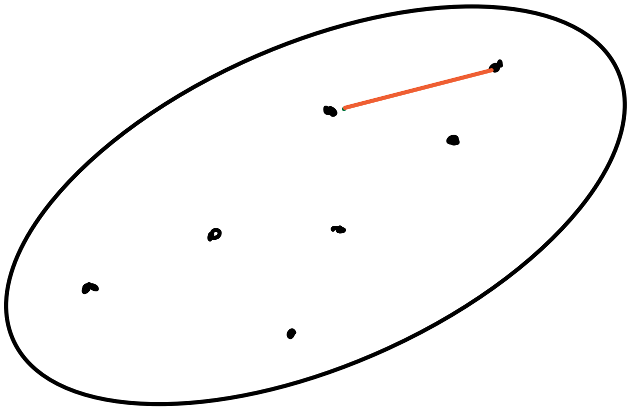
Tail Bounds comparison

Let $\underbrace{Z = X_1 + \cdots + X_n}_{\text{i.i.d. Bernoulli}(1/2)}$

$E[Z] = \frac{n}{2}$ $\qquad$ $Var[Z] = \frac{n}{4}$

	$\mathbb{P}[Z \geq \frac{3n}{4}] \leq ?$	$\mathbb{P}[Z \geq \frac{n}{2} + \sqrt{n}] \leq ?$
Markov $\mathbb{P}[Z \geq t] \leq \frac{E[Z]}{t}$ (for $Z \geq 0$)	$\frac{E[Z]}{\frac{3n}{4}} = \frac{n/2}{3n/4} = \frac{2}{3}$	$\frac{n/2}{\frac{n}{2} + \sqrt{n}} = \frac{n}{n + 2\sqrt{n}} \simeq 1 - \frac{2}{\sqrt{n}}$
Chebyshev $\mathbb{P}[\lvert Z - \mu\rvert \geq t] \leq \frac{Var[Z]}{t^2}$	$\frac{Var[Z]}{(n/4)^2} = \frac{(n/4)}{(n/4)^2} = \frac{4}{n}$	$\frac{(n/4)}{(\sqrt{n})^2} = \frac{1}{4}$

Threshold Phenomena

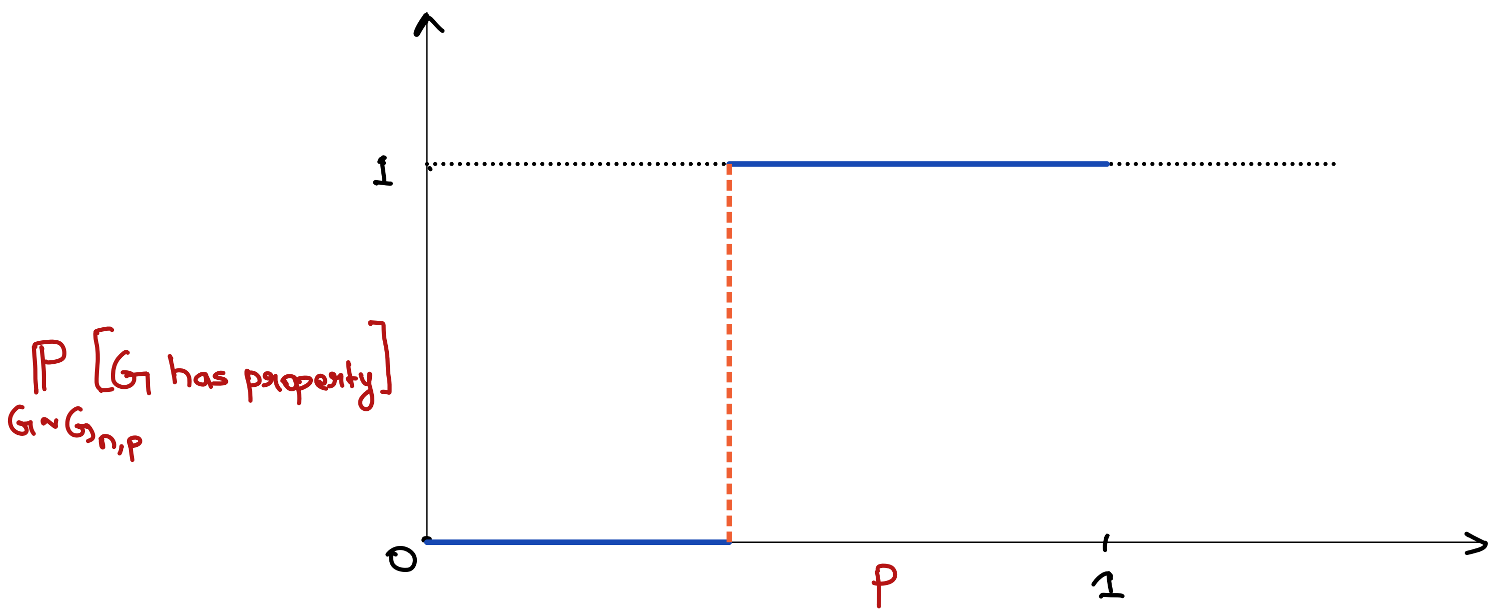


$G = (V, E)$

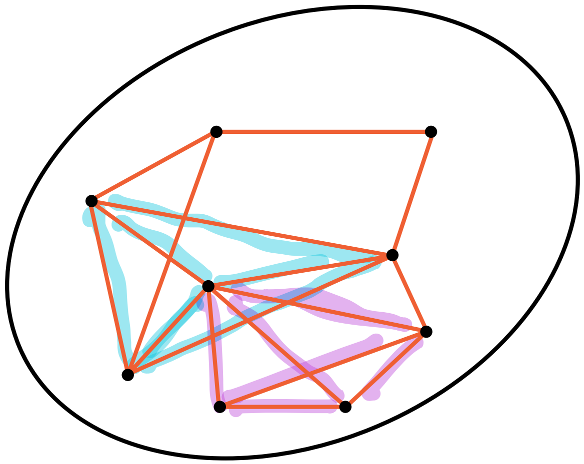
$|V| = n$

$$\left.\begin{array}{l}(i,j) \in E \text{ with probability } p \\ (i,j) \notin E \text{ with probability } 1-p \\ (\text{independently for each } (i,j))\end{array}\right\} G_{n,p}$$

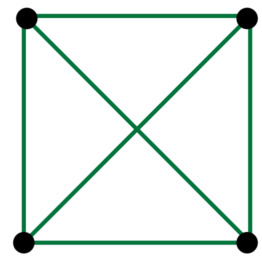
$$\#\text{edges} = Y = \sum X_{\{i,j\}} = \text{Binomial}\left(\binom{n}{2}, p\right)$$



Does random G "contain" a fixed H



$G \sim G_{n,p}$



$H \equiv$ Complete graph on 4 vertices (K_4)

- Let $G \sim G_{n,p}$. Then

  - $p \ll n^{-2/3} \Rightarrow \mathbb{P}[G \text{ contains copy of } K_4] \to 0$ as $n \to \infty$

    $\lim_{n\to\infty} \frac{p}{n^{-2/3}} = 0$

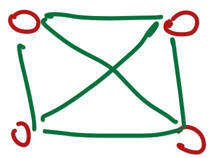
  - $p \gg n^{-2/3} \Rightarrow \mathbb{P}[G \text{ contains copy of } K_4] \to 1$ as $n \to \infty$

    $\lim_{n\to\infty} \frac{p}{n^{-2/3}} = \infty$

Random Variables

$$Z = \#\text{ copies of } K_4 = \sum_{\substack{C \subseteq [n] \\ |C|=4}} X_C \qquad X_C = \begin{cases} 1 & \text{if } C \text{ is a copy of } K_4 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}[Z] = \sum_C \mathbb{E}[X_C] = \sum_C p^6 = \binom{n}{4} \cdot p^6$$



Markov: $\mathbb{P}[Z > 0] = \mathbb{P}[Z \geq 1] \leq \frac{\mathbb{E}[Z]}{1} = \binom{n}{4} \cdot p^6$

$$= O(n^4 \cdot p^6)$$

$$\lim_{n \to \infty} n^4 p^6 = \lim_{n \to \infty} \left(\frac{p}{n^{-2/3}}\right)^6 = 0 \quad \text{for } p \ll n^{-2/3}$$

$$\therefore \ \mathbb{P}[Z > 0] \to 0 \quad \text{for } p \ll n^{-2/3}$$

Existence of K_4 for $p \gg n^{-2/3}$

Chebyshev : $\mathbb{P}[Z=0] \leq \mathbb{P}\big[|Z-\mathbb{E}[Z]| \geq \mathbb{E}[Z]\big]$

$$\leq \frac{\mathrm{Var}[Z]}{(\mathbb{E}[Z])^2}$$

$$Z = \sum_c X_c$$

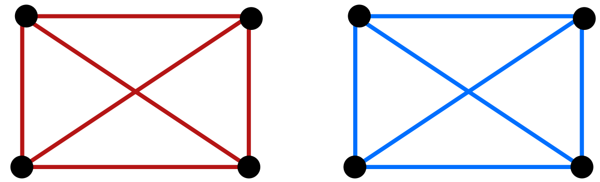
$$\mathrm{Var}[Z] = \mathbb{E}\left[\left(\sum_c X_c - \sum_c \mathbb{E}[X_c]\right)^2\right]$$

$$= \sum_{C,D} \underbrace{\mathbb{E}\left[(X_C - \mathbb{E}X_C)(X_D - \mathbb{E}X_D)\right]}_{= \mathrm{Cov}(X_C,\, X_D)}$$

Understanding Covariances

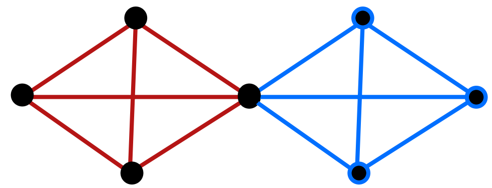
$$\mathbb{E}[X_C X_D] - \mathbb{E}[X_C] \cdot \mathbb{E}[X_D]$$

- $|C \cap D| = 0$



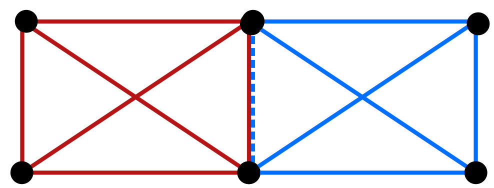
$= 0$

- $|C \cap D| = 1$



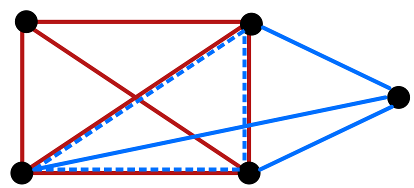
$= 0$

- $|C \cap D| = 2$



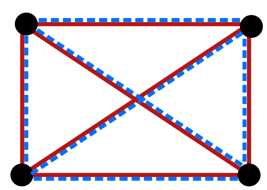
$= p^{11} - p^{12}$

- $|C \cap D| = 3$



$= p^{9} - p^{12}$

- $|C \cap D| = 4$



$\mathrm{Var}[X_C] = p^{6} - p^{12}$

Finishing the computation

$$\mathrm{Var}[Z] = \underbrace{\binom{n}{6}\binom{6}{4}}_{|C\cap D|=2}\cdot\left(p^{11}-p^{12}\right) + \underbrace{\binom{n}{5}\binom{5}{4}}_{|C\cap D|=3}\cdot\left(p^{9}-p^{12}\right) + \underbrace{\binom{n}{4}}_{|C\cap D|=4}\cdot\left(p^{6}-p^{12}\right)$$

$$= O\left(n^6\cdot p^{11}\right) + O\left(n^5\cdot p^9\right) + O\left(n^4\cdot p^6\right)$$

$$P[Z=0] \leq \frac{\mathrm{Var}[Z]}{\left(E[Z]\right)^2} = \frac{O\left(n^6\cdot p^{11}\right) + O\left(n^5\cdot p^9\right) + O\left(n^4\cdot p^6\right)}{\left(\binom{n}{4}\cdot p^6\right)^2}$$

$$= O\left(\frac{1}{n^2\cdot p}\right) + O\left(\frac{1}{n^3\cdot p^3}\right) + O\left(\frac{1}{n^4\cdot p^6}\right)$$

Chernoff-Hoeffding Bounds

$$Z = X_1 + \cdots + X_n \qquad X_1 \ldots X_n \text{ mutually independent}$$

$$X_i \sim \text{Bernoulli}(p_i)$$

$$\mu = \mathbb{E}[Z] = \sum_i \mathbb{E}[X_i] = \sum_i p_i$$

- (Chernoff/Hoeffding)
  - $\forall \delta > 0 \quad \mathbb{P}[Z \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$
  - $\forall \delta \in (0,1) \quad \mathbb{P}[Z \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$

$\leq e^{-\delta^2 \mu/3}$ for $\delta \in (0,1)$

Tail Bounds comparison

Let $Z = X_1 + \cdots + X_n$ i.i.d. Bernoulli(½)

	$\mathbb{P}[Z \geq \frac{3n}{4}] \leq ?$	$\mathbb{P}[Z \geq \frac{n}{2} + \sqrt{n}] \leq ?$
Markov	$\frac{2}{3}$	$\frac{n}{n+2\sqrt{n}} \simeq 1 - \frac{2}{\sqrt{n}}$
Chebyshev	$\frac{4}{n}$	$\frac{1}{4}$
Chernoff-Hoeffding $\mathbb{P}[Z \geq (1+\delta)\mu] \leq e^{-\delta^2\mu/3}$	$\mathbb{P}[Z \geq (1+\frac{1}{2})\cdot\frac{n}{2}] \leq e^{-\frac{1}{4}\cdot\frac{n}{2}\cdot\frac{1}{3}} = e^{-n/24}$	$\mathbb{P}[Z \geq (1+\frac{2}{\sqrt{n}})\cdot\frac{n}{2}] \leq e^{-\frac{4}{n}\cdot\frac{n}{2}\cdot\frac{1}{3}} \leq e^{-2/3}$

Upper tail proof

$$Z = X_1 + \dots + X_n$$

$X_1 \dots X_n$ independent
$X_i \sim \text{Bernoulli}(p_i)$

$\forall \lambda > 0$

$$\mathbb{P}[Z \geq (1+\delta)\mu] = \mathbb{P}[e^{\lambda Z} \geq e^{\lambda \cdot (1+\delta) \cdot \mu}]$$

$$\leq \frac{\mathbb{E}[Z]}{e^{\lambda (1+\delta)\mu}}$$

$$\mathbb{E}[e^{\lambda Z}] = \mathbb{E}[e^{\lambda (X_1 + \dots + X_n)}] = \prod_i \mathbb{E}[e^{\lambda X_i}]$$

$$\mathbb{E}[e^{\lambda X_i}] = p_i \cdot e^{\lambda \cdot 1} + (1 - p_i) \cdot e^{\lambda \cdot 0} = 1 + p_i (e^{\lambda} - 1)$$

Choosing λ

$$\mathbb{E}[e^{\lambda Z}] = \prod_{i=1}^{n} \mathbb{E}[e^{\lambda X_i}] = \prod_{i=1}^{n} (1 + p_i(e^{\lambda}-1))$$

$$\leq \prod_{i=1}^{n} \left(e^{p_i \cdot (e^{\lambda}-1)} \right)$$

$$= e^{(\sum p_i)\cdot(e^{\lambda}-1)} = e^{\mu\cdot(e^{\lambda}-1)}$$

$$\forall \lambda > 0 \quad \mathbb{P}[Z \geq (1+\delta)\cdot\mu] \leq \frac{e^{\mu\cdot(e^{\lambda}-1)}}{e^{\lambda(1+\delta)\mu}} = \left(\frac{e^{e^{\lambda}-1}}{e^{\lambda(1+\delta)}}\right)^{\mu}$$

$$f(\lambda) = \ln \frac{e^{e^{\lambda}-1}}{e^{\lambda(1+\delta)}} = e^{\lambda} - 1 - \lambda(1+\delta) \qquad \frac{df}{d\lambda} = e^{\lambda} - (1+\delta)$$

Best $\lambda = \ln(1+\delta) \quad (e^{\lambda} = 1+\delta)$

Plugging back in : $\mathbb{P}[Z \geq (1+\delta)\cdot\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$

The lower tail

$Z = X_1 + \cdots + X_n$

$X_1, \ldots, X_n$ independent
$X_i \sim \text{Bernoulli}(p_i)$

$\forall \lambda > 0$

$$\mathbb{P}[Z \leq (1-\delta)\cdot\mu] = \mathbb{P}[e^{-\lambda Z} \geq e^{-\lambda(1-\delta)\mu}]$$

<u>Ex:</u> Obtain bound using Markov

Commonly used form: $\mathbb{P}[|Z-\mu| \geq \delta\mu] \leq 2 \cdot e^{-\delta^2 \mu/3}$

(for $\delta \in (0,1)$)