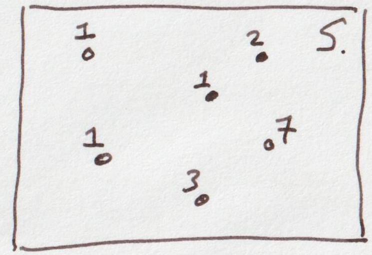
Probabilistic Method (Erdős and Renyi)

Version 1: For any event E, if $Pr(E) > 0$ then $E \neq \emptyset$.



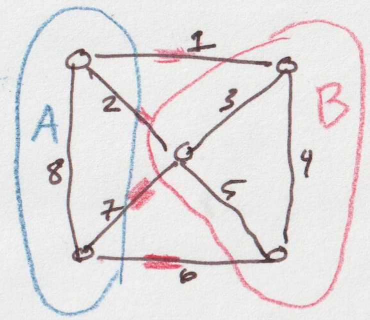
Version 2: For any random variable $X : S \to \mathbb{R}$, there exists at least one $\omega \in S$ such that $X(\omega) \geq E(X)$.

Proof by contradiction: If $X(\omega) < E(X)$ for every $\omega \in S$ then

$$E(X) = \sum_{\omega \in S} Pr(\omega) \cdot X(\omega) < \sum_{\omega \in S} Pr(\omega) \cdot E(X) = E(X) \cdot \sum_{\omega \in S} Pr(\omega) = E(X). \ \lightning$$

Example 1: $G = (V, E)$ is a graph. If (A, B) is a partition of V we say that an edge $e \in E$ crosses (A, B) if e has one endpoint in A and one endpoint in B.

Theorem: Every graph $G = (V, E)$ has a vertex-partition (A, B) such that the number of edges in G that cross (A, B) is at least $|E|/2$.



Proof: For each vertex $v \in V$, toss a coin to decide if v goes into A or v goes into B.

For each edge $e \in E$, define $X_e = \begin{cases} 1 & \text{if } e \text{ crosses } (A,B) \\ 0 & \text{otherwise.} \end{cases}$

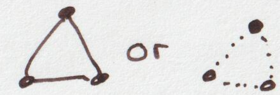
Define X = "the number of edges of G that cross (A, B)"

$$E(X) = E\Big(\sum_{e \in E} X_e\Big) = \sum_{e \in E} E(X_e) = \sum_{e \in E} Pr(X_e = 1) = \sum_{e \in E} \frac{1}{2} = \frac{1}{2} \cdot |E|.$$

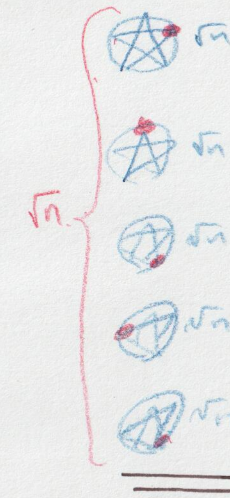
∴ There exist at least one vertex partition (A, B) of G such that at least $\frac{1}{2}|E|$ edges cross (A, B).

Theorem: Every graph with $\overset{n}{6}$ vertices contains a clique of size $\overset{\log_4 n}{\underset{\sqrt{n}}{3}}$ or contains an independent set of size $\underset{\sqrt{n}}{3}$.

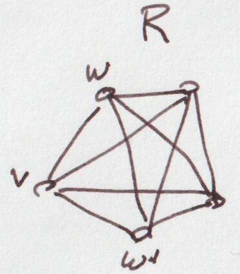
Theorem: For every positive integer n and every integer $K \geq 2\log_2 n + 3$, there exists a graph with no clique of size K and no independent set of size K.



Proof: Let $G = G_{n,\frac{1}{2}}$ with n vertices and, for each pair of vertices v and w, toss a coin to decide if vw is an edge of G.



- Let R be a set of K vertices in G.



- Let A_R = "the vertices in R form a clique"

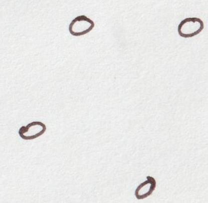
$$Pr(A_R) = \left(\frac{1}{2}\right)^{\binom{K}{2}} = \frac{1}{2^{K(K-1)/2}} \leq \frac{1}{2^{K(2\log_2 n+2)/2}} = \frac{1}{2^{K\log_2 n + K}}$$

$$\leq \frac{1}{2^{K\log_2 n + 1}} = \frac{1}{2}\cdot\frac{1}{n^K}.$$

$$K-1 \geq 2\log_2 n + 2$$

- Let B_R = "the vertices in R form an indep. set"

$$Pr(B_R) \leq \frac{1}{2}\cdot\frac{1}{n^K}.$$



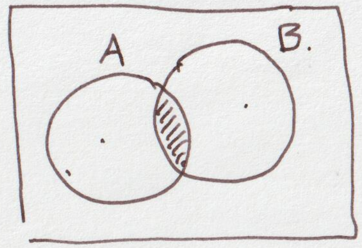
- Let $\{V_1, V_2, \dots V_{\binom{n}{K}}\}$ be the set of K-~~element~~ vertex subset of vertices of G.

- $$Pr\left(\begin{matrix}G \text{ has a clique or}\\ \text{IS of size } K.\end{matrix}\right) = Pr\left(\bigcup_{i=1}^{\binom{n}{K}}(A_{V_i}\cup B_{V_i})\right) \leq \sum_{i=1}^{\binom{n}{K}} Pr(A_{V_i}) + Pr(B_{V_i}) \leq \sum_{i=1}^{\binom{n}{K}} \frac{1}{n^K} = \frac{\binom{n}{K}}{n^K} < 1.$$

$$Pr\left(\begin{matrix}G \text{ does not have}\\ \text{clique or IS of size } K\end{matrix}\right) > 0$$

$\therefore$ there exists n-vertex G with no clique or IS of size K.

Jaccard Distance.

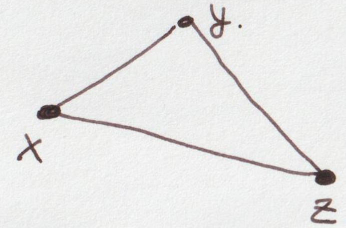


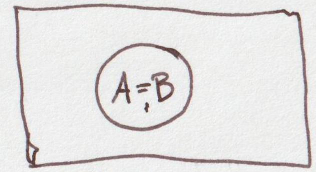
For any two ~~sets~~ non-empty sets A and B.

$$J(A,B) = 1 - \frac{|A \cap B|}{|A \cup B|}$$

Theorem: For any threee non-empty sets X, Y, Z.

$$J(X,Z) \le J(X,Y) + J(Y,Z).$$

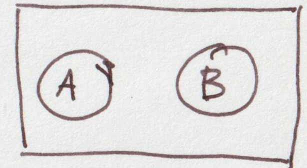




$J(A,B) = 1-1 = 0$

Proof: Let $x_1 \ldots x_m$ be a random permutation of $X \cup Y \cup Z$.

$dist(x,z) \le dist(x,y) + dist(y,z).$



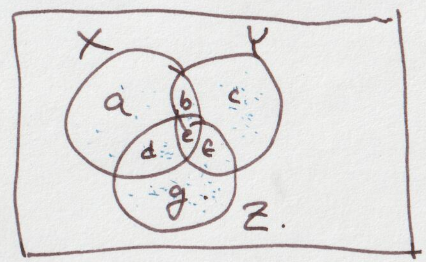
$J(A,B) = 1-0 = 1$

Define $i = \min\{i : x_i \in X\}$. $j = \min\{j : x_j \in Y\}$ $K = \min\{K : x_K \in Z\}$.

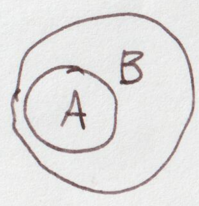
$$Pr(i=j) = \frac{|X \cap Y|}{|X \cup Y|}, \quad Pr(i \ne j) = 1 - \frac{|X \cap Y|}{|X \cup Y|} = J(X,Y).$$

$$Pr(j \ne K) = 1 - \frac{|Y \cap Z|}{|Y \cup Z|} = J(Y,Z).$$

$$Pr(i \ne K) = J(X,Z).$$



Fact 1: If A implies B then $Pr(A) \le Pr(B)$.
$A \subseteq B$



$$J(X,Y) = 1 - \frac{b+e}{a+b+c+d+e+f}.$$

$\vdots$

Fact 2: For any three numbers i, j, and K.

If $i \ne K$ then $i \ne j$ or $j \ne K$.

$$\underbrace{Pr(i \ne K)}_{J(X,Z)} \le Pr(i \ne j \text{ or } j \ne K) \le \underbrace{Pr(i \ne j) + Pr(j \ne K)}_{J(X,Y) + J(Y,Z)}.$$

$J(X,Z) \le J(X,Y) + J(Y,Z).$