Pravděpodobnostní metoda

Nástroje a odhady

T(Union-bound**):** $P[\bigcup A_i] \leq \sum_{i=0}^{n} P[A_i]$.

T(Markov**):** X nezáporná, $a > 0$: $P[X > \alpha E[X]] < \frac{1}{\alpha}$.

T(Čebyšev**):** $P[|X - E[X]| \geq \lambda\sigma] \leq \frac{1}{\lambda^2}$.

P: $\sigma^2 = \mathrm{Var}[X] = E[(X - E[X])^2] \geq \lambda^2\sigma^2 P[|X - E[X]|] \geq \lambda\sigma$.

T(Černov**):** Nechť S_n je součtová náhodná proměnná pro n mincí s hodnotami $+1, -1$. Pak: $P[S_n > a] < e^{-a^2/2n}$.

P: Vytvoříme náhodnou proměnnou $Y = e^{uX}$.
$E[Y] = \prod_{i=1}^{n} E[e^{uX_i}] = \frac{e^u + e^{-u}}{2}^n \leq e^{nu^2/2}$. Použitím Markova pak $P[X \geq t] = P[Y \geq e^{ut}] \leq E[Y]/e^{ut} \leq e^{nu^2/2 - ut}$. Zvol $u = t/n$.

Základní použití

T(Hrubý Ramsey**):** $\forall k \geq 3 : R(k,k) > 2^{k/2-1}$.

P: Vezmi $G(n, 1/2)$, použij union-bound na jev „G obsahuje kliku nebo nezávislou velikosti k".

D: $m(k)$ bude označovat minimálni počet hran k-unif. hypergrafu, který není 2-obarvitelný. $m(2) = 3$ – Fano.

O: $m(3) \geq 7$.

P: 6 hran hypergrafu. Nejvýš 6 vrcholů: rozeber. Aspoň 7: dva nejsou spojené hranou, tak je spoj do jednoho vrcholu.

T(Odhad $m(k)$**):** $m(k) \geq 2^{k-1}$.

P: Vezmi menší náhodný hypergraf. Zvol náhodné obarvení, spočti pravděpodobnost že existuje jednobarevná (a tedy špatná) hrana.

T: Pokud $\binom{n}{k}(1 - 2^{-k})^{n-k} < 1$, tak existuje turnaj s vlastností, že každých k vrcholů má společného vst. souseda.

P: Union bound.

Náhodné permutace

D: Systém množin F *má průnik* pokud mají průnik všechny dvojice množin z F.

T(Slunečnicové lemma**):** Nosná množina $n \geq 2k$. Pokud má systém k-prvkových množin F průnik, tak $|F| \leq \binom{n-1}{k-1}$.

L(Lemma pro lemma**):** $X = Z_n$, zadefinuj $A_s = \{s, s+1, \dots s+k-1\}$ pro $s \in X$, n buď větší jak $2k$. Pak F s průnikem obsahuje nanejvýš k z množin A_s.

P: Když tam patří jedna A_i, tak tam patří nanejvýš ty, co se s ní pronikají $(2k-2)$, a ty se dají rozdělit ještě do disjunktních dvojic.

P(Slunečnicové lemma**):** Budeme volit náhodné s a náhodnou σ, co přepermutuje A_s. Díky lemmatu je nejvýš k obraz $\sigma(A_s)$ uvnitř F, čili $P[\sigma(A_s) \in F] \leq k/n$. Každá k-tice má ale stejně permutací, ve kterých je za sebou, čili taky $P[\sigma(A_s) \in F] = |F|/\binom{n}{k}$. Počítej dvěma způsoby.

D: Definujme $n(k,l)$ jako maximální n s.t. $\exists A_{[1,n]} \,\&\, B_{[1,n]}$ s.t.

1. $|A_i| = k, |B_i| = l$,
2. $A_i \cap B_i = \emptyset$,
3. $i \neq j \to A_i \cap B_j \neq \emptyset$.

T(Bollobásovo lemma o průnicích**):** $n(k,l) = \binom{k+l}{k}$.

P: Náhodně přepermutuj prvky. Zadefinuj jev: $U_i \equiv$ všechny prvky A_i přechází všechno z B_i. $P[U_i] = 1/\binom{k+l}{k}$. U_i a U_j nemohou nastat zároveň, čili $1 \geq P[\bigcup U_i] = \sum_i P[U_i] = n\binom{k+l}{k}$.

R: Spernerova věta.

Linearita střední hodnoty

T: Existuje turnaj, který má alespoň $n!/2^{n-1}$ Ham. cest.

P: Rozděl jev na jednotlivé permutace.

T: Každý graf s m hranami obsahuje bipartitní graf s $m/2$ hranami.

P: Derandomizací nebo i deterministicky.

Alterace

T(Slabý Turán**):** $\alpha(G) \geq n/2d$.

P: Vezmi náhodnou podmnožinu S, každý vrchol vlož s pravd. p. $E[|S|] = np$. $E[||S||] = \frac{ndp^2}{2}$. $|S| - ||S||$ odpovídá tomu, že smažeme z každé hrany jeden vrchol. Zvolme p, ať maximalizujeme $E[|S| - ||S||]$.

T(Další Ramsey**):** $\forall k, n : R(k,k) > n - \binom{n}{k}2^{1-\binom{k}{2}}$.

P: Vezmi náhodné obarvení grafu na n vrcholech. Střední hodnota počtu jednobarevných k-klik je $\binom{n}{k}2^{1-\binom{k}{2}}$. Existuje obarvení, co to dosahuje, z každé kliky smaž jeden vrchol.

T(Erdös, obvod neurčuje barevnost**):** $\forall k, l \exists G$ s.t. $\chi(G) > k \,\&\, g(G) > l$.

P: Nastav $\varepsilon = 1/2l, p = n^{\varepsilon - 1}$. Střední hodnota počtu cyklů délky l je

$$E[X] \leq \sum_{i=3}^{l} n^{\varepsilon i} = o(n).$$

Zvol n s.t. $E[X] < n/4$. To nám dá $E[X] < 1/2$, použij Markova na $P[X \geq n/2] < 1/2$.

Teď počitej barevnost pomocí největší nezávislé. (Barevné třídy jsou nezávislé.) a buď $\lceil 3/p \ln n \rceil$, pak máme

$$P[\alpha \geq a] \leq \binom{n}{a}(1-p)^{\binom{a}{2}} \leq n^a e^{-p\binom{a}{2}} = e^{(\ln n - p(a-1)/2)a}.$$

To jde do nuly, tedy zase máme $P[\alpha] < 1/2$. Smaž jeden vrchol ze všech krátkých cyklů. Zbyde půlka vrcholů, barevnost odhadni velikostí α.

T: Existuje množina S n vrcholů v jednotkovém čtverci, kde největší trojúhelník má obsah alespoň $T(S) \geq 1/(100n^2)$.

P: Zvol tři body A, B, C uniformně z množiny všech bodů ve čtverci. Nechť je $|AB| = x$. Chceme odhadnout $P[\mu(PQR) < c]$. Nejdřív najdeme vzorec pro $P[b \leq x \leq b + \Delta b]$. Kdyby to platilo, tak jsou A a B od sebe někde mezi dvěma kruhy, čili $P[b \leq x \leq b + \Delta b] \leq \pi(b + \Delta b)^2 - \pi b^2$. V limitě $P[b \leq x \leq b + db] \leq 2\pi b db$.

No a pokud vzdálenost je kolem b, pak výška musí být nejvýš $2c/b$, což znamená, že R je v pásu šířky $4c/b$ a šířky $\sqrt{2}$. Nyní zintegrujeme možnosti podle b:

$$\int_0^{\sqrt{2}} (2\pi b)(4\sqrt{2}c/b) db = 16\pi c.$$

Zvolíme $c = 1/(100n^2)$ a spočítáme střední hodnotu počtu trojic, které splňují, že jejich obsah je víc než c. Je to $E[T] \leq \binom{2n}{3}(0.6n^{-2}) < n$. Alterujeme.

T(Komlós, Pintz, Szemerédi**):** $T(S) = \Omega(\frac{\log n}{n^2})$.

Dependent Random Choice

L(Základní DRC**):** G graf, $|G| = n, d$ prům. stupeň. Pokud existuje přirozené číslo t takové, že $\frac{d^t}{n^{t-1}} - \binom{n}{r}\frac{m}{n}^t \geq a$, pak G obsahuje podmnožinu vcholů $|U| = a$ s.t. $\forall r$ vrcholů z ní má alespoň m společných sousedů (značíme $|N_s(R)| = m$.

P: Zvol t vrcholů T náhodně s opakováním. $A = N_s(T)$, $X = |A|$. Použijeme konvexitu z^t pro nerovnost: $E[X] = \sum_{v \in V(G)} \left(\frac{|N_s(v)|}{n}^t\right) = n^{-t}\sum\left(|N_s(v)|^t\right) = n^{-t}n(\sum |N_s(v)|/n)^t = \frac{d^t}{n^{t-1}}$.
Y buď počet špatných podmnožin A, tj. s velikostí r ale s méně jak m společnými sousedy. $P[S \subseteq A] = (|N(S)|/n)^t$, a tedy $E[Y] < \binom{n}{r}(m/n)^t$.
Alteruj, $E[X - Y] \geq a$, smaž špatné vrcholy.

L: Pokud máme podmnožinu $U \subseteq V(G), |U| = a$ takovou, že každá r prvková podmnožina má alespoň $a + b$ společnych sousedů, tak sem umíme vnořit H bipartitní s partitami velikosti a, b, kde jedna z nich má $\Delta = r$.

P: Indukcí, společných sousedů je $a + b$ a to stačí pro vnořování vrcholů s max. stupněm b (společní sousedé mohou být opět v U, to nevadí).

T: H bipartitní graf s $\Delta = r$ v jedné partitě $\to ex(n, H) \leq cn^{2-1/r}$.

P: Vol $m = a + b$, $t = r$, $c = \max\left(a^{1/r}, \frac{3a+b}{r}\right)$, $||G|| > ex$. Pak $d \geq 2cn^{1-1/r}$, odhadni $r!$ pomocí $(r/e)^r$ a konečně:

$$(2c)^r - \frac{n^r}{r!}\frac{(a+b)^r}{n}^r \geq c^r \geq a.$$

Pro celkové vítězství aplikujeme předchozí lemma.

Rozptyl

T: $\forall m \geq 1 : \binom{2m}{m} \geq 2^{2m}/(4\sqrt{m} + 2)$.

P: Vezmi náhodný součet $2m$ 0/1-proměnných s $p = 1/2$. $E[X] = m, \mathrm{Var}[X] = m/2$. Čebyšev dává: $P[|X - m| < \sqrt{m}] \geq 1/2$. Pravděpodobnost, že X nabyde přesně $m + k$ pro $|k| < \sqrt{m}$, je nejvýš $\binom{2m}{m-k}2^{-2m} \leq \binom{2m}{m}2^{-2m}$. Celkem tedy:

$$1/2 \leq \sum_{|k| < \sqrt{m}} P[X = m + k] \leq (2\sqrt{m} - 1)\binom{2m}{m}2^{-2m}.$$

L(O malém rozptylu.**):** Nechť X_i je posloupnost proměnných, pro které platí $\lim_{n\to\infty} \frac{\mathrm{Var}(X_n)}{E[X_n]^2} = 0$. Pak $\lim_{n\to\infty} P[X_n > 0] = 1$.

L: Pokud $k(n)$ je funkce splňující $\lim_{n\to\infty} \binom{n}{k(n)}2^{-\binom{k(n)}{2}} = \infty$, pak platí: $\lim_{n\to\infty} P[\Omega(G(n, 1/2)) > k(n)] = 1$.

P: V podstatě dokazujeme pro $k(n) = (2-\varepsilon)\log n$. $E(n,k) \equiv$ střední počet k-klik v $G(n,1/2)$. $k(n) < 2\log n$, protože $E(n, 2\log n) \to 0$. $k(n) > 3/2\log n$, protože
$\log E(n, 3/2\log n) = 3/2\log^2 n - 9/8\log^2 n \to \infty$.

Budeme chtít použit lemma o malém rozptylu. X (X_n) buď proměnná počítající počet $k(n)$-klik. Její rozptyl je součet kovariancí pro každou $k(n)$-tici vrcholů zvlášť, navíc to ještě sečteme podle velikostí průniků t, $C(t)$ bude daný součet pro průniky velikosti t.

Kovarianci dvou k-tic S,T mohu odhadnout pomocí $E[X_S X_T] \leq 2^{\binom{t}{2}-2\binom{k}{2}}$. $C(t)$ pak kombinatoricky jako

$$C(t) \leq \binom{n}{k}\binom{k}{t}\binom{n-k}{k-t}E[X_S X_T].$$

Budeme nyní chtít ulimitit $\sum_{t\geq 2}^{k} \frac{C(t)}{E[X]^2}$ do nuly, trikem ji rozložíme na dvě sumy: do $k/2$ a nad $k/2$.

První část půjde do 0, pokud $k < 2\log n$. Výpočet:

$\frac{C(t)}{E[X]^2} \leq \frac{\binom{k}{t}\binom{n-k}{k-t}}{\binom{n}{k}} 2^{\binom{t}{2}} \leq k^2 t(n^{\underline{t}} t!)^{-1} 2^{t^2/2} \leq (k^2 2^{-k/2} 2^{t/2})^t$.

Sumu pak omez jako geometrickou řadu. Druhá část počítá jen s jednou mocninou $E[X]$. Použijeme, že $3/2\log n \leq k$.

$$\begin{aligned}\frac{C(t)}{E[X]} &\leq \binom{k}{k-t}\binom{n}{k-t} 2^{\binom{t}{2}-\binom{k}{2}} \leq (kn2^{-(k+t-1)/2})^{k-t} \\ &\leq (2^{\log k + 2k/3 - 3k/4})^{k-t}.\end{aligned}$$

Opět omez geometrickou řadou a použij fakt, že $E[X]$ jde do nekonečna, čili to platí i pro druhé mocniny.

D: H graf je *balancovaný*, pokud jeho hustota je větší nebo rovna všem jeho podgrafům.

T: Pokud je H balancovaný, ρ jeho hustota, pak funkce $n^{-1/\rho}$ je jeho prahová funkce.

D: Množina $R \subseteq [n]$ bude mít *různé sumy*, pokud její všechny sumy jsou různé. $f(n)$ je maximální velikost takové R.

T(Menší zlepšení rozdílných sum**):** $f(n) \leq \log_2 n + (1/2)\log_2\log_2 n + O(1)$.

P:Máme zadané $r_{[1,k]}$, vezmi náhodnou sumu (každé číslo vezmi s pravd. $1/2$). Spočítej střední hodnotu, rozptyl. (Indikátory jsou nezávislé.) Čebyšev nám dává: $P[|X-\mu| > \lambda n\sqrt{k}/2] \leq \lambda^{-2}$, čili i $1-\lambda^{-2} \leq P[|X-\mu| \leq \lambda n\sqrt{k}/2]$. Všimneme si dále, že jedné konkrétní hodnoty nabude X s pravděpodobností 0 n. 2^-k, čili $P[\ldots] < 2^{-k}(\lambda n\sqrt{k}+1)$. Spojíme nerovnosti a upravíme.

Lovászovo lokální lemma

T(Symetrické LLL**):** Buď $A_1, \ldots, A_n$ s.t. $P[A_i] \leq p$ a všechny výstupně v závislostním orientovaném grafu jsou nanejvýš d. Pokud $ep(d+1) \leq 1$, pak

$$P[\bigcup_{i=1}^{n} \overline{A_i}] > 0.$$

T(Asymetrické LLL**):** Buď A_i, nastav k nim reálné 0-1-proměnné x_i. Pokud platí

$$P[A_i] \geq x_i \prod_{(i,j)\in E}(1-x_j), \text{pak}$$

$$P[\bigcap_i \overline{A_i}] \geq \prod_i (1-x_i) > 0.$$

P:Je třeba počítat A_i (to, co nechceme, aby se stalo) za předpokladu, že už se nějaka podmnožina S jevů nestala. Indukcí podle S dokazujeme: $P[A_i | \bigcap_{j\in S}\overline{A_j}] \leq x_i$.

Základ indukce je hned z předpokladů. Dále rozdělíme S na zajímavé (S_1) a nezávislé na A_i (S_2). Rozepíšeme:

$$P[A_i | \bigcap_{j\in S}\overline{A_j}] = \frac{P[A_i \cap \bigcap_{j\in S_1}\overline{A_j} | \bigcap_{l\in S_2}\overline{A_l}]}{P[\bigcap_{j\in S_1}\overline{A_j} | \bigcap_{l\in S_2}\overline{A_l}]}.$$

Nahoře můžeme odhadnout pomocí $P[A_i|\ldots]$, což dá x_i z předpokladů věty, neboť S_2 byly nezávislé na A_i. Dole použijeme indukci: rozdělíme jmenovatel na součin nezávislých jevů, které všechny mají menší S: $P[\bigcap_{j\in S_1}\overline{A_j} | \bigcap_{l\in S_2}\overline{A_l}] = P[A_{j_1} | \bigcap_{l\in S_2}\overline{A_l}] \cdot P[A_{j_2} | \overline{A_{j_1}} \cap \bigcap_{l\in S_2}\overline{A_l}] \cdots \leq (1-x_{j_1})(1-x_{j_2})\cdots(1-x_{j_r})$.

P(Symetrické LLL**):**Nastav x_i na $1/(d+1)$. Pak $x_i \prod_{(i,j)\in E}(1-x_j) = (d+1)^{-1}(1-1/(d+1))^d \leq (e(d+1))^{-1} \geq p$ a můžeme použít asymetrické lemma.

T(Barvení hypergrafů**):**H hypergraf, kde každá hrana má alespoň k vrcholů a proniká nanejvýš d hran. Pokud $e(d+1) \leq 2^{k-1}$, tak H je 2-obarvitelné.

P:Použij symetrickou verzi LLL na jev, že hrana je jednobarevná.

T(Cykly dlouhé násobek**):**D digraf, minimální výstupeň δ, maximální vstupeň Δ. Pak platí-li $k \leq \frac{\delta}{1+\ln(1+\delta\Delta)}$, tak D obsahuje orientovaný cyklus dělitelný k.

P:Přiřaď vrcholům čísla a jev A_v bude jev, kdy $\exists z \in N_+(v)$ s.t. $k(z) = k(v) + 1 \bmod k$. A_v pak nezávisí na barvě v, jen na barvách $N_+(v)$, čili závisí jen na takových A_w, kteří jsou buď přímo následníci v nebo mají s v společného následníka. Těch je celkem $\delta\Delta$, čili $ep(d+1) \leq e(1-1/k)^\delta(\delta\Delta+1) \leq 1$.

D:Obarvěme reálná čísla k barvami. $S \subseteq R$ je *duhová* $\equiv c(S) = [k]$.

T:$\forall k \exists m$ s.t. $\forall S \subseteq R, |S| = m$, umíme nabarvit reálná čísla k barvami tak, aby jakékoli posunutí m bylo duhové.

P:Konečně mnoho posunutí: jedno posunutí závisí na méně než m^2 jiných, pravděpodobnost neduhovosti je $k(1-1/k)^m$. Zvol m správně.

Nekonečně mnoho posunutí: kompaktnost.

Černov a spol.

O(Černov pro $\{0,1\}$ nez. p.**)**:$P[X \geq E[X]+t] \leq e^{-2t^2/n}$ a symetricky pod střední hodnotou.

O(Černov pro $(0,1)$ nez. p.**)**:$P[X \geq E[X]+t] < \exp(\frac{-t^2}{2(\sigma^2+t/3)})$ a symetricky pod střední hodnotou.

D(k-pakování**):**Máme rodinu $F, |F| = m$ podmnožin z $[n]$, hledáme max. podrodinu, že každý element je nejvýše k-krát obsažen.

k-pakování můžeme reprezentovat jako matici $n \times m$ s hodnotami 0/1, hledáme největší podmnožinu sloupců, co se nasčítají na nejvýš k v každé složce.

Vyřeš lineárně, zaokrouhluj podle j-té složky optima (x_j^*) na 1 nebo 0. $E[y] = OPT^*, E[(Ay)_i] \leq k$, ale to nemusí být přípustné – nastavíme $y_j = 1$ s pravd. $(1-\varepsilon/2)x_j^*$.

T:$\varepsilon \in [0,1]$, $k \geq \frac{10}{\varepsilon}\ln(2n+2)$. Pak s pravděpodobností $> 1/2$ je y přípustné s alespoň $(1-\varepsilon)OPT^*$ množinami.

P: Nastav X součet v j-té složce matice. $OPT* > k$, $E[X] = (1\varepsilon/2)OPT^*$, $\mathrm{Var}[X] \leq E[X]$ jako obvykle.

Pak $P[X < (1-\varepsilon/2)OPT^*] \leq \exp(-\frac{\varepsilon^2}{10}OPT^*) \leq \exp(-\frac{\varepsilon^2}{10}k \leq \frac{1}{2n+2}$.

Druhou stranu uděláme totožně.