# Instructions

- The homework is due on **Friday 3/24 at 5pm ET**.
- No extension will be provided, unless for serious documented reasons.
- Start early!
- Study the material taught in class, and feel free to do so in small groups, but the solutions should be a product of your own work.
- This is not a multiple choice homework; reasoning, and mathematical proofs are required before giving your final answer.

# 1 $G(n, p)$ [50 points]

1. (5pts) In the limit as $n$ goes to infinity, how does $(1 - \frac{1}{n})^{n \ln n}$ behave?
2. (5pts) How many labeled graphs on $n$ nodes have exactly $m$ edges, where $0 \leq m \leq \binom{n}{2}$
3. (10pts) Consider a graph $G$ sampled from the $G(n, p)$ model. Prove that conditioned on $G$ having $m$ edges, it is equally likely among all graphs that have $m$ edges.
4. (10pts) Suppose that $p = \frac{c}{n}$ where $c$ is a constant. Prove that the number of vertices of degree $k$ is asymptotically equal to $\frac{c^k e^{-c}}{k!} n$ for any fixed positive integer $k$.
5. (10pts) Consider generating the edges of a random graph by flipping two coins, one with probability $p_1$ of heads and the other with probability $p_2$ of heads. For each pair of nodes add an edge between them if either of the coins comes down heads. Show that this is equivalent to generating a graph $G$ from $G(n, p)$ for an appropriate value of $p$. What is this value $p$?
6. (10pts) Consider $G$ sampled from $G(n, 0.1)$. How does the Central limit theorem apply to the degree of any node in $G$? Specifically, within what range will the degree of a node lie with probability at least 99%?

# 2 Coding [50 points]

Check the Jupyter notebook on our Git repo.

1. (5pts) In the limit as $n$ goes to infinity, how does $(1-\frac{1}{n})^{n\ln n}$ behave?

$$\lim_{n\to\infty}(1-\frac{1}{n})^{n\ln n} \qquad \because e^{\ln x} = x$$

$$= \lim_{n\to\infty} e^{\ln((1-\frac{1}{n})^{n\ln n})} = \lim_{n\to\infty} e^{n\cdot\ln n\cdot\ln(1-\frac{1}{n})}$$

$$\lim_{n\to\infty} n\cdot\ln n \ln(1-\frac{1}{n})$$

$$= \lim_{n\to\infty} n\cdot\ln(n)\cdot[(-n)^{-1} - \frac{(-n)^{-2}}{2} + \frac{(-n)^{-3}}{3} + \cdots] \qquad \text{Taylor series of } \ln(1+\tfrac{1}{x})$$

$$= \lim_{n\to\infty} n\cdot\ln(n)\cdot[-\frac{1}{n} - O(1)]$$

$$= \lim_{n\to\infty} \ln(n)\cdot(-1)$$

$$\therefore \lim_{n\to\infty} e^{n\ln(n)\cdot\ln(1-\frac{1}{n})} = \lim_{n\to\infty} e^{-1\cdot\ln(n)} = \lim_{n\to\infty} e^{-\infty} = 0$$

2. (5pts) How many labeled graphs on $n$ nodes have exactly $m$ edges, where $0 \le m \le \binom{n}{2}$

total number : $2^{\binom{n}{2}}$

# Graphs with exactly m edges : $\binom{\binom{n}{2}}{m}$

3. (10pts) Consider a graph $G$ sampled from the $G(n,p)$ model. Prove that conditioned on $G$ having $m$ edges, it is equally likely among all graphs that have $m$ edges.

Graph with V=n, E=m

$$P(G=H \mid E=m) = \frac{P(E=m \mid G=H)\cdot \Pr(G=H)}{\Pr(E=m)}$$

$$P(G=H) = p^m(1-p)^{\binom{n}{2}-m}$$

$$P(E=m) = \binom{\binom{n}{2}}{m}p^m(1-p)^{\binom{n}{2}-m}$$

$$P(E=m \mid G=H) = 1$$

$$P(G=H \mid E=m) = \frac{1\cdot p^m(1-p)^{\binom{n}{2}-m}}{\binom{\binom{n}{2}}{m}p^m(1-p)^{\binom{n}{2}-m}} = \frac{1}{\binom{\binom{n}{2}}{m}} \Rightarrow \text{equally likely}$$

4. (10pts) Suppose that $p = \frac{c}{n}$ where $c$ is a constant. Prove that the number of vertices of degree $k$ is asymptotically equal to $\frac{c^k e^{-c}}{k!} n$ for any fixed positive integer $k$.

$$P_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

$$E(X) = n \cdot P_k = n \cdot \binom{n-1}{k} p^k (1-p)^{n-1-k} = n \binom{n-1}{k} \left(\frac{c}{n}\right)^k \left(1 - \frac{c}{n}\right)^{n-1-k}$$

$$\because \binom{n-1}{k} = \frac{(n-1)!}{(n-1-k)!\,k!} \approx \frac{(n-1)^k}{k!}$$

$$\therefore P_k = \frac{(n-1)^k}{k!} p^k (1-p)^{n-1-k} = \frac{[(n-1)\cdot p]^k}{k!} (1-p)^{n-1-k}$$

$$\lim_{n\to\infty} n P_k = \lim_{n\to\infty} \frac{(n \cdot p)^k}{k!} \frac{(1-p)^n}{(1-p)^k} = \lim_{n\to\infty} \frac{(n \cdot \frac{c}{n})^k}{k!} \frac{(1-\frac{c}{n})^n}{(1-\frac{c}{n})^k}$$

$$= \lim_{n\to\infty} \frac{c^k \cdot (1-\frac{c}{n})^n}{k!\, o(1)} = \frac{c^k e^{-c}}{k!}$$

5. (10pts) Consider generating the edges of a random graph by flipping two coins, one with probability $p_1$ of heads and the other with probability $p_2$ of heads. For each pair of nodes add an edge between them if either of the coins comes down heads. Show that this is equivalent to generating a graph $G$ from $G(n, p)$ for an appropriate value of $p$. What is this value $p$?

$$\begin{array}{cccc} & & (1-p_1)\cdot p_2 & \\ HH, & HT, & TH, & TT \\ p_1 p_2 & p_1(1-p_2) & & (1-p_1)(1-p_2) \end{array}$$

$P(\text{edge} = 1) = 1 - P(\text{both } T) = 1 - (1-p_1)(1-p_2) = p_1 + p_2 - p_1 p_2$

$p = p_1 + p_2 - p_1 p_2$

In $G(n, p)$, $p = P(\text{generate an edge}) = p_1 + p_2 - p_1 p_2$

$1 - p = P(\text{no edge}) = 1 - (p_1 + p_2 - p_1 p_2)$

$= (1-p_1)(1-p_2) = P(\text{both } T)$

Therefore, 2 models are equivalent,

and the appropriate $p$ is $p_1 + p_2 - p_1 p_2$

6. (10pts) Consider $G$ sampled from $G(n, 0.1)$. How does the Central limit theorem apply to the degree of any node in $G$? Specifically, within what range will the degree of a node lie with probability at least 99%?

The degree in $G(n,p)$ follows a binomial distribution.

$\mu = np = 0.1n$, $\quad var = np(1-p) = 0.09n \quad \sigma = \sqrt{\frac{0.09n}{n}} = 0.3$

$P(\mu - 3\sigma \le k \le \mu + 3\sigma) = 99\%$ according to CLT confidence interval.

$\mu - 3\sigma = 0.1n - 0.9$, $\quad \mu + 3\sigma = 0.1n + 0.9$

$\Rightarrow$ within range $[0.1n - 0.9, 0.1n + 0.9]$