February 21

Friday, February 21, 2020 2:56 PM

Comp Sci 212

1. Chernoff bounds

2. Graphs

Chernoff bounds (law of large number, CS)

Thm. If $R_1, \dots, R_n$ independent random variables, s.t. $R_i : S \to [-1, 1]$, $\mathbb{E}[R_i] = 0$, positive u

$\Pr[R_1 + \dots + R_n \geq u\sqrt{n}] \leq e^{-\frac{u^2}{4e}}$

Bell curve

Proof. By other Chernoff

$\Pr[R_1 + \dots + R_n \geq u\sqrt{n}] \leq \frac{\mathbb{E}[e^{t(R_1 + \dots + R_n)}]}{e^{tu\sqrt{n}}}$

anything positive, but <1 choose later

If $u > \sqrt{n}$

$R_1 + \dots + R_n \geq n$

If $R_1, \dots, R_n$ are ind., so is $e^{tR_1}, \dots, e^{tR_n}$

$\mathbb{E}[e^{t(R_1 + \dots + R_n)}] = \mathbb{E}[e^{tR_1}]\mathbb{E}[e^{tR_2}] \dots \mathbb{E}[e^{tR_n}]$

$\mathbb{E}[e^{tR_1}] = \mathbb{E}[1 + tR_1 + \frac{t^2R_1^2}{2} + \frac{t^3R_1^3}{3!} + \dots]$

Taylor series

$= \mathbb{E}[1] + \underbrace{\mathbb{E}[tR_1]}_{0} + \mathbb{E}[\frac{t^2R_1^2}{2}] + \dots$

(linearity of expectation)

$= 1 + \sum_{i=2}^{\infty} \mathbb{E}[\frac{t^iR_1^i}{i!}]$

$\leq 1 + \sum_{i=2}^{\infty} \frac{t^i}{i!}$ $(|R| \leq 1)$

$\leq 1 + t^2 \sum_{i=2}^{\infty} \frac{1}{i!}$ $(t < 1)$

$\leq 1 + t^2 e$

$\leq e^{t^2 e}$ $(1 + x \leq e^x)$

$\mathbb{E}[e^{t(R_1 + \dots + R_n)}] \leq e^{nt^2 e}$

$\Pr[R_1 + \dots + R_n \geq u\sqrt{n}] \leq e^{nt^2 e - tu\sqrt{n}}$

let $t = \frac{u}{2e\sqrt{n}}$

$= e^{\frac{u^2}{4e} - \frac{u^2}{2e}} = e^{\frac{-u^2}{4e}}$

$t \leq 1, \to$

$u \leq 2e\sqrt{n}$

□

Coins - $S' = \{H, T\}$, $S = S' \times \dots \times S'$, uniform dist.

$R : S \to \mathbb{R}$, $R(s) = \#$ of heads in s $(s \in S)$ (n = 100)

$\Pr[R \geq 75] \leq ?$

$R = \mathbb{1}_{E_1} + \mathbb{1}_{E_2} + \dots + \mathbb{1}_{E_n}$, E_i = event that the i^{th} coin is heads

$R_i = \mathbb{1}_{E_i} - \frac{1}{2}$ $\mathbb{E}[R_i] = 0$

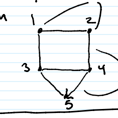
$R \geq 75$ iff $R_1 + \dots + R_n \geq 25 = 2.5\sqrt{100}$

$\Pr[R_1 + \dots + R_n \geq 25] \leq e^{\frac{-2.5^2}{4e}} \approx 0.56$

Graphs Def. (Graph) Pair, Vertex set V, Edgeset $E \subseteq V \times V$

$G = (V, E)$ ← graph; vertices elements of V

Ex.



edges, elements of E

$V = \{1, 2, 3, 4, 5\}$

$E = \{(1,2), (1,3), (2,1), (2,4), (3,1), (3,4), (3,5), (4,2), (4,3), (4,5), (5,3)\}$

no $(5,4)$

10,000 coins (instead of 100)

$\Pr[R_1 + \dots + R_n \geq \underbrace{1000}_{10 \cdot \sqrt{10{,}000}}] \leq$ 0.0001 (Chernoff)

0.025 (Chebyshev)

Comparison to Chebyshev

$\text{Var}[R_1 + \dots + R_n] = \sum_{i=1}^{n} \text{Var}[R_i] = \sum_{i=1}^{n} \mathbb{E}[R_i^2]$ (because $\mathbb{E}[R_i] = 0$)

$\leq n$ (because $|R| \leq 1$)

$\Pr[|R_1 + \dots + R_n| \geq u\sqrt{n}] \leq \frac{n}{u^2 n} = \frac{1}{u^2}$ (Chebyshev)

$e^{\frac{-u^2}{4e}}$ (Chernoff)

$e^{\frac{-u^2}{4e}} = o(\frac{1}{u^2})$

Graphs can represent

Social networks, V = people, E = friendships

Transportation networks V = stops, E = routes

Applications V = people, jobs, E = applications

Types of graphs

Directed graphs - $(v, u) \in E$ does not imply $(u, v) \in E$ (arrows for edges)

(Instagram)

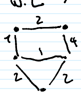
Undirected graphs - $(v, u) \in E$ does imply $(u, v) \in E$ (no arrows for edges)

(Facebook)

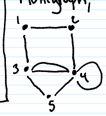
Weighted graphs, $G = (V, E, w)$

$w : E \to \mathbb{R}$

$(w(e),$ length route, strength of a friendship$)$



Multigraph, E a multi-set

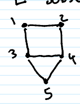


- self-loop

Def. self-loop an edge of the form (v, v)

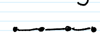
Def. Simple graphs - undirected, unweighted graphs with no self-loops (also not a multi-graph)

E subset of the set of subsets of V of size 2.



$E = \{\{1,2\}, \{1,3\}, \{2,4\}, \{3,4\}, \{3,5\}, \{4,5\}\}$

Common graphs $V = [n]$, (i.e. $\{1, 2, \dots, n\}$)



path

$E = \{\{1,2\}, \{2,3\}, \dots, \{n-1, n\}\}$

$|E| = n - 1$

cycle

$E = \{\{1,2\}, \{2,3\}, \dots, \{n-1, n\}, \{1, n\}\}$

$|E| = n$

Star

$E = \{\{1,2\}, \{1,3\}, \dots, \{1, n\}\}$

$|E| = n - 1$

Complete graph

E = all possible edges

$|E| = \binom{n}{2}$

(tree)

(talk about Wednesday)