Eötvös Loránd Tudományegyetem
Természettudományi Kar

Wolosz János András
Matematika BSc
Matematikus szakirány

Valószínűségi módszerek

Szakdolgozat

Témavezető: Móri Tamás egyetemi docens

Budapest, 2012

Köszönetnyilvánítás

Szeretném megköszönni témavezetőmnek, Móri Tamásnak a dolgozat igen alapos ellenőrzését, a konzultációkat és a kiváló szakirodalom ajánlását.

Tartalomjegyzék

Bevezető

Valószínűségi módszerek a matematika nagyon sok ágában használhatók és hasznosak. Különböző egzisztencia bizonyítások, konstrukciók, becslések születtek a valószínűségszámítás eszközeit felhasználva például számelméletben, analízisben, véges matematikában. A kapott eredmények ereje annak köszönhető, hogy a matematika két különböző ágának eredményeit ötvözi. Látni fogjuk a dolgozatban is, hogy ha a valószínűségszámítási ismeretekből többet használunk, az egyszerűbbé teheti a tételek bizonyítását, illetve erősebbé azok állítását. Meg kell említenünk a véletlen algoritmusok gyakorlati jelentőségét is, olyan fontos feladatok esetén, mint az utazó ügynök probléma vagy a gépi tanulás problémája – eddig ezek teljesítenek a legjobban.

Látszik, hogy a valószínűségi módszerek témaköre hatalmas. Ebben a szakdolgozatban főképp véges matematikai alkalmazásokat mutatunk be. A dolgozat felépítése a következő:

– Az első fejezetben a valószínűségi módszer klasszikus alkalmazásait mutatjuk be.

– A második fejezetben az Azuma-egyenlőtlenség alkalmazásait mutatjuk be, ennek segítségével megvizsgáljuk egy gráf véletlen részgráfjának különböző paramétereit, illetve egy érdekes tételt bizonyítunk a $G(n, \frac{1}{2})$ véletlen gráf klikkszámáról.

– A harmadik fejezetben a Talagrand-elmélet segítségével, amely komolyabb mértékelméleti aparátust igényel, diszkrét optimalizálásban felmerülő valószínűségi változókat vizsgálunk.

– A negyedik fejezetben bebizonyítjuk a lokális lemmát, egy fontos véges matematikai eszközt egzisztencia tételek bizonyítására. Ezután megnézzük a lokális lemma egyszerűbb alkalmazásai lehetőségeit, illetve egy felmerülő algoritmizálhatósági problémát tárgyalunk.

1. Klasszikus alkalmazások

Ebben a fejezetben a valószínűségi módszer alapgondolatát és klasszikus alkalmazásait mutatjuk be négy tételen keresztül. Az első alkalmazás Erdős Pál 1947-es eredménye lesz a Ramsey-számok alsó becslésére [1]. Második alkalmazásként a véges halmazrendszerekkel kapcsolatos egyik legnevezetesebb tételt, a Sperner tételt bizonyítjuk be. Itt a valószínűségi módszerrel éles becslést fogunk tudni bizonyítani. Harmadik alkalmazásként Erdős és Guy sejtésére adunk bizonyítást a metszési számról. A bizonyítás [2]-ből származik. Végül egy kombinatorikus geometriai eredményt bizonyítunk be véletlen konstrukció segítségével.

1.1. Alsó becslés az átlós Ramsey-számokra

1.1.1. Tétel. *Ha n és k olyan egészek, hogy $\binom{n}{k}2^{1-\binom{k}{2}} < 1$, akkor $R(k,k) > n$.*

Bizonyítás. Legyen az n csúcsú teljes gráf G. Színezzük G éleit egymástól függetlenül egyenlő valószínűséggel pirosra illetve kékre. Azt fogjuk megmutatni, hogy annak a valószínűsége, hogy van k csúcsú monokromatikus részgráf, kisebb, mint 1. Az összes k csúcsú K részgráfra a teljes gráfban legyen A_K az az esemény, hogy K monokromatikus. Ekkor

$$P\Bigg(\bigcup_{I\subset G,|I|=k} A_I\Bigg) \leq \sum_{I\subset G,|I|=k} P(A_I)$$

miatt, és mert $P(A_I) = 2^{1-\binom{k}{2}}$ azt kaptuk, hogy

$$P\Bigg(\bigcup_{I\subset G,|I|=k} A_I\Bigg) \leq \binom{n}{k}2^{1-\binom{k}{2}} < 1,$$

és éppen ezt akartuk. □

Az előző eredményt és az $\binom{n}{k} < (\frac{ne}{k})^k$ összefüggést kombinálva az $R(k,k) > > \frac{k}{e}\cdot 2^{\frac{k-1}{2}-\frac{1}{k}}$ összefüggést kapjuk. Később látni fogjuk, hogy a lokális lemma segítségével ez az egyenlőtlenség erősíthető.

A fenti bizonyítás egzisztencia bizonyítás, azonban kaptunk egy gyakorlatban jól használható eljárást is egy konkrét, monokromatikus k részgráf nélküli színezés megtalálásához. Nevezetesen, ha az n csúcsú teljes gráf éleit véletlenszerűen színezzük egyenlő valószínűséggel kékre vagy pirosra, akkor annak a valószínűsége, hogy nincs monokromatikus k csúcsú részgráf, legalább $1-\binom{n}{k}2^{1-\binom{k}{2}}$. Így például $n = 1024$, $k = 20$ esetén annak a valószínűsége, hogy a véletlen színezés után van 20 csúcsú monokromatikus részgráf, kevesebb mint

$$\binom{1024}{20}2^{1-\binom{20}{2}} < \frac{2^{200}}{20!}2^{-189} = \frac{2^{11}}{20!} < 10^{-15} \ll 1.$$

Nem mindig kapunk ilyen könnyen implementálható eljárást a valószínűségi okoskodásunkból, egy bonyolultabb esettel a 4.2 szakaszban foglalkozunk.

1.2. Sperner-tétel

1.2.1. Tétel. *Legyen adott egy n elemű alaphalmaz és azon egy F halmazrendszer. Ha igaz az, hogy tetszőleges $A, B \in F$ esetén $A \not\subset B$ és $B \not\subset A$, akkor $|F| \le \binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Bizonyítás. Vegyük az n elemű halmaz egy véletlen permutációját, minden permutációt egyenlő, $\frac{1}{n!}$ valószínűséggel választjuk. Egy $(s_1, s_2, \ldots, s_n)$ permutáció i-edik kezdőszeletének nevezzük bármely $i \in \{1, 2, \ldots, n\}$ esetén az $\{s_1, s_2, \ldots, s_i\}$ halmazt. X valószínűségi változó legyen a véletlen permutációban kezdőszeletként előálló F-beli halmazok száma. X nem lehet 1-nél nagyobb, hiszen bármely két kezdőszelet tartalmazás tekintetében relációban áll, míg ez semelyik két elemre sem igaz F-ben a feltételünk szerint. Tehát $E(X) \le 1$. Másrészt F egy t számosságú eleme $t!\,(n-t)!$ darab permutációban kezdőszelet, így

$$E(X) = \sum_{A \in F} |A|!\,(n-|A|)! \cdot \frac{1}{n!} = \sum_{A \in F} \frac{1}{\binom{n}{|A|}} \le 1.$$

Tudjuk, hogy az $\binom{n}{k}$ binomiális együttható $k = \lfloor \frac{n}{2} \rfloor$ esetén maximális, így $\frac{1}{\binom{n}{|A|}}$ legalább $\frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$, tehát a fenti összeg legfeljebb $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ tagú, azaz $|F| \le \binom{n}{\lfloor \frac{n}{2} \rfloor}$. Ezzel a tételt igazoltuk. □

1.3. Metszéspont-lemma

Ismert, hogy a $K_{3,3}$ és K_5 gráfok nem síkbarajzolhatók, azaz ezen gráfok tetszőleges síkbaágyazása esetén valamely élek metszeni fogják egymást. Egy gráf metszési számát úgy definiáljuk, hogy a lehető legkevesebb élmetszéspontot tartalmazó síkbaágyazásában a metszéspontok száma. A síkbaágyazás során 3 vagy több él egy élmetszéspontban nem találkozhat.

Jelöljük G metszési számát $cr(G)$-vel. A következő tétel a gráf v csúcsszáma és e élszáma segítségével ad alsó becsést $cr(G)$-re.

1.3.1. Tétel. *Metszéspont-lemma: Ha $1 \ge \frac{4v}{e}$, akkor $cr(G) \ge \frac{1}{64}\frac{e^3}{v^2}$.*

Bizonyítás. Vegyük G minimális metszéspontú síkbaágyazását. Tegyük fel, hogy a $cr(G)$ darab metszéspont is a gráf csúcsa, két metszéspont közötti síkbaágyazott élrész pedig a gráf új éle. Így megkaptuk a G' síkgráfot. G'-nek $v + cr(G)$ csúcsa van. Mivel egy metszésponton két él megy át, G' éleinek száma $e+2cr(G)$, ez látszik például a fokszámokból is, hiszen minden új csúcs 4 fokú.

A G' síkgráfra alkalmazva az $e \le 3v-6$ ismert egyenlőtlenséget azt kapjuk, hogy $cr(G) \ge e - 3v + 6$.

Rögzítsünk egy p valószínűséget. Válasszuk ki G minden csúcsát a többi csúcs választástól függetlenül p valószínűséggel. A kapott véletlen gráfot jelöljük G_p-vel. Ha G_p csúcsainak száma v_p, éleinek száma e_p, akkor az előbbiek miatt $cr(G_p) \ge$ $\ge e_p - 3v_p + 6$. Innen azonban az $E(cr(G_p)) \ge E(e_p) - 3E(v_p) + 6$ egyenlőtlenséget kaptuk a várható értékekre.

A várható értékeket rendre kiszámolhatjuk. $E(v_p) = vp$, $E(e_p) = ep^2$, hiszen egy adott él p^2 valószínűséggel kerül be G_p-be, pontosan akkor, ha a két végpontját választjuk. Végül $E(cr(G_p)) = cr(G)p^4$, hiszen egy adott metszéspont p^4 valószínűséggel lesz G_p-ben, pontosan akkor, amikor a két metsző élhez tartozó 4 végpontot beválasztjuk. Innen

$$cr(G)p^4 \ge ep^2 - 3vp + 6 > ep^2 - 3vp.$$

Ha p-t $\frac{4v}{e}$-nek választjuk, akkor a fenti egyenlőtlenségből éppen a bizonyítani kívánt $cr(G) \ge \frac{1}{64}\frac{e^3}{v^2}$ egyenlőtlenséget kapjuk. □

1.4. Egy kombinatorikus geometriai konstrukció

Tekintsük a következő problémát. Legyen S egy n elemű ponthalmaz az egységnyi oldalhosszúságú négyzetben. Jelöljük $T(S)$-sel a legkisebb területű háromszöget, melynek csúcsai S-beliek. Legyen $M(n) = \max_{|S|=n} T(S)$.

1.4.1. Tétel. $M(n) \ge \frac{1}{100n^2}$.

Bizonyítás. Egy hasznos becslést bizonyítunk először. Válasszuk a P, Q, R pontokat az egységnégyzetben egymástól függetlenül, egyenletes eloszlás szerint. Legyen X a PQR háromszög területét felvevő valószínűségi változó. Becsüljük felülről a $P(X <$ $< t)$ valószínűséget. Legyen az Y valószínűségi változó a PQ oldal hossza. Ekkor $P(\ell < Y < \ell+d\ell) \le ((\ell+d\ell)^2-\ell^2)\pi$ az egyenletes eloszlás miatt. Y sűrűségfüggvénye így az ℓ pontban legfeljebb $2\ell\pi$. Ha a PQ oldalhossz ℓ, akkor a PQR háromszög területe akkor kisebb mint t, ha R a PQ egyenestől $\dfrac{2t}{\ell}$-nél kisebb távolságra van. Ennek a sávnak a területe az egységnégyzetben felülről becsülhető $\sqrt{2}\cdot\dfrac{4t}{\ell}$-lel, így annak az eseménynek a valószínűsége, hogy $PQ = \ell$ esetén PQR területe kisebb mint t, felülről becsülhető $\sqrt{2} \cdot \dfrac{4t}{\ell}$-lel. Mivel Y ismeretében $P(X < t)$-re már van felső becslésünk, és Y sűrűségfüggvényét is tudjuk pontonként felülről becsülni, a

$$P(X < t) \leq \int_0^{\sqrt{2}} (2\ell\pi)(\sqrt{2} \cdot \frac{4t}{\ell})d\ell = 16t\pi$$

felső becslés igaz.

$M(n)$ alsó becsléséhez válasszunk ki az egységnégyzeten egyenletes eloszlás szerint egymástól függetlenül $2n$ pontot. Legyen a Z valószínűségi változó ekkor a $2n$ pont által meghatározott háromszögek közül az $\frac{1}{100n^2}$-nél kisebb területűek száma.

Felülről becsüljük $E(Z)$-t. Annak a valószínűsége, hogy egy háromszög területe kisebb, mint $\frac{1}{100n^2}$, felülről becsülhető az előbb bizonyított becslés miatt $16\frac{1}{100n^2}\pi$-vel. Így

$$E(Z) \leq \binom{2n}{3} \cdot 16\frac{1}{100n^2}\pi < \frac{8n^3}{6} \cdot 16\frac{1}{100n^2}\pi < n.$$

Tehát van olyan $2n$ pont az egységnégyzetben, melyek maximum n kicsi háromszöget határoznak meg. Minden kicsi háromszögnek hagyjuk el egy tetszőleges csúcsát. Maximum n csúcsot hagytunk el, tehát maradt legalább n csúcsunk az egységnégyzetben, melyek közül bármely három által alkotott háromszög területe nem kisebb, mint $\frac{1}{100n^2}$. Ezzel a tétel bizonyítása kész. $\square$

A fenti bizonyításban is megfigyelhető, hogy bár a véletlen választással kapott $2n$ elemű halmaz nem egy konkrét konstrukció, különböző paraméterei jól kiszámíthatók, mint például a kicsi háromszögek számának várható értéke. Ez az átmeneti tulajdonsága a véletlen objektumoknak nagyon hasznos, amikor nehéz egy adott tulajdonságot kielégítő konkrét konstrukciót megadnunk.

Jellemző a valószínűségi módszerre, ahogy a bizonyításban nem rögtön a konkrét becslést szolgáltató konstrukciót próbáltuk kiválasztani, hanem egy nagyobb konstrukciót igazítottunk megfelelőre.

Komlós, Pintz és Szemerédi 1982-ben ugyancsak valószínűségi módszerrel a fentinél erősebb $M(n) \geq c \cdot \frac{\ln n}{n^2}$ egyenlőtlenséget bizonyították, ahol c pozitív konstans [3].

2. Az Azuma-egyenlőtlenség alkalmazásai

Ebben a fejezetben az Azuma-egyenlőtlenség alkalmazásait fogjuk megvizsgálni. A fejezet bizonyításai Lovász László [4] előadásán elhangzottakon alapulnak. Az alaptételt 1967-ben publikálta Kazuoki Azuma az [5] cikkben.

2.0.2. Tétel. *Azuma-egyenlőtlenség: Legyen $X_0, X_1, \ldots, X_n$ olyan martingál, melyre $k \in [n]$ esetén $|X_k - X_{k-1}| \leq c_k$ teljesül 1 valószínűséggel. Ekkor a következő egyenlőtlenségek érvényesek:*

$$P(X_n - X_0 \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{2\sum_1^n c_k^2}\right),$$

$$P(|X_n - X_0| \geq \lambda) \leq 2 \cdot \exp\left(-\frac{\lambda^2}{2\sum_1^n c_k^2}\right).$$

A továbbiakban $c_k = 1$ lesz $\forall k \in [n]$ esetén.

2.1. Véletlen leképezés képterének mérete

Adott egy n elemű S véges halmaz. A véletlen $\phi : S \to S$ leképezés S minden elemének egyenlő $\frac{1}{n}$ valószínűséggel rendeli S valamely elemét, a többi hozzárendeléstől függetlenül.

2.1.1. Tétel. $P\Big(\big||\phi(S)| - E(|\phi(S)|)\big| \geq \lambda\Big) \leq 2 \cdot e^{-\frac{\lambda^2}{2n}}$

Bizonyítás. Legyen $S = [n]$ és $X_k = E(|\phi(S)|\big|\phi(1), \phi(2), \ldots, \phi(k))$ valószínűségi változó. Tehát X_k a képtér méretének feltételes várható értéke, ha már tudjuk, ϕ hova képzi S első k elemét.

Az előbb definiált valószínűségi változó sorozat martingál, hiszen a feltételes várható értékben a σ-algebra egyre bővebb, és $E((X|G)|F) = E(X|F), F \subset G$ esetén, továbbá nyilván értelmes, hiszen $E(|\phi(S)|)$ véges.

Igaz továbbá, hogy $|X_k - X_{k-1}| \leq 1$, hiszen egy elem képének megismerése 1-nél többel nem módosíthat a feltételes várható értéken.

Az Azuma-egyenlőtlenség szerint tehát $P(|X_n - X_0| \geq \lambda) \leq 2 \cdot e^{-\frac{\lambda^2}{2n}}$.

Azonban $X_n = |\phi(S)|$ és $X_0 = E(|\phi(S)|\big|F_0) = E(|\phi(S)|)$, ahol F_0 a triviális σ-algebra. A tétel állítását ezzel bebizonyítottuk. $\square$

A bizonyításban használt martingál képzési módot a fejezet későbbi alkalmazásaiban is látni fogjuk majd. Az így kapott martingálok a Doob-martingálok.

Az $E(|\phi(S)|)$ érték könnyen meghatározható, $E(|\phi(S)|) = n \cdot (1 - (\frac{n-1}{n})^n)$, tehát $\frac{E(|\phi(S)|)}{n} \to 1 - \frac{1}{e}$.

2.2. Véletlen részgráf élsűrűsége

Legyen G v csúcsú és f élű gráf. Válasszunk ki véletlenszerűen k darab csúcsot G-ből, és nézzük az így kapott feszített véletlen részgráfot, H-t. H éleinek számát jelöljük Y-nal. A következő tétel annak a valószínűségét becsüli felülről, hogy a feszített részgráf élsűrűsége és az eredeti gráf élsűrűsége λ-nál jobban eltér egymástól.

2.2.1. Tétel. $P\left(\left|\frac{Y}{\binom{k}{2}} - \frac{f}{\binom{v}{2}}\right| \geq \lambda\right) \leq 2 \cdot e^{\frac{-k\lambda^2}{8}}$.

Bizonyítás. H csúcsait a következőképpen választhatjuk ki. Először kiválasztunk egy v_1 csúcsot $V(G)$-ből. Aztán egy v_2 csúcsot $V(G) - \{v_1\}$-ből. Ezt folytatjuk, végül v_k-t $V(G) - \{v_1, v_2, \ldots, v_{k-1}\}$-ből választjuk. Egy adott kiválasztás a már kiválasztott csúcsoktól függetlenül, a lehetséges csúcsok között egyenlő valószínűséggel történik meg.

Vezessük be a következő valószínűségi változókat:

$$X_i = \frac{1}{k-1} E(Y|v_1, v_2, \ldots, v_i),\ i = 0{,}1, \ldots, k.$$

Ekkor a definiált valószínűségi változó sorozat martingál, továbbá mivel egy csúcs maximum $k-1$ élen van rajta, $|X_i - X_{i-1}| \leq 1$. Az Azuma-egyenlőtlenség szerint tehát

$$P(|X_k - X_0| \geq \lambda \cdot \frac{k}{2}) \leq 2 \cdot e^{\frac{-k\lambda^2}{8}}.$$

Határozzuk meg az X_k és X_0 valószínűségi változókat. Mivel Y $\sigma\{v_1, v_2, \ldots, v_k\}$-mérhető, $X_k = \frac{Y}{k-1}$. Továbbá $E(Y) = \binom{k}{2} \cdot \frac{f}{\binom{v}{2}}$, hiszen az élsűrűség G-ben $\frac{f}{\binom{v}{2}}$ és H-ban $\binom{k}{2}$ él van. Tehát X_0-t is meghatároztuk, $X_0 = \frac{E(Y)}{k-1} = \frac{k}{2} \cdot \frac{f}{\binom{v}{2}}$.

A kapott eredményeket a fenti egyenlőtlenségbe beírva és rendezve éppen a tétel állítását kapjuk. □

2.3. Véletlen részgráf függetlenségi száma

A G gráf csúcsaiból véletlenszerűen kiválasztunk k darabot, az így kapott feszített véletlen részgráf legyen H. Jelöljük $\alpha(H)$-val H maximális független csúcshalmazának méretét, azaz a függetlenségi számát. A következő mondható ekkor az $\alpha(H)$ valószínűségi változóról.

2.3.1. Tétel. $P\big(|\alpha(H) - c(k)| \geq \lambda\big) \leq 2 \cdot e^{-\frac{\lambda^2}{k}}$, *valamilyen $c(k)$ konstansra.*

Bizonyítás. Az előbbi tétel bizonyításához hasonlóan járhatunk el. H csúcsait ugyanúgy válasszuk egyenként a korábban nem választott csúcsok közül, és ha v_i jelöli az

i-ediknek beválasztott csúcsot, akkor vezessük be a következő valószínűségi változókat:

$$X_i = E(\alpha(H)|v_1, v_2, \ldots, v_i),\ i = 0{,}1, \ldots, k.$$

Ekkor a fenti valószínűségi változó sorozat martingál. $|X_i - X_{i-1}| \leq 1$ teljesül, hiszen egy új csúcs beválasztása H-ba nem módosíthatja a maximális független csúcshalmaz feltételes várható értékét 1-nél többel.

Az Azuma-egyenlőtlenség szerint

$$P\big(|X_k - X_0| \geq \lambda\big) \leq 2 \cdot e^{-\frac{\lambda^2}{k}}.$$

Mivel X_0 1 valószínűséggel konstans, $X_k = \alpha(H)$, ezért a tétel állítását bebizonyítottuk. □

A bizonyításból látszik, hogy a $c(k)$ konstans éppen $E(\alpha(H))$.

Teljesen hasonló gondolatmenettel becsülhetnénk hasonló valószínűségeket a véletlen feszített részgráf más paramétereire nézve is, például a kromatikus számot, vagy a legnagyobb teljes részgráf méretét. Ha nem a csúcsokat választjuk véletlenszerűen, hanem a véletlen részgráf éleit, akkor a keletkező véletlen k élű részgráf maximális független élhalmazára is adhatunk fenti típusú becslést az Azuma-egyenlőtlenség segítségével.

2.4. Véletlen gráf klikkjei

$G(n,p)$ az n csúcsú véletlen gráf, azaz egy olyan gráf, melyre $V(G(n,p)) = [n]$, az élhalmazt pedig úgy kapjuk, hogy bármely két csúcsot egymástól függetlenül p valószínűséggel összekötünk. Azt szeretnénk megbecsülni, hogy mennyi a valószínűsége $G(n, \frac{1}{2})$-ben egy k-klikk kialakulásának. Egy G gráf maximális klikkjének méretét $\omega(G)$-vel jelöljük.

Szükségünk lesz a fejezet eddigi bizonyításaiban használt gondolatmenet általánosítására és egy gráfok függetlenségi számával kapcsolatos tételre.

2.4.1. Tétel. *McDiarmid-egyenlőtlenség: Adott a* $(\mathcal{X}, \mathcal{B})$ *mérhető tér, és egy* $f : \mathcal{X}^n \to \mathbb{R}$ *mérhető függvény, amelyre*

$$|f(x_1, x_2, \ldots x_n) - f(y_1, y_2, \ldots, y_n)| \leq 1,$$

amikor $(x_1, x_2, \ldots x_n)$ *és* $(y_1, y_2, \ldots, y_n)$ *vektorok csak egy koordinátában térnek el. Legyenek* $X_1, X_2, \ldots, X_n$ $\mathcal{X}$*-be képező független valószínűségi változók. Ekkor*

$$P(f(X_1, X_2, \ldots, X_n) - E(f(X_1, X_2, \ldots, X_n)) \geq \lambda) \leq e^{-\frac{\lambda^2}{2n}}$$

és

$$P(|f(X_1, X_2, \dots, X_n) - E(f(X_1, X_2, \dots, X_n))| \geq \lambda) \leq 2 \cdot e^{-\frac{\lambda^2}{2n}}.$$

Bizonyítás. Ugyanúgy járunk el, mint korábban. Vezessük be az

$$Y_k = E(f(X_1, X_2, \dots, X_n)|X_1, X_2, \dots, X_k)$$

martingált. Nyilván $Y_n = f(X_1, X_2, \dots, X_n)$ és $Y_0 = E(f(X_1, X_2, \dots, X_n)))$. Végül $|X_k - X_{k-1}| \leq 1$ is teljesül: a lemma feltétele éppen ezt biztosítja, hiszen a különbségben felmerülő helyek legfeljebb egy koordinátában, a k-adikban térnek el egymástól. □

2.4.2. Tétel. *Legyen a v csúcsú és e élű G gráf függetlenségi száma $\alpha(G)$. Ekkor tetszőleges $p \in [0,1]$ számra*

$$\alpha(G) \geq vp - ep^2.$$

Bizonyítás. Tetszőleges H gráfra teljesül, hogy $\alpha(H) \geq |V(H)| - |E(H)|$, hiszen ha egy csúcs nincs benne a maximális független csúcshalmazban, akkor biztos, hogy él köti össze annak valamelyik elemével. Így, ha p valószínűséggel, egymástól függetlenül választunk ki csúcsokat G-ből, és a feszített részgráf éleinek, csúcsainak, és függetlenségi számának várható értékét rendre a, b, c-vel jelöljük, akkor egyrészt $c \geq b - a$ másrészt $b = vp$, $a = ep^2$ és $\alpha(G) \geq c$ miatt a tétel állítását bebizonyítottuk. □

A következő tétel szerint a $G(n, \frac{1}{2})$ gráfban, ha a $k+1$-es klikkek várható értéke legalább 1, akkor szinte biztos, hogy van k méretű klikk a gráfban.

2.4.3. Tétel. *Ha k a legnagyobb olyan egész szám, melyre*

$$\binom{n}{k+1} 2^{-\binom{k+1}{2}} \geq 1 \tag{2.4.1}$$

akkor

$$P(\omega(G(n, \tfrac{1}{2})) < k) \leq e^{-\frac{n^2}{256k^8}}$$

elég nagy n-re.

Bizonyítás. Első megfigyelésünk, hogy (2.4.1) szerint $k \sim 2\log_2 n$. Valóban, az $\binom{n}{k+1} < n^{k+1}$ becslést alkalmazva a (2.4.1) egyenlőtlenségre, majd logaritmust véve és rendezve a $k < 2\log_2 n$ egyenlőtlenséget kapjuk.

Az alsó becsléshez $\binom{n}{k+2} 2^{-\binom{k+2}{2}} < 1$-ből indulunk, és alkalmazzuk az $\left(\frac{n}{k+2}\right)^{k+2} < < \binom{n}{k+2}$ egyenlőtlenséget. Logaritmusvétel és rendezés után $2\log_2 n - 2\log_2(k+2) < < k+1$-et kapjuk. Ezzel $k \sim 2\log_2 n$-t igazoltuk.

Tekintsük azt a H gráfot, melynek csúcsai a $G(n, \frac{1}{2})$ k méretű klikkjei, és két csúcsot akkor kötünk össze H-ban, ha a megfelelő klikkeknek van közös élük. Ekkor

$$P(\omega(G(n, \tfrac{1}{2})) < k) = P(\alpha(H) = 0). \tag{2.4.2}$$

Legyen $X_{(i,j)}$ annak az eseménynek az indikátora, hogy i és j össze vannak kötve $G(n, \frac{1}{2})$-ben. Ekkor van olyan f mérhető leképezés, mely az $X_{(i,j)}$ $i, j = 1,2\ldots, n$ $i \neq j$ változóktól függ, és az $\alpha(H)$ értéket veszi fel. Nyilván egy $X_{(i,j)}$ megváltozása az $\alpha(H)$ értéket 1-nél nagyobb mértékben nem változtatja. Vezessük be az $s = E(\alpha(H))$ jelölést. A 2.4.1. Tételt alkalmazva:

$$P(\alpha(H) = 0) = P(\alpha(H) \leq 0) \leq P(\alpha(H) - s \leq -s) \leq e^{\frac{-s^2}{n(n-1)}} < e^{\frac{-s^2}{n^2}}. \tag{2.4.3}$$

Már csak s-et kéne alulról becsülnünk. Legyen $v = E(|V(H)|)$ és $e = E(L(H))$, ahol $L(H)$ a H éleinek száma. Ekkor $v = \binom{n}{k} \cdot 2^{-\binom{k}{2}}$.

Kiszámoljuk e-t is. Annak a valószínűsége, hogy H két csúcsa, A és B, össze van kötve, feltételezve, hogy $|A \cap B| = i$, éppen $2^{-\left(2\binom{k}{2} - \binom{i}{2}\right)}$, ahol $2 \leq i \leq k-1$. Így megszámolva, hogy $|A \cap B| = i$ hányszor fordul elő, és súlyozva a megfelelő valószínűségekkel azt kapjuk, hogy

$$e = \sum_{i=2}^{k-1} \binom{n}{k}\binom{k}{i}\binom{n-k}{k-i} 2^{-\left(2\binom{k}{2} - \binom{i}{2}\right)}.$$

Írjuk fel a 2.4.2. Tétel állítását a $G(n, \frac{1}{2})$ gráfra. Tetszőleges $p \in [0,1]$ esetén

$$\alpha(G(n, \tfrac{1}{2})) \geq |V(G(n, \tfrac{1}{2}))|p - |E(G(n, \tfrac{1}{2}))|p^2.$$

A fenti egyenlőtlenség a várható értékekre $s \geq vp - ep^2$ egyenlőtlenséget adja. Az alsó becslés $p = \dfrac{v}{2e}$ esetén a legerősebb, és $1 \geq p$, hiszen $e \geq v$ látszik e összegalakjából, már az $i = k-1$-hez tartozó tag is nagyobb v-nél, ha n elég nagy. Így $s \geq \dfrac{v^2}{4e}$. Felülről becsüljük az

$$\frac{e}{v^2} = \sum_{i=2}^{k-1} \frac{1}{\binom{n}{k}} \binom{k}{i}\binom{n-k}{k-i} 2^{\binom{i}{2}}$$

összeget. Azt szeretnénk megmutatni, hogy az összeg két szélső tagja nagyon nagy a többi taghoz képest, ezért az összeg felülről becsülhető a két szélső tag összegének kétszeresével. Két szomszédos tag hányadosa az összegben

$$\left(\frac{1}{\binom{n}{k}}\binom{k}{i+1}\binom{n-k}{k-i-1} 2^{\binom{i+1}{2}}\right) \Big/ \left(\frac{1}{\binom{n}{k}}\binom{k}{i}\binom{n-k}{k-i} 2^{\binom{i}{2}}\right) = \frac{2^i (k-i)^2}{(i+1)(n-2k+i+1)}.$$

Innen látható, hogy nagy n esetén az összeg tagjai először csökkennek, majd egy indexen túl növekednek.

A csökkenő szakasz tagjainak összegét felülről becsüljük az első tag kétszereresével. Ez korrekt, hiszen a csökkenő szakasz második tagja $\frac{4(k-2)}{3(n-2k+3)}$-szerese az első tagnak, a többi tag pedig még kisebb. Mivel a tagok száma legfeljebb k, összegük legfeljebb $\frac{4k(k-2)}{3(n-2k+3)}$-szerese az elsőnek, ami $k^2 < \frac{3}{4}n$ esetén kisebb, mint 1. Ez $k \sim \sim 2\log_2 n$ miatt minden elég nagy n-re teljesül, tehát az első tag kétszerese valóban jó felső becslés a csökkenő szakasz tagjainak összegére.

Hasonlóan bizonyítjuk, hogy az utolsó tag kétszerese nagyobb, mint a növekvő szakasz elemeinek összege. Az utolsó tag $\frac{2^k}{(k-1)(n-k-1)}$-szerese az utolsó előttinek, ami $k \sim 2\log_2 n$ miatt nagyobb, mint $n^{1-\epsilon}$. Tehát az utolsó tag legalább $n^{1-\epsilon}$-szorosa a növekvő szakasz bármelyik másik elemének. Mivel a növekvő szakasz legfeljebb k elemű, így ezek összege valóban felülről becsülhető az utolsó tag kétszeresével.

Az első tagot felülről becsüljük a következőképpen:

$$\frac{1}{\binom{n}{k}}\binom{k}{2}\binom{n-k}{k-2}2^{\binom{2}{2}} = k^2(k-1)^2\frac{(n-k)!^2}{n!\,(n-2k+2)!} \leq \frac{k^2(k-1)^2}{n(n-1)} < \frac{k^4}{n^2}.$$

Az $\binom{n}{k} \cdot 2^{-\binom{k}{2}} > \frac{n^3}{k^3}2^{1-k}$ egyenlőtlenség segítségével az utolsó tagot is felülről becsüljük

$$\frac{1}{\binom{n}{k}}k(n-k)2^{\binom{k-1}{2}} < kn\frac{2^{1-k}}{\binom{n}{k}\cdot 2^{-\binom{k}{2}}} < \frac{k^4}{n^2}.$$

Az eddigi becsléseinket felhasználva $\frac{e}{v^2} < \frac{4k^4}{n^2}$, így s-re a következő becslést kaptuk: $\frac{n^2}{16k^4} < s$. Tehát a (2.4.2), (2.4.3) összefüggések szerint

$$P(\omega(G(n, \tfrac{1}{2})) < k) \leq e^{-\frac{n^2}{256k^8}}.$$

□

3. Talagrand-elmélet

3.1. Bevezető

Michel Talagrand 1995-ben publikálta az itt ismertetett elméletet alapjait [6]. Adott (Ω, A, μ) valószínűségi mező esetén bevezetjük Ω^n-en az úgynevezett Talagrand-féle konvex távolságot. Azt a jelölési konvenciót követjük, hogy a szorzattér elemeinek indexelés nélküli változóneveket adunk, tehát például $x \in \Omega^n$, és ekkor x i-edik koordinátáját x_i-vel jelöljük. Ω^n-en a μ^n valószínűségi mértéket P-vel jelöljük.

3.1.1. Definíció. *Legyen $x \in \Omega^n$ és $A \subseteq \Omega^n$. Ekkor x és A Talagrand-féle konvex távolsága a következő*

$$d_T(x, A) = \inf \left\{ t \in \mathbb{R} : \forall \alpha(x) \in \mathbb{R}^n \ \exists y \in A \ \sum_{i=1}^{n} \alpha_i(x) 1(x_i \neq y_i) \leq t ||\alpha(x)|| \right\}.$$

Ha α minden koordinátája egyenlő és nem nulla, akkor a $\sum_{i=1}^{n} \alpha_i 1(x_i \neq y_i)$ kifejezés azon $y \in A$ esetén minimális, ahol x és A Hamming-távolsága felvétetik. Innen x és A Hamming-távolságát $d_H(x, A)$-val jelölve a $\frac{d_H(x,A)}{\sqrt{n}} \leq d_T(x, A)$ összefüggést kapjuk. A Hamming-távolságot bizonyos speciális $A \subseteq \Omega^n$ halmazoktól a 3.4 szakaszban vizsgáljuk.

A fejezet további részében Talagrand izoperimetrikus egyenlőtlenségének különböző alkalmazási lehetőségeit mutatjuk be. Az alaptételt bizonyítás nélkül mondjuk ki, egy bizonyítás megtalálható például [7]-ben.

3.1.2. Tétel. *Bármely mérhető $A \subseteq \Omega^n$ esetén*

$$\int_{\Omega^n} \exp\left(\frac{1}{4} d_T^2(x, A)\right) dP(x) \leq \frac{1}{P(A)}.$$

3.1.3. Tétel. $P(d_T(x, A) > t) \leq \exp(-\frac{t^2}{4})/P(A)$.

Bizonyítás. Legyen $\Omega_1 = \{x \in \Omega^n : d_T(x, A) > t\}$. A következő egyszerű átalakítás bizonyítja az állításunkat. A 3.1.2. Tételből kiindulva

$$\frac{1}{P(A)} \geq \int_{\Omega^n} \exp\left(\frac{1}{4} d_T^2(x, A)\right) dP(x) \geq$$
$$\geq \int_{\Omega_1} \exp\left(\frac{1}{4} d_T^2(x, A)\right) dP(x) \geq \exp\left(\frac{t^2}{4}\right) P(\Omega_1).$$

□

3.2. Steiner-fa probléma

Legyen adott n pont, $x = \{x_1, x_2, \ldots, x_n\}$, a d dimenziós térben. Ekkor az x-beli pontok Steiner-fája az a minimális élhosszösszegű fa, mely minden x_i pontot tartalmaz. Tehát az $ST(x)$ Steiner-fának a csúcshalmaza nem feltétlenül esik egybe x-szel. Jelöljük $ST(x)$ élhosszösszegét $E_{ST(x)}$-szel. Hasonlóan definiálhatjuk x halmaz minimális élhosszösszegű feszítő fáját, $MST(x)$-et, azzal a különbséggel, hogy $MST(x)$ csúcshalmaza x. Jelöljük $MST(x)$ éleinek összegét $E_{MST(x)}$-szel. Nyilván $E_{ST(x)} \leq E_{MST(x)}$.

Ebben a szakaszban azt fogjuk vizsgálni, hogy mi mondható az $E_{ST(X)}$ valószínűségi változóról, ha X elemeit egymástól függetlenül egyenletes eloszlás szerint választjuk az egységnégyzetből.

3.2.1. Lemma. *Legyen L a $[0,1]^d$ véges részhalmazain értelmezett valós értékű függvény, mely monoton a következő értelemben:*

$$L(\{x_1, x_2, \ldots, x_n\}) \leq L(\{x_1, x_2, \ldots, x_n, x_{n+1}\}). \tag{3.2.1}$$

Tegyük fel továbbá, hogy tetszőleges véges $x = \{x_1, x_2, \ldots, x_n\}$ halmazhoz vannak olyan $\alpha_i(x), i = 1,2\ldots, n$ nemnegatív számok, hogy tetszőleges $y = \{y_1, y_2, \ldots, y_n\}$ esetén

$$L(x) \leq L(y) + \sum_{i=1}^{n} \alpha_i(x) 1(x_i \neq y_i). \tag{3.2.2}$$

Ekkor ha létezik olyan c szám, hogy tetszőleges $x \subset \mathbb{R}^d$ véges halmaz esetén az előbb definiált $\alpha_i(x)$-ek teljesítik a

$$\sum_{i=1}^{|x|} \alpha_i(x)^2 \leq c^2 \tag{3.2.3}$$

egyenlőtlenséget, akkor véve n darab független, egyenletes eloszlású X_i valószínűségi változót $[0,1]^d$-ben

$$P(|L(\{X_1, X_2 \ldots, X_n\}) - M_n| \geq t) \leq 4 \exp\left(-\frac{t^2}{4c^2}\right),$$

ahol M_n az $L(\{X_1, X_2 \ldots, X_n\})$ valószínűségi változó mediánja.

Bizonyítás. A lemmát a 3.1.3. Tétel segítségével látjuk be. A Talagrand-féle konvex távolság alkalmazhatóságához be kell vezetnünk egy valószínűségi mezőt, ez a természetes $[0,1]^d$ lesz a Lebesgue-mértékkel ellátva. Legyen $x = \{x_1, x_2, \ldots, x_n\}$ és

$$A(a) = \{\{y_1, y_2, \ldots, y_n\} : L(\{y_1, y_2, \ldots, y_n\}) \leq a\}.$$

Ekkor a (3.2.2) egyenlőtlenség jobb oldalán álló összeget y-ban minimalizálva $A(a)$-n azt kapjuk, hogy

$$L(x) \leq L(y) + \sum_{i=1}^{n} \alpha_i(x) 1(x_i \neq y_i) \leq a + \min_{y \in A(a)} \sum_{i=1}^{n} \alpha_i(x) 1(x_i \neq y_i) =$$
$$= a + ||\alpha|| \min_{y \in A(a)} \sum_{i=1}^{n} \frac{\alpha_i}{||\alpha||}(x) 1(x_i \neq y_i) \leq a + c \cdot d_T(x, A(a)),$$

ahol az utolsó egyenlőtlenség egyszerűen a Talagrand-féle konvex távolság definíciója és (3.2.3) feltétel miatt igaz. Átrendezve az $(L(x)-a)/c \leq d_T(x, A(a))$ egyenlőtlenséget kapjuk, tetszőleges x véges halmaz esetén. Így a 3.1.3. Tételt az $\{X_1, X_2 \ldots, X_n\}$ véletlen halmazra alkalmazva

$$P\big((L(\{X_1, X_2 \ldots, X_n\}) - a)/c \geq t\big) \leq \exp\left(-\frac{t^2}{4}\right) / P(A(a)). \tag{3.2.4}$$

Ha a (3.2.4) egyenlőtlenségben elvégezzük az $u = ct$ helyettesítést, és rendre $a = M_n$-et, illetve $a = M_n - u$-t helyettesítünk, akkor a

$$P\big(L(\{X_1, X_2 \ldots, X_n\}) - M_n \geq u\big) \leq \exp\left(-\frac{u^2}{4c^2}\right) / P(A(M_n)) \leq 2\exp\left(-\frac{u^2}{4c^2}\right)$$

illetve

$$\frac{1}{2} \leq P\big(L(\{X_1, X_2 \ldots, X_n\}) \geq M_n\big) \leq \exp\left(-\frac{u^2}{4c^2}\right) / P(A(M_n - u))$$

összefüggéseket kapjuk. Ebből a két összefüggésből pedig már következik a tétel állítása,

$$P(|L(\{X_1, X_2 \ldots, X_n\}) - M_n| \geq t) \leq 4\exp\left(-\frac{t^2}{4c^2}\right),$$

$A(a)$ és M_n definíciója miatt. Ezzel a lemma bizonyítása kész. □

A bizonyított lemma segítségével a következőt mondhatjuk véletlen Steiner-fákra vonatkozóan.

3.2.2. Tétel. *Legyenek X_i, $i = 1,2 \ldots n$ az egységnégyzetben egyenletes eloszlású, független valószínűségi változók. Ekkor létezik olyan c abszolút konstans, hogy minden $t > 0$ esetén*

$$P\left(|(E_{ST(\{X_1, X_2, \ldots, X_n\})}) - M_n| \geq t\right) \leq 4\exp\left(-\frac{t^2}{4c^2}\right),$$

ahol M_n az $E_{ST(\{X_1, X_2, \ldots, X_n\})})$ valószínűségi változó mediánja.

Bizonyítás. Elegendő lenne bebizonyítani, hogy a 3.2.1. Lemma feltételeit az $E_{ST(x)}$ függvény teljesíti. A monotonitás triviális. Adott x véges halmaz elemeihez definiálnunk kell $\alpha_i(x)$ értékeket. Általában jó ötlet, ha az $\alpha_i(x)$ számokat úgy választjuk, hogy $\alpha_i(x)$ azt mérje, mennyire változtat $E_{ST(x)}$ értékén az, hogy $x_i \in x$. Ennek megfelelően $\alpha_i(x)$ legyen az x_i csúcsból induló élek összegének kétszerese x minimális feszítő fájában, $MST(x)$-ben.

Ellenőrizzük a (3.2.2) feltételt. Legyen y egy tetszőleges n elemű ponthalmaz az egységnégyzetben. Tekintsük azt a G gráfot, melynek $x \cup y$ a csúcshalmaza, az élhalmaza pedig

$$E(ST(y)) \cup \{(x_i, x_j) \in E(MST(x)) : (x_i \notin y \vee x_j \notin y)\}.$$

Ekkor a G gráfban az élhosszak összegénél nagyobb $E_{ST(y)} + \sum_{i=1}^{n} \alpha_i(x) 1(x_i \neq y_i)$, ami a (3.2.2) feltétel jobb oldala. Mivel $E_{ST(x)} \leq E_{ST(x \cup y)}$, elegendő megmutatnunk, hogy G összefüggő.

Ha $x = y$ vagy x és y diszjunktak, akkor (3.2.2) triviálisan igaz. Az y-beli csúcsok összefüggők G-ben, hiszen $ST(y)$ részgráfja G-nek. Belátjuk, hogy tetszőleges x_i csúcs elérhető egy $x \cap y$-beli z_0 csúcsból G-ben, ebből már következik, hogy G összefüggő. Tegyük fel, hogy $MST(x)$ összes élét hozzávesszük G-hez, a kapott gráf legyen a G' gráf. Ekkor G'-ben van z_0-t x_i-vel összekötő út. Ez az út csak akkor tartalmaz olyan élet, mely nem G-beli, ha az két y-beli csúcsot köt össze. Ekkor azonban ezt az élet helyettesíthetjük egy G-ben haladó úttal y összefüggősége miatt. Így G valóban összefüggő, $E_{ST(x)}$ a (3.2.2) feltételt teljesíti.

Még ellenőriznünk kell a (3.2.3) feltételt. Mivel $MST(x)$ minden csúcsának a foka legfeljebb 6, a $\sum_{i=1}^{n} \alpha_i(x)^2$ összeg felülről becsülhető az $MST(x)$-beli élek négyzetösszegének konstansszorosával. Az pedig, hogy az élek négyzetösszege n-től függetlenül felülről becsülhető, egy ismert geometriai tény, amit mi nem bizonyítunk. Egy bizonyítása ennek az állításnak megtalálható például a [8] cikkben. □

3.3. A leghosszabb monoton részsorozat probléma

Adott $x \in [0,1]^n$ sorozat esetén legyen $I_n(x)$ a leghosszabb monoton növekedő részsorozat hossza x-ben. Legyenek adottak az X_i $(i = 1,2,\ldots,n)$ független, $[0,1]$ intervallumon egyenletes eloszlású valószínűségi változók és legyen $X = (X_1, X_2, \ldots, X_n)$. Ebben a szakaszban $I_n(X)$ eloszlását fogjuk vizsgálni.

3.3.1. Tétel. *Legyen az $I_n(X)$ valószínűségi változó mediánja M_n. Ekkor $t > 0$ esetén*

$$P(I_n(X) \geq M_n + t) \leq 2 \exp\left(-\frac{t^2}{4(M_n + t)}\right), \tag{3.3.1}$$

valamint

$$P(I_n(X) \leq M_n - t) \leq 2\exp\left(-\frac{t^2}{4M_n}\right). \tag{3.3.2}$$

Bizonyítás. Először a (3.3.1) egyenlőtlenséget látjuk be. $I_n(x) = k$ esetén legyen $J(x) = \{i_1, i_2, \dots i_k\}$ a k hosszú monoton növekedő részsorozat tagjainak indexe. Ekkor tetszőleges $y \in [0,1]^d$-re igaz a következő összefüggés:

$$I_n(y) \geq I_n(x) - \sum_{i \in J(x)} 1(x_i \neq y_i). \tag{3.3.3}$$

Definiáljuk az $\alpha_i(x)$ értékeket. Legyen $\alpha_i(x) = \frac{1}{\sqrt{I_n(x)}}$, ha $i \in J(x)$, különben pedig 0. Vezessük be az $A(a) = \{y \in [0,1]^n : I_n(y) \leq a\}$ halmazokat. Felülről becsüljük a $\sum_{i \in J(x)} 1(x_i \neq y_i)$ összeg $A(a)$-n vett minimumát:

$$\min_{y \in A(a)} \sum_{i \in J(x)} 1(x_i \neq y_i) = \min_{y \in A(a)} \sqrt{I_n(x)} \sum_{i \in J(x)} \frac{1}{\sqrt{I_n(x)}} 1(x_i \neq y_i) \leq \sqrt{I_n(x)} d_T(x, A(a)).$$

Ezt az eredményt beírva a (3.3.3) egyenlőtlenségbe azt kapjuk, hogy

$$a \geq I_n(x) - \sqrt{I_n(x)} d_T(x, A(a)) \text{ azaz } d_T(x, A(a)) \geq \frac{I_n(x) - a}{\sqrt{I_n(x)}}. \tag{3.3.4}$$

Felhasználva, hogy $f(t) = \frac{t-a}{\sqrt{t}}$ $t > 0$ esetén monoton nő, továbbá a (3.3.4) egyenlőtlenséget és a 3.1.3. Tételt, a következőt kapjuk:

$$\begin{aligned} P(I_n(X) \geq a + t) &= P\left(\frac{I_n(X) - a}{\sqrt{I_n(X)}} \geq \frac{t}{\sqrt{a+t}}\right) \leq \\ &\leq P\left(d_T(x, A(a)) \geq \frac{t}{\sqrt{a+t}}\right) \leq \exp\left(-\frac{t^2}{4(a+t)}\right) / P(A(a)). \end{aligned}$$

Innen $a = M_n$ helyettesítést elvégezve éppen a (3.3.1) egyenlőtlenséget kapjuk. A (3.3.2) egyenlőtlenség bizonyításához is a (3.3.4) egyenlőtlenséget és a 3.1.3. Tételt használjuk.

$$\begin{aligned} \frac{1}{2} \leq P(I_n(X) \geq M_n) &= P\left(\frac{I_n(X) - a}{\sqrt{I_n(X)}} \geq \frac{M_n - a}{\sqrt{M_n}}\right) = \\ &= P\left(\frac{I_n(X) - M_n + t}{\sqrt{I_n(X)}} \geq \frac{t}{\sqrt{M_n}}\right) \leq \\ &\leq P\left(d_T(x, A(M_n - t)) \geq \frac{t}{\sqrt{M_n}}\right) \leq \exp\left(-\frac{t^2}{4M_n}\right) / P(A(M_n - t)) \end{aligned}$$

Átrendezve a kapott egyenlőtlenséget éppen a (3.3.2)-t kapjuk. Ezzel a 3.3.1. Tétel bizonyítása kész. □

3.4. H-halmazrendszerek

Adott a $V = \{0,1\}^n$ véges halmaz és ezen a majorizálás részbenrendezés, amit a következő tulajdonság definiál: ha $x, y \in V$, akkor $x \prec y \Leftrightarrow x_i \leq y_i \ \forall i \in [n]$. Azt mondjuk, hogy $A \subseteq V$ egy H-halmazrendszer, ha $q \in A$ és $p \prec q$ esetén $p \in A$.

Legyen μ egy valószínűségi mérték a diszkrét $\{0,1\}$ halmazon. Ennek a diszkrét valószínűségi mezőnek az n-edik hatványa egy V alaphalmazú, P valószínűségi mértékkel rendelkező diszkrét valószínűségi mező. Ekkor a kapott valószínűségi mezőben a következő becslést tudjuk adni H-halmazrendszerekkel kapcsolatban:

3.4.1. Tétel. *Tetszőleges $t, k \in \mathbb{N}^+$ és $A \subseteq V$ H-halmazrendszer esetén*

$$P(d_H(x, A) \geq t) \leq \frac{1}{P(A)} \exp\left(-\frac{t^2}{4k}\right) + P\left(\sum_{i=1}^{n} 1(x_i = 1) \geq k\right).$$

Bizonyítás. Mivel A H-halmazrendszer, a $d_H(x, A)$ Hamming-távolságot felírhatjuk olyan típusú összegként, amelyekkel a Talagrand-féle konvex távolságot definiáltuk.

$$d_H(x, A) := \min_{y \in A} \sum_{i=1}^{n} 1(x_i \neq y_i) = \min_{y \in A} \sum_{i=1}^{n} 1(x_i = 1) 1(x_i \neq y_i).$$

Így ha $\alpha_i(x)$ definíciója a következő:

$$\alpha_i(x) = 1(x_i = 1) \cdot \left(\sum_{i=1}^{n} 1(x_i = 1)\right)^{-1/2},$$

akkor a Talagrand-féle konvex távolság definíciója miatt

$$d_H(x, A) \left(\sum_{i=1}^{n} 1(x_i = 1)\right)^{-1/2} \leq d_T(x, A). \tag{3.4.1}$$

Felhasználjuk, hogy $P(A) \leq P(A \cap B) + P(\overline{B})$ tetszőleges A, B eseményekre $A = \{x : d_H(x, A) \geq t\}$, $B = \{x : \sum_{i=1}^{n} 1(x_i = 1) \leq k\}$ szereposztással. Innen (3.4.1) miatt

$$P(d_H(x,A) \geq t) \leq P\left(d_T(x,A)\left(\sum_{i=1}^{n} 1(x_i = 1)\right)^{1/2} \geq t\right) \leq$$

$$\leq P\left(d_T(x,A) \geq \frac{t}{\sqrt{k}}\right) + P\left(\sum_{i=1}^{n} 1(x_i = 1) \geq k\right). \quad (3.4.2)$$

Tehát a 3.1.3. Tétel miatt

$$P(d_H(x,A) \geq t) \leq \frac{1}{P(A)} \exp\left(-\frac{t^2}{4k}\right) + P\left(\sum_{i=1}^{n} 1(x_i = 1) \geq k\right).$$

Ezzel a tétel bizonyítása kész. □

Vegyük észre, hogy a 3.4.1. Tételben szereplő összeg egy binomiális eloszlás faroköszege, ezekre az összegekre pedig jó közelítések ismertek.

4. A lokális lemma

4.1. A lokális lemma és egyszerű alkalmazásai

A következő lemmát Lovász László és Erdős Pál publikálta 1975-ben [9]. A tétel bizonyítása és az alkalmazások a [10] műből származnak.

Definíció: Legyenek $A_1, A_2, \ldots, A_n$ események egy tetszőleges valószínűségi mezőben. Ekkor egy $D = (V, E)$ irányított gráfot, melynek csúcsait rendre az A_i $(i = 1 \ldots n)$ eseményeknek feleltetjük meg, az $A_1, A_2, \ldots, A_n$ események függőségi gráfjának nevezzük, ha A_i független az $\{A_j | (A_i, A_j) \notin E\}$ eseményrendszertől.

4.1.1. Tétel. *Lokális lemma: Ha $D = (V, E)$ az $A_1, A_2, \ldots, A_n$ események függőségi gráfja, és a gráf csúcsaihoz léteznek olyan $0 \leq x_i < 1$ súlyok, hogy*

$$P(A_i) \leq x_i \prod_{(A_i, A_j) \in E} (1 - x_j),$$

akkor

$$P\left(\bigcap_{i=1}^{n} \overline{A_i}\right) \geq \prod_{i=1}^{n} (1 - x_i),$$

tehát speciálisan nem 0.

Bizonyítás. Először a következőt látjuk be: bármely $i \in [n]$ és $S \subset [n]$ esetén, ha $i \notin S$, akkor

$$P\left(A_i \middle| \bigcap_{j \in S} \overline{A_j}\right) \leq x_i. \tag{4.1.1}$$

Ezt S mérete szerinti indukcióval látjuk be. $|S| = 0$ esetén (4.1.1) igaz, hiszen a feltétel szerint $P(A_i) \leq x_i$. Vegyünk most egy S-et, és tegyük fel, hogy (4.1.1) igaz $[n]$-nek minden S-nél kisebb számosságú részhalmazára. Osszuk S-et két részre: $S = S_1 \cup S_2$, ahol $j \in S_1$, ha $j \in S$ és $(A_i, A_j) \in E$ és $S_2 = S - S_1$, tehát A_i független az S_2-beli eseményektől. Vezessük be a $B = \bigcap_{j \in S_2} \overline{A_j}$ jelölést. Ekkor egyszerű számolással:

$$
\begin{aligned}
P(A_i|\bigcap_{j\in S}\overline{A_j}) = \frac{P(A_i\cap\bigcap_{j\in S_1}\overline{A_j}|B)}{P(\bigcap_{j\in S_1}\overline{A_j}|B)} &\le \\
\le \frac{P(A_i|B)}{P(\bigcap_{j\in S_1}\overline{A_j}|B)} = \frac{P(A_i)}{P(\bigcap_{j\in S_1}\overline{A_j}|B)} &\le \\
\le \frac{x_i\prod_{(A_i,A_j)\in E}(1-x_j)}{P(\bigcap_{j\in S_1}\overline{A_j}|B)}&.
\end{aligned}
$$

A nevezőt kellene még alulról becsülnünk. Ezt az indukciós feltevés segítségével a következőképpen tehetjük. Legyen $S_1 = \{j_1, j_2, \dots j_r\}$, ekkor:

$$
\begin{aligned}
P\left(\bigcap_{j\in S_1}\overline{A_j}\Big|B\right) = \big(1-P(A_{j_1}|B)\big)\big(1-P(A_{j_2}|\overline{A_{j_1}}\cap B)\dots\times \\
\times\,(1-P(A_{j_r}|\overline{A_{j_1}}\cap\dots\cap\overline{A_{j_{r-1}}}\cap B)\big).
\end{aligned}
$$

A tényezőkre rendre alkalmazhatjuk az indukciós feltevést, hiszen maximum $|S|-$ $-\,1$ darab esemény metszete van a feltételes valószínűségek feltétel részében. Így

$$
P\left(\bigcap_{j\in S_1}\overline{A_j}\Big|B\right) \ge \prod_{j\in S_1}(1-x_j) \ge \prod_{(A_i,A_j)\in E}(1-x_j).
$$

Ezzel a segédtétel bizonyítása kész. Innen a lokális lemma bizonyítása:

$$
\begin{aligned}
P\left(\bigcap_{i=1}^{n}\overline{A_i}\right) = (1-P(A_1))(1-P(A_2|\overline{A_1}))\dots\times \\
\times\,(1-P(A_n|\bigcap_{i=1}^{n-1}\overline{A_i})) \ge \prod_{i=1}^{n}(1-x_i) > 0.
\end{aligned}
$$

□

A fenti tétel következő speciális esete gyakran hasznos:

4.1.2. Tétel. *Szimmetrikus lokális lemma: Ha a D függőségi gráfban minden fokszám legfeljebb d és bármely eseményre $P(A_i)\le p$, valamint ha $ep(d+1)\le 1$, akkor $P(\bigcap_{i=1}^{n}\overline{A_i}) > 0$.*

Bizonyítás. Az előző lemmát kell alkalmazni $x_i = \frac{1}{d+1}$ választással, és kihasználni, hogy $(1-\frac{1}{d+1})^d > 1/e$. □

A lokális lemma kényelmesen ellenőrizhető elégséges feltételt ad arra nézve, hogy az A_i események egyikének sem teljesülése nem lehetetlen esemény. Emiatt a lemma jól alkalmazható egzisztenciális állítások bizonyítására, amint az alábbi példák is illusztrálják. Érdemes megjegyezni, hogy míg az első fejezetben bemutatott valószínűségi módszer direkt alkalmazása általában azt adta, hogy a tulajdonság, amelynek az egzisztenciáját bizonyítani akartuk, szinte biztosan bekövetkezik, addig a lokális lemma az események számában exponenciálisan kicsi esélyt ad csak erre. Ez egyrészt ígéretes az elméleti alkalmazások szempontjából, hiszen erősebb egzisztenciális tételekre számíthatunk a lokális lemma segítségével, azonban gyakorlati szempontból nem ad rögtön jól működő algoritmust egy konkrét konstrukció megtalálásához. A lokális lemma algoritmikus alkalmazási lehetőségeit a következő szakaszban tárgyaljuk.

4.1.3. Tétel. *Legyen a $H = (V, E)$ hipergráf olyan, hogy minden éle legalább k csúcsot tartalmaz, és minden él maximum d másik élt metsz. Ekkor, ha $e(d+1) < 2^{k-1}$, akkor a gráf csúcsainak van olyan kétszínezése, hogy nincs monokromatikus él.*

Bizonyítás. A gráf csúcsait színezzük egymástól függetlenül és egyenlő valószínűséggel kékre vagy pirosra. Legyen A_i az az esemény, hogy az i-edik éle a gráfnak monokromatikus. Ekkor éppen azt szeretnénk megmutatni, hogy $P(\bigcap_{i=1}^{n} \overline{A_i}) > 0$. Mivel minden él legalább k elemű, így $P(A_i) \leq 2^{1-k}$. Meg kell még határoznunk az összefüggőségek egy felső korlátját. A második feltétel miatt ez éppen d. A harmadik feltétel szerint $2^{1-k}e(d+1) < 1$, így a 4.1.2. Tétel feltételei teljesülnek, tehát $P(\bigcap_{i=1}^{n} \overline{A_i}) > 0$, és éppen ezt akartuk belátni. □

Egy hipergráf k-uniform, ha minden éle k csúcsot tartalmaz, k-reguláris, ha minden csúcsot k él tartalmaz. Egy k-uniform k-reguláris gráfban egy élt $k(k-1)$ másik metsz legfeljebb, mégpedig az adott él k darab csúcsán átmenő $k-1$ másik él. Az előző tétel szerint $e(k(k-1)+1) < 2^{k-1}$, azaz azt kapjuk, hogy $k \geq 9$ esetén nincs k-uniform k-reguláris hipergráf.

A lokális lemma második alkalmazásaként adjunk alsó korlátot a Ramsey-számokra!

A kiinduló gráfunk csúcsainak száma n. Legyen A_i az az esemény, hogy a gráf i-edik k csúcsú teljes részgráfja piros. Legyen B_j az az esemény, hogy a j-edik l csúcsú teljes részgráfja kék (összesen $\binom{n}{k} + \binom{n}{l}$ eseményt definiáltunk). Színezzük a gráf csúcsait egymástól függetlenül pirosra és kékre p, illetve $1-p$ valószínűséggel. Ekkor $P(A_i) = p^{-\binom{k}{2}}$. Hasonlóan $P(B_j) = (1-p)^{-\binom{l}{2}}$. A lokális lemma alkalmazásához a függőségi gráf csúcsaira meg kell adnunk egy súlyozást. A függőségi gráf A_i eseményeknek megfelelő csúcsait nevezzük A típusúnak, a többit B típusúnak. Legyen az A típusú csúcsok súlya x, a B típusú csúcsoké y. Becsüljük meg a függőségi gráfban a fokszámokat. Egy A típusú csúcs akkor függ egy másik csúcstól, ha a megfelelő részgráfoknak van közös éle. Az adott közös élt $\binom{k}{2}$-féleképpen választhatjuk ki. Így egy A típusú csúcs legfeljebb $\binom{k}{2}\binom{n-2}{k-2}$ A típusútól függhet, és legfeljebb $\binom{k}{2}\binom{n-2}{l-2}$ B típusútól. Hasonlóan meghatározható a B típusú csúcsok függősége. Így a lokális lemma szerint, ha van olyan $0 \leq p, x, y < 1$, hogy a

$$p^{-\binom{k}{2}} \leq x(1-x)^{\binom{k}{2}\binom{n-2}{k-2}}(1-y)^{\binom{k}{2}\binom{n-2}{l-2}},$$
$$(1-p)^{-\binom{l}{2}} \leq y(1-x)^{\binom{l}{2}\binom{n-2}{k-2}}(1-y)^{\binom{l}{2}\binom{n-2}{l-2}}$$

egyenlőtlenségek teljesülnek, akkor az n csúcsú gráf élei színezhetőek úgy, hogy ne legyen se teljes k csúcsú piros részgráf, se teljes l csúcsú kék részgráf. Innen következnek például az $R(k,3) > \frac{ck^2}{(\ln k)^2}$, $R(k,4) > k^{\frac{5}{2}+o(1)}$ becslések komoly mennyiségű számítás után.

Nézzük még meg a 4.1.2. Tétel segítségével, milyen alsó becslést kapunk $R(k,k)$-ra. Függetlenül színezzük az n csúcsú gráf csúcsait egyenlő valószínűséggel pirosra és kékre. A_i az az esemény, hogy az i-edik teljes k csúcsú részgráf monokromatikus. Ekkor $P(A_i) = 2^{1-\binom{k}{2}}$. A_i csak olyan A_j-től függhet, amivel van közös éle, így a függőségi gráf maximális fokszáma legfeljebb $\binom{k}{2}\binom{n-2}{k-2}$. Így a 4.1.2. Tétel szerint, ha $e(\binom{k}{2}\binom{n-2}{k-2}+1)2^{1-\binom{k}{2}} < 1$, akkor van olyan színezése a csúcsoknak, hogy nincs monokromatikus k csúcsú részgráf. Felhasználva az $\binom{n}{k} < (\frac{ne}{k})^k$ összefüggést, azt kapjuk, hogy $R(k,k) > \frac{k}{e} \cdot 2^{\frac{k+1}{2}}$. Látható, hogy ez az eredmény nem sokkal jobb, mint a direkt valószínűségi módszerrel kapott korlát, ennek az az oka, hogy a függőségi gráfunk fokszámai nagyok.

4.2. A lokális lemma algoritmikus szempontból

Az előző részben látott lokális lemma nem adott a kezünkbe gyakorlatban jól használható véletlen algoritmust, csupán egzisztencia-tételeket. A következőkben a hipergráf kétszínezési probléma egy speciális esetét vizsgáljuk meg algoritmikus szempontból.

Legyenek n és d rögzített egész számok. $\Omega = [m]$ a hipergráf csúcsainak alaphalmaza. Adottak $A_1, A_2, \ldots, A_N \subseteq \Omega$ a hipergráf élei, és a hipergráf n-uniform, azaz minden él n elemű. Tegyük még fel, hogy egyik él sem metsz több, mint d másik élt, és hogy $e(d+1) < 2^{n-1}$. A következő adatstruktúrával rendelkezünk: minden csúcshoz ismerjük az őt tartalmazó élhez indexhalmazát, illetve minden élnek ismerjük a benne lévő csúcsok indexhalmazát. Keressük a csúcsok egy olyan kétszínezését, hogy ne legyen monokromatikus él. A feladat méretét N határozza meg.

A fenti problémát Beck oldotta meg 1991-ben [11].

A 4.1.3. Tételt felhasználva tudjuk, hogy a keresett színezés létezik.

4.2.1. Tétel. *Ha n, d rögzítettek, $D = d(d-1)^3$, és létezik n_1, n_2, n_3, hogy $n = n_1 + n_2 + n_3$ és*

$$16D(1+d) < 2^{n_1},$$

$$16D(1+d) < 2^{n_2},$$

$$2e(1+d) < 2^{n_3},$$

akkor létezik véletlen algoritmus, amely megtalálja a fenti hipergráf egy nem élkromatikus színezését $O(N(\ln N)^c)$ időben, ahol $c > 0$ konstans.

Bizonyítás. Gondoljunk az $A_1, A_2, \dots, A_N$ élekre mint eseményekre, ahol A_i azt jelenti, hogy a neki megfelelő él monokromatikus, és jelöljük $A_1, A_2, \dots, A_n$ függőségi gráfját G-vel (azaz A_i és A_j akkor van összekötve G-ben, ha a megfelelő élek metszik egymást).

Az algoritmus lényege a következő: 3 lépésben színezzük a hipergráf csúcsait. Az egyes színezési szakaszok működését rendre n_1, n_2, n_3 szabályozza. Ügyes véletlen színezéssel nagy valószínűséggel elérhetjük, hogy az első két lépésben a probléma logaritmikusan egyszerűsödjön, így a végül kapott $O(\ln \ln N)$ méretű probléma már brutal force algoritmussal is gyorsan megoldható.

A következőkben egy esemény szinte biztos teljesülése alatt azt értjük, hogy az esemény valószínűsége 1-hez tart, ha $N \to \infty$.

Első színezési szakasz: Kezdetben egyik csúcsnak sincs színe. Vegyük a csúcsok egy véletlen sorrendjét, és kezdjük el ebben a sorrendben színezni a hipergráf csúcsait egyenlő valószínűséggel kékre, illetve pirosra. Minden csúcs kiszínezése után megnézzük, hogy nem keletkezett-e a gráfban olyan él, ami még monokromatikus (azaz a benne kiszínezett csúcsok csak kékek vagy csak pirosak), és amiben már n_1 csúcsot kiszíneztünk. Ha ilyen keletkezett, azt veszélyes élnek nevezzük, hiszen veszélyes, hogy monokromatikus lesz. Egy veszélyessé vált él maradék csúcsait az 1-es színezési szakszban már nem színezzük, ha ezek következnek színezésre, hagyjuk őket színezetlenül. Legyen S azon élek (véletlen) halmaza a hipergráfban, amelyek az 1-es színezési szakasz végén még monokromatikusak. Tekintsünk S-re, mint G részgráfjára. Ekkor a következő állítás igaz.

Állítás: S minden komponense szinte biztosan $O(\ln N)$ méretű.

Bizonyítás. Egy $T \subseteq V(G)$ részgráf 4-csoport, ha a következő feltételek teljesülnek: T bármely két csúcsának a távolsága G-ben legalább 4, és igaz, hogy ha csak a pontosan 4 távolságra lévő T-beli csúcsokat kötjük össze, összefüggő gráfot kapunk. Ezt a gráfot a T 4-csoport 4-hosszú élei által alkotott gráfnak nevezzük. Szeretnénk megmutatni, hogy szinte biztosan nincs $c_1 \ln N$-nél nagyobb 4-csoport.

Felülről becsüljük az s méretű 4-csoportok számát G-ben. Rögzítsünk egy s számozott csúcsú F fát, melynek csúcsai rendre $t_1, t_2, \ldots, t_s$, és $i > 1$ esetén létezik olyan $j < i$, hogy t_i és t_j össze vannak kötve a fában. Számoljuk meg, hány olyan 4-csoport van, amelynek a csúcsai számozhatók úgy, hogy a 4-csoport 4-hosszú élei által alkotott gráfnak F éppen feszítőfája legyen. Jó felső becslést kapunk, ha azt nézzük meg, F-et hányféleképpen tehetjük bele G-be. F t_1 csúcsát N-féle helyre tehetjük G-ben. A további csúcsok F már valamelyik kisebb indexű rögzített csúcsától legfeljebb 4 távolságra lehetnek. Mivel G-ben a maximális fokszám d, ez legfeljebb $D = d(d-1)^3$ lehetséges hely. Tehát adott F legfeljebb ND^{s-1} 4-csoportban lehet a 4-hosszú élekből alkotott gráf feszítőfája. Minden számozott fának van olyan átszámozása, mint amit F-től elvártunk. Ez például indukcióval következik, ha felhasználjuk, hogy mindig van 1 fokú csúcs. A Cayley-tétel szerint legfeljebb 4^s számozott fa van s csúcson. Így az eddigi eredményeink felhasználásával $ND^{s-1}4^s$-nel felülről becsülhetjük a 4-csoportok számát.

Most megbecsüljük annak a valószínűségét, hogy S tartalmaz legalább egy T 4-csoportot. Észrevehetjük, annak a valószínűsége, hogy A_i él veszélyessé válik az első színezési szakaszban, legfeljebb 2^{1-n_1}. Így egy s méretű független csúcshalmaz esetén annak a valószínűsége, hogy minden csúcs veszélyessé válik, legfeljebb $2^{s(1-n_1)}$. Ha van 4-csoport S-ben, akkor egy csúcsa és annak G-beli szomszédai mind függetlenek a 4-csoport egy másik csúcsától és annak szomszédaitól. Ez összesen $(d+1)^s$ független csúcshalmaz G-ben. Mivel $T \subseteq S$, a 4-csoport egyik csúcsa vagy annak valamelyik szomszédja az első színezési szakaszban veszélyessé vált. Tehát az előző $(d+1)^s$ független csúcshalmaz valamelyikének minden eleme veszélyes. Ennek az esélye a színezés elkezdésekor legfeljebb $2^{s(1-n_1)}(d+1)^s$ volt. Mivel egy esemény következményének valószínűsége nem kisebb az esemény valószínűségénél, azt kaptuk, hogy annak a valószínűsége, hogy S tartalmaz s méretű 4-csoportot, legfeljebb $2^{s(1-n_1)}(d+1)^s$.

Eddigi eredményeinket összevetve elmondhatjuk, hogy S-ben az s méretű 4-csoportok számának várható értéke legfeljebb

$$2^{s(1-n_1)}(d+1)^s ND^{s-1}4^s < N(8D(d+1)2^{-n_1})^s < N2^{-s}.$$

Az utolsó egyenlőtlenséget a tétel első feltételéből kaptuk. Tehát ha $s = c_1 \ln N$, akkor szinte biztos, hogy nincs 4-csoport S-ben, hiszen ezek számának várható értéke 0-hoz tart, amint $N \to \infty$ (c_1-et úgy választjuk, hogy $c_1 > 1/\ln 2$).

Innen S komponenseinek méretét már könnyen becsülhetjük. Ha egy adott C komponensben T egy maximális 4-csoport, akkor C minden csúcsa a 4-csoport valamelyik csúcsától maximum 3 távolságra van.

Mivel minden fokszám legfeljebb d, $|C| < |T|d^3$. Tudjuk, hogy szinte biztosan $|T| < c_1 \ln N$, tehát $|C| < c_2 \ln N$, és éppen ezt akartuk bebizonyítani. □

A probléma megoldását az első színezési szakasszal kezdjük. Ez várhatóan $O(N)$ időben lefut, és olyan színezést ad, hogy S-nek nincs $c_2 \ln N$-nél nagyobb összefüggőségi komponense. Az első színezési szakasz szinte biztosan sikeresen végződik, ha mégsem, kezdjük előlről az egész algoritmus futtatását. A sikeres első szakasz után rögzítjük a már színt kapott csúcsokat. A második színezési szakaszban már csak az S-beli élekkel, illetve az S-beli színtelen csúcsokkal foglalkozunk.

Minden S-beli A_i élben jelöljünk ki egy $n - n_1$ elemű, még színezetlen B_i csúcshalmazt. A második színezési szakaszban ugyanazt csináljuk, mint az első szakaszban, a B_i halmazokat tekintjük éleknek, és egy B_i-t akkor nevezünk veszélyesnek, ha még monokromatikus, és n_2 darab csúcsát már kiszíneztük. Itt is lehetséges, hogy egyes komponenseket esetleg többször kell színeznünk, de el fogjuk érni várhatóan lineáris időben, hogy a második színezés végén a még mindig monokromatikus élek S' gráfjában minden komponens maximum $O(\ln \ln N)$ méretű legyen, hiszen egy komponens újraszínezésének valószínűsége $o(1)$.

A harmadik színezési szakaszban minden $B_i \subseteq S'$ éleiben kijelölünk egy n_3 elemű, még színezetlen C_i csúcshalmazt. A tétel harmadik feltétele és a 4.1.3. Tétel szerint létezik C_i-beli csúcsoknak olyan színezése, hogy nem lesz monokromatikus C_j él. S' $O(\ln \ln N)$ méretű komponenseiben ezt a színezést brutal force algoritmussal keressük meg. A komponenseink száma $O(N)$, egy $O(\ln \ln N)$ méretű halmazon a brutal force futási ideje $O((\ln \ln N) \cdot 2^{n(\ln \ln N)})$. Így a harmadik színezési szakasz $O(N(\ln \ln N) \cdot 2^{n(\ln \ln N)})$ időben lefut. Mivel n rögzített, ez éppen $O(N(\ln N)^c)$-nagyságrendű futási idő, a tételt tehát beláttuk. $\square$

Hivatkozások

[1] P. Erdős (1947), Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.*, **53**, 292–294.

[2] M. Aigner, G. Ziegler (2009), *Proofs from THE BOOK* (4th ed.), Springer-Verlag, Berlin, New York.

[3] J. Komlós, J. Pintz, E. Szemerédi (1982), A lower bound for Heilbronn's problem, *J. London Math. Soc.*, **25**, 13–24.

[4] Lovász L. (2010), *Véletlen struktúrák és alkalmazásaik*, előadásjegyzet. http://www.cs.elte.hu/∼lovasz/random-2010.pdf

[5] K. Azuma (1967), Weighted Sums of Certain Dependent Random Variables, *Tohoku Math. Joun.*, **19**, 357–367.

[6] M. Talagrand (1995), Concentration of measure and isoperimetric inequalities in product spaces, *Publications I.H.E.S.*, **81**, 73–205.

[7] M. J. Steele (1997), *Probability theory and combinatorial optimization*, SIAM, Philadelphia.

[8] D. Aldous, M. J. Steele (1992), Asymptotics of Euclidean Minimal Spanning Trees on Random Samples, *Probab. Th. Rel. Fields*, **92**, 247–258.

[9] P. Erdős, L. Lovász (1975), Problems and results on 3-chromatic hypergraphs and some related questions, In: A. Hajnal et al. (eds), *Infinite and Finite Sets*, North Holland, Amsterdam, 609–628.

[10] N. Alon and J. H. Spencer (1992), *The Probabilistic Method*, Wiley, New York.

[11] Beck J. (1991), An algorithmic approach to the Lovász local lemma, I, *Random Structures* and *Algorithms*, **2**, 343–365.