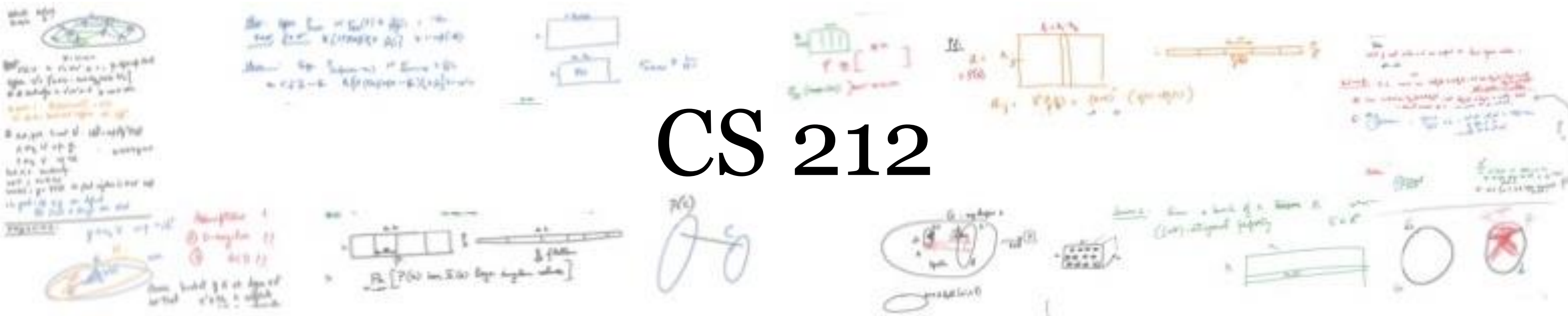


CS 212

Mathematical Foundations of Computer Science

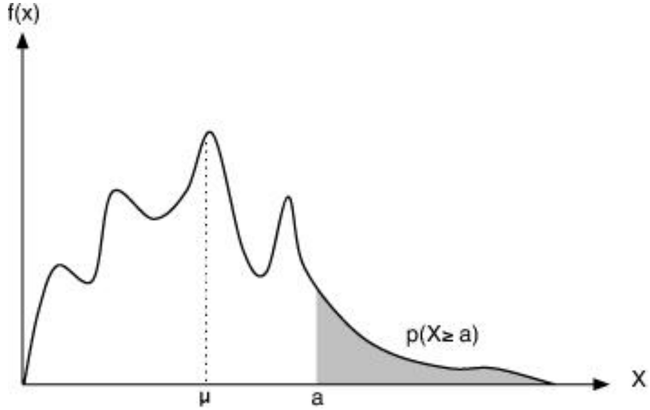
Lecture 16: More Deviation Bounds and Graph Theory

Announcements

1. Midterm scores are out
2. Average score: 12.4
3. Contact TAs Michalis or Vaidehi for grading concerns
4. Eric will hold office hours today 4:00 to 5:30 in Mudd 3011. No office hours Monday 9:00 to 11:00 a.m. Can collect gradebooks during this times.

Markov's inequality

If X is a *non-negative r.v.* with mean E[X], then



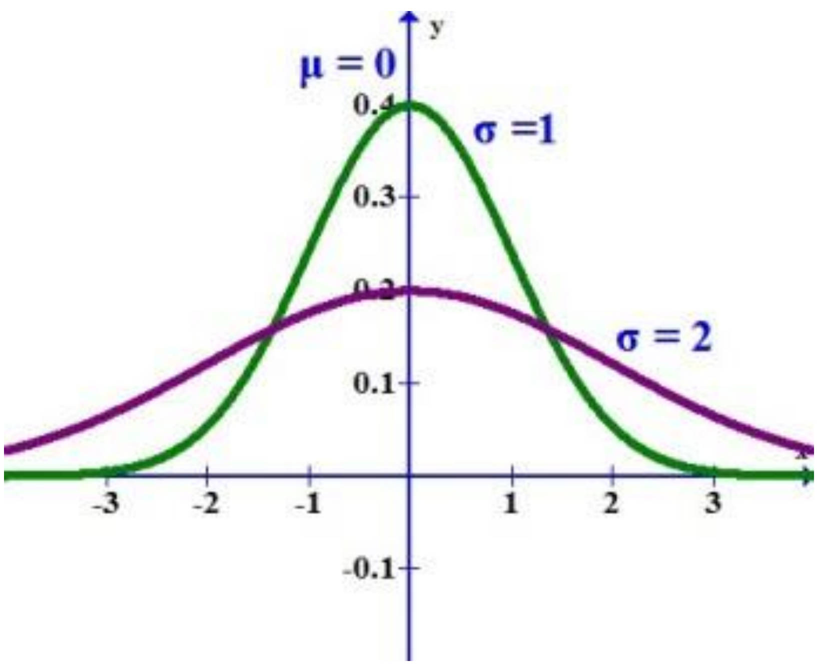
$$\Pr[\, X \geq 2E[X] \,] \;\leq\; ½$$

$$\forall k \geq 1, \qquad \Pr[\, X \geq k \cdot E[X] \,] \;\leq\; 1/k$$

Recap: Variance of random variable X

How much is the deviation from the mean $\mu = E[X]$?

Let $Y = X - E[X]$. What about $E[Y]$? $= 0$



$$Var[X] = E[(X - EX)^2] = E[Y^2]$$

Fact: $\mathrm{Var}[\mathrm{X}] = E[X^2] - (E[X])^2$

Standard deviation $\sigma(X) = \sqrt{\mathrm{Var}(X)}$

Recall X, Y independent

$\Rightarrow E[XY] = E[X]E[Y]$

Properties of Variance

$$\mathrm{Var}[X] = E[(X - EX)^2] = E[X^2] - E[X]^2$$

1. $\mathrm{Var}[aX] = a^2 \times \mathrm{Var}[X]$

$= E[(aX)^2] - (E[aX])^2 = E[a^2X^2] - (aE[X])^2$

$= a^2E[X^2] - a^2E[X]^2$

2. *Independent* random variables X, Y:

$$\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

$= E[(X+Y)^2] - (E[X+Y])^2 = E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2$

$= E[X^2] + 2E[XY] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2$

$= E[X^2] + 2E[X]E[Y] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2$

*remember: $E[X + Y] = E[X] + E[Y]$ for all X, Y

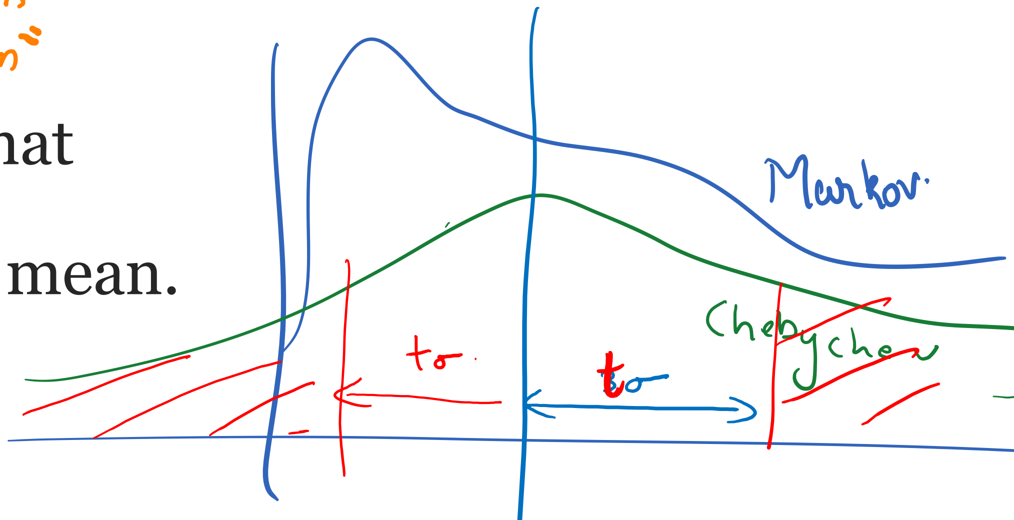
Chebychev Inequality

Bounds deviation around mean on both sides: X *any* random variable with mean $E[X]$, standard deviation $\sigma = \sqrt{\mathrm{Var}[X]}$.

$$\forall t > 0, \underbrace{\Pr[\overbrace{|X - E[X]|}^{\text{distance from } X \text{ to its mean}} \geq t\sigma]}_{\text{"probability that } X \text{ is far from its mean"}} \leq 1/t^2$$

Eg. At most probability 1/9 that

X is more than 3σ away from mean.



Proof of Chebychev's inequality

To prove: $\Pr[|X - E[X]| \geq t\sigma] \leq 1/t^2$

$$= \Pr[(X - E[X])^2 \geq (t\sigma)^2]$$

$$= \Pr[Y \geq t^2\sigma^2]$$

$$= \Pr[Y \geq t^2 E[Y]] \geq 1/t^2$$

by Markov

Set $Y = (X - E[X])^2$
and note that
$E[Y] = Var(X)$
$\sigma^2 = Var(X)$

Probabilities of three similar events

Suppose you have a fair coin. If you flip the coin 20 times, what is the probability you get 9 to 11 heads? 0.497

If you flip the coin 200 times, what is the probability you get 90 to 110 heads? 0.863

If you flip the coin 2000 times, what is the probability you get 900 to 1100 heads? 0.99993

Averaging of Identical RVs
Variance goes down

Suppose $X_1, X_2, X_3, \dots, X_n$ are *independent RVs* (unbiased coins)

$$\mathrm{E}[\mathrm{X_i}] = \Pr[X_i = 1] = \frac{1}{2}, \qquad E\left[X_i^2\right] = \frac{1}{2} * 1 + \frac{1}{2} * 0 = \frac{1}{2}.$$

$$Var[X_i] = E\left[X_i^2\right] - E[X_i]^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

$$Z = \frac{1}{n}(X_1 + X_2 + \cdots + X_n), \qquad E[Z] = \frac{1}{n}\sum E[X_i] = \frac{1}{n} \cdot n \cdot \frac{1}{2} = \frac{1}{2}$$

$$Var[X_1 + X_2 + \cdots + X_n] = n \times Var[X_i] = n/4.$$

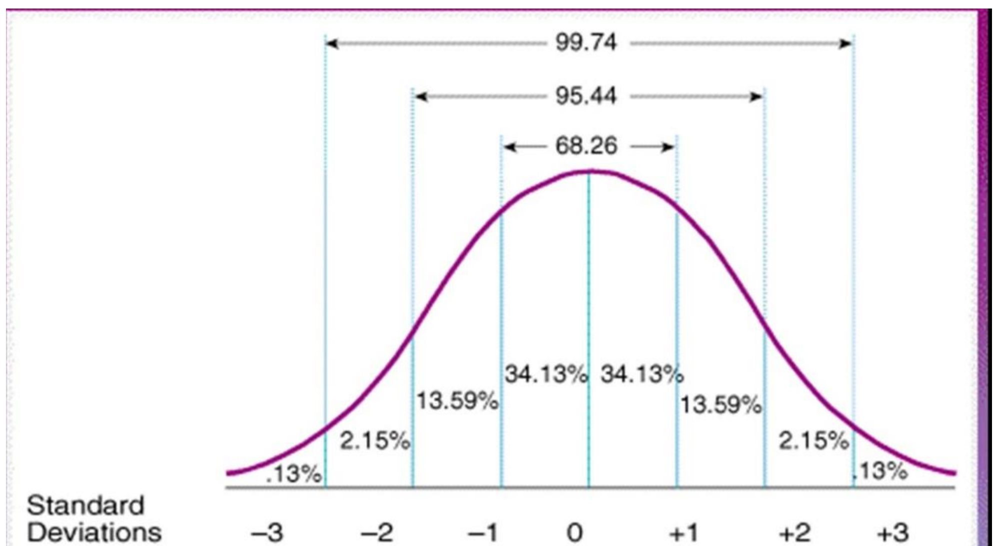
$$Var[Z] = Var\left[\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right] = \frac{1}{n^2} \cdot \frac{n}{4} = \frac{1}{4n}$$

Moral: Variance does down by averaging

Aside: Central Limit Theorem (CLT)

Suppose $X_1, X_2, X_3, \dots, X_n$ are *independent RVs* with $E[X_i] = 0$

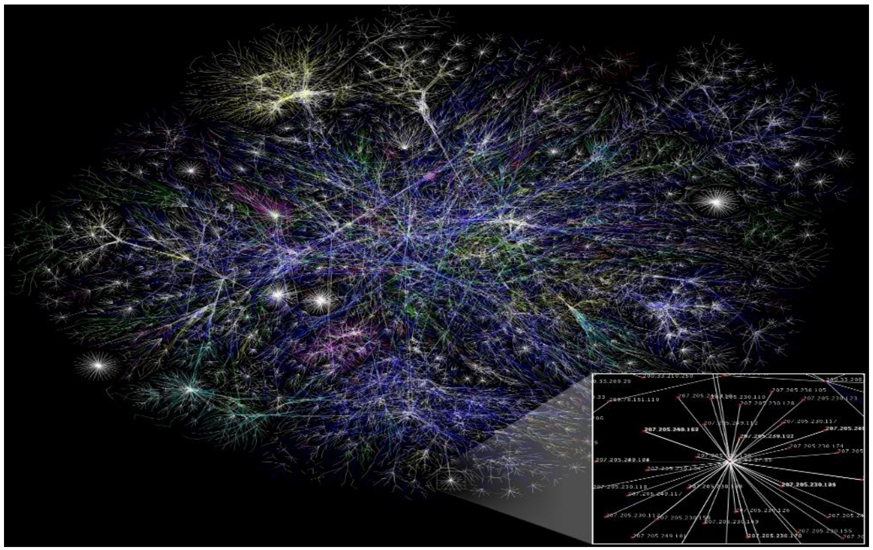
$$Lim_{n\to\infty} \frac{1}{\sqrt{n}}(X_1 + X_2 + \cdots + X_n) \to \text{normal distribution}$$



A fundamental law of nature. Many natural distributions are normal/ Gaussian because of CLT.

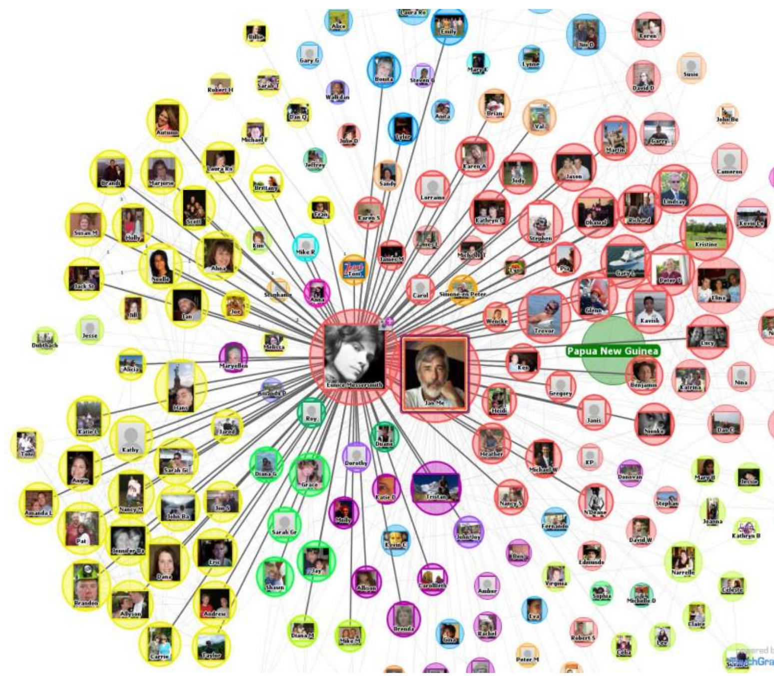
Introduction to Graphs

Networks



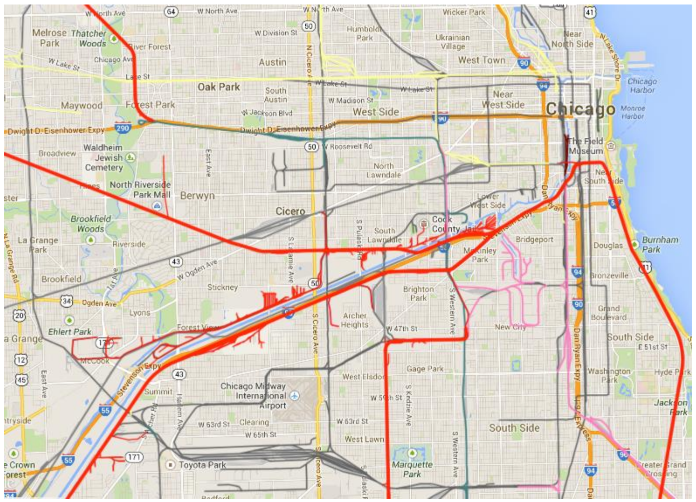
Partial Map of the Internet

Social Networks



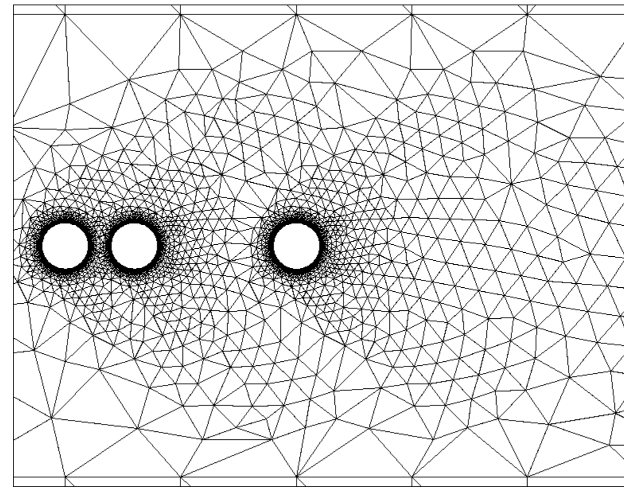
Mobile Networks

Networks in Other Sciences

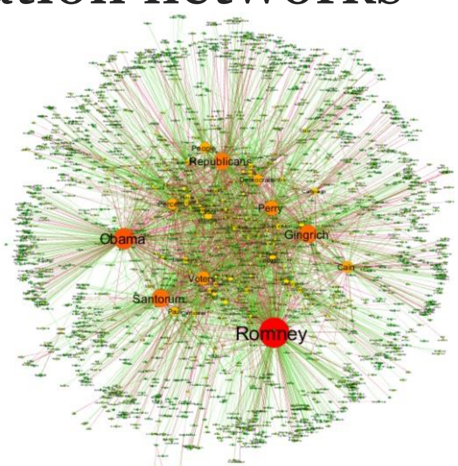


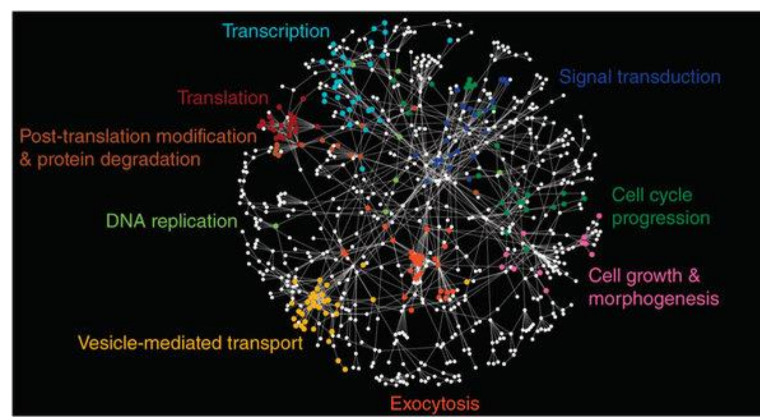
Transportation networks

Scientific computing



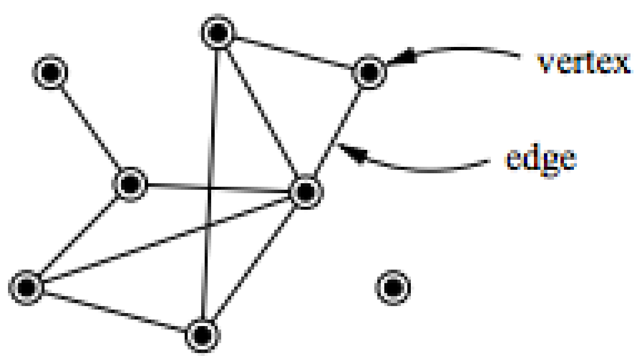
Politics: US 2012 elections

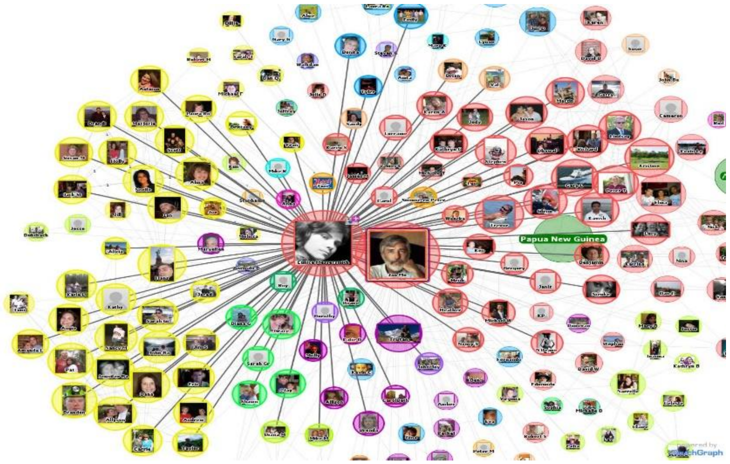




Biology: gene interaction networks

Abstraction: Graph





Representing objects and relations/ connections between objects

- Entities = Vertices
- Relations/ Connections = Edges

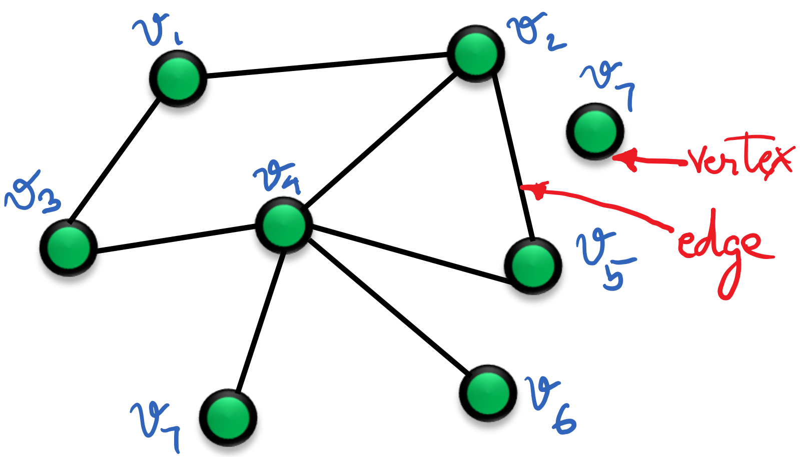
Social Networks :

- Vertices = People.
- Edges = Friendships.

$|V| = n = 8$

$|E| = m = 8$

Graphs



$G(V, E)$

Graph $G = (V, E)$ is a pair of sets

$V = set\ of\ vertices,$

$E \subseteq V \times V$ = set of edges

$V \times V = \{(\mathrm{u}, \mathrm{v}) | \mathrm{u} \in V\ and\ v \in V\}$

Undirected graph: For any $u, v \in V$, if $(u, v) \in E$, then $(v, u) \in E$

$V = \{v_1, v_2, ..., v_8\}$

$E = \{(v_1, v_2), (v_2, v_5), (v_5, v_4), (v_4, v_3), (v_3, v_1), (v_4, v_2), (v_7, v_4), (v_7, v_6)\}$

$|V| = n$ will denote the number of nodes in a graph

$|E| = m$ will denote the number of edges in a graph