The Probabilistic Method

Week 9: Random Graphs

Joshua Brody
CS49/Math59
Fall 2015

Reading Quiz

What is a **G(n,p)**?

(A) **a probability distribution**

(B) **a random variable**

(C) **a random graph**

(D) **a probability space**

(E) **multiple answers correct**

Reading Quiz

What is a **G(n,p)**?

(A) a probability distribution

(B) a random variable

(C) a random graph

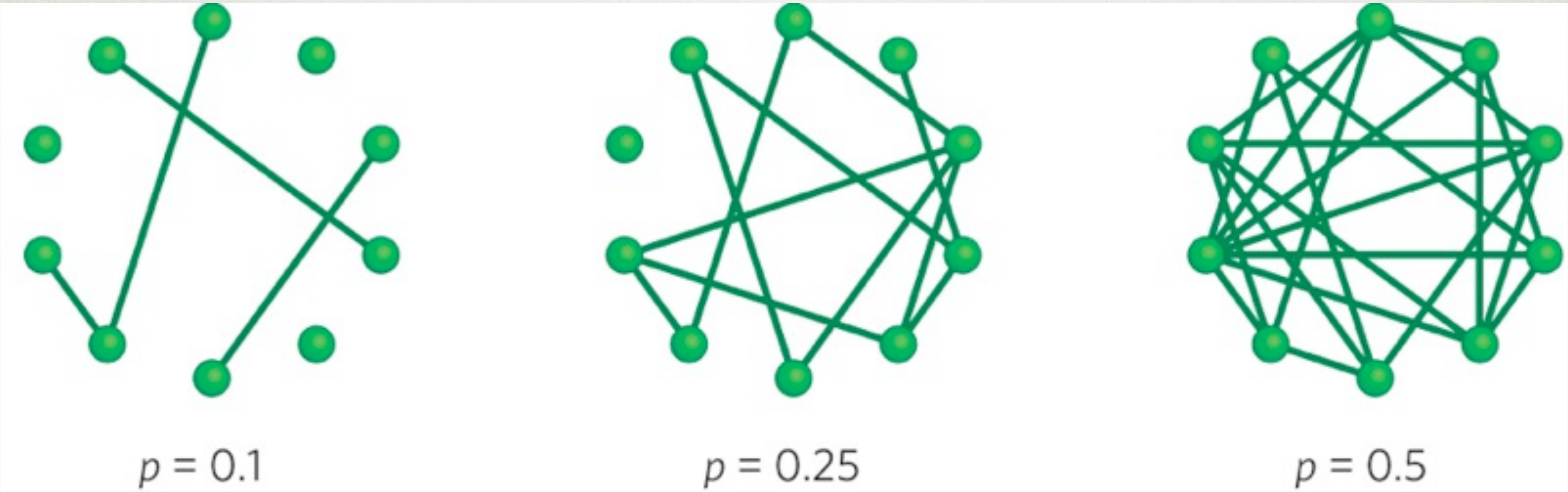
(D) a probability space

(E) multiple answers correct

Random Graphs

[Erdős-Rényi 60]

G ~ G(n,p) : random graph on **n** vertices **V = {1, ..., n}**
each edge **(i,j) ∈ E** independently with prob. **p**



G(n,p) : probability distribution
G : random variable

Clicker Question

When are A_S and A_T ***not*** independent?

(A) S, T share at least one vertex

(B) S, T share at least one edge

(C) S, T share at least two vertices

(D) (A) and (B)

(E) (B) and (C)

Clicker Question

When are A_S and A_T ***not*** independent?

(A) **S, T share at least one vertex**

(B) **S, T share at least one edge**

(C) **S, T share at least two vertices**

(D) **(A) and (B)**

(E) **(B) and (C)**

Clicker Question

If $|S \cap T| = 2$, what is $\Pr[A_T | A_S]$?

(A) p^6

(B) p^5

(C) p^4

(D) p^3

(E) 1

Clicker Question

If $|S \cap T| = 2$, what is $Pr[A_T|A_S]$?

(A) p^6

(B) p^5

(C) p^4

(D) p^3

(E) 1

Clicker Question

How many $T \subseteq V$ ($|T|=4$) such that $|S \cap T| = 3$?

(A) n choose 4

(B) n choose 3

(C) n choose 2

(D) 4(n-4)

(E) None of the above

Clicker Question

How many $T \subseteq V$ ($|T|=4$) such that $|S \cap T| = 3$?

(A) n choose 4

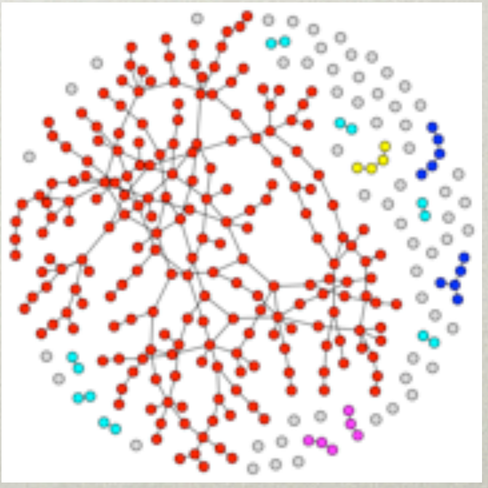
(B) n choose 3

(C) n choose 2

(D) 4(n-4)

(E) None of the above

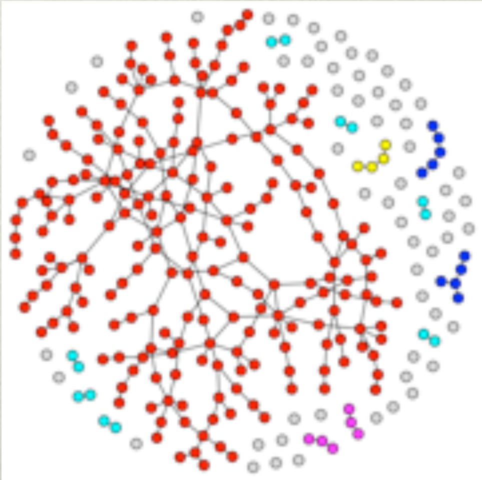
Clique Threshold recap



Clique Threshold recap

Theorem: If **$p \gg n^{-2/3}$** then **G** almost surely has a **4-clique**.

proof: **X := # 4-cliques** in **G**

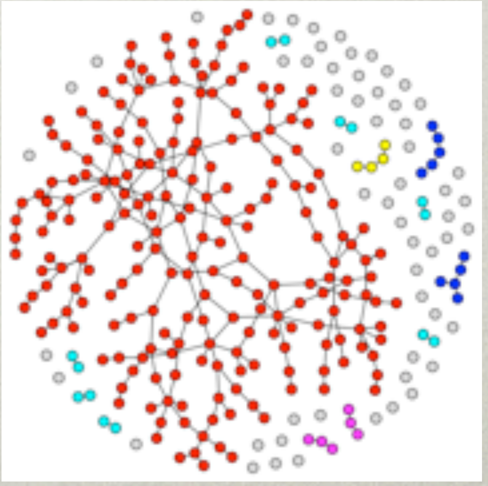


Clique Threshold recap

Theorem: If $p >> n^{-2/3}$ then G almost surely has a **4-clique**.

proof: X := # 4-cliques in G

- $E[X] \sim n^4p^6/24 \to \infty$
- $Var[X] \leq E[X] + \sum_{S\sim T} Pr[A_S \cap A_T]$

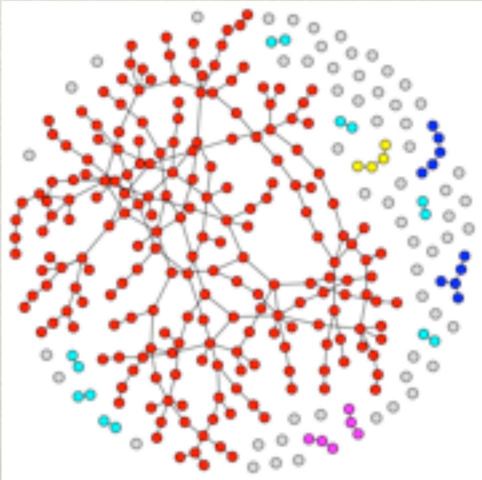


Clique Threshold recap

Theorem: If $p \gg n^{-2/3}$ then G almost surely has a **4-clique**.

proof: $X :=$ # 4-cliques in G

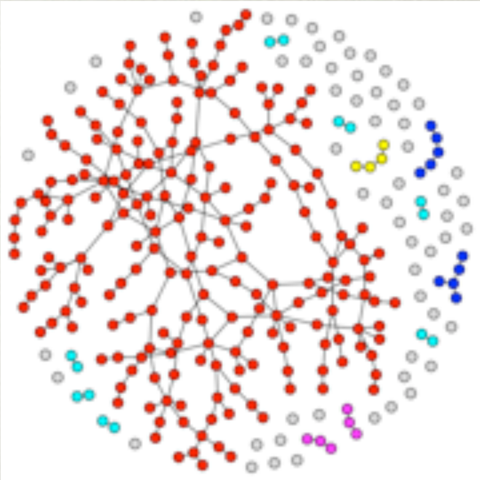
- $E[X] \sim n^4p^6/24 \rightarrow \infty$
- $Var[X] \leq E[X] + \sum_{S\sim T} Pr[A_S \cap A_T]$
- For any S:
  - $O(n^2)$ T with $|S \cap T| = 2$; $Pr[A_T|A_S] = p^5$
  - $O(n)$ T with $|S \cap T| = 3$; $Pr[A_T|A_S] = p^3$
  - $\sum_{T\sim S} Pr[A_S \cap A_T] = O(n^2p^5) + O(np^3) = o(E[X])$



Clique Threshold recap

Theorem: If $p >> n^{-2/3}$ then G almost surely has a 4-clique.
proof: $X :=$ # 4-cliques in G

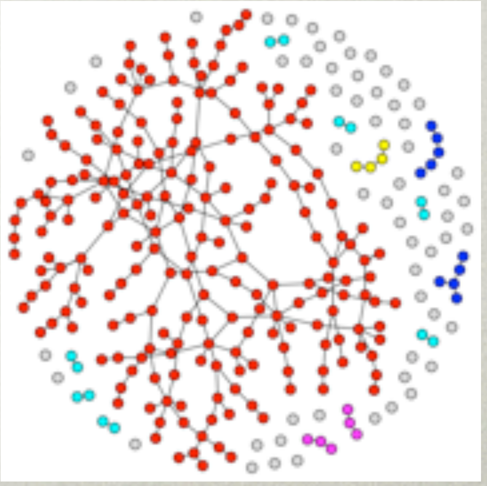
- $E[X] \sim n^4p^6/24 \to \infty$
- $Var[X] \leq E[X] + \sum_{S\sim T} Pr[A_S \cap A_T]$
- For any S:
  - $O(n^2)$ T with $|S \cap T| = 2$; $Pr[A_T|A_S] = p^5$
  - $O(n)$ T with $|S \cap T| = 3$; $Pr[A_T|A_S] = p^3$
  - $\sum_{T\sim S} Pr[A_S \cap A_T] = O(n^2p^5) + O(np^3) = o(E[X])$
- $\sum_{S\sim T} Pr[A_S \cap A_T] = \sum_S Pr[A_S] \sum_{T\sim S} Pr[A_T \mid A_S]$
  $= \sum_S Pr[A_S] o(E[X])$
  $= o(E[X]^2)$



Clique Threshold recap

Theorem: If $p \gg n^{-2/3}$ then G almost surely has a **4-clique**.
proof: $X :=$ # 4-cliques in G

- $E[X] \sim n^4p^6/24 \to \infty$
- $Var[X] \le E[X] + \sum_{S\sim T} Pr[A_S \cap A_T]$
- For any S:
  - $O(n^2)$ T with $|S \cap T| = 2$; $Pr[A_T|A_S] = p^5$
  - $O(n)$ T with $|S \cap T| = 3$; $Pr[A_T|A_S] = p^3$
  - $\sum_{T\sim S} Pr[A_S \cap A_T] = O(n^2p^5) + O(np^3) = o(E[X])$
- $\sum_{S\sim T} Pr[A_S \cap A_T] = \sum_S Pr[A_S] \sum_{T\sim S} Pr[A_T \mid A_S]$
  $= \sum_S Pr[A_S] o(E[X])$
  $= o(E[X]^2)$



- $Var[X] \le E[X] + o(E[X]^2) = o(E[X]^2)$
- Therefore $X \sim E[X] \gg 0$ almost always.

Threshold Functions

Definition: $r(n)$ is a threshold function for graph property P if

(1) If $p \ll r(n)$ then $G(n,p)$ almost never has P

(2) If $p \gg r(n)$ then $G(n,p)$ almost always has P

Threshold Functions

Definition: **r(n)** is a threshold function for graph property **P** if

(1) If **p << r(n)** then **G(n,p)** almost never has **P**

(2) If **p >> r(n)** then **G(n,p)** almost always has **P**

Examples:

- **CL(G) ≥ 4** has threshold function **r(n) =** $n^{-2/3}$
- **Connected** has threshold function **r(n) = ln(n)/n**

Threshold Functions

Definition: $r(n)$ is a threshold function for graph property P if

(1) If $p \ll r(n)$ then $G(n,p)$ almost never has P

(2) If $p \gg r(n)$ then $G(n,p)$ almost always has P

Examples:

- $CL(G) \geq 4$ has threshold function $r(n) = n^{-2/3}$
- **Connected** has threshold function $r(n) = \ln(n)/n$
- **Largest Component**:
  - ★ $p < (1-c)/n$ then largest component is $O(\log n)$
  - ★ $p = 1/n$ then largest component is $n^{2/3}$
  - ★ $p > (1+c)/n$ then largest component is $> n/2$

Concentration of Measure

Concentration of Measure

Examples:

- **degrees:** **deg(v) ~ pn** almost always *(Chernoff Bounds)*

Concentration of Measure

Examples:

- **degrees: deg(v) ~ pn** almost always *(Chernoff Bounds)*
- **clique number: CL(G) ~ 2log(n)** almost always

Concentration of Measure

Examples:

- **degrees:** **deg(v) ~ pn** almost always *(Chernoff Bounds)*
- **clique number:** **CL(G) ~ 2log(n)** almost always
  - ★ there is **r ~2log(n)/log(1/p)** s.t. **CL(G) = r** or **r+1** almost always

Concentration of Measure

Examples:

- **degrees:** **$deg(v) \sim pn$** almost always *(Chernoff Bounds)*
- **clique number:** **$CL(G) \sim 2\log(n)$** almost always
  - ★ there is **$r \sim 2\log(n)/\log(1/p)$** s.t. **$CL(G) = r$** or **$r+1$** almost always
- **chromatic number:** χ**(G) =** # colors needed to color **G**

Concentration of Measure

Examples:

- **degrees:** **deg(v) ~ pn** almost always *(Chernoff Bounds)*
- **clique number:** **CL(G) ~ 2log(n)** almost always
  - ★ there is **r ~2log(n)/log(1/p)** s.t. **CL(G) = r** or **r+1** almost always
- **chromatic number:** **χ(G) =** # colors needed to color **G**
  - ★ **χ(G) ~ n log(1/1-p) / 2log(n)** almost always

Concentration of Measure

Examples:

- **degrees:** **$deg(v) \sim pn$** almost always *(Chernoff Bounds)*
- **clique number:** **$CL(G) \sim 2\log(n)$** almost always
  - ★ there is **$r \sim 2\log(n)/\log(1/p)$** s.t. **$CL(G) = r$** or **$r+1$** almost always
- **chromatic number:** **$\chi(G) =$** # colors needed to color **G**
  - ★ **$\chi(G) \sim n \log(1/1-p) / 2\log(n)$** almost always
- **diameter:** max distance between nodes

Concentration of Measure

Examples:

- **degrees:** **deg(v) ~ pn** almost always *(Chernoff Bounds)*
- **clique number:** **CL(G) ~ 2log(n)** almost always
  - ★ there is **r ~2log(n)/log(1/p)** s.t. **CL(G) = r** or **r+1** almost always
- **chromatic number:** **χ(G) =** # colors needed to color **G**
  - ★ **χ(G) ~ n log(1/1-p) / 2log(n)** almost always
- **diameter:** max distance between nodes
  - ★ **diam(G) ~ log(n)/log(pn)** almost always (if **p >> 1/n**)

Concentration of Measure

Examples:

- **degrees:** **$deg(v) \sim pn$** almost always *(Chernoff Bounds)*
- **clique number:** **$CL(G) \sim 2\log(n)$** almost always
  - ★ there is **$r \sim 2\log(n)/\log(1/p)$** s.t. **$CL(G) = r$** or **$r+1$** almost always
- **chromatic number:** **$\chi(G)$ =** # colors needed to color **G**
  - ★ **$\chi(G) \sim n \log(1/1-p) / 2\log(n)$** almost always
- **diameter:** max distance between nodes
  - ★ **$diam(G) \sim \log(n)/\log(pn)$** almost always (if **$p >> 1/n$**)
  - ★ if **$p = \Omega(\log(n)/n)$** then concentration is on **$O(1)$** values

Zero-One Laws

Definition: **fix $0 < p < 1$.** Property P obeys 0-1 law if
$\lim_{n\to\infty} \Pr[G(n,p) \text{ has } P] = 0 \text{ or } 1$.

Zero-One Laws

Definition: **fix $0 < p < 1$.** Property P obeys 0-1 law if
$\lim_{n\to\infty} \Pr[G(n,p) \text{ has } P] = 0 \text{ or } 1$.

Examples:

- **G** has **triangle**
- **G** has **no isolated vertex**
- **G** has **diameter < 2**

Zero-One Laws

Definition: **fix $0 < p < 1$.** Property P obeys 0-1 law if **$\lim_{n\to\infty} \Pr[G(n,p) \text{ has } P] = 0 \text{ or } 1$**.

Examples:

- **G** has **triangle**
- **G** has **no isolated vertex**
- **G** has **diameter < 2**

Theorem: fix **$0 < p < 1$.** **Any property** expressed in **first-order theory of graphs** obeys **0-1** law.

The Probabilistic Method