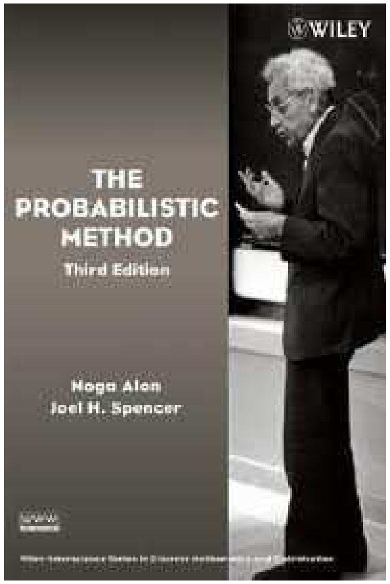
Topic Course on Probabilistic Methods
(Week 10)
Poisson Paradigm

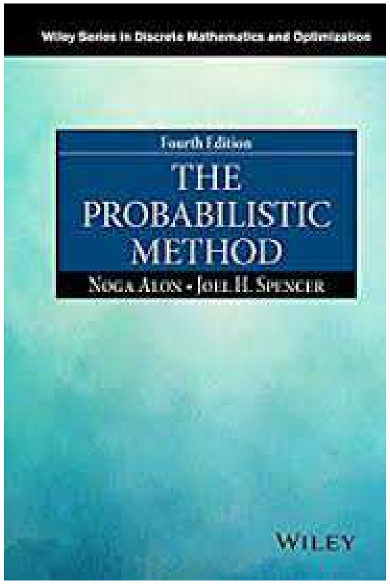
Linyuan Lu

University of South Carolina

Introduction

The topic course is mostly based the textbook "The probabilistic Method" by Noga Alon and Joel Spencer (third edition 2008, John Wiley & Sons, Inc. ISBN 9780470170205 or fourth edition ISBN-13: 978-1119061953.)





Selected topics

- Linearity of Expectation (2 weeks)
- Alterations (1 week)
- The second moment method (1 week)
- The Local Lemma (1-2 weeks)
- Correlation Inequalities (1 week)
- Large deviation inequalities (3 weeks)
- Poisson Paradigm (1 week)
- Random graphs (2 weeks)
- Discrepancy (1 week)
- Entropy (1 week)

Subtopics

Poisson paradigm

- Poisson Paradigm
- Janson's inequality
- Brun's sieve
- An application–EPIT
- Large deviation (sparse cases)
- Suen's theorem

Poisson Paradigm

Let $\{B_i\}_{i \in I}$ be a set of bad events. We will estimate $\Pr(\wedge_{i \in I} \bar{B}_i)$.

Poisson Paradigm

Let $\{B_i\}_{i\in I}$ be a set of bad events. We will estimate $\Pr(\wedge_{i\in I}\bar{B}_i)$.

- If B_i's are "mostly independent", then one may expect

$$\Pr(\wedge_{i\in I}\bar{B}_i) \approx \prod_{i\in I} \Pr(\bar{B}_i).$$

Poisson Paradigm

Let $\{B_i\}_{i\in I}$ be a set of bad events. We will estimate $\Pr(\wedge_{i\in I}\bar{B}_i)$.

- If B_i's are "mostly independent", then one may expect

$$\Pr(\wedge_{i\in I}\bar{B}_i) \approx \prod_{i\in I}\Pr(\bar{B}_i).$$

- Let X_i be the random indicator of the event B_i and $X = \sum_{i\in I} X_i$. If $\Pr(B_i)$'s are small and "mostly independent", then one may expect X follows "Poisson-like distribution". In particular,

$$\Pr(X = 0) \approx e^{-\mathrm{E}(X)}.$$

Setting

- U: a finite universal set.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.
- $\{A_i\}_{i \in I}$: a family of subsets of U.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.
- $\{A_i\}_{i \in I}$: a family of subsets of U.
- $i \sim j$ if $A_i \cap A_j \neq \emptyset$.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.
- $\{A_i\}_{i \in I}$: a family of subsets of U.
- $i \sim j$ if $A_i \cap A_j \neq \emptyset$.
- B_i: the event $A_i \subset R$ for $i \in I$.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.
- $\{A_i\}_{i \in I}$: a family of subsets of U.
- $i \sim j$ if $A_i \cap A_j \neq \emptyset$.
- B_i: the event $A_i \subset R$ for $i \in I$.
- X_i: the indicator random variable for B_i.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.
- $\{A_i\}_{i \in I}$: a family of subsets of U.
- $i \sim j$ if $A_i \cap A_j \neq \emptyset$.
- B_i: the event $A_i \subset R$ for $i \in I$.
- X_i: the indicator random variable for B_i.
- $X := \sum_{i \in I} X_i$; $\mu = \mathrm{E}(X) = \sum_{i \in I} \Pr(B_i)$.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.
- $\{A_i\}_{i \in I}$: a family of subsets of U.
- $i \sim j$ if $A_i \cap A_j \neq \emptyset$.
- B_i: the event $A_i \subset R$ for $i \in I$.
- X_i: the indicator random variable for B_i.
- $X := \sum_{i \in I} X_i$; $\mu = \mathrm{E}(X) = \sum_{i \in I} \Pr(B_i)$.
- $\Delta = \sum_{i \sim j} \Pr(B_i \wedge B_j)$; sum over all ordered pairs $i \sim j$.

Setting

- U: a finite universal set.
- R: a random subset of U given by $\Pr(r \in R) = p_r$.
- $\{A_i\}_{i \in I}$: a family of subsets of U.
- $i \sim j$ if $A_i \cap A_j \neq \emptyset$.
- B_i: the event $A_i \subset R$ for $i \in I$.
- X_i: the indicator random variable for B_i.
- $X := \sum_{i \in I} X_i$; $\mu = \mathrm{E}(X) = \sum_{i \in I} \Pr(B_i)$.
- $\Delta = \sum_{i \sim j} \Pr(B_i \wedge B_j)$; sum over all ordered pairs $i \sim j$.
- $M = \prod_{i \in I} \Pr(\bar{B}_i)$.

Janson inequality

The Janson inequality: Assume all $\Pr(B_i) \leq \epsilon$. Then

$$M \leq \Pr(\wedge_{i \in I} \bar{B}_i) \leq M e^{\frac{\Delta}{2(1-\epsilon)}},$$

and, further,

$$\Pr(\wedge_{i \in I} \bar{B}_i) \leq e^{-\mu + \frac{\Delta}{2}}.$$

Janson inequality

The Janson inequality: Assume all $\Pr(B_i) \leq \epsilon$. Then

$$M \leq \Pr(\wedge_{i\in I}\bar{B}_i) \leq Me^{\frac{\Delta}{2(1-\epsilon)}},$$

and, further,

$$\Pr(\wedge_{i\in I}\bar{B}_i) \leq e^{-\mu+\frac{\Delta}{2}}.$$

The Extended Janson inequality: If further $\Delta \geq \mu$, then

$$\Pr(\wedge_{i\in I}\bar{B}_i) \leq e^{\frac{-\mu^2}{2\Delta}}.$$

Proof

Proof given by Boppana and Spencer: We will use the following correlation inequality.

- For all $J \subset I$, $i \notin J$,

$$\Pr(B_i \mid \wedge_{j \in J} \bar{B}_j) \leq \Pr(B_i).$$

- For $J \subset I$, $i, k \notin J$,

$$\Pr(B_i \mid B_k \wedge \wedge_{j \in J} \bar{B}_j) \leq \Pr(B_i \mid B_k).$$

Proof

Proof given by Boppana and Spencer: We will use the following correlation inequality.

- For all $J \subset I$, $i \notin J$,

$$\Pr(B_i \mid \wedge_{j \in J} \bar{B}_j) \leq \Pr(B_i).$$

- For $J \subset I$, $i, k \notin J$,

$$\Pr(B_i \mid B_k \wedge \wedge_{j \in J} \bar{B}_j) \leq \Pr(B_i \mid B_k).$$

Order the index set $I = \{1, 2, \ldots, m\}$.

$$\Pr(\wedge_{i \in I} \bar{B}_i) = \prod_{i=1}^{m} \Pr\left(\bar{B}_i \mid \wedge_{1 \leq j < i} \bar{B}_j\right) \geq \prod_{i=1}^{m} \Pr(\bar{B}_i).$$

Continue

For a given i renumber, for convenience, so that $i \sim j$ for $1 \leq j \leq d$ and not for $d+1 \leq j < i$. Let $A = B_i$, $B = \bar{B}_1 \wedge \cdots \wedge \bar{B}_d$, and $C = \bar{B}_{d+1} \wedge \cdots \wedge \bar{B}_{i-1}$,

$$\begin{aligned} \Pr(B_i \mid \wedge_{1 \leq j < i} \bar{B}_j) &= \Pr(A \mid B \wedge C) \\ &\geq \Pr(A \wedge B \mid C) \\ &= \Pr(A \mid C) \Pr(B \mid A \wedge C). \end{aligned}$$

Note $\Pr(A \mid C) = \Pr(A)$ and

$$\Pr(B \mid A \wedge C) \geq 1 - \sum_{j=1}^{d} \Pr(B_j \mid B_i \wedge C) \geq 1 - \sum_{j=1}^{d} \Pr(B_j \mid B_i).$$

Continue

For a given i renumber, for convenience, so that $i \sim j$ for $1 \leq j \leq d$ and not for $d+1 \leq j < i$. Let $A = B_i$, $B = \bar{B}_1 \wedge \cdots \wedge \bar{B}_d$, and $C = \bar{B}_{d+1} \wedge \cdots \wedge \bar{B}_{i-1}$,

$$\begin{aligned}
\Pr(B_i \mid \wedge_{1\leq j<i}\bar{B}_j) &= \Pr(A \mid B \wedge C) \\
&\geq \Pr(A \wedge B \mid C) \\
&= \Pr(A \mid C)\Pr(B \mid A \wedge C).
\end{aligned}$$

Note $\Pr(A \mid C) = \Pr(A)$ and

$$\Pr(B \mid A \wedge C) \geq 1 - \sum_{j=1}^{d} \Pr(B_j \mid B_i \wedge C) \geq 1 - \sum_{j=1}^{d} \Pr(B_j \mid B_i).$$

$$\Pr(B_i \mid \wedge_{1\leq j<i}\bar{B}_j) \geq \Pr(B_i) - \textstyle\sum_{j=1}^{d} \Pr(B_j \wedge B_i).$$

Continue

$$
\begin{aligned}
\Pr(\bar{B}_i \mid \wedge_{1\le j<i}\bar{B}_j) &\le \Pr(\bar{B}_i) + \sum_{j=1}^{d} \Pr(B_j \wedge B_i) \\
&\le \Pr(\bar{B}_i)\left(1 + \frac{1}{1-\epsilon}\sum_{j=1}^{d} \Pr(B_j \wedge B_i)\right) \\
&\le \Pr(\bar{B}_i) e^{\frac{1}{1-\epsilon}\sum_{j=1}^{d} \Pr(B_j \wedge B_i)}.
\end{aligned}
$$

Continue

$$\begin{aligned}
\Pr(\bar{B}_i \mid \wedge_{1\le j<i}\bar{B}_j) &\le \Pr(\bar{B}_i) + \sum_{j=1}^{d} \Pr(B_j \wedge B_i) \\
&\le \Pr(\bar{B}_i)\left(1 + \frac{1}{1-\epsilon}\sum_{j=1}^{d} \Pr(B_j \wedge B_i)\right) \\
&\le \Pr(\bar{B}_i) e^{\frac{1}{1-\epsilon}\sum_{j=1}^{d} \Pr(B_j \wedge B_i)}.
\end{aligned}$$

Plug it into $\Pr(\wedge_{i\in I}\bar{B}_i) = \prod_{i=1}^{m} \Pr\left(\bar{B}_i \mid \wedge_{1\le j<i}\bar{B}_j\right)$; we get the first inequality. The second inequality use the following estimation.

$$\begin{aligned}
\Pr(\bar{B}_i \mid \wedge_{1\le j<i}\bar{B}_j) &\le \Pr(\bar{B}_i) + \sum_{j=1}^{d} \Pr(B_j \wedge B_i) \\
&\le e^{-\Pr(B_i)+\sum_{j=1}^{d} \Pr(B_j \wedge B_i)}.
\end{aligned}$$

Proof of second Theorem

From the Jansen inequality, we have

$$-\ln(\Pr(\wedge_{i\in I}\bar{B}_i)) \geq \sum_{i\in I}\Pr(B_i) - \frac{1}{2}\sum_{i\sim j}\Pr(B_i \wedge B_j).$$

Proof of second Theorem

From the Jansen inequality, we have

$$-\ln(\Pr(\wedge_{i\in I}\bar{B}_i)) \geq \sum_{i\in I}\Pr(B_i) - \frac{1}{2}\sum_{i\sim j}\Pr(B_i \wedge B_j).$$

For any set $S \subset I$, the same inequality applied to $\{B_i\}_{i\in S}$:

$$-\ln(\Pr(\wedge_{i\in S}\bar{B}_i)) \geq \sum_{i\in S}\Pr(B_i) - \frac{1}{2}\sum_{i,j\in S, i\sim j}\Pr(B_i \wedge B_j).$$

Proof of second Theorem

From the Jansen inequality, we have

$$-\ln(\Pr(\wedge_{i\in I}\bar{B}_i)) \geq \sum_{i\in I}\Pr(B_i) - \frac{1}{2}\sum_{i\sim j}\Pr(B_i \wedge B_j).$$

For any set $S \subset I$, the same inequality applied to $\{B_i\}_{i\in S}$:

$$-\ln(\Pr(\wedge_{i\in S}\bar{B}_i)) \geq \sum_{i\in S}\Pr(B_i) - \frac{1}{2}\sum_{i,j\in S, i\sim j}\Pr(B_i \wedge B_j).$$

Now take S be a random subset of I given by $\Pr(i \in S) = p$, and take the expectation.

$$E\left[-\ln(\Pr(\wedge_{i\in S}\bar{B}_i))\right] \geq p\mu - p^2\frac{\Delta}{2}.$$

Continue

Now choose $p = \mu/\Delta$.

$$E\left[-\ln(\Pr(\wedge_{i\in S}\bar{B}_i))\right] \geq \frac{\mu^2}{2\Delta}.$$

Continue

Now choose $p = \mu/\Delta$.

$$E\left[-\ln(\Pr(\wedge_{i\in S}\bar{B}_i))\right] \geq \frac{\mu^2}{2\Delta}.$$

Then there is a specific $S \subset I$ for which

$$-\ln(\Pr(\wedge_{i\in S}\bar{B}_i)) \geq \frac{\mu^2}{2\Delta}.$$

Continue

Now choose $p = \mu/\Delta$.

$$E\left[-\ln(\Pr(\wedge_{i\in S}\bar{B}_i))\right] \geq \frac{\mu^2}{2\Delta}.$$

Then there is a specific $S \subset I$ for which

$$-\ln(\Pr(\wedge_{i\in S}\bar{B}_i)) \geq \frac{\mu^2}{2\Delta}.$$

$$\Pr(\wedge_{i\in S}\bar{B}_i) \leq e^{-\frac{\mu^2}{2\Delta}}.$$

Continue

Now choose $p = \mu/\Delta$.

$$E\left[-\ln(\Pr(\wedge_{i\in S}\bar{B}_i))\right] \geq \frac{\mu^2}{2\Delta}.$$

Then there is a specific $S \subset I$ for which

$$-\ln(\Pr(\wedge_{i\in S}\bar{B}_i)) \geq \frac{\mu^2}{2\Delta}.$$

$$\Pr(\wedge_{i\in S}\bar{B}_i) \leq e^{-\frac{\mu^2}{2\Delta}}.$$

$$\Pr(\wedge_{i\in I}\bar{B}_i) \leq \Pr(\wedge_{i\in S}\bar{B}_i) \leq e^{-\frac{\mu^2}{2\Delta}}. \qquad \square$$

Brun's sieve

- X_i: the indicator random variable for B_i, for $i \in I$.
- $X := \sum_{i=1}^{m} X_i$.
- $m = m(n)$, $B_i = B_i(n)$, and $X = X(n)$.
- Let
$$S^{(r)} = \sum \Pr(B_{i_1} \wedge \cdots \wedge B_{i_r}),$$
where the sum is over all sets $\{i_1, \ldots, i_r\} \subset \{1, 2 \ldots, m\}$.
- Let
$$X^{(r)} = X(X-1)\cdots(X-r+1).$$

Brun's sieve

- X_i: the indicator random variable for B_i, for $i \in I$.
- $X := \sum_{i=1}^m X_i$.
- $m = m(n)$, $B_i = B_i(n)$, and $X = X(n)$.
- Let
$$S^{(r)} = \sum \Pr(B_{i_1} \wedge \cdots \wedge B_{i_r}),$$
where the sum is over all sets $\{i_1, \ldots, i_r\} \subset \{1, 2 \ldots, m\}$.
- Let
$$X^{(r)} = X(X-1)\cdots(X-r+1).$$

By inclusion-exclusion principle,

$$\Pr(X=0) = \Pr(\bar{B}_1 \wedge \cdots \wedge \bar{B}_m) = \sum_{r\geq 0} (-1)^r S^{(r)}.$$

Brun's sieve

Theorem: Suppose there is a constant μ so that for every fixed r,

$$\mathrm{E}\binom{X}{r} = S^{(r)} \to \frac{\mu^r}{r!}.$$

Then

$$\Pr(X=0) \to e^{-\mu},$$

and for every t

$$\Pr(X=t) \to \frac{\mu^t}{t!}e^{-\mu}.$$

Proof

Proof: We only prove the case $t = 0$. Fix $\epsilon > 0$. Choose s so that

$$\left|\sum_{r=0}^{2s}(-1)^r\frac{\mu^r}{r!} - e^{-\mu}\right| \leq \frac{\epsilon}{2}.$$

Select n_0 so that for $n \geq n_0$,

$$|S^{(r)} - \frac{\mu^r}{r!}| \leq \frac{\epsilon}{2s(2s+1)}$$

for $0 \leq r \leq 2s$.

Continue

For such n,

$$\begin{aligned}
\Pr[X=0] &\leq \sum_{r=0}^{2s}(-1)^r S^{(r)} \\
&\leq \sum_{r=0}^{2s}(-1)^r \frac{\mu^r}{r!} + \frac{\epsilon}{2} \\
&\leq e^{-\mu} + \epsilon.
\end{aligned}$$

Continue

For such n,

$$\begin{aligned}
\Pr[X=0] &\leq \sum_{r=0}^{2s}(-1)^r S^{(r)} \\
&\leq \sum_{r=0}^{2s}(-1)^r \frac{\mu^r}{r!} + \frac{\epsilon}{2} \\
&\leq e^{-\mu} + \epsilon.
\end{aligned}$$

Similarly, taking the sum to $2s+1$, we can find n_0 so that for $n \geq n_0$,

$$\Pr[X=0] \geq e^{-\mu} - \epsilon.$$

As ϵ was arbitrary $\Pr(X=0) \to e^{-\mu}$. $\square$

An application

Let $G = G(n,p)$, and EPIT represent the statement that every vertex lies in a triangle.

Theorem (a special case of Spencer's Theorem): Let $c > 0$ be fixed and let $p = p(n)$, $\mu = \mu(n)$ satisfy

$$\begin{aligned} \binom{n-1}{2} p^3 &= \mu, \\ e^{-\mu} &= \frac{c}{n}. \end{aligned}$$

Then

$$\lim_{n\to\infty} Pr(G(n,p) \text{ satisfies } EPIT) = e^{-c}.$$

Proof

First fix $x \in V(G)$. For each unordered $y, z \neq x$ let B_{xyz} be the event that $\{x, y, z\}$ is a triangle of G. Let C_x be the event $\wedge_{y,z} \bar{B}_{xyz}$ and X_x the corresponding indicator random variable. Apply Janson's Inequality to bound $\mathrm{E}(X_x) = \Pr(\wedge_{y,z} \bar{B}_{xyz})$.

Proof

First fix $x \in V(G)$. For each unordered $y, z \neq x$ let B_{xyz} be the event that $\{x, y, z\}$ is a triangle of G. Let C_x be the event $\wedge_{y,z} \bar{B}_{xyz}$ and X_x the corresponding indicator random variable. Apply Janson's Inequality to bound $\mathrm{E}(X_x) = \Pr(\wedge_{y,z} \bar{B}_{xyz})$.

$$\Delta = \sum_{y,z,z'} \Pr(B_{xyz} \wedge B_{xyz'}) = O(n^3 p^5) = o(1)$$

since $p = n^{-2/3+o(1)}$.

Proof

First fix $x \in V(G)$. For each unordered $y, z \neq x$ let B_{xyz} be the event that $\{x, y, z\}$ is a triangle of G. Let C_x be the event $\wedge_{y,z} \bar{B}_{xyz}$ and X_x the corresponding indicator random variable. Apply Janson's Inequality to bound $\mathrm{E}(X_x) = \Pr(\wedge_{y,z} \bar{B}_{xyz})$.

$$\Delta = \sum_{y,z,z'} \Pr(B_{xyz} \wedge B_{xyz'}) = O(n^3 p^5) = o(1)$$

since $p = n^{-2/3+o(1)}$. Thus

$$\mathrm{E}(X_x) \approx e^{-\mu} = \frac{c}{n}.$$

continue

Let $X = \sum_x X_x$, which is the number of vertices x no lying a triangle.

$$\mathrm{E}(X) = \sum_x \mathrm{E}(X_x) \to c.$$

We need to show that the Poisson Paradigm applies to X.

continue

Let $X = \sum_x X_x$, which is the number of vertices x no lying a triangle.

$$\mathrm{E}(X) = \sum_x \mathrm{E}(X_x) \to c.$$

We need to show that the Poisson Paradigm applies to X. Fix r and consider

$$\mathrm{E}\binom{X}{r} = S^{(r)} = \sum \Pr(C_{x_1} \wedge \cdots \wedge C_{x_r}),$$

where the sum is over all sets $\{x_1, \ldots, x_r\}$.

continue

Let $X = \sum_x X_x$, which is the number of vertices x no lying a triangle.

$$\mathrm{E}(X) = \sum_x \mathrm{E}(X_x) \to c.$$

We need to show that the Poisson Paradigm applies to X. Fix r and consider

$$\mathrm{E}\binom{X}{r} = S^{(r)} = \sum \Pr(C_{x_1} \wedge \cdots \wedge C_{x_r}),$$

where the sum is over all sets $\{x_1, \ldots, x_r\}$. Note

$$C_{x_1} \wedge \cdots \wedge C_{x_r} = \wedge_{1 \le i \le r, y, z} \overline{B_{x_i y z}}.$$

Continue

We apply Janson's Inequality again.

$$\sum \Pr(B_{x_i yz}) = p^3 \left(r\binom{n-1}{2} - O(n) \right) = r\mu + O(n^{-1+o(1)}).$$

As before Δ is p^5 times the number of pairs $x_i yz \sim x_j yz$; $\Delta = O(n^3 p^5) = o(1)$.

Continue

We apply Janson's Inequality again.

$$\sum \Pr(B_{x_i yz}) = p^3 \left(r\binom{n-1}{2} - O(n) \right) = r\mu + O(n^{-1+o(1)}).$$

As before Δ is p^5 times the number of pairs $x_i yz \sim x_j yz$;
$\Delta = O(n^3 p^5) = o(1)$.

$$\Pr(C_{x_1} \wedge \cdots \wedge C_{x_r}) \sim e^{-r\mu}$$

$$\mathrm{E}\binom{X}{r} \approx \binom{n}{r} e^{-r\mu} \approx \frac{c^r}{r!}.$$

Continue

We apply Janson's Inequality again.

$$\sum \Pr(B_{x_i yz}) = p^3 \left(r \binom{n-1}{2} - O(n) \right) = r\mu + O(n^{-1+o(1)}).$$

As before Δ is p^5 times the number of pairs $x_i yz \sim x_j yz$; $\Delta = O(n^3 p^5) = o(1)$.

$$\Pr(C_{x_1} \wedge \cdots \wedge C_{x_r}) \sim e^{-r\mu}$$

$$\mathrm{E}\binom{X}{r} \approx \binom{n}{r} e^{-r\mu} \approx \frac{c^r}{r!}.$$

Applying Brun's Sieve method, we have $\Pr(X = 0) \to e^{-c}$.

Large deviations

Let $X = \sum_{i,j} t_{xi} t_{xj} t_{ij}$ be the number of triangles containing x in $G(n,p)$. Let $\mu = \mathrm{E}(X)$.

- Kim-Vu's inequality implies " if $\mu \gg \ln^6 n$, then with probability $1 - o(1)$, $(1-\epsilon)\mu \le X \le (1+\epsilon)\mu$."

Large deviations

Let $X = \sum_{i,j} t_{xi} t_{xj} t_{ij}$ be the number of triangles containing x in $G(n,p)$. Let $\mu = \mathrm{E}(X)$.

- Kim-Vu's inequality implies " if $\mu \gg \ln^6 n$, then with probability $1 - o(1)$, $(1-\epsilon)\mu \leq X \leq (1+\epsilon)\mu$."
- With disjoint family (of Poisson paradigm), one can lower the condition to $\mu \gg \ln n$.

Disjoint family

For a fixed random set R, an index $J \subseteq I$ is a **disjoint family** (disfam) if

- $B_j \subset R$ for every $j \in J$.
- For no j, $j' \in J$ is $j \sim j'$.

Disjoint family

For a fixed random set R, an index $J \subseteq I$ is a **disjoint family** (disfam) if

- $B_j \subset R$ for every $j \in J$.
- For no j, $j' \in J$ is $j \sim j'$.

J is a **maximal disjoint family** (maxdisfam) if in addition

- If $j' \notin J$ and $B_{j'}$ the $j \sim j'$ for some $j \in J$.

Disjoint family

For a fixed random set R, an index $J \subseteq I$ is a **disjoint family** (disfam) if

- $B_j \subset R$ for every $j \in J$.
- For no $j, j' \in J$ is $j \sim j'$.

J is a **maximal disjoint family** (maxdisfam) if in addition

- If $j' \notin J$ and $B_{j'}$ the $j \sim j'$ for some $j \in J$.

Lemma 8.4.1:
$\Pr(\text{there exists a maxdisfam } J, |J| = s) \leq \frac{\mu^s}{s!}$.

Disjoint family

For a fixed random set R, an index $J \subseteq I$ is a **disjoint family** (disfam) if

- $B_j \subset R$ for every $j \in J$.
- For no j, $j' \in J$ is $j \sim j'$.

J is a **maximal disjoint family** (maxdisfam) if in addition

- If $j' \notin J$ and $B_{j'}$ the $j \sim j'$ for some $j \in J$.

Lemma 8.4.1:
$\Pr(\text{there exists a maxdisfam } J, |J| = s) \leq \frac{\mu^s}{s!}$.

Lemma 8.4.2: Let $\nu = \max_{j \in I} \sum_{i \sim j} \Pr(B_i)$. Then
$\Pr(\text{there exists a maxdisfam } J, |J| = s) \leq \frac{\mu^s}{s!} e^{-\mu} e^{s\nu} e^{\Delta/2}$.

Proof of Lemma 8.4.1

Let $\sum^*$ denote the sum over all s-sets $J \subseteq I$ with no $j \sim j'$.
Let $\sum^o$ denote the sum over ordered distinct s-tuples. Then

$$\begin{aligned}
&\Pr(\text{there exists a maxdisfam } J, |J| = s) \\
&\leq \sum\nolimits^* \Pr\left(\wedge_{j\in J} B_j\right) \\
&= \sum\nolimits^* \prod_{j\in J} \Pr(B_j) \leq \frac{1}{s!} \sum\nolimits^o \Pr(B_{j_1}) \cdots \Pr(B_{j_s}) \\
&\leq \frac{1}{s!} \left(\sum_{i\in I} \Pr(B_i) \right)^s = \frac{\mu^s}{s!}.
\end{aligned}$$

Proof of Lemma 8.4.2

Let μ_s denote the minimum, over all $j_1, \ldots, j_s \in I$ of $\sum \Pr(B_i)$, the sum over all $i \in I$ except those i with $i \sim j_l$ for some $1 \leq l \leq s$. We have $\mu_s \geq \mu - s\nu$.

$$\Pr(J \text{ maxdisfam}) = \Pr(J \text{ disfam})\Pr(\wedge^* \bar{B}_i).$$

Applying Janson's inequality, we get

$$\Pr(\wedge^* \bar{B}_i) \leq e^{-\mu_s} e^{\Delta/2}.$$

$$\begin{aligned} \sum{}^* \Pr(J \text{ maxdisfam}) &\leq e^{-\mu_s} e^{\Delta/2} \sum{}^* \Pr(J \text{ disfam}) \\ &\leq \frac{\mu^s}{s!} e^{-\mu_s} e^{\Delta/2} \leq \frac{\mu^s}{s!} e^{-\mu} e^{s\nu} e^{\Delta/2}. \end{aligned}$$

Conclusion

Let P be the Poisson distribution with mean μ. When $\mu \gg \ln n$, $\Delta = o(1)$, and $\nu\mu = o(1)$, then

Conclusion

Let P be the Poisson distribution with mean μ. When $\mu \gg \ln n$, $\Delta = o(1)$, and $\nu\mu = o(1)$, then

$$\Pr(\text{there exists a maxdisfam } J, |J| \le \mu(1-\epsilon)) \\ \le (1+o(1))\Pr(P \le \mu(1-\epsilon));$$

$$\Pr(\text{there exists a maxdisfam } J, \mu(1+\epsilon) \le |J| \le 3\mu) \\ \le (1+o(1))\Pr((1+\epsilon)\mu \le P \le 3\mu);$$

$$\Pr(\text{there exists a maxdisfam } J, |J| \ge 3\mu) \le \sum_{s=3\mu}^{\infty} \frac{\mu^s}{s!} = o(n^{-1}).$$

Conclusion

Let P be the Poisson distribution with mean μ. When $\mu \gg \ln n$, $\Delta = o(1)$, and $\nu\mu = o(1)$, then

$$\begin{aligned}
&\Pr(\text{there exists a maxdisfam } J, |J| \leq \mu(1-\epsilon)) \\
&\qquad \leq (1+o(1))\Pr(P \leq \mu(1-\epsilon)); \\
&\Pr(\text{there exists a maxdisfam } J, \mu(1+\epsilon) \leq |J| \leq 3\mu) \\
&\qquad \leq (1+o(1))\Pr((1+\epsilon)\mu \leq P \leq 3\mu);
\end{aligned}$$

$$\Pr(\text{there exists a maxdisfam } J, |J| \geq 3\mu) \leq \sum_{s=3\mu}^{\infty} \frac{\mu^s}{s!} = o(n^{-1}).$$

With probability $1 - o(n^{-1})$, every maxdisfam J has size between $(1-\epsilon)\mu$ and $(1+\epsilon)\mu$.

Application

Let X be the number of triangles containing x in $G(n, p)$.
Let $\mu = \mathrm{E}(X) \sim \frac{1}{2}n^2p^3$. Assume $\mu \gg \ln n$.

Application

Let X be the number of triangles containing x in $G(n,p)$. Let $\mu = \mathrm{E}(X) \sim \frac{1}{2}n^2p^3$. Assume $\mu \gg \ln n$.

We have $\Delta \leq n^3p^5 = o(1)$ and $\mu\nu = n^3p^5 = o(1)$. Thus, with probability $1 - o(n^{-1})$, every maxdisfam J has size between $(1-\epsilon)\mu$ and $(1+\epsilon)\mu$.

Application

Let X be the number of triangles containing x in $G(n,p)$. Let $\mu = \mathrm{E}(X) \sim \frac{1}{2}n^2p^3$. Assume $\mu \gg \ln n$.

We have $\Delta \leq n^3p^5 = o(1)$ and $\mu\nu = n^3p^5 = o(1)$. Thus, with probability $1 - o(n^{-1})$, every maxdisfam J has size between $(1-\epsilon)\mu$ and $(1+\epsilon)\mu$.

Construct a graph $H = (V, E)$ with $V = \{$ all triangles containing $x\}$ and two triangles is adjacent if they share an edge. The with probability $1 - o(1)$, each vertex xyz has degree at most 9 and no set of four disjoint edges. This implies, for any J, $|J| \geq X - 27$. Thus,

$$X \leq (1+\epsilon)\mu + 27 \leq (1+\epsilon')\mu.$$

Generalization

A sufficient condition for Janson's Inequality:

- I: a dependency digraph; if for each $i \in I$ the event B_i is mutually independent of $\{B_j : i \not\sim j\}$.
- $\Delta := \sum_{i \sim j} \Pr(B_i \wedge B_j)$.
- For all $J \subset I$, $i \notin J$,

$$\Pr(B_i \mid \wedge_{j \in J} \bar{B}_j) \leq \Pr(B_i).$$

- For $J \subset I$, $i, k \notin J$,

$$\Pr(B_i \mid B_k \wedge \wedge_{j \in J} \bar{B}_j) \leq \Pr(B_i \mid B_k).$$

Then Janson's inequality holds.

Suen's theorem

An binary relation $\sim$ on I is **superdenpendency digraph** if the following holds:
Suppose that $J_1, J_2 \subset I$ are disjoint subsets so that there is no edge between J_1 and J_2. Let B^1 be any Boolean combination of the events $\{B_j\}_{j \in J_1}$ and B^2 be any Boolean combination of the events $\{B_j\}_{j \in J_2}$. Then B^1 and B^2 are independent.

Suen's theorem

An binary relation $\sim$ on I is **superdenpendency digraph** if the following holds:
Suppose that $J_1, J_2 \subset I$ are disjoint subsets so that there is no edge between J_1 and J_2. Let B^1 be any Boolean combination of the events $\{B_j\}_{j\in J_1}$ and B^2 be any Boolean combination of the events $\{B_j\}_{j\in J_2}$. Then B^1 and B^2 are independent.
Theorem [Suen]: Under the above conditions,

$$\left|\Pr(\wedge_{i\in I}\bar{B}_i) - M\right| \leq M(e^{\sum_{i\sim j} y(i,j)} - 1),$$

where
$y_{i,j} = (\Pr(B_i \wedge B_j) + \Pr(B_i)\Pr(B_j)) \prod_{l\sim i \text{ or } l\sim j}(1 - \Pr(B_l))^{-1}.$