$\boxed{\text{Lecture 5}}$ Chernoff inequality & random graphs

One downside of Hoeffding's iniquality is that it involves σ^2 but not the actual variance. A stronger inequality is possible for Binomial(n,p)

Thm (Chernoff) Let $X_i \sim \mathrm{Ber}(p_i)$ be independent and set $S_n = \sum_{i=1}^n X_i$, $\mu = \mathbb{E}S$. Then

$$P(S_n > t) \leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t \qquad (*)$$

Let's simplify this. Recall

$$\mathrm{Var}(S_n) = \sum_{i=1}^n \mathrm{Var}(X_i) = \sum_{i=1}^n p_i(1-p_i)$$

$$\approx \mu \quad \text{for } p_i \in (0, \tfrac{2}{3})$$

So let's set $t = (1+\delta)\mu$ and then

$$\log (*) = \log\left(e^{-\mu}\left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}\right)$$

$$= \delta\mu - (1+\delta)\mu \log(1+\delta)$$

$$= \mu\left[\delta - (1+\delta)\log(1+\delta)\right]$$

Define the function $h(\delta) = (1+\delta)\log(1+\delta)$

Then $h'(\delta) = 1 + \log(1+\delta)$

$$h''(\delta) = \frac{1}{1+\delta} \geq \frac{1}{2} \quad \forall\, \delta \in (0,1)$$

So the mean value thm says that for any $\delta \geq 0$ $\exists\, \hat{\delta} \in [0, \delta]$ s.t.

$$h(\delta) = h(0) + h'(0)\delta + \frac{1}{2}\delta^2 \underbrace{h''(\hat{\delta})}_{\geq \frac{1}{2}}$$

$$\geq \delta + \frac{\delta^2}{4}$$

So $\log(*) \leq -\frac{\mu\delta^2}{4}$, that is

$$\boxed{P(S_n - \mu \geq \delta\mu) \leq \exp\left(-\frac{\mu\delta^2}{4}\right)}$$

for small deviations $\delta \in (0,1)$.

A similar argument gives the two sided bound

$$\boxed{P(|S_n - \mu| \geq \delta\mu) \leq 2\exp\left(-\frac{\mu\delta^2}{4}\right)}$$

Pf of Chernoff's inequality:

Proceed as usual

$$P(S_n \geq t) \leq \frac{\prod_i \mathbb{E}\exp(\lambda X_i)}{\exp(\lambda t)}$$

$$= \frac{\prod_i \left[1 + p_i(\exp(\lambda) - 1)\right]}{\exp(\lambda t)}$$

The inequality $1 + x \leq e^x$ gives

$$1 + p_i(\exp(\lambda) - 1) \leq \exp(p_i(\exp(\lambda) - 1))$$

$$\leq \exp\Big(\underbrace{\sum p_i}_{= \mu}(e^\lambda - 1)\Big) \cdot e^{-\lambda t}$$

$$= \exp(\mu(e^\lambda - 1) - \lambda t)$$

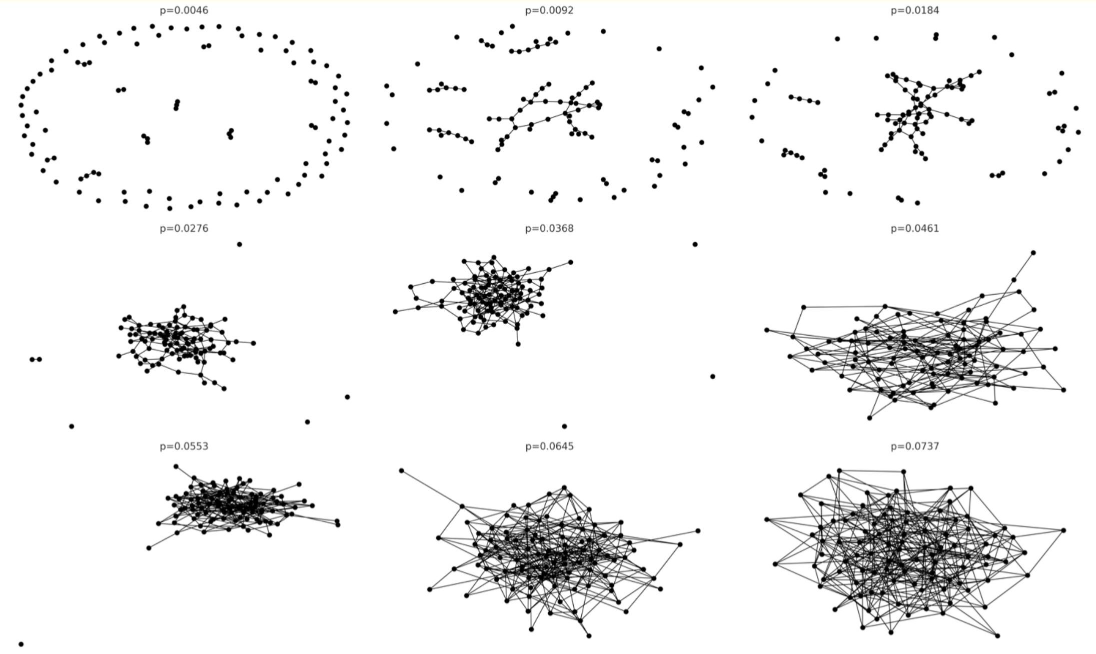
Minimizing in λ gives $\lambda = \ln\left(\frac{t}{\mu}\right)$

$$\leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t$$

as claimed $\boxtimes$

Ex (Regularity of random graphs)

An Erdös-Renyi $G(n,p)$ is a random graph on n vertices obtained by adding each edge independently with probability p.



Some facts: As $n \to \infty$, the following are true.

① $np < 1 \implies$ every connected component has size $O(\log(n))$

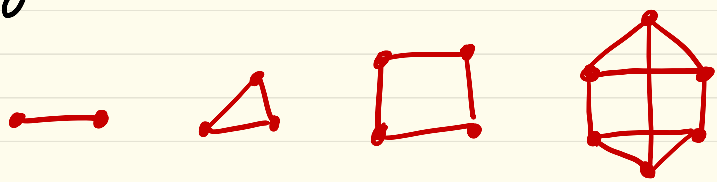
② $np = 1 \implies \exists$ connected component of size $n^{2/3}$

③ $np \to c > 1 \implies \exists$ one connected component of size $\Omega(n)$ and all other components have size $O(\log(n))$

④ $np < (1-\epsilon)\log(n) \implies \exists$ isolated vertices

⑤ $np > (1+\epsilon)\log(n) \implies$ graph is connected

Def: A graph is d-regular if $\deg(i) = d$ for all i.



Thm: For any $\delta \in (0,1)$, if $n(p-1) \geq \frac{8}{\delta^2}\log(2n)$, then $G(n,p)$ is almost d-regular with $d \triangleq (n-1)p$, meaning

$$P\left[(1-\delta)d \leq \deg(i) \leq (1+\delta)d \quad \forall i\right] \geq 1 - \frac{1}{2n}$$

pf: For each i, we may write

$$\deg(i) = \sum_{j \neq i} \mathbb{1}_{ij \in E}$$

where E is the set of edges

So $\mathbb{E}\deg(i) = n-1$ and Chernoff gives

$$P\left[\,|\deg(i) - d| > \delta d\,\right] \leq 2\exp(-\delta^2 d/4)$$

Taking a union bound gives

$$P\left[\,|\deg(i) - d| \leq \delta d \quad \forall i\right] \quad A$$

$$\geq 1 - 2n\exp(-\delta^2 d/4)$$

$$= 1 - \exp\left(\log(2n) - \frac{\delta^2 d}{4}\right)$$

$$\geq 1 - \exp(-\log(2n)) = 1 - \frac{1}{2n} \qquad \square$$