Let G be a graph with $e(G)$ edges. There exists a bipartite subgraph with at least $e(G)/2$ edges.

This is only tight when $e(G) = 0$.

Suppose you have an optimal partition with exactly $e(G)/2$ edges between the parts $V = V_1 \dot\cup V_2$. If there are exactly $e(G)/2$ in the graph induced by (V_1, V_2), then for every $v \in V_1$, $\deg_V v = \deg_{V_2} v$.

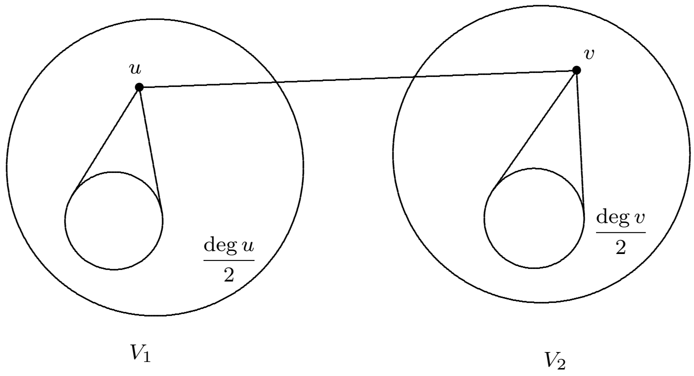
Then,

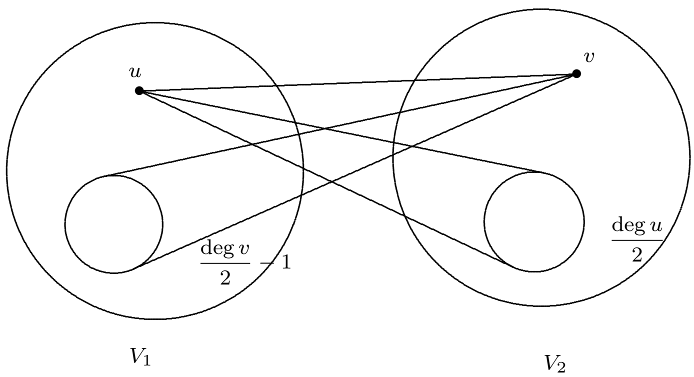
$$\begin{aligned}
e(G) &= \frac{1}{2}\sum_{v\in V} \deg v \\
&= \frac{1}{2}\sum_{v} \deg_{V_1} v + \deg_{V_2} v \\
&= \frac{1}{2}\sum_{v\in V_1} \deg_{V_1} v + \frac{1}{2}\sum_{v\in V_1} \deg_{V_2} v + \frac{1}{2}\sum_{v\in V_2} \deg_{V_1} v + \frac{1}{2}\sum_{v\in V_2} \deg_{V_2} v \\
&= \frac{1}{2}\sum_{v\in V_1} \deg_{V_1} v + \frac{1}{2}\sum_{v\in V_2} \deg_{V_2} v + \frac{e(G)}{2}
\end{aligned}$$

Thus,

$$\begin{aligned}
e(G) &= \sum_{v\in V_1} \deg_{V_1} v + \sum_{v\in V_2} \deg_{V_2} v \\
&\le \sum_{v\in V_1} \frac{\deg v}{2} + \sum_{v\in V_2} \frac{\deg v}{2} \\
&= \frac{1}{2}\sum_{v\in V} \deg v \\
&= e(G)
\end{aligned}$$

So, $e(V_1, V_2) = \dfrac{e(G)}{2}$, $\deg_{V_1} v = \dfrac{\deg v}{2}$ for all $v \in V_1$, and $\deg_{V_2} v = \dfrac{\deg v}{2}$ for all $v \in V_2$.
Suppose $u \in V_1$, $v \in V_2$ and $u \sim v$.





So if there exists an edge, then there exists a partition with strictly more than $e(G)/2$ edges. ■

HOMEWORK: For every graph G, there exists a bipartition with at least $e(G) \cdot \frac{\lfloor n^2/4 \rfloor}{\binom{n}{2}}$ edges.

Probability

Discrete Distribution – a random variable in which integers take on values according to a probability mass function $p(i) = \Pr[X = i]$. Notice that $\sum_i p(i) = 1$.

Expectation: $\mathbb{E}[X] = \sum_i i \cdot p(i)$ represents the average

If you have

$$p(i) = \begin{cases} 1/n, & i \in \{1, 2, \ldots, n\}; \\ 0, & \text{else.} \end{cases},$$

then, $\mathbb{E}[X] = \sum_i i \cdot p(i) = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n}\binom{n+1}{2} = \frac{n+1}{2}$.

If X is a random variable and f is a real-valued function, then $f(x)$ is a random variable.
$\Pr[f(x) = i] = \Pr[x \in f^{-1}(i)] = \sum_{j \in f^{-1}(i)} p(j)$.

$\mathbb{E}[f(X)] = \sum_i f(i)p(i)$

In particular, $\mathbb{E}[X^2] = \sum i^2 p(i)$.
$\mathrm{Var}[x] = \mathbb{E}[(X - E[X])^2]$.
Let $\mathbb{E}[X] = \mu$. Then, $\mathrm{Var}[X] = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2]$
If a is a constant, then $\mathbb{E}[a] = a$.

HOMEWORK: Let $X_1, X_2, \ldots, X_n$ be random variables and $a_1, a_2, \ldots, a_n$ be constants. Then, $\mathbb{E}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i \mathbb{E}[X_i]$. This is called the Linearity of Expectation.

$\mathrm{Var}[X] = \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - (\mathbb{E}[x])^2$.
$\sigma^2 = \mathrm{Var}[X]$

For continuous random variables, you have a probability density for $p(x)$ and $\Pr[X \geq a] = \int_a^\infty p(x)\ dx$.

$\Pr[a \leq X < a+1] = \int_a^{a+1} p(x)\ dx$

$\mathbb{E}[X] = \int_{-\infty}^{\infty} xp(x)\ dx$

$\mathbb{E}[X] = \int_{-\infty}^{\infty} x^2 p(x)\ dx$

Markov's Inequality

Let Z be a positive random variable, $a > 0$. Then, $\Pr[Z > a] \leq \frac{\mathbb{E}[Z]}{z}$.

PROOF: Let Z have a probability mass function p. Then,

$$\begin{aligned}
\mathbb{E}[Z] &= \sum_{i\geq 0} i \cdot p(i) \\
&= \sum_{i=0}^{a-1} i \cdot p(i) + \sum_{i=a}^{\infty} i \cdot p(i) \\
&\geq \sum_{i=0}^{a-1} 0 \cdot p(i) + \sum_{i=a}^{\infty} i \cdot p(i) \\
&\geq 0 + \sum_{i=a}^{\infty} a \cdot p(i) \\
&= a \sum_{i=a}^{\infty} \Pr[Z = i]
\end{aligned}$$

So, $\mathbb{E}[Z] \geq a \Pr[Z \geq a]$.

In the continuous case, $\mathbb{E}[X] = \int_0^a xp(x)\ dx + \int_a^\infty xp(x)\ dx \geq a \int_a^\infty p(x)\ dx$ ■

Random Graphs

$G_{n,p}$ is a random variable.

Label n vertices. Each graph on these variables will occur with probability $p^{e(G)}(1-p)^{\binom{n}{2}-e(G)}$ by flipping a coin where Heads has probability p and Tails has probability $1-p$. You get the edges by flipping the coin $\binom{n}{2}$ times independently, once for each pair of vertices. If the coin comes up Heads, you add the edge between the vertices; if the coin comes up Tails, there is no edge between the vertices.

p can be function of n.

$p = \frac{1}{2}$ means that all labeled graphs are equally likely.

$G_{n,p}$ has $2^{\binom{n}{2}}$ possible values, which are graphs.

THEOREM: Fix $p \in (0,1)$. Then, $\lim_{n\to\infty} \Pr\left[\omega(G_{n,p}) > \frac{2\ln n}{\ln(1/p)}\right] > 0$, where $\omega(H)$ is the order of the largest complete subgraph of H.

PROOF: Let Y_r be the number of cliques of order r in $G_{n,p}$. So, $Y_r = Y_r(G_{n,p})$.

$\Pr\left[\omega(G_{n,p}) > \frac{2\ln n}{\ln(1/p)}\right] = \Pr\left[\vee_{r\geq r_0}\{Y_r \geq 1\}\right] = \Pr[Y_{r_0} \geq 1]$, where $r_0 = \left\lceil \frac{2\ln n}{\ln(1/p)} \right\rceil - 1$.

By Markov,

$$\begin{aligned}
\Pr[Y_{r_0} \geq 1] &\leq \mathbb{E}[Y_{r_0}] \\
&= \mathbb{E}\left[\sum_{S \subseteq V, |S| = r_0} \mathbf{1}_{\{S \text{ clique}\}}\right] \\
&= \sum_{S \subseteq V, |S| = r_0} \mathbb{E}\left[\mathbf{1}_{\{S \text{ clique}\}}\right] \\
&= \sum_S \Pr[S \text{ clique}] \\
&= \sum_S p^{\binom{r_0}{2}} \\
&= \binom{n}{r_0} p^{\binom{r_0}{2}} \\
&\leq \left(\frac{en}{r_0}\right)^{r_0} \left(p^{(r_0-1)/2}\right)^{r_0} \\
&= \left(\frac{en}{r_0} p^{(r_0-1)/2}\right)^{r_0}
\end{aligned}$$

Check that $\frac{en}{r_0} p^{(r_0-1)/2} < 1$ for $r_0 = \frac{2 \ln n}{\ln(1/p)}$ and n large. Therefore, $\Pr[Y_{r_0} \geq 1] \to 0$ and $n \to \infty$. ■