Today

Probability.

Concentration.

Minimum Cut

A minimum cut of a graph $G = (V, E)$ is a minimum sized subset of edges that disconnects the graph.

Or, a partition $(S, V - S)$ where $|E \cap (S \times V - S)|$ is minimized.

Observation: Any graph with minimum cut δ has $m \geq \delta n/2$.

Randomized min-cut algorithm:
Collapse a random edge repeatedly until two vertices.
Report corresponding cut.

Minimum Cut: Analysis.

Observation: Any graph with minimum cut δ has $m \geq \delta n/2$.

Analysis: Consider cut C of size δ.
In step i, $n - i$ vertices remain.

$$Pr[\text{edge in } C \text{ collapsed in step } i] \leq \frac{\delta}{\delta(n-i)/2} \leq \frac{2}{n-i}$$

Probability C survives step i is $\geq 1 - \frac{2}{n-i} = \frac{n-i-2}{n-i}$.

The probability that C survives all steps:

$$\frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdot \frac{n-4}{n-2} \cdots \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3}$$

Which is $\geq \frac{2}{(n)(n-1)}$.

Thm: Min-cut process yields a minimum cut with probability $\geq \frac{2}{(n)(n-1)}$.

Implications.

Thm: Min-Cut process yields a minimum cut with probability $\geq \frac{2}{(n)(n-1)}$.

What is the maximum number of minimum cuts? $\binom{n}{2}$.

Algorithm gives a “structural” theorem on the number of min-cuts!

Analysis shows one solution could survives to bound the number of solutions.

Still, there is no proof that one gets a minimum cut. One just hopes.

In 2000, Karger eventually gives a linear time randomized min-cut algorithm.

Introduced sparsification: central to linear time algorithms.

Kawarabashi and Thorup (and...) give a near-linear proof and deterministic algorithm for this.

Probability.

Min Cut Algorithm finds min-cut C with probability $\geq \frac{1}{\binom{n}{2}}$.

Discrete Probability.
Probability space: set of outcomes with probability for each.
Probabilities add to 1.

Outcome Space: cuts.
Sum of probabilities of outcomes is 1.
Total number of minimum cuts is at most $\binom{n}{2}$.

Note: Event A, $A \subseteq \Omega$, $Pr[A] = \sum_{\omega \in A} Pr[w]$.

Tight example graph? The cycle.

Repeat $k = \binom{n}{2}$ times, and choose smallest cut.

Probability of not seeing C, $\leq (1 - 1/k)^k \approx 1/e$.

Using indpendence of trials.

Repeat $k = \binom{n}{2} \ln\left(\frac{1}{p}\right)$ times, and choose best: $(1/e)^{\ln 1/p}$
Probability of not seeing $C \approx p$.

Random Variables

Probability space: Ω, $Pr : \Omega \to \Re$.

$X : \Omega \to \Re$.

Flip n coins. Assign 1 for heads and 0 for tails.

$Pr[\geq n/2 + \rho n]$ heads?

Event view:
$Pr[\{\omega \text{ s.t. } \omega \text{ has } \geq n/2 + \rho n \text{ heads}]$.
One expression: $\frac{1}{2^n} \sum_{i \geq n/2 + \rho n} \binom{n}{i}$.

Random Variable View:
X_i is 1 if coordinate heads, 0 otherwise.

An useful expression for $Pr[X \geq n/2 + \rho n/2]$?

Random Variable View.

$E[X] = \sum_a aPr[X(\omega) = a]$, typically say μ.

Markov: X is a positive random variable, $Pr[X \geq x] \leq \frac{E[X]}{x}$.

Proof:
"Average at least x times the probability of at least x."
$E[X] = \sum_a a \cdot Pr[X = a] \geq \sum_{a \geq x} a \cdot Pr[X = a]$
$\geq \sum_{a \geq x} x \cdot Pr[X = a] = x \cdot \sum_{a \geq x} Pr[X = x] = x \cdot Pr[X \geq x]$ □

Generalized: Strictly increasing function, $g(\cdot)$. $Pr[X \geq x] \leq \frac{E[g(X)]}{g(x)}$.
Proof: "$X > x$" $\equiv$ "$g(X) \geq g(x)$". Use Markov. □

Coin flipping: For n fair coins, let X be number of heads:
$Pr[X \geq 3n/4] \leq \frac{n/2}{3n/4} = 2/3$.

Moments.

$E[X^k]$ is kth moment of the function. $E[X^2]$ is second moment.

Variance or second central moment: $Var(X) = E[(X - E[X])^2]$
Also $Var(X) = E[X^2] - 2E[X]E[X] + E[X]^2 = E[X^2] - E[X]^2$.

The standard deviation $\delta(x) = \sqrt{Var(X)}$.

We want to know how far from average a variable is: $E[|X - E[X]|]$.

$Var(X) = E[(X - E[X])^2]$ in "nicer."
E.g. $Var(X + Y) = Var(X) + Var(Y)$ for independent random variables.

Chebyshev: $Pr[|X - E[X]| \geq x] \leq \frac{Var(X)}{x^2}$

Chebyshev: Proof.

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

Chebyshev: $Pr[|X - E[X]| \geq x] \leq \frac{Var(X)}{x^2}$

Proof: Markov on non-neg variable $(x - E(X))^2$:
$Pr[|X - E[X]| \geq x] = Pr[(X - E[X])^2 \geq x^2] \leq \frac{E[(X-E[X])^2}{x^2} = \frac{Var(X)}{x^2}$ □.

Chebyshev: in terms of standard deviation.
$Pr[|X - E[X]| > t\sigma] \leq \frac{1}{t^2}$.

Coin Flipping: X is number of heads for coin flips.
$X = X_1 + \cdots X_n$ where

$$X_i = \begin{cases} 1 \text{ if flip } i \text{ is heads.} \\ 0 \text{ otherwise.} \end{cases}$$

$E[X_i] = p$. $E[X] = np$
$Var(X_i) = p(1-p)^2 + (1-p)(p)^2 = p(1-p)((1-p)+p) = (1-p)p$
$\implies Var(X) = np(1-p)$

Fair coin: $Pr[X \geq \frac{3}{4}n] \leq \frac{n/4}{(n/4)^2} \leq \frac{4}{n}$.

Comment on Chebyshev.

Fair coin: $Pr[X \geq \frac{3}{4}n] \leq \frac{n/4}{(n/4)^2} \leq \frac{4}{n}$.

Variance: $n/4$, standard deviation $\frac{\sqrt{n}}{2}$.
$n/4$ deviation is $t = \frac{\sqrt{n}}{2}$ std. deviations.

Note: So deviates by t std. deviations with prob $O(1/t^2)$.

Way better at tail ($\leq \frac{4}{n}$) than Markov bound of $2/3$.

Also good for pairwise independent sums of variables.

Sums of independent variables.

X is sum of n independent random variables with mean μ.

$E[X] = \mu n$.

Bound: $Pr[X \geq (1+\varepsilon)E[X]]$ or $Pr[X \leq (1-\varepsilon)E[X]]$.

Central limit theorem: $\frac{X - n\mu}{\sqrt{n\sigma^2}} \to N(0,1)$.

For $X \sim N(0,1)$ $Pr[|X - E[X]| > t] < e^{\frac{-t^2}{2}}$

For 0-1 variables: $Pr[X \geq x] = \sum_{i \geq k} \binom{n}{i} p^i (1-p)^i$.

Central Limit Theorem: $\frac{X - np}{\sqrt{np(1-p)}} \to N(0,1)$.

Which would imply: $Pr[|X - np| > t\sqrt{np(1-p)}] < e^{\frac{-t^2}{2}}$.

Prove this convergence?

Higher Moments?

Chebyshev was better than Markov, so...?

Thm: $Pr[|X - E[X]| \geq x] \leq \frac{E[(X-E[X])^k]}{x^k}$ for even k.

Coin Flips: $E[(X - E[X])]^4 = \frac{1}{16}(\binom{4}{2}\binom{n}{2} + \binom{n}{1})$.

Proof Idea:
Use $Y_i = 2X_i - 1$ to get ± 1 variables with mean zero.
Consider expansion $E[(\sum_i Y_i)^4]$.
Anything with odd power of Y_i has expectation 0.
Others have expectation 1. ..then count. □Yuck.

For coins: $Pr[X \geq 3/4n] = O(1/n^2)$.

In terms of std. dev: $Pr[X - \mu \geq t\sqrt{n}/2] \leq O(1/t^4)$
Versus Gaussians: $\leq e^{\frac{-t^2}{2}}$.

Worse constants. And kind of a mess.

Chernoff: $Pr[X \geq x] = Pr[e^{tX} \geq e^{tx}] \leq \frac{E[e^{tX}]}{e^{tx}}$.

Why the exponential?

Chernoff: $Pr[X \geq x] = Pr[e^{tX} \geq e^{tx}] \leq \frac{E[e^{tX}]}{e^{tx}}$.

Why e^{tx}?

$M_X(t) = E[e^{tx}] = E[\sum_{i \geq 0} \frac{t^i}{i!} x^i] = \sum_{i \geq 0} \frac{t^i}{i!} E[X^i]$.

kth derivative of $M_X(t)$ at $t = 0 \equiv$ "kth moment of X"

For independent X and Y:
$E[e^{t(X+Y)}] = E[e^{tX} e^{tY}] = E[e^{tX}]E[e^{tY}]$.

Chernoff Bounds.

Chernoff: $Pr[X \geq x] = Pr[e^{tX} \geq tx] \leq \frac{E[e^{tX}]}{e^{tx}}$.

Heterogenous Coin Flips: 0/1 variables $X_1, \ldots, X_n$ with $E[X_i] = p_i$.

Note: Compute 4th moment with different p_i now? No.

$X = \sum_i X_i$, $\mu = E[X] = \sum_i p_i$.

$$\begin{aligned}
E[e^{tX}] &= E[e^{tX_1} e^{tX_2} \cdots e^{tX_n}] \\
&= \Pi_i E[e^{tX_i}] \text{ by independence} \\
&= \Pi_i p_i e^t + (1 - p_i) e^0 \\
&= \Pi_i (1 + p_i(e^t - 1)) \\
&\leq \Pi_i e^{p_i(e^t - 1)} \quad \text{since } (1 + x) \leq e^x. \\
&= e^{\sum_i p_i (e^t - 1)} = e^{\mu(e^t - 1)}.
\end{aligned}$$

Chernoff.

"Moments" Bound: $E[e^{tX}] \leq e^{\mu(e^t - 1)}$.

Recall $M_X(t) = \sum_{i \geq 0} \frac{t^i}{i!} E[X^i]$.

Can use t to pick out which range of moments to use for bound!

Markov: $Pr[X > a] \leq \frac{E[e^{tX}]}{e^{ta}}$

$Pr[X \geq (1 + \delta)\mu] \leq \frac{e^{\mu(e^t - 1)}}{e^{t(1+\delta)\mu}}$.

Some bounds: picking t.

Exponential expection plus Markov:

$$Pr[X \geq (1 + \delta)\mu] \leq \frac{e^{\mu(e^t - 1)}}{e^{t(1+\delta)\mu}}.$$

If $\delta > 0$, choose $t = \ln(1 + \delta)$ (calculus), to get

$$Pr[X \geq (1 + \delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$$

If $0 < \delta < 1$, $\frac{e^{\delta}}{(1+\delta)^{1+\delta}} \leq e^{-\delta^2/3}$:

$$Pr[X \geq (1 + \delta)\mu) \leq e^{-\frac{\delta^2 \mu}{3}}.$$

If $R > 6\mu$, $1 + \delta > 6$, $e^{\delta} < e^{1+\delta}$, and $e/6 \leq 1/2$,

$$Pr[X \geq R] \leq \left(\frac{1}{2}\right)^R.$$

Lower tails bound.

$Pr[X \leq (1 - \delta)\mu] \leq e^{-\mu\delta^2/2}$ for $0 < \delta < 1$

Chernoff, Coins, and Central Limit Theorem.

For $\delta < 1$,

$Pr[X \geq (1 + \delta)\mu) \leq e^{-\frac{\delta^2 \mu}{3}}$.

X is sum of n independent fair coins.

$\mu = n/2$. $\sigma = \sqrt{n}/2$.

Number of std. dev: $t = \delta\mu/\sigma = \delta n/\sqrt{n} = \delta\sqrt{n}$.

$t^2 = \delta^2 n = 2\delta^2 \mu$.

$Pr[X - E[X] \geq t\sigma] \leq e^{\frac{-t^2}{6}}$.

Exponential dropoff in t^2, like random variable from $N(0, 1)$.

For coins:

$Pr[X \geq \frac{3}{4}n] \leq e^{-\frac{(\frac{1}{2})^2 \frac{n}{2}}{3}} = e^{-\frac{n}{24}}$.

Versus Chebyshev: $\frac{4}{n}$.

Graph Sparsification.

Given graph G, is there a weighted graph with fewer edges so that forall cuts, the cutsize is within $1 + \varepsilon$ of the original?

Process: include each edge in G with probability p with weight $\frac{1}{p}$.

Expected number of edges: pm.

Let $\tilde{C}$ be number of sampled edges in cut C.

$E[\tilde{C}] = \sum_{e \in C} p = p|C|$, and their expected weight is $|C|$.

Good algorithm in expectation for any cut. Could be off.

How small can we set p?

Set $p = 9 \log n / \varepsilon^2 \delta$ where δ is min-cut size.

For a cut C or size k.

$Pr[|\tilde{C} - pk| > \varepsilon pk] \leq 2e^{\frac{-pk\varepsilon^2}{3}} \leq 2e^{-\frac{3k \ln n}{\delta}}$.

This is less than $2/n^3$, since $k \geq \delta$.

Union bound on 2^n cuts is problematic.

Using Karger's process.

For a cut C or size k.
$Pr[|\tilde{C} - pk| > \epsilon pk] \le 2e^{\frac{-3k\ln n}{\delta}}$.

Thm: There are at most $n^{2\alpha}$ cuts of size $\alpha\delta$.

Proof Idea:
What is probability that a cut of size $\alpha\delta$ survives Karger's process? □.

$$Pr[\text{any bad cut}] \le \sum_{cutsC} Pr[\tilde{C} \text{ is bad}]$$
$$\le \sum_{\alpha \ge 1} \sum_{C:|C|<\alpha\delta} Pr[\tilde{C} \text{ is bad}||C| < \alpha\delta] \le \sum_{\alpha\ge 1} n^{2\alpha} 2e^{-3\alpha \ln n}$$
$$\le \sum_{\alpha \ge 1} 2n^{-\alpha} \le \frac{4}{n}$$

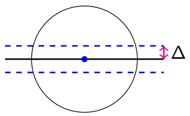
Note: Fancy stuff allows non-uniform sampling to get yet sparser graph.
$O(n \log n/\epsilon^2)$ and even $O(n/\epsilon^2)$ edges.

Sampling with gaussian tails.

z is uniformly random unit vector.
Random point on the unit d dimensional sphere. $E[\sum_{i\in[k]} z_i^2] = \frac{k}{d}$.

Claim: $\Pr[|z_1| > \frac{t}{\sqrt{d}}] \le e^{-t^2/2}$

Sphere view: surface "far" from equator defined by e_1.



$|z_1| \ge \Delta$ if
$z \ge \Delta$ from equator of sphere.
Point on "Δ-spherical cap".

Area of caps
$\le$ S.A. of sphere of radius $\sqrt{1-\Delta^2}$
$\propto r^{d-1} = \left(1-\Delta^2\right)^{(d-1)/2}$
$\propto \left(1 - \frac{t^2}{d}\right)^{(d-1)/2} \approx e^{\frac{-t^2}{2}}$
Constant of $\propto$ is unit sphere area. □

$\Pr[\text{any } z_i^2 > (2\log d)E[z_i^2]]$ is small.

Concentration.

Complicated to count distribution and moments.

The moment generating function is nice.

Get effect of moments and concentration.

Other views: KL distance.