# Large deviations in random graphs

Wojciech Samotij

(joint work with Matan Harel, Gady Kozma, and Frank Mousset)

The 63rd Annual Meeting of the Australian Mathematical Society
Melbourne, December 6, 2019

# Sums of i.i.d. random variables

Suppose that $Y_1, \ldots, Y_N$ are independent, identically distributed (i.i.d.) random variables (each taking finitely many values).

# Sums of i.i.d. random variables

Suppose that $Y_1, \ldots, Y_N$ are independent, identically distributed (i.i.d.) random variables (each taking finitely many values).

Define

$$X = X_N := Y_1 + \cdots + Y_N.$$

# Sums of i.i.d. random variables

Suppose that $Y_1, \ldots, Y_N$ are independent, identically distributed (i.i.d.) random variables (each taking finitely many values).

Define

$$X = X_N := Y_1 + \cdots + Y_N.$$

Letting $\mu := \mathbb{E}[Y_1]$, we have

$$\mathbb{E}[X_N] = \mathbb{E}[Y_1] + \cdots + \mathbb{E}[Y_N] = \mu N.$$

# Sums of i.i.d. random variables

Suppose that $Y_1, \ldots, Y_N$ are independent, identically distributed (i.i.d.) random variables (each taking finitely many values).

Define

$$X = X_N := Y_1 + \cdots + Y_N.$$

Letting $\mu := \mathbb{E}[Y_1]$, we have

$$\mathbb{E}[X_N] = \mathbb{E}[Y_1] + \cdots + \mathbb{E}[Y_N] = \mu N.$$

**Question**

How 'concentrated' is $X_N$ around its expectation?

# The (Weak) Law of Large Numbers

## The (Weak) Law of Large Numbers (Bernoulli 1713)

For every fixed $\varepsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mu N| \geqslant \varepsilon N\right) = 0.$$

# The (Weak) Law of Large Numbers

**The (Weak) Law of Large Numbers (Bernoulli 1713)**

For every fixed $\varepsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mu N| \geqslant \varepsilon N\right) = 0.$$

In other words, $X_N$ typically takes values between $(\mu - \varepsilon)N$ and $(\mu + \varepsilon)N$.

# The (Weak) Law of Large Numbers

**The (Weak) Law of Large Numbers (Bernoulli 1713)**

For every fixed $\varepsilon > 0$,

$$\lim_{N \to \infty} \Pr\left(|X_N - \mu N| \geqslant \varepsilon N\right) = 0.$$

In other words, $X_N$ typically takes values between $(\mu - \varepsilon)N$ and $(\mu + \varepsilon)N$.

Two possible ways to strengthen this result:

# The (Weak) Law of Large Numbers

## The (Weak) Law of Large Numbers (Bernoulli 1713)

For every fixed $\varepsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mu N| \geqslant \varepsilon N\right) = 0.$$

In other words, $X_N$ typically takes values between $(\mu - \varepsilon)N$ and $(\mu + \varepsilon)N$.

Two possible ways to strengthen this result:

- How fast can $\varepsilon$ tend to zero as $N \to \infty$?

# The (Weak) Law of Large Numbers

## The (Weak) Law of Large Numbers (Bernoulli 1713)

For every fixed $\varepsilon > 0$,

$$\lim_{N \to \infty} \Pr\left(|X_N - \mu N| \geqslant \varepsilon N\right) = 0.$$

In other words, $X_N$ typically takes values between $(\mu - \varepsilon)N$ and $(\mu + \varepsilon)N$.

Two possible ways to strengthen this result:

- How fast can $\varepsilon$ tend to zero as $N \to \infty$?
- What is the rate of convergence?

# Typical deviations – Central Limit Theorem

The standard deviation $\sigma$ of $Y_1$ is defined by

$$\sigma := \sqrt{\mathsf{Var}(Y_1)} = \left(\mathbb{E}[Y_1^2] - \mathbb{E}[Y_1]^2\right)^{1/2}.$$

# Typical deviations – Central Limit Theorem

The standard deviation $\sigma$ of $Y_1$ is defined by

$$\sigma := \sqrt{\mathrm{Var}(Y_1)} = \left(\mathbb{E}[Y_1^2] - \mathbb{E}[Y_1]^2\right)^{1/2}.$$

## The Central Limit Theorem (Laplace 1810)

For every fixed $x \geqslant 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mu N| \geqslant x \cdot \sigma\sqrt{N}\right) = F(x) := \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-u^2/2}\, du.$$

# Typical deviations – Central Limit Theorem

The standard deviation $\sigma$ of $Y_1$ is defined by

$$\sigma := \sqrt{\mathrm{Var}(Y_1)} = \left(\mathbb{E}[Y_1^2] - \mathbb{E}[Y_1]^2\right)^{1/2}.$$

## The Central Limit Theorem (Laplace 1810)

For every fixed $x \geqslant 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mu N| \geqslant x \cdot \sigma\sqrt{N}\right) = F(x) := \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-u^2/2}\, du.$$

The quantity $\sigma\sqrt{N}$ is the standard deviation of $X_N$.

# Typical deviations – Central Limit Theorem

The standard deviation $\sigma$ of $Y_1$ is defined by

$$\sigma := \sqrt{\mathrm{Var}(Y_1)} = \left(\mathbb{E}[Y_1^2] - \mathbb{E}[Y_1]^2\right)^{1/2}.$$

## The Central Limit Theorem (Laplace 1810)

For every fixed $x \geqslant 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mu N| \geqslant x \cdot \sigma\sqrt{N}\right) = F(x) := \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-u^2/2}\, du.$$

The quantity $\sigma\sqrt{N}$ is the standard deviation of $X_N$.

It is already unlikely that $|X_N - \mu N| \gg \sqrt{N}$.

# Typical deviations – Central Limit Theorem

The standard deviation $\sigma$ of $Y_1$ is defined by

$$\sigma := \sqrt{\mathrm{Var}(Y_1)} = \left(\mathbb{E}[Y_1^2] - \mathbb{E}[Y_1]^2\right)^{1/2}.$$

## The Central Limit Theorem (Laplace 1810)

For every fixed $x \geqslant 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mu N| \geqslant x \cdot \sigma\sqrt{N}\right) = F(x) := \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-u^2/2}\, du.$$

The quantity $\sigma\sqrt{N}$ is the standard deviation of $X_N$.

It is already unlikely that $|X_N - \mu N| \gg \sqrt{N}$.

The limiting behaviour depends only on $\mathbb{E}[Y_1]$ and $\mathbb{E}[Y_1^2]$.

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} \colon (0, \infty) \to (0, \infty]$ such that

$$\Pr\left(X_N \geqslant (\mu + \varepsilon)N\right) = \exp\left(-\big(I(\varepsilon) + o(1)\big) \cdot N\right).$$

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} \colon (0,\infty) \to (0,\infty]$ such that

$$\Pr\left(X_N \geqslant (\mu+\varepsilon)N\right) = \exp\left(-\left(I(\varepsilon)+o(1)\right)\cdot N\right).$$

## Proof of the upper bound (sketch).

For every $\lambda > 0$, the function $x \mapsto e^{\lambda x}$ is (strictly) increasing. Thus

$$\Pr(X_N \geqslant t)$$

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} : (0, \infty) \to (0, \infty]$ such that

$$\Pr\left(X_N \geqslant (\mu + \varepsilon)N\right) = \exp\left(-\left(I(\varepsilon) + o(1)\right) \cdot N\right).$$

## Proof of the upper bound (sketch).

For every $\lambda > 0$, the function $x \mapsto e^{\lambda x}$ is (strictly) increasing. Thus

$$\Pr(X_N \geqslant t) = \Pr\left(e^{\lambda X_N} \geqslant e^{\lambda t}\right)$$

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} \colon (0, \infty) \to (0, \infty]$ such that

$$\Pr\left(X_N \geqslant (\mu + \varepsilon)N\right) = \exp\left(-\big(I(\varepsilon) + o(1)\big) \cdot N\right).$$

## Proof of the upper bound (sketch).

For every $\lambda > 0$, the function $x \mapsto e^{\lambda x}$ is (strictly) increasing. Thus

$$\Pr(X_N \geqslant t) = \Pr\left(e^{\lambda X_N} \geqslant e^{\lambda t}\right) \leqslant e^{-\lambda t} \cdot \mathbb{E}\left[e^{\lambda X_N}\right],$$

by Markov's inequality.

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} \colon (0,\infty) \to (0,\infty]$ such that

$$\Pr\left(X_N \geqslant (\mu+\varepsilon)N\right) = \exp\left(-\left(I(\varepsilon)+o(1)\right)\cdot N\right).$$

## Proof of the upper bound (sketch).

For every $\lambda > 0$, the function $x \mapsto e^{\lambda x}$ is (strictly) increasing. Thus

$$\Pr(X_N \geqslant t) = \Pr\left(e^{\lambda X_N} \geqslant e^{\lambda t}\right) \leqslant e^{-\lambda t}\cdot \mathbb{E}\left[e^{\lambda X_N}\right],$$

by Markov's inequality. Morevoer,

$$\mathbb{E}\left[e^{\lambda X_N}\right] = \mathbb{E}\left[e^{\sum_{i=1}^N \lambda Y_i}\right]$$

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} \colon (0,\infty) \to (0,\infty]$ such that

$$\Pr\left(X_N \geqslant (\mu+\varepsilon)N\right) = \exp\left(-\left(I(\varepsilon)+o(1)\right)\cdot N\right).$$

## Proof of the upper bound (sketch).

For every $\lambda > 0$, the function $x \mapsto e^{\lambda x}$ is (strictly) increasing. Thus

$$\Pr(X_N \geqslant t) = \Pr\left(e^{\lambda X_N} \geqslant e^{\lambda t}\right) \leqslant e^{-\lambda t} \cdot \mathbb{E}\left[e^{\lambda X_N}\right],$$

by Markov's inequality. Morevoer,

$$\mathbb{E}\left[e^{\lambda X_N}\right] = \mathbb{E}\left[e^{\sum_{i=1}^{N} \lambda Y_i}\right] = \mathbb{E}\left[\prod_{i=1}^{N} e^{\lambda Y_i}\right]$$

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} \colon (0,\infty) \to (0,\infty]$ such that

$$\Pr\left(X_N \geqslant (\mu+\varepsilon)N\right) = \exp\left(-\left(I(\varepsilon)+o(1)\right)\cdot N\right).$$

## Proof of the upper bound (sketch).

For every $\lambda > 0$, the function $x \mapsto e^{\lambda x}$ is (strictly) increasing. Thus

$$\Pr(X_N \geqslant t) = \Pr\left(e^{\lambda X_N} \geqslant e^{\lambda t}\right) \leqslant e^{-\lambda t}\cdot \mathbb{E}\left[e^{\lambda X_N}\right],$$

by Markov's inequality. Morevoer,

$$\mathbb{E}\left[e^{\lambda X_N}\right] = \mathbb{E}\left[e^{\sum_{i=1}^N \lambda Y_i}\right] = \mathbb{E}\left[\prod_{i=1}^N e^{\lambda Y_i}\right] = \prod_{i=1}^N \mathbb{E}\left[e^{\lambda Y_i}\right] = \mathbb{E}\left[e^{\lambda Y_1}\right]^N,$$

as $Y_1,\dots,Y_N$ are i.i.d.

# Large deviations – Cramér's theorem

## Theorem (Cramér 1938)

There is a function $I = I_{Y_1} \colon (0, \infty) \to (0, \infty]$ such that

$$\Pr\left(X_N \geqslant (\mu + \varepsilon)N\right) = \exp\left(-\left(I(\varepsilon) + o(1)\right) \cdot N\right).$$

## Proof of the upper bound (sketch).

For every $\lambda > 0$, the function $x \mapsto e^{\lambda x}$ is (strictly) increasing. Thus

$$\Pr(X_N \geqslant t) = \Pr\left(e^{\lambda X_N} \geqslant e^{\lambda t}\right) \leqslant e^{-\lambda t} \cdot \mathbb{E}\left[e^{\lambda X_N}\right],$$

by Markov's inequality. Morevoer,

$$\mathbb{E}\left[e^{\lambda X_N}\right] = \mathbb{E}\left[e^{\sum_{i=1}^{N} \lambda Y_i}\right] = \mathbb{E}\left[\prod_{i=1}^{N} e^{\lambda Y_i}\right] = \prod_{i=1}^{N} \mathbb{E}\left[e^{\lambda Y_i}\right] = \mathbb{E}\left[e^{\lambda Y_1}\right]^N,$$

as $Y_1, \ldots, Y_N$ are i.i.d. We choose the optimal value of $\lambda$ (...) □

# A word of motivation

The proof of Cramér's theorem crucially uses the assumption that $X$ is a linear function of independent random variables.

# A word of motivation

The proof of Cramér's theorem crucially uses the assumption that $X$ is a linear function of independent random variables.

What happens if we take away the linearity property and assume that $X$ is a more complicated function of the $Y_i$s?

# A word of motivation

The proof of Cramér's theorem crucially uses the assumption that $X$ is a linear function of independent random variables.

What happens if we take away the linearity property and assume that $X$ is a more complicated function of the $Y_i$s?

Perhaps a low degree polynomial?

# A word of motivation

The proof of Cramér's theorem crucially uses the assumption that $X$ is a linear function of independent random variables.

What happens if we take away the linearity property and assume that $X$ is a more complicated function of the $Y_i$s?

Perhaps a low degree polynomial?

A natural example coming from random graph theory:

$$X_N = \#\text{triangles in } G_{n,p};$$

# A word of motivation

The proof of Cramér's theorem crucially uses the assumption that $X$ is a linear function of independent random variables.

What happens if we take away the linearity property and assume that $X$ is a more complicated function of the $Y_i$s?

Perhaps a low degree polynomial?

A natural example coming from random graph theory:

$$X_N = \#\text{triangles in } G_{n,p};$$

here, $N = \binom{n}{2}$ and $X_N$ may be expressed as degree-three polynomial in $N$ independent Bernoulli random variables.

# Triangles in the binomial random graph

The binomial random graph $G_{n,p}$ has vertex set $\llbracket n \rrbracket := \{1, \ldots, n\}$ and

$$\Pr\left(ij \in G_{n,p}\right) = p \quad \text{for all } i, j \in \llbracket n \rrbracket,$$

independently of all other pairs.

# Triangles in the binomial random graph

The binomial random graph $G_{n,p}$ has vertex set $[\![n]\!] := \{1, \ldots, n\}$ and

$$\Pr\left(ij \in G_{n,p}\right) = p \quad \text{for all } i, j \in [\![n]\!],$$

independently of all other pairs.

A triangle in $G_{n,p}$ is a triple $\{i, j, k\}$ of vertices such that $ij, ik, jk \in G_{n,p}$.

# Triangles in the binomial random graph

The binomial random graph $G_{n,p}$ has vertex set $[\![n]\!] := \{1, \dots, n\}$ and

$$\Pr\left(ij \in G_{n,p}\right) = p \quad \text{for all } i, j \in [\![n]\!],$$

independently of all other pairs.

A triangle in $G_{n,p}$ is a triple $\{i, j, k\}$ of vertices such that $ij, ik, jk \in G_{n,p}$.

Let $X_N$ denote the number of triangles in $G_{n,p}$ and note that

$$X_N = \sum_{i,j,k} Y_{ij} Y_{ik} Y_{jk} \qquad \text{and} \qquad \mathbb{E}[X_N] = \binom{n}{3} p^3,$$

where $Y_{ij} = \mathbf{1}_{ij \in G_{n,p}} \sim \text{Bernoulli}(p)$.

# Triangles in the binomial random graph

The binomial random graph $G_{n,p}$ has vertex set $[\![n]\!] := \{1, \ldots, n\}$ and

$$\Pr\left(ij \in G_{n,p}\right) = p \quad \text{for all } i, j \in [\![n]\!],$$

independently of all other pairs.

A triangle in $G_{n,p}$ is a triple $\{i, j, k\}$ of vertices such that $ij, ik, jk \in G_{n,p}$.

Let $X_N$ denote the number of triangles in $G_{n,p}$ and note that

$$X_N = \sum_{i,j,k} Y_{ij} Y_{ik} Y_{jk} \qquad \text{and} \qquad \mathbb{E}[X_N] = \binom{n}{3} p^3,$$

where $Y_{ij} = \mathbf{1}_{ij \in G_{n,p}} \sim \text{Bernoulli}(p)$.

**Remark**

We will allow $p$ to depend on $n$. In fact, assume $p = p(n) \to 0$ as $n \to \infty$.

# Typical deviations of triangle count

If $\mathbb{E}[X_N] \to \infty$, then $X_N$ obeys a Central Limit Theorem.

# Typical deviations of triangle count

If $\mathbb{E}[X_N] \to \infty$, then $X_N$ obeys a Central Limit Theorem.

**Theorem (Ruciński 1988)**

If $p \gg 1/n$, then, for every fixed $x \geqslant 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mathbb{E}[X_N]| \geqslant x \cdot \sigma_N\right) = F(x) := \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-u^2/2}\, du,$$

where $\sigma_N$ is the standard deviation of $X_N$.

# Typical deviations of triangle count

If $\mathbb{E}[X_N] \to \infty$, then $X_N$ obeys a Central Limit Theorem.

**Theorem (Ruciński 1988)**

If $p \gg 1/n$, then, for every fixed $x \geqslant 0$,

$$\lim_{N\to\infty} \Pr\left(|X_N - \mathbb{E}[X_N]| \geqslant x \cdot \sigma_N\right) = F(x) := \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-u^2/2}\, du,$$

where $\sigma_N$ is the standard deviation of $X_N$.

The standard deviation of $X_N$ is straightforward to compute:

$$\sigma_N^2 = \mathrm{Var}(X_N) = \binom{n}{3} p^3(1-p^3) + \binom{n}{4}\binom{4}{2} p^5(1-p).$$

# Large deviations of triangle count

## Problem

For a given $\delta > 0$, detemine the asymptotics of

$$\Pr\left(|X - \mathbb{E}[X]| \geqslant \delta\mathbb{E}[X]\right).$$

# Large deviations of triangle count

**Problem**

For a given $\delta > 0$, detemine the asymptotics of

$$\Pr\left(|X - \mathbb{E}[X]| \geqslant \delta\mathbb{E}[X]\right).$$

**Problem (upper tail)**

For a given $\delta > 0$, detemine the asymptotics of

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right).$$

# Large deviations of triangle count

## Problem

For a given $\delta > 0$, detemine the asymptotics of

$$\Pr\left(|X - \mathbb{E}[X]| \geqslant \delta\mathbb{E}[X]\right).$$

## Problem (upper tail)

For a given $\delta > 0$, detemine the asymptotics of

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right).$$

## Problem (lower tail)

For a given $\delta \in (0, 1]$, detemine the asymptotics of

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right).$$

# Upper tail – lower bounds

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.

By classical large deviation theory, the 'cost' is $\exp(-c_\delta n^2 p)$.

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.
By classical large deviation theory, the 'cost' is $\exp(-c_\delta n^2 p)$.

**(2nd guess)** Plant a subgraph with many triangles!

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.
By classical large deviation theory, the 'cost' is $\exp(-c_\delta n^2 p)$.

**(2nd guess)** Plant a subgraph with many triangles!
If $G_{n,p}$ contains a graph $G$ with $(1+\delta)\mathbb{E}[X]$ triangles, then $X \geqslant (1+\delta)\mathbb{E}[X]$.

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.

By classical large deviation theory, the 'cost' is $\exp(-c_\delta n^2 p)$.

**(2nd guess)** Plant a subgraph with many triangles!

If $G_{n,p}$ contains a graph $G$ with $(1+\delta)\mathbb{E}[X]$ triangles, then $X \geqslant (1+\delta)\mathbb{E}[X]$.

The 'cost' of planting any $G$ in $G_{n,p}$ is $p^{e(G)}$.

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.

By classical large deviation theory, the 'cost' is $\exp(-c_\delta n^2 p)$.

**(2nd guess)** Plant a subgraph with many triangles!

If $G_{n,p}$ contains a graph $G$ with $(1+\delta)\mathbb{E}[X]$ triangles, then $X \geqslant (1+\delta)\mathbb{E}[X]$.

The 'cost' of planting any $G$ in $G_{n,p}$ is $p^{e(G)}$.

Letting $G$ be the complete graph with $(1+\delta)^{1/3}np$ vertices (which has the required number of triangles), we get a lower bound of $p^{c_\delta n^2 p^2}$.

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.

By classical large deviation theory, the 'cost' is $\exp(-c_\delta n^2 p)$.

**(2nd guess)** Plant a subgraph with many triangles!

If $G_{n,p}$ contains a graph $G$ with $(1+\delta)\mathbb{E}[X]$ triangles, then $X \geqslant (1+\delta)\mathbb{E}[X]$.

The 'cost' of planting any $G$ in $G_{n,p}$ is $p^{e(G)}$.

Letting $G$ be the complete graph with $(1+\delta)^{1/3}np$ vertices (which has the required number of triangles), we get a lower bound of $p^{c_\delta n^2 p^2}$.

If $p \ll 1$, then $p^2 \log(1/p) \ll p$ and the second strategy is more effective!

# Upper tail – lower bounds

**(1st guess)** Increase the number of edges by a factor of $(1+\delta)^{1/3}$.
By classical large deviation theory, the 'cost' is $\exp(-c_\delta n^2 p)$.

**(2nd guess)** Plant a subgraph with many triangles!
If $G_{n,p}$ contains a graph $G$ with $(1+\delta)\mathbb{E}[X]$ triangles, then $X \geqslant (1+\delta)\mathbb{E}[X]$.
The 'cost' of planting any $G$ in $G_{n,p}$ is $p^{e(G)}$.
Letting $G$ be the complete graph with $(1+\delta)^{1/3} np$ vertices (which has the required number of triangles), we get a lower bound of $p^{c_\delta n^2 p^2}$.

If $p \ll 1$, then $p^2 \log(1/p) \ll p$ and the second strategy is more effective!

We conclude that

$$\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big) \geqslant \exp\big(-c_\delta n^2 p^2 \log(1/p)\big).$$

# Upper tail – upper bounds

Progression of upper bounds on $\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$:

# Upper tail – upper bounds

Progression of upper bounds on $\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$:

| | |
|---|---|
| Vu (2001) | $\exp\big(-c_\delta (np)^{3/2}\big)$ |

# Upper tail – upper bounds

Progression of upper bounds on $\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$:

| | |
|---|---|
| Vu (2001) | $\exp\big(-c_\delta (np)^{3/2}\big)$ |
| Janson–Ruciński (2002) | $\exp\big(-c_\delta n^2 p^3\big)$ |

# Upper tail – upper bounds

Progression of upper bounds on $\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$:

| | |
|---|---|
| Vu (2001) | $\exp\big(-c_\delta (np)^{3/2}\big)$ |
| Janson–Ruciński (2002) | $\exp\big(-c_\delta n^2 p^3\big)$ |
| Kim–Vu (2004)<br>Janson–Oleszkiewicz–Ruciński (2004) | $\exp\big(-c_\delta n^2 p^2\big)$ |

# Upper tail – upper bounds

Progression of upper bounds on $\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$:

| | |
|---|---|
| Vu (2001) | $\exp\big(-c_\delta (np)^{3/2}\big)$ |
| Janson–Ruciński (2002) | $\exp\big(-c_\delta n^2 p^3\big)$ |
| Kim–Vu (2004)<br>Janson–Oleszkiewicz–Ruciński (2004) | $\exp\big(-c_\delta n^2 p^2\big)$ |
| Chatterjee (2012)<br>DeMarco–Kahn (2012) | $\exp\big(-c_\delta n^2 p^2 \log(1/p)\big)$ |

# Upper tail – upper bounds

Progression of upper bounds on $\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$:

| | |
|---|---|
| Vu (2001) | $\exp\big(-c_\delta (np)^{3/2}\big)$ |
| Janson–Ruciński (2002) | $\exp\big(-c_\delta n^2 p^3\big)$ |
| Kim–Vu (2004)<br>Janson–Oleszkiewicz–Ruciński (2004) | $\exp\big(-c_\delta n^2 p^2\big)$ |
| Chatterjee (2012)<br>DeMarco–Kahn (2012) | $\exp\big(-c_\delta n^2 p^2 \log(1/p)\big)$ |

**Theorem (Chatterjee / DeMarco–Kahn)**

If $p \gg \log n/n$, then, for every fixed $\delta > 0$,

$$\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big) = \exp\big(-\Theta_\delta\big(n^2 p^2 \log(1/p)\big)\big).$$

# Upper tail – upper bounds

Progression of upper bounds on $\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$:

| | |
|---|---|
| Vu (2001) | $\exp\big(-c_\delta (np)^{3/2}\big)$ |
| Janson–Ruciński (2002) | $\exp\big(-c_\delta n^2 p^3\big)$ |
| Kim–Vu (2004)<br>Janson–Oleszkiewicz–Ruciński (2004) | $\exp\big(-c_\delta n^2 p^2\big)$ |
| Chatterjee (2012)<br>DeMarco–Kahn (2012) | $\exp\big(-c_\delta n^2 p^2 \log(1/p)\big)$ |

**Theorem (Chatterjee / DeMarco–Kahn)**

If $p \gg \log n/n$, then, for every fixed $\delta > 0$,

$$\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big) = \exp\big(-\Theta_\delta\big(n^2 p^2 \log(1/p)\big)\big).$$

The assumption $p \gg \log n/n$ is necessary.

# Upper tail – lower bounds (revisited)

**Question**

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically?

# Upper tail – lower bounds (revisited)

**Question**

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically? (What for?)

# Upper tail – lower bounds (revisited)

**Question**

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically? (What for?)

Define

$$\psi(\delta) = \min\big\{e(G) : \mathbb{E}[X \mid G \subseteq G_{n,p}] \geqslant (1+\delta)\mathbb{E}[X]\big\}.$$

# Upper tail – lower bounds (revisited)

## Question

Can we compute $\log \Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right)$ asymptotically? (What for?)

Define

$$\psi(\delta) = \min\left\{e(G) : \mathbb{E}[X \mid G \subseteq G_{n,p}] \geqslant (1+\delta)\mathbb{E}[X]\right\}.$$

## Proposition (easy)

If $\psi(\delta) \to \infty$, then

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) \geqslant p^{(1+o(1))\cdot\psi(\delta)}.$$

# Upper tail – lower bounds (revisited)

**Question**

Can we compute $\log \Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right)$ asymptotically? (What for?)

Define

$$\psi(\delta) = \min\left\{e(G) : \mathbb{E}[X \mid G \subseteq G_{n,p}] \geqslant (1+\delta)\mathbb{E}[X]\right\}.$$

**Proposition (easy)**

If $\psi(\delta) \to \infty$, then

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) \geqslant p^{(1+o(1))\cdot\psi(\delta)}.$$

**Theorem (Lubetzky–Zhao 2014)**

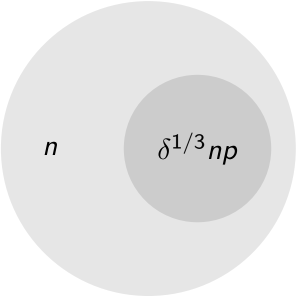
$$\psi(\delta)/n^2p^2 \to \begin{cases} \delta^{2/3}/2 & \text{if } n^{-1} \ll p \ll n^{-1/2}, \\ \min\{\delta^{2/3}/2, \delta/3\} & \text{if } n^{-1/2} \ll p \ll 1. \end{cases}$$

# Optimal planted subgraphs

The constants $\delta^{2/3}/2$ and $\delta/3$ come from the following:

# Optimal planted subgraphs

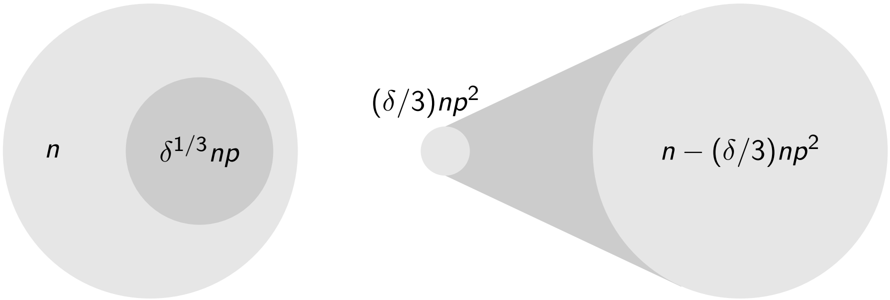
The constants $\delta^{2/3}/2$ and $\delta/3$ come from the following:




clique containing
$\delta\mathbb{E}[X]$ triangles

# Optimal planted subgraphs

The constants $\delta^{2/3}/2$ and $\delta/3$ come from the following:

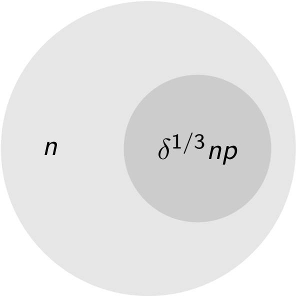



clique containing $\delta\mathbb{E}[X]$ triangles

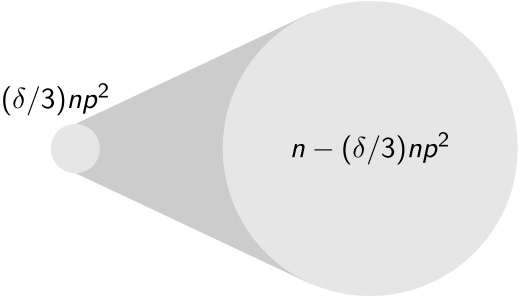
$K_3$-free 'hub' with $\delta\mathbb{E}[X]/p$ copies of $K_{1,2}$

# Optimal planted subgraphs

The constants $\delta^{2/3}/2$ and $\delta/3$ come from the following:



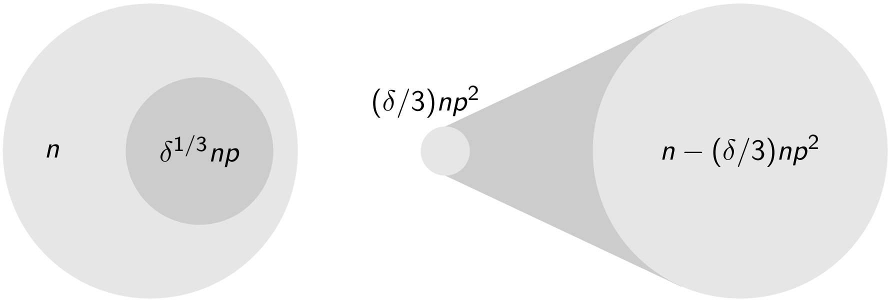

clique containing
$\delta\mathbb{E}[X]$ triangles




$K_3$-free 'hub' with $\delta\mathbb{E}[X]/p$ copies of $K_{1,2}$
(each becomes a triangle with probability $p$)

# Optimal planted subgraphs

The constants $\delta^{2/3}/2$ and $\delta/3$ come from the following:




clique containing $\delta\mathbb{E}[X]$ triangles

$K_3$-free 'hub' with $\delta\mathbb{E}[X]/p$ copies of $K_{1,2}$ (each becomes a triangle with probability $p$)

The 'hub' works only when $np^2 \gg 1$, as $(\delta/3)np^2$ is assumed an integer.

# Upper tail – upper bounds (revisited)

We expect the following to be true (the assumption $p \ll 1$ is needed):

## Theorem

If $n^{-\alpha} \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) \leqslant p^{(1-o(1))\cdot\psi(\delta)}.$$

# Upper tail – upper bounds (revisited)

We expect the following to be true (the assumption $p \ll 1$ is needed):

**Theorem**

If $n^{-\alpha} \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big) \leqslant p^{(1-o(1))\cdot\psi(\delta)}.$$

A short summary of the progression:

| authors | on arXiv | assumption |
|---|---|---|
| | | |

# Upper tail – upper bounds (revisited)

We expect the following to be true (the assumption $p \ll 1$ is needed):

**Theorem**

If $n^{-\alpha} \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) \leqslant p^{(1-o(1))\cdot\psi(\delta)}.$$

A short summary of the progression:

| authors | on arXiv | assumption |
|---|---|---|
| Chatterjee–Dembo | 2014 | $\alpha = 1/42$ |

# Upper tail – upper bounds (revisited)

We expect the following to be true (the assumption $p \ll 1$ is needed):

**Theorem**

If $n^{-\alpha} \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) \leqslant p^{(1-o(1))\cdot\psi(\delta)}.$$

A short summary of the progression:

| authors | on arXiv | assumption |
|---|---|---|
| Chatterjee–Dembo | 2014 | $\alpha = 1/42$ |
| Eldan | 2016 | $\alpha = 1/18$ |

# Upper tail – upper bounds (revisited)

We expect the following to be true (the assumption $p \ll 1$ is needed):

**Theorem**

If $n^{-\alpha} \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) \leqslant p^{(1-o(1))\cdot\psi(\delta)}.$$

A short summary of the progression:

| authors | on arXiv | assumption |
|---|---|---|
| Chatterjee–Dembo | 2014 | $\alpha = 1/42$ |
| Eldan | 2016 | $\alpha = 1/18$ |
| Cook–Dembo | Sep 2018 | $\alpha = 1/3$ |

# Upper tail – upper bounds (revisited)

We expect the following to be true (the assumption $p \ll 1$ is needed):

**Theorem**

If $n^{-\alpha} \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) \leqslant p^{(1-o(1))\cdot\psi(\delta)}.$$

A short summary of the progression:

| authors | on arXiv | assumption |
|---|---:|---:|
| Chatterjee–Dembo | 2014 | $\alpha = 1/42$ |
| Eldan | 2016 | $\alpha = 1/18$ |
| Cook–Dembo | Sep 2018 | $\alpha = 1/3$ |
| Augeri | Oct 2018 | $\alpha = 1/2$ |

# Our contribution

## Theorem (Harel–Mousset–S. 2019)

If $\log n/n \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big) = p^{(1+o(1))\cdot\psi(\delta)}.$$

# Our contribution

**Theorem (Harel–Mousset–S. 2019)**

If $\log n/n \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big) = p^{(1+o(1))\cdot\psi(\delta)}.$$

Moreover, conditioned on the upper tail event, $G_{n,p}$ typically contains either an 'almost-clique' or an 'almost-hub' of the right size

# Our contribution

## Theorem (Harel–Mousset–S. 2019)

If $\log n/n \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) = p^{(1+o(1))\cdot\psi(\delta)}.$$

Moreover, conditioned on the upper tail event, $G_{n,p}$ typically contains either an 'almost-clique' or an 'almost-hub' of the right size ... or a combination of the two if $p = \Theta(n^{-1/2})$.

# Our contribution

**Theorem (Harel–Mousset–S. 2019)**

If $\log n/n \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) = p^{(1+o(1))\cdot\psi(\delta)}.$$

Moreover, conditioned on the upper tail event, $G_{n,p}$ typically contains either an 'almost-clique' or an 'almost-hub' of the right size ... or a combination of the two if $p = \Theta(n^{-1/2})$.

The assumption $p \gg \log n/n$ is necessary.

# Our contribution

## Theorem (Harel–Mousset–S. 2019)

If $\log n/n \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) = p^{(1+o(1))\cdot\psi(\delta)}.$$

Moreover, conditioned on the upper tail event, $G_{n,p}$ typically contains either an 'almost-clique' or an 'almost-hub' of the right size ... or a combination of the two if $p = \Theta(n^{-1/2})$.

The assumption $p \gg \log n/n$ is necessary. But what if $p \ll \log n/n$?

# Our contribution

**Theorem (Harel–Mousset–S. 2019)**

If $\log n/n \ll p \ll 1$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) = p^{(1+o(1))\cdot\psi(\delta)}.$$

Moreover, conditioned on the upper tail event, $G_{n,p}$ typically contains either an 'almost-clique' or an 'almost-hub' of the right size ... or a combination of the two if $p = \Theta(n^{-1/2})$.

The assumption $p \gg \log n/n$ is necessary. But what if $p \ll \log n/n$?

**Theorem (Harel–Mousset–S. 2019)**

If $1/n \ll p \ll \log n/n$, then, for every $\delta > 0$,

$$\Pr\left(X \geqslant (1+\delta)\mathbb{E}[X]\right) = \exp\left(-(1+o(1))\mathrm{Po}(\delta)\cdot\mathbb{E}[X]\right).$$

where $\mathrm{Po}(\delta) = (1+\delta)\log(1+\delta) - \delta$.

# Lower tail

## Problem (lower tail)

For a given $\delta \in (0, 1]$, detemine the asymptotics of

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right).$$

# Lower tail

## Problem (lower tail)

For a given $\delta \in (0, 1]$, detemine the asymptotics of

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right).$$

Harris's correlation inequality implies

$$\Pr(X = 0) \geqslant \max\left\{(1-p^3)^{\binom{n}{3}}, (1-p)^{\binom{n}{2}}\right\}.$$

# Lower tail

## Problem (lower tail)

For a given $\delta \in (0,1]$, detemine the asymptotics of

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right).$$

Harris's correlation inequality implies

$$\Pr(X = 0) \geqslant \max\left\{(1-p^3)^{\binom{n}{3}}, (1-p)^{\binom{n}{2}}\right\}.$$

On the other hand, Janson's inequality gives, for every $\delta \in (0,1]$,

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right) \leqslant \exp\left(-c_\delta \cdot \min\left\{n^2 p, n^3 p^3\right\}\right).$$

# Lower tail

## Problem (lower tail)

For a given $\delta \in (0,1]$, detemine the asymptotics of

$$\Pr\big(X \leqslant (1-\delta)\mathbb{E}[X]\big).$$

Harris's correlation inequality implies

$$\Pr(X=0) \geqslant \max\left\{(1-p^3)^{\binom{n}{3}}, (1-p)^{\binom{n}{2}}\right\}.$$

On the other hand, Janson's inequality gives, for every $\delta \in (0,1]$,

$$\Pr\big(X \leqslant (1-\delta)\mathbb{E}[X]\big) \leqslant \exp\big(-c_\delta \cdot \min\{n^2 p, n^3 p^3\}\big).$$

## Corollary (Harris 1960 / Janson 1990)

If $p < .99$, then, for every $\delta \in (0,1]$,

$$\Pr\big(X \leqslant (1-\delta)\mathbb{E}[X]\big) = \exp\big(-\Theta_\delta\big(\min\{n^2 p, n^3 p\}\big)\big).$$

# Lower tail (revisited)

## Question

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically?

# Lower tail (revisited)

**Question**

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically?

Here, we assume that $p \gg n^{-1/2}$, so that $n^3p^3 \gg n^2p$.

# Lower tail (revisited)

**Question**

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically?

Here, we assume that $p \gg n^{-1/2}$, so that $n^3p^3 \gg n^2p$.

If $G$ has no triangles, then

$$\Pr(X = 0) \geqslant \Pr(G_{n,p} \subseteq G) = (1-p)^{\binom{n}{2} - e(G)}.$$

# Lower tail (revisited)

**Question**

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically?

Here, we assume that $p \gg n^{-1/2}$, so that $n^3p^3 \gg n^2p$.

If $G$ has no triangles, then

$$\Pr(X = 0) \geqslant \Pr(G_{n,p} \subseteq G) = (1-p)^{\binom{n}{2} - e(G)}.$$

The right-hand side is maximised when $G$ is complete bipartite, giving

$$\Pr(X = 0) \geqslant (1-p)^{n^2/4}.$$

# Lower tail (revisited)

## Question

Can we compute $\log \Pr\big(X \geqslant (1+\delta)\mathbb{E}[X]\big)$ asymptotically?

Here, we assume that $p \gg n^{-1/2}$, so that $n^3p^3 \gg n^2p$.

If $G$ has no triangles, then

$$\Pr(X = 0) \geqslant \Pr(G_{n,p} \subseteq G) = (1-p)^{\binom{n}{2}-e(G)}.$$

The right-hand side is maximised when $G$ is complete bipartite, giving

$$\Pr(X = 0) \geqslant (1-p)^{n^2/4}.$$

## Theorem (Łuczak 2000)

If $p \gg n^{-1/2}$, then $\Pr(X = 0) \leqslant (1-p)^{n^2/4 - o(n^2)}$.

# Lower tail (revisited)

If $\delta < 1$, then we could consider a graph $G_\delta$ with at most $(1-\delta)\binom{n}{3}$ triangles and as many edges as possible to obtain

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right) \gtrapprox \Pr(G_{n,p} \subseteq G_\delta) = (1-p)^{\binom{n}{2} - e(G_\delta)}.$$

# Lower tail (revisited)

If $\delta < 1$, then we could consider a graph $G_\delta$ with at most $(1-\delta)\binom{n}{3}$ triangles and as many edges as possible to obtain

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right) \gtrapprox \Pr(G_{n,p} \subseteq G_\delta) = (1-p)^{\binom{n}{2} - e(G_\delta)}.$$

Choose $q\colon \binom{[\![n]\!]}{2} \to [0,1]$ and let $G_{n,q}$ be the random graph on $[\![n]\!]$ s.t.:

$$\Pr(ij \in G_{n,q}) = q_{ij} \quad \text{for all } i,j \in [\![n]\!].$$

# Lower tail (revisited)

If $\delta < 1$, then we could consider a graph $G_\delta$ with at most $(1-\delta)\binom{n}{3}$ triangles and as many edges as possible to obtain

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right) \gtrapprox \Pr(G_{n,p} \subseteq G_\delta) = (1-p)^{\binom{n}{2} - e(G_\delta)}.$$

Choose $q \colon \binom{[\![n]\!]}{2} \to [0,1]$ and let $G_{n,q}$ be the random graph on $[\![n]\!]$ s.t.:

$$\Pr(ij \in G_{n,q}) = q_{ij} \quad \text{for all } i,j \in [\![n]\!].$$

## Proposition

Suppose that $q$ is such that $\mathbb{E}\left[\#K_3(G_{n,q})\right] \leqslant (1-\delta)\mathbb{E}[X] = (1-\delta)\binom{n}{3}p^3$. Then,

$$\Pr\left(X \leqslant (1-\delta)\mathbb{E}[X]\right) \geqslant \exp\left(-(1+o(1)) \cdot \sum_{i,j} I_p(q_{ij})\right),$$

where $I_p(q) = q \log \frac{q}{p} + (1-q)\log\frac{1-q}{1-p}$.

# Our contribution

Define, for every $\delta \in (0,1]$,

$$\Phi(\delta) = \min \left\{ \sum_{i,j} I_p(q_{ij}) : \mathbb{E}[\# K_3(G_{n,q})] \leqslant (1-\delta)\mathbb{E}[X] \right\}.$$

# Our contribution

Define, for every $\delta \in (0, 1]$,

$$\Phi(\delta) = \min \left\{ \sum_{i,j} I_p(q_{ij}) : \mathbb{E}[\# K_3(G_{n,q})] \leqslant (1-\delta)\mathbb{E}[X] \right\}.$$

We have

$$\frac{\Phi(1)}{\log(1-p)} = \mathrm{ex}(n, K_3) - \binom{n}{2} = \left\lfloor \frac{n^2}{4} \right\rfloor - \binom{n}{2},$$

# Our contribution

Define, for every $\delta \in (0,1]$,

$$\Phi(\delta) = \min \left\{ \sum_{i,j} I_p(q_{ij}) : \mathbb{E}[\# K_3(G_{n,q})] \leqslant (1-\delta)\mathbb{E}[X] \right\}.$$

We have

$$\frac{\Phi(1)}{\log(1-p)} = \mathrm{ex}(n, K_3) - \binom{n}{2} = \left\lfloor \frac{n^2}{4} \right\rfloor - \binom{n}{2},$$

but computing the function $\Phi(\delta)$ for all $\delta$ seems very hard.

# Our contribution

Define, for every $\delta \in (0,1]$,

$$\Phi(\delta) = \min\left\{ \sum_{i,j} I_p(q_{ij}) : \mathbb{E}[\#K_3(G_{n,q})] \leqslant (1-\delta)\mathbb{E}[X] \right\}.$$

We have

$$\frac{\Phi(1)}{\log(1-p)} = \mathrm{ex}(n, K_3) - \binom{n}{2} = \left\lfloor \frac{n^2}{4} \right\rfloor - \binom{n}{2},$$

but computing the function $\Phi(\delta)$ for all $\delta$ seems very hard.

**Theorem (Kozma–S. 2019+)**

If $n^{-1/2} \ll p \leqslant .99$, then, for every $\delta \in (0,1]$,

$$\Pr\big(X \leqslant (1-\delta)\mathbb{E}[X]\big) = \exp\big(-(1+o(1)) \cdot \Phi(\delta)\big).$$

**Thank you for your attention!**