UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Matematika – 1. stopnja

Timotej Stibilj

VERJETNOSTNA METODA

Delo diplomskega seminarja

Mentor: izr. prof. dr. Mihael Perman

Ljubljana, 2024

Kazalo

Verjetnostna metoda

Povzetek

Verjetnostno metodo uporabljamo za nekostrukcijsko dokazovanje obstaja kombinatoričnih objektov z določenimi lastnostmi. Spoznamo osnove verjetnostnih algoritmov in povezavo z verjetnostno metodo; dokaz z verjetnostno metodo lahko pogosto prevedemo na verjetnostni algoritem. Spoznamo več načinov uporabe verjetnostne metode: osnovno metodo, metodo izbrisa, uporabo linearnosti pričakovane vrednosti in metodo drugega momenta. Pri vsaki metodi je predstavljen najmanj en primer. Predstavljena so Ramseyeva števila, barvanje hipergrafov in turnirji. Na primeru maksimalnega prereza grafa prikažemo prevedbo dokaza na verjetnostni algoritem in njegovo derandomizacijo. Dokažemo Erdősev izrek, ki pravi, da obstaja graf s poljubno veliko ožino in poljubno velikim kromatičnim številom. Spoznamo pojem slučajnega grafa in pragovne funkcije ter poiščemo pragovno funkcijo za vsebovanost danega podgrafa.

Probabilistic method

Abstract

The probabilistic method is used for nonconstructive proofs of existence of combinatorial objects with certain properties. We present the basic ideas of randomized algorithms and show their connection to the probabilistic method. Proofs using the probabilistic method often lead to randomized algorithms for such objects. Several ways of using the probabilistic method are shown, including the basic method, alteration, the use of linearity of expectation and the second-moment method. There is at least one example for each method. We present Ramsey numbers, hypergraph coloring and tournaments. In the example of the maximum cut problem, we present the randomized algorithm and its derandomization. We prove Erdős' theorem, which states that there exists a graph with arbitrarily large girth and arbitrarily large chromatic number. Lastly, we introduce the idea of random graphs, their threshold functions and find the threshold function for containing a given subgraph.

Math. Subj. Class. (2020): 05D40, 05C20, 05D10, 05C80, 68W20

Ključne besede: verjetnostna metoda, verjetnostni algoritem, metoda izbrisa, pričakovana vrednost, neenakost Markova, neenakost Čebiševa, slučajni graf

Keywords: probabilistic method, randomized algorithm, alteration, expected value, Markov's inequality, Chebyshev's inequality, random graph

1 Uvod

Verjetnostna metoda je matematično orodje za dokazovanje obstoja kombinatoričnih objektov. Konstrukcijski dokaz obstoja kombinatoričnega objekta z določenimi lastnostmi je lahko težaven. Verjetnostna metoda podaja alternativno pot, to je nekonstrukcijski način dokazovanja obstoja kombinatoričnega objekta s pomočjo verjetnosti. Zanimivo je, da lahko verjetnost uporabimo pri dokazovanju izrekov, ki sami po sebi niso verjetnostne narave. Na področju diskretne matematike in/ali verjetnostne metode je v 20. stoletju delovalo veliko madžarskih matematikov. Omenimo nekatere, ki so se ukvarjali s temami predstavljenimi v diplomski nalogi: L. Redei, T. Szele, L. Lovász in P. Erdős (1913-1996), ki ima največ zaslug za razvoj verjetnostne metode in ga smatramo za njenega začetnika.
Verjetnostna metoda se naravno povezuje tudi z algoritmičnim razmišljanjem, saj lahko dokaz pridobljen z verjetnostno metodo pogosto prevedemo na konstrukcijo želenega objekta s pomočjo verjetnostnega algoritma. Algoritmi s katerimi se najverjetneje vsak sprva seznani in bi jih lahko imeli za 'običajne', so deterministični. To pomeni, da ob izbranih vhodnih podatkih vedno izvedejo isto zaporedje korakov, ki privedejo do istega rezultata. Verjetnostni algoritem po drugi strani vrne rezultat, ki je odvisen od naključnih vrednosti uporabljenih tekom izvajanja algoritma in je zato naključen.[1] Cilj diplomskega dela je skozi primere predstaviti različne načine uporabe verjetnostne metode in prikazati povezavo verjetnostne metode z verjetnostnimi algoritmi. Verjetnostna metoda je močno orodje na številnih področjih matematike, recimo v teoriji števil in kombinatorični geometriji. Izbrani primeri predstavljeni v diplomskem delu v večini izvirajo iz teorije grafov.
V uvodnih dveh poglavjih predstavimo osnove verjetnostnih algoritmov in računske zahtevnosti problemov. V poglavju Koncentracija slučajnih spremenljivk dokažemo neenakost Markova in neenakost Čebiševa, ki podajata oceni odstopanja nenegativne slučajne spremenljivke od določene vrednosti. Sledi osrednji del diplomskega dela, tj. Verjetnostna metoda. Uvodoma na primerih iz teorije grafov predstavimo uporabo osnovne metode in algoritmični vidik verjetnostne metode. V 5.2 poglavju predstavimo uporabo linearnosti pričakovane vrednosti. Na primeru maksimalnega prereza grafa prikažemo, kako lahko ponekod iz nekostrukcijskega dokaza pridobimo verjetnostni algoritem ter ga derandomiziramo. Vpeljemo tudi pojem slučajnega grafa. V poglavju Metoda izbrisa pokažemo, kako lahko z alteracijo dokažemo rezultate, za katere neposredna uporaba osnovne metode odpove. Dokažemo tudi Erdősev izrek, ki pove, da obstaja graf s poljubno veliko ožino in poljubno velikim kromatičnim številom. V poglavju Metoda drugega momenta prikažemo uporabo neenačbe Čebiševa, vpeljemo pojem pragovne funkcije slučajnih grafov in poiščemo pragovno funkcijo za vsebovanost ciklov v grafu.

2 Verjetnostni algoritmi

Verjetnostni algoritem je vrsta algoritma, ki med svojim izvajanjem uporablja naključne vrednosti, na primer psevdonaključna števila. Posledično je izvedba takega

[1] Lahko imamo tudi drugačne oblike verjetnostnih algoritmov, kar je predstavljeno v naslednjem poglavju.

algoritma naključna, natančneje čas izvajanja ali sam rezultat algoritma sta slučajni spremenljivki. Razloga za izbiro verjetnostnega algoritma pri reševanju določenega problema sta predvsem naslednja dva. Velikokrat je koda v verjetnostnem algoritmu preprostejša kakor v deterministični različici algoritma za isti problem. Poleg tega lahko z uporabo verjetnostnih algoritmov pogosto pričakujemo hitrejši čas izvajanja. Slabost je v tem, da lahko kdaj tak algoritem vrne nepravilen odgovor.

Primeri področij, kjer se verjetnostni algoritmi pojavljajo, so denimo kombinatorična optimizacija, testi praštevilskosti in Monte Carlo simulacije.

Poznamo dve vrsti verjetnostnih algoritmov. ***Monte Carlo algoritem*** je verjetnostni algoritem, ki poda pravilni odgovor z določeno verjetnostjo. Verjetnost napačnega odgovora je običajno majhna. Rezultat algoritma je slučajna spremenljivka odvisna od izbire naključnih vrednosti skozi izvajanje algoritma. Pri tem je čas izvajanja določen z vhodnimi podatki. ***Las Vegas algoritem*** je verjetnostni algoritem, ki vedno poda pravilen odgovor, vendar je čas izvajanja slučajna spremenljivka.

Primer 2.1 (Quicksort algoritem)**.** Urejevalni algoritem za hitro urejanje (quicksort algoritem) deluje po načelu 'deli in vladaj'. Na vsakem koraku izbere pivot, elemente manjše od pivota položi v levi podseznam, preostale v desni, oba podseznama rekurzivno uredi in ju nato stakne skupaj. Pivot lahko izberemo na različne načine, na primer izberemo vedno prvi element seznama. Lahko ga izberemo tudi naključno in tako dobimo verjetnostni algoritem za hitro urejanje. Na primeru padajoče urejenega seznama dolžine n bo quicksort algoritem, ki za pivot vedno izbire prvi element v seznamu naredil $O(n^2)$ korakov. Po drugi strani pa lahko pokažemo, da verjetnosti quicksort algoritem, ki izbira pivot naključno uredi seznam v pričakovanih $O(n \log(n))$ korakih. Dokaz najdemo denimo v [5, str. 35-36]. □

3 Računska zahtevnost problemov

Ena izmed ključnih stvari pri obravnavi algoritmov je njihova časovna zahtevnost. Poleg tega, lahko govorimo tudi o težavnosti problemov samih.

Definicija 3.1. *Razred P (polinomski)* je razred vseh problemov za katere obstaja algoritem, ki poda rešitev v polinomskem času. *Razred NP (nedeterministično polinomski)* je razred vseh problemov, za katere lahko pravilnost rešitve preverimo v polinomskem času. Problem je *NP-težek*, če lahko vsak drugi problem iz razreda NP v polinomskem času prevedemo nanj.[2] Problem je *NP-poln*, če je *NP-težek* in *NP*.

Za nekatere probleme, za katere lahko preverimo pravilnost rešitev v polinomskem času, še nihče ni našel polinomskega algoritma. Zdi se, da velja $P \neq NP$, vendar to ostaja odprto vprašanje.[3]

[2] Če za en sam NP-težek problem pokažemo, da je v razredu P, takoj sledi $P = NP$.

[3] Problem $P \neq NP$ je en izmed sedmih problemov iz nabora The Millennium Prize Problems, za rešitev katerih je skupno razpisanih 7 milijonov dolarjev nagrade.

4 Koncentracija slučajnih spremenljivk

Včasih točnega odgovora za določeno verjetnost ne znamo podati ali ga je zahtevno izračunati, zato podamo ocene za določene verjetnosti, odstopanja od pričakovane vrednosti. V nadaljevanju bomo potrebovali naslednje neenakosti.

Pred formulacijo se spomnimo, da je za diskretno slučajno spremenljivko X pričakovana vrednost definirana kot $E[X] = \sum_x x \cdot P(X = x)$, varianca pa kot $\mathrm{Var}(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$.

Trditev 4.1 (Neenakost Markova)**.** *Naj bo X nenegativna slučajna spremenljivka in naj bo $a > 0$. Potem velja:*

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Dokaz.

$$E[X] = \sum_k k \cdot P(X = k) \geq \sum_{k \geq a} a \cdot P(X = k) = a \cdot P(X \geq a) \tag{4.1}$$

□

Trditev 4.2 (Neenakost Čebiševa)**.** *Naj bo $a > 0$. Potem velja:*

$$P(|X - E[X]| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

Dokaz. Za $a > 0$ velja:

$$P(|X - E[X]| \geq a) = P((X - E[X])^2 \geq a^2).$$

Na slučajni spremenljivki $(X - E[X])^2$ lahko uporabimo neenakost Markova, saj je nenegativna. Dobimo:

$$P((X - E[X])^2 \geq a^2) \leq \frac{E[(X - E[X])^2]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}.$$

□

Definicija 4.3. Vrednosti $E[X^n]$ pravimo n-ti moment slučajne spremenljivke X.

Opomba 4.4. Obe oceni veljata tudi za zvezne slučajne spremenljivke. Za uporabo obeh neenakosti ni potrebno poznati porazdelitve slučajne spremenljivke, dovolj je, da poznamo 1. oziroma 2. moment. Poleg tega obe neenakosti podajata najmočnejšo oceno, če poznamo le 1. oziroma 2. moment.

Če o slučajni spremenljivki vemo več, kot le prvi in drugi moment, lahko podamo močnejšo oceno.

Trditev 4.5 (Bernsteinova neenakost)**.** *Naj bodo I_i neodvisne enako porazdeljene slučajne spremenljivke, $I_i \sim Bernoulli(p)$, ki predstavljajo indikatorje dogodkov $A_i, i = 1, \ldots, n$. Potem za $S_n = \sum_{i=1}^n I_i$ in vsak $\epsilon > 0$ velja:*
$P(\frac{S_n}{n} - p \geq \epsilon) \leq e^{-\frac{1}{4}n\epsilon^2}$.

Dokaz izhaja iz [4, str. 31].

Dokaz. Vemo, da je S_n porazdeljen binomsko $Bin(n,p)$. Sledi, da je $P(\frac{S_n}{n} \geq p+\epsilon) = \sum_{k=m}^{n} \binom{n}{k} p^k (1-p)^{n-k}$, kjer je $m = \lceil (n(p+\epsilon) \rceil$. Eksponentna funkcija je naraščajoča, zato za vsak $\lambda > 0$ velja $e^{\lambda k} \geq e^{\lambda n(p+\epsilon)}$, čim je $k \geq m$. Pišimo $q := 1-p$. Velja:

$$e^{\lambda k - \lambda n(p+\epsilon)} = e^{-\lambda n\epsilon} e^{\lambda k} e^{-\lambda pn} = e^{-\lambda n\epsilon} e^{\lambda qk} e^{-\lambda p(n-k)}.$$

Torej je:

$$\begin{aligned} P\left(\frac{S_n}{n} \geq p+\epsilon\right) &\leq \sum_{k=m}^{n} e^{\lambda k - \lambda n(p+\epsilon)} \binom{n}{k} p^k q^{n-k} \\ &= e^{-\lambda n\epsilon} \sum_{k=m}^{n} e^{\lambda qk} e^{-\lambda p(n-k)} \binom{n}{k} p^k q^{n-k} \\ &\leq e^{-\lambda n\epsilon} \sum_{k=0}^{n} \binom{n}{k} (pe^{\lambda q})^k (qe^{-\lambda p})^{(n-k)} \\ &= e^{-\lambda n\epsilon} (pe^{\lambda q} + qe^{-\lambda p})^n. \end{aligned}$$

Zadnja enakost sledi iz binomskega izreka. Sedaj uporabimo neenakost $e^x \leq x + e^{x^2}$ in dobimo:
$P\left(\frac{S_n}{n} \geq p+\epsilon\right) \leq e^{-\lambda n\epsilon} \left(pe^{\lambda^2 q^2} + qe^{\lambda^2 p^2}\right)^n \leq e^{\lambda^2 n - \lambda n\epsilon}$. Ocena velja za vsak $\lambda \in \mathbb{R}$. Izberemo λ, pri katerem je izraz najmanjši. Funkcija $g : \mathbb{R} \to \mathbb{R}, g(\lambda) = e^{\lambda^2 n - \lambda n\epsilon}$ doseže minimum pri $\lambda = \frac{\epsilon}{2}$. Pri tej vrednosti λ dobimo želeno:
$P\left(\frac{S_n}{n} \geq p+\epsilon\right) \leq e^{-\frac{1}{4}n\epsilon^2}$ za vsak $\epsilon > 0$.

□

Opomba 4.6. Simetričen argument bi lahko uporabili za $P\left(\frac{S_n}{n} \leq p-\epsilon\right)$ in s tem pokazali, da slučajna spremenljivka $\frac{S_n}{n}$ skoraj gotovo konvergira k p, kar je poseben primer zakona velikih števil za indikatorske slučajne spremenljivke z verjetnostjo p.

Na zgledu prikažimo moč ocen neenakosti Markova, neenakosti Čebiševa in Bernsteinove neenakosti.

Zgled 4.7. Mečemo pošten kovanec. Neodvisnost med meti lahko privzamemo. Zanima nas verjetnost, da je v n metih več kot 75% cifer. Označimo z X število cifer v n metih. Neenakost Markova nam da $P(X \geq \frac{3n}{4}) \leq \frac{\frac{n}{2}}{\frac{3n}{4}} = \frac{2}{3}$. Z indikatorji dogodkov, da je i-ti met cifra dobimo $\mathrm{Var}(X) = \frac{n}{4}$. Sledi $P(X \geq \frac{3n}{4}) = P(X - \frac{n}{2} \geq \frac{n}{4}) \leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) \leq \frac{\mathrm{Var}(X)}{(\frac{n}{4})^2} = \frac{4}{n}$. Dodajmo še Bernsteinovo neenakost. Število cifer X zadošča predpostavkam Bernsteinove neenakosti, saj je X vsota neodvisnih indikatorjev. Velja $P(\frac{X}{n} \geq \frac{1}{2} + \epsilon) \leq e^{-\frac{1}{4}n\epsilon^2}$. Če vzamemo $\epsilon = \frac{1}{4}$, dobimo $P(S_n \geq \frac{3n}{4}) \leq e^{-\frac{n}{64}}$.
Vidimo, da pri dovolj velikem številu metov n upoštevanje več informacij o slučajni spremenljivki res podaja strožjo oceno za verjetnost. □

5 Verjetnostna metoda

5.1 Osnovna metoda

Osnovni princip verjetnostne metode je sledeč. Iščemo objekt iz neke množice X z dano lastnostjo L. Množici X priredimo končen verjetnostni prostor in pokažemo, da je verjetnost, da X vsebuje element z lastnostjo L strogo pozitivna. Od tod sledi nekonstrukcijski dokaz obstoja takega objekta.

Besedilo v poglavju Osnovne metode je osnovano na [6] in [8].

Pred obravnavo primerov si poglejmo naslednjo elementarno lemo, ki je kljub preprostosti pogosto uporabna.

Lema 5.1 (Boolova neenakost)**.** *Za dogodke $A_i \subseteq \Omega$, $i = 1, ..., n$, velja:*

$$P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i).$$

Dokaz. Za $n = 2$: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq P(A_1) + P(A_2)$.

Za $n \rightsquigarrow n + 1$:

$$\begin{aligned} P(\bigcup_{i=1}^{n+1} A_i) &= P((A_1 \cup \ldots \cup A_n) \cup A_{n+1}) \leq P(A_1 \cup \ldots \cup A_n) + P(A_{n+1}) \quad \text{[baza]} \\ &\leq \sum_{i=1}^{n} P(A_i) + P(A_{n+1}) \quad \text{[i.p.]} \\ &= \sum_{i=1}^{n+1} P(A_i) \end{aligned} \tag{5.1}$$

□

5.1.1 Ramseyeva števila

Ramseyeva števila se v literaturi pogosto vpelje s sledečim motivacijskim primerom. Prirejamo zabavo, na katero moramo povabiti določeno število ljudi. Za vsako trojico gostov si želimo, da ni sestavljena iz samih prijateljev in niti iz samih neznancev. Največ koliko ljudi lahko povabimo, da bo to še možno? Z drugimi besedami: pri katerem številu povabljencev to ne bo več možno in bo zagotovo obstajala trojica neznancev ali trojica prijateljev? [4]

V tem poglavju sta dodatno uporabljena še vira [1] in [2].

Definicija 5.2. *Neodvisna množica* $A \subset V$ grafa G je podmnožica vozlišč, med katerimi ni skupnih povezav.

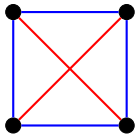
Definicija 5.3. *Poln graf* je graf, v katerem je vsako izmed vozlišč povezano z vsemi preostalimi vozlišči. *K-klika* je podgraf v grafu G, ki je izomorfen polnemu grafu s k vozlišči.

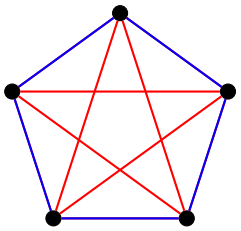
[4] Pri tem predpostavljamo, da med zbranimi ni zvezdnikov, torej da je poznanostvo simetrična relacija.

Ramseyevo število $R(k,l)$ je najmanjše naravno število $n \in \mathbb{N}$, pri katerem poljuben graf na n točkah vsebuje k-kliko ali neodvisno množico velikosti l. Ekvivalentno lahko zapišemo:

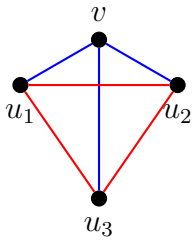
$$R(k,l) = \min(n \in \mathbb{N}, \text{ pri vsakem barvanju povezav polnega grafa } K_n \text{ z rdečo in modro barvo obstaja rdeča } k\text{-klika ali modra } l\text{-klika}).$$

Poglejmo si odgovor za začetni primer. Poiskati želimo najmanjše število vozlišč n, da bomo pri poljubnem barvanju povezav z rdečo in modro barvo zagotovo lahko našli vsaj en enobarvni trikotnik. Vozlišča predstavljajo ljudi, modra barva povezav predstavlja poznanstvo, rdeče povezave pa povezujejo neznance. Videli bomo, da je iskano število oseb 6, tj. $R(3,3) = 6$. Iz sledečih barvanj brez enobarvnih trikotnikov vidimo, da je $R(3,3) \geq 6$.





Pokažimo še, da je $R(3,3) \leq 6$. Izberimo poljubno osebo/vozlišče $v \in V$. Ostane še 5 oseb. Bodisi se v pozna z vsaj tremi od teh bodisi se z vsaj tremi ne pozna. Brez škode za splošnost se v pozna z vsaj tremi. Vzamemo 3 od teh, to so u_1, u_2, u_3 in povezave med v in u_i pobarvamo z modro. Če se želimo izogniti modrim trikotnikom, potem nobena izmed povezav u_iu_j ne sme biti modra, torej so 3 povezave med u_1, u_2, u_3 rdeče in imamo rdeč trikotnik. Od tod sledi odgovor na začetno uganko res $R(3,3) = 6$.



Ramseyev izrek preko zveze $R(s,t) \leq R(s-1,t) + R(s,t-1)$ zagotavlja obstoj $R(k,l)$ za poljubna $k \in \mathbb{N}$, $l \in \mathbb{N}$ (dokaz najdemo v [1, str. 7]). Računanje točnih vrednosti za Ramsayeva števila pa je računsko zahtevno.[5] Za diagonalna Ramseyeva števila, to so števila $R(k,k)$ poznamo $R(3,3) = 6$, $R(4,4) = 18$ (dokaz v [1, str. 12]), točna vrednost za $R(5,5)$ pa ni več znana. Z verjetnostno metodo podajmo spodnjo mejo za $R(k,k)$, ki jo je prvi zapisal P. Erdős leta 1947.

Trditev 5.4. *Naj bo* $\binom{n}{k} 2^{-\binom{k}{2}+1} < 1$. *Potem obstaja barvanje polnega grafa* K_n *z dvema barvama brez monokromatskih* k*-klik.*

[5] P. Erdős je dejal, naj si predstavljamo, da na Zemlji pristanejo nezemljani in od nas zahtevajo vrednost $R(5,5)$, sicer bodo uničili naš planet. V tem primeru trdi, da bi morali združiti vse svoje računalnike in matematike ter poskušati najti to vrednost. Toda recimo, da zahtevajo vrednost $R(6,6)$. V tem primeru nam ne preostane drugega kot boj.

Dokaz. Pobarvajmo povezave grafa K_n naključno in neodvisno od ostalih. Natančneje, definiramo $\Omega = \{\text{barvanja povezav v polnem grafu } K_n \text{ z dvema barvama}\}$. Velja $|\Omega| = 2^{\binom{n}{2}}$ in $P(\text{poljubno barvanje}) = 2^{-\binom{n}{2}}$. V grafu K_n je $\binom{n}{k}$ k-klik, ki jih poljubno oštevilčimo. Za $i = 1, ..., \binom{n}{k}$ definiramo dogodek $A_i = \{\text{i-ta } k\text{-klika je monokromatska}\}$. Očitno velja $P(A_i) = 2 \cdot (\frac{1}{2})^{\binom{k}{2}} = 2^{-\binom{k}{2}+1}$. Z Boolovo neenakostjo 5.1 ocenimo:

$$P\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) \leq \sum_{i=1}^{\binom{n}{k}} P(A_i) = \binom{n}{k} \cdot 2^{-\binom{k}{2}+1} < 1.$$

Zadnja neenakost sledi iz predpostavke trditve. Od tod pa dobimo obstoj takega barvanja:

$$P\left(\bigcap_{i=1}^{\binom{n}{k}} A_i^C\right) = P\left(\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right)^C\right) = 1 - P\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) > 0.$$

□

Trditev 5.4 uporabimo za dokaz, da števila $R(k,k)$ rastejo eksponentno v k.[6]

Lema 5.5. *Naj bo $k \geq 3$. Potem velja $\frac{2^{1+\frac{k}{2}}}{k!} < 1$.*

Dokaz. Dokazujemo z indukcijo na k.
Baza indukcije za $k = 3$:

$$\frac{2^{1+\frac{3}{2}}}{3!} = \frac{2 \cdot \sqrt{2}}{3} < 1.$$

Predpostavimo, da neenakost velja za k. Potem je:

$$\frac{2^{1+\frac{k+1}{2}}}{(k+1)!} = \frac{2^{1+\frac{k}{2}}}{k!} \cdot \frac{2^{\frac{1}{2}}}{k+1} < 1$$

□

Posledica 5.6. *Za $k \geq 3$ je $R(k,k) > 2^{k/2}$.*

Dokaz. Naj bo $k \geq 3$. Vzemimo $n \coloneqq 2^{k/2}$. Velja:

$$\binom{n}{k} 2^{-\binom{k}{2}+1} \leq \frac{n^k}{k!} \cdot 2^{1-\frac{k\cdot(k-1)}{2}} \leq \frac{(2^{\frac{k}{2}})^k}{k!} \cdot 2^{1-\frac{k^2}{2}+\frac{k}{2}} = \frac{2^{1+\frac{k}{2}}}{k!} < 1.$$

Torej je izpolnjen pogoj iz trditve 5.4. □

Opomba 5.7. Verjetnosti bi se lahko pri dokazovanju spodnje meje za Ramseyeva števila popolnoma izognili in podoben dokaz napravili le s preštevanjem. Pokazati bi morali, da je število barvanj polnega grafa K_n z monokromatskimi k-klikami manjše od števila vseh barvanj K_n. V splošnem primeru dokazovanja obstoja kombinatoričnega objekta, ko se poslužujemo denimo metode drugega momenta, pa je zgolj preštevanje lahko brezupno in je verjetnostni pristop ključnega pomena za preprostejši dokaz, zato je bil le-ta prikazan že v zgornjem primeru. Poleg tega iz verjetnostne metode naravno sledi algoritmičen pristop dokazovanja.

[6] V resnici s tem pokažemo, da rastejo vsaj eksponentno v k. Potrebovali bi še zgornjo mejo. Znana je na primer meja $R(k,k) < 4^k$ (dokaz v [1, str. 14])

Definicija 5.8. Za funkcijo f pravimo: $f(n) = o(g(n))$, če $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.
Torej velja $f = o(1)$, če je $\lim_{n\to\infty} f(n) = 0$.

5.1.2 Algoritmični vidik

Zanima nas ali obstaja polinomski način konstrukcije objekta, katerega obstoj smo dokazali z verjetnostno metodo. Za primer Ramseyevih števil bi to pomenilo iskanje učinkovitega algoritma za barvanje polnega grafa K_n brez monokromatske k-klike. Preštevanje vseh možnih barvanj zahteva eksponentno mnogo časa v k, želimo pa si polinomskega algoritma. V splošnem, če lahko, postopamo na sledeč način.

Poiščemo učinkovito vzorčenje za želen objekt.

Denimo, da je verjetnost, da izbrani vzorec zadostuje želenim lastnostim, enaka p. Potem je $X = \{$število vzorcev do dobre rešitve$\}$ slučajna spremeneljivka, $X \sim Geom(p)$. Sledi $E[X] = \frac{1}{p}$. Za polinomski pričakovani čas izvajanja algoritma mora biti $\frac{1}{p}$ polinomsko odvisen od podatkov problema.

Za $p = 1 - o(1)$ vzorčimo le enkrat, t.j. konstruiramo le en objekt in imamo Monte Carlo algoritem, ki poda pravilno rešitev z veliko verjetnostjo. Verjetnost pravilne rešitve z enim samim vzorčenjem je lahko majhna. To lahko izboljšamo z večkratnim vzorčenjem.

Za algoritem, ki bo vedno podal pravilen odgovor, t. j. Las Vegas algoritem, potrebujemo polinomski način preverjanja, da objekt zadostuje iskanim pogojem. Nato preverjamo vzorce/konstruirane objekte, dokler ne dobimo 'dobrega'. Če časovno zahtevnost preverjanja pravilnosti vzorca zmnožimo s časovno zahtevnostjo za generiranje le tega in pričakovanim številom potrebnih vzorcev, dobimo skupno časovno zahtevnost takega Las Vegas algoritma.

Na primeru Ramseyevih števil bi bil premislik sledeč. V skladu s posledico 5.6 za $n = \lfloor 2^{\frac{k}{2}} \rfloor$ obstaja barvanje polnega grafa K_n z dvema barvama brez monokromatske k-klike. Dovolj je, da vzorčimo le enkrat, tj. pobarvamo vsako povezavo K_n z verjetnostjo $\frac{1}{2}$ z rdečo in z verjetnostjo $\frac{1}{2}$ z modro barvo. V dokazu posledice 5.6 smo videli, da je za $n = \lfloor 2^{\frac{k}{2}} \rfloor$ verjetnost dogodka A, da obstaja monokromatska k-klika: $P(A) \leq \frac{2^{1+\frac{k}{2}}}{k!} = o(1)$. Tako dobimo Monte Carlo algoritem za barvanje K_n brez monokromatske k-klike z veliko verjetnostjo. Še en primer verjetnostnega algoritma je prikazan v razdelku 5.27.

5.1.3 Barvanje hipergrafa

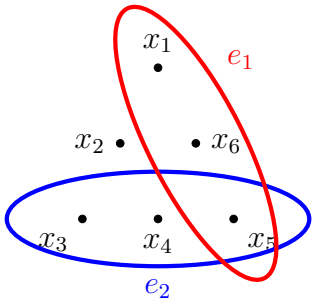
Hipergraf predstavlja posplošitev pojma enostavnega grafa.

Definicija 5.9. *Hipergraf* je par (V, E), kjer je množica vozlišč V končna množica točk in množica hiperpovezav E družina poljubnih podmnožic množice V.

Definicija 5.10. Hipergraf (V, E) je *k-enoličen*, če za vsako povezavo $e \in E$ velja $|e| = k$.

Definicija 5.11. Hipergraf (V, E) je *2-obarvljiv*, če obstaja barvanje vozlišč V z dvema barvama brez monokromatskih hiperpovezav.

Naj bo $m(k)$ najmanjše število hiperpovezav, pri katerem obstaja k-enoličen hipergraf, ki ni 2 obarvljiv.



Slika 1: 3-enoličen hipergraf

Primer 5.12. Enostavni graf je 2-enolični hipergraf. 2-obarvljivost enostavnega grafa je ekvivalentna *dobremu barvanju* vozlišč enostavnega grafa, tj. barvanju, kjer sta poljubni sosednji vozlišči različnih barv. Kromatično število $\chi(G)$ je najmanjše število potrebnih barv za dobro barvanje enostavnega grafa G. Hitro vidimo, da velja $\omega(G) \leq \chi(G)$, kjer je $\omega(G)$ velikost največje klike v grafu G. Posledično je $m(2) = 3$, saj očitno velja, da poln graf K_3 ni 2-obarvljiv, edina dva grafa z dvema povezavama pa sta 2-obarvljiva. □

Natančnih vrednosti $m(k)$ za večje k ne poznamo. Ponovno pa lahko podamo spodnjo mejo za $m(k)$ z osnovno verjetnostno metodo.

Trditev 5.13. *Za vsak $k \geq 2$ velja:*

$$m(k) \geq 2^{k-1}.$$

Dokaz. Vzemimo poljuben k-enoličen hipergraf $G = (V, E)$ z manj kot 2^{k-1} povezavami in pokažimo, da je 2-obarvljiv. Vsako vozlišče v grafu pobarvamo z verjetnostjo $\frac{1}{2}$ rdeče in z verjetnostjo $\frac{1}{2}$ z modro neodvisno od ostalih. Verjetnost P_{e_i}, da je i-ta povezava v hipergrafu G monokromatska je: $P_{e_i} = 2 \cdot (\frac{1}{2})^k = 2^{1-k}$. Z Boolovo neenakostjo 5.1 ocenimo verjetnost dogodka A, da pri takem barvanju obstaja monokromatska hiperpovezava.

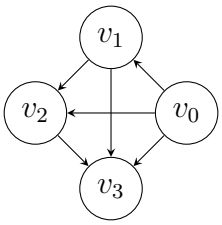
$$P(A) = P(\bigcup_{e_i \in E} P_{e_i}) \leq 2^{1-k} \cdot |E| < 2^{1-k} \cdot 2^{k-1} = 1.$$

Pri tem smo upoštevali predpostavko, da je $|E| < 2^{k-1}$. Torej je verjetnost dogodka A^C, da pri takem barvanju ne obstaja monokromatska povezava $P(A^C) = 1 - P(A) > 0$ in posledično obstaja barvanje z 2 barvama. Torej potrebujemo vsaj 2^{k-1} hiperpovezav. □

5.1.4 Turnirji z lastnostjo $\mathbf{P_k}$

Definicija 5.14. *Turnir* $T = (V, E)$ je usmerjen poln graf, kjer je V končna množica točk in E množica usmerjenih povezav. Za vsaki vozlišči $x, y \in V$ natanko en izmed parov (x, y) oz. (y, x) pripada množici E.

Turnir lahko, kot ime pove, interpretiramo kot igro, kjer vsak par igralcev odigra med sabo eno igro in vedno en izmed igralcev zmaga. Vozlišča grafa so igralci,



Slika 2: Turnir s 4 igralci

povezave pa predstavljajo odigrano igro. Usmerjeno povezavo (a, b) po dogovoru interpretiramo kot zmago igralca a nad igralcem b.

Imejmo igro z n igralci. Za določeno množico igralcev, denimo da jih izberemo $k < n$, se lahko vprašamo ali obstaja kakšen izmed preostalih igralcev, ki premaga vseh izbranih k igralcev. Če za poljubno množico s k igralci tak igralec obstaja, pravimo, da ima turnir lastnost P_k. Zapišimo še formalno definicijo:

Definicija 5.15. Turnir $T = (V, E)$ ima *lastnost* P_k, če za vsako podmnožico $S \subseteq V$, za katero je $|S| = k$, obstaja $v \in V \setminus S$, da je $(v, x) \in E$ za vsak $x \in S$.

Lastnost P_1 denimo pomeni, da je v danem turnirju vsak igralec premagan vsaj enkrat. Poglejmo si lastnost P_1 za majhno število igralcev. V primeru 2 igralcev turnir z lastnostjo P_1 ne obstaja, saj je vedno en neporažen. V primeru 3 igralcev obstaja tako turnir z lastnostjo P_1 kot turnir brez nje. Izrek 5.17 s pomočjo naslednje trditve 5.16 pove, da za poljuben k obstaja turnir z lastostjo P_k ob primerno dovolj velikem številu igralcev n.

Trditev 5.16. *Če velja* $\binom{n}{k}(1 - 2^{-k})^{n-k} < 1$*, potem obstaja turnir z n vozlišči in lastostjo* P_k*.*

Dokaz. Ponovno se dokazovanja lotimo na prvo žogo s tem, da konstuiramo slučajen turnir na n vozliščih, ki ga dobimo tako, da povezavo med vsakima dvema vozliščema (igralcema) $x, y \in V$ usmerimo v eno smer z verjetnostjo $\frac{1}{2}$ in v drugo smer z verjetnostjo $\frac{1}{2}$ neodvisno od ostalih. Naj bodo $S_i, i = 1, ..., \binom{n}{k}$ vse možne množice k igralcev. Definiramo $A_{S_i} =$ noben igralec iz $V \setminus S_i$ ne premaga vseh igralcev iz S_i. Dogodek A_{S_i} je enak preseku neodvisnih dogodkov, da posamezen igralec $v \in V \setminus S_i$ ne premaga vseh igralcev iz S_i. Verjetnost posamičnega takega dogodka je $1 - P(\text{v premaga vse igralce iz } S_i) = 1 - (\frac{1}{2})^k$, takih dogodkov pa je $n - k$, kolikor je igralcev izven množice S_i.
Sledi, da je:

$$P(A_{S_i}) = \left(1 - (\frac{1}{2})^k\right)^{n-k} = (1 - 2^{-k})^{n-k}.$$

Zdaj pa z Boolovo neenakostjo 5.1 za $A = \bigcup_{i=1}^{\binom{n}{k}} A_{S_i}$ ocenimo:

$$P(A) \leq \sum_{i=1}^{\binom{n}{k}} P(A_{S_i}) = \binom{n}{k}(1 - 2^{-k})^{n-k} < 1.$$

Naj dogodek P_k simbolizira lastnost P_k, torej, da za vsako k-elementno množico igralcev $S_i \subset V$ obstaja igralec $v \in V \setminus S_i$, ki premaga vse igralce iz S_i. Dogodek

P_k, se zgodi natanko takrat, ko se ne zgodi noben izmed dogodkov A_{S_i}. Velja torej:

$$P(P_k) = P((\bigcup_{i=1}^{\binom{n}{k}} A_{S_i})^C) = P(A^C) = 1 - P(A) > 0, \text{ kar smo želeli pokazati.}$$

□

Naslednji izrek s pomočjo prejšnje trditve predstavi ekspliciten rezultat o dovoljšnem številu igralcu v turnirju za izpolnjevanje lastnosti P_k.

Izrek 5.17. *Naj bo $n \geq k^2 \cdot 2^{k+1}$. Potem obstaja turnir z n igralci in lastnostjo P_k.*

Kot pri 5.6 moramo za nazornejši rezultat tudi tu poleg pristopa z verjetnostno metodo uporabiti nekaj spretnosti z neenakostmi.

Lema 5.18. *Za vsak $x \leq 1$ in $r \geq 0$ velja $(1-x)^r \leq e^{-xr}$.*

Dokaz. Hitro vidimo, da za vsak $x \in R$ velja $1 - x \leq e^{-x}$. To je ekvivalentno temu, da je funkcija $f(x) = e^{-x} + x - 1$ nenegativna za vsak $x \in R$. Računamo:

$f'(x) = 1 - e^{-x}$. Velja $f'(x) = 0$ za $e^{-x} = 1,$ torej za $x = 1$. Za $x > 0$ je $f'(x) > 0$, za $x < 0$ je $f'(x) < 0$ in $f(0) = 0$, torej je funkcija f nenegativna za vsak $x \in R$ in zato v posebnem primeru tudi za $x \geq 0$.

Funkcije $f(x) = x^r$ za $r > 0$ so naraščajoče za $x > 0$, od koder sledi želena neenakost. □

Lema 5.19. *Za vsak $k \geq 2$ velja $e^{\frac{k}{2^k}} < k!$.*

Dokaz. Uporabimo indukcijo na k. Za $k = 2$ dobimo: $e^{\frac{1}{2}} \approx 1.6487 < 2$. Indukcijski korak:

$$e^{\frac{k+1}{2^{k+1}}} = e^{\frac{k}{2^{k+1}} + \frac{1}{2^{k+1}}} = e^{\frac{k}{2^{k+1}}} \cdot e^{\frac{1}{2^{k+1}}} < e^{\frac{k}{2^k}} \cdot e^{\frac{1}{2^{k+1}}}$$

Velja $\frac{1}{2^{k+1}} < \frac{1}{2}$ oziroma $e^{\frac{1}{2^{k+1}}} < e^{\frac{1}{2}} \approx 1.6487$. Če upoštevamo še indukcijsko predpostavko dobimo: $e^{\frac{k+1}{2^{k+1}}} < k! \cdot e^{\frac{1}{2}} < (k+1)!$. □

Lema 5.20. *Za vsak $k \geq 2, k \in \mathbb{N}$ velja $2\log k + (k+1)\log 2 - 2k < 0$.*

Dokaz. Pokažimo, da za funkcijo $g(k) = 2\log k + (k+1)\log(2) - 2k$ velja $g(k) < 0$ za $k \geq 2$. Računajmo: $g'(k) = \frac{2}{k} + \log 2 - 2 = 0$, kar nam da $k_0 = \frac{2}{2-\log(2)} \approx 1.530$. Za vrednosti $k > k_0$ je funkcija $g'(k)$ negativna, torej je $g(k)$ padajoča na intervalu (x_0, ∞). V točki $k_0 \approx 1.530$ je $g(k_0) \approx -0.456 < 0$. Posledično za vsak $k \geq 2$ velja $g(k) < 0$ in v posebnem to velja za $k \in \mathbb{N}, k \geq 2$. □

Lema 5.21. *Za vsak $k \in \mathbb{N}, k \geq 2$ in vsak $n \geq k^2 2^{k+1}$ velja $\frac{n^k}{e^{\frac{n}{2^k}}} < 1$.*

Dokaz. Pokažimo, da za funkcije $f_k(x) = x^k e^{\frac{-x}{2^k}}, k \in \mathbb{N}, k \geq 2$ velja: $f_k(x) < 1$ za vsak $x \geq k^2 2^{k+1}$. Izračunajmo odvod:

$$\begin{aligned}
\frac{d}{dx} f_k(x) &= k x^{k-1} e^{\frac{-x}{2^k}} + x^k e^{\frac{-x}{2^k}} \left(\frac{-1}{2^k}\right) \\
&= e^{\frac{-x}{2^k}} x^{k-1} \left(k - \frac{x}{2^k}\right) \\
&= 2^{-k} e^{\frac{-x}{2^k}} x^{k-1} (2^k k - x).
\end{aligned}$$

Zanima nas obnašanje funkcije za pozitivne x, natančneje za $x \geq k^2 2^{k+1}$. Vidimo, da ima $f_k'(x)$ ničlo pri $x = 2^k k$, za $x > 2^k k$ pa je vrednost $f_k'(x)$ negativna, zato je $f_k(x)$ padajoča za $x > 2^k k$. Za $k \geq 2$ velja $k^2 2^{k+1} > 2^k k$, zato je torej dovolj pokazati, da je $f_k(k^2 2^{k+1}) < 1$ za vsak $k \geq 2$. Računamo:

$$f_k(k^2 2^{k+1}) = (k^2 2^{k+1})^k e^{-\frac{k^2 2^{k+1}}{2^k}} = k^{2k} 2^{k(k+1)} e^{-2k^2}.$$

To, da za vsak $k \geq 2$ velja $f_k(k^2 2^{k+1}) < 1$, lahko zapišemo na ekvivalenten način kot:

$$\begin{aligned} k^{2k} 2^{k(k+1)} e^{-2k^2} &< 1 \ / \ e^{2k^2} \\ k^{2k} 2^{k(k+1)} &< e^{2k^2} \ / \ \log \\ \log k^{2k} + \log 2^{k(k+1)} &< 2k^2 \\ 2k \log k + k(k+1) \log 2 - 2k^2 &< 0 \ / \ : k \\ 2 \log k + (k+1) \log 2 - 2k &< 0, \text{ kar drži po lemi 5.20.} \end{aligned}$$

□

Dokaz izreka 5.17. Dokazujemo, da za vsak $n \geq k^2 \cdot 2^{k+1}$ obstaja turnir z n igralci in lastnostjo P_k. Za $k \geq 2$ bomo uporabili zgornje leme, primer $k = 1$ pokažimo samostojno.
Pri $k = 1$ nas zanima ali pri n igralcih, kjer je $n \geq 1^2 2^2 = 4$ obstaja turnir z lastnostjo P_1. Kot že omenjeno, lastnost P_1 pomeni, da je vsak igralec premagan vsaj enkrat. Primer želenega turnirja je recimo poln graf na vozliščih $v_0, \ldots, v_{n-1}$, med katerimi so usmerjene povezave $(v_i, v_{i+1 \bmod n})$ za $i = 1, 2, ..., n-1$, to je usmerjen cikel na vseh vozliščih grafa (Hamiltonov cikel), ostale povezave pa so usmerjene poljubno.
Za $k \geq 2$ uporabimo oceno $\binom{n}{k} < \frac{n^k}{k!}$ in nato zaporedoma uporabimo leme 5.18, 5.19 in 5.21 ter dobimo:

$$\binom{n}{k}(1 - 2^{-k})^{n-k} \leq \frac{n^k}{k!} e^{-\frac{n-k}{2^k}} \leq \frac{n^k}{e^{\frac{k}{2^k}}} e^{-\frac{n}{2^k}} e^{\frac{k}{2^k}} = \frac{n^k}{e^{\frac{n}{2^k}}} < 1.$$

S tem je izrek dokazan. □

5.2 Uporaba lastnosti pričakovane vrednosti

Spomnimo se: pričakovana vrednost diskretne slučajne spremenljivke X je $E[X] = \sum_x x \cdot P(X = x)$.

Ideja osnovne verjetnostne metode je pokazati, da je verjetnost želenega objekta v skonstruiranem verjetnostnem prostoru strogo pozitivna. Včasih si pri tem lahko pomagamo s pričakovaneo vrednostjo. Naslednja preprosta lema pravi, da diskretna slučajna spremenljivka X ne more doseči zgolj vrednosti strogo večjih oziroma strogo manjših od $E[X]$.

Lema 5.22. *Za slučajno spremenljivko X velja:*

$$P(X \geq E[X]) > 0 \text{ in } P(X \leq E[X]) > 0.$$

Dokaz. S protislovjem. Recimo, da velja $P(X \geq E[X]) = 0$. Potem velja:

$$\begin{aligned} E[X] &= \sum_x x \cdot P(X = x) = \sum_{x<E[X]} x \cdot P(X = x) \leq \\ &< \sum_{x<E[X]} E[X] \cdot P(X = x) = \sum_x E[X] \cdot P(X = x) = E[X]. \end{aligned} \tag{5.2}$$

Simetrično za $P(X \leq E[X]) = 0$. □

Poleg tega je pri uporabi pričakovane vrednosti pogosto uporabna njena linearnost, kar je znana lastnost. To pomeni, da za poljubne $a_i \in \mathbb{R}$ in slučajne spremenljivke X_i velja $E[\sum_{i=1}^n a_i \cdot X_i] = \sum_{i=1}^n a_i \cdot E[X_i]$ ne glede na neodvisnost med slučajnimi spremenljivkami X_i. Najpogosteje linearnost uporabimo tako, da slučajno spremenljivko zapišemo kot vsoto indikatorjev.

5.2.1 Hamiltonske poti v turnirjih

V prejšnjem razdelku smo si ogledali lastnost P_k na turnirjih. Poglejmo še eno izmed tem, ki jih lahko preučujemo na turnirjih.

Definicija 5.23. Hamiltonska pot v turnirju $T = (V, E)$ je usmerjena pot skozi vse točke. To je taka permutacija $\pi \in S_n$ igralcev $v_1, ..., v_n$, da za vsak $i = 1, ..., n$ velja: $(v_{\pi(i)}, v_{\pi(i+1)}) \in E$.

Madžarski matematik Szele je z verjetnostjo metodo leta 1943 pokazal, da vedno obstaja turnir z veliko hamiltonskimi potmi. Izrek 5.25 je natančneje formuliran v nadaljevanju. Pred tem si poglejmo izrek, ki pravi, da vsak turnir vsebuje hamiltonsko pot.

Izrek 5.24. *(Rédei 1934) V vsakem turnirju lahko najdemo vsaj eno hamiltonsko pot.*

Dokaz. Pri tem dokazu ne uporabljamo verjetnostne metode, ampak preprosto indukcijo na število vozlišč n.
Baza indukcije za $n = 1$ je na prazno izpolnjena, saj je (v_1) hamiltonska pot. Indukcijski korak je sledeč: Denimo, da v vsakem turnirju z $n - 1$ vozlišči obstaja hamiltonska pot. Vzemimo poljuben turnir $T = (V, E)$ z n vozlišči in naj bo v eno izmed vozlišč. Po indukcijski predpostavki v turnirju $T - v$ (to je turnir, v katerem iz T odstranimo vozlišče v in vse povezave, ki imajo v za krajišče), obstaja hamiltonska pot. Brez škode za splošnost naj bo $(v_1, \ldots, v_{n-1})$ ta pot. Sedaj imamo 2 možnosti. V primeru, da igralec v premaga vse preostale igralce, je iskana hamiltonska pot kar $(v, v_1, \ldots, v_{n-1})$. V nasprotnem primeru obstaja vsaj en igralec, ki ga v ne premaga. Naj bo $j = \min\{i \in \{1, \ldots, n-1\} |\ v \text{ premaga } v_i\}$. Potem lahko konstruiramo hamiltonsko pot $(v_1, v_2, \ldots, v_{i-1}, v, v_i, \ldots, v_{n-1})$. □

Izrek 5.25. *(Szele, 1943) Obstaja turnir na n točkah z vsaj $\frac{n!}{2^{n-1}}$ hamiltonskimi potmi.*

Dokaz. Za razliko od dokaza izreka 5.24 tokrat uporabimo verjetnostni pristop.
Vzemimo slučajni turnir $T = (V, E)$ z n igralci (tj. $|V| = n$) kot v dokazu trditve

5.16. Definirajmo slučajno spremenljivko $X =$ število Hamiltonskih poti v slučajnem turnirju z n igralci. X lahko zapišemo kot $X = \sum_{\pi \in S_n} I_\pi$, kjer z I_π označimo indikator dogodka, da permutacija π določa hamiltonsko pot v turnirju.

$$I_\pi = \begin{cases} 1, & \text{če je } (v_{\pi(i)}, v_{\pi(i+1)}) \in E, \text{ za vsak } 1 \leq i \leq n-1, \\ 0, & \text{sicer.} \end{cases}$$

Velja $E[I_\pi] = P(I_\pi = 1) = (\frac{1}{2})^{n-1}$. Sledi:

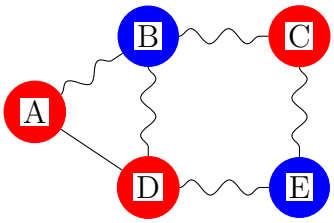
$$E[X] = E[\sum_{\pi \in S_n} I_\pi] = \sum_{\pi \in S_n} E[I_\pi] = \sum_{\pi \in S_n} (\frac{1}{2})^{n-1} = n! \cdot (\frac{1}{2})^{n-1}.$$

Po lemi 5.22 sledi želen rezultat. □

5.2.2 Maksimalni prerez grafov

Pomemben NP-težek algoritmični problem na grafih je iskanje maksimalnega prereza. Osnova za rezultate o maksimalnem prerezu je [5, 6. poglavje].

Definicija 5.26. Naj bo dan graf $G = (V, E)$. Naredimo particijo množice vozlišč V, tj. razdelimo V na dve disjunktni neprazni množici A in B, $V = A \cup B$. *Prerez je množica povezav* med A in B, število teh povezav pa je *velikost prereza.*



Slika 3: Primer grafa z maksimalnim prerezom velikosti 5, kjer imamo particijo na modra in rdeča vozlišča. Povezave v prerezu so vijugaste.

Z verjetnostno metodo zdaj pokažimo, da je v poljubnem grafu G velikost maksimalnega prereza vsaj polovica povezav grafa G.

Trditev 5.27. *Za vsak graf $G = (V, E)$, $|E| = m$ obstaja dvodelen podgraf z vsaj $\frac{m}{2}$ povezavami oz. ekvivalentno, maksimalen prerez v poljubnem grafu je vsaj $\frac{m}{2}$.*

Dokaz. Naključno skonstruiramo množici A in B, tako da vsako vozlišče grafa G damo z verjetnostjo $\frac{1}{2}$ v A in $\frac{1}{2}$ v B neodvisno od ostalih. Oštevilčimo povezave v G in jih označimo z $e_1, \ldots, e_m$. Za vsak $i = 1, \ldots, m$ definirajmo indikatorsko slučajno spremenljivko za dogodek, da povezava e_i povezuje A in B:

$$X_i = \begin{cases} 1, & e_i := u_i v_i \text{ povezuje } A \text{ in } B \\ 0, & \text{sicer.} \end{cases} \tag{5.3}$$

Računamo:

$$\begin{aligned} E[X_i] = P(X_i = 1) \quad &= P(e_i \text{ povezuje } A \text{ in } B) \\ &= P((u_i \in A \text{ in } v_i \in B) \cup (u_i \in B \text{ in } v_i \in A)) \\ &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}. \end{aligned} \tag{5.4}$$

Očitno je velikost prereza $C(A,B) = \sum_{i=1}^{m} X_i$. Velja:

$$E[C(A,B)] = E(\sum_{i=1}^{m} X_i) = \sum_{i=1}^{m} E(X_i) = \frac{m}{2}.$$

Po lemi 5.22 v grafu G obstaja prerez velikosti vsaj $\frac{m}{2}$. □

V skladu s premislekom v 5.1.2 poiščimo Las Vegas algoritem za prerez velikosti vsaj $\frac{m}{2}$. Sprva moramo učinkovito pridobiti vzorec, kar je v tem primeru nezahtevno s particijo kot v dokazu zgornje trditve 5.27. Kakšna je verjetnost p, da bo maksimalni prerez takega vzorca grafa večji ali enak $\frac{m}{2}$?

$$\begin{aligned}
\frac{m}{2} &= E(C(A,B)) \\
&= \sum_{i=1}^{\lfloor \frac{m}{2}-1 \rfloor} i \cdot P(C(A,B)=i) + \sum_{i=\frac{m}{2}}^{m} i \cdot P(C(A,B)=i) \\
&\leq \sum_{i=1}^{\lfloor \frac{m}{2}-1 \rfloor} i \cdot (1-p) + \sum_{i=\frac{m}{2}}^{m} i \cdot p \qquad (5.5) \\
&\leq (1-p) \cdot (\frac{m}{2} - 1) + pm.
\end{aligned}$$

Od tod dobimo:

$$p \geq \frac{1}{1+\frac{m}{2}}.$$

Označimo z X število potrebnih generiranj vzorcev do prvega, ki zadostuje pogoju $C(A,B) \geq \frac{m}{2}$. Slučajna spremenljivka X je porazdeljena $Geom(p)$, tj. $P(X=k) = (1-p)^{k-1}p$. Pričakovano število potrebnih vzorcev je potem enako $E[X] = \frac{1}{p} \leq \frac{m}{2}+1$. Vse kar manjka, je učinkovito preverjanje pravilnosti vzorca, t. j. velikosti prereza, kar naredimo preprosto v linearnem času (v številu povezav) s preštevanjem števila povezav skozi prerez. Skupaj dobimo polinomski Las Vegas algoritem.

5.2.3 Derandomizacija

Nekatere verjetnostne algoritme lahko derandomiziramo. Poglejmo si, kako lahko pridemo do deterministčnega algoritma za prerez velikosti vsaj $\frac{m}{2}$ v grafu z m povezavami.

Po 5.27 vemo, da za velikost prereza pri naključni particiji, kjer vsako vozlišče neodvisno damo v množico A oz. B z verjetnostjo $\frac{1}{2}$ velja $E[C(A,B)] = \frac{m}{2}$.

Poljubno oštevilčimo vozlišča v grafu: $v_1, \ldots, v_n$. Označimo z X_k slučajno spremenljivko, ki predstavlja možno izbiro položaja vozlišča v_k. Denimo, da jih deterministično damo v množico A ali B in označimo to izbiro z $X_i = x_i$ (x_i je A ali B). Denimo, da smo prvim k vozliščem dodelili množici A oz. B. Zanima nas pogojna pričakovana vrednost velikosti prereza $E[C(A,B) \mid X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k]$. Pokažimo, kako izbrati položaj za vozlišče v_{k+1}, da bomo povečali pričakovano velikost prereza.
Če na vsakem koraku izberemo tak x_{k+1}, $k = 1, \ldots, n-1$, da bo veljalo:

$$E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k] \leq E[C(A,B) \mid X_1 = x_1, \ldots, X_{k+1} = x_{k+1}],$$

potem bo sledilo:

$$\frac{m}{2} \leq E[C(A,B)] \leq E[C(A,B) \mid X_1 = x_1, \ldots, X_n = x_n].$$

Izraz na desni pa je končna velikost prereza ob takšnem načinu zaporednih izbir položaja vozlišč.

Naredimo indukcijo, v kateri bo jasno, kakšen mora biti način dodeljevanja vozlišč množicam A in B. Baza indukcije velja, saj zaradi simetrije položaj prvega vozlišča ni pomemben:

$$E[C(A,B) \mid X_1 = x_1] = E[C(A,B)].$$

V indukcijskem koraku uporabimo popolno pričakovano vrednost, ki pravi, da za particijo verjetnostnega prostora na dogodke $A_i, i = 1, \ldots, n$ velja $E[X] = \sum_i^n P(A_i)E[X \mid A_i]$:

$$\begin{aligned} &E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k] \\ &= \frac{1}{2}E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = A] + \\ &\quad + \frac{1}{2}E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = B]. \end{aligned} \tag{5.6}$$

Takoj vidimo:

$$\begin{aligned} &\max(E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = A], \\ &\qquad E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = B]) \\ &\qquad \geq E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k]. \end{aligned} \tag{5.7}$$

Torej preprosto izračunamo $E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = A]$ in $E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = B]$ in postavimo vozlišče v_{k+1} v tisto množico, ki da večjo pogojno pričakovano velikost prereza. Za izračun $E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = A]$ (oziroma $E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = B]$) potrebujemo linearno mnogo časa. Položaji prvih $k+1$ vozlišč so določeni. Preštejemo povezave med njimi, ki gredo iz A v B. Vsaka povezava, ki ni med prvimi $k+1$ vozlišči, pa prispeva k prerezu z verjetnostjo $\frac{1}{2}$, tj. $P(X_i = 1) = \frac{1}{2}$ za vse povezave, ki niso med prvimi $k+1$ vozlišči. Skupaj potem dobimo $E[C(A,B) \mid X_1 = x_1, \ldots, X_k = x_k, X_{k+1} = A]$ = (število dosedanjih povezav med A in B) + polovica preostalih povezav. Način izbire množice A oziroma B za vozlišče v_{k+1}, $j = 2, \ldots, m-1$ je torej odvisen le od števila sosednjih vozlišč v množici A in B. Izberemo tisto, kjer je sosedov manj.

Definicija 5.28. Optimizacijski problem Π je družina optimizacijskih nalog 'istega tipa': $\Pi = \pi_1, \pi_2, \ldots, \pi_i = (D_i, f_i, \text{opt})$, kjer je $D_i \neq \varnothing$ dopustno območje, $f_i : D_i \to \mathbb{R}^+$ funkcija, za katero iščemo optimalno vrednost in $\text{opt} \in \min, \max$.

Definicija 5.29. Naj bo Π optimizacijski problem in A algoritem, ki za dano nalogo iz Π vrne dopustno rešitev. Naj bo $\epsilon > 0$. Algoritem A je ϵ-aproksimacijski, če za vsako nalogo $\pi \in \Pi$ velja: $\left|\frac{\text{opt}(\pi) - f(A(\pi))}{\text{opt}(\pi)}\right| \leq \epsilon$, oz. drugače, za maksimizacijski problem: $f(A(\pi)) \geq (1-\epsilon) \cdot \text{opt}(\pi)$.

Iskanje maksimalnega prereza v danem grafu lahko torej opišemo kot optimizacijski problem na množici vseh particij vozlišč na dve množici A in B, kjer želimo maksimizirati število povezav med njima. Opisan postopek 5.2.3 za prerez grafa nam da $\frac{1}{2}$-aproksimacijski požrešni algoritem za maksimalni prerez v grafu, saj je maksimalni prerez navzgor omejen s številom povezav m.

Opomba 5.30. Za reševanja problemov iz razreda NP se poleg verjetnostnih algoritmov pogosto uporabljajo tudi aproksimacijski algoritmi.

5.2.4 Slučajni grafi

Veliko NP problemov izvira iz teorije grafov, kot na primer iskanje Hamiltonovega cikla, neodvisnostnega števila (glej 5.38), kromatičnega števila (glej 5.42) ipd. To je ena izmed motivacij za vpeljavo definicije slučajnega grafa.

Definicija 5.31. Verjetnostni prostor slučajnih grafov $G(n,p)$ ima za elementarne dogodke grafe na n vozliščih, kjer je verjetnost grafa z izbranimi m povezavami $P(G) = p^m(1-p)^{\binom{n}{2}-m}$. Verjetnost vsake povezave je p, neodvisno od ostalih. Vsako množico grafov na n vozliščih lahko smatramo za dogodek.

Opomba 5.32. Formalno bi se morali prepričati, da je na množici vseh grafov na n vozliščih Ω in družino $\mathcal{F} = \mathcal{P}(\Omega)$ z $P(G) = p^m(1-p)^{\binom{n}{2}-m}$ res definirana verjetnostna mera. Veljavnosti aksiomov Kolmogorova ni težko videti. Pokažimo le $P(\Omega) = 1$. Iz binomskega izreka sledi:

$$\sum_{G\in G(n,p)} P(G) = \sum_{m=0}^{\binom{n}{2}} \sum_{G_m} P(G_m) = \sum_{m=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m} = 1.$$

Invariante grafa so nenegativne diskretne slučajne spremenljivke, npr. število povezav, kromatično število, povezanost, neodvistnostno število... Za ilustracijo: z uporabo indikatorjev hitro ugotovimo, da je pričakovano število povezav enako $\binom{n}{2}p$ in pričakovana stopnja poljubnega vozlišča enaka $(n-1)p$.

Lema 5.33 (Pričakovani cikli)**.** *Naj bo slučajna spremenljivka $X : G(n,p) \to \mathbb{R}$ število k-ciklov v $G = (V,E) \in G(n,p)$. Potem je :*

$$E(X) = \frac{n^{\underline{k}}}{2k} p^k.$$

Dokaz povzemamo iz [8, str. 28].

Dokaz. Vpeljimo oznako $x^{\underline{n}} := \binom{x}{n} n! = x \cdot (x-1) \cdot \ldots \cdot (x-n+1)$ (beremo kot 'x na n padajoče').
Preštejmo število možnih k-ciklov na $n = |V|$ vozliščih. Pomagamo si z zaporedji vozlišč $C = v_1, v_2, \ldots, v_k$. Število zaporedij z različnimi elementi dolžine k na n elementni množici je $n^{\underline{k}}$, vsak cikel pa ustreza $2k$ takim zaporedjim, saj imamo k začetnih točk in dve smeri premikanja. Število možnih ciklov je torej $\frac{n^{\underline{k}}}{2k}$. Za vsak cikel $C_i, i = 1, \ldots, \frac{n^{\underline{k}}}{2k}$, definiramo indikatorsko slučajno spremenljivko X_i:

$$X_i = \begin{cases} 1, & C_i \subseteq G \\ 0, & \text{sicer} \end{cases}. \tag{5.8}$$

Velja:

$$E[X_i] = P(X_i = 1) = p^k, \text{ saj je verjetnost odvisna le od števila povezav.}$$

Slučajno spremenljivko X, ki predstavlja število vseh k-ciklov, zapišemo kot $X = \sum_{i=1}^{\frac{n^{\underline{k}}}{2k}} X_i$. Uporabimo linearnost pričakovane vrednosti in dobimo:

$$E[X] = E[\sum_{i=1}^{\frac{n^{\underline{k}}}{2k}} X_i] = \sum_{i=1}^{\frac{n^{\underline{k}}}{2k}} E[X_i] = \sum_{i=1}^{\frac{n^{\underline{k}}}{2k}} p^k = \frac{n^{\underline{k}}}{2k} \cdot p^k.$$

□

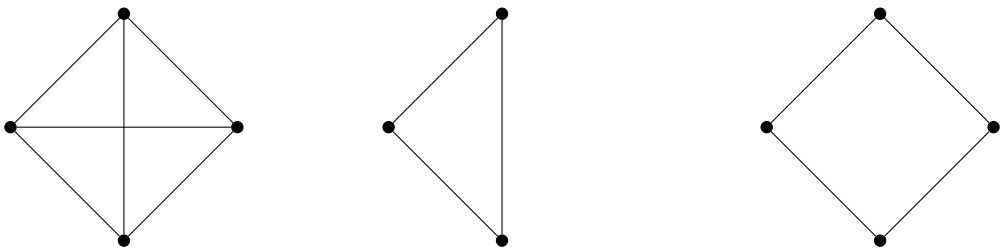
Pri slučajnih grafih se lahko vprašamo, kolikšna bo verjetnost določene lastnosti (npr. ali je graf brez povezav) za veliko število vozlišč.

Definicija 5.34. *Lastnost grafa* je razred grafov, ki za vsak vsebovan graf G vsebuje tudi vse njemu izomorfne grafe.

V prostoru $G(n, p)$ vzemimo za p neko fiksno funkcijo odvisno od števila vozlišč, $p = p(n)$ in naj bo L lastnost grafa. Zanima nas obnašanje verjetnosti $P(G \in L)$[7] za $G \in G(n, p)$, ko gre $n \to \infty$.

Definicija 5.35. Če gre za $n \to \infty$ verjetnost $P(G \in L)$ proti 1, rečemo, da ima *skoraj vsak* graf $G \in G(n, p)$ lastnost L. Če gre $P(G \in L)$ za $n \to \infty$ proti 0, rečemo, da *skoraj noben* graf $G \in G(n, p)$ nima lasnosti L.

Definicija 5.36. Graf $H = (V', E')$ je podgraf grafa $G = (V, E)$, če je $V' \subset V$ in $E' \subset E$. Podgraf $H = (V', E')$ je induciran podgraf, če ima povezana vsa vozlišča iz V', ki so povezana tudi v grafu G. Pišemo $H = G[V']$.



Slika 4: Poln graf $G = K_4$ (levo), induciran 3-cikel (center), in ne-induciran 4-cikel (desno).

Trditev 5.37. *Za vsako vrednost $p \in (0, 1)$ in vsak fiksen graf H skoraj vsak graf $G \in G(n, p)$ vsebuje induciran podgraf H.*

Dokaz povzemamo iz [8, str. 29].

[7] Neformalno bi lahko pisali P(lastnost L velja za graf G).

Dokaz. Naj bo H dan (fiksen) graf s k vozlišči. Za $n \geq k$ in fiksno množico $U \subset \{v_1, \ldots, v_n\}$, $|U| = k$, tj. fiksno podmnožico k vozlišč iz G, je $G[U]$ izomorfen H z neko verjetnostjo $r > 0$. Pri tem je verjetnost r odvisna od p, število n pa nanjo ne vpliva. Dovolj bo, da se osredotočimo na disjunktne množice U_i v grafu G, teh pa je $\lfloor \frac{n}{k} \rfloor$. Po Boolovi neenakosti 5.1 je verjetnost, da G vsebuje induciran podgraf H kvečjemu večja od verjetnosti, da je kakšen od $G[U_i]$ izomorfen H. Verjetnost P, da noben graf $G[U_i]$ ni izomorfen H pa je $(1-r)^{\lfloor \frac{n}{k} \rfloor}$, saj so ti dogodki neodvisni. Torej je:

$$P(G \text{ ne vsebuje induciranega podgrafa } H) \leq (1-r)^{\lfloor \frac{n}{k} \rfloor} \to 0, \text{ za } n \to \infty,$$

od koder sledi želen rezultat. □

5.3 Metoda izbrisa

Osnovna verjetnostna metoda deluje v redkih primerih. Pri metodi izbrisa pokažemo obstoj objekta, ki skoraj zadošča želenim pogojem in ga v nekaj korakih popravimo do želenega.

Nadaljujemo s slučajnimi grafi.

5.3.1 Največja neodvisna množica

V razdelku 5.1.1 smo definirali neodvisno množico v grafu kot množico vozlišč, med katerimi ni povezav. Dopolnimo to definicijo.

Definicija 5.38. *Neodvisnostno število* $\alpha(G)$ je moč največje neodvisne množice. Vozlišča iz dane neodvisne množice imenujemo *neodvisna vozlišča*.

Iskanje neodvisnostnega števila $\alpha(G)$ v danem grafu G je NP-težek problem. Z metodo izbrisa podajmo spodnjo mejo za $\alpha(G)$.

Trditev 5.39. *Naj bo dan graf* $G = (V, E)$, $|E| = m$, $|V| = n$. *Potem velja:* $\alpha(G) \geq \frac{n^2}{4m}$.

Dokaz je prirejen iz [5, str. 133-134].

Dokaz. Uporabimo še eno idejo, ki se pogosto uporablja pri dokazovanju z verjetnostno metodo. Izpeljemo dokaz z verjetnostjo p kot parametrom in ga določimo kasneje tako, da optimizira rezultat. Naredimo sledeče:
1. korak: Odstranimo vsako vozlišče (ter pripadajoče povezave) iz G z verjetnostjo $1-p$ neodvisno od ostalih. To pomeni, da vsako vozlišče ostane v grafu z verjetnostjo p.
2. korak: Odstranimo vsako preostalo povezavo ter eno poljubno izmed njenih krajišč. S tem dobimo neodvisno množico, saj smo odstranili vse povezave. Poglejmo si njeno velikost. Označimo $X =$ število vozlišč, ki ostanejo v grafu po 1. koraku. Po linearnosti pričakovane vrednosti je: $E[X] = n \cdot E(\text{vozlišče } v_1 \text{ ostane}) = n \cdot E[\mathbb{1}(p)] = np$. Označimo še $Y =$ število povezav v grafu G po 1. koraku. Začeli smo z grafom z m povezavami. Povezava ostane v grafu, če v njem ostaneta obe krajišči. Sledi: $E[Y] = E[\sum_{i=1}^{m} \mathbb{1}(p^2)] = m \cdot p^2$.

V 2. koraku odstranimo preostalih Y povezav in največ Y vozlišč. Dobljena neodvisna množica ima velikost vsaj $X - Y$. Pri tem je:

$$E[X - Y] = np - mp^2.$$

Vzemimo verjetnost p, ki maksimizira pričakovano vrednost neodvistnostnega števila. $E[X - Y]$ je navzdol obrnjena parabola z maksimumom v temenu. Odvajamo po p: $\frac{d}{dp}(np - mp^2) = n - 2mp$. Maksimum bo dosežen pri $p = \frac{n}{2m}$ in bo enak: $E[X - Y] = np - mp^2 = \frac{n^2}{2m} - \frac{n^2}{4m} = \frac{n^2}{4m}$. Po lemi 5.22 taka neodvisna množica res obstaja. □

Dodajamo še en rezultat povezan z neodvisnostnim številom.

Lema 5.40. *Za vsa naravna števila n in k, za katere velja $n \geq k \geq 2$, je verjetnost, da $G \in G(n, p)$ vsebuje množico s k neodvisnimi vozlišči, največ $P(\alpha(G) \geq k) \leq \binom{n}{k}(1-p)^{\binom{k}{2}}$.*

Dokaz. Takih množic je $\binom{n}{k}$. Verjetnost vsake je enaka $(1-p)^{\binom{k}{2}}$. Rezultat sledi po Boolovi neenakosti 5.1. □

5.3.2 Erdősev izrek

Kot že omenjeno, je bil madžarski matematik Paul Erdős vodilni matematik na področju verjetnostne metode. Predstavili bomo enega izmed mnogih po njem poimenovanih izrekov, ki govori o grafih z veliko ožino in velikim kromatičnim številom. Osnova za rezultate tega poglavja je [3, str. 299-301].

Definicija 5.41. Ožina grafa G je velikost najmanjšega cikla, ki je vsebovan v G.

Definicija 5.42. Kromatično število $\chi(G)$ grafa G je najmanjše število barv s katerimi lahko dobro pobarvamo vsa vozlišča. Barvanje je dobro, če sta vsaki dve sosednji vozlišči pobarvani z različnima barvama.

Izrek 5.43 (Erdős, 1959)**.** *Za vsaki naravni števili $g \in \mathbb{N}$, $k \in \mathbb{N}$ obstaja graf G z ožino $g(G) > g$ in kromatičnim številom $\chi(G) > k$.*

Pred dokazom sledi še nekaj pripravljanih rezultatov.

Iz osnov teorije grafov je znano, da lahko kromatično število ocenimo z: $\omega(G) \leq \chi(G) \leq \Delta(G) + 1$, kjer je $\omega(G)$ velikost največje klike v G in $\Delta(G)$ največja stopnja vozlišča v G.[8] Podamo lahko tudi drugačno spodnjo mejo, ki bo prav prišla v dokazu Erdősevega izreka.

Lema 5.44. *Kromatično število grafa G lahko ocenimo s $\chi(G) \geq \frac{n}{\alpha(G)}$.*

Dokaz. Pri barvanju vozlišč grafa z neodvisnostnim številom $\alpha(G)$ lahko z isto barvo pobarvamo največ $\alpha(G)$ vozlišč, potrebnih barv pa je $\chi(G)$. Torej bo $\chi(G)\alpha(G) \geq n$. □

[8] Zgornja meja predstavlja požrešno barvanje in jo dokažemo z indukcijo na število vozlišč, spodnja meja pa sledi direktno iz same definicije dobrega barvanja.

Opomba 5.45. Velja $\alpha(G) = \omega(G^c)$, kjer je G^c komplementni graf grafa G.

Cikel dolžine k lahko dobro pobarvamo s tremi barvami v primeru lihega cikla oziroma z dvema barvama v primeru sodega cikla, medtem ko je za poln graf K_n z ožino $g(G) = 3$, potrebno n barv. Intuitivno bi lahko sklepali, da veliko kromatično število zahteva dovolj veliko število povezav, da v grafu nujno pride do majhnih ciklov. Erdősev izrek pokaže, da tak intutiven premislek ne drži.

Kako se lotiti dokaza Erdősevega izreka? V skladu z lemo 5.44 je dovolj poiskati graf G brez majhnih ciklov in z majhnim neodvisnostnim številom $\alpha(G)$, natančneje z $\omega(G) > g$ in $\alpha(G) < \frac{n}{k}$. Do tega bi lahko hitro prišli z osnovno metodo, če bi uspeli najti verjetnost p pri kateri bi za slučajni graf $G \in G(n,p)$ veljalo $P(\omega(G) \leq g) < \frac{1}{2}$ in $P(\alpha(G) \geq \frac{n}{k}) < \frac{1}{2}$. Majhna vrednost p nam z veliko verjetnostjo zagotovi, da nimamo majhnih ciklov, vendar imamo z veliko verjetnostjo veliko neodvisno množico. Po drugi strani nam velika vrednost p z veliko verjetnostjo zagotovi, da je neodvisnostno število majhno, ampak se z veliko verjetnostjo tudi pojavijo majhni cikli. Kot bo pokazano kasneje pri pragovnih funkcijah (glej 5.57), moramo vzeti $p < \frac{1}{n}$, če želimo majhno verjetnost pojavitve kratkih ciklov. Hkrati se pri taki vrednosti p z majhno verjetnostjo pojavi kakršenkoli cikel (glej 5.58) in je posledično graf z veliko verjetnostjo brez ciklov. V posebnem primeru je z veliko verjetnostjo brez lihih ciklov, kar iz osnov teorije grafov vemo, da pomeni, da je dvodelen, torej ima z veliko verjetnostjo neodvisno množico večjo ali enako $\frac{n}{2}$. Postopali bomo na sledeč način. Vzemimo malenkost večji $p = n^{\epsilon-1}$. S tem pričakujemo, da bo velika neodvisna množica malo verjetna, a hkrati pričakujemo nekaj kratkih ciklov. Za dovolj majhen ϵ jih pričakujemo dovolj malo, da lahko naredimo alteracijo in izbrišemo kratke cikle ter bo tako dobljen graf še vedno dovolj velik, da bo imel veliko kromatično število. Idejo imamo, sedaj jo formalizirajmo.

Dokaz izreka 5.43. 1. Fiksirajmo števili $\epsilon \in \mathbb{R}$, $n \in \mathbb{N}$ in naj bo $p = n^{\epsilon-1}$. Primeren ϵ in n bomo določili tekom dokaza. Vzemimo sedaj poljuben graf $G \in G(n,p)$. Definiramo slučajno spremenljivko $X =$ število ciklov dolžine največ g v grafu G in X_i število ciklov dolžine i v grafu G, za $i = 1, \ldots, g$. Po lemi 5.33 je matematično upanje slučajne spremenljivke X enako:

$$\begin{aligned} E[X] &= \sum_{i=3}^{g} E[X_i] = \sum_{i=3}^{g} \frac{n^{\underline{i}}}{2i} p^i = \sum_{i=3}^{g} \frac{n^{\underline{i}}}{2i} \cdot \frac{n^{\epsilon i}}{n^i} \\ &\leq \sum_{i=3}^{g} \frac{n^{\epsilon i}}{2i} = \frac{1}{2i} \cdot (n^{3\epsilon} + \ldots + n^{g\epsilon}) \leq \frac{1}{2}(g-2) \cdot n^{g\epsilon}. \end{aligned}$$

Uporabimo neenakost Markova za X:

$$P\left(X \geq \frac{n}{2}\right) \leq \frac{E[X]}{\frac{n}{2}} \leq \frac{(g-2) \cdot n^{g\epsilon}}{n} = (g-2) \cdot n^{g\epsilon-1}.$$

Do zdaj vrednosti ϵ še nismo določili. Izberimo tak ϵ, da bo veljalo:

$$\lim_{n\to\infty} P\left(X \geq \frac{n}{2}\right) = 0.$$

Da zgornje drži, mora veljati $g\epsilon - 1 < 0$, torej vzamemo poljuben $\epsilon < \frac{1}{g}$. Pri takem ϵ za skoraj noben graf $G \in G(n,p)$ ne velja lastnost $X \geq \frac{n}{2}$. To lahko zapišemo tudi kot $P(X \geq \frac{n}{2}) = o(1)$. V posebnem primeru dobimo, da za dovolj velik $n = n_1$ velja $P(X \geq \frac{n_1}{2}) < \frac{1}{2}$.

2. V lemi 5.44 smo videli, da lahko kromatično število ocenimo z neodvisnostnim številom. Poglejmo si $P(\alpha(G) \geq r)$.

Fiksirajmo $r \in \mathbb{N}$, ki ga bomo določili kasneje. Upoštevajmo lemo 5.40, tj. $P(\alpha(G) \geq k) \leq \binom{n}{k}(1-p)^{\binom{k}{2}}$ ter neenakost iz leme 5.18 in računajmo:

$$\begin{aligned} P(\alpha(G) \geq r) &\leq \binom{n}{r}(1-p)^{\binom{r}{2}} \leq n^r (1-p)^{\frac{r(r-1)}{2}} \\ &= (n(1-p)^{\frac{r-1}{2}})^r \leq (ne^{-p\frac{(r-1)}{2}})^r \end{aligned}$$

Želimo si, da je $P(\alpha(G) \geq r) \to 0$, za $n \to \infty$. To bo veljalo recimo za tak r, pri katerem bo $ne^{-p\frac{(r-1)}{2}} < \frac{1}{\sqrt{n}}$. Ekvivalentno to zapišemo kot:

$$\begin{aligned} n^{\frac{3}{2}} &< e^{p\frac{(r-1)}{2}} \ / \ \log \\ \frac{3}{2}\log(n) &< p\frac{(r-1)}{2} \ / \ \cdot \frac{2}{3} \\ \log(n) &< \frac{p(r-1)}{3} \ / \ \cdot \frac{3}{p} \\ r &> 1 + \frac{3}{p}\log(n). \end{aligned}$$

Za velike n lahko število 1 zanemarimo in vzamemo $r = \lceil \frac{3}{p}\log(n) \rceil$. Pri takem r je torej $P((G) \geq r) = o(1)$.

V vrednosti r upoštevamo prej določeno verjetnost $p = n^{\epsilon-1}$, $\epsilon < \frac{1}{g}$ in zapišemo kot $r = \lceil \frac{3\log(n)}{n^{\epsilon-1}} \rceil$. Pokazali smo, da lastnost $P(\alpha(G) \geq r)$ ne velja za skoraj noben graf $G \in G(n,p)$. V posebnem primeru to ponovno pomeni, da lahko vzamemo dovolj velik $n = n_2$, da bo $P(\alpha(G) \geq r) < \frac{1}{2}$. Izbrati moramo dovoljšnje število vozlišč, da velja tudi $P(X \geq \frac{n}{2}) < \frac{1}{2}$ iz točke 1. Sprva vzamemo tak $n = n_1$, da bo $P(X \geq \frac{n_1}{2}) < \frac{1}{2}$; nato pa vzamemo $n_3 := \max(n_1, n_2)$ s čimer zadostimo obema pogojema. Pri takem n_3 obstaja nek graf $G \in G(n_3, p = n_3^{\epsilon-1})$ z manj kot $\frac{n_3}{2}$ cikli dolžine kvečjemu g in neodvisnostnim številom manjšim od $r = 3\log(n_3)n_3^{1-\epsilon}$. Res, iz Boolove neenakosti 5.1 sledi:

$$P\left((X < \frac{n_3}{2}) \cap (\alpha(G) < r)\right) = 1 - P\left((X \geq \frac{n_3}{2}) \cup (\alpha(G) \geq r)\right) > 0.$$

3. Sedaj sledi alteracija. Graf G popravimo v želenega, tako da se znebimo kratkih ciklov. Iz vsakega izmed manj kot $\frac{n_3}{2}$ ciklov velikosti kvečjemu g odstranimo eno poljubno vozlišče. S tem dobimo graf G', ki ima najmanj $\frac{n_3}{2}$ vozlišč in je brez kratkih ciklov, natančneje ima ožino večjo od g. Poleg tega očitno velja $\alpha(G') \leq \alpha(G)$, saj je vsaka neodvisna množica v G' tudi neodvisna množica v G. 4. Sedaj za kromatično število grafa G' uporabimo lemo 5.44 in dobimo:

$$\chi(G') \geq \frac{|V(G')|}{\alpha(G')} \geq \frac{\frac{n_3}{2}}{\alpha(G)} \geq \frac{\frac{n_3}{2}}{\frac{3\log(n_3)}{n_3^{\epsilon-1}}} = \frac{n_3^{\epsilon}}{6\log(n_3)}.$$

Hitro vidimo, da je $\lim_{n\to\infty} \frac{n^\epsilon}{6\log(n)} = \infty$. Torej bo pri dovolj velikem n veljalo tudi $\chi(G') > k$. Če n_3 še ne zadošča temu pogoju, ga povečamo dovolj, da mu in s tem dobimo potrebno število vozlišč n, ki smo ga fiksirali na začetku dokaza. Pri takem n skozi alteracijo torej dobimo graf G', ki ima tako ožino $g(G') > g$, kakor tudi kromatično število $\chi(G') > k$. □

Opomba 5.46. Eksplicitna konstrukcija takih grafov obstaja, vendar je zahtevna. Najdemo jo lahko v [9].

5.4 Metoda drugega momenta

V poglavju 4 smo pokazali neenakost Čebiševa, ki jo bomo uporabljali v nadaljevanju pri metodi drugega momenta. Ponovno bo poudarek na slučajnih grafih.

V poglavju o slučajnih grafih smo vpeljali pojem lastnosti skoraj vseh grafov. V primeru 5.37 smo videli, da nekatere lastnosti veljajo za skoraj vsak graf ne glede na vrednost p. Lastnost, da graf G vsebuje nek (fiksen) induciran podgraf H je namreč lastnost skoraj vseh grafov ne glede na izbran $p \in (0,1)$, velja na primer tako za $G \in G(n, 0.9)$ kakor tudi za $G \in G(n, \frac{1}{10^5})$. Pri nekaterih lastnostih pa obstaja meja za p, okoli katere se lastnost ravno pojavi oziroma se ne pojavi. To idejo formalizira pojem pragovne funkcije.

5.4.1 Pragovne funkcije slučajnih grafov

Definicija 5.47. Funkcija $r = r(n)$ je pragovna funkcija za lastnost L, če velja:

$$\lim_{n\to\infty} P(G \in L) = \begin{cases} 0 & \text{če: } \frac{p(n)}{r(n)} \to 0, \text{ za } n \to \infty, \\ 1 & \text{če: } \frac{p(n)}{r(n)} \to \infty, \text{ za } n \to \infty. \end{cases}$$

Opomba 5.48. Iz definicije ne moremo reči ali vedno obstaja. Da se pokazati, da obstaja za naraščajoče monotone lastnosti, to so take, pri katerih se lastnost ohrani z dodajanjem povezav v graf, npr. vsebovanost k-cikla, vsebovanost k-klike. Dokaz najdemo v [10, str. 37]. Prav tako ni enolična, saj je za pragovno funkcijo $r(n)$ tudi $k \cdot r(n)$ pragovna funkcija za isto lastnost, kjer je k poljubna konstanta.

Opomba 5.49. Za krajši zapis namesto $\lim_{n\to\infty} \frac{p(n)}{r(n)} = 0$ včasih pišemo raje $p(n) \ll r(n)$.

Pri metodi drugega momenta uporabljamo neenačbo Čebiševa, v kateri nastopa $\mathrm{Var}(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$, zato se spomnimo lastnosti variance. Za razliko od pričakovane vrednosti varianca ni linearna, aditivna je le v primeru nekoreliranih slučajnih spremenljivk. V splošnem velja:

$$\mathrm{Var}(\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n) = \sum_{i=1}^{n} \alpha_i^2 \mathrm{Var}(X_i) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_i \alpha_j \mathrm{Cov}(X_i, X_j),$$

v primeru neodvisnih slučajnih spremenljivk X_i pa dobimo:

$$\mathrm{Var}(\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n) = \sum_{i=1}^{n} \alpha_i^2 \mathrm{Var}(X_i).$$

Pri tem je kovarianca definirana kot: $\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y]$.

Iz neenačbe Čebiševa sledi naslednja trditev, ki bo uporabna v nadaljevanju.

Trditev 5.50. *Za nenegativno slučajno spremenljivko X velja:*

$$P(X = 0) \leq \frac{\mathrm{Var(X)}}{E[X]^2}.$$

Dokaz. Uporabimo neenakost Čebiševa in dobimo:

$$\begin{aligned} P(X = 0) = P(X - E[X] = -E[X]) &\leq P(|X - E[X]| = E[X]) \\ &\leq P(|X - E[X]| \geq E[X]) \leq \frac{\mathrm{Var}(X)}{(E[X])^2}. \end{aligned}$$

□

Od tod direktno sledi naslednja posledica.

Posledica 5.51. *Naj bo $\{X_n\}_{n=1}^{\infty}$ zaporedje nenegativnih slučajnih spremenljivk. Če zanj velja:*
$\lim_{n\to\infty} \frac{\mathrm{Var(X_n)}}{E[X_n]^2} = 0$, *tj.* $\mathrm{Var(X_n)} = o(E[X_n]^2)$, *potem je* $\lim_{n\to\infty} P(X_n > 0) = 1$.

Opomba 5.52. V nadaljevanju (v trditvi 5.53 in izreku 5.56) uporabimo zlorabo te notacije s tem, da zaporedje slučajnih spremenljivk predstavimo kar kot eno samo slučajno spremenljivko. Na primer, namesto zaporedja $\{X_n\}_{n=1}^{\infty}$ na pišemo kar X.

Poglejmo si primer pragovne funkcije za vsebovanost trikotnika, tj. 3-cikla v grafu. Kasneje bomo to posplošili na vsebovanost poljubnega k-cikla in vsebovanost poljubnega fiksnega grafa H.

Trditev 5.53. *Pragovna funkcija za vsebovanost trikotnika je $r(n) = \frac{1}{n}$.*

Dokaz izhaja iz [8, str. 40].

Dokaz. Označimo slučajno spremenljivko - število trikotnikov v grafu $G \in G(n,p)$ s T. Po lemi 5.33 je pričakovano število 3-ciklov v grafu $G \in G(n,p)$ enako $E[T] = \frac{n^3}{6}p^3$. Lastnost vsebovanosti trikotnika zapišemo kot $T \geq 1$ in z neenakostjo Markova dobimo: $P(T \geq 1) \leq \frac{E[T]}{1} = \frac{n^3}{6}p^3$. Če vzamemo $p \ll \frac{1}{n}$, bo res veljalo $P(T \geq 1) \leq \frac{n^3}{6}p^3 \to 0$, $n \to \infty$, s čimer je dokazan prvi del.
Pokazati moramo še drugo stran, tj. da za $p \gg \frac{1}{n}$ res velja: $P(T \geq 1) \to 1$, $n \to \infty$. Na tem mestu bomo uporabili neenačbo Čebiševa. V ta namen ocenimo varianco $\mathrm{Var}(T)$. Število trikotnikov T zapišemo kot $T = \sum_{i=1}^{\binom{n}{3}} T_i$, kjer je T_i indikator dogodka pojavitve posameznega določenega trikotnika. Varianco T lahko potem zapišemo kot

$$\mathrm{Var}(T) = \sum_{i=1}^{\binom{n}{3}} \mathrm{Var}(T_i) + \sum_{i=1}^{\binom{n}{3}} \sum_{\substack{j=1 \\ j\neq i}}^{\binom{n}{3}} \mathrm{Cov}(T_i, T_j).$$

Za vsak posamičen trikotnik T_i ocenimo $\mathrm{Var}(T_i) = E[T_i^2] - E[T_i]^2 \leq E[T_i^2] = p^3$.
Trikotnike lahko razdelimo v dve skupini, tj. na trikotnike z eno skupno povezavo

(več jih ne more biti) in na trikotnike brez skupnih povezav. Če vzamemo dva trikotnika T_i, T_j brez skupnih povezav, sta neodvisna in bo $\mathrm{Cov}(T_i, T_j) = 0$. V primeru ene skupne povezave pa ocenimo kovarianco s $\mathrm{Cov}(T_i, T_j) = E[T_iT_j] - E[T_i]E[T_j] \leq E[T_i \cdot T_j] = p^5$, saj imata skupnih 5 povezav.
Skupaj dobimo

$$\mathrm{Var}(T) \leq \binom{n}{3}p^3 + 12\binom{n}{4}p^5 \leq n^3p^3 + n^4p^5.$$

Pri tem smo upoštevali še to, da je število parov trikotnikov enako $12\binom{n}{4}$ (izberemo 4 vozlišča in eno izmed 6 možnih skupnih povezav, nato pa upoštevamo še 2 možna vrstne reda trikotnikov).
Sedaj lahko ocenimo:

$$\begin{aligned}\frac{\mathrm{Var}(T)}{E[T]^2} &\leq \frac{n^3p^3 + n^4p^5}{\left(\frac{n^3}{6}p^3\right)^2} = \frac{n^3p^3 + n^4p^5}{\left(\frac{n^3}{6}\right)^2 p^6} \\ &= \frac{n^3 + n^4p^2}{(\frac{n^3}{6})^2p^3} = \frac{n^3}{(\frac{n^3}{6})^2p^3} + \frac{n^4}{(\frac{n^3}{6})^2p}.\end{aligned}$$

Zanima nas, za kakšen p zgornja dva ulomka konvergirata k 0. Ekvivalentno, kdaj je $\lim_{n\to\infty} \frac{1}{n^3p^3} = 0$ in $\lim_{n\to\infty} \frac{1}{n^2p} = 0$. Veljati mora $p \gg \frac{1}{n}$ in $p \gg \frac{1}{n^2}$, torej je dovolj $p \gg \frac{1}{n}$. Po trditvi 5.50 je torej pri takem p res $\lim_{n\to\infty} P(T > 0) = 1$. S tem smo pokazali, da je $r(n) = \frac{1}{n}$ res pragovna funkcija za vsebovanost trikotnika. □

Opomba 5.54. Če lahko navzdol omejimo pričakovano vrednost lastnosti L, še ne nujno sledi, da ima skoraj vsak graf lastnost L. Za $p \gg \frac{1}{n}$ je pričakovano število trikotnikov $\lim_{n\to\infty} E[T] = \infty$, vendar od tod še ne sledi direktno, da ima skoraj vsak graf kak trikotnik. Lahko bi nekateri grafi imeli ogromno trikotnikov, preostali pa nobenega. Lema 5.51 pove, da se to ne more zgoditi, če je zagotovljeno tudi le majhno odstopanje od pričakovane vrednosti, kar pri lastnosti vsebovanosti trikotnika, kot smo lahko videli v dokazu, tudi je.

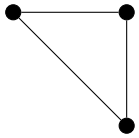
Definicija 5.55. Naj bo $G(V, E)$ graf z $v = |V(G)|$ vozlišči in $e = |E(G)|$ povezavami. *Gostota grafa* G je število $\rho(G) = \frac{e}{v}$.
Graf H je *uravnotežen*, če za vsak njegov podgraf H' velja $\rho(H') \leq \rho(H)$. Primeri uravnoteženih grafov so recimo cikli in polni grafi.

Izrek 5.56. *Naj bo H dan (fiksen) uravnotežen graf z gostoto $\rho(H) = \rho$. Potem je $r(n) = n^{\frac{-1}{\rho}} = n^{\frac{-v}{e}}$ pragovna funkcija za lastnost, da je H (ne nujno induciran) podgraf grafa $G \in G(n, p)$.*

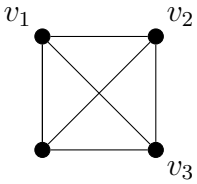
Dokaz izhaja iz [7, str. 28-29].

Dokaz. Potek dokaza bo podoben kot v 5.53. Označimo z L lastnost vsebovanosti grafa H. Pokazati moramo dve stvari in sicer, da za $p \ll n^{\frac{-v}{e}}$ velja $\lim_{n\to\infty} P(G \in L) = 0$ in za $p \gg n^{\frac{-v}{e}}$ velja $\lim_{n\to\infty}(G \in L) = 1$. Označimo z Y število podgrafov H v grafu $G \in G(n, p)$. Lastnost L lahko potem pišemo kot $L = \{G, Y(G) \geq 1\}$. Za prvo lastnost je dovolj, da je $\lim_{n\to\infty} E[Y] = 0$, saj je po neenakosti Markova $P(Y \geq 1) \leq E[Y]$. Za drugo lastnost pa uporabimo posledico 5.51. Za dokaz druge

lastnosti moramo torej pokazati, da je $\mathrm{Var(Y)} = \mathrm{o}((\mathrm{E[Y]})^2)$.
Začnimo tako, da označimo vozlišča grafa H z $\alpha = a_1, \ldots, a_v$. Vzemimo poljuben $G \in G(n,p)$. Nato vzemimo urejeno v-terico $\beta = (b_1, \ldots, b_v)$ vozlišč iz $V(G)$. Z A_β označimo dogodek, da je H podgraf na tej množici β. Natančneje s tem mislimo, da če sta povezani vozlišči a_i in a_j, sledi, da sta povezani tudi b_i in b_j oziroma še drugače; preslikava $f : a_i \mapsto b_i$, $f : \alpha \to \beta$ je homomorfizem med grafoma H in G.
Naj bo X_β indikator dogodka A_β. Definirajmo $X = \sum_\beta X_\beta$. Pazljivi moramo biti pri sklepanju, da je X pravzaprav Y torej število podgrafov H v G. To ni res, saj smo lahko kakšen graf zaradi morebitne simetrije grafa G šteli večkrat. Prikažimo na primeru. Vidimo, da podgraf H na vozliščih v_1, v_2 in v_3 dobimo večkrat, saj



(a) Graf H



(b) Graf G

lahko za množico β vzamemo poljubno permutacijo vozlišč v_1, v_2 in v_3. Kljub temu pa velja, da je $X = 0$ natanko tedaj, ko $Y = 0$ oziroma $X > 0$ natanko tedaj, ko je $Y > 0$, to pa za naše potrebe zadostuje.
Pokažimo sedaj, da za $p \ll r(n)$ velja $\lim_{n\to\infty} P(G \in L) = 0$. Po uvodnem premisleku vemo, da je smiselno izračunati $E[X]$. Za indikator X_β velja $E[X_\beta] = p^e$, saj imamo opravka s slučajnimi grafi in so povezave med seboj neodvisne. Vseh možnih v-teric različnih vozlišč v grafu $G \in G(n,p)$ pa je $n^{\underline{v}}$.

$$E[X] = E\left[\sum_\beta X_\beta\right] = n^{\underline{v}} \cdot p^e.$$

Pri tem je le verjetnost $p = p(n)$ odvisna od n, saj sta v in e konstanti. Vidimo, da za $p(n) \ll n^{\frac{-v}{e}}$ velja $\lim_{n\to\infty} E[Y] = 0$, s čimer je prvi del dokazan.
Sedaj želimo pokazati, da je $\mathrm{Var(Y)} = \mathrm{o}((\mathrm{E[Y]})^2)$. Vzemimo $p(n) \gg n^{\frac{-v}{e}}$ in si oglejmo varianco:

$$\mathrm{Var}(X) = \sum_\beta \mathrm{Var}(X_\beta) + \sum_\beta \sum_{\beta\neq\gamma} \mathrm{Cov}(X_\beta, X_\gamma) = \sum_\beta \sum_\gamma \mathrm{Cov}(X_\beta, X_\gamma).$$

Nadaljujemo podobno kot v 5.53. Vpeljimo še oznako H_β za podgraf H na množici β. Zgornje kovariance bodo neničelne le v primeru, ko si grafa H_β in H_γ delita kakšno skupno povezavo.
Denimo sedaj, da imata H_β in H_γ $t \geq 2$ skupnih vozlišč. Predpostavka, da je H uravnotežen nam pove, da imata potem H_β in H_γ največ $t\rho$ skupnih povezav. Res jih ne moreta imeti več, saj bi potem graf T, ki ga dobimo z njunim presekom imel gostoto $\rho(T) = \frac{|E(T)|}{|V(T)|} < \frac{t\rho}{t} = \rho$. Obenem pa je T podgraf H in bi moral imeti po predpostavki uravnoteženosti H manjšo gostoto kakor H.
Če imata H_β in H_γ skupnih t vozlišč ima njuna unija najmanj $2e - t\rho$ povezav. Posledično lahko njuno varianco ocenimo z:

$$\mathrm{Cov}(X_\beta, X_\gamma) \leq E[X_\beta X_\gamma] \leq p^{2e-t\rho}.$$

Preštejmo število parov množic β in γ s t skupnimi vozlišči. Iz množice n vozlišč jih izbiramo $2v-t$, kar je $\binom{n}{2v-t}$, nato pa iz te izberemo β in γ , teh možnosti pa je konstantno mnogo in neodvisno od n, ker je graf H fiksen in njegova velikost ne narašča z n. Posledično je število parov množic β in γ s t skupnimi vozlišči enako $O(\binom{n}{2v-t})$, kar pa lahko zapišemo tudi kot $O(n^{2v-t})$.

V nadaljevanju uporabimo lastnosti notacije veliki O, ki sledijo neposredno iz definicije in sicer: $O(f(n))+O(g(n)) = O(f(n)+g(n)),\ O(f(n))\cdot O(g(n)) = O(f(n)\cdot g(n))$ in $lim_{n\to\infty}O(f(n)) = O(lim_{n\to\infty}f(n))$.
Za vsako fiksno število t potem dobimo:

$$\sum_{\beta\cap\gamma=t} \mathrm{Cov}(X_\beta, X_\gamma) = O(n^{2v-t}p^{2e-t\rho}) = O((n^v p^e)^{2-\frac{t}{v}}).$$

Pri maksimalnem preseku množic β in γ, torej pri $t = v$, seveda dobimo kar varianco $\mathrm{Var}(X_\beta) = O(n^v p^e)$.
Varianco slučajne spremenljivke X nato dobimo kot:

$$\mathrm{Var}(X) = O\left(\sum_{t=2}^{v}(n^v p^e)^{2-\frac{t}{v}}\right)$$

Sedaj upoštevamo še izračunan $E[X] = n^{\underline{v}} \cdot p^e$ iz prvega dela dokaza in dobimo:

$$\lim_{n\to\infty}\frac{\mathrm{Var(X)}}{E[X]^2} = \lim_{n\to\infty} O\left(\sum_{t=2}^{v}(n^v p^e)^{-\frac{t}{v}}\right) = 0.$$

Pri tem smo upoštevali, da je $\lim_{n\to\infty} n^v p^e = \infty$, saj imamo $p(n) \gg n^{\frac{-v}{e}}$. V skladu z lemo 5.51 je $\lim_{n\to\infty} P(X > 0) = 1$ in dokaz je zaključen. □

Posledica 5.57. *Za $k \geq 3$ je $r(n) = \frac{1}{n}$ pragovna funkcija za lastnost vsebovanosti k-cikla v $G(n,p)$, saj so cikli uravnoteženi grafi in ima očitno vsak k-cikel gostoto enako 1 ne glede na k.*

Posledica 5.58. *Funkcija $r(n) = \frac{1}{n}$ je pragovna funkcija tudi za lastnost vsebovanosti kakršnegakoli cikla v $G(n,p)$.*

Dokaz. Po lemi 5.33 vemo, da za X_k - število k-ciklov velja: $E(X_k) = \frac{n^{\underline{k}}}{2k}p^k$. Potem je $X = \sum_{k=3}^{n} X_k$ število vseh ciklov za katerega velja:

$$E[X_k] = \sum_{k=3}^{n}\frac{n^{\underline{k}}}{2k}p^k \leq \sum_{k=3}^{n} n^k p^k \leq \sum_{k=3}^{\infty}(np)^k = \frac{n^3p^3}{1-np}.$$

Za $p(n) \ll \frac{1}{n}$, tj. pn = o(1), je $\lim_{n\to\infty} E[X_k] = 0$, kar po neenakosti Markova pomeni, da je $\lim_{n\to\infty} P(X \geq 1) = 0$. To da za $p(n) \gg \frac{1}{n}$ skoraj vsak graf vsebuje nek cikel pa takoj sledi iz posledice 5.57.

□

6 Zaključek

Predstavili smo osnovne načine uporabe verjetnostne metode. Začeli smo z osnovno metodo, ki z neničelno verjetnostjo obstoja nekega objekta dokazuje njegov obstoj. Nadaljevali smo z metodo izbrisa, objekt, ki ga sprva vzorčimo, nekoliko popravimo, da dobimo želene lastnosti. Pri obravnavi problemov nam prav pride tudi znanje diskretne matematike, v predstavljenih primerih gre predvsem za osnove teorije grafov. Pomagali smo si tudi z dejstvi iz verjetnosti, slučajna spremenljivka ne more biti vedno strogo večja oz. strogo manjša od pričakovane vrednosti, uporabljali smo linearnost matematičnega upanja in neenakost Markova ter neenakost Čebiševa. Ponekod smo uporabili tudi nekaj spretnosti z neenakostmi, da smo direkten rezultat z verjetnostno metodo spravili v nazornejšega, npr. v primeru 5.17. Videli smo, kako lahko iz dokaza pridobimo verjetnostni algoritem. Dokazali smo Erdősev izrek, ki se ga v literaturi pogosto navaja kot enega lepših zgledov uporabe verjetnostne metode. Začetni premisleki pri dokazovanju Erdősevega izreka so nas pripeljali do pojma slučajnega grafa in pragovnih funkcij. Nadaljnjo teorijo slučajnih grafov in mnoge druge primere uporabe verjetnostne metode lahko najdemo denimo v [6].

Slovar strokovnih izrazov

alteration method/sample and modify method metoda izbrisa
balanced (graph) uravnotežen (graf)
Chebyshev's inequality neenakost Čebiševa
chromatic number kromatično število
clique klika
girth ožina
graph property lastnost grafa
hamiltonian path hamiltonska pot
hypergraph hipergraf
independence number neodvisnostno število
independent set neodvisna množica
induced subgraph induciran podgraf
linearity of expectation linearnost pričakovane vrednosti
Markov's inequality neenakost Markova
maximum cut maksimalni prerez
probabilistic method verjetnostna metoda
random graph slučajni graf
randomized algorithm verjetnostni algoritem
Ramsey's numbers Ramseyeva števila
second moment method metoda drugega momenta
threshold function pragovna funkcija
tournament turnir

Literatura

[1] L. Barton. Ramsey theory. *Walla walla Whitman College*, 2016.

[2] C. Chachamis. Ramsey numbers. 2018. URL: `https://api.semanticscholar.org/CorpusID:264765248`.

[3] R. Diestel. *Graph theory*. Springer, Heidelberg, 2016.

[4] G. Grimmett in D. Stirzaker. *Probability and random processes*. Oxford University Press Inc., New York, 2001.

[5] M. Mitzenmacher in E. Upfal. *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press, Cambridge, 2017.

[6] N. Alon in J. H. Spencer. *The probabilistic method*. John Wiley & Sons, New Jersey, 2016.

[7] J. Vondrák in J. Matoušek. The probabilistic method. *Lecture Notes, Department of Applied Mathematics, Charles University, Prague*, 2001.

[8] B. Lužar in R. Škrekovski. Verjetnostna metoda in algoritmi, 2011. URL: `https://users.fmf.uni-lj.si/skreko/Gradiva/Verjetnostna_Metoda.pdf`.

[9] C. To Tsui. Graphs with large girth and large chromatic number. URL: `https://math.uchicago.edu/~may/REU2016/REUPapers/Tsui.pdf`.

[10] Y. Zhao. Probabilistic methods in combinatorics. URL: `https://ocw.mit.edu/courses/18-226-probabilistic-method-in-combinatorics-fall-2020/mit18_226f20_full_notes.pdf`.