Vjerojatnosna metoda u kombinatorici

Lukić, Mateja

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:757853

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2025-03-25**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Matematički odsjek

Mateja Lukić

Vjerojatnosna metoda u kombinatorici

Diplomski rad

Voditelj rada:
izv. prof. dr. sc. Vedran Krčadinac

Zagreb, srpanj 2021.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____ , predsjednik
2. _____ , član
3. _____ , član

Povjerenstvo je rad ocijenilo ocjenom _____ .

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Sadržaj

1 Uvod **1**

2 Osnove vjerojatnosti **2**

3 Primjeri vjerojatnosne metode **7**
3.1 Teorija grafova 7
3.2 Hipergrafovi 12
3.3 Kombinatorna teorija brojeva 15
3.4 Disjunktni parovi 17
3.5 Erdős-Ko-Radov teorem 19
3.6 Dominantni skupovi 21

4 Linearnost očekivanja **23**
4.1 Metoda preinaka 29

5 Drugi moment **36**

Literatura **42**

Sažetak **44**

Summary **45**

Životopis **46**

1 Uvod

Vjerojatnosna metoda predstavlja jedan od moćnih i široko korištenih alata primijenjenih u kombinatorici. Jedan od glavnih razloga za brzi razvoj metode je veliki značaj slučajnosti u teorijskoj računarnoj znanosti i statističkoj fizici. Metoda se najjednostavnije može objasniti ovako: u svrhu dokazivanja postojanja kombinatorne strukture s određenim svojstvima, konstruira se odgovarajući vjerojatnosni prostor u kojem se struktura nalazi te se pokaže kako slučajno odabrani element tog vjerojatnosnog prostora ima željena svojstva s pozitivnom vjerojatnošću. Obzirom da je metoda inicirana i razvijana radom matematičara Paula Erdősa, prikladno je metodu nazvati i „Erdősovom metodom".

U ovom radu ćemo se prvo prisjetiti nekih osnovnih definicija i teorema iz teorije vjerojatnosti koje su potrebne za razumijevanje metode. Nakon toga, u trećem poglavlju ćemo prvi puta upotrijebiti metodu na primjerima iz teorije grafova, hipergrafova, kombinatorne teorije brojeva, disjunktnih parova te dominantnih skupova. Također ćemo pogledati Erdős-Ko-Radov teorem koji se bavi presijecajućim familijama i njihovim veličinama. U četvrtom poglavlju ćemo korištenjem svojstva linearnosti očekivanja obraditi još neke primjere. Spomenut ćemo i primjere korištenja metode preinaka, kojom ćemo eliminirati nedostatke nasumičnih struktura. Na kraju, u petom poglavlju, pogledat ćemo još kako se drugi moment može iskoristiti u vjerojatnosnoj metodi.

2 Osnove vjerojatnosti

Da bismo mogli koristiti vjerojatnosnu metodu, potrebno je prisjetiti se osnovnih pojmova i rezultata iz teorije vjerojatnosti.

Definicija 2.1. *Familija $\mathcal{F}$ podskupova od Ω jest σ-*algebra skupova *ako je*

1. $\emptyset \in \mathcal{F}$,
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$,
3. $A_i \in \mathcal{F}, i \in \mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definicija 2.2. *Neka je $\mathcal{F}$ σ-algebra na skupu Ω. Uređen par $(\Omega, \mathcal{F})$ zove se* izmjeriv prostor.

Definicija 2.3. *Neka je $(\Omega, \mathcal{F})$ izmjeriv prostor. Funkcija $\mathbb{P}\colon \mathcal{F} \to \mathbb{R}$ jest* vjerojatnost *ako vrijedi*

1. $\mathbb{P}(A) \geq 0, A \in \mathcal{F}$ *i* $\mathbb{P}(\Omega) = 1$,
2. $A_i \in \mathcal{F}, i \in \mathbb{N}$ *i* $A_i \cap A_j = \emptyset$ *za* $i \neq j \Rightarrow \mathbb{P}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Definicija 2.4. *Uređena trojka $(\Omega, \mathcal{F}, \mathbb{P})$ gdje je $\mathcal{F}$ σ-algebra na Ω i $\mathbb{P}$ vjerojatnost na $\mathcal{F}$, zove se* vjerojatnosni prostor.

Navedimo i neka svojstva vjerojatnosti:

Teorem 2.5. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor. Tada vrijedi:*

1. $\mathbb{P}(\emptyset) = 0$.
2. *Ako su $A_1, ..., A_n \in \mathcal{F}$ međusobno disjunktni, tada je*
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$$
3. $A, B \in \mathcal{F}$, $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
4. $A_n \in \mathcal{F}$, $n \in \mathbb{N} \Rightarrow \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

Definicija 2.6. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor i $A, B \in \mathcal{F}$. Događaji A i B su* nezavisni *ako vrijedi*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B).$$

U ovom radu ćemo se fokusirati na diskretne vjerojatnosne prostore, odnosno na prostore gdje je $\Omega = \{\omega_1, \omega_2, ...\}$ konačan ili prebrojiv skup te ćemo uzimati $\mathcal{F} = \mathcal{P}(\Omega)$.

Definicija 2.7. *Neka je $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ diskretan vjerojatnosni prostor.* Slučajna varijabla *X je realna funkcija definirana na Ω, odnosno $X\colon \Omega \to \mathbb{R}$.*

Definicija 2.8. *Neka je $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ diskretan vjerojatnosni prostor i neka su $X_1, X_2, \ldots, X_k$ slučajne varijable na Ω. Kažemo da su $X_1, X_2, \ldots, X_k$* nezavisne slučajne varijable *ako za proizvoljne $B_i \subset \mathbb{R}, i = 1, \ldots, k$, vrijedi*

$$\mathbb{P}(X_1 \in B_1, \ldots, X_k \in B_k) = \mathbb{P}\left(\bigcap_{i=1}^{k} X_i \in B_i\right) = \prod_{i=1}^{k} \mathbb{P}(X_i \in B_i).$$

Definicija 2.9. *Neka je $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ diskretan vjerojatnosni prostor, $\Omega = \{\omega_1, \omega_2, \ldots\}$ i X slučajna varijabla na tom vjerojatnosnom prostoru. Ako red $\sum_{k=1}^{\infty} X(\omega_k)\mathbb{P}(\{\omega_k\})$ apsolutno konvergira, onda njegovu sumu zovemo* (matematičkim) očekivanjem *slučajne varijable X i označavamo sa:*

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} X(\omega_k)\mathbb{P}(\{\omega_k\}).$$

Teorem 2.10 (Linearnost očekivanja)**.** *Neka su $X_1, \ldots, X_n$ slučajne varijable na vjerojatnosnom prostoru $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ koje imaju konačno očekivanje i $c_i \in \mathbb{R}, i \in \{1, \ldots, n\}$. Tada i slučajna varijabla $X = \sum_{i=1}^{n} c_i X_i$ ima konačno očekivanje i vrijedi*

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} c_i \mathbb{E}[X_i]. \tag{1}$$

Teorem 2.11. *Neka su slučajne varijable $X_1, \ldots X_n$ nezavisne i neka postoji $\mathbb{E}[X_k]$ $(k = 1, \ldots, n)$. Tada slučajna varijabla $\prod_{k=1}^{n} X_k$ ima očekivanje i vrijedi*

$$\mathbb{E}\left[\prod_{k=1}^{n} X_k\right] = \prod_{k=1}^{n} \mathbb{E}[X_k].$$

Vrijednosti $X(\omega_k)$ označavat ćemo često a_k, a $\mathbb{P}(\{\omega_k\})$ kao p_k.

Teorem 2.12. *Neka je X slučajna varijabla i neka je $g\colon \mathbb{R} \to \mathbb{R}$ proizvoljna funkcija. Tada vrijedi*

$$\mathbb{E}[g(X)] = \sum_{i=1}^{\infty} g(a_i)p_i$$

uz pretpostavku da red apsolutno konvergira.

Definicija 2.13. *Neka je X slučajna varijabla i neka postoji $\mathbb{E}[X]$.* Varijancu *od X definiramo kao*

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

ako to očekivanje postoji.

Teorem 2.14. *Neka je X slučajna varijabla i postoji $\mathrm{Var}X$. Tada vrijedi:*

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

Dokaz.

$$\begin{aligned}\mathrm{Var}X =& \mathbb{E}(X - \mathbb{E}[X])^2\\ =& \mathbb{E}\big(X^2 - 2X \cdot \mathbb{E}[X] - \mathbb{E}[X]^2\big)\\ =& \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2\\ =& \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2\\ =& \mathbb{E}[X^2] - \mathbb{E}[X]^2,\end{aligned}$$

gdje smo koristili linearnost očekivanja i činjenicu da je očekivanje konstante jednako toj konstanti, odnosno $\mathbb{E}\big[\mathbb{E}[X]^2\big] = \mathbb{E}[X]^2$.

□

Definicija 2.15. *Neka su X_1 i X_2 slučajne varijable na Ω i neka postoje $\mathbb{E}X_1^2$ i $\mathbb{E}X_2^2$.* Kovarijancu *slučajnih varijabli X_1 i X_2 definiramo kao*

$$\mathrm{Cov}(X_1, X_2) = \mathbb{E}\left[(X_1 - \mathbb{E}X_1)(X_2 - \mathbb{E}X_2)\right].$$

Kovarijancu ćemo koristiti i u idućem obliku

$$\mathrm{Cov}(X_1, X_2) = \mathbb{E}(X_1X_2) - \mathbb{E}X_1 \cdot \mathbb{E}X_2.$$

Definicija 2.16. *Kažemo da su slučajne varijable X_1 i X_2 na Ω* nekorelirane *ako vrijedi* $\mathrm{Cov}(X_1, X_2) = 0$.

Napomena 2.17. *Iz definicija o nekoreliranosti i nezavisnosti dvije slučajne varijable može se vidjeti da nezavisnost povlači nekoreliranost, ali ne vrijedi i obratno.*

Navedimo još osnovne definicije vezane uz neprekidne slučajne varijable. U nastavku vjerojatnosni prostor ne mora biti diskretan.

Definicija 2.18. *Neka je $\mathbb{R}$ skup realnih brojeva. Tada σ-algebru generiranu familijom svih otvorenih skupova na $\mathbb{R}$ zovemo* σ-algebra Borelovih skupova na $\mathbb{R}$ *i označavamo sa $\mathcal{B}$. Elemente σ-algebre $\mathcal{B}$ zovemo* Borelovi skupovi.

Definicija 2.19. *Funkciju* $g : \mathbb{R} \to \mathbb{R}$ *zovemo* Borelovom funkcijom *ako je* $g^{-1}(B) \in \mathcal{B}$ *za svaki* $B \in \mathcal{B}$.

Definicija 2.20. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor. Funkcija* $X : \Omega \to \mathbb{R}$ *je* slučajna varijabla *na* Ω *ako je* $X^{-1}(B) \in \mathcal{F}$ *za proizvoljno* $B \in \mathcal{B}$*, to jest* $X^{-1}(\mathcal{B}) \subset \mathcal{F}$.

Definicija 2.21. *Neka je* X *slučajna varijabla na* Ω. Funkcija distribucije *od* X *je funkcija* $F_X : \mathbb{R} \to [0, 1]$ *definirana sa*

$$\begin{aligned} F_X(x) = \mathbb{P}_X\left(\langle -\infty, x]\right) &= \mathbb{P}(X^{-1}\langle -\infty, x]) \\ &= \mathbb{P}\{\omega \in \Omega : X(\omega) \leq x\} = \mathbb{P}\{X \leq x\}, x \in \mathbb{R}. \end{aligned}$$

Definicija 2.22. *Neka je* X *slučajna varijabla na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathbb{P})$ *i neka je* F_X *njezina funkcija distribucije. Kažemo da je* X *apsolutno neprekidna ili kraće,* neprekidna slučajna varijabla *ako postoji nenegativna realna Borelova funkcija* $f : \mathbb{R} \to \mathbb{R}^+$ *takva da je*

$$F_X(x) = \int_{-\infty}^{x} f(t) d\lambda(t), \quad x \in \mathbb{R}.$$

Funkciju f *nazivamo* funkcijom gustoće *slučajne varijable* X.

Definicija 2.23. *Neka su* $m, \sigma \in \mathbb{R}, \sigma > 0$. *Neprekidna slučajna varijabla* X *ima* normalnu distribuciju *s parametrima* m *i* σ^2 *ako je njena funkcija gustoće*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

To označavamo kao $X \sim N(m, \sigma^2)$.

Idući teorem je jedan od klasičnih centralnih graničnih teorema.

Teorem 2.24 (*Levy*)**.** *Neka je* $(X_n, n \in \mathbb{N})$ *niz nezavisnih jednako distribuiranih slučajnih varijabli s očekivanjem* m *i varijancom* $\sigma^2, 0 < \sigma^2 < \infty$ *i neka je* $S_n = \sum_{k=1}^{n} X_k$. *Tada vrijedi*

$$\frac{S_n - \mathbb{E}S_n}{\sigma\sqrt{n}} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{za } n \to \infty$$

Teorem 2.25. *Neka je* $Y \sim N(m, \sigma^2)$ *te neka je slučajna varijabla* X *zadana kao* $X = |Y|$. *Kažemo da* X *ima* presavijenu normalnu distribuciju. *Tada vrijedi*

$$\mathbb{E}X = m_f = \sqrt{\frac{2}{\pi}} \sigma e^{-\frac{m^2}{2\sigma^2}} + m\left[1 - 2F\left(-\frac{m}{\sigma}\right)\right],$$

gdje je

$$F(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{t^2}{2}} dt.$$

Također vrijedi i

$$\mathrm{Var}X = \sigma_f^2 = m^2 + \sigma^2 - m_f^2.$$

Napomena 2.26. *Ako je $Y \sim N(0,1)$, odnosno jedinična normalna distribucija, tada za $X = |Y|$ vrijedi $m_f = \sqrt{\frac{2}{\pi}}$.*

Teorem 2.27 (**Čebiševljeva nejednakost**)**.** *Neka je X slučajna varijabla s konačnim očekivanjem m i konačnom varijancom σ^2. Tada za svaki $\lambda > 0$ vrijedi*

$$\mathbb{P}\Big(|X - m| \geq \lambda\sigma\Big) \leq \frac{1}{\lambda^2}.$$

3 Primjeri vjerojatnosne metode

U uvodu je rečeno da je metoda moćan alat te je ugrubo opisana, međutim metodu je najbolje ilustrirati kroz primjere. Stoga ćemo pogledati nekoliko primjera iz različitih područja kombinatorike.

3.1 Teorija grafova

Prvo ćemo se zadržati na par primjera iz teorije grafova vezanih uz Ramseyeve brojeve i turnire.

Definicija 3.1. Ramseyev broj $R(k, l)$ *je najmanji prirodan broj* n *takav da u bilo kojem bojenju bridova potpunog grafa* K_n *u crvenu i plavu boju postoji potpuni plavi podgraf* K_k *ili potpuni crveni podgraf* K_l.

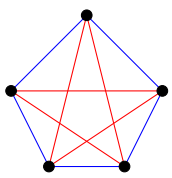
Frank P. Ramsey (1903. – 1930.) je pokazao kako je $R(k, l)$ konačan broj za bilo koje prirodne brojeve l i k, međutim bilo bi dobro znati i donju granicu za $R(k, l)$. Pogledajmo primjere nekih Ramseyevih brojeva.

$R(1, 1) = 1$ jer potpun podgraf K_1 nema bridova.

$R(2, 2) = 2$ jer su sva bojenja potpunog grafa K_2:

Slika 1: Bojenja potpunog grafa K_2

$R(3, 3) = 6$, što ćemo pokazati preko dvije nejednakosti: $R(3, 3) \leq 6$ i $R(3, 3) > 5$. Dokažimo prvo $R(3, 3) \leq 6$. Ako imamo potpun graf K_6 tada za bilo koji vrh v znamo da je susjedan s točno 5 vrhova. Barem tri brida koja su susjedni s v su iste boje, neka su to bridovi $\{v, s\}, \{v, t\}$ i $\{v, u\}$. Ako je bar jedan od bridova $\{s, t\}, \{s, u\}$ i $\{t, u\}$ te iste boje, onda imamo jednobojni K_3. Ako nijedan od tih bridova nije te boje, to znači da su sva tri brida iste boje, stoga oni zajedno čine jednobojni K_3. Drugu nejednakost dokažemo tako da nađemo bojenje od K_5 koje ne sadrži jednobojni K_3:



Slika 2: Bojenje potpunog grafa K_5

Propozicija 3.2. *Ako je* $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$*, tada je* $R(k,k) > n$*. Stoga je* $R(k,k) > \lfloor 2^{k/2} \rfloor$ *za sve* $k \geq 3$*.*

Dokaz. Uzmimo nasumično neko bojenje bridova potpunog grafa K_n u dvije boje tako da je svaki brid neovisno obojan s jednakom vjerojatnošću u plavu ili crvenu boju.

Za bilo koji fiksirani skup S od k vrhova, neka A_S označava događaj da je inducirani graf K_k na S jednobojan. Vrijedi

$$\mathbb{P}(A_S) = \frac{2}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$$

jer ukupno imamo $\binom{k}{2}$ bridova i svaki se može obojati na 2 načina, a događaj A_S vrijedi u 2 slučaja, ili su svi bridovi plavi ili su svi bridovi crveni. Skup S možemo odabrati na $\binom{n}{k}$ načina pa je vjerojatnost da će se dogoditi bar jedan događaj A_S najviše $\binom{n}{k}2^{1-\binom{k}{2}} < 1$. To znači da se s pozitivnom vjerojatnošću ne događa ni jedan A_S, odnosno $R(k,k) > n$.

Ako uzmemo $k \geq 3$ i $n = \lfloor 2^{k/2} \rfloor$ tada imamo:

$$\begin{aligned}\binom{n}{k}2^{1-\binom{k}{2}} &= \frac{n(n-1)\cdots(n-k+1)}{k!} \cdot 2^{1-\frac{k^2}{2}+\frac{k}{2}} < \frac{n^k}{k!} \cdot \frac{2^{1+k/2}}{2^{k^2/2}} \\ &= \frac{2^{1+k/2}}{k!} \cdot \frac{n^k}{2^{k^2/2}} = \frac{2^{1+k/2}}{k!} \cdot \left(\frac{n}{2^{k/2}}\right)^k < 1 \cdot 1^k = 1,\end{aligned}$$

gdje smo koristili da je $f(k) = \frac{2^{1+k/2}}{k!}$ padajuća funkcija za $k \in \mathbb{N}$ te za sve $k \geq 3$ vrijedi $f(k) \leq f(3) = \frac{2\sqrt{2}}{3} < 1$. Stoga je $R(k,k) > \lfloor 2^{k/2} \rfloor$.

□

Ovaj jednostavan primjer prikazuje bit vjerojatnosne metode. U primjeru prije propozicije 3.2 dokazali smo $R(3,3) > 5$ tako da smo eksplicitno konstruirali bojenje od K_5 koje ne sadrži jednobojni trokut K_3. S druge strane, u propoziciji dokazujemo $R(k,k) > n$ bez da eksplicitno konstruiramo bojenje od K_n koje ne sadrži jednobojni K_k, nego samo pokazujemo vjerojatnosnom metodom da takvo bojenje postoji. Tako za $k = 3$ i $n = 3$ iz propozicije dobijemo $\binom{3}{3} \cdot 2^{1-\binom{3}{2}} = \frac{1}{4}$ stoga $R(3,3) > 3$.

Možemo primijetiti kako vjerojatnost nije potrebna u ovom primjeru. Propoziciju 3.2 možemo dokazati prebrojavanjem: provjerimo da je broj svih bojenja bridova K_n u dvije boje veći od broja onih koji sadrže jednobojni K_k. Bojenja bridova u K_n u dvije boje ukupno ima $2^{\binom{n}{2}}$. Odabrati k vrhova koji tvore jednobojni K_k možemo na $2\binom{n}{k}$ načina. Ostane još $\binom{n}{2} - \binom{k}{2}$ bridova

koje trebamo obojati. Neka bojenja koja sadrže jednobojni K_k bojali smo više puta, ali svako takvo bojenje se može dobiti na ovaj način. Stoga imamo da je broj bojenja K_n u kojima postoji jednobojni K_k manji ili jednak od:

$$2 \cdot \binom{n}{k} \cdot 2^{\binom{n}{2}-\binom{k}{2}} = \binom{n}{k} \cdot 2^{1-\binom{k}{2}} \cdot 2^{\binom{n}{2}}.$$

Iskoristimo li nejednakost iz uvjeta teorema $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$ i pomnožimo ju s $2^{\binom{n}{2}}$ dobivamo:

$$2 \cdot \binom{n}{k} \cdot 2^{\binom{n}{2}-\binom{k}{2}} < 2^{\binom{n}{2}}.$$

Odnosno, dobivamo da je broj svih bojenja grafa K_n u dvije boje veći od broja onih koji sadrže jednobojni K_k.

Kako su većinom vjerojatnosni prostori u kombinatornim problemima konačni, većina dokaza vjerojatnosne metode može se formulirati kao prebrojavanje. Međutim, u praksi je praktičnije koristiti vjerojatnost.

Promotrimo problem generiranja prebrojavanja koji bi pokazivao da postoji bojenje bridova K_n u dvije boje bez jednobojnog $K_{2\log_2 n}$. S obzirom da je ukupan broj bojenja konačan, možemo tražiti po svim bojenjima. Međutim, takva procedura može zahtijevati $2^{\binom{n}{2}}$ koraka, vrijeme koje je eksponencijalno u odnosu na veličinu problema, koja je $\binom{n}{2}$. Algoritmi čije je vrijeme izvršavanja veće od polinomijalne složenosti smatraju se nepraktičnima. Iscrpno pretraživanje svih bojenja u tom smislu nije prihvatljivo i zbog toga je korisno da je dokaz propozicije nekonstruktivan. Međutim, možemo uočiti kako dokaz može koristiti da se efektivno nađe bojenje koje je vrlo vjerojatno onakvo kakvo tražimo. To je zato što, za velike k, ako je $n = \lfloor 2^{k/2} \rfloor$, tada

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

Stoga nasumično bojenje od K_n vrlo vjerojatno neće sadržavati jednobojni $K_{2\log n}$. Ako iz nekog razloga moramo predstaviti bojenje bridova K_{1024} u dvije boje bez jednobojnog K_{20} možemo proizvesti nasumično bojenje bacanjem pravilnog novčića $\binom{1024}{2}$ puta. Vjerojatnost da sadrži jednobojni K_{20} je manja ili jednaka od:

$$\frac{2\binom{n}{k}2^{\binom{n}{2}-\binom{k}{2}}}{2^{\binom{n}{2}}} = \binom{n}{k}2^{1-\binom{k}{2}} \leq \frac{2^{1+\frac{k}{2}}}{k!}\left(\frac{n}{2^{k/2}}\right)^k = \frac{2^{11}}{20!} \approx 8.42 \cdot 10^{-16},$$

a ona je vjerojatno manja nego vjerojatnost da napravimo pogrešku u nekom rigoroznom dokazu da je to bojenje dobro. Stoga u mnogim slučajevima vjerojatnosna, nekonstruktivna metoda daje efektivne vjerojatnosne algoritme.

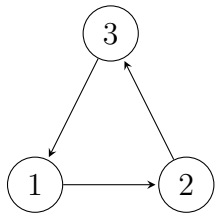
Metoda ima široku primjenu u kombinatorici, a neke jednostavne primjere ćemo prikazati u ostatku ovog poglavlja.

Definicija 3.3. Turnir *na skupu V s n elemenata je usmjerenje $T = (V, E)$ bridova potpunog grafa na skupu vrhova V.*

Stoga, za svaka dva različita $x, y \in V$, ili je $(x, y) \in E$, ili je $(y, x) \in E$. Naziv *turnir* dolazi prirodno ako skup V gledamo kao skup igrača u kojem se svaka dva igrača natječu te samo jedan od njih pobjeđuje. Ako je $(x, y) \in E$ onda kažemo da igrač x pobjeđuje igrača y i obrnuto.

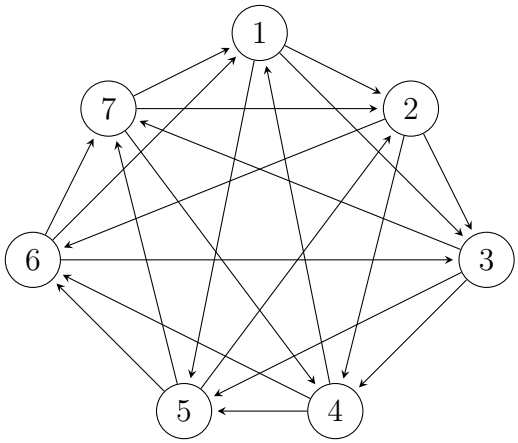
Definicija 3.4. *Kažemo da turnir T* ima svojstvo S_k *ako, za svaki skup od k igrača, postoji igrač koji pobjeđuje tih k igrača.*

Primjer 3.5. *Pogledajmo usmjereni trokut $T_3 = (V, E)$ gdje Primijetimo da*



Slika 3: Usmjereni trokut T_3

ovaj usmjereni trokut ima svojstvo S_1. Navedimo i primjer turnira na sedam vrhova koji ima svojstvo S_2:



Slika 4: Turnir sa svojstom S_2

Postavlja se pitanje: postoji li za svaki $k \in \mathbb{N}$ turnir T koji ima svojstvo S_k?

Teorem 3.6. *Ako vrijedi* $\binom{n}{k}(1-2^{-k})^{n-k} < 1$, *tada postoji turnir s* n *vrhova koji ima svojstvo* S_k.

Dokaz. Uzmimo na slučajan način turnir na skupu $V = \{1, \ldots, n\}$. Za svaki fiksan podskup K od V veličine k i $v \in V \setminus K$, neka je A_K^v događaj da v ne pobjeđuje sve vrhove iz K, a A_K događaj da niti jedan $v \in V \setminus K$ ne pobjeđuje sve vrhove iz K. Tada je $\mathbb{P}(A_K^v) = 1 - 1/2^k = 1 - 2^{-k}$, a s obzirom da u $V \setminus K$ postoji $n-k$ vrhova, zbog nezavisnosti događaja A_K^v vrijedi: $\mathbb{P}(A_K) = \mathbb{P}(\bigcap_{v \in V} A_K^v) = (1-2^{-k})^{n-k}$. Iz toga slijedi:

$$\mathbb{P}\left[\bigcup_{\substack{K \subset V \\ |K|=k}} A_K\right] \leq \sum_{\substack{K \subset V \\ |K|=k}} \mathbb{P}(A_K) = \binom{n}{k}(1-2^{-k})^{n-k} < 1,$$

odnosno s pozitivnom vjerojatnošću ne događa se ni jedan događaj A_K iz čega slijedi da postoji turnir s n vrhova koji ima svojstvo S_k.

□

Za $k = 2$ možemo uvrštavanjem brojeva primijetiti kako teorem garantira postojanje turnira sa svojstvom S_2 tek za $n = 21$, a u primjeru smo vidjeli da postoji već za $n = 7$. Međutim, za fiksni k i za dovoljno velik n uvjet teorema će biti zadovoljen. Naime, vrijedi

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!} = \frac{n^k}{k^k} \cdot \frac{k^k}{k!} < \frac{n^k}{k^k} \cdot \sum_{j=0}^{\infty} \frac{k^j}{j!} = \left(\frac{en}{k}\right)^k.$$

Zbog nejednakosti $1 - x < e^{-x}$,

$$(1-2^{-k})^{n-k} < \left(\sum_{j=0}^{\infty} \frac{\left(-2^{-k}\right)^j}{j!}\right)^{n-k} = \left(e^{-2^{-k}}\right)^{n-k} = e^{-(n-k)/2^k}.$$

Sada imamo funkciju $g(n) = \left(\frac{en}{k}\right)^k \cdot e^{-(n-k)/2^k}$ za koju će uvijek vrijediti $g(n) > f(n) = \binom{n}{k}(1-2^{-k})^{n-k} > 0$. Isto tako, uz primjenu L'Hôspitalovog pravila, za funkciju $g(n)$ vrijedi:

$$\begin{aligned}
\lim_{n\to\infty} \left(\frac{en}{k}\right)^k \cdot e^{-(n-k)/2^k} &= \lim_{n\to\infty} \frac{\left(\frac{en}{k}\right)^k}{e^{-(n-k)/2^k}} \\
&= \lim_{n\to\infty} \frac{k \cdot \frac{e}{k} \cdot \left(\frac{en}{k}\right)^{k-1}}{-\frac{1}{2^k} \cdot e^{-(n-k)/2^k}} \\
&= \cdots = \lim_{n\to\infty} \frac{k! \cdot \left(\frac{e}{k}\right)^k}{\left(-\frac{1}{2^k}\right)^k \cdot e^{-(n-k)/2^k}} = 0,
\end{aligned}$$

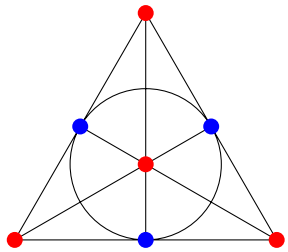
što nam govori kako sigurno postoji prirodan broj n za koji vrijedi $f(n) < g(n) < 1$.

G. i E. Szekeres [16] su dokazali da vrijedi $f(k) \geq c_1 \cdot k \cdot 2^k$. U istom radu su pokazali kako je $f(3) = 19$ te postavili hipotezu da je $f(k) = (k-1) \cdot 2^k + 3$ za svaki $k > 1$.

3.2 Hipergrafovi

Definicija 3.7. Hipergraf *je par* $H = (V, E)$ *gdje je* V *konačan skup čije elemente nazivamo* vrhovima*, a* E *je familija podskupova od* V *koje nazivamo* bridovima*. Hipergraf je* n-uniforman *ako svaki njegov brid sadrži točno* n *vrhova. Za hipergraf kažemo da je* dvoobojiv *ako postoji bojenje vrhova od* V *u dvije boje takvo da nijedan brid nije jednobojan.*

Označimo još s $m(n)$ minimalni mogući broj bridova n-uniformnog hipergrafa koji nije dvoobojiv. Lako se vidi da je $m(1) = 1$ i $m(2) = 3$, a Erdős i Hajnal [9] su pokazali kako je $m(3) = 7$. Prikažimo i primjer takvog hipergrafa te dokažimo da on zaista nije dvoobojiv:



Slika 5: 3-uniforman hipergraf koji nije dvoobojiv

.

Hipergraf koji se nalazi na slici 5 se sastoji od 7 vrhova i 7 bridova koji sadrže po 3 vrha. Poznat je kao Fanova ravnina, a to je i najmanja projektivna ravnina reda 2. Ako preuzmemo terminologiju u projektivnoj geometriji, vrhove nazivamo točkama, a bridove pravcima. Fanova ravnina ima svojstvo da kroz svaku točku prolaze tri pravca, a kroz svake dvije točke prolazi samo jedan pravac. Da dokažemo da Fanova ravnina nije dvoobojiva, zapravo trebamo dokazati da za svako bojenje 7 točaka u Fanovoj ravnini u crvenu ili plavu boju, uvijek postoji barem jedan jednobojni pravac, odnosno pravac kojemu su sve točke koje se na njemu nalaze u istoj boji. U bilo kojem bojenju 7 točaka, po Dirichletovom principu, postoje barem 4 točke koje su bojane istom bojom. Bez smanjenja općenitosti, neka je to crvena boja. Ako su među

te 4 crvene točke neke 3 na istome pravcu, tada imamo crveni pravac. U suprotnom, te 4 točke čine potpuni četverovrh, a svaka od preostale 3 točke leži na dva od 6 pravaca koji prolaze kroz vrhove četverovrha. Ako je barem jedna od te 3 točke crvena, tada se onda, zajedno sa 2 vrha četverovrha, nalazi na crvenom pravcu. U suprotnom su sve 3 točke plave boje, a one leže na sedmom pravcu Fanove ravnine, te stoga imamo plavi pravac.

Propozicija 3.8 (Erdős [10]). *Svaki n-uniformni hipergraf s manje od 2^{n-1} bridova je dvoobojiv. Stoga, $m(n) \geq 2^{n-1}$.*

Dokaz. Neka je $H = (V, E)$ n-uniformni hipergraf s manje od 2^{n-1} bridova. Obojimo V na slučajan način u dvije boje. Za svaki brid $e \in E$, neka je A_e događaj da je brid e jednobojan. Očito, $\mathbb{P}[A_e] = \frac{2}{2^n} = 2^{1-n}$. Tada,

$$\mathbb{P}\left[\bigcup_{e \in E} A_e\right] \leq \sum_{e \in E} \mathbb{P}[A_e] < 2^{n-1} \cdot 2^{1-n} = 1$$

te stoga s pozitivnom vjerojatnošću se zbiva događaj da bojenje nema jednobojnih bridova, odnosno H je dvoobojiv.

□

Mnogi rezultati koji se susreću u vjerojatnosnoj metodi su asimptotski, stoga navodimo standardnu asimptotsku notaciju. Za svake dvije funkcije $f, g : \mathbb{N} \to \mathbb{R}$ pišemo:

- $f = O(g)$ ako postoji pozitivna konstanta c i $n_0 \in \mathbb{N}$ takvi da je $f(n) \leq c \cdot g(n)$ za sve $n \geq n_0$.
- $f = \Omega(g)$ ako vrijedi $g = O(f)$.
- $f = \Theta(g)$ ako vrijedi $f = O(g)$ i $f = \Omega(g)$.
- $f = o(g)$ ako kvocijent f/g teži prema 0 kada varijabla teži prema ∞.
- $f \sim g$ ako $f = (1 + o(1))g$, odnosno f/g teži prema 1 kada varijabla teži prema ∞.

Najpoznatiju gornju ocjenu za $m(n)$ dobijemo korištenjem vjerojatnosnog argumenta „naopačke". Bridove biramo nasumično, a bojanje vrhova je zadano. Fiksiramo V s v vrhova, a kasnije ćemo optimizirati v. Neka je χ bojenje od V takvo da je a vrhova u jednoj boji, a $b = v - a$ u drugoj. Neka je $S \subset V$ uniformno odabran n-skup. Tada

$$\mathbb{P}[S \text{ je jednobojan pod } \chi] = \frac{\binom{a}{n} + \binom{b}{n}}{\binom{v}{n}}. \tag{2}$$

Pretpostavimo radi jednostavnosti da je v paran broj, te bez smanjenja općenitosti pretpostavimo da je $a \leq b$. Pokazat ćemo da se minimum gornjeg izraza postiže za $a = b$. Ako uvedemo oznaku $c = \frac{v}{2}$ i $d = c - a$ možemo našu pretpostavku zapisati u obliku

$$\binom{c-d}{n} + \binom{c+d}{n} \geq 2\binom{c}{n}.$$

Uvedimo li niz $a_j = \binom{c-j}{n} + \binom{c+j}{n}$ vidimo da za $j = 0$ dobivamo desnu stranu izraza. Dokazat ćemo općenitije, da vrijedi $a_{j+1} \geq a_j$ za svaki $j = 0, 1, \ldots, n$, a onda slijedi da $a_j \geq a_0$ za svaki $j = 0, 1, \ldots, n$.

$$\begin{aligned}
a_{j+1} - a_j &= \binom{c+j+1}{n} - \binom{c+j}{n} - \binom{c-j}{n} + \binom{c-j-1}{n} \\
&= \binom{c+j}{n-1} - \binom{c-j-1}{n-1} \\
&= \binom{c+j}{n-1} - \binom{c+j-1}{n-1} + \binom{c+j-1}{n-1} + \cdots + \binom{c-j}{n-1} - \binom{c-j-1}{n-1} \\
&= \binom{c+j-1}{n-2} + \binom{c+j-2}{n-2} + \cdots + \binom{c-j-1}{n-2} \geq 0,
\end{aligned}$$

gdje smo iskoristili Pascalov identitet $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$. Ovo povlači da izraz (2) minimum postiže u $a = b$. Stoga,

$$\mathbb{P}[S \text{ je jednobojan pod } \chi] \geq p,$$

gdje smo definirali

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}}.$$

Neka su sada $S_1, \ldots S_m$ uniformno i nezavisno odabrani n-skupovi, gdje ćemo m tek odrediti. Može se dogoditi da neki od skupova S_i budu odabrani više puta, međutim tada bismo imali još manju familiju koja nije dvoobojiva. Za bojenje χ, neka je A_χ^i događaj da S_i nije jednobojan, a A_χ događaj da nijedan S_i nije jednobojan. Tada je $\mathbb{P}[A_\chi^i] \leq 1 - p$, a po nezavisnosti događaja S_i, $\mathbb{P}[A_\chi] = \prod_{i=1}^m \mathbb{P}[A_\chi^i] \leq (1-p)^m$.

Postoji 2^v bojenja, stoga

$$\mathbb{P}\Big[\bigcup_\chi A_\chi\Big] \leq 2^v(1-p)^m.$$

Kada izraz s desne strane nejednakosti iznosi manje od 1, tada postoje $S_1, \ldots S_m$ takvi da ne vrijedi ni jedan A_χ, odnosno, $m(n) \leq m$.

Možemo iskoristiti nejednakost $1-p \leq e^{-p}$, koja vrijedi za sve pozitivne p, a za male vrijednosti p su vrijednosti približno jednake. Kad je

$$m = \left\lceil \frac{v \ln 2}{p} \right\rceil,$$

tada vrijedi $2^v(1-p)^m < 2^v e^{-pm} \leq 2^v e^{-p\frac{v\ln 2}{p}} = 1$ pa je $m(n) \leq m$. Sada trebamo pronaći v tako da minimiziramo v/p, da bi m bio što manji. Vrijednost p možemo interpretirati kao dvostruku vjerojatnost izvlačenja n bijelih kuglica iz vreće u kojoj se nalazi $v/2$ bijelih i $v/2$ crnih kuglica, bez vraćanja u vreću. Koristimo aproksimaciju drugog reda

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}} = 2^{1-n}\prod_{i=0}^{n-1}\frac{v-2i}{v-i} \approx 2^{1-n}\prod_{i=0}^{n-1} e^{-\frac{i}{v}} \approx 2^{1-n}e^{-n^2/2v}$$

gdje smo procijenili članove produkta dok god je $v \gg n^{3/2}$:

$$\frac{v-2i}{v-i} = 1 - \frac{i}{v-i} = 1 - \frac{i}{v} + \frac{i^2}{v^2 - vi} = 1 - \frac{i}{v} + O\left(\frac{i^2}{v^2}\right) = e^{-i/v + O(i^2/v^2)}.$$

Traženje minimuma od p/v sada postaje pitanje traženja minimuma funkcije $h(v) = \frac{2^{1-n}e^{-n^2/2v}}{v}$. Deriviranjem funkcije $h(v)$ dobivamo

$$\begin{aligned}
h'(v) &= \frac{2^{1-n}e^{-n^2/2v} \cdot \frac{-n^2}{2} \cdot \frac{-1}{v^2} \cdot v - 2^{1-n}e^{-n^2/2v}}{v^2} \\
h'(v) &= \frac{2^{1-n}e^{-n^2/2v} \cdot (\frac{n^2}{2v} - 1)}{v^2} \\
h'(v) &= 0 \iff v = n^2/2 \\
h''(v) &= \frac{2^{1-n}e^{-n^2/2v} \cdot (n^4 - 8n^2v + 8v^2)}{4v^5} \\
h''(n^2/2) &= -\frac{2^{4-n}}{n^6 e} < 0,
\end{aligned}$$

odnosno $v = n^2/2$ kao točku minimuma. Ovo dokazuje sljedeći rezultat Erdősa [11]:

Teorem 3.9. $m(n) < (1+o(1))\dfrac{e\ln 2}{4}n^2 2^n$.

3.3 Kombinatorna teorija brojeva

Definicija 3.10. *Podskup A Abelove grupe G zovemo* skupom bez sume *(„sum-free set") ako je $(A+A)\cap A = \emptyset$, odnosno ako ne postoje $a_1, a_2, a_3 \in A$ takvi da je $a_1 + a_2 = a_3$.*

Teorem 3.11 (Erdős [12]). *Svaki skup $B = \{b_1, \cdots, b_n\}$ od n cijelih brojeva različitih od nule sadrži podskup A koji je skup bez sume u grupi $(\mathbb{Z}, +)$ takav da je $|A| > \frac{1}{3}n$.*

Dokaz. Neka je $p = 3k + 2$ prost broj koji zadovoljava $p > 2\max\{|b_i|, i = 1, \ldots, n\}$. Znamo da takav p postoji jer postoji beskonačno mnogo prostih brojeva oblika $3k + 2$. Pretpostavimo suprotno, odnosno da postoji samo konačno mnogo prostih brojeva $p_1, \ldots p_l$ oblika $3k + 2$, odnosno $3k - 1$. Označimo broj $N = 3p_1p_2 \cdots p_l - 1$, koji je istog oblika. Zbog pretpostavke ne može biti prost broj pa je on produkt prostih faktora, odnosno $N = q_1q_2 \cdots q_m$. Barem jedan od prostih faktora q_i mora biti oblika $3k + 2$ jer je produkt brojeva oblika $3k + 1$ također oblika $3k + 1$. No, s obzirom da N nije djeljiv ni sa jednim p_i, a oni su svi postojeći prosti brojevi oblika $3k+2$, dolazimo do kontradikcije, odnosno postoji beskonačno mnogo prostih brojeva oblika $3k + 2$.

Stavimo $C = \{k + 1, k + 2, \cdots, 2k + 1\}$. Uočimo da je najmanji zbroj dva elementa iz C $k + 1 + k + 1 = 2k + 2$, što je veće od bilo kojeg elementa skupa C. Također, najveći zbroj elemenata iz C je $2k + 1 + 2k + 1 = 4k + 2$, što je zapravo $k(\text{mod } p)$, a k je nedovoljno velik da bude element skupa C. To znači da je C skup bez sume u cikličkoj grupi $\mathbb{Z}_p = \{0, 1, 2, \cdots, 3k + 1\}$ i vrijedi

$$\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}.$$

Odaberimo na slučajan način cijeli broj x, $1 \leq x < p$ i definirajmo $d_1, \cdots, d_n$ kao $d_i \equiv xb_i(\text{mod } p), 0 \leq d_i < p$. Za svaki fiksan $i, 1 \leq i \leq n$, kako x postiže bilo koju vrijednost od $1, 2, \cdots, p - 1$, tako d_i može biti bilo koji element iz $\mathbb{Z}_p$ različit od nule. Zaista, očito d_i poprima vrijednosti iz $\mathbb{Z}_p$, a ne može postići vrijednost nula jer je p prost broj veći od x i svih b_i, tako da xb_i nikad neće davati višekratnik broja p. Zatim, vrijednosti d_i su sve različite jer kada bi postojala ista vrijednost, tada bi postojali b_i i b_j takvi da $(xb_i - xb_j)(\text{mod } p) \equiv 0 \Rightarrow x(b_i - b_j)(\text{mod } p) \equiv 0$, što je opet kontradiktorno tome da je p prost broj veći od svih b_i i x. Zbog toga vrijedi $\mathbb{P}[d_i \in C] = |C|/(p - 1) > 1/3$. Označimo s X_i indikatorsku slučajnu varijablu takvu da $X_i = 1$ ako je b_i takav da je $d_i \in C$. Uvedimo i slučajnu varijablu X koja predstavlja broj elemenata b_i takvih da je $d_i \in C$, odnosno $X = \sum_{i=1}^{n} X_i$. Sada, zbog linearnosti očekivanja, vrijedi

$$\mathbb{E}X = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}X_i > n \cdot \frac{1}{3}.$$

Posljedično tome, postoji $x, 1 \leq x < p$ i skup $A \subset B$ kardinaliteta $|A| > n/3$ takav da $xa(\text{mod } p) \in C, \forall a \in A$. Ovaj skup A je očito skup bez sume, jer,

ako $a_1 + a_2 = a_3$ za $a_1, a_2, a_3 \in A$, tada $xa_1 + xa_2 \equiv xa_3 (\mathrm{mod}\ p)$, što je kontradikcija s činjenicom da je C skup bez sume u grupi $\mathbb{Z}_p$.

$\square$

Dokaz gornjeg teorema je djelotvoran kad god je p prost broj koji ne dijeli ni jednog b_i. Ovo možemo iskoristiti za stvaranje efikasnog determinističkog algoritma za pronalazak skupa bez sume A veličine veće od $|B|/3$ iz danog skupa B. Alon i Kleitman [3] su dokazali kako se konstanta $1/3$ u gornjem teoremu ne može zamijeniti s $12/29$. Pretpostavili su da je $\frac{1}{3}$ najbolja moguća konstanta što su kasnije dokazali Eberhard, Green i Manners [8]. Zatim su pokazali da svaki skup s n elemenata različitih od nule iz proizvoljne Abelove grupe sadrži skup bez sume s više od $2n/7$ elemenata i da je konstanta $2/7$ najbolja moguća. Nije još riješeno pitanje postoji li za svaki skup s n cijelih brojeva različitih od nule, podskup koji je skup bez sume takav da mu je kardinalnost barem $n/3 + w(n)$ gdje $w(n)$ teži prema ∞ kako n teži prema ∞. Bilo bi vrlo iznenađujuće da takav $w(n)$ zapravo ne postoji.

3.4 Disjunktni parovi

Vjerojatnosna metoda je najzanimljivija kada se koristi u dokazivanju teorema čije tvrdnje na prvi pogled ne zahtijevaju korištenje vjerojatnosti. Većina primjera koje smo dokazali su upravo takvi, a u idućem primjeru ćemo opisati malo složeniji rezultat Alona i Franka [2], koji riješava hipotezu Daykina i Erdősa.

Neka je $\mathcal{F}$ familija m različitih podskupova od $X = \{1, 2, \ldots, n\}$. Označimo s $d(\mathcal{F})$ broj disjunktnih parova u $\mathcal{F}$, odnosno:

$$d(\mathcal{F}) = |\{\{F, F'\} : F, F' \in \mathcal{F}, F \cap F' = \emptyset\}|.$$

Daykin i Erdős su pretpostavili da, ako je $m = 2^{(1/2+\delta)n}$, tada za svaki fiksan $\delta > 0$, $d(\mathcal{F}) = o(m^2)$, kako n teži u beskonačnost. Ovaj rezultat slijedi iz idućeg teorema, koji je poseban slučaj općenitijeg rezultata:

Teorem 3.12. *Neka je $\mathcal{F}$ familija $m = 2^{(1/2+\delta)n}$ podskupova od $X = \{1, 2, \ldots, n\}$ gdje $\delta > 0$. Tada je*

$$d(\mathcal{F}) < m^{2-\delta^2/2}. \tag{3}$$

Dokaz. Prepostavimo da ne vrijedi (3); tada odaberemo nezavisno t članova $A_1, A_2, \ldots, A_t$ iz $\mathcal{F}$ s ponavljanjem na slučajan način, gdje je t neki veliki prirodni broj kojeg ćemo kasnije odrediti. Pokazat ćemo da s pozitivnom vjerojatnošću vrijedi $|A_1 \cup A_2 \cup \cdots \cup A_t| > n/2$ i da je ta unija disjunktna s više od $2^{n/2}$ različitih podskupova od X. To pokazuje da ta unija sadrži više

od pola elemenata iz X i u isto vrijeme ne sadrži više od pola elemenata iz X, što je očito kontradikcija. Ta kontradikcija dokazuje valjanost izraza (3).

Vrijedi

$$\begin{aligned}\mathbb{P}[|A_1 \cup A_2 \cup \cdots \cup A_t| \leq n/2] &\leq \sum_{S \subset X, |S|=n/2} \mathbb{P}[A_i \subset S, i = 1, \ldots, t] \\ &\leq 2^n \left(\frac{2^{n/2}}{2^{(1/2+\delta)n}} \right)^t = 2^{n(1-\delta t)}.\end{aligned} \tag{4}$$

Druga nejednakost slijedi iz idućih činjenica. Broj $n/2$-članih podskupova $S \subseteq X$ je $\binom{n}{n/2}$, a to je manje ili jednako od ukupnog broja podskupova od X, to jest 2^n. Vjerojatnost da slučajni element $A_i \in \mathcal{F}$ bude podskup od S je manja ili jednaka od $\frac{2^{n/2}}{2^{(1/2+\delta)n}}$ jer postoji $2^{n/2}$ mogućih podskupova od S.

Želimo dobiti da vrijedi $\mathbb{P}[|A_1 \cup A_2 \cup \cdots \cup A_t| \leq n/2] \leq 1$ jer tada s pozitivnom vjerojatnošću vrijedi $|A_1 \cup A_2 \cup \cdots \cup A_t| > n/2$. Zbog rezultata (4), zapravo želimo $1 - \delta t < 0$, odnosno $t > 1/\delta$. Iz tog razloga uzimamo $t = \lceil 1 + \frac{1}{\delta} \rceil$.

Za fiksni $B \in \mathcal{F}$ definiramo

$$v(B) = |\{A \in \mathcal{F} : B \cap A = \emptyset\}|.$$

Očito,

$$\sum_{B \in \mathcal{F}} v(B) = 2d(\mathcal{F}) \geq 2m^{2-\delta^2/2}.$$

Neka je Y slučajna varijabla čija vrijednost je broj članova $B \in \mathcal{F}$ koji su disjunktni sa svim $A_i, 1 \leq i \leq t$. Zbog konveksnosti od $f(z) = z^t$ (što se može provjeriti diferenciranjem), očekivana vrijednost od Y zadovoljava

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{B \in \mathcal{F}} \left(\frac{v(B)}{m} \right)^t = \frac{1}{m^t} \cdot m \left(\frac{\sum v(B)^t}{m} \right) \\ &\geq \frac{1}{m^t} \cdot m \left(\frac{\sum v(B)}{m} \right)^t = \frac{1}{m^t} \cdot m \left(\frac{2d(\mathcal{F})}{m} \right)^t \\ &\geq \frac{1}{m^t} \cdot m \left(\frac{2m^{2-\delta^2/2}}{m} \right)^t = \frac{1}{m^t} \cdot m \cdot 2^t \cdot m^{t-t\delta^2/2} \geq 2m^{1-t\delta^2/2}.\end{aligned}$$

S obzirom da je $Y \leq m$,

$$\begin{aligned}\mathbb{E}[Y] = \sum_{i=1}^{m} \mathbb{P}(Y = i) \cdot i &= \sum_{i=1}^{\lceil m^{1-t\delta^2/2} - 1 \rceil} \mathbb{P}(Y = i) \cdot i + \sum_{\lceil m^{1-t\delta^2/2} \rceil}^{m} \mathbb{P}(Y = i) \cdot i \\ &\leq m^{1-t\delta^2/2} \cdot \mathbb{P}(Y < m^{1-t\delta^2/2}) + m \cdot \mathbb{P}(Y \geq m^{1-t\delta^2/2}) \\ &\leq m^{1-t\delta^2/2} + m \cdot \mathbb{P}(Y \geq m^{1-t\delta^2/2}).\end{aligned}$$

Stoga zaključujemo:

$$\mathbb{P}[Y \geq m^{1-t\delta^2/2}] \geq m^{-t\delta^2/2}. \tag{5}$$

Za $t = \lceil 1 + \frac{1}{\delta} \rceil$ i $\delta < 0.5$ vrijedi $t < 2 + 1/\delta$ i iz toga slijedi

$$\begin{aligned} m^{1-t\delta^2/2} \geq m^{1-(2+1/\delta)\delta^2/2} &= m^{1-\delta^2-\delta/2} = 2^{(1/2+\delta)n(1-\delta^2-\delta/2)} \\ &= 2^{n/2+n\delta(1/2-\delta)(\delta+3/2)} \geq 2^{n/2}. \end{aligned}$$

Stoga je desna strana u nejednakosti (5) veća od desne strane u nejednakosti (4). Zato vrijedi da s pozitivnom vjerojatnošću $|A_1 \cup A_2 \cup \cdots \cup A_t| > n/2$ i ta unija je disjunktna s više od $2^{n/2}$ članova iz $\mathcal{F}$. Ova kontradikcija implicira nejednakost (3).

□

3.5 Erdős-Ko-Radov teorem

Definicija 3.13. *Familiju skupova $\mathcal{F}$ nazivamo* presijecajućom *ako $A, B \in \mathcal{F}$ povlači $A \cap B \neq \emptyset$.*

Neka je $\mathcal{F}$ familija koja sadrži podskupove skupa $\{1, 2 \ldots, n\}$. Prvo nas zanima koliko najviše elemenata može sadržavati familija $\mathcal{F}$ koja je presijecajuća. S jedne strane, znamo da vrijedi

$$A \in \mathcal{F} \Rightarrow A^c \notin \mathcal{F},$$

odnosno $\mathcal{F}$ ne može sadržavati više od $2^n/2$ elemenata. S druge strane, ako odaberemo neki element $x \in \{1, 2, \ldots, n\}$ i tvorimo familiju sa svim podskupovima koji sadrže element x, tada znamo da je $|\mathcal{F}| = 2^{n-1}$, odnosno postoji presijecajuća familija s barem 2^{n-1} elemenata. Iz toga dvoje slijedi da je najveći broj elemenata koje može sadržavati presijecajuća familija jednak 2^{n-1}.

Pitanje koje možemo iz ovog rezultata postaviti je: ako presijecajuća familija sadrži 2^{n-1} elemenata, je li ona nužno oblika $\mathcal{F} = \{A \subseteq \{1, 2, \ldots, n\} : x \in A\}$? Odgovor je ne, a kao protuprimjer navodimo dva primjera, kada je n neparan i kada je n paran. Ako je n neparan, neka je $\mathcal{F} = \{A \subset \{1, 2, \ldots, n\} : |A| > n/2\}$. Zbog neparnosti, to je točno pola svih podskupova, a zbog veličine podskupova znamo da je familija nepresijecajuća. Ako je n paran, prvo odaberemo u $\mathcal{F}$ sve skupove A takve da $|A| > n/2$ i još od svih skupova A veličine $n/2$ odaberemo po jedan iz svakog para (A, A^c). Tada je opet veličina familije jednaka 2^{n-1}.

Što ako su svi skupovi u $\mathcal{F}$ veličine k? Koliko velika tada familija može biti? Pogledajmo tri slučaja, od kojih će se samo jedan pokazati zanimljivim

i on će biti motivacija za slijedeći teorem. Ako je $k > \frac{n}{2}$, tada svaka dva skupa te veličine imaju neprazan presjek pa ukoliko ih sve uzmemo u familiju $\mathcal{F}$, ona je presijecajuća. Odnosno, $|\mathcal{F}| = \binom{n}{k}$. Ukoliko je $k = \frac{n}{2}$, tada odaberemo iz svakog disjunktnog para (A, A^c) po jedan skup i dobijemo presijecajuću familiju veličine $\frac{1}{2}\binom{n}{k}$. Ostaje nam najzanimljiviji slučaj, kada je $k < \frac{n}{2}$. Ako odaberemo sve skupove koji sadrže neki x, tada je $|\mathcal{F}| = \binom{k-1}{n-1}$. Idući teorem pokazuje da je to i najveći mogući broj elemenata u familiji koja je presijecajuća.

Uvedimo, radi jednostavnosti, oznake $[n] = \{1, 2, \ldots, n\}$ i neka je $\mathcal{P}_k(S)$ familija svih k-članih podskupova od S.

Teorem 3.14 (Erdős-Ko-Rado)**.** *Neka je $n > 2k$ i neka je $\mathcal{F}$ presijecajuća familija koja sadrži k-člane podskupove skupa $\{1, 2, \ldots, n\}$. Tada je $|\mathcal{F}| \leq \binom{n-1}{k-1}$. Jednakost je zadovoljena ako i samo ako je $\mathcal{F}$ oblika $\{A \in \mathcal{P}_k([n]) : x \in A\}$ za neki $x \in [n]$.*

Dokaz. Neka je $(x_1, \ldots, x_n)$ neka nasumična permutacija skupa $[n]$ koju promatramo kao ciklus (odnosno x_1 promatramo kao sljedbenika elementa x_n. Za svaki j, označimo sa I_j skup $\{x_j, x_{j+1}, \ldots, x_{j+k-1}\}$ gdje je zbrajanje u smislu modulo n. Prvo ćemo pokazati kako najviše k ovakvih skupova može biti sadržano u presijecajućoj familiji.

Pretpostavimo, bez smanjenja općenitosti, da skup $I_1 = \{x_1, x_2, \ldots, x_k\}$ pripada familiji $\mathcal{F}$. Tada svaki skup koji pripadaja familiji $\mathcal{F}$ mora imati presjek sa skupom I_1, a to su skupovi I_j za $j \in \{2, 3, \ldots, k\}$ i skupovi $I_j^- = \{x_{j-k}, \ldots, x_{j-1}\}$ za $j \in \{2, 3, \ldots, k\}$. Zbog disjunktnosti skupova I_j i I_j^-, iz svakog takvog para skupova možemo najviše uzeti jedan da bi familija bila presijecajuća. Takvih parova je $k - 1$, što zajedno sa skupom I_1 daje da u presijecajućoj familiji postoji najviše k takvih skupova. Neka je X_j indikatorska slučajna varijabla takva da je $X_j = 1$ ako vrijedi $I_j \in \mathcal{F}$, a $X_j = 0$ inače. Neka je također $X = \sum_{j=1}^{n} X_j$. Tada vrijedi

$$\mathbb{P}[X_j] = \frac{|\mathcal{F}|}{\binom{n}{k}}$$

$$\mathbb{E}[X] = \sum_{j=1}^{n} \mathbb{E}[X_j] = n\frac{|\mathcal{F}|}{\binom{n}{k}}.$$

Vidjeli smo kako najviše k skupova I_j može pripadati familiji $\mathcal{F}$ pa vrijedi $\mathbb{E}[X] \leq k$, to jest $n\frac{|\mathcal{F}|}{\binom{n}{k}} \leq k$, odnosno $|\mathcal{F}| \leq \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}$.

Sada ćemo dokazati drugi dio tvrdnje, vezan za jednakost. Jednu stranu smo već dokazali u diskusiji prije teorema, a sada želimo dokazati da ako vrijedi jednakost $|\mathcal{F}| = \binom{n-1}{k-1}$, tada je $\mathcal{F}$ oblika $\{A \in \mathcal{P}_k([n]) : x \in A\}$. Ako

jednakost vrijedi, tada u svakoj permutaciji, mora postojati k uzastopnih skupova da bi $\mathcal{F}$ bila presijecajuća familija. Za $1 < j < k$ ne možemo uzeti u familiju i I_{j+1}^- i I_{j+1}, ali moramo odabrati jedan od ta sva skupa kako bismo u familiji imali k skupova. Ukoliko odaberemo I_{j+1}^-, tada smo prisiljeni (jer skupovi moraju biti uzastopni) uzeti I_{j+2}^- i tako do I_k^-. Odnosno, tada postoji $1 \leq r \leq k$ takav da su skupovi koje smo odabrali $I_1, I_2, \ldots, I_r, I_{r+1}^-, \ldots, I_k^-$, od kojih svaki skup sadrži element x_r.

Sada odgovaramo na pitanje, je li svaki skup koji sadrži x_k element familije. Neka su $x_1, \ldots, x_{2k-1}$ takvi da svi skupovi $\{x_j, \ldots, x_{j+k-1}\}$, $1 \leq j \leq k$ pripadaju familiji $\mathcal{F}$ te neka je u element koji je različit od svih $x_1, \ldots, x_{2k-1}$. On postoji jer je $n > 2k$, a poslužit će nam za konstruiranje permutacije skupa $[n]$. Neka je A n-člani skup koji sadrži element x_k i označimo njegove elemente sa $\{y_1, \ldots, y_k\}$ tako da prvih r elemenata pripada skupu $\{x_1, \ldots, x_k\}$, $y_r = x_k$, a ostali ne pripadaju. Sada napravimo permutaciju skupa $[n]$ gdje je početni element u, zatim su elementi skupa $\{x_1, \ldots, x_k\}$ poredani tako da su $y_1, \ldots, y_r$ u poretku zadnji, zatim elementi $y_{r+1}, \ldots, y_k$, a zatim svi ostali elementi skupa $[n]$ gdje je poredak proizvoljan. Skup $\{u, x_1, \ldots, x_{k-1}\}$ ne pripada familiji $\mathcal{F}$ jer ima prazan presjek sa skupom $\{x_k, \ldots, x_{2k-1}\}$. S obzirom da mora postojati k uzastopnih intervala, tada znamo da skup A pripada familiji $\mathcal{F}$. Stoga, familija $\mathcal{F}$ sadrži svaki skup veličine k koji sadrži element x_k. S obzirom da je ovo maksimalna presijecajuća familija i ima veličinu $\binom{n-1}{k-1}$, tvrdnja slijedi.

□

3.6 Dominantni skupovi

Definicija 3.15. Dominantni skup *neusmjerenog grafa $G = (V, E)$ je skup $U \subset V$ takav da svaki vrh $v \in V \setminus U$ ima bar jedan susjedni vrh u U.*

Stupanj vrha $v \in V$ u grafu $G = (V, E)$ je broj bridova iz E koji su incidentni s v.

Teorem 3.16. *Neka je $G = (V, E)$ graf na n vrhova s minimalnim stupnjem $\delta > 1$. Tada graf G sadrži dominantni skup koji sadrži najviše $n \cdot \dfrac{1 + \ln(\delta + 1)}{\delta + 1}$ vrhova.*

Dokaz. Neka je $p \in [0, 1]$ proizvoljan. Na slučajan način nezavisno odabiremo svaki vrh od V s vjerojatnošću p. Neka je X skup svih odabranih vrhova i neka je $Y = Y_X$ skup svih vrhova iz $V \setminus X$ takvih da nemaju ni jedan susjedni vrh u X. Očekivana vrijednost od $|X|$ je očito np jer smo birali n vrhova, svaki s vjerojatnošću p. Za svaki vrh $v \in V$,

$\mathbb{P}[v \in Y] = \mathbb{P}[v$ i svi njegovi susjedni vrhovi nisu u $X] \leq (1-p)^{\delta+1}$. Varijablu $|Y|$ možemo zapisati kao sumu n indikatorskih varijabli χ_v $(v \in V)$ gdje je $\chi_v = 1$ ako je $v \in Y$ i $\chi_v = 0$ u suprotnom. Zbog svojstva linearnosti očekivanja sada možemo zaključiti:

$$\begin{aligned}\mathbb{E}\big[|X| + |Y|\big] = \mathbb{E}[|X|] + \mathbb{E}[|Y|] &= \mathbb{E}[|X|] + \mathbb{E}\Big[\sum_{v \in V} \chi_v\Big] \\ &= \mathbb{E}[|X|] + \sum_{v \in V} \mathbb{E}[\chi_v] \leq np + n(1-p)^{\delta+1}.\end{aligned}$$

Stoga, postoji bar jedan odabir skupa X takav da vrijedi nejednakost $|X| + |Y_X| \leq np + n(1-p)^{\delta+1}$. Skup $U = X \cup Y_X$ je očito dominantni skup od G čija kardinalnost je najviše $np + n(1-p)^{\delta+1}$.

Gornje tvrdnje vrijede za sve $p \in [0,1]$. Za optimizaciju rezultata, prvo ograničimo $1 - p \leq e^{-p}$, što dovodi do jednostavnijeg ograničenja $|U| \leq np + ne^{-p(\delta+1)}$. Traženjem lokalnog minimuma desne strane nejednakosti, dobijemo

$$p = \frac{\ln(\delta+1)}{\delta+1}.$$

Ovaj rezultat uvrstimo u ocjenu za $|U|$ i dobijemo $|U| \leq n\frac{1+\ln(\delta+1)}{\delta+1}$ iz čega slijedi tvrdnja teorema.

□

U ovom dokazu nalaze se tri jednostavne, ali važne ideje. Prva je linearnost očekivanja, a na tom svojstvu ćemo se posebno zadržati u idućem poglavlju. Druga ideja je metoda preinaka, koju koristimo u slučajevima kada nasumičan odabir ne osigurava željeni skup. Tako je u dokazu nasumičnim odabirom dobiven skup X koji je preinačen dodavanjem skupa Y_X kako bismo dobili traženi dominantni skup U. Tu, drugu ideju, ćemo detaljnije razložiti u četvrtom poglavlju. Treća ideja je pronaći optimalan p tako da se na početku p koristi kao varijabla, a na kraju se odabire p koji je optimalan za dobivenu ocjenu od $|U|$.

U ovom dokazu se javlja i četvrta ideja o pomicanju i prilagođavanju ocjena. S obzirom da je u većini slučajeva teško naći točan minimum, postavlja se prilagođena ocjena (u našem primjeru $1 - p \leq e^{-p}$) koja bitno pojednostavljuje pronalazak približnog minimuma.

4 Linearnost očekivanja

Kao što smo naveli u prošlom poglavlju, mnogo rezultati koji se dokazuju vjerojatnosnom metodom koriste svojstvo linearnosti očekivanja. U ovom poglavlju ćemo prvo opisati od kakve koristi linearnost očekivanja može biti, a zatim ćemo kroz primjere vidjeti koliko je primijenjiva.

Ovo svojstvo je vrlo korisno jer ne zahtijeva od slučajnih varijabli nezavisnost. U mnogim slučajevima, $\mathbb{E}[X]$ može se jednostavno izračunati preko razumne dekompozicije u jednostavne (često indikatorske) slučajne varijable.

Neka je σ nasumična permutacija skupa $\{1, \ldots, n\}$ uniformno izabrana. Neka je $X(\sigma)$ broj fiksnih točaka od σ. Da bismo pronašli očekivanje $\mathbb{E}[X]$ napravimo dekompoziciju $X = X_1 + \cdots + X_n$, gdje je X_i indikatorska slučajna varijabla za događaj $\sigma(i) = i$. Tada je

$$\mathbb{E}[X_i] = \mathbb{P}[\sigma(i) = i] = \frac{1}{n}$$

tako da vrijedi

$$\mathbb{E}[X] = \frac{1}{n} + \cdots + \frac{1}{n} = 1.$$

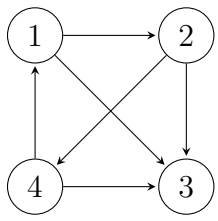
U praksi, često koristimo činjenicu da postoji element ω u vjerojatnosnom prostoru za koji vrijedi $X(\omega) \geq \mathbb{E}[X]$ i također element ω' za koji vrijedi $X(\omega') \leq \mathbb{E}[X]$. Ovu činjenicu smo već koristili u dokazu teorema 3.11.

Sljedeći teorem se često navodi kao prvi primjer u korištenju vjerojatnosne metode, a rezultat je rada Szelea [17].

Definicija 4.1. Hamiltonov put *u grafu $G = (V, E)$ je put koji prolazi svakim vrhom grafa točno jednom.*

Teorem 4.2. *Postoji turnir T s n igrača koji ima barem $n!2^{-(n-1)}$ Hamiltonovih puteva.*

Prije dokaza teorema, nađimo za $n = 4$ turnir koji ima barem $4! \cdot 2^{-3} = 3$ Hamiltonova puteva.



Slika 6: Turnir s 3 Hamiltonovih puteva.

Možemo primjetiti kako su tri Hamiltonova puta $(1, 2, 4, 3)$, $(4, 1, 2, 3)$ i $(2, 4, 1, 3)$.

Dokaz. U nasumično odabranom turniru, neka je X broj Hamiltonovih puteva. Za svaku permutaciju σ, neka je X_σ indikatorska slučajna varijabla da σ daje Hamiltonov put, odnosno vrijedi $X_\sigma(T) = 1$ ako je $(\sigma(i), \sigma(i+1)) \in T$ za $1 \leq i < n$, $X_\sigma(T) = 0$ inače. Tada je $X = \sum X_\sigma$ i vrijedi

$$\mathbb{E}[X] = \sum \mathbb{E}[X_\sigma] = n! \frac{2^{n(n-1)/2-(n-1)}}{2^{n(n-1)/2}} = n! 2^{-(n-1)}.$$

Stoga mora postojati turnir koji ima barem $\mathbb{E}[X]$ Hamiltonovih puteva.

□

Definicija 4.3. *Kažemo da je graf $G = (V, E)$* bipartitan *ako se skup V može razdvojiti u dva skupa B i C tako da svaki brid iz E spaja vrh iz B s vrhom iz C.*

Teorem 4.4. *Neka je $G = (V, E)$ graf s n vrhova i e bridova. Tada postoji bipartitan podgraf od G koji ima minimalno $e/2$ bridova.*

Dokaz. Neka je $T \subseteq V$ slučajno odabran podskup takav da je $\mathbb{P}[v \in T] = \frac{1}{2}, \forall v \in V$ nezavisno od ostalih vrhova. Neka je $B = V \setminus T$. Brid $\{v_1, v_2\}$ ćemo nazivati *prijelaznim* ako je točno jedan v_1 ili v_2 u T. Neka je X sada broj svih prijelaznih bridova. Napravimo dekompoziciju skupa X u indikatorske slučajne varijable X_{xy} gdje je $X_{xy} = 1$ ukoliko je brid $\{x, y\}$ prijelazan, a u suprotnom je $X_{xy} = 0$. Pomoću svojstva linearnosti očekivanja tada možemo zaključiti:

$$\mathbb{E}[X] = \mathbb{E}\Big[\sum_{\{x,y\} \in E} X_{xy}\Big] = \sum_{\{x,y\} \in E} \mathbb{E}[X_{xy}] = e \cdot \frac{1}{2} = \frac{e}{2}.$$

Stoga, $X \geq e/2$ za neke izbore podskupa T i skup prijelaznih bridova tvori bipartitan podgraf.

□

Ovaj rezultat možemo i poboljšati ako malo izmijenimo pretpostavke teorema:

Teorem 4.5. *Ako graf $G = (V, E)$ sadrži $2n$ vrhova i e bridova, tada postoji bipartitni podgraf grafa G koji ima minimalno $en/(2n-1)$ bridova. U slučaju kada graf sadrži neparni broj vrhova $2n+1$ i e bridova, postoji bipartitni graf s minimalno $e(n+1)/(2n+1)$ bridova.*

Dokaz. Neka je skup T uniformno odabran između svih podskupova skupa V s n članova. Kada imamo slučaj da je broj vrhova paran, svaki brid $\{x, y\}$ je prijelazan s vjerojatnošću

$$\frac{n}{2n} \cdot \frac{n}{2n-1} + \frac{n}{2n} \cdot \frac{n}{2n-1} = \frac{n}{2n-1}.$$

Dokaz tvrdnje slijedi kao u gornjem teoremu, koristeći linearnost očekivanja. U slučaju kada je broj vrhova neparan, svaki brid $\{x, y\}$ je prijelazan s vjerojatnošću

$$\frac{n}{2n+1} \cdot \frac{n+1}{2n} + \frac{n+1}{2n+1} \cdot \frac{n}{2n} = \frac{n+1}{2n+1},$$

pa tvrdnja slijedi kao i za slučaj parnog broja vrhova.

□

Teorem 4.6. *Postoji bojenje bridova potpunog grafa K_n u dvije boje s najviše $\binom{n}{a}2^{1-\binom{a}{2}}$ jednobojnih podgrafova K_a.*

Dokaz. Uzmimo neko slučajno bojenje bridova grafa K_n u dvije boje i neka je slučajna varijabla X broj jednobojnih K_a. Definirajmo i indikatorsku slučajnu varijablu X_{K_a} koja poprima vrijednost 1 kada je K_a jednobojan, a inače 0. Tada vrijedi

$$\begin{aligned}\mathbb{E}[X] = \mathbb{E}\left[\sum_{K_a \text{ podgraf od } K_n} X_{K_a}\right] &= \sum_{K_a \text{ podgraf od } K_n} \mathbb{E}\left[X_{K_a}\right] \\ &= \binom{n}{a}\frac{2}{2^{\binom{a}{2}}} = \binom{n}{a}2^{1-\binom{a}{2}}.\end{aligned}$$

Iz ovoga slijedi da postoje vrijednosti slučajne varijable X koje su manje ili jednake od vrijednosti očekivanja slučajne varijable X. Odnosno postoji bojenje grafa K_n tako da postoji najviše $\binom{n}{a}2^{1-\binom{a}{2}}$ jednobojnih podgrafova K_a.

□

Teorem 4.7. *Neka su $v_1, \ldots, v_n \in \mathbb{R}^n$ takvi da $\forall i \in \{1, \ldots, n\}, ||v_i|| = 1$. Tada postoje $\epsilon_1, \ldots, \epsilon_n = \pm 1$ takvi da vrijedi*

$$||\epsilon_1 v_1 + \cdots + \epsilon_n v_n|| \leq \sqrt{n}.$$

Također, postoje $\epsilon_1, \ldots, \epsilon_n = \pm 1$ takvi da vrijedi

$$||\epsilon_1 v_1 + \cdots + \epsilon_n v_n|| \geq \sqrt{n}.$$

Dokaz. Neka su $\epsilon_1, \ldots, \epsilon_n$ slučajno odabrani na uniforman način nezavisno od drugih iz skupa $\{-1, 1\}$. Neka je $X = ||\epsilon_1 v_1 + \cdots + \epsilon_n v_n||^2$ slučajna varijabla. Za nju tada vrijedi

$$X = \sum_{i=1}^{n} \epsilon_i v_i \cdot \sum_{j=1}^{n} \epsilon_j v_j = \sum_{i=1}^{n}\sum_{j=1}^{n} \epsilon_i v_i \cdot \epsilon_j v_j.$$

Stoga vrijedi

$$\mathbb{E}X = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i \cdot v_j \cdot \mathbb{E}[\epsilon_i \epsilon_j].$$

Za slučajnu varijablu ϵ_i, s obzirom da je izabrana na uniforman način, vrijedi $\mathbb{E}[\epsilon_i] = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (+1) = 0$. Ako je $i \neq j$, tada, s obzirom da su one nezavisne, vrijedi $\mathbb{E}[\epsilon_i \epsilon_j] = \mathbb{E}[\epsilon_i]\mathbb{E}[\epsilon_j] = 0$. Ako pak vrijedi $i = j$, tada je $\mathbb{E}[\epsilon_i^2] = \frac{1}{2} \cdot (-1)^2 + \frac{1}{2} \cdot (+1)^2 = 1$. Stoga je

$$\mathbb{E}X = \sum_{i=1}^{n} v_i \cdot v_i = \sum_{i=1}^{n} 1 = n,$$

odnosno, postoje $\epsilon_1, \ldots, \epsilon_n = \pm 1$ kojima je vrijednost slučajne varijable $X \geq \mathbb{E}[X] = n$ i također oni kojima je vrijednost slučajne varijable $X \leq \mathbb{E}[X] = n$. Korijenujući te nejednakosti dobijemo tvrdnju teorema.

□

U prošlom teoremu slučajne varijable smo odabirali na uniforman način i vektori v_i su zadovoljavali $||v_i|| = 1$. Idući teorem će biti svojevrsna linearna translacija prošlog teorema, gdje ćemo gledati i neuniformne načine odabira slučajnih varijabli ϵ i vektore v_i koji zadovoljavaju $||v_i|| \leq 1$.

Teorem 4.8. *Neka su $v_1, \ldots, v_n \in \mathbb{R}^n$ takvi da $\forall i \in \{1, \ldots, n\}, ||v_i|| \leq 1$ i neka su $p_1, \ldots, p_n \in [0, 1]$ i označimo $w = p_1 v_1 + \cdots + p_n v_n$. Tada postoje $\epsilon_1, \ldots, \epsilon_n \in \{0, 1\}$ i $v = \epsilon_1 v_1 + \cdots + \epsilon_n v_n$ takvi da vrijedi*

$$||w - v|| \leq \frac{\sqrt{n}}{2}.$$

Dokaz. Neka su $\epsilon_1, \ldots, \epsilon_n$ slučajno odabrani, nezavisno jedan od drugih, s vjerojatnostima $\mathbb{P}[\epsilon_i = 1] = p_i$ i $\mathbb{P}[\epsilon_i = 0] = 1 - p_i$. Slučajan odabir varijabli ϵ_i daje slučajne varijable v i $X = ||w - v||^2$. Raspišemo li taj izraz, dobijemo:

$$X = \left|\left|\sum_{i=1}^{n}(p_i - \epsilon_i)v_i\right|\right|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i \cdot v_j \cdot (p_i - \epsilon_i)(p_j - \epsilon_j),$$

što daje:

$$\mathbb{E}[X] = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i \cdot v_j \cdot \mathbb{E}[(p_i - \epsilon_i)(p_j - \epsilon_j)].$$

Znamo da vrijedi $\mathbb{E}[p_i - \epsilon_i] = (1 - p_i) \cdot p_i + p_i \cdot (p_i - 1) = 0$, stoga za $i \neq j$ vrijedi

$$\mathbb{E}[(p_i - \epsilon_i)(p_j - \epsilon_j)] = \mathbb{E}[p_i - \epsilon_i] \cdot \mathbb{E}[p_j - \epsilon_j] = 0.$$

Kada vrijedi $i = j$, tada je

$$\mathbb{E}[(p_i - \epsilon_i)^2] = (1 - p_i)^2 \cdot p_i + p_i^2 \cdot (p_i - 1) = p_i(1 - p_i).$$

Ako pogledamo funkciju $f(x) = x(1-x) = -x^2+x$, znamo da je to kvadratna funkcija koja postiže globalni maksimum $\frac{1}{4}$ za $x = \frac{1}{2}$. Stoga vrijedi

$$\mathbb{E}[X] = \sum_{i=1}^{n} p_i(1 - p_i)||v_i||^2 \leq \frac{1}{4} \sum_{i=1}^{n} ||v_i||^2 \leq \frac{n}{4}.$$

Tvrdnja teorema sada slijedi na analogan način kao u dokazu prošlog teorema.

□

Teorem 4.9. *Neka su $a_{ij} = \pm 1, 1 \leq i, j \leq n$. Tada postoje $x_i, y_i = \pm 1, 1 \leq i, j \leq n$ takvi da vrijedi*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j \geq \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{\frac{3}{2}}.$$

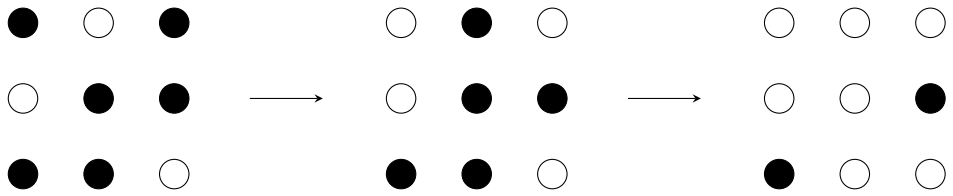
Ovaj rezultat se može interpretirati na zanimljiv način. Uzmimo neku matricu A dimenzije $n \times n$ u kojoj se na svakom mjestu nalazi lampica, odnosno a_{ij} predstavlja jednu lampicu, koja može biti upaljena ($a_{ij} = 1$) ili ugašena ($a_{ij} = -1$). Ako postoji prekidač za svaki red lampica (x_i) i za svaki stupac lampica (y_j), možemo reći da će se aktivacijom prekidača ($x_i = -1$ za red, $y_j = -1$ za stupac) svaka lampica koja je (u tom redu ili stupcu) bila upaljena ugasiti, a svaka koja je bila ugašena upaliti. Teorem sada govori da za bilo koje početno stanje upaljenih i ugašenih lampica, moguće je izvesti određene aktivacije prekidača tako da broj upaljenih lampica minus broj ugašenih lampica bude barem $(\sqrt{2/\pi} + o(1))n^{3/2}$.

Primjer 4.10. *Neka je* $a_{ij} = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$, *on tada ima 3 upaljene lampice. Ako postavimo vrijednosti $x = (-1, 1, 1)$ i $y = (1, -1, 1)$, prekidači koji se aktiviraju su x_1, odnosno prvi red i y_2 odnosno drugi stupac. Nakon aktivacija, matrica lampica izgleda ovako*

Odnosno, upaljenih lampica je 7, a ugašene su dvije.

Dokaz. Neka su $y_1 \dots y_n \in \{-1, 1\}$ nezavisno odabrani na slučajan, uniforman način i označimo

$$R_i = \sum_{j=1}^{n} a_{ij} y_j, \quad R = \sum_{i=1}^{n} |R_i|.$$



Slika 7: Matrica lampica

Neovisno o a_{ij}, vjerojatnost da je $a_{ij}y_j = 1$ je $\frac{1}{2}$ (analogno i za -1) i vrijednosti su nezavisne, odnosno, bez obzira je li u redu i došlo do aktivacije prekidača, on je i dalje uniformno distribuiran redak, gdje je 2^n mogućnosti jednako vjerojatno. Stoga R_i prati distribuciju slučajne varijable S_n koja predstavlja distribuciju sume od n nezavisnih uniformnih $\{-1, 1\}$ varijabli, $S_n = \sum_{j=1}^{n} X_j, X_j = \pm 1$. Po centralnom graničnom teoremu, za $n \to \infty$ vrijedi

$$\frac{S_n - \mathbb{E}[X_j]}{\sqrt{\mathrm{Var}X_j}\sqrt{n}} \longrightarrow N(0,1).$$

S obzirom da je $\mathbb{E}[X_j] = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$ i $\mathrm{Var}X_j = \mathbb{E}[(X_j - \mathbb{E}[X_j])^2] = \mathbb{E}[X_j^2] = 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1$, zapravo vrijedi:

$$\frac{S_n}{\sqrt{n}} \longrightarrow N(0,1),$$

odnosno $S_n = \sqrt{n}(Z + o(1))$, pri čemu Z ima standardnu normalnu distribuciju. Stoga $|S_n|$ možemo aproksimirati pomoću presavijene standardne normalne distribucije čije je očekivanje $\sqrt{\frac{2}{\pi}}$. Stoga vrijedi

$$\mathbb{E}[|R_i|] = \mathbb{E}[|S_n|] = \sqrt{n} \cdot \left(\sqrt{\frac{2}{\pi}} + o(1)\right),$$

a zbog linearnosti očekivanja vrijedi

$$\mathbb{E}[R] = \sum_{i=1}^{n} \mathbb{E}[|R_i|] = n^{3/2} \cdot \left(\sqrt{\frac{2}{\pi}} + o(1)\right).$$

Znamo da tada postoje $y_1, \dots y_n \in \{-1, 1\}$ takvi da je R najmanje te vrijednosti. Odaberimo takve x_i da je predznak jednak predznaku od R_i. Tada vrijedi

$$\sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} y_j = \sum_{i=1}^{n} x_i R_i = \sum_{i=1} |R_i| = R \geq n^{3/2} \cdot \left(\sqrt{\frac{2}{\pi}} + o(1)\right).$$

$\square$

4.1 Metoda preinaka

Osnovnu vjerojatnosnu metodu smo opisali u prošlim poglavljima tako da smo za dokazivanje postojanja strukture sa željenim svojstvima prvo definirali vjerojatnosni prostor struktura i pokazali da željena svojstva postoje unutar tog prostora s pozitivnom vjerojatnošću. Sada ćemo se fokusirati na primjere u kojima „nasumična" struktura nema sva željena svojstva, nego ima neke „nedostatke". Uz preinake, eliminaramo nedostatke i tako dobivamo željenu strukturu.

Prisjetimo se iz prvog primjera sa Ramseyevim brojevima kako $R(k,l) > n$ znači da postoji bojenje bridova grafa K_n u dvije boje tako da ne postoji potpuni crveni podskup K_k ni potpuni plavi podskup K_l.

Teorem 4.11. *Za bilo koji cijeli broj n vrijedi $R(k,k) > n - \binom{n}{k}2^{1-\binom{k}{2}}$.*

Dokaz. Uzmimo neko nasumično bojenje bridova potpunog grafa K_n u dvije boje tako da svaki brid bojimo crveno ili plavo s jednakom vjerojatnošću i neovisno o bojenju drugih bridova. Za bilo koji skup R koji sadrži k vrhova, neka je X_R indikatorska slučajna varijabla za događaj da je inducirani potpuni podgraf grafa K_n definiran na vrhovima iz skupa R jednobojan. Definirajmo također slučajnu varijablu X koja označava broj jednobojnih podgrafova u nasumičnom bojenju od K_n. Tada je $X = \sum X_R$, gdje suma prolazi kroz sve moguće skupove R. Tada, zbog linearnosti očekivanja, znamo da je

$$\mathbb{E}[X] = \sum_R \mathbb{E}[X_R] = \binom{n}{k} \cdot \frac{2}{2^{\binom{k}{2}}} = \binom{n}{k} \cdot 2^{1-\binom{k}{2}}.$$

Zbog ovog rezultata možemo zaključiti da postoji bojenje takvo da vrijedi $X \leq m$ gdje je $m = \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$, to jest da je broj jednobojnih K_k manji ili jednak od m. Odaberimo sada upravo takvo bojenje grafa K_n te maknimo iz K_n jedan vrh iz svakog jednobojnog k-skupa. Izbacili smo najviše m vrhova. Neki vrhovi smiju se nalaziti u više jednobojnih podgrafova K_k, ali to za naš argument ne smeta. Stoga nam ostaje s vrhova, gdje je $s \geq n - \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$. Bojenje na ovih s vrhova ne sadrži ni jedan jednobojni k-skup, te stoga vrijedi $R(k,k) > s$.

□

Dosada smo se kod Ramseyevih brojeva fokusirali na dijagonalne brojeve, odnosno one kojima je $k = l$. Iduća dva rezultata, vezana uz sve Ramseyeve brojeve, koristit ćemo kako bismo prikazali korištenje preinaka u vjerojatnosnoj metodi.

Teorem 4.12. *Ako postoji $p \in [0,1]$ takav da vrijedi*

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}} < 1,$$

tada je $R(k,l) > n$.

Dokaz. Neka je ϕ neko slučajno bojenje bridova skupa K_n nezavisno u dvije boje, crvenu i plavu, tako da je p vjerojatnost da brid bude obojen crvenom bojom. Neka je R skup koji sadrži k vrhova iz K_n, te P skup koji sadrži l vrhova iz K_n. Označimo sa X_R indikatorsku slučajnu varijablu koja postiže vrijednost 1 kada je inducirani podgraf od K_n definiran s vrhovima iz R crvene boje, te analogno definiramo X_P kao indikatorsku slučajnu varijablu koja postiže vrijednost 1 kada je inducirani podgraf od K_n definiran vrhovima iz P plave boje. Definirajmo i slučajnu varijablu X kao ukupni broj svih crvenih k-skupova plus broj svih plavih l-skupova. Tada, zbog linearnosti očekivanja vrijedi:

$$\begin{aligned} \mathbb{E}[X] = \mathbb{E}\left[\sum_R X_R + \sum_P X_P\right] &= \sum_R \mathbb{E}\left[X_R\right] + \sum_P \mathbb{E}\left[X_P\right] \\ &= \binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}}. \end{aligned} \tag{6}$$

Iz pretpostavke teorema slijedi da je postoji p takav da $\mathbb{E}[X] < 1$, odnosno postoji bojenje od K_n takvo da je $X = 0$. Stoga je $R(k,l) > n$.

□

Teorem 4.13. *Za sve cijele brojeve n i sve $p \in [0,1]$ vrijedi*

$$R(k,l) > n - \binom{n}{k} p^{\binom{k}{2}} - \binom{n}{l} (1-p)^{\binom{l}{2}}.$$

Dokaz. Kao i u prošlom teoremu, zaključujemo da vrijedi jednakost (6) te označimo desnu stranu jednakosti sa s iz čega slijedi da postoji takvo bojenje da je $\mathbb{E}[X] \leq s$, odnosno postoji najviše s problematičnih skupova koji su ili crveni k-skupovi ili plavi l-skupovi. Izbacimo li iz svakog od tih problematičnih skupova po jedan vrh, dobijemo bojenje potpunog grafa K_{n-s} bez problematičnih skupova, odnosno $n - s < R(k,l)$.

□

Definicija 4.14. *Neka je $G = (V,E)$ graf.* Nezavisan podskup vrhova *od G je podskup od V u kojem ni jedna dva vrha ne čine brid iz E.* Broj nezavisnosti grafa G *je kardinalni broj najvećeg nezavisnog podskupa vrhova od G.*

Broj nezavisnosti označavamo s $\alpha(G)$, a $\alpha(G) \geq t$ znači da postoji t vrhova grafa G koji međusobno nisu povezani bridom.

Teorem 4.15. *Neka je $G = (V, E)$ graf s n vrhova i $nd/2$ bridova, pri čemu je $d \geq 1$. Tada je $\alpha(G) \geq n/2d$.*

Dokaz. Neka je $S \subseteq V$ nasumičan podskup gdje je $\mathbb{P}[v \in S] = p$. Vjerojatnost p ćemo kasnije odrediti, a događaji $v \in S$ su međusobno nezavisni. Tada S inducira podgraf $G_S = (S, E_S)$ gdje su E_S svi bridovi iz E koji povezuju vrhove iz S. Označimo s $X = |S|$ i $Y = |E_S|$. Za svaki brid $e = \{i, j\} \in E$, neka je Y_e indikatorska slučajna varijabla za događaj $i, j \in S$ (odnosno, $e \in E_S$), što znači da je $Y = \sum_{e \in E} Y_e$. Tada, po linearnosti očekivanja vrijedi

$$\mathbb{E}[Y] = \sum_{e \in E} \mathbb{E}[Y_e] = \frac{nd}{2} \cdot \mathbb{P}[i, j \in S] = \frac{nd}{2} p^2.$$

Očekivanje slučajne varijable X analogno, uz pomoć linearnosti očekivanja, izračunamo i dobijemo $\mathbb{E}[X] = np$, te stoga vrijedi

$$\mathbb{E}[X - Y] = np - \frac{nd}{2} p^2.$$

Uvrstimo sada vrijednost $p = 1/d$ i dobijemo da je

$$\mathbb{E}[X - Y] = \frac{n}{2d}.$$

Stoga, postoji takav podskup S za koji je razlika između broja vrhova i broja bridova u S barem $n/2d$. Odaberimo za svaki brid iz S po jedan vrh koji mu je incidentan i sve takve vrhove izbacimo. Tako smo izbacili jednak broj vrhova i bridova, pri čemu smo izbacili sve bridove i dobili novi podgraf S^* koji ima barem $n/2d$ vrhova. S obzirom da S^* nema ni jedan brid, taj podskup je sigurno nezavisan podskup vrhova, stoga je $\alpha(G) \geq n/2d$.

□

Za skup S od n točaka u jediničnom kvadratu $U = [0, 1] \times [0, 1]$, neka je $T(S)$ najmanja površina trokuta čiji su vrhovi elementi iz S. Označimo $T(n) = \max T(S)$ gdje maksimum uzimamo po svim mogućim skupovima $S \subseteq U$, $|S| = n$. Dokazat ćemo u idućem teoremu da vrijedi $T(n) = \Omega(1/n^2)$, odnosno $\frac{1}{n^2} = O(T(n))$. Dovoljno je naći pozitivan broj c takav da vrijedi $\frac{1}{n^2} \leq c \cdot T(n)$.

Teorem 4.16. *Postoji skup S koji se sastoji od n točaka iz jediničnog kvadrata U takav da vrijedi $T(S) \geq 1/(100n^2)$.*

Dokaz. Neka su P, Q, R nezavisno i uniformno odabrane točke iz U te neka je $\mu = P(PQR)$ površina trokuta PQR te želimo doći do vjerojatnosti $\mathbb{P}[\mu \leq \epsilon]$. Neka je x udaljenost između točaka P i Q, odnosno Q se nalazi na kružici čije je središte točka P, a radijus joj je x. S obzirom da se nalazimo u jediničnom kvadratu, vjerojatnost da se Q nalazi unutar određenog skupa je jednaka površini tog skupa, stoga možemo zaključiti da vrijedi

$$\mathbb{P}[b \leq x \leq b + \Delta b] = \mathbb{P}[x \leq b + \Delta b] - \mathbb{P}[x \leq b] \leq \pi(b + \Delta b)^2 - \pi b^2,$$

što daje $\mathbb{P}[b \leq x \leq b + \Delta b] \leq 2\pi b\, \Delta b$. Ako imamo P, Q na udaljenosti b, tada je udaljenost točke R od dužine PQ zapravo visina trokuta PQR, te pomoću formule za površinu možemo zaključiti da, u slučaju $\mu \leq \epsilon$ mora vrijediti $h \leq 2\epsilon/b$. Jer se R može nalaziti udaljena samo za h, to može biti unutar pravokutnika kojemu je širina $2 \cdot h$ (pri čemu je dužina PQ sredina tog pravokutnika) i duljine $\sqrt{2}$, s obzirom da je to najveća duljina dužine unutar jediničnog kvadrata. Stoga je vjerojatnost događaja da je R unutar tog pravokutnika opet površina, koja je jednaka $4\sqrt{2}\epsilon/b$. Kako mora vrijediti $0 \leq b \leq \sqrt{2}$, vjerojatnost da je površina trokuta manja od ϵ je manja ili jednaka od

$$\int_0^{\sqrt{2}} (2\pi b)(4\sqrt{2}\epsilon/b)db = \int_0^{\sqrt{2}} (8\pi\sqrt{2}\epsilon)db = 16\pi\epsilon.$$

Neka su sada $P_1, \ldots, P_{2n}$ nezavisno i uniformno odabrane točke u U i neka X označava broj trokuta $P_iP_jP_k$ takvih da je njihova površina manja od $1/(100n^2)$. Za svaki i, j, k vjerojatnost takvog događaja je:

$$\mathbb{P}\left[\mu \leq \frac{1}{100n^2}\right] \leq 16\pi\frac{1}{100n^2} = \frac{0.16\pi}{n^2} < 0.6n^{-2}.$$

Koristeći linearnost matematičkog očekivanja, možemo zaključiti da vrijedi:

$$\mathbb{E}[X] \leq \binom{2n}{3}(0.6n^{-2}) = \frac{8n}{10} - \frac{12}{10} + \frac{4}{10n} < n - 1 + \frac{4}{10} < n.$$

Iz toga slijedi da postoji skup od $2n$ točaka gdje manje od n trokuta ima površinu manju od $1/(100n^2)$. Kada iz svakog takvog trokuta obrišemo po jedan vrh, ostane nam barem n vrhova, odnosno n točaka gdje nijedne tri točke ne čine trokut površine manje od $1/(100n^2)$.

□

Kada slučajno bojenje nije samo po sebi dovoljno za rješavanje problema, postoji nekoliko načina da se s rješavanjem problema nastavi. Jedan od tih

načina je slučajno prebojenje koje može popraviti „nesavršenosti” koje su se pojavile nakon prvotnog bojenja. Takve metode su korisne u proučavanju dvoobojivosti hipergrafa iz 3. poglavlja. Prisjetimo se, $m(n)$ označava minimalni mogući broj bridova n-uniformnog hipergrafa koji nije dvoobojiv, odnosno minimalni mogući broj bridova takav da za n-uniforman hipergraf svako bojenje vrhova sadrži barem jedan jednobojan brid. Proučit ćemo dokaz Cherkashina i Kozika [7], zasnovan na algoritmu pohlepnog bojanja.

Teorem 4.17. *Ako postoji $p \in [0,1]$ takav da vrijedi $k(1-p)^n + k^2 p < 1$, tada vrijedi $m(n) > 2^{n-1}k$.*

Korolar 4.18. *Vrijedi $m(n) = \Omega\left(2^n (n/\ln n)^{(1/2)}\right)$.*

Dokaz. Koristimo ocjenu $1 - p \leq e^{-p}$ te neka je $f(p) = ke^{-pn} + k^2 p$ za neki pozitivan broj k koji ćemo kasnije odrediti. Deriviranjem možemo zaključiti kako funkcija postiže minimum kada vrijedi

$$-nke^{-p_0 n} + k^2 = 0 \implies p_0 = \frac{\ln \frac{n}{k}}{n}.$$

Ako uvrstimo vrijednost p_0 nazad u funkciju dobijemo da vrijedi

$$k(1-p_0)^n + k^2 p_0 \leq ke^{-p_0 n} + k^2 p_0 = \frac{k^2}{n}\left(1 + \ln \frac{n}{k}\right).$$

Ako odaberemo k takav da je gornji izraz manji od 1, tada vrijede pretpostavke teorema te možemo zaključiti $m(n) > 2^{n-1}k$. Nejednakost je zadovoljena kada vrijedi $k = c\sqrt{\frac{n}{\ln n}}$, pri čemu je $c < \sqrt{2}$ i n je dovoljno velik. Zaista, vrijedi

$$\begin{aligned}\lim_{n\to\infty} \frac{c^2}{\ln n}(1 + \ln n - \ln k) =& c^2 - \lim_{n\to\infty} \frac{c^2}{\ln n}(\ln c + \frac{1}{2}\ln n - \frac{1}{2}\ln\ln n) \\ =& c^2 - \frac{1}{2}c^2 = \frac{1}{2}c^2 < 1.\end{aligned}$$

□

Dokaz teorema 4.17. Uzmimo $n-$uniforman hipergraf $H = (V, E)$ s $m = 2^{n-1}k$ bridova i neka je p takav da zadovoljava pretpostavku teorema. Opišimo nasumičan proces bojanja vrhova iz V.

Za svaki vrh $v \in V$ neka je x_v slučajno odabrana uniformna varijabla iz $[0,1]$ koju ćemo nazivati oznakom vrha v, pri čemu su odabiri oznaka neovisni među različitim vrhovima. Algoritam prolazi kroz sve vrhove hipergrafa H redom po njihovim oznakama, od manjih prema većih, bojeći na način da će

bojati svaki vrh v_i plavom bojom, osim ako postoji brid e_j iz E takav da je $v_i \in e_j$ te su svi drugi vrhovi brida e_j obojeni plavom bojom. U tom slučaju će se vrh v_i bojati crvenom bojom. Takvim algoritmom dolazimo do bojenja u kojem ne postoji brid e kome su svi vrhovi obojeni plavom bojom. Jedini jednobojni bridovi tada mogu biti crveni, a to se može dogoditi samo ako su svi njegovi vrhovi posljednji obojeni vrhovi nekih drugih bridova.

Nazovimo par bridova (e, f) konfliktnim parom bridova ako je posljednji obojeni vrh brida e ujedno i prvi obojeni vrh brida f. Ukoliko u hipergrafu ne postoji konfliktnih parova bridova, tada bojenje generirano algoritmom ne proizvodi jednobojne bridove. Pokažimo da konfliktni parovi bridova ipak postoje s pozitivnom vjerojatnošću.

U svrhu dokazivanja ove tvrdnje, podijelimo interval $[0, 1]$ u tri podintervala L, S i D (lijevi, srednji i desni podinterval) tako da vrijedi $L = [0, \frac{1-p}{2}\rangle$, $S = [\frac{1-p}{2}, \frac{1+p}{2}\rangle$ i $D = [\frac{1+p}{2}, 1]$. Vjerojatnost da postoji konfliktan par bridova (e, f) gdje je $e \subset L$ ili $f \subset D$ je ograničena vjerojatnošću da postoji brid hipergrafa H sadržan u L ili u D, što je najviše

$$\begin{aligned}\sum_{e\in E} \mathbb{P}[e \subset L] + \sum_{e\in E} \mathbb{P}[e \subset D] =& \sum_{e\in E} \left(\frac{1-p}{2}\right)^n + \sum_{e\in E} \left(1 - \frac{1+p}{2}\right)^n \\ =& 2m \left(\frac{1-p}{2}\right)^n = k(1-p)^n.\end{aligned}$$

Za bilo koji drugi konfliktni par (e, f), postoji jedinstveni vrh $v = e \cap f$ takav da $x_v \in S$. (Vrh v je jedinstven jer je on posljednji obojeni vrh brida e, a prvi obojeni vrh brida f.) Dodatno, za svaki vrh $u \in e \setminus \{v\}$, s obzirom da je obojan prije vrha v, vrijedi $x_u < x_v$. Također, analogno, za svaki vrh $w \in f \setminus \{v\}$ vrijedi $x_w > x_v$. Vjerojatnost da (e, f) tvore ovako navedeni konfliktni par je ograničena vjerojatnošću događaja $x_v \in S$. Ta vjerojatnost je jednaka $\frac{1+p}{2} - \frac{1-p}{2} = p$. U tom slučaju, vjerojatnost da vrijedi $x_u < x_v$ za svaki vrh $u \in e - v$ i $x_w > x_v$ za svaki vrh $w \in f - v$ je

$$\begin{aligned}\sum_{u\in e\setminus v} \mathbb{P}[x_u < x_v] \cdot \sum_{w\in f\setminus v} \mathbb{P}[x_w > x_v] =& \sum_{u\in e\setminus v} (x_v - 0) \cdot \sum_{w\in f\setminus v} (1 - x_v) \\ =& x_v^{n-1}(1 - x_v)^{n-1} \leq \left(\frac{1}{4}\right)^{n-1},\end{aligned}$$

gdje posljednja nejednakost slijedi iz činjenice da je $g(x) = x - x^2$ parabola negativnog vodećeg koeficijenta koja svoj globalni maksimum postiže kada je $x = \frac{1}{2}$. Stoga je vjerojatnost da je (e, f) takav konfliktni par bridova najviše $p \cdot \left(\frac{1}{4}\right)^{n-1}$. S obzirom da postoji manje od m^2 uređenih parova bridova od H,

ukupna vjerojatnost da postoji konfliktni par bridova je manja od

$$k(1-p)^n + (k2^{n-1})^2 p \left(\frac{1}{4}\right)^{n-1} = k(1-p)^n + k^2 p,$$

što je po pretpostavci teorema manje od 1. To znači da je vjerojatnost postojanja bojenja bez konfliktnih bridova, odnosno s jednobojnim bridom, veća od 0, dakle $m(n) > m$.

$\square$

Algoritam koji smo koristili u dokazu teorema možemo formulirati i kao proceduru prebojenja. Nakon što smo na slučajan način odabrali oznake x_v, sve vrhove v gdje je $x_v \in L \cup S$ bojimo plavom bojom, a preostale, gdje je $x_v \in D$ bojimo crvenom bojom. Nakon toga prolazimo kroz vrhove iz S od najmanje oznake prema najvećoj tako da svaki vrh koji je posljednji u plavo obojenom bridu prebojimo u crveno.

5 Drugi moment

Kao što smo u prošlom poglavlju nalazili primjere korištenja svojstva linearnosti očekivanja, tako ćemo se sada osvrnuti na korištenje drugog mometa, odnosno varijance. Započet ćemo s primjerom iz teorije brojeva. Prvo ćemo dokazati jedan rezultat koji će biti korišten u dokazu idućeg teorema.

Teorem 5.1. *Neka je $n \in \mathbb{N}$, tada za sumu svih recipročnih vrijednosti prostih brojeva manjih od n vrijedi*

$$\sum_{p \leq n} \left(\frac{1}{p}\right) = \ln \ln n + O(1).$$

Dokaz. Neka je $p \in \{1, \ldots, n\}$ prost broj i neka $\gamma(p)$ predstavlja stupanj tog faktora u raspisu broja $n!$ na proste faktore. Faktor p ima stupanj najmanje $\lfloor n/p \rfloor$ jer se pojavljuje u svakom svojem višekratniku, a njih je točno toliko. Također, p se pojavljuje „duplo" u svakom višekratniku broja p^2 koji je manji od n, te, s obzirom da smo dodali jedan stupanj faktora, za svaki takav višekratnik faktor dobiva stupanj više. Analognim razmišljanjem i za ostale potencije dobijemo da vrijedi

$$\gamma(p) = \sum_{p \leq n} \left\lfloor \frac{n}{p^i} \right\rfloor.$$

Sada možemo zaključiti da vrijedi

$$\ln n! = \ln p_1^{a_1} \cdot p_2^{a_2} \cdots p_k^{a_k} = \sum_{j=1}^{k} a_j \cdot \ln p_j = \sum_{j=1}^{k} \sum_{p_j^i \leq n} \left\lfloor \frac{n}{p_j^i} \right\rfloor \cdot \ln p_j.$$

Zbog svojstva $y - 1 \leq \lfloor y \rfloor \leq y$, zapravo vrijedi

$$\ln n! = n \cdot \sum_{j=1}^{k} \frac{1}{p_j} \cdot \ln p_j + O\left(\sum_{j=1}^{k} \ln p_j\right),$$

što zbog $c \cdot \sum \ln p_j = c \cdot \ln p_1 \cdots p_k \leq c \cdot \ln n^k \leq c \cdot k \cdot n$ daje rezultat

$$\ln n! = n \cdot \sum_{j=1}^{k} \frac{\ln p_j}{p_j} + O(n).$$

Koriteći Stirlingovu formulu $\ln n! = n \ln n + O(n)$ dobijemo

$$\sum_{j=1}^{k} \frac{\ln p_j}{p_j} = \ln n + O(1).$$

Radi jednostavnosti zapišimo $\sum_{j=1}^{k} \frac{\ln p_j}{p_j}$ kao $\sum_{p \le n} \frac{\ln p}{p}$. Sada možemo iskoristiti Abelovu sumaciju i označimo u tu svrhu $A(n) = \sum_{p \le n} \frac{\ln p}{p}$. Tada vrijedi

$$\begin{aligned}\sum_{p \le n} \frac{1}{p} = \sum_{p \le n} \frac{\ln p}{p} \cdot \frac{1}{\ln p} &= A(n)\frac{1}{\ln n} + \int_1^n A(t) \cdot \frac{1}{t \cdot \ln t^2} dt \\ &= (\ln n + O(1))\frac{1}{\ln n} + \int_1^n (\ln(t) + O(1)) \cdot \frac{1}{t \cdot \ln t^2} dt \\ &= O(1) + \int_1^n \frac{1}{t \cdot \ln t} dt = \ln \ln n + O(1),\end{aligned}$$

gdje se zadnja jednakost može dobiti tako da u integral uvrstimo supstituciju $u = \ln t$.

□

Neka $\nu(n)$ označava broj prostih brojeva p koji dijele n. Rezultat koji slijedi, zasnovan na dokazu Turána [18], govori kako je skoro svim n broj prostih faktora blizu broja $\ln \ln n$.

Teorem 5.2. *Neka $\omega(n) \to \infty$ proizvoljno sporo. Tada je broj x-eva iz $\{1, \ldots, n\}$ za koje vrijedi*

$$|\nu(x) - \ln \ln n| > \omega(n)\sqrt{\ln \ln n}$$

jednak $o(n)$.

Dokaz. Neka je x slučajno odabran iz $\{1, \ldots, n\}$. Za prost broj p, označimo

$$X_p = \begin{cases} 1 & \text{ako } p \text{ dijeli } x, \\ 0 & \text{inače.} \end{cases}$$

Neka je $M = n^{1/10}$ i $X = \sum X_p$ gdje suma prolazi kroz sve proste brojeve $p \le M$. Tada za svaki $x \le n$ vrijedi da ima najviše 10 prostih faktora većih od M jer bi se inače vrijedilo

$$x = c \cdot p_1^{a_1} \cdot p_2^{a_2} \cdots p_{10}^{a_{10}} \ge c \cdot p_1 \cdot p_2 \cdots p_{10} > M^{10} = n,$$

pri čemu je $c \in \mathbb{N}$, a p_i predstavljaju 10 prostih faktora većih od 10. Tada očito vrijedi $\nu(x) - 10 \le X(x) \le \nu(x)$. Sada vrijedi

$$\mathbb{E}[X_p] = \mathbb{P}[X_p = 1] = \frac{\{\text{Koliko ima brojeva djeljivih s p}\}}{\{\text{Koliko ima ukupno brojeva}\}} = \frac{\lfloor \frac{n}{p} \rfloor}{n}.$$

Zbog svojstva $y - 1 \leq \lfloor y \rfloor \leq y$, zapravo vrijedi

$$\mathbb{E}[X_p] = \frac{\frac{n}{p}}{n} + O\left(\frac{1}{p}\right) = \frac{1}{p} + O\left(\frac{1}{p}\right).$$

Koristimo li svojstvo linearnosti očekivanja i prethodni teorem dobijemo

$$\mathbb{E}[X] = \sum_{p \leq M} \left(\frac{1}{p} + O\left(\frac{1}{p}\right)\right) = \ln \ln n + O(1).$$

Znamo da varijanca može biti izračunata preko formule

$$\mathrm{Var}[X] = \sum_{p \leq M} \mathrm{Var}[X_p] + \sum_{p \neq q} \mathrm{Cov}[X_p, X_q]. \tag{7}$$

Izračunajmo prvo lijevu sumu. Znamo da je

$$\begin{aligned}\mathrm{Var}[X_p] = \mathbb{E}X^2 - (\mathbb{E}X)^2 &= \frac{1}{p} + O\left(\frac{1}{p}\right) - \left(\frac{1}{p} + O\left(\frac{1}{p}\right)\right)^2 \\ &= \frac{1}{p}\left(1 - \frac{1}{p}\right) + O\left(\frac{1}{p}\right).\end{aligned}$$

Iz toga slijedi

$$\sum_{p \leq M} \mathrm{Var}[X_p] = \left(\sum_{p \leq M} \frac{1}{p}\left(1 - \frac{1}{p}\right)\right) + O(1) = \left(\sum_{p \leq M} \frac{1}{p}\right) + O(1) = \ln \ln n + O(1),$$

gdje smo koristili činjenicu kako red $\sum \frac{1}{p^2}$ konvergira prema vrijednosti manjoj od $\pi^2/6$ kao što je dokazano u [6]. Promotrimo sada drugu sumu u (7). Kada su p, q dva različita prosta broja, $X_p X_q = 1$ ako i samo ako i p i q dijele x, što je ekvivalentno tome da pq dijeli x (jer su u pitanju prosti brojevi). Stoga vrijedi

$$\begin{aligned}\mathrm{Cov}[X_p, X_q] &= \mathbb{E}[X_p X_q] - \mathbb{E}[X_p]\mathbb{E}[X_q] \\ &= \frac{\left\lfloor \frac{n}{pq} \right\rfloor}{n} - \frac{\left\lfloor \frac{n}{p} \right\rfloor}{n} \cdot \frac{\left\lfloor \frac{n}{q} \right\rfloor}{n} \\ &\leq \frac{1}{pq} - \left(\frac{\frac{n}{p} - 1}{n}\right)\left(\frac{\frac{n}{q} - 1}{n}\right) \\ &= \frac{1}{pq} - \left(\frac{1}{p} - \frac{1}{n}\right)\left(\frac{1}{q} - \frac{1}{n}\right) \\ &= \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right) - \frac{1}{n^2} \leq \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right).\end{aligned}$$

Jer sumiramo po prostim brojevima p manjim od M zaključujemo da vrijedi

$$\sum_{p\neq q} \mathrm{Cov}[X_p, X_q] \leq \frac{1}{n} \sum_{p\neq q} \left(\frac{1}{p} + \frac{1}{q}\right) \leq \frac{2M}{n} \sum \frac{1}{p}.$$

Uvrštavanjem $M = n^{1/10}$ to daje

$$\sum_{p\neq q} \mathrm{Cov}[X_p, X_q] \leq O(n^{-9/10} \ln\ln n) = o(1),$$

pri čemu smo koristili da izraz funkcija $n^{-9/10} \ln\ln n$ konvergira prema 0 kako n raste prema ∞. S druge strane, vrijedi

$$\begin{aligned}
\mathrm{Cov}[X_p, X_q] &= \mathbb{E}[X_p X_q] - \mathbb{E}[X_p]\mathbb{E}[X_q] \\
&= \frac{\left\lfloor \frac{n}{pq} \right\rfloor}{n} - \frac{\left\lfloor \frac{n}{p} \right\rfloor}{n} \cdot \frac{\left\lfloor \frac{n}{q} \right\rfloor}{n} \\
&\geq \frac{\frac{n}{pq} - 1}{n} - \frac{1}{p} \cdot \frac{1}{q} = -\frac{1}{n}
\end{aligned}$$

Iz toga slijedi da vrijedi

$$\sum_{p\neq q} \mathrm{Cov}[X_p, X_q] \geq -\frac{1}{n} \sum_{p\neq q} 1 \geq -\frac{M}{n} = -o(1).$$

Zaključujemo da kovarijanca ne utječe na varijancu $\mathrm{Var}[X] = \ln\ln n + O(1)$, a korištenjem Čebiševljeve nejednakosti dobijemo da vrijedi

$$\mathbb{P}\left[|X - \ln\ln n| > \lambda\sqrt{\ln\ln n}\right] < \frac{1}{\lambda^2} + o(1),$$

za svaki $\lambda > 0$. Kako je $|X - \nu| \leq 10$, vrijedi $|\nu - \ln\ln n| \leq 10 + |X - \ln\ln n|$ iz čega slijedi tvrdnja teorema.

$\square$

Kao još jedan primjer korištenja drugog momenta u vjerojatnosnoj metodi, navodimo primjer o različitim sumama. Kažemo da skup $\{x_1, \ldots, x_k\}$ koji se sastoji od prirodnih brojeva ima različite sume, ako su sve sume oblika

$$\sum_{i\in S} x_i, \quad S \subseteq \{1, \ldots, k\}$$

međusobno različite. Označimo sa $f(n)$ maksimalan k za koji postoji skup

$$\{x_1, \ldots, x_k\} \subset \{1, \ldots, n\}$$

koji ima različite sume. Odnosno, $f(n)$ predstavlja najveći mogući broj elemenata skupa prirodnih brojeva manjih od n koji ima različite sume.

Primjer 5.3. *Neka je* $A = \{2^i : i \leq \log_2 n\}$. *Svi članovi skupa* A *su manji ili jednaki od* $2^{\log_2 n} = n$. *Osim toga, članova ovoga skupa ima* $1 + \lfloor \log_2 n \rfloor$ *jer* A *uvijek sadrži bar element* 1. *Štoviše, cijeli skup* A *je skup s različitim sumama. Zaista, ako pretpostavimo suprotno, da različiti elementi skupa* A *u sumi daju jednak zbroj te ih bez smanjenja općenitosti poredamo od najmanjeg prema najvećem, možemo dobiti zapis oblika*

$$2^{i_1} + \cdots + 2^{i_m} = 2^{j_1} + \cdots + 2^{j_n}.$$

Tvrdnja slijedi iz jedinstvenosti zapisa prirodnih brojeva u binarnom zapisu. Iz prethodnog primjera slijedi da za svaki n *vrijedi*

$$f(n) \geq 1 + \lfloor \log_2 n \rfloor.$$

Ostaje otvoreno pitanje gornjeg ograničenja za $f(n)$. Erdős ponudio novčanu nagradu za dokaz kako vrijedi

$$f(n) \leq \log_2 n + C,$$

za neku konstantu C, međutim ta pretpostavka još nije dokazana. Primjer 5.3 daje ocjenu $f(n) < \log_2 n + \log_2 \log_2 n + O(1)$, a idući teorem će dati malo poboljšanje ove ocjene.

Teorem 5.4. *Za svaki* $n \in \mathbb{N}$ *vrijedi* $f(n) \leq \log_2 n + \frac{1}{2} \cdot \log_2 \log_2 n + O(1)$.

Dokaz. Neka je $\{x_1, \ldots, x_k\} \subset \{1, \ldots, n\}$ skup s različitim sumama. Neka su $\epsilon_1, \ldots, \epsilon_k$ nezavisno odabrane indikatorske varijable takve da vrijedi

$$\mathbb{P}[\epsilon_i = 1] = \mathbb{P}[\epsilon_i = 0] = \frac{1}{2}.$$

Definirajmo i slučajnu varijablu

$$X = \epsilon_1 x_1 + \cdots + \epsilon_k x_k,$$

odnosno X je suma elemenata skupa, a indikatorska varijabla ϵ_i određuje je li element i u sumi ili nije. Odredimo očekivanje i varijancu slučajne varijable X:

$$\mu = \mathbb{E}[X] = \frac{x_1 + \cdots + x_k}{2}$$

$$\begin{aligned}\sigma^2 = \mathrm{Var}[X] &= \sum_{i=1}^{k} \mathrm{Var}[\epsilon_i x_i] + \sum_{i \neq j} \mathrm{Cov}[\epsilon_i x_i, \epsilon_j x_j] \\ &= \frac{x_1^2 + \cdots + x_k^2}{4} \leq \frac{n^2 k}{4}.\end{aligned}$$

Iz toga i zbog nenegativnosti standardne devijacije slijedi da vrijedi $\sigma \leq \frac{n\sqrt{k}}{2}$. Sada možemo iskoristiti Čebiševljevu nejednakost, za bilo koji $\lambda > 1$ vrijedi

$$\mathbb{P}\left[|X-\mu| \geq \lambda\frac{n\sqrt{k}}{2}\right] \leq \mathbb{P}\big[|X-\mu| \geq \lambda\sigma\big] \leq \frac{1}{\lambda^2}.$$

Tu istu nejednakost možemo zapisati i preko vjerojatnosti suprotnog događaja:

$$1-\frac{1}{\lambda^2} \leq \mathbb{P}\left[|X-\mu| < \lambda\frac{n\sqrt{k}}{2}\right]. \tag{8}$$

Kako je X suma elemenata skupa koji ima različite sume, znamo da X poprima svaku vrijednost s vjerojatnošću 0 ili $\frac{1}{2^k}$. Iz toga slijedi:

$$\begin{aligned}
\mathbb{P}\left[|X-\mu| < \lambda\frac{n\sqrt{k}}{2}\right] &\leq \mathbb{P}\left[|X-\mu| \leq \lambda\frac{n\sqrt{k}}{2}\right] \\
&= \mathbb{P}\left[\mu-\lambda\frac{n\sqrt{k}}{2} \leq X \leq \mu+\lambda\frac{n\sqrt{k}}{2}\right] \\
&\leq \frac{1}{2^k}\cdot\left(\mu+\lambda\frac{n\sqrt{k}}{2}-\mu+\lambda\frac{n\sqrt{k}}{2}+1\right) \\
&= \frac{1}{2^k}\cdot(\lambda n\sqrt{k}+1).
\end{aligned} \tag{9}$$

Kombiniranjem nejednadžbi (8) i (9) dobijemo da vrijedi:

$$n \geq \frac{2^k(1-\lambda^{-2})-1}{\sqrt{k}\lambda}.$$

□

Literatura

[1] Abel's summation formula, dostupno na *https://en.wikipedia.org/wiki/Abel%27s_summation_formula* (prosinac 2020.)

[2] N. Alon, P. Frankl, *The maximum number of disjoint pairs in a family of subsets*, Graphs and Combinatorics 1 (1985), 13-21.

[3] N. Alon, D.J. Kleitman, *Sum-free subsets*, A Tribute to Paul Erdős (A. Baker, B. Bollobás i A. Hajnál, ur.), Cambridge University Press, Cambridge, 1990, 13-26.

[4] N. Alon, J. H. Spencer, *The probabilistic method*, Wiley Publishing, New Jersey, 2015.

[5] A. P. Burger, E. J. Cockayne, C. M. Mynhardt, *Domination and irredundance in the queens' graph*, Discrete Mathematic 163 (1997), 47-66.

[6] R. Champan, Evaluating $\zeta(2)$, dostupno na *http://empslocal.ex.ac.uk/people/staff/rjchapma/etc/zeta2.pdf* (prosinac 2020.)

[7] D. D. Cherkashin, J. Kozik, A note on random greedy coloring of uniform hypergraphs,
dostupno na *https://arxiv.org/abs/1310.1368* (siječanj 2021.)

[8] S. Eberhard, B.Green, F. Manners, *Sets of integers with no large sum-free subset*, Annals of Mathematics 180 (2014), 621–652.

[9] P. Erdős, A. Hajnal, *On a property of families of sets*, Acta Mathematica Academiae Scientiarum Hungarica 12 (1961), 87-123.

[10] P. Erdős, *On a combinatorial problem*, Nordisk Matematisk Tidskrift 11 (1963) 5-10.

[11] P. Erdős, (1964) *On a combinatorial problem II*, Acta Mathematica Academiae Scientiarum Hungarica 15 (1964), 445-447.

[12] P. Erdős, *Extremal problems in number theory*, Proceedings of Symposia in Pure Mathematics 8 (1965), 181-189.

[13] W.T.Gowers, Topics in combinatorics, dostupno na *https://drive.google.com/file/d/1V778zHQTx4XE8FxDgznt2jTshZzxAFot/view* (lipanj 2021.)

[14] F. C. Leone, L. S. Nelson, R. B. Nottingham, *The Folded Normal Distribution*, Technometrics 3 (1961), 543-550.

[15] N. Sarapa, *Teorija vjerojatnosti*. Školska knjiga, Zagreb, 2002.

[16] E. Szekeres, G. Szekeres, *On a problem of Schütte and Erdős*, The Mathematical Gazette 49 (1965), 290–293.

[17] T. Szele, *Kombinatorikai vizsgálatok az iránytott teljes gráffal kapcsolatban*, Matematikai és Fizikai Lapok 50 (1943), 223-256.

[18] P. Turán, *On a theorem of Hardy and Ramanujan*, Journal of the London Mathematical Society 9 (1934), 274–276.

Sažetak

Ovaj rad obrađuje osnovne ideje i primjere vjerojatnosne metode u kombinatorici. Na početku rada se podsjećamo definicija i rezultata iz teorije vjerojatnosti potrebnih za razumijevanje vjerojatnosne metode. Zatim obrađujemo primjere iz različitih područja kombinatorike u kojima se vidi koliko je, iako ponekad nije nužna, vjerojatnosna metoda praktična. U četvrtom poglavlju obrađujemo dodatne primjere uz pomoć indikatorskih slučajnih varijabli i svojstva linearnosti očekivanja. Također spominjemo i metodu preinaka kojom eliminiramo nedostatke nasumičnih struktura. U zadnjem dijelu rada spominjemo nekoliko primjera korištenja vjerojatnosne metode uz pomoć drugog momenta.

Summary

This thesis deals with basic ideas and examples of the probabilistic method in combinatorics. First, we recall the definitions and results from probability theory that are necessary for understanding the probabilistic method. We, then, deal with examples from different areas of combinatorics showing how, even if it's not always necessary, probabilistic method can be practical. In the fourth chapter, we deal with additional examples using indicator random variables and linearity property of expectation. We also mention the alteration method by which we remove blemishes of random structures. In the last part of the thesis, we mention a few examples of using probabilistic method with the help of second moment.

Životopis

Rođena sam 7. srpnja 1995. godine u Zagrebu gdje završavam Osnovnu školu Tituša Brezovačkog te 2010. godine upisujem prirodoslovno-matematički smjer u gimnaziji Lucijana Vranjanina. Preddiplomski studij Matematike upisujem 2014. godine na Prirodoslovno-Matematičkom fakultetu u Zagrebu. Nakon stjecanja prvostupničke diplome, 2018. godine upisujem Diplomski sveučilišni studij Matematička statistika na istom fakultetu.