6. Probabilistinen menetelmä

- ⋆ Probabilistista menetelmää sovelletaan olioiden olemassaolon todistamiseen
- ⋆ Jotkin ominaisuudet omaavan olion olemassaolo voidaan todeta jos voimme osoittaa olioiden perusjoukon, jossa satunnaisesti vedetyllä oliolla on vaaditut ominaisuudet positiivisella tn.:llä
- ⋆ Koska vaaditut ominaisuudet omaavan olion valintatn. on positiivinen, niin perusjoukossa täytyy olla tuollainen olio
- ⋆ Näin ollen vaaditut ominaisuudet omaava olio on olemassa
- ⋆ Usein saadaan satunnaisalgortimi halutun olion kostruoimiseksi ja joskus se voidaan determinisoida

Laskemismenetelmä

- ⋆ Merk. n:n solmun *täydellistä* — kaikki $\binom{n}{2} = n(n-1)/2$ kaarta sisältävää — verkkoa K_n:llä
- ⋆ k:n solmun klikki verkossa K_n on täydellinen verkko K_k

Lause 6.1: *Jos* $\binom{n}{2}2^{-\binom{k}{2}+1} < 1$*, niin verkon* K_n *kaaret voidaan värittää kahdella värillä s.e. se ei sisällä yksiväristä* k*-klikkiä* K_k*.*

Todistus. Mahdollisia kaarten värityksiä on $2^{n(n-1)/2}$ kpl, joten yksittäisen värityksen tn. tasaisen jakauman mukaan on $2^{-n(n-1)/2}$. Samaan jakaumaan päädytään valitsemalla kunkin kaaren väri riippumattomasti s.e. molempien värivaihtoehtojen tn. on $1/2$.

Mahdollisia k-klikkejä verkossa K_n on $\binom{n}{k}$, kiinnitetään niille jokin järjestys.

Olk. A_i tapahtuma "klikki i on yksivärinen". Klikin ensinnä väritettävä kaari voi olla kumpaa väriä vain, loppujen $\binom{k}{2} - 1$ kaaren tulee olla sen kanssa samanvärisiä, joten $\Pr(A_i) = 2^{-\binom{k}{2}+1}$. Yhdisteen tn.:n perusteella

$$\Pr\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) \le \sum_{i=1}^{\binom{n}{k}} \Pr(A_i) = \binom{n}{k} 2^{-\binom{k}{2}+1},$$

joka lauseen oletuksen perusteella on aidosti yhtä pienempi. Näin ollen

$$\Pr\left(\bigcap_{i=1}^{\binom{n}{k}} \bar{A}_i\right) = 1 - \Pr\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) > 0.$$

Koska sellaisen värityksen, joka ei sisällä yksiväristä k-klikkiä, valitsemisen tn. on aidosti positiivinen, niin tällainen väritys on olemassa. □

- ⋆ Ed. lauseen ehto pätee esim. kun $n \le 2^{k/2}$ ja $k \ge 3$, sillä tällöin

$$\begin{aligned}\binom{n}{k}2^{-\binom{k}{2}+1} &\le \frac{n^k}{k!}2^{-(k(k-1)/2)+1} \\ &\le \frac{2^{k/2+1}}{k!} < 1\end{aligned}$$

- ⋆ Esim. $1000 \le 2^{10}$, joten K_{1000} voidaan värittää kahdella värillä s.e. siinä ei ole yksiväristä klikkiä K_{20}
- ⋆ Valitsemalla kunkin kaaren väri symmetrisesti ja riippumattomasti saadaan Monte Carlo -algoritmi vaaditun värityksen löytämiseksi
- ⋆ Todistuksen arvion perusteella tässä tapauksessa onnistumistn. on ainakin

$$1 - \frac{2^{20/2+1}}{20!} \ge 1 - 8,5 \cdot 10^{-16}$$

- ⋆ Edellä verkkojen perusjoukosta otanta on helppoa, mutta aina tehokas otanta ei ole mahdollista
- ⋆ Jos tehokas otanta on mahdollista ja onnistumistn. kullakin otospisteellä on p, niin tarvittavien toistojen lkm on geometrisesti jakautunut s-muuttuja, jonka odotusarvo on $1/p$
- ⋆ Täten $(1/p)$:n tulisi olla syötteen pituuden suhteen polynominen
- ⋆ Las Vegas -algoritmia halutut ominaisuudet omaavan olion konstruoimiseksi varten tarvitaan vielä polynomiaikainen tunnistaja etsitylle oliolle
- ⋆ Tällöin otospisteitä voidaan testata kunnes vaaditut ominaisuudet omaava otospiste tulee vedettyä

Odotusarvomenetelmä

- ⋆ Jos diskreetillä s-muuttujalla on positiivinen tn., niin sen täytyy saada ainakin yksi korkeintaan ja yksi vähintään odotusarvon suuruinen arvo

Lemma 6.2: *Olk.* $\mathbf{E}[X] = \mu$, *tällöin* $\Pr(X \geq \mu) > 0$ *ja* $\Pr(X \leq \mu) > 0$.

Todistus. Jos $\Pr(X \geq \mu) = 0$, niin

$$\begin{aligned} \mathbf{E}[X] &= \sum_x x \Pr(X = x) \\ &= \sum_{x<\mu} x \Pr(X = x) \\ &< \sum_{x<\mu} \mu \Pr(X = x) = \mu, \end{aligned}$$

joka on ristiriita.

Tapaus $X \leq \mu$ todetaan vastaavasti. □

MAKSIMILEIKKAUS

- ⋆ NP-kovassa maksimileikkaus-ongelmassa verkon solmut jaetaan kahteen pistevieraaseen joukkoon
- ⋆ Olk. solmujoukosta toiseen menevät kaaret *leikkauskaaret*
- ⋆ Suuntaamattoman verkon leikkauksen arvo on leikkauskaarten yhteenlaskettu paino
- ⋆ Seur. kunkin kaaren paino on aina 1
- ⋆ Seur. tulos osoittaa, että maksimileikkaus on ainakin puolet verkon kaarista

Lause 6.3: *Olk.* $G = (V, E)$ *suuntaamaton verkko, jossa on* n *solmua ja* m *kaarta. Solmujoukko* V *voidaan jakaa kahteen erilliseen joukkoon* A *ja* B *s.e. ainakin* $m/2$ *kaarta kulkee* A*:n solmusta* B*:n solmuun.*

Todistus. Muodostetaan joukot A ja B sijoittamalla kukin solmu riippumattomasti ja satunnaisesti toiseen joukoista.

Kiinnitetään verkon kaarille jokin numerointi. Olk. s-muuttujat X_i s.e.

$$X_i = \begin{cases} 1, & \text{jos kaari } i \text{ kulkee } A\text{:sta } B\text{:hen} \\ 0, & \text{muuten.} \end{cases}$$

Tn., että kaari $e_i \in E$ kulkee joukon A solmusta joukon B solmuun on $1/2$, joten $\mathbf{E}[X_i] = 1/2$.

Joukkojen A ja B määräämän leikkauksen arvon $C_{A,B}$ odotusarvo on lineaarisuuden perusteella

$$\mathbf{E}[C_{A,B}] = \sum_{i=1}^{m} \mathbf{E}[X_i] = m \cdot \frac{1}{2} = \frac{m}{2}.$$

Lemman 6.2 perusteella on olemassa G:n leikkaus, jonka arvo on vähintään $m/2$. □

- ⋆ Solmujen jako joukkoihin A ja B on helppo vetää
- ⋆ Odotusarvomenetelmä ei anna alarajaa tn.:lle $p = \Pr(C_{A,B} \geq m/2)$

$$
\begin{aligned}
\frac{m}{2} &= \mathbf{E}[C_{A,B}] \\
&= \sum_{i \leq m/2-1} i\Pr(C_{A,B} = i) \\
&\quad + \sum_{i \geq m/2} i\Pr(C_{A,B} = i) \\
&\leq (1-p)\left(\frac{m}{2} - 1\right) + pm
\end{aligned}
$$

- ⋆ Täten $p \geq 1/(m/2+1)$, joten vetämällä odotusarvoisesti $m/2+1$ jakoa joukkoihin löydetään leikkaus, jonka arvo on väh. $m/2$
- ⋆ On helppo testata täyttääkö leikkaus ehdon, joten saamme LV-algoritmin leikkauksen löytämiseksi

MAKSIMITOTEUTUVUUS (MAXSAT)

- ⋆ Tiedämme lausekalkyylin kaavojen toteutumisongelman (SAT) olevan NP-täydellisen
- ⋆ Olk. tarkastellut kaavat CNF-muotoisia, s.o. tekijöiden (clause) konjunktioita, missä kukin tekijä on literaalien disjunktio
- ⋆ MAXSAT-ongelmassa on annettu SAT-kaava ja pyrkimyksenä on toteuttaa mahdollisimman monta sen tekijöistä
- ⋆ Oletamme, ettei mikään tekijä sisällä sekä muuttujaa että sen negaatiota

Lause 6.4: *Olk. annettu m:n tekijän joukko. Olk. k_i tekijän i, $i = 1, \ldots, m$, literaalien lkm ja $k = \min k_i$. On olemassa Boolen-muuttujien arvoasetus, jonka toteuttamien*

tekijöiden lkm on ainakin

$$\sum_{i=1}^{m}(1 - 2^{-k_i}) \geq m(1 - 2^{-k}).$$

Todistus. Asetetaan muuttujille arvot riippumattomasti ja symmetrisesti. Tn., että is tekijä toteutuu on väh. $1 - 2^{-k_i}$. Toteutuvien tekijöiden lkm on siten väh.

$$\sum_{i=1}^{m}(1 - 2^{-k_i}) \geq m(1 - 2^{-k}),$$

joten lemman 6.2 perusteella on olemassa arvoasetus, joka toteuttaa ainakin näin monta tekijää. □

- ⋆ Näin saamme helposti satunnaisalgoritmin MAXSAT-ongelmalle

Derandomisointi

- ⋆ Sovelletaan ehdollisten odotusarvojen menetelmää maksimileikkausalgoritmin determinisoimiseksi
- ⋆ Sijoitetaan verkon solmut deterministisesti yksi kerrallaan joukkoihin A ja B mv. järjestyksessä $v_1, \ldots, v_n$
- ⋆ Olk. $x_i \in \{A, B\}$ se joukko, johon v_i sijoitetaan
- ⋆ Tarkastellaan leikkauksen odotusarvoa $\mathbf{E}[C_{A,B} \mid x_1, \ldots, x_k]$ kun ensimmäiset k solmua on sijoitettu ja jäljellä olevat sijoitetaan riippumattomasti tasaisen jakauman mukaan
- ⋆ Induktiolla osoitamme kuinka sijoittaa v_{k+1} s.e.

$$\mathbf{E}[C_{A,B} \mid x_1, \ldots, x_k] \leq \mathbf{E}[C_{A,B} \mid x_1, \ldots, x_{k+1}]$$

- ⋆ Induktiosta seuraa lopulta

$$\frac{m}{2} \leq \mathbf{E}[C_{A,B}] \leq \mathbf{E}[C_{A,B} \mid x_1, \ldots, x_n],$$

missä oikea puoli on deterministisen sijoittelualgoritmin valitseman leikkauksen arvo

- ⋆ Induktion perustapaus on $\mathbf{E}[C_{A,B}] = \mathbf{E}[C_{A,B} \mid x_1]$, joka pätee, koska ensimmäisen solmun sijoittamisella ei ole merkitystä
- ⋆ Tarkastellaan v_{k+1}:n satunnaista ja symmetristä sijoittamista joukkoon A tai B
- ⋆ Olk. Y_{k+1} valittua joukkoa kuvaava s-muuttuja. Nyt $\mathbf{E}[C_{A,B} \mid x_1, \ldots, x_k]$

$$= \frac{1}{2}\mathbf{E}[C_{A,B} \mid x_1, \ldots, x_k, Y_{k+1} = A] + \frac{1}{2}\mathbf{E}[C_{A,B} \mid x_1, \ldots, x_k, Y_{k+1} = B]$$

- ⋆ Täten

$$\max\{\, \mathbf{E}[C_{A,B} \mid x_1, \ldots, x_k, Y_{k+1} = X] \mid X \in \{\, A, B \,\}\,\} \geq \mathbf{E}[C_{A,B} \mid x_1, \ldots, x_k]$$

- ⋆ Riittää siis laskea ehdolliset odotusarvot, joissa Y_{k+1}:n valinta on mukana ja sijoittaa v_{k+1} paremman odotusarvon antavaan joukkoon
- ⋆ Tämän jälkeen ind.askel on osoitettu
- ⋆ Ehd. odotusarvoissa ensimmäisten $k+1$ solmun joukot on kiinnitetty
- ⋆ Voimme siis laskea moniko niiden välisistä kaarista tulee leikkauksen arvoon laskettua
- ⋆ kaikille lopuille kaarille pätee, että tn.:llä $1/2$ sen päätepisteet päätyvät eri joukkoon ja kaari tulee leikkauksen arvoon mukaan laskettavaksi

- ⋆ Odotusarvon lineaarisuuden perusteella ehd. odotusarvot siis ovat havaitut kaaret joukkojen välillä plus puolet jäljelle jäävistä kaarista
- ⋆ Ehd. odotusarvot voidaan helposti laskea lineaarisessa ajassa
- ⋆ Itse asiassa kahden ehd. odotusarvon järjestys määräytyy sen mukaan kummassa joukossa v_{k+1}:lla on enemmän naapureita
- ⋆ Deterministinen algoritmi:
  - ◇ Sijoita ensimmäinen solmu (esim.) joukkoon A
  - ◇ Sijoita seuraavat solmut siihen joukkoon, jossa niillä on vähemmän naapureita (tasatilanteet mv.)
- ⋆ Tämän algoritmin tuottaman leikkauksen arvo on vähintään $m/2$

Valitse ja muokkaa

- ⋆ Nyt vedetään satunnainen olio ja muokataan sitä s.e. sillä on halutut ominaisuudet
- ⋆ Tarkastellaan ensin tunnetusti NP-kovaa riippumaton joukko -ongelmaa
- ⋆ Verkon G *riippumaton joukko* (IS) on sellaisten solmujen joukko, joiden välillä ei ole laisinkaan kaaria

Lause 6.5: *Olk.* $G = (V, E)$ *verkko, jossa on* n *solmua ja* m *kaarta.* G*:ssä on riippumaton joukko, jossa on ainakin* $n^2/4m$ *solmua.*

Todistus. Olk. G:n solmujen keskim. aste $d = 2m/n$. Satunnaisalgoritmi:

1. Riippumattomasti tn.:llä $1 - 1/d$ poista kukin solmuista yhdessä siihen liittyvien kaarten kanssa.
2. Poista kukin jäljelle jäänyt kaari ja

toinen siihen liittyvistä solmuista.

Jäljelle jäävät solmut muodostavat IS:n, koska kaikki kaaret on poistettu.

Olk. X askelen 1 jälkeen jäljellä olevien solmujen lkm. Lineaarisuuden perusteella $\mathbf{E}[X] = n/d$.

Olk. Y askelen 1 jälkeen jäljellä olevien kaarten lkm. Lähtötilanteessa kaaria on $nd/2$ kpl ja ne jäävät jäljelle, joiden molemmat päät selviävät askelesta:

$$\mathbf{E}[Y] = \frac{nd}{2}\left(\frac{1}{d}\right)^2 = \frac{n}{2d}.$$

Askelessa 2 poistetaan kaikki jäljelle jääneet kaaret ja kork. Y solmua. Lopulta jäljelle jääviä solmuja on siis väh. $X - Y$, ja

$$\mathbf{E}[X - Y] = \frac{n}{d} - \frac{n}{2d} = \frac{n}{2d} = \frac{n^2}{4m},$$

koska $d = 2m/n$. □

VERKON YMPÄRYSMITTA

- ⋆ Verkon *ympärysmitta* (girth) on sen lyhimmän syklin pituus
- ⋆ Tiheillä verkoilla pitäisi intuitiivisesti olla pieni ympärysmitta
- ⋆ Voidaan kuitenkin osoittaa, että tiheälläkin verkolla voi olla melko suuri ympärysmitta

Lause 6.6: *Millä tahansa $k \geq 3$ on olemassa verkko, jossa on n solmua ja väh. $n^{1+1/k}/4$ kaarta, jonka ympärysmitta on väh. k.*

Todistus. Vedetään ensin verkko $G \in G_{n,p}$ parametrilla $p = n^{1/k-1}$. Olk. X verkon kaarten lkm. Tällöin

$$\mathbf{E}[X] = p\binom{n}{2} = \frac{1}{2}\left(1 - \frac{1}{n}\right)n^{1/k+1}.$$

Olk. Y verkon kork. $(k-1)$-pituisten syklien lkm. Tietty i-sykli on verkossa tn.:llä p^i.

Valitsemalla i solmua ja niille järjestys (+ käänteinen sykli on sama), nähdään että i-syklejä on kaikkiaan $\binom{n}{i}(i-1)!/2$ kpl.

$$\begin{aligned}\mathbf{E}[Y] &= \sum_{i=3}^{k-1}\binom{n}{i}\frac{(i-1)!}{2}p^i \leq \sum_{i=3}^{k-1} n^i p^i \\ &= \sum_{i=3}^{k-1} n^{i/k} < kn^{(k-1)/k}.\end{aligned}$$

Poistetaan satunnaisesti valitun verkon G jokaisesta kork. $(k-1)$-pituisesta syklistä yksi kaari. Näin muodostuvan verkon ympärysmitta on siis väh. k. Kun n on riittävän suuri, niin kaarten lkm:n odotusarvo muodostuvassa verkossa on

$$\begin{aligned}\mathbf{E}[X - Y] &\geq \frac{1}{2}\left(1 - \frac{1}{n}\right)n^{1/k+1} - kn^{(k-1)/k} \\ &\geq \frac{1}{4}n^{1/k+1}.\end{aligned}$$

Lemman 6.2 perusteella on siis olemassa väitetyn kaltaisia verkkoja. □

Toisen momentin menetelmä

- ⋆ Tšebyševin epäyhtälöllä saadaan

Lause 6.7: *Jos X on ei-negatiivisia kokonaislukuarvoja saava s-muuttuja, niin*

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

Todistus.

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

□

- ⋆ Tarkastellaan nyt satunnaisverkkojen kynnyskäyttäytymistä
- ⋆ Usein mallin $G_{n,p}$ verkoilla on kynnysfunktio f s.e.
  - ⋄ $p < f(n)$: lähes millään verkolla ei ole kaivattua ominaisuutta
  - ⋄ $p > f(n)$: lähes kaikilla verkoilla on kaivattu ominaisuus

Lause 6.8:

1. *Mallissa $G_{n,p}$ olk. $p = f(n)$, missä $f(n) = o(n^{-2/3})$. Tällöin kaikilla $\varepsilon > 0$ ja riittävän suurilla n tn., että satunnaisesti mallista $G_{n,p}$ valitussa verkossa on väh. neljän solmun klikki, on kork. ε.*
2. *Vastaavasti jos $f(n) = \omega(n^{-2/3})$, niin riittävän suurilla n tn., että satunnaisesti mallista $G_{n,p}$ valitussa verkossa ei ole väh. neljän solmun klikkiä, on kork. ε.*

Todistus.

1. Asetetaan jokin numerointi kaikille $k = \binom{n}{4}$ neljän solmun joukoille C_i ja otetaan s-muuttujat

$$X_i = \begin{cases} 1, & \text{jos } C_i \text{ on 4-klikki} \\ 0, & \text{muuten.} \end{cases}$$

Olk. $X = \sum_{i=1}^{k} X_i$. Koska $\Pr(X_i = 1) = p^6$, niin

$$\begin{aligned} \mathbf{E}[X] &= \binom{n}{4} p^6 = \Theta(n^4 p^6) \\ &= \Theta((pn^{2/3})^6) = o(1), \end{aligned}$$

sillä $p = o(n^{-2/3})$. Näin ollen, riittävän suurilla n, $\mathbf{E}[X] < \varepsilon$. Koska X on ei-negatiivisia kokonaislukuarvoja saava s-muuttuja, niin $\Pr(X \geq 1) \leq \mathbf{E}[X] < \varepsilon$. Täten tn., että satunnaisesti $G_{n,p}$:stä valitussa verkossa on väh. neljän solmun klikki on kork. ε.

2. Nyt $\lim_{n\to\infty} \mathbf{E}[X] = \infty$. Tämä ei kuitenkaan vielä yksin riitä väitteen todistamiseen. Lauseen 6.7 perusteella voimme kuitenkin osoittaa, että $\Pr(X = 0) = o(1)$ tässä

tapauksessa. Tämä edellyttää, että $\mathbf{Var}[X] = o(\mathbf{E}[X]^2)$.

Lauseen 6.7 soveltamista varten tarvitsemme varianssia. Seuraava lemma voidaan todistaa lauseen 3.2 yleistyksen avulla.

Lemma 6.9: *Olk. Y_i:t indikaattoreita ja $Y = \sum_i Y_i$. Tällöin*

$$\mathbf{Var}[Y] \leq \mathbf{E}[Y] + \sum_{i \neq j} \mathbf{Cov}(Y_i, Y_j)\,.$$

Laskeaksemme $\mathbf{Var}[X]$:n lemman 6.9 perusteella, meidän tulee huomioida kovarianssit. Kun $C_i \cap C_j = \emptyset$, niin klikit ovat erillisiä ja X_i ja X_j ovat riippumattomia. Täten $\mathbf{Cov}(X_i, X_j) = 0$. Sama pätee kun $|C_i \cap C_j| = 1$.

Kun $|C_i \cap C_j| = 2$, niin klikeillä on yksi yhteinen kaari. Jotta molemmat klikit

voivat olla verkossa yhtä aikaa, niin tarvittavien $6 + 6 - 1 = 11$ kaaren tulee olla verkossa. Täten

$$\begin{aligned} \mathbf{Cov}(X_i, X_j) &= \mathbf{E}[X_i X_j] - \mathbf{E}[X_i]\,\mathbf{E}[X_j] \\ &\leq \mathbf{E}[X_i X_j] \leq p^{11}. \end{aligned}$$

Tarvittavat kuusi solmua pitää valita s.e. kaksi valitaan klikkien yhteisiksi ja kaksi kummankin klikin omaksi, joten vaihtoehtoja on

$$\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2} = \Theta(n^6).$$

Kun $|C_i \cap C_j| = 3$, niin klikeillä on kolme yhteistä kaarta ja kaikkiaan 9:n kaaren tulee olla verkossa. Nyt

$$\mathbf{Cov}(X_i, X_j) \leq \mathbf{E}[X_i X_j] \leq p^9.$$

Solmuvaihtoehtoja on nyt

$$n(n-1)\binom{n-2}{3} = \Theta(n^5).$$

Muistetaan, että $\mathbf{E}[X] = \binom{n}{4}p^6 = \Theta(n^4p^6)$, joten $\mathbf{E}[X]^2 = \Theta(n^8p^{12})$. Varianssi $\mathbf{Var}[X]$ puolestaan on kork.

$$\begin{aligned} &\leq \mathbf{E}[X] + \sum_{i \neq j} \mathbf{Cov}(X_i, X_j) \\ &\leq \Theta(n^4p^6) + \Theta(n^6p^{11}) + \Theta(n^5p^9) \\ &= o(n^8p^{12}), \end{aligned}$$

sillä esim.

$$\frac{n^4p^6}{n^8p^{12}} = \frac{1}{n^4p^6} = o\left(\frac{1}{n^4(n^{-2/3})^6}\right) = o(1).$$

Täten siis $\mathbf{Var}[X] = o(\mathbf{E}[X]^2)$, joten kohta 2 seuraa nyt lauseesta 6.7. □

- ⋆ Lauseen 6.8 kohdan 2 todistus olisi voinut perustua myös seuraavaan tulokseen, joka pätee vaikka muuttujat X_i eivät olisi riippumattomia

Lause 6.10: *Olk.* $X = \sum_{i=1}^{n} X_i$, *missä kukin* X_i *on indikaattorimuuttuja. Tällöin*

$$\Pr(X > 0) \geq \sum_{i=1}^{n} \frac{\Pr(X_i = 1)}{\mathbf{E}[X \mid X_i = 1]}.$$

- ⋆ Todistuksen merkinnöin $X = \sum_{i=1}^{k} X_i$, missä $X_i = 1$ vain jos neljän solmun joukko C_i on 4-klikki
- ⋆ $\Pr(X_j = 1) = p^6$
- ⋆ Odotusarvon lineaarisuuden perust. $\mathbf{E}[X \mid X_j = 1] = \sum_{i=1}^{k} \mathbf{E}[X_i \mid X_j = 1] = \sum_{i=1}^{k} \Pr(X_i \mid X_j = 1)$, koska X_i on indikaattori

- ⋆ Neljän solmun joukkoja C_i, jotka eivät leikkaa joukkoa C_j on $\binom{n-4}{4}$ kpl ja kullakin tn., että $X_i = 1$ on p^6
- ⋆ Joukkoja, joilla $|C_i \cap C_j| = 1$ on $\binom{4}{1}\binom{n-4}{3} = 4\binom{n-4}{3}$ kpl ja niilläkin tn., että $X_i = 1$ on p^6
- ⋆ Joukkoja, joilla $|C_i \cap C_j| = 2$ on $\binom{4}{2}\binom{n-4}{2} = 6\binom{n-4}{2}$ kpl ja kullakin tn., että $X_i = 1$ on p^5, jne.
- ⋆ Kaikkiaan $\mathbf{E}[X \mid X_j = 1] = \binom{n-4}{4}p^6 + 4\binom{n-4}{3}p^6 + 6\binom{n-4}{2}p^5 + 4\binom{n-4}{1}p^3 + 1$
- ⋆ Lauseen 6.10 perusteella

$$\Pr(X > 0) \geq \frac{\binom{n}{4}p^6}{\mathbf{E}[X \mid X_j = 1]}$$

- ⋆ $\lim_{n\to\infty} \Pr(X > 0) = 1$