^{2v-i}) \) and \( \Pr[A_S \land A_T] = O(p^{2v-i}) \) where \( e_i \) is the maximum number of edges of \( H \) in \( i \) points.

Since \( H \) is balanced \( \frac{e}{v} \leq \frac{e_i}{v_i} \). So given \( i \):

\[ \sum = O(n^{2v-i}p^{2v-e_i}) = O\left(\frac{(n^vp^v)^2}{n^ip^{e_i}}\right) = o(1) \]

End of Proof.

**2.2 Example:**

The number of Hamiltonian cycles: \( \frac{n!}{2^n} \) with probability \( p^n \) so the expected number is \( \frac{n!p^n}{2^n} \) which reaches 1 at around \( p = \frac{e}{n} \) but this is not the threshold function as at that point there are lots of isolated vertices. The actual threshold function is \( \frac{\ln n}{n} \) which took many years to show.

**2.3 Special Case:**

Suppose the events \( A_S \) are symmetric. Then

\[ \Delta = \sum \Pr[A_S \land A_T] = \sum_{S} \sum \Pr[A_S] \Pr[A_T|A_S] = E[X]\Delta^* \]

where

\[ \Delta^* = \sum \Pr[A_T|A_S] \]

for any given \( S \).
2.4 Claim:

$$\Delta = o(E[X]^2) \leftrightarrow \Delta^* = o(E[X])$$

3 Connectivity

Perhaps the most famous of the Erdős-Rényi results.

**Theorem:** If

$$p = \frac{\ln n}{n} + \frac{c}{n}$$

then $G(n, p)$ is connected with limiting probability $e^{-e^{-c}}$.

Why? What stops $G$ from being connected in this range is isolated points. Let WEIRD be the event that $G$ has no isolated points but is not connected. We claim WEIRD (for $p$ in this range) has limiting probability zero. Its probability is at most the sum over $2 \leq k \leq n/2$ of the probability of having a $k$ point component. For fixed $k$ we say there are $\Theta(n^k)$ possible components, probability $O(n^{1-k})$ that those $k$ points are connected and probability $(1-p)^{k(n-k)} = \Theta(n^{-k})$ that there are no edges crossing the border so this is $o(1)$. Larger $k$ (and summing over $k$) is moderately technically challenging and we’ll skip it.

Let $X_v$ be the indicator for $v$ being isolated and $Z = \sum X_v$ the number of isolated points. So $E[X_v] = (1-p)^n \sim e^{-c}n^{-1}$ and $E[Z] \sim e^{-c}$. Set $\mu := E[Z]$. What we want is that $Z$ is asymptotically Poisson, in particular that $\Pr[Z = 0] \sim e^{-\mu}$. For this it suffices that each fixed $r$-th factorial moment $E[(\sum X_v)^r] \sim \frac{\mu^r}{r!}$ where This is like inclusion-exclusion. We have

$$\sum E[X_{v_1} \cdots X_{v_r}] = \binom{n}{r} E[X_1 \cdots X_r]$$

The probability that a given $r$ points are connected is $(1-p)^r(n-1)(1-p)^{-\binom{r}{2}}$ but since $r$ is fixed this is asymptotic to the result if they were independent and so $E[(\sum X_v)^r]$ is asymptotically what we want.

Or: Let $A_v$ be the indicator of $v$ being isolated and we want the probability that no $A_v$ holds. By Inclusion-Exclusion it is $1 - S_1 + S_2 - S_3 + \ldots$ where $S_r$ is the sum over $r$-sets of $r$ points being isolated. So $S_r \to \frac{n^r}{r!}$.

To interchange the limits we use the *Bonferroni Inequalities* which say that Inclusion-Exclusion alternately over- and under- estimates the final answer. For any fixed $r$ we can bound between the sum up to $r$ and the sum up to $r+1$ and with $r$ fixed we interchange the limits. Then the sum gets more and more sandwiched as we let $r \to \infty$. 
