Erdős-Rényijev model slučajnih grafova

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Monika Bajcer

ERDŐS-RÉNYIJEV MODEL SLUČAJNIH GRAFOVA

Diplomski rad

Voditelj rada:
Doc.dr.sc. Vjekoslav Kovač

Zagreb, studeni, 2016.

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

Uvod

Polovicom prošlog stoljeća poznati matematičari Paul Erdős i Alfred Rényi dali su koncept slučajnog grafa i počeli s proučavanjem njegovih svojstava te tako probudili interes za tom temom koji traje još i danas. Slučajni grafovi savršen su primjer dobre matematičke teorije—jednostavno su definirani, a imaju zanimljivu strukturu, koja kao rezultat daje primjene u mnogim područjima, od kojih je možda najvažnija njihova primjena u računarstvu. Spomenuti model se definira kao vjerojatnosni prostor nad skupom svih grafova s fiksiranim skupom vrhova, pri čemu se bridovi u grafu pojavljuju nezavisno s vjerojatnosti p. Cilj ovog rada je upoznati čitatelja s definicijom i nekim svojstvima tog vjerojatnosnog prostora.

Ovaj rad sastoji se od četiri poglavlja. U prvom poglavlju navodimo osnovne definicije i rezultate iz teorije grafova i vjerojatnosti koji se ne tiču direktno slučajnih grafova, ali će nam poslužiti kao pomoć tijekom njihovog proučavanja. Poglavlje je podijeljeno na tri manja odjeljka. U prvom odjeljku dajemo osnovne definicije vezane uz determinističke grafove, a u drugom rezultate iz teorije vjerojatnosti. Osobito će nam biti važne Poissonova i binomna distribucija te njihova veza. Treći odjeljak daje samo kratak opis jednostavnog procesa grananja, kojeg ćemo kasnije usporediti s procesom grananja kod grafova.

U drugom poglavlju uvodimo preciznu definiciju Erdős-Rényijevog modela $G(n, e)$, ali i modificiranu verziju tog modela $G(n, p)$ koju je ponudio Edgar Gilbert i na koju ćemo se oslanjati. Promatrat ćemo monotona svojstva koja ostaju vrijediti ako determinističkom grafu dodamo nove bridove, kao što je postojanje trokuta u grafu. Svi naši rezultati bit će asimptotski, a za vjerojatnost pojavljivanja brida slučajnog grafa $p = p_n$ dozvoljavamo da bude funkcija koja ovisi o n. Važnu ulogu imat će prijelazna funkcija $r(n)$ monotonog svojstva, koja je zanimljiva zbog toga što je za $p_n \gg r(n)$ vjerojatnost tog svojstva asimptotski jednaka 1, dok je za $p_n \ll r(n)$ ona asimptotski jednaka 0. Također, u ovom poglavlju dokazujemo da za $p_n = p$ fiksiran i bilo koje svojstvo grafa koje se može izraziti logikom prvog reda, vrijedi zakon nula-jedan.

U trećem poglavlju uvodimo definiciju klike u grafu te proučavamo vjerojatnost pojavljivanja klike veličine $k = k(n)$. Pokazat ćemo da postoje funkcije $k(n)$ za koje je aimptotska vjerojatnost pojavljivanja klike te veličine jednaka 1, ali i da postoje funkcije $k(n)$ za koje je ta vjerojatnost 0. Na kraju poglavlja dajemo i primjere prijelaznih funkcija za neka svojstva vezana uz postojanje unaprijed zadanih podgrafova.

Konačno, u četvrtom poglavlju dolazimo do Erdős-Rényijeve promjene faze, koja je izazvala znatiželju mnogih matematičara. Radi se o vrijednosti od p_n oko koje se događaju značajne promjene u veličini najveće komponente grafa. Tako će u slučajnom grafu $G(n, p_n)$ za vrijednosti p_n manje od $\frac{1}{n}$ vrijediti da su sve komponente u grafu jednake veličine, reda najviše $\ln n$, dok će se za vrijednosti p_n koje su veće od $\frac{1}{n}$ pojaviti komponenta koju ćemo zvati divovskom i koja će biti reda n.

Veoma sam zahvalna voditelju rada doc. dr. sc. Vjekoslavu Kovaču na strpljenju i pomoći te na svim uočenim greškama i komentarima. Također, svojim roditeljima dugujem posebnu zahvalnost zbog razumijevanja i podrške tijekom cijelog školovanja.

Poglavlje 1

Preliminarni rezultati

U ovom poglavlju navodimo pojmove i rezultate na koje ćemo se referirati u ostatku rada. Za nizove pozitivnih brojeva $(a_n)_{n\in\mathbb{N}}$ i $(b_n)_{n\in\mathbb{N}}$ definiramo:

- $a_n = O(b_n)$ ako vrijedi $a_n \leq cb_n$, za $c > 0$ i sve dovoljno velike n,
- $a_n = \Omega(b_n)$ ako vrijedi $b_n = O(a_n)$,
- $a_n = \Theta(b_n)$ ako vrijedi $a_n = O(b_n)$ i $a_n = \Omega(b_n)$,
- $a_n = o(b_n)$ ako vrijedi $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$,
- $a_n \sim b_n$ ako vrijedi $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$,
- $a_n \ll b_n$ ako vrijedi $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$,
- $a_n \gg b_n$ ako vrijedi $\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$.

1.1 Osnovni pojmovi teorije grafova

Ponovimo ukratko osnovne pojmove vezane uz determinističke grafove. *Neusmjereni graf* ili kratko samo *graf* je uređeni par $G = (V, E)$, gdje je $\emptyset \neq V = V(G)$ skup vrhova, $E = E(G)$ skup bridova, a svaki brid $e \in E$ spaja dva vrha, $i, j \in V$, koji se zovu krajevi od e. Kažemo još da su tada vrhovi i i j *incidentni* s e, a vrhovi i i j *susjedni* i pišemo $i \sim j$ ili $e = \{i, j\}$. Katkada se ova definicija proširuje tako da se dozvole *petlje* (bridovi koji spajaju vrh sa samim sobom) i *višestruki bridovi* (više bridova između para vrhova). Graf je *jednostavan* ako ne sadrži višestruke bridove i petlje. Kažemo da je G *prazan graf* ako je $E(G) = \emptyset$. Jednostavan graf u kojem je svaki par vrhova spojen svojim bridom zove se *potpun graf*.

Očito je da jedan graf možemo prikazati na različite načine. Stoga definiramo kada dva grafa smatramo istim, tj. izomorfnim. Kažemo da su grafovi G i H *izomorfni*, i pišemo $G \approx H$, ako postoje bijekcije $\theta : V(G) \to V(H)$ i $\rho : E(G) \to E(H)$ takve da je vrh v incidentan s bridom e u G ako i samo ako je $\theta(v)$ incidentan s $\rho(e)$ u H. Uređeni par $f = (\theta, \rho)$ se tada zove *izomorfizam* iz G u H. *Automorfizam* grafa G je izomorfizam iz G u G.

Neka su G i G' grafovi. Ako je $V(G') \subseteq V(G)$ i $E(G') \subseteq E(G)$, a svaki brid iz G' ima iste krajeve u G' kao što ih ima u G, onda kažemo da je G' *podgraf* od G i pišemo $G' \subseteq G$, a G zovemo *nadgraf* od G'. Dvije vrste podgrafa su naročito važne. *Inducirani podgraf* grafa G induciran skupom V' je podgraf G' sa skupom vrhova V' i skupom bridova E', gdje se E' sastoji od svih bridova od G čija oba kraja leže u V', dok je V' podskup od V. *Razapinjujući podgraf* G' je podgraf sa skupom vrhova V i skupom bridova E'.

Put u grafu G je niz $(v_o, e_1, v_1, e_2, v_2, \dots, e_n, v_n)$ gdje je e_i brid $\{v_{i-1}, v_i\}$ za $i = 1, \dots, n$, a svi vrhovi v_i su različiti. Definiramo relaciju ekvivalencije $\equiv$ na skupu vrhova V grafa G: $x \equiv y$ ako postoji put od x do y. Ova relacija ekvivalencije definira jednu particiju skupa V, te definiramo *komponente povezanosti*, ili kraće, *komponente* grafa G kao podgrafove inducirane klasama ekvivalencije. Kažemo da je graf *povezan* ako postoji samo jedna komponenta, a povezan graf bez ciklusa zovemo *stablo*. Za grafove $G_1 = (V_1, E_1)$ i $G_2 = (V_2, E_2)$ definiramo graf $G = G_1 \cup G_2$ sa skupom vrhova $V = V_1 \cup V_2$ i skupom bridova $E = E_1 \cup E_2$. Graf $G = (V, E)$ zovemo *unijom grafova* G_1 i G_2.

Sljedeći teorem daje nam važan rezultat teorije grafova čiji se dokaz može pronaći u članku P. Shora, [10].

Teorem 1.1.1. *(Cayleyjeva formula) Za $n \in \mathbb{N}$ i $V = \{1, 2, \dots, n\}$ postoji n^{n-2} stabala čiji je skup vrhova V.*

Osim toga, često će se javiti potreba za približnom ocjenom faktorijela nekog broja. U tu svrhu ćemo koristiti *Stirlingovu formulu*

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{1.1}$$

1.2 Vjerojatnosne metode

U teoriji grafova mnogi se problemi mogu napasti vjerojatnosnim postupcima i metodama. Recimo da želimo dokazati da postoji određena struktura. Tada konstruiramo vjerojatnosni prostor za takve strukture i pokažemo da u tom prostoru tražena struktura ima strogo pozitivnu vjerojatnost. Zato takva struktura postoji. U ovom radu mi ćemo proučavati vjerojatnosni model slučajnih grafova, koji se onda može primijeniti i kao metoda dokazivanja determinističkih tvrdnji korištenjem vjerojatnosti. Za početak, uvodimo osnovne pojmove.

Neka je n prirodan broj te $0 \leq p \leq 1$. *Slučajni graf* $G(n, p)$ je vjerojatnosni prostor nad skupom svih jednostavnih grafova s vrhovima $\{1, \dots, n\}$ određen s

$$\mathbb{P}(\{i, j\} \in G) = p, \quad i, j \in \{1, \dots, n\},$$

pri čemu su ti događaji međusobno nezavisni. U posebnom slučaju $p = \frac{1}{2}$ svi grafovi, promatrani kao elementarni događaji tog vjerojatnosnog prostora, imaju jednaku vjerojatnost pojavljivanja, $\frac{1}{2^{\binom{n}{2}}}$. Cilj ovog rada je proučiti svojstva slučajnih grafova $G(n, p)$.

Pogledajmo sada jedan dinamički model slučajnih grafova. Za sve parove i, j neka je $x_{i,j}$ izabran uniformno iz intervala $[0, 1]$, pri čemu su ti izbori međusobno nezavisni. Zamislimo sada da p raste od 0 prema 1. Na početku, za $p = 0$, graf se sastoji samo od n vrhova, bez bridova. Brid $\{i, j\}$ biva uključen u graf kada postane $p \geq x_{i,j}$. Konačno, kad p naraste na 1, svi bridovi su uključeni u graf. U trenutku p, graf sa svim dotad uključenim bridovima ima distribuciju $G(n, p)$. Kako p raste, sve je više bridova uključeno u graf.

Iz teorije vjerojatnosti

U našoj analizi slučajnih grafova često će nam trebati pojmovi i rezultati iz teorije vjerojatnosti. Za početak navodimo teorem koji daje vezu između binomne i Poissonove distribucije. Dokaz teorema može se pronaći u knjizi N. Sarape [8].

Teorem 1.2.1. *(Poisson) Neka je $X_n \sim B(n, p_n)$ za $n \in \mathbb{N}$ i neka je $\lim_{n\to\infty} p_n = 0$, $\lim_{n\to\infty} np_n = \lambda$, gdje je λ fiksan broj, $0 < \lambda < \infty$. Tada za svako $k = 0, 1, 2, \dots$ vrijedi*

$$\lim_{n\to\infty} \mathbb{P}[X_n = k] = \lim_{n\to\infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Za ograničenu funkciju distribucije F na $\mathbb{R}$ definiramo karakterističnu funkciju.

Definicija 1.2.2. Karakteristična funkcija *od F jest funkcija ρ definirana sa*

$$\rho(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x) = \int_{-\infty}^{\infty} \cos tx\, dF(x) + i \int_{-\infty}^{\infty} \sin tx\, dF(x), \quad t \in \mathbb{R}.$$

Definicija 1.2.3. *Neka je X slučajna varijabla s funkcijom distribucije F_X.* Karakteristična funkcija ρ_X od X *je karakteristična funkcija od F_X.*

Napomena 1.2.4. *Iz definicije očekivanja slučajne varijable X slijedi*

$$\rho_X(t) = \int_{-\infty}^{\infty} e^{itx} dF_X(x) = \mathbb{E}[e^{itX}], \quad t \in \mathbb{R}.$$

Sljedeći teorem pokazuje da je korespondencija između funkcija distribucije i karakterističnih funkcija 1–1 korespondencija.

Teorem 1.2.5. *(Teorem jedinstvenosti) Neka su F_1 i F_2 funkcije distribucije na $\mathbb{R}$ i neka one imaju istu karakterističnu funkciju, tj. za sve $t \in \mathbb{R}$ vrijedi*

$$\int_{-\infty}^{\infty} e^{itx} dF_1(x) = \int_{-\infty}^{\infty} e^{itx} dF_2(x).$$

Tada je $F_1 = F_2$.

Pomoću karakterističnih funkcija lako se pokaže sljedeći teorem koji će nam biti važan u nastavku ovog rada.

Teorem 1.2.6. *Neka su $X_1, \dots, X_n$ nezavisne slučajne varijable, neka je $X_k \sim P(\lambda_k)$ za $k = 1, \dots, n$ i neka je $X = \sum_{k=1}^n X_k$. Tada je $X \sim P(\sum_{k=1}^n \lambda_k)$.*

Dokaz. Za proizvoljnu slučajnu varijablu $Y \sim P(\lambda)$ izračunajmo njezinu karakterističnu funkciju. Iz napomene 1.2.4 i definicije očekivanja diskretne slučajne varijable je

$$\varphi_Y(t) = \mathbb{E}[e^{itY}] = \sum_{l=0}^{\infty} e^{itl} \frac{\lambda^l}{l!} e^{-\lambda} = e^{-\lambda} \sum_{l=0}^{\infty} \frac{(\lambda e^{it})^l}{l!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}, \quad t \in \mathbb{R}.$$

Sada zbog napomene 1.2.4 i nezavisnosti slučajnih varijabli slijedi

$$\begin{aligned} \varphi_X(t) &= \mathbb{E}[e^{it\sum_{k=1}^n X_k}] = \mathbb{E}[\prod_{k=1}^n e^{itX_k}] = (\text{nezavisnost}) = \prod_{k=1}^n \mathbb{E}[e^{itX_k}] \\ &= \prod_{k=1}^n \varphi_{X_k}(t) = \prod_{k=1}^n e^{\lambda_k(e^{it}-1)} = e^{(e^{it}-1)\sum_{k=1}^n \lambda_k} = \varphi_{P(\sum_{k=1}^n \lambda_k)}(t), \quad t \in \mathbb{R}. \end{aligned}$$

Zbog teorema jedinstvenosti je $X \sim P(\sum_{k=1}^n \lambda_k)$. □

Teorem 1.2.7. *Neka su zadane slučajne varijable $X \sim B(n, p)$ i $Y \sim B(X, q)$, $p \in [0, 1], q \in [0, 1]$ i $n \in \mathbb{N}$. Tada je $Y \sim B(n, pq)$.*

Dokaz. Za $n, k, l \in \mathbb{N}$ takve da je $n \geq k \geq l$ vrijedi

$$\binom{n}{k}\binom{k}{l} = \binom{n}{l}\binom{n-l}{k-l}. \tag{1.2}$$

Tvrdnju (1.2) dokazujemo kombinatornim argumentima. Lijeva strana: neka je dan n-člani skup S. Biramo uređeni par (A, B), gdje je A k-člani podskup od S, a B je l-člani podskup od A. Desna strana: najprije odabiremo l-člani podskup B od S, onda od preostalih elemenata biramo $k-l$-člani podskup C te kreiramo uređeni par $(Z \cup Y, Y)$. Tvrdnja teorema sad slijedi iz

$$\begin{aligned}
\mathbb{P}[Y = l] &= \mathbb{P}[Y = l, X \in \{0, 1, \dots, n\}] \\
&= \sum_{k=0}^{n} \mathbb{P}[X = k]\mathbb{P}[Y = l | X = k] \\
&= \sum_{k=l}^{n} \mathbb{P}[X = k]\mathbb{P}[B(k, q) = l] \\
&= \sum_{k=l}^{n} \binom{n}{k} p^k (1-p)^{n-k} \binom{k}{l} q^l (1-q)^{k-l} \\
&= \sum_{k=l}^{n} \binom{n}{l}\binom{n-l}{k-l} p^k (1-p)^{n-k} q^l (1-q)^{k-l} \\
&= \sum_{k=0}^{n-l} \binom{n}{l}\binom{n-l}{k} p^{l+k} (1-p)^{n-l-k} q^l (1-q)^{k} \\
&= \binom{n}{l} (pq)^l \sum_{k=0}^{n-l} \binom{n-l}{k} (p-pq)^k (1-p)^{n-l-k} \\
&= \binom{n}{l} (pq)^l (p - pq + 1 - p)^{n-l} \\
&= \binom{n}{l} (pq)^l (1-pq)^{n-l}.
\end{aligned}$$

□

Razlikujemo više tipova konvergencije slučajnih varijabli, a u ovom radu najvažnije će biti konvergencija po vjerojatnosti i konvergencija po distribuciji.

Definicija 1.2.8. *Kažemo da niz $(X_n, n \in \mathbb{N})$ slučajnih varijabli* konvergira po vjerojatnosti *prema slučajnoj varijabli X ako za svako $\epsilon > 0$ vrijedi*

$$\lim_{n \to \infty} \mathbb{P}[|X_n - X| \geq \epsilon] = 0$$

To označujemo $(\mathbb{P}) \lim_n X_n = X$.

Definicija 1.2.9. *Kažemo da niz $(X_n, n \in \mathbb{N})$ slučajnih varijabli* konvergira po distribuciji *prema slučajnoj varijabli X ako je*

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x), \quad x \in C(F_X),$$

gdje je $C(F_X)$ skup svih točaka neprekidnosti funkcije F_X. To označujemo $(\mathcal{D}) \lim_n X_n = X$.

Iz definicije konvergencije po distribuciji jednostavno slijedi sljedeći teorem.

Teorem 1.2.10. *Neka je* $(X_n, n \in \mathbb{N})$ *niz slučajnih varijabli, pri čemu za svaki* $n \in \mathbb{N}$ *vrijedi* $X_n \sim B(m, p_n)$ *za fiksan* $m \in \mathbb{N}$. *Ako je* $\lim_{n\to\infty} p_n = p$, *tada niz* $(X_n, n \in \mathbb{N})$ *slučajnih varijabli konvergira po distribuciji prema slučajnoj varijabli* $X \sim B(m, p)$.

Navodimo još dvije važne nejednakosti čiji se dokazi mogu pronaći u knjizi N. Sarape, [8].

Teorem 1.2.11. *(Markovljeva nejednakost) Neka je* X *slučajna varijabla,* $r > 0$ *i* $\mathbb{E}[|X|^r] < \infty$. *Tada za proizvoljno* $\epsilon > 0$ *vrijedi*

$$\mathbb{P}[|X| \geq \epsilon] \leq \frac{\mathbb{E}[|X|^r]}{\epsilon^r}.$$

Teorem 1.2.12. *(Čebiševljeva nejednakost) Neka je* X *slučajna varijabla s konačnim očekivanjem i varijancom. Tada za proizvoljno* $\epsilon > 0$ *vrijedi*

$$\mathbb{P}[|X - \mathbb{E}X| \geq \epsilon] \leq \frac{Var\,X}{\epsilon^2}.$$

Teorem 1.2.1 dao nam je vezu između binomne i Poissonove distribucije pa je opravdano binomnu slučajnu varijablu aproksimirati Poissonovom. To će se pokazati korisnim (između ostalog) i zbog sljedećeg teorema.

Teorem 1.2.13. *(Chernoffove nejednakosti) Neka je* X *slučajna varijabla s Poissonovom distribucijom i očekivanjem* λ. *Za* $\epsilon > 0$

$$\mathbb{P}[X \leq \lambda(1-\epsilon)] \leq e^{-\epsilon^2 \frac{\lambda}{2}},$$
$$\mathbb{P}[X \geq \lambda(1+\epsilon)] \leq [e^{\epsilon}(1+e)^{-(1+\epsilon)}]^{\lambda}.$$

Dokaz se može pronaći u knjizi Alon-Spencer, [1].

1.3 Jednostavni proces grananja

Proces grananja opisuje populaciju koja započinje s jednom jedinkom u trenutku $t = 0$. U trenutku $t = 1$ ta jedinka daje slučajan broj potomaka s nekom distribucijom na $\{0, 1, 2, \dots\}$, a sama nestaje iz populacije. U trenutku $t = 2$, svaka od jedinki u populaciji daje slučajan broj potomaka, s istom distribucijom, nezavisno jedna od druge. Postupak se nastavlja sve dok u populaciji postoje jedinke. Označimo sa Y_t broj jedinki u populaciji u trenutku $t \geq 0$. Niz slučajnih varijabli $(Y_t, t \geq 0)$ zove se *proces grananja* ili *Galton-Watsonov proces*.

Definirajmo formalno proces grananja. Neka je $\{Z_{n,j}, n \geq 1, j \geq 1\}$ familija nezavisnih, jednako distribuiranih slučajnih varijabli s vrijednostima u $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. Stavimo $Y_0 = 1$, te za $n \geq 1$,

$$Y_n = \sum_{j=1}^{Y_{n-1}} Z_{n,j} = Z_{n,1} + \cdots + Z_{n,Y_{n-1}},$$

uz konvenciju da je $Y_n = 0$ ukoliko je $Y_{n-1} = 0$. Interpretacija gornje definicije je sljedeća: u trenutku $n-1$ broj jedinki u populaciji je Y_{n-1}. Jedinka j daje $Z_{n,j}$ potomaka i nestaje iz populacije. Broj jedinki u populaciji u trenutku n jednak je zbroju svih potomaka jedinki $n-1$-ve generacije, tj. gornjoj sumi. Broj pribrojnika u toj sumi jednak je slučajnom broju jedinki u $n-1$-voj generaciji. U ovom ćemo radu definirati jednostavni proces grananja kod grafova koji ćemo usporediti s Galton-Watsonovim procesom grananja.

Poglavlje 2

Slučajni grafovi

2.1 Erdős-Rényijev model

Prisjetimo se definicije slučajnog grafa. Neka je n prirodan broj i neka je $0 \leq p \leq 1$. Slučajni graf $G(n, p)$ definirali smo kao vjerojatnosni prostor nad skupom svih jednostavnih grafova sa skupom vrhova $\{1, \dots, n\}$, pri čemu je vjerojatnost pojavljivanja svakog brida $\{i, j\}$ jednaka p, tj.

$$\mathbb{P}[\{i, j\} \in G] = p,$$

gdje su ti događaji međusobno nezavisni. Nekad se dozvoljava da $p = p_n$ bude funkcija od n. Ovaj model često se koristi u vjerojatnosnim metodama za dokazivanje postojanja određenog grafa. U ovom poglavlju proučavat ćemo svojstva vjerojatnosnog prostora $G(n, p)$ kao teme od zasebnog interesa.

U svojim ranim originalnim člancima, Erdős i Rényi predstavili su svoj, ponešto drukčiji, model slučajnog grafa, u oznaci $G(n, e)$ sa n vrhova i točno e bridova. I oni su imali dinamički model: Počinjemo sa n vrhova i bez bridova i dodajemo na slučajan način jedan po jedan brid. Ipak, mi ćemo u daljnjim razmatranjima promatrati prvotno opisani, moderniji model slučajnih grafova $G(n, p_n)$, koji će imati vrlo slična svojstva kao $G(n, e)$ kada je $p_n \sim \frac{e}{\binom{n}{2}}$.

2.2 Monotona svojstva grafova

Slučajni graf $G(n, p)$ definirali smo kao vjerojatnosni prostor nad skupom svih grafova s n vrhova, pri čemu je vjerojatnost pojavljivanja svakog brida jednaka p. Za bilo koje dano svojstvo A grafa G možemo izračunati vjerojatnost da $G(n, p)$ ima svojstvo A, što ćemo označavati sa $\mathbb{P}[G(n, p) \models A]$.

Definicija 2.2.1. *Za svojstvo A reći ćemo da je* padajuće svojstvo *ako svaki podgraf grafa G sa svojstvom A također ima svojstvo A. Svojstvo A je* rastuće *ako iz činjenice da neki podgraf od G ima svojstvo A slijedi da i cijeli graf G ima svojstvo A. Padajuća i rastuća svojstva zajedno nazivamo* monotona svojstva.

Intuitivno, padajuće svojstvo je svako svojstvo koje ostaje sačuvano ako grafu obrišemo neki brid, dok rastuće svojstvo ostaje sačuvano dodavanjem nekih bridova. Jedan primjer padajućeg svojstva je $A = \{G \text{ ne sadrži trokut}\}$, dok je svojstvo $\neg A = \{G \text{ sadrži trokut}\}$ rastuće.

Propozicija 2.2.2. *Za svako rastuće svojstvo A i za $p_1 \leq p_2$ vrijedi*

$$\mathbb{P}[G(n, p_1) \models A] \leq \mathbb{P}[G(n, p_2) \models A].$$

Dokaz. Ako je $p_1 \neq 1$, definirajmo p_0 tako da vrijedi

$$p_1 + (1 - p_1)p_0 = p_2.$$

Ako je baš $p_1 = 1$, tada po pretpostavci mora biti i $p_2 = 1$ te gornja relacija vrijedi uz bilo koji odabir od p_0, npr. $p_0 = 0$.

Inače je $p_0 \in [0, 1]$ i za $G_0 \sim G(n, p_0)$ te $G_1 \sim G(n, p_1)$ vrijedi $G_0 \cup G_1 \sim G(n, p_2)$. Naime, pojedini brid se pojavljuje u $G_0 \cup G_1$ ako se ili pojavljuje u G_1, ili se ne pojavljuje u G_1, ali se pojavljuje u G_0. Vjerojatnost prvog događaja je p_1, a vjerojatnost drugog (zbog nezavisnosti) iznosi $(1 - p_1)p_0$. Po našem odabiru zbroj te dvije vjerojatnosti iznosi upravo p_2. Stoga je

$$\mathbb{P}[G_1 \models A] \leq \mathbb{P}[G_0 \cup G_1 \models A].$$ □

Iz navedenog teorema slijedi da, ako je A monotono svojstvo, tada je $\mathbb{P}[G(n, p) \models A]$ monotona funkcija od p. Pogledajmo sada jedan konstruktivan primjer. Neka je $A = \{G \text{ ne sadrži trokut}\}$ te neka je X_n broj trokuta sadržanih u grafu $G(n, p_n)$. Maksimalni broj trokuta u grafu s n vrhova je $t = \binom{n}{3}$. Neka su $Y_1, \ldots, Y_t$ slučajne varijable koje detektiraju sadrži li graf G i-ti trokut (u tom slučaju slučajna varijabla Y_i poprima vrijednost 1, a inače 0). Vjerojatnost pojavljivanja svakog brida je p_n, a zbog nezavisnosti slijedi da je vjerojatnost pojavljivanja svakog trokuta jednaka p_n^3. Iz definicije očekivanja diskretne slučajne varijable slijedi da je $E[Y_i] = p_n^3$, za svaki i. Zbog linearnosti očekivanja tada je

$$\mathbb{E}[X_n] = \mathbb{E}\Big[\sum_{i=1}^{t} Y_i\Big] = \sum_{i=1}^{t} \mathbb{E}[Y_i] = \binom{n}{3} p_n^3.$$

Ako uvedemo parametrizaciju $p_n = \frac{c}{n}$, za $c > 0$, tada vrijedi

$$\lim_{n\to\infty} \mathbb{E}(X_n) = \lim_{n\to\infty} \binom{n}{3} p_n^3 = \lim_{n\to\infty} \frac{n(n-1)(n-2)}{6} \left(\frac{c}{n}\right)^3 = \frac{c^3}{6}.$$

Može se dokazati da niz slučajnih varijabli $(X_n, n \in \mathbb{N})$ konvergira po distribuciji Poissonovoj slučajnoj varijabli $P(\lambda)$ s parametrom $\lambda = \frac{c^3}{6}$. Iz toga slijedi

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \models A] = \lim_{n\to\infty} \mathbb{P}[X_n = 0] = \mathbb{P}[X = 0] = e^{-\frac{c^3}{6}}.$$

Primijetimo da vrijedi

$$\lim_{c\to 0} e^{-\frac{c^3}{6}} = 1, \qquad \lim_{c\to\infty} e^{-\frac{c^3}{6}} = 0. \tag{2.1}$$

Ako za svaki $n \in \mathbb{N}$ stavimo $p_n = \frac{10^6}{n}$, tada zbog (2.1) zaključujemo da je vjerojatnost da će $G(n, p_n)$ sadržavati trokut jako mala, dok je za $p_n = \frac{10^{-6}}{n}$ ta vjerojatnost velika. Primijetimo još da je u drugom slučaju za male $n \in \mathbb{N}$, $p_n > 1$, dok za dovoljno velike n vrijedi $p_n \in [0, 1]$. Budući da su naši rezultati uvijek asimptotski, pretpostavit ćemo da je $G(n, p_n)$ definiran samo za dovoljno velike n.

U dinamičkom modelu prvi trokut se gotovo uvijek pojavi za $p_n = \Theta(\frac{1}{n})$. Ako definiramo $p_n = n^{-0.9}$, zbog $p_n \gg n^{-1}$ zaključujemo da će $G(n, p_n)$ s velikom vjerojatnosti sadržavati trokut, tj. preciznije,

$$\lim_{n\to\infty} \mathbb{P}[G(n, n^{-0.9}) \models A] = 1.$$

Slično, za $p_n \ll n^{-1}$, npr. $p_n = \frac{1}{n \ln n}$, vrijedit će da $G(n, p_n)$ s malom vjerojatnosti sadrži trokut.

Postoje mnoga svojstva grafova koja postaju istinita za vrlo uski interval vrijednosti od p, što nas navodi na sljedeću definiciju.

Definicija 2.2.3. *Funkciju r, tj. $n \mapsto r(n)$, zovemo* prijelaznom funkcijom *za svojstvo A grafa $G(n, p_n)$ ako vrijedi*

$$\lim_{n\to\infty} \mathbb{P}[G(n,p_n) \models A] = 0, \quad \text{za } p_n \ll r(n),$$
$$\lim_{n\to\infty} \mathbb{P}[G(n,p_n) \models A] = 1, \quad \text{za } p_n \gg r(n)$$

ili obratno.

U našem primjeru, $r(n) = \frac{1}{n}$ je prijelazna funkcija za $A = \{\text{graf ne sadrži trokut}\}$. Primijetimo još da prijelazna funkcija za neko svojstvo, ukoliko postoji, nije jedinstvena — npr. $r(n) = \frac{10}{n}$ je također prijelazna funkcija za A.

2.3 Zakoni nula-jedan

U ovom poglavlju pažnju usmjeravamo na svojstva grafova koja možemo izraziti u jeziku logike prvog reda, koji se sastoji od varijabli $(x, y, z, \dots)$, koje uvijek predstavljaju vrhove grafa, znakova jednakosti i susjedstva $(=, \sim)$, uobičajenih Booleovih operatora $(\vee, \wedge, \neg)$ te univerzalnih i egzistencijalnih kvantifikatora $(\forall_x, \exists_x)$. Takvo svojstvo A zvat ćemo *svojstvo prvog reda*. Kao primjer, izrazimo simbolima postojanje trokuta u grafu:

$$\exists_x \exists_y \exists_z (x \sim y \wedge y \sim z \wedge x \sim z),$$

te svojstvo da ne postoji izolirana točka:

$$\forall_x \exists_y (x \sim y).$$

Za svako svojstvo A te svako n, promatramo vjerojatnost da slučajni graf $G(n, p_n)$ zadovoljava svojstvo A, što ćemo pisati kao

$$\mathbb{P}[G(n, p_n) \models A].$$

Reći ćemo da svojstvo A *vrijedi s velikom vjerojatnosti* ili da je *asimptotski ispunjeno* ako je limes niza ovih vjerojatnosti jednak 1. Nadalje, reći ćemo da neko svojstvo *A vrijedi s malom vjerojatnosti* ako je taj limes jednak 0, što upravo znači da suprotno svojstvo vrijedi s velikom vjerojatnosti. Objekt našeg razmatranja u ovom poglavlju bit će sljedeći teorem kojeg su 1969. dokazali Glebskii, Kogan, Liagonkii i Talanov [5] te 1976., nezavisno od njih i Fagin [4].

Teorem 2.3.1. *Ako za svako $n \in \mathbb{N}$ stavimo $p_n = p$, za fiksirani $0 < p < 1$, tada za svako svojstvo prvog reda A vrijedi*

$$\lim_{n \to \infty} \mathbb{P}[G(n, p) \models A] = 0 \text{ ili } 1.$$

Navedeni teorem ima zanimljivu interpretaciju kada je $p = \frac{1}{2}$, jer je tada vjerojatnost pojavljivanja svakog grafa iz $G(n, p)$ jednaka. Tada iz teorema 2.3.1 slijedi da svako svojstvo A prvog reda s velikom vjerojatnosti vrijedi za sve grafove ili s velikom vjerojatnosti ono ne vrijedi ni za koji graf.

Kako bismo dokazali teorem 2.3.1, definiramo prvo *Ehrenfeuchtovu* igru često zvana *Igra naprijed-nazad*. Neka su G, H dva jednostavna grafa čiji su skupovi vrhova disjunktni te neka je $t \geq 0$ cijeli broj. Definiramo igru s potpunom informacijom, u oznaci $EHR(G, H, t)$ sa dva igrača, nazovimo ih Kvaritelj i Duplikator. Igra se sastoji od t krugova od kojih svaki krug ima dva dijela. Prvo Kvaritelj bira vrh $x \in V(G)$ ili vrh $y \in V(H)$. Dakle, sam bira iz kojeg će grafa odabrati vrh. Tada Duplikator odgovara izborom vrha iz drugog grafa. Nakon t krugova imamo po t odabranih vrhova iz oba grafa. Označimo sa $x_1, x_2, \dots, x_t$ vrhove odabrane iz $V(G)$ te $y_1, y_2, \dots, y_t$ vrhove iz $V(H)$, gdje su x_i, y_i vrhovi odabrani u i-tom krugu. Na kraju Duplikator pobjeđuje ako i samo ako su grafovi inducirani odabranim vrhovima izomorfni, tj. ako i samo ako za sve $1 \leq i < j \leq t$, vrijedi

$$\{x_i, x_j\} \in E(G) \iff \{y_i, y_j\} \in E(H).$$

Opisana igra koristi se kao tehnika u matematici kojom se otkriva jesu li dvije strukture ekvivalentne. Glavna ideja leži u tome kako imamo dvije strukture i dva igrača, od kojih jedan pokušava pokazati da su te strukture različite, dok drugi pokušava pokazati da su na neki način slične (prema logici prvog reda). Kako nema nikakvih skrivenih poteza te igra ne može biti neriješena, jedan od igrača mora imati pobjedničku strategiju pa ćemo reći da je taj igrač pobijedio u igri $EHR(G, H, t)$.

Lema 2.3.2. *Za svako svojstvo prvog reda A postoji $t = t(A)$ takav da ako su grafovi G, H takvi da $G \models A$ te $H \models \neg A$, tada Kvaritelj sigurno pobjeđuje u igri $EHR(G, H, t)$.*

Za dokaz ove leme potrebno je ulaziti u logiku prvog reda pa ga navodimo bez dokaza, no promotrimo zato sljedeći primjer. Neka je A svojstvo $\forall_x \exists_y (x \sim y)$, tj. neka je $A = \{\text{Graf nema izolirani vrh}\}$, neka su grafovi G, H takvi da $G \models A$ te $H \models \neg A$ te stavimo $t = 2$. Kvaritelj počinje birajući izolirani vrh $y_1 \in V(H)$, što je moguće zbog pretpostavke $H \models \neg A$. Duplikator mora odabrati $x_1 \in V(G)$, no zbog pretpostavke $G \models A$, x_1 sigurno nije izoliran vrh pa u drugom krugu postoji $x_2 \in V(G)$ takav da $x_1 \sim x_2$, Kvaritelj će odabrati njega i Duplikator više ne može odabrati vrh u grafu H koji bi mu donio pobjedu.

Propozicija 2.3.3. *Ako postoji $0 < p < 1$ takav da za svako $t \geq 0$ te za $G(n, p)$, $H(m, p)$ nezavisno izabrane slučajne grafove čiji su skupovi vrhova međusobno disjunktni, vrijedi*

$$\lim_{m,n\to\infty} \mathbb{P}[\textit{Duplikator je pobijedio u igri } EHR(G(n, p), H(m, p), t)] = 1,$$

tada za taj p vrijedi zakon nula-jedan.

Napomena 2.3.4. *Primijetimo još jednom da za svaki izbor od G, H igra $EHR(G, H, t)$ ne može biti neriješena i netko od dvojice igrača svakako mora pobijediti.*

Dokaz. Pretpostavimo suprotno, tj. da p ne zadovoljava zakon nula-jedan. To znači da postoji svojstvo prvog reda A za koje ili vrijedi da je

$$\lim_{n\to\infty} \mathbb{P}[G(n, p) \models A] = c, \tag{2.2}$$

za $0 < c < 1$, ili limes ne postoji. Pokazat ćemo da oba slučaja vode do kontradikcije.

Iz leme 2.3.2 slijedi da postoji $t = t(A)$, takav da ako G zadovoljava svojstvo A, a H ne zadovoljava svojstvo A, Kvaritelj pobjeđuje u igri $EHR(G, H, t)$. U prvom slučaju, budući da su $G(n, p)$ i $H(m, p)$ nezavisno odabrani slučajni grafovi, zbog (2.2) slijedi da je vjerojatnost da točno jedan od navedenih grafova zadovoljava svojstvo A jednaka $2c(1 - c) > 0$. U tom slučaju bi Kvaritelj pobijedio u igri $EHR(G(n, p), H(m, p), t)$ što je kontradikcija s pretpostavkom propozicije.

U slučaju da limes ne postoji, zbog ograničenosti niza $\mathbb{P}[G(n, p) \models A]$, on ima barem jedan konvergentan podniz. Tada postoji podniz čiji je limes $c \in (0, 1)$ ili postoje dva podniza čiji su limesi 0 i 1. U oba slučaja na isti način zaključujemo da Kvaritelj pobjeđuje u igri $EHR(G, H, t)$ s pozitivnom vjerojatnosti. Kontradikcija. □

Propozicija 2.3.3 omogućuje nam da se sa svijeta logike prvog reda prebacimo na svijet grafova. Kako bismo dokazali zakon nula-jedan, više ne trebamo znati logiku prvog reda — dovoljno je samo pronaći dobru strategiju za Duplikatora.

Definicija 2.3.5. *Kažemo da graf G ima svojstvo proširenja razine s ako za svake različite vrhove $u_1, \dots, u_a, v_1, \dots, v_b \in V(G)$ takve da je $a + b \leq s$ postoji vrh $x \in V(G)$ takav da vrijedi $\{x, u_i\} \in E(G)$, za svaki $1 \leq i \leq a$ te $\{x, v_j\} \notin E(G)$, za svaki $1 \leq j \leq b$.*

Pretpostavimo sada da grafovi G i H oba imaju svojstvo proširenja razine $t - 1$. Tada postoji strategija kojom bi Duplikator pobijedio u $EHR(G, H, t)$ igri. U i-tom krugu, neka su izabrani vrhovi $x_1, \dots, x_{i-1}, y_1, \dots, y_{i-1}$ te bez smanjenja općenitosti pretpostavimo da Kvaritelj bira vrh x_i. Duplikator tada bira vrh y_i koji ima iste susjede među $y_j,\ j < i$ kao što x_i ima među vrhovima $x_j,\ j < i$. Svojstvo proširenja razine $t - 1$ garantira nam da će takav y_i sigurno postojati.

Propozicija 2.3.6. *Za svako fiksno p, $0 < p < 1$, te svako s, slučajni graf $G(n, p)$ s velikom vjerojatnosti ima svojstvo proširenja razine s.*

Dokaz. Za sve $u_1, \dots, u_a, v_1, \dots, v_b, x \in V(G)$ takve da je $a + b \leq s$, neka je $E_{u_1,\dots,u_a,v_1,\dots,v_b,x}$ događaj definiran sa

$$E_{u_1,\dots,u_a,v_1,\dots,v_b,x} = \Big\{ \{x, u_i\} \in E(G),\ \text{za } 1 \leq i \leq a, \{x, v_j\} \notin E(G),\ \text{za } 1 \leq j \leq b \Big\}.$$

Tada zbog nezavisnosti događaja $\{u_i, u_j\} \in G(n, p)$ vrijedi

$$\mathbb{P}[\ E_{u_1,\dots,u_a,v_1,\dots,v_b,x}\ ] = p^a(1 - p)^b.$$

Stavimo sada

$$E_{u_1,\dots,u_a,v_1,\dots,v_b} = \bigcap_{\substack{x \\ x \neq u_1,\dots,u_a \\ x \neq v_1,\dots,v_b}} E^c_{u_1,\dots,u_a,v_1,\dots,v_b,x}.$$

Ponovno, zbog nezavisnosti događaja $\{u_i, u_j\} \in G(n, p)$ slijedi da su događaji u presjeku međusobno nezavisni (presjek po različitim x). Iz toga slijedi da je

$$\mathbb{P}\Big[\bigcap_{\substack{x \\ x \neq u_1, \dots, u_a \\ x \neq v_1, \dots, v_b}} E^c_{u_1, \dots, u_a, v_1, \dots, v_b, x} \Big] = \text{(vjerojatnost suprotnog događaja)} = [1 - p^a(1 - p)^b]^{n-a-b}.$$

Postavimo sada $\epsilon = \min\{p, 1 - p\}^s$. Zbog nejednakosti

$$1 - p^a(1 - p)^b \leq 1 - \min\{p, 1 - p\}^a \min\{p, 1 - p\}^b \leq 1 - \min\{p, 1 - p\}^s = 1 - \epsilon$$

vrijedi

$$\mathbb{P}\Big[\bigcap_{\substack{x \\ x \neq u_1, \dots, u_a \\ x \neq v_1, \dots, v_b}} E^c_{u_1, \dots, u_a, v_1, \dots, v_b, x} \Big] \leq (1 - \epsilon)^{n-(a+b)} \leq (1 - \epsilon)^{n-s}.$$

Označimo sada sa

$$E = \bigcup_{\substack{u_1, \dots, u_a \\ v_1, \dots, v_b \\ a+b \leq s}} E_{u_1, \dots, u_a, v_1, \dots, v_b}$$

uniju po svim različitim vrhovima $u_1, \dots, u_a, v_1, \dots, v_b$ takvim da je $a + b \leq s$. Budući da parova (a, b) koji bi zadovoljavali uvjet $a + b \leq s$ možemo odabrati na manje od s^2 načina, dok vrhove možemo odabrati na manje od n^s načina, ukupno imamo manje od $s^2 n^s$ izbora. Iz toga, zbog konačne subaditivnosti vjerojatnosne mjere, slijedi

$$\mathbb{P}[E] \leq s^2 n^s (1 - \epsilon)^{n-s}.$$

Budući da naše rezultate gledamo asimptotski, zbog

$$\lim_{n \to \infty} s^2 n^s (1 - \epsilon)^{n-s} = 0$$

i teorema o sendviču zaključujemo da E asimptotski ima vjerojatnost 0. Stoga suprotan događaj, E^c, na limesu kada $n \to \infty$ ima vjerojatnost 1, a on odgovara upravo svojstvu iz iskaza, koje je posljedično asimptotski ispunjeno. Dakle, vrijedi tvrdnja da $G(n, p)$ s velikom vjerojatnosti ima svojstvo proširenja razine s. □

Ovime je dokazan teorem 2.3.1. Za svako fiksno $p_n = p$, $0 < p < 1$ i s, asimptotski gledano (kada $m, n \to \infty$) će vjerojatnost da $G(n, p)$ i $H(m, p)$ oba imaju svojstvo proširenja razine s biti 1. Dokazali smo da tada slijedi da će Duplikator pobijediti u igri $EHR(G(n, p), H(m, p), s)$ s vjerojatnosti 1, tj. iz propozicije 2.3.3 slijedi da vrijedi zakon nula-jedan.

Zanimljivo je kako slučaj kad je $p_n = n^{-\alpha}$ zahtijeva puno složeniji pristup. O rezultatu govori sljedeći teorem, kojeg su 1988. dokazali Shelah i Spencer i kojeg nećemo dokazivati. Njegov dokaz se može naći u [9] i recimo samo da je donekle sličan dokazu teorema 2.3.1.

Teorem 2.3.7. *Neka je α, $0 < \alpha < 1$ iracionalan te za svako $n \in \mathbb{N}$ stavimo $p_n = n^{-\alpha}$. Tada za svako svojstvo prvog reda A vrijedi*

$$\lim_{n \to \infty} \mathbb{P}[G(n, p_n) \models A] = 0 \text{ } ili \text{ } 1.$$

Poglavlje 3

Svojstva slučajnih grafova

3.1 Osnovni rezultati

Neka je X slučajna varijabla koja može poprimiti samo nenegativne cjelobrojne vrijednosti. Pretpostavimo da želimo ograničiti $\mathbb{P}[X = 0]$ obzirom na vrijednosti $\mu = E[X]$ i $\sigma^2 = Var[X]$. Kao prvo, vidimo da je

$$\mathbb{P}[X > 0] \leq \mathbb{E}[X].$$

To vrijedi zbog

$$\mathbb{E}[X] = \mathbb{E}[X\,\mathbf{1}_{\{X=0\}} + X\,\mathbf{1}_{\{X>0\}}] = \mathbb{E}[X\,\mathbf{1}_{\{X\geq 1\}}] \geq \mathbb{E}[\mathbf{1}_{X\geq 1}] = \mathbb{P}[X > 0].$$

Pretpostavimo li sada da je $(X_n)_{n\in\mathbb{N}}$ niz slučajnih varijabli za koji vrijedi $\mathbb{E}[X_n] \to 0$, tada je (zbog nenegativnosti vjerojatnosti i teorema o sendviču)

$$\lim_{n\to\infty} \mathbb{P}(X_n > 0) = 0, \text{ tj. } \lim_{n\to\infty} \mathbb{P}(X_n = 0) = 1,$$

a u skladu s terminologijom iz prethodnog poglavlja, možemo reći i da je $X_n = 0$ s velikom vjerojatnosti. S druge strane, iz $\mathbb{E}[X_n] \to \infty$ na gornji način ne možemo ništa zaključiti. Za tu svrhu će nam poslužiti rezultati koji slijede.

Teorem 3.1.1. *Za nenegativnu slučajnu varijablu X s konačnom varijancom i strogo pozitivnim očekivanjem vrijedi*

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

Dokaz. Označimo li $\mu = \mathbb{E}[X]$, $\sigma^2 = Var[X]$, zbog Čebiševljeve nejednakosti vrijedi

$$\mathbb{P}[X = 0] \leq \mathbb{P}[|X - \mu| \geq \mu] \leq \frac{Var[X]}{\mu^2}. \qquad \square$$

Ovaj rezultat primijenit ćemo na naše rezultate koji su asimptotski.

Korolar 3.1.2. *Pretpostavimo li da za niz nenegativnih slučajnih varijabli* $(X_n)_{n\in\mathbb{N}}$ *vrijedi* $Var[X_n] = o(\mathbb{E}[X_n]^2)$, *tada je*

$$\lim_{n\to\infty} \mathbb{P}(X_n > 0) = 1,$$

tj. $X_n > 0$ *s velikom vjerojatnosti.*

Iz dokaza teorema 3.1.1 slijedi da za $\epsilon > 0$ i sasvim proizvoljnu slučajnu varijablu X imamo i

$$\mathbb{P}[\ |X - \mathbb{E}[X]| \geq \epsilon\mathbb{E}[X]\ ] \leq \frac{Var[X]}{\epsilon^2\mathbb{E}[X]^2} \tag{3.1}$$

pa vrijedi sljedeći korolar.

Korolar 3.1.3. *Ako je* $(X_n)_{n\in\mathbb{N}}$ *niz slučajnih varijabli za koji vrijedi* $Var[X_n] = o(\mathbb{E}[X_n]^2)$, *tada* $X_n \sim \mathbb{E}[X_n]$ *s velikom vjerojatnosti, što nam ovdje znači da omjer* $X_n/\mathbb{E}[X_n]$ *po vjerojatnosti konvergira u* 1.

Neka je sada $X = Y_1 + \cdots + Y_m$, gdje je Y_i slučajna varijabla koja detektira događaj A_i, tj. $Y_i = \mathbf{1}_{A_i}$ je indikatorska varijabla događaja A_i. Za i, j za koje vrijedi $i \neq j$ i događaji A_i i A_j nisu nezavisni uvodimo oznaku $i \sim j$. Označimo još

$$\triangle := \sum_{\substack{i,j \\ i\sim j}} \mathbb{P}[A_i \cap A_j]. \tag{3.2}$$

Primijetimo da iz definicije slučajnih varijabli Y_i slijedi da za $i \sim j$ imamo

$$Cov[Y_i, Y_j] = \mathbb{E}[Y_iY_j] - \mathbb{E}[Y_i]\mathbb{E}[Y_j] \leq \mathbb{E}[Y_iY_j] = \mathbb{P}[A_i \cap A_j],$$

dok za $i \neq j$ i $i \nsim j$, zbog nezavisnosti vrijedi $Cov[Y_i, Y_j] = 0$. Osim toga je, zbog definicije slučajne varijable Y_i, $Var[Y_i] = \mathbb{E}[Y_i^2] - \mathbb{E}[Y_i]^2 \leq \mathbb{E}[Y_i]$ pa slijedi

$$\begin{aligned}
Var[X] &= Var[Y_1 + \ldots Y_m] \\
&= \sum_{i,j} Cov[Y_i, Y_j] \\
&= \sum_{\substack{i,j \\ i\sim j}} Cov[Y_i, Y_j] + \sum_{i} Cov[Y_i, Y_i] + \sum_{\substack{i,j \\ i\neq j \\ i\nsim j}} Cov[Y_i, Y_j] \\
&\leq \triangle + \sum_{i} \mathbb{E}[Y_i] \\
&= \triangle + \mathbb{E}[X]. \qquad (3.3)
\end{aligned}$$

Korolar 3.1.4. *Neka je $(X_n)_{n\in\mathbb{N}}$ niz slučajnih varijabli gornjeg oblika, što znači da svaka varijabla X_n ima rastav u obliku sume indikatorskih varijabli nekih događaja i da joj je pridružena pripadna veličina $\triangle_n$. Ako vrijedi $\mathbb{E}[X_n] \to \infty$ i $\triangle_n = o(\mathbb{E}[X_n]^2)$, tada je $X_n > 0$ s velikom vjerojatnosti. Osim toga, $X_n \sim \mathbb{E}[X_n]$ s velikom vjerojatnosti.*

Dokaz. Dokaz direktno slijedi iz teorema 3.1.1 te formula (3.1) i (3.3). □

Definicija 3.1.5. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor. Kažemo da je preslikavanje $T : \Omega \to \Omega$* preslikavanje koje čuva mjeru $\mathbb{P}$ *ako za svako $B \in \mathcal{F}$ vrijedi $T^{-1}B \in \mathcal{F}$ i $\mathbb{P}(T^{-1}B) = \mathbb{P}(B)$.*

Za slučajne varijable $Y_1, \dots, Y_m$ reći ćemo da su *simetrične* ako za svake $i \neq j$ postoji preslikavanje T koje čuva pripadnu vjerojatnosnu mjeru i takvo je da vrijedi $T^{-1}(A_i) = A_j$. Uvijek imamo

$$\triangle = \sum_{\substack{i,j \\ i\sim j}} \mathbb{P}[A_i \cap A_j] = \sum_i \mathbb{P}[A_i] \sum_{\substack{j \\ j\sim i}} \mathbb{P}[A_j | A_i], \tag{3.4}$$

a u spomenutom slučaju zbog svojstva simetrije zaključujemo da unutarnja suma u (3.4) ne ovisi o i pa za proizvoljno fiksno i možemo definirati

$$\triangle^* := \sum_{\substack{j \\ j\sim i}} \mathbb{P}[A_j | A_i],$$

iz čega slijedi

$$\triangle = \sum_i \mathbb{P}[A_i] \triangle^* = \triangle^* \sum_i \mathbb{P}[A_i] = \triangle^* \mathbb{E}[X].$$

Korolar 3.1.6. *Neka je $(X_n)_{n\in\mathbb{N}}$ niz slučajnih varijabli gornjeg oblika. Ako vrijedi $\mathbb{E}[X_n] \to \infty$ i $\triangle_n^* = o(\mathbb{E}[X_n])$, tada je $X_n > 0$ s velikom vjerojatnosti. Osim toga, $X_n \sim \mathbb{E}[X_n]$ s velikom vjerojatnosti.*

Dokaz. Slijedi iz teorema 3.1.1 te formula (3.1) i (3.3). □

3.2 Najveća klika u slučajnom grafu $G(n, \frac{1}{2})$

Definicija 3.2.1. Klika *u grafu G je potpun podgraf od G.* Veličina najveće klike*, u oznaci $\omega(G)$, je broj vrhova najveće klike u grafu G.*

U ovom poglavlju pažnju ćemo posvetiti veličini najveće klike. Fiksirajmo vjerojatnost $p_n = p = \frac{1}{2}$ i promatrajmo veličinu najveće klike $\omega(G)$ u grafu $G(n, \frac{1}{2})$. Za dani $k \in \mathbb{N}$, neka je $X_k^{(n)}$ slučajna varijabla koja broji klike u $G(n, \frac{1}{2})$ veličine k. Za dani skup vrhova A veličine k, označimo sa $X_A^{(n)}$ slučajnu varijablu koja broji klike u $G(n, \frac{1}{2})$ sa skupom vrhova upravo A. Očito je da slučajna varijabla $X_A^{(n)}$ poprima vrijednosti 0 ili 1, iz čega slijedi da je $\mathbb{P}[X_A^{(n)} = 1] = \mathbb{E}[X_A^{(n)}]$, a zbog nezavisnosti je $\mathbb{P}[X_A^{(n)} = 1] = 2^{-\binom{k}{2}}$, budući da svih $\binom{k}{2}$ bridova mora biti uključeno u $G(n, \frac{1}{2})$. Također je jasno da je $X_k^{(n)} = \sum_{|A|=k} X_A^{(n)}$, pri čemu sumiramo po svim A takvim da je $A \subseteq V(G)$ i $|A| = k$. Zbog linearnosti matematičkog očekivanja i jednakosti $\mathbb{P}[X_A^{(n)} = 1] = \mathbb{E}[X_A^{(n)}]$, vrijedi

$$\mathbb{E}[X_k^{(n)}] = \mathbb{E}\Big[\sum_{|A|=k} X_A^{(n)}\Big] = \sum_{|A|=k} \mathbb{E}[X_A^{(n)}] = \sum_{|A|=k} 2^{-\binom{k}{2}} = \binom{n}{k} 2^{-\binom{k}{2}}. \tag{3.5}$$

Pretpostavimo da je $k = k(n)$ funkcija koja ovisi o broju vrhova n slučajnog grafa. Zanima nas vjerojatnost pojavljivanja klike veličine k u slučajnom grafu $G(n, \frac{1}{2})$ kada $n \to \infty$.

Gornja ograda za veličinu klike u slučajnom grafu

Teorem 3.2.2. *Neka je $\epsilon > 0$ i $k = k(n) \geq 2(1+\epsilon)\log_2 n$, za svaki $n \in \mathbb{N}$. Tada vjerojatnost da slučajni graf $G(n, \frac{1}{2})$ sadrži kliku veličine k teži prema 0 kada $n \to \infty$. Dakle, $G(n, 1/2)$ s malom vjerojatnosti sadrži kliku veličine $k(n)$.*

Dokaz. Promatramo vjerojatnost da $G(n, \frac{1}{2})$ ima bar jednu kliku veličine k, tj. $\mathbb{P}[X_k^{(n)} \neq 0]$. Budući da je $X_k^{(n)}$ nenegativna cjelobrojna slučajna varijabla, iz Markovljeve nejednakosti i (3.5) slijedi

$$\begin{aligned}
\mathbb{P}[X_k^{(n)} \neq 0] = \mathbb{P}[X_k^{(n)} \geq 1] \leq \mathbb{E}[X_k^{(n)}] &= \binom{n}{k} 2^{-\binom{k}{2}} \\
&\leq n^k 2^{-\binom{k}{2}} = (n 2^{-\frac{k-1}{2}})^k \leq (n 2^{-(1+\epsilon)\log_2 n} \sqrt{2})^k \\
&= (n n^{-(1+\epsilon)} \sqrt{2})^k = (n^{-\epsilon} \sqrt{2})^k \to 0
\end{aligned}$$

kada $n \to \infty$. □

Donja ograda za veličinu klike u slučajnom grafu

Za $0 < \epsilon < \frac{1}{2}$, pretpostavimo da je $k = k(n)$ funkcija takva da vrijedi

$$2(1-2\epsilon)\log_2 n \leq k(n) \leq 2(1-\epsilon)\log_2 n. \tag{3.6}$$

Primijetimo da je drugi moment slučajne varijable $X_k^{(n)}$ jednak očekivanju slučajne varijable $(X_k^{(n)})^2$. Očito je $\mathbb{P}[(X_k^{(n)})^2 = l] = \mathbb{P}[X_k^{(n)} = j]$ ako je $l = j^2$, za neki $j \in \mathbb{N}_0$, inače je $\mathbb{P}[(X_k^{(n)})^2 = l] = 0$. Nadalje, $(X_k^{(n)})^2$ je slučajna varijabla koja predstavlja broj uređenih parova (A, B), gdje su A i B klike u grafu $G(n, \frac{1}{2})$ veličine k. Slično kao u prethodnom poglavlju, za dani uređeni par skupova vrhova (A, B), gdje su A i B veličine k, označimo sa $X_{(A,B)}^{(n)}$ slučajnu varijablu koja detektira broj parova klika u grafu $G(n, \frac{1}{2})$, sa skupovima vrhova A i B. Tada je

$$(X_k^{(n)})^2 = \sum_{|A|=k} \sum_{|B|=k} X_{(A,B)}^{(n)}$$

pa zbog linearnosti matematičkog očekivanja vrijedi

$$\mathbb{E}[(X_k^{(n)})^2] = \sum_{|A|=k} \sum_{|B|=k} \mathbb{E}[X_{(A,B)}^{(n)}].$$

Preostaje još izračunati $\mathbb{E}[X_{(A,B)}^{(n)}]$, što ovisi o skupu $A \cap B$. U slučaju da skup $A \cap B$ ima j elemenata, tada je broj parova vrhova u A jednak $\binom{k}{2}$, a isto vrijedi i za skup B, dok je u presjeku $\binom{j}{2}$ parova vrhova. Formula uključivanja i isključivanja tada kaže da je ukupan broj parova vrhova $2\binom{k}{2} - \binom{j}{2}$, a da bi A i B činili par klika, svi ti parovi moraju biti međusobno spojeni. Zbog nezavisnosti i definicije slučajne varijable $X_{(A,B)}^{(n)}$ tada vrijedi

$$\mathbb{E}[X_{(A,B)}^{(n)}] = \mathbb{P}[X_{(A,B)}^{(n)} = 1] = 2^{-2\binom{k}{2}+\binom{j}{2}}.$$

Broj načina da odaberemo skupove A i B čiji presjek $A \cap B$ ima j elemenata je $\binom{n}{k}\binom{k}{j}\binom{n-k}{k-j}$. Naime, prvo biramo vrhove skupa A, što možemo učiniti na $\binom{n}{k}$ načina, zatim na $\binom{k}{j}$ načina biramo vrhove presjeka $A \cap B$ i konačno imamo $\binom{n-k}{k-j}$ načina da odaberemo preostale vrhove skupa B. Zbog linearnosti očekivanja tada vrijedi

$$\begin{aligned}
\mathbb{E}[(X_k^{(n)})^2] &= \sum_{j=0}^{k} \sum_{|A|=k} \sum_{\substack{B_1 \subseteq A \\ |B_1|=j}} \sum_{\substack{B_2 \cap A = \emptyset \\ |B_2|=k-j}} \mathbb{E}\left[X_{(A,B_1 \cup B_2)}^{(n)}\right] \\
&= \sum_{j=0}^{k} \binom{n}{k}\binom{k}{j}\binom{n-k}{k-j} 2^{-2\binom{k}{2}+\binom{j}{2}} \qquad (3.7)
\end{aligned}$$

pa zbog jednakosti (3.5), (3.7) te teorema 3.1.1 vrijedi

$$\begin{aligned}
\mathbb{P}[X_k^{(n)} = 0] &\leq \frac{Var[(X_k^{(n)})^2]}{\mathbb{E}[X_k^{(n)}]^2} = \frac{\mathbb{E}[(X_k^{(n)})^2]}{\mathbb{E}[X_k^{(n)}]^2} - 1 \\
&= \frac{\sum_{j=0}^{k} \binom{n}{k}\binom{k}{j}\binom{n-k}{k-j} 2^{-2\binom{k}{2}+\binom{j}{2}}}{\left(\binom{n}{k} 2^{-\binom{k}{2}}\right)^2} - 1 \\
&= \frac{\sum_{j=0}^{k} \binom{k}{j}\binom{n-k}{k-j} 2^{\binom{j}{2}}}{\binom{n}{k}} - 1.
\end{aligned}$$

Ako uvedemo oznaku

$$\alpha_j^{(n)} := \frac{\binom{k}{j}\binom{n-k}{k-j} 2^{\binom{j}{2}}}{\binom{n}{k}},$$

tada znamo da vrijedi

$$\mathbb{P}[X_k^{(n)} = 0] \leq \sum_{j=0}^{k} \alpha_j^{(n)} - 1. \tag{3.8}$$

Naglasimo da za $m, n \in \mathbb{N}$ takve da je $m < n$ vrijedi $\binom{m}{n} = 0$. Tada je

$$\alpha_0^{(n)} = \frac{\binom{k}{0}\binom{n-k}{k} 2^{\binom{0}{2}}}{\binom{n}{k}} = \frac{\binom{n-k}{k}}{\binom{n}{k}} \to 1 \qquad \text{kada } n \to \infty,$$

što povlači $\alpha_0^{(n)} - 1 \to 0$ kada $n \to \infty$. Zbog

$$\begin{aligned}
k\alpha_1^{(n)} &= k\frac{\binom{k}{1}\binom{n-k}{k-1} 2^{\binom{1}{2}}}{\binom{n}{k}} = k\frac{k\binom{n-k}{k-1} 2^0}{\frac{n}{k}\binom{n-1}{k-1}} \\
&= \frac{k^3}{n} \frac{\binom{n-k}{k-1}}{\binom{n-1}{k-1}} = \frac{k^3}{n} \frac{\frac{(n-k)\dots(n-2k+2)}{(k-1)!}}{\frac{(n-1)\dots(n-k+1)}{(k-1)!}} \\
&= \frac{k^3}{n} \frac{(n-k)\dots(n-2k+2)}{(n-1)\dots(n-k+1)} \\
&\leq \frac{8\log_2^3 n}{n} \frac{(n-k)\dots(n-2k+2)}{(n-1)\dots(n-k+1)},
\end{aligned}$$

a budući da n raste puno brže nego bilo koja potencija od $\log_2 n$, slijedi $\frac{8\log_2^3 n}{n} \to \infty$. Očito je da vrijedi $\frac{(n-k)\dots(n-2k+2)}{(n-1)\dots(n-k+1)} \to 1$ pa zaključujemo $k\alpha_1^{(n)} \to 0$.

Pokažimo sada da vrijedi

$$\alpha_k^{(n)} = \frac{\binom{k}{k}\binom{n-k}{0}2^{\binom{k}{2}}}{\binom{n}{k}} = \frac{2^{\binom{k}{2}}}{\binom{n}{k}} = \frac{1}{\binom{n}{k}2^{-\binom{k}{2}}} = \frac{1}{\mathbb{E}[X_k^{(n)}]} \to 0.$$

Kako bismo dokazali navedenu tvrdnju, dovoljno je pokazati da $\mathbb{E}[X_k^{(n)}] \to \infty$. Primijetimo prvo da za dovoljno velike n vrijedi

$$(1-\epsilon)\log_2 n \leq \log_2 n \leq \frac{n}{4},$$

što povlači

$$n-k \geq n - 2(1-\epsilon)\log_2 n \geq \frac{n}{2},$$

a iz toga slijedi

$$\frac{n-k}{k} \geq \frac{n}{2k}. \tag{3.9}$$

Osim toga, za dovoljno velike n je

$$n^{\frac{\epsilon}{2}} \geq 8\log_2^2 n.$$

Iz svega navedenog slijedi

$$\begin{aligned}
\mathbb{E}[X_k^{(n)}] &= \binom{n}{k}2^{-\binom{k}{2}} = \frac{n(n-1)\dots(n-k+1)}{1\cdot 2\cdot\dots\cdot k}2^{-\frac{k(k-1)}{2}} \\
&\geq \frac{(n-k)^k}{k^k}2^{-\frac{k^2}{2}} = \left(\frac{n-k}{k}2^{-\frac{k}{2}}\right)^k \\
&\geq \left(\frac{n}{2k}2^{-(1-\epsilon)\log_2 n}\right)^k = \left(\frac{n}{2k}\frac{1}{n^{1-\epsilon}}\right)^k \\
&= \left(\frac{n^\epsilon}{2k}\right)^k = \left(\frac{n^{\frac{\epsilon}{2}}}{2k^2}kn^{\frac{\epsilon}{2}}\right)^k \\
&\geq \left(\frac{n^{\frac{\epsilon}{2}}}{8\log_2^2 n}kn^{\frac{\epsilon}{2}}\right)^k \geq \left(kn^{\frac{\epsilon}{2}}\right)^k \\
&= k^k n^{\frac{\epsilon k}{2}} \to \infty,
\end{aligned}$$

čime smo dokazali tvrdnju da $\alpha_k^{(n)} \to 0$ kada $n \to \infty$.

Primijetimo da čak i $\frac{\mathbb{E}[X_k^{(n)}]}{k} \to \infty$, što povlači $k\alpha_k^{(n)} \to 0$. Dokažemo li još

$$\max\{\alpha_2^{(n)}, \ldots, \alpha_{k-1}^{(n)}\} \leq \max\{\alpha_1^{(n)}, \alpha_k^{(n)}\},$$

tada će slijediti

$$\sum_{i=1}^{k} \alpha_i^{(n)} \leq \alpha_1^{(n)} + k \max(\alpha_1^{(n)}, \alpha_k^{(n)}) + \alpha_k^{(n)} \leq \alpha_1^{(n)} + k(\alpha_1^{(n)} + \alpha_k^{(n)}) + \alpha_k^{(n)} \to 0.$$

Da bismo pokazali da vrijedi $\max\{\alpha_2^{(n)}, \ldots, \alpha_{k-1}^{(n)}\} \leq \max\{\alpha_1^{(n)}, \alpha_k^{(n)}\}$, najprije definiramo $\beta_j^{(n)} = \frac{\alpha_{j+1}^{(n)}}{\alpha_j^{(n)}}$ za $1 \leq j \leq k$. Vrijedi da je

$$\begin{aligned}
\beta_j^{(n)} = \frac{\alpha_{j+1}^{(n)}}{\alpha_j^{(n)}} &= \frac{\binom{k}{j+1}\binom{n-k}{k-j-1}}{\binom{k}{j}\binom{n-k}{k-j}} 2^{\binom{j+1}{2}-\binom{j}{2}} \\
&= \frac{\frac{k!}{(j+1)!(k-j-1)!}\frac{(n-k)!}{(n-2k+j+1)!(k-j-1)!}}{\frac{k!}{(k-j)!j!}\frac{(n-k)!}{(n-2k+j)!(k-j)!}} 2^{\binom{j+1}{2}-\binom{j}{2}} \\
&= \frac{(k-j)!j!(n-2k+j)!(k-j)!}{(j+1)!(k-j-1)!(n-2k+j+1)!(k-j-1)!} 2^{j\frac{(j+1)-(j-1)}{2}} \\
&= \frac{k-j}{j+1}\frac{k-j}{n-2k+j+1} 2^j.
\end{aligned}$$

Sada za $2 \leq j \leq k-1$ računamo sljedeći omjer:

$$\frac{\alpha_{j-1}^{(n)}\alpha_{j+1}^{(n)}}{(\alpha_j^{(n)})^2} = \frac{\beta_j^{(n)}}{\beta_{j-1}^{(n)}} = 2\Big(1 - \frac{1}{k-j+1}\Big)^2\Big(1 - \frac{1}{j+1}\Big)\Big(1 - \frac{1}{n-2k+j+1}\Big).$$

Za dovoljno veliki n, radi $k - j + 1 \geq 2$, $j + 1 \geq 3$ i pretpostavke (3.6), imamo da je gornji izraz veći od 1, što daje $(\alpha_j^{(n)})^2 \leq \alpha_{j-1}^{(n)}\alpha_{j+1}^{(n)}$. U tom slučaju se kaže se da je konačan niz $\alpha_1^{(n)}, \alpha_2^{(n)}, \ldots, \alpha_{k-1}^{(n)}, \alpha_k^{(n)}$ *log-konveksan*, a jednostavna posljedica je da on najveću vrijednost postiže svojim prvim ili posljednjim članom. Dakle, doista je $\max\{\alpha_2^{(n)}, \ldots, \alpha_{k-1}^{(n)}\} \leq \max\{\alpha_1^{(n)}, \alpha_k^{(n)}\}$. Zaključujemo da zbog (3.8) i teorema o sendviču

$$\lim_{n\to\infty} \mathbb{P}[X_k^{(n)} = 0] = 0, \quad \text{tj.} \lim_{n\to\infty} \mathbb{P}[X_k^{(n)} \neq 0] = 1$$

Ovime smo dokazali sljedeći teorem.

Teorem 3.2.3. *Neka je $0 < \epsilon < \frac{1}{2}$ i neka niz $k = k(n)$ zadovoljava (3.6). Tada vjerojatnost da slučajni graf $G(n, \frac{1}{2})$ sadrži kliku veličine k teži prema 1. Dakle, $G(n, \frac{1}{2})$ s velikom vjerojatnosti sadrži kliku veličine $k(n)$.*

Malo drugačiji dokaz (ali koji se oslanja na iste prethodne račune), mogli smo dati korištenjem korolara 3.1.6.

Dokaz. Za svaki skup vrhova A veličine k neka je $X_A^{(n)}$ slučajna varijabla koja detektira čine li elementi od A kliku u grafu G. Neka je

$$X_k^{(n)} = \sum_{|A|=k} X_A^{(n)},$$

tako da vrijedi da je $\omega(G) \geq k$ ako i samo ako je $X_k^{(n)} > 0$. Promatramo $\triangle_n^*$ iz korolara 3.1.6. Za fiksirani skup vrhova A i proizvoljan skup vrhova B uvodimo oznaku $A \sim B$ ako i samo ako vrijedi $|A \cap B| = i$, za $2 \leq i \leq k - 1$. Tada je

$$\triangle_n^* = \sum_{i=2}^{k-1} \binom{k}{i}\binom{n-k}{k-i} 2^{\binom{i}{2}-\binom{k}{2}},$$

iz čega slijedi

$$\frac{\triangle_n^*}{\mathbb{E}[X_k^{(n)}]} = \sum_{i=2}^{k-1} \alpha_i^{(n)},$$

pri čemu je

$$\alpha_i^{(n)} = \frac{\binom{k}{i}\binom{n-k}{k-i} 2^{\binom{i}{2}}}{\binom{n}{k}}.$$

Iz rasprave s početka odjeljka slijedi da

$$\mathbb{E}[X_k^{(n)}] \to \infty \quad \text{i} \quad \frac{\triangle_n^*}{\mathbb{E}[X_k^{(n)}]} \to 0.$$

Iz korolara 3.1.6 sada opet slijedi tvrdnja teorema. □

Obzirom da je $\omega(G) \geq k$ ako i samo ako G sadrži kliku veličine k, iz teorema 3.2.2 i teorema 3.2.3 posebno slijedi

$$\lim_{n\to\infty} \mathbb{P}[\omega(G(n, 1/2)) \leq 2(1+\epsilon)\log_2 n] = 1, \quad \lim_{n\to\infty} \mathbb{P}[\omega(G(n, 1/2)) \geq 2(1-\epsilon)\log_2 n] = 1$$

za svaki $0 < \epsilon < 1/2$, što zajedno možemo pisati

$$\lim_{n\to\infty} \mathbb{P}[|\omega(G(n, 1/2)) - 2\log_2 n| \leq 2\epsilon \log_2 n] = 1.$$

Time smo dobili sljedeći rezultat.

Teorem 3.2.4. *Veličina najveće klike slučajnog grafa $G = G(n, 1/2)$ zadovoljava $\omega(G) \sim 2\log_2 n$ s velikom vjerojatnosti, tj. omjer $\frac{\omega(G)}{2\log_2 n}$ po vjerojatnosti konvergira prema 1.*

3.3 Prijelazne funkcije nekih svojstava

U ovom dijelu proučavat ćemo prijelazne funkcije za neka svojstva slučajnih grafova. Prisjetimo se da za funkciju r, tj. $n \mapsto r(n)$ kažemo da je prijelazna funkcija za svojstvo A grafa $G(n, p_n)$ ako za $p = p(n) \ll r(n)$ graf $G(n, p_n)$ s velikom vjerojatnosti ima svojstvo A, dok za $p = p(n) \gg r(n)$ graf $G(n, p_n)$ s velikom vjerojatnosti nema svojstvo A, ili obrnuto.

Prvi primjer vezan je za broj vrhova u maksimalnoj kliki grafa G, kojeg smo i do sad označavali $\omega(G)$.

Teorem 3.3.1. *Svojstvo $\omega(G) \geq 4$ ima prijelaznu funkciju $r(n) = n^{-\frac{2}{3}}$.*

Dokaz. Za svaki skup vrhova S grafa $G(n, p_n)$ veličine 4, neka je $A_S^{(n)}$ događaj

$$A_S^{(n)} = \{\text{vrhovi u skupu S čine kliku u grafu } G(n, p_n)\}$$

te neka je $X_S^{(n)}$ Bernoullijeva slučajna varijabla koja detektira vrijedi li za graf $G(n, p_n)$ događaj $A_S^{(n)}$, tj. $X_S^{(n)} = \mathbf{1}_{A_S^{(n)}}$ je indikatorska varijabla događaja $A_S^{(n)}$. Iz definicije očekivanja Bernoullijeve slučajne varijable tada slijedi

$$\mathbb{E}[X_S^{(n)}] = \mathbb{P}[A_S^{(n)}] = p_n^6,$$

budući da svih $\binom{4}{2} = 6$ bridova mora biti sadržano u grafu $G(n, p_n)$. Stavimo još

$$X^{(n)} = \sum_{|S|=4} X_S^{(n)},$$

tako da je $X^{(n)}$ slučajna varijabla koja grafu $G(n, p_n)$ pridružuje broj klika veličine 4. Tada je $X^{(n)} > 0$ ako i samo ako je $\omega(G(n, p_n)) \geq 4$. Zbog linearnosti matematičkog očekivanja vrijedi

$$\mathbb{E}[X^{(n)}] = \sum_{|S|=4} \mathbb{E}[X_S^{(n)}] = \binom{n}{4}\mathbb{E}[X_S^{(n)}] = \binom{n}{4}p_n^6 \sim \frac{n^4 p_n^6}{24}.$$

Sada vidimo da za $p_n \ll n^{-\frac{2}{3}}$ vrijedi $\mathbb{E}[X^{(n)}] = o(1)$ pa zbog rasprave u odjeljku 3.1 slijedi da je $X^{(n)} = 0$ s velikom vjerojatnosti.

Pretpostavimo da je $p_n \gg n^{-\frac{2}{3}}$, tako da $\mathbb{E}[X^{(n)}] \to \infty$. Ako bismo pokazali da za $\triangle_n^*$, definirano sa (3.2), vrijedi $\triangle_n^* = o(\mathbb{E}[X^{(n)}])$, tada bi iz korolara 3.1.6 slijedilo $X^{(n)} > 0$ s velikom vjerojatnosti. Primijetimo da iz definicije od $X_S^{(n)}$ slijedi da su to simetrične slučajne varijable. Za skupove vrhova S i T, takve da je $S \neq T$ te S i T imaju zajedničke bridove, pisat ćemo $S \sim T$. Skupovi S i T imaju zajedničke bridove ako i samo ako je $|S \cap T| = 2$ ili $|S \cap T| = 3$, pri čemu za fiksirani skup vrhova S postoji $O(n^2)$ skupova vrhova T sa svojstvom $|S \cap T| = 2$ i za svaki takav skup je uvjetna vjerojatnost $\mathbb{P}[A_T^{(n)} | A_S^{(n)}] = p_n^5$, budući da svih 5 preostalih bridova čiji su vrhovi u T moraju ležati u grafu $G(n, p_n)$. Slično, postoji $O(n)$ skupova vrhova T sa svojstvom $|S \cap T| = 3$ i za svaki takav skup je uvjetna vjerojatnost $\mathbb{P}[A_T^{(n)} | A_S^{(n)}] = p_n^3$. Odavde slijedi

$$\triangle_n^* = O(n^2 p_n^5) + O(n p_n^3) = o(n^4 p_n^6) = o(\mathbb{E}[X^{(n)}]). \tag{3.10}$$

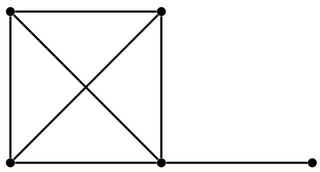
Naime, iz pretpostavke $p_n \gg n^{-\frac{2}{3}}$ slijedi da $p_n n^{\frac{2}{3}} \to \infty$, stoga vrijedi

$$\frac{n^2 p_n^5 + n p_n^3}{n^4 p_n^6} \to 0,$$

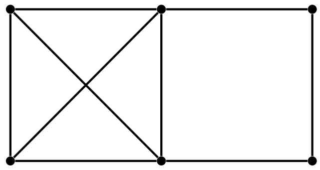
iz čega slijedi druga jednakost u (3.10). Korolar 3.1.6 sada povlači $X^{(n)} > 0$ s velikom vjerojatnosti, tj. s velikom vjerojatnosti postoji klika veličine 4, što je i trebalo pokazati. □

Definicija 3.3.2. *Neka je H graf sa v vrhova i e bridova. Definiramo* gustoću grafa *H, u oznaci $\rho(H)$, kao $\rho(H) = \frac{e}{v}$. Reći ćemo da je graf H* uravnotežen *ako za svaki podgraf H' grafa H vrijedi $\rho(H') \leq \rho(H)$. Graf H je* strogo uravnotežen *ako za svaki pravi podgraf H' grafa H vrijedi $\rho(H') < \rho(H)$.*

Primjer 3.3.3. *Lako se pokaže da je graf K_4 strogo uravnotežen. Općenito, svaki graf K_k, $k \in \mathbb{N}$ je strogo uravnotežen.*



Graf na slici iznad nije uravnotežen jer ima gustoću $\frac{7}{5}$, dok podgraf K_4 ima gustoću $\frac{3}{2}$, istu kao i cijeli graf.



Graf na slici iznad je uravnotežen, ali nije strogo uravnotežen jer njegov pravi podgraf K_4 ima gustoću $\frac{3}{2}$.

Teorem 3.3.4. *Neka je H uravnotežen graf sa v vrhova i e bridova, $G(n, p_n)$ slučajan graf te neka je $A = A^{(n)}$ događaj $\{$Graf H je podgraf grafa $G(n, p_n)\}$. Tada je $r(n) = n^{-\frac{v}{e}}$ prijelazna funkcija za niz događaja $A^{(n)}$, tj. za svojstvo sadržavanja podgrafa izomorfnog s H.*

Dokaz. Ideja dokaza je slična kao kod dokaza teorema 3.3.1. Za svaki skup vrhova S veličine v neka je $A_S^{(n)}$ događaj da je graf H podgraf od $G|_S$. Takav podgraf H sa v vrhova možemo smjestiti u podgraf $G|_S$ na najviše $v!$ načina, a vjerojatnost pojavljivanja svakog načina je (zbog nezavisnosti) jednaka p_n^e. Iz toga slijedi

$$p_n^e \leq \mathbb{P}[A_S^{(n)}] \leq v! p_n^e.$$

Označimo sa $X_S^{(n)}$ Bernoullijevu slučajnu varijablu koja detektira vrijedi li događaj $A_S^{(n)}$. Iz definicije očekivanja Bernoullijeve slučajne varijable imamo

$$\mathbb{E}[X_S^{(n)}] = \mathbb{P}[A_S^{(n)}].$$

Stavimo još

$$X^{(n)} = \sum_{|S|=v} X_S^{(n)},$$

tako da je $X^{(n)} > 0$ ako i samo ako vrijedi $A^{(n)}$. Zbog linearnosti matematičkog očekivanja vrijedi

$$\mathbb{E}[X^{(n)}] = \sum_{|S|=v} \mathbb{E}[X_S^{(n)}] = \binom{n}{v}\mathbb{E}[X_S^{(n)}] = \binom{n}{v}\mathbb{P}[A_S^{(n)}] \sim n^v p_n^e. \tag{3.11}$$

Ako je $p_n \ll n^{-\frac{v}{e}}$, tj. ako $pn^{\frac{e}{v}} \to 0$, tada iz teorema o sendviču i (3.11) slijedi $\mathbb{E}[X^{(n)}] = o(1)$, pa iz odjeljka 3.1 slijedi da je $X^{(n)} = 0$ s velikom vjerojatnosti. Pretpostavimo li da je $p_n \gg n^{-\frac{v}{e}}$, tj. $\mathbb{E}[X^{(n)}] \to \infty$ te ako za $\triangle_n^*$ definirane sa (3.2) vrijedi $\triangle_n^* = o(\mathbb{E}[X^{(n)}])$, tada bi iz korolara 3.1.6 slijedilo $X^{(n)} > 0$ s velikom vjerojatnosti. Budući da svi skupovi vrhova S veličine v izgledaju jednako, iz definicije $X_S^{(n)}$ slijedi da su $X_S^{(n)}$ simetrične slučajne varijable. Za skupove vrhova S i T, takve da je $S \neq T$ te S i T imaju zajedničke bridove, pisat ćemo $S \sim T$. Skupovi S i T imaju zajedničke bridove ako i samo ako je $|S \cap T| = i$, za $2 \leq i \leq v - 1$. Ako fiksiramo skup S, tada možemo računati

$$\triangle_n^* = \sum_{T \sim S} \mathbb{P}[A_T^{(n)} | A_S^{(n)}] = \sum_{i=2}^{v-1} \sum_{|T \cap S|=i} \mathbb{P}[A_T^{(n)} | A_S^{(n)}].$$

Za svako i iz gornje sume postoji $O(n^{v-i})$ izbora za skup T. Za fiksirane skupove S i T izračunajmo sada uvjetnu vjerojatnost $\mathbb{P}[A_T^{(n)}|A_S^{(n)}]$. Primijetimo prvo da postoji $O(1)$ mogućih kopija grafa H na skupu vrhova T. Označimo sa H' inducirani podgraf grafa H sa vrhovima u presjeku skupova S i T (broj takvih vrhova je i) te označimo sa e' broj bridova od H'. Iz pretpostavke o uravnoteženosti grafa H slijedi

$$\frac{e'}{i} \leq \frac{e}{v},$$

tj. postoji najviše $\frac{ie}{v}$ bridova čija su oba kraja u skupu S, dok je preostalih bridova grafa H barem $e - \frac{ie}{v}$. Iz svega navedenog zaključujemo da je

$$\mathbb{P}[A_T^{(n)}|A_S^{(n)}] = O(p_n^{e-\frac{ie}{v}}),$$

pa dalje računamo

$$\begin{aligned}
\triangle_n^* &= \sum_{i=2}^{v-1} O(n^{v-i} p_n^{e-\frac{ie}{v}}) \\
&= \sum_{i=2}^{v-1} O((n^v p_n^e)^{1-\frac{i}{v}}) \\
&= \sum_{i=2}^{v-1} o(n^v p_n^e) \\
&= o(\mathbb{E}[X^{(n)}]),
\end{aligned}$$

gdje treća jednakost slijedi iz pretpostavke $p_n \gg n^{-\frac{v}{e}}$, što znači da $n^v p_n^e \to \infty$, a zadnja nejednakost iz (3.11). Sada iz korolara 3.1.6 slijedi tvrdnja teorema. □

Propozicija 3.3.5. *Koristeći notaciju iz teorema 3.3.4, ako H nije uravnotežen graf, tada $r(n) = n^{-\frac{v}{e}}$ nije prijelazna funkcija za svojstvo A.*

Dokaz. Dokaz propozicije sličan je dokazu teorema 3.3.4, no ovdje ćemo pronaći funkciju $p_n \gg n^{-\frac{v}{e}}$ takvu da svojstvo A vrijedi s malom vjerojatnosti.

Iz pretpostavke da graf H nije uravnotežen slijedi da postoji podgraf H_1 grafa H sa v_1 vrhova i e_1 bridova, tako da vrijedi $\frac{e_1}{v_1} > \frac{e}{v}$. Neka je α takav da je $\frac{v_1}{e_1} < \alpha < \frac{v}{e}$ te stavimo $p_n = n^{-\alpha}$. Očito je tako definirana funkcija $p_n \gg n^{-\frac{v}{e}}$. Slično kao u prethodnom dokazu, ako računamo očekivani broj kopija grafa H_1, zbog $n^{v_1} p_n^{e_1} = n^{v_1 - \alpha e_1} \to 0$ dobivamo da je on $o(1)$. Odavde zaključujemo da graf $G(n, p_n)$ s velikom vjerojatnosti ne sadrži kopije grafa H_1, a iz toga slijedi das velikom vjerojatnosti ne sadrži ni kopije grafa H. □

Napomena 3.3.6. *Za dani graf H, neka je H_1 podgraf sa maksimalnom gustoćom $\rho(H_1) = \frac{e_1}{v_1}$. U slučaju kad je H uravnotežen neka je $H_1 = H$. Može se pokazati da je prijelazna funkcija za svojstvo A iz teorema 3.3.4 $r(n) = n^{-\frac{v_1}{e_1}}$, a dokaz se može pronaći u radu Erdősa i Rényia, [3].*

Teorem 3.3.7. *Neka je H strogo uravnotežen graf s v vrhova, e bridova i a automorfizama. Neka je $X^{(n)}$ slučajna varijabla koja broji kopije od H u slučajnom grafu $G(n, p_n)$. Ako je $p_n \gg n^{-\frac{v}{e}}$, tada s velikom vjerojatnosti vrijedi*

$$X^{(n)} \sim \frac{n^v p^e}{a}.$$

Dokaz. Označimo vrhove grafa H sa $1, \dots, v$ te neka je za svaku uređenu v-torku $(x_1, \dots, x_v)$ različitih vrhova grafa $G(n, p_n)$, $A^{(n)}_{x_1,\dots,x_v}$ događaj da podgraf grafa $G(n, p_n)$ induciran vrhovima $x_1, \dots, x_v$ sadrži kopiju od H. Preciznije, $A^{(n)}_{x_1,\dots,x_v}$ je događaj

$$\Big\{\{i, j\} \in E(H) \Rightarrow \{x_i, x_j\} \in E(G)\Big\}.$$

Neka je $I^{(n)}_{x_1,\dots,x_v}$ Bernoullijeva slučajna varijabla koja detektira vrijedi li događaj $A^{(n)}_{x_1,\dots,x_v}$, te definiramo klasu ekvivalencije na v-torkama $(x_1, \dots, x_v)$. Za $(x_1, \dots, x_v)$ i $(y_1, \dots, y_v)$ vrijedi $(x_1, \dots, x_v) \equiv (y_1, \dots, y_v)$ ako postoji automorfizam σ od H na skupu vrhova $V(H)$ tako da $y_{\sigma(i)} = x_i$, $1 \le i \le v$. Sumiramo slučajne varijable $I_{x_1,\dots,x_v}$ po klasama, tj.

$$X^{(n)} = \sum I^{(n)}_{x_1,\dots,x_v},$$

gdje u sumu ulazi po jedan reprezentant svake klase. Tada je $X^{(n)}$ slučajna varijabla koja broji kopije od H u G. Obzirom da navedena suma ima $\frac{\binom{n}{v}v!}{a}$ članova, zbog linearnosti matematičkog očekivanja je

$$\mathbb{E}[X^{(n)}] = \frac{\binom{n}{v}v!}{a}\mathbb{E}[I^{(n)}_{x_1,\dots,x_v}] = \frac{\binom{n}{v}v!p_n^e}{a} \sim \frac{n^v p_n^e}{a}. \tag{3.12}$$

Zbog pretpostavke $p_n \gg n^{-\frac{v}{e}}$ i (3.12) slijedi $\mathbb{E}[X^{(n)}] \to \infty$. Dovoljno je još pokazati $\triangle_n^* = o(\mathbb{E}[X^{(n)}])$. Za fiksiranu v-torku $x_1, \dots, x_v$, izračunajmo

$$\triangle_n^* = \sum_{(y_1,\dots,y_v)\sim(x_1,\dots,x_v)} \mathbb{P}[A^{(n)}_{(y_1,\dots,y_v)} | A^{(n)}_{(x_1,\dots,x_v)}], \tag{3.13}$$

gdje suma u (3.13) ide po svim v-torkama za koje vrijedi $(y_1, \dots, y_v) \neq (x_1, \dots, x_v)$, ali događaji $A^{(n)}_{(y_1,\dots,y_v)}$ i $A^{(n)}_{(x_1,\dots,x_v)}$ nisu nezavisni. Razlikujemo dva slučaja. U prvom slučaju, neka je $\{x_1, \ldots, x_v\} = \{y_1, \ldots, y_v\}$. Broj takvih članova u sumi (3.13) je $\frac{v!}{a} = O(1)$, a zbog ograničenosti (uvjetne) vjerojatnosti, takvi sumirani članovi su reda $O(1) = o(\mathbb{E}[X^{(n)}])$. U drugom slučaju, pretpostavimo $\{x_1, \ldots, x_v\} \neq \{y_1, \ldots, y_v\}$, tj. $|\{x_1, \ldots, x_v\} \cap \{y_1, \ldots, y_v\}| = i$, $2 \leq i \leq v - 1$. Ponovimo li dokaz teorema 2.3.3, zaključujemo kako su takvi sumirani članovi reda $o(E[X^{(n)}])$. Iz navedenog slijedi i da je $\triangle_n^* = o(\mathbb{E}[X^{(n)}])$ pa zbog korolara 3.1.6 slijedi tvrdnja teorema. $\square$

Poglavlje 4

Erdős-Rényijeva promjena faze

U svom su radu [3] Erdős i Rényi posebnu pažnju posvetili slučajnim grafovima $G(n, N(n))$ u slučaju kada je broj bridova $N(n)$ blizu $\frac{n}{2}$. Primijetili su da u slučaju kada vrijedi $\frac{e}{n} \to c$, $c < \frac{1}{2}$, veličina najveće komponente je reda $\log n$, za $\frac{e}{n} \to c$, $c \sim \frac{1}{2}$ reda $n^{\frac{2}{3}}$, a za $\frac{e}{n} \to c$, $c > \frac{1}{2}$ veličina najveće komponente je reda n. Skok veličine najveće komponente kada limes prijeđe vrijednost $\frac{1}{2}$ činjenica je koja je dobila puno pažnje i tome ćemo se posvetiti u ovom poglavlju. Taj događaj zovemo Erdős-Rényijeva promjena faze.

4.1 Uvod i osnovni pojmovi

Iako su u svojim ranim originalnim člancima Erdős i Rényi predstavili svoj model slučajnog grafa sa n vrhova i $N(n)$ bridova, u oznaci $G(n, N(n))$, mi ćemo se kao i do sad držati modela $G(n, p_n)$, pri čemu za

$$p_n = \frac{N(n)}{\binom{n}{2}},$$

model $G(n, N(n))$ heuristički korespondira sa $G(n, p_n)$. U ovom poglavlju pretpostavljamo da je $p_n = \Theta(n^{-1})$ i pokazujemo da tada dolazi do promjene faze.

Uvodimo više različitih parametrizacija—prvo, *grubu parametrizaciju* definiranu sa

$$p_n = \frac{c}{n}.$$

Bit će zanimljivo promatrati kako se mijenjaju svojstva slučajnog grafa kada je c blizu 1, pa će tako za vrijednosti $c > 1$ i $c < 1$ slučajni grafovi $G(n, p_n)$ imati potpuno različita svojstva. Osim toga, uvodimo i *finu parametrizaciju*

$$p_n = \frac{1}{n} + \lambda n^{-\frac{4}{3}},$$

čija važnost nije a priori jasna, a često ju koristimo i u obliku

$$p_n = \frac{1+\epsilon}{n},$$

ili, kada želimo izbjeći negativne vrijednosti od ϵ,

$$p_n = \frac{1-\epsilon}{n},$$

gdje je $\epsilon = \lambda n^{-\frac{1}{3}}$.

Kada ćemo opisativati svojstva slučajnog grafa $G(n, p_n)$, ponekad ćemo se referirati na kompleksnost komponente.

Definicija 4.1.1. *Za komponentu grafa G sa v vrhova i e bridova reći ćemo da je* kompleksnosti $e - v + 1$. *Komponente kompleksnosti nula i jedan zovemo* jednostavne komponente*, dok komponente kompleksnosti veće od jedan zovemo* složene komponente.

Primijetimo da kompleksnosti nula i jedan odgovaraju redom komponenti bez ciklusa (stablu) i komponenti s točno jednim ciklusom.

Označimo sa $C^{(n)}(v)$ komponentu slučajnog grafa $G(n, p_n)$ koja sadrži vrh v, a veličinu te komponente sa $|C^{(n)}(v)|$. Za svaki vrh v, $|C^{(n)}(v)|$ ima određenu distribuciju, a zbog simetrije grafa $G(n, p_n)$ te su distribucije jednake za svaki vrh v. Nas će zanimati veličina najveće komponente u grafu $G(n, p_n)$. Intuitivno, ako p_n raste, raste i veličina najveće komponente. Nadalje, sa $C_i^{(n)}$ označavat ćemo i-tu najveću komponentu, a sa $L_i^{(n)}$ broj vrhova i-te najveće komponente grafa $G(n, p_n)$. Očito je tada $L_1^{(n)} = \max_v |C^{(n)}(v)|$. Promatrat ćemo $L_1^{(n)}$ i $L_2^{(n)}$ te njihov međusobni odnos, točnije, jesu li blizu po veličini.

Uz pretpostavku $p_n = \Theta(n^{-1})$, ovisno o definiciji od p_n, graf $G(n, p_n)$ imat će različita svojstva. Razlikujemo pet slučaja—jako ispod kritičnog, slabo ispod kritičnog, kritični, slabo iznad kritičnog, jako iznad kritičnog. Ukratko opisujemo svaki od tih slučajeva. Za svaku od tvrdnji iz nastavka zapravo želimo reći da vrijedi s velikom vjerojatnosti, tj. da je ispunjena asimptotski kada $n \to \infty$.

Jako ispod kritičnog

Ovaj slučaj opisujemo pomoću grube parametrizacije $p_n = \frac{c}{n}$, $c < 1$, npr. $p_n = \frac{1}{2n}$.

- Sve komponente su jednostavne, tj. složenosti nula ili jedan.
- $L_1^{(n)} = \Theta(\ln n)$.
- $L_k^{(n)} \sim L_1^{(n)}$, za svaki k.

Slabo ispod kritičnog

Ovdje koristimo finu parametrizaciju $p_n = \frac{1-\epsilon}{n}$, $\epsilon = \lambda n^{-\frac{1}{3}}$. Pretpostavka je da je $\epsilon = o(1)$ te $\lambda(n) = \lambda \to \infty$, npr. $p_n = \frac{1}{n} - n^{-\frac{4}{3}} n^{0.01}$.

- Sve komponente su jednostavne, tj. složenosti nula ili jedan.
- $L_1^{(n)} = \Theta(\epsilon^{-2} \ln \lambda) = \Theta(n^{\frac{2}{3}} \lambda^{-2} \ln \lambda)$.
- $L_k^{(n)} \sim L_1^{(n)}$, za svaki k.

Kritični

Slično kao u prethodnom slučaju, koristimo finu parametrizaciju $p_n = \frac{1-\epsilon}{n}$, $\epsilon = \lambda n^{-\frac{1}{3}}$, no sada je λ realna konstanta, npr. $p_n = \frac{1}{n} \pm 2n^{-\frac{4}{3}}$.

- Prvih k najvećih komponenti je veličine $L_k^{(n)} = \Theta(n^{\frac{2}{3}})$.

Slabo iznad kritičnog

Koristimo finu parametrizaciju, $p_n = \frac{1+\epsilon}{n}$, $\epsilon = \lambda n^{-\frac{1}{3}}$. Pretpostavljamo da je $\epsilon = o(1)$ te $\lambda(n) = \lambda \to \infty$, npr. $p_n = \frac{1}{n} + n^{-\frac{4}{3}} n^{0.01}$.

- $L_1^{(n)} \sim 2\epsilon n = 2\lambda n^{\frac{2}{3}}$.
- Najveća komponenta je kompleksnosti koja se približava beskonačnosti kako n raste.
- Sve ostale komponente su jednostavne.
- $L_2^{(n)} = \Theta(\epsilon^{-2} \ln \lambda) = \Theta(n^{\frac{2}{3}} \lambda^{-2} \ln \lambda)$.

Primijetimo kako ovdje omjer $\frac{L_1^{(n)}}{L_2^{(n)}}$ teži u beskonačnost kako n raste. U ovom slučaju najveću komponentu zovemo *dominantnom komponentom*.

Jako iznad kritičnog

Ovdje koristimo grubu parametrizaciju $p_n = \frac{c}{n}$, $c > 1$.

- $L_1^{(n)} \sim yn$, gdje je $y = y(c)$ pozitivan realan broj koji zadovoljava jednadžbu

$$e^{-cy} = 1 - y. \tag{4.1}$$

- Najveća komponenta je kompleksnosti koja se približava beskonačnosti kako n raste.

- Sve ostale komponente su jednostavne.
- $L_2^{(n)} = \Theta(\ln n)$.

U ovom slučaju, najveću komponentu zovemo *divovskom komponentom*.

Dokazat ćemo neke od navedenih tvrdnji, dok se ostale mogu pronaći u radovima Bollobása [2] i Jansona [7].

Pretraživanje prvih susjeda

Opisujemo algoritam *Pretraživanja prvih susjeda* (PPS) koji ćemo koristiti kod traženja komponente slučajnog grafa $G(n, p_n)$ koja sadrži vrh v. Pretpostavljamo da su vrhovi grafa nekako poredani te počinjemo od danog vrha v. Razlikovat ćemo tri vrste vrhova: *aktivni*, *neutralni* i *neaktivni* vrhovi, pri čemu su aktivni vrhovi poredani u *slijed Q*. Algoritam počinje u trenutku $t = 0$ i u tom trenutku su svi vrhovi neutralni, osim početnog vrha v koji je aktivan i nalazi se u slijedu Q. U trenutku t, $t \geq 1$, razlikujemo dva slučaja. Ako je slijed Q neprazan, vrh w koji se nalazi na početku slijeda Q iz njega izlazi i postaje neaktivan. Osim toga, pretražujemo po svim neutralnim vrhovima w' i stavljamo na kraj slijeda Q sve vrhove w' za koje vrijedi da su w i w' susjedni. U slučaju da je slijed Q prazan, algoritam prestaje, a svi neaktivni vrhovi zajedno čine komponentu $C^{(n)}(v)$ grafa $G(n, p_n)$. Označimo sa $T^{(n)}$ vrijeme u kojem je algoritam stao. Tada je $T^{(n)} = |C^{(n)}(v)|$. Primijetimo da dani algoritam nalazi samo komponente u obliku stabla, što je dovoljno u našem slučaju kada nas zanima samo veličina komponente, tj. broj vrhova.

Označimo sa $Z_t^{(n)}$ broj vrhova dodanih u slijed Q u trenutku t te s $Y_t^{(n)}$ broj vrhova koji se nalaze u slijedu Q u trenutku t, uz dogovor $Y_0^{(n)} = 1$. Prema algoritmu, u trenutku t, jedan vrh izlazi iz slijeda Q, a $Z_t^{(n)}$ novih ulazi u Q pa vrijedi rekurzija

$$Y_t^{(n)} = Y_{t-1}^{(n)} + Z_t^{(n)} - 1. \tag{4.2}$$

Sa $N_t^{(n)}$ označimo broj neutralnih vrhova u trenutku t. Očito vrijedi $N_0^{(n)} = n - 1$, $N_t^{(n)} = N_{t-1}^{(n)} - Z_t^{(n)}$ ili, budući da u trenutku t imamo t neaktivnih i $Y_t^{(n)}$ aktivnih vrhova, $N_t^{(n)} = n - t - Y_t^{(n)}$. $Z_t^{(n)}$ ćemo pronaći tako da među $N_{t-1}^{(n)}$ neutralnih vrhova provjerimo koji je susjedan sa vrhom w s početka slijeda Q, a kako je vjerojatnost da su dva različita vrha susjedna jednaka p_n, slijedi da je

$$Z_t^{(n)} \sim B(N_{t-1}^{(n)}, p_n) \sim B(n - (t-1) - Y_{t-1}^{(n)}, p_n). \tag{4.3}$$

4.2 Procesi grananja kod grafova i njihova analiza

Galton-Watsonov proces grananja

Oznake za proces grananja u ovom će radu ipak biti malo drugačije od opisanog u poglavlju 1.3. Budući da će nam oznaka n biti rezervirana za broj vrhova u grafu, za vrijeme koristimo oznaku t. Osim toga, niz nezavisnih slučajnih varijabli označavat ćemo sa $\{Z_t, t \geq 1\}$, pri čemu svaka slučajna varijabla Z_t ima distribuciju Z. Proces počinje sa jednom jedinkom $Z_1 = 1$ koja zatim ima potomke $Z_2, \ldots Z_{1+Z_1}$ i koje tim redom ulaze u slijed Q. Potomke t-te jedinke označavat ćemo sa Z_t. Jedinka s početka slijeda Q ima neki broj potomaka koji ulaze u Q, a sama jedinka zatim nestaje. Jedinke će u ovom slučaju predstavljati vrhove grafa. Sa Y_t označavamo broj aktivnih jedinki u trenutku t (ili broj jedinki u slijedu Q) i stavimo $Y_0 = 1$. Označimo sa T ukupan broj jedinki stvorenih u opisanom procesu (u slučaju da proces teče beskonačno je $T = \infty$). Vrijedi rekurzija

$$Y_t = Y_{t-1} + Z_t - 1. \tag{4.4}$$

Razlikujemo dva slučaja.

- $Y_t > 0$ za sve $t \geq 0$. U tom slučaju se Galton-Watsonov proces nastavlja beskonačno i $T = \infty$.
- $Y_t = 0$ za neke $t \geq 0$. U tom slučaju, neka je T najmanji cijeli broj za koji $Y_T = 0$. Galton-Watsonov proces tada staje sa T-tom jedinkom i T je ukupan broj jedinki nastalih u tom procesu.

Nadalje, T ćemo zvati *vrijeme izumiranja*. U ovom radu od interesa će nam biti slučajevi kada Z ima Poissonovu ili binomnu distribuciju i takave procese zvat ćemo *Poissonov proces grananja* i *binomni proces grananja*. Osim toga, uvodimo i proces grananja kod grafova.

Proces grananja kod grafova

Za $n \in \mathbb{N}$, opisujemo *proces grananja kod grafa* G s n vrhova. Neka je $p_n \in [0, 1]$, $Z_1^{(n)}, \ldots, Z_n^{(n)}$ niz nezavisnih slučajnih varijabli, pri čemu $Z_t^{(n)}$ ima binomnu distribuciju s parametrima $N_{t-1}^{(n)}$ i p_n, a parametar $N_t^{(n)}$, $t \geq 0$ je definiran pomoću rekurzije

$$N_t^{(n)} = N_{t-1}^{(n)} - Z_t^{(n)}, \qquad \text{uz početni uvjet } N_0^{(n)} = n - 1.$$

Proces grananja sličan je binomnom procesu grananja, uz razliku da je niz slučajnih varijabli ovdje konačan, a slučajne varijable $Z_t^{(n)}$ nisu jednako distribuirane. Kao i ranije, $Y_t^{(n)}$, $t \geq 0$ označava broj aktivnih jedinki u trenutku t ili duljinu slijeda Q u trenutku t s početnim uvjetom $Y_0^{(n)} = 1$. Tada je $Y_t^{(n)} = Y_{t-1}^{(n)} + Z_t^{(n)} - 1$. Sa $T^{(n)}$ označavamo minimalni cijeli broj t za koji $Y_t^{(n)} = 0$. Zbog $N_t^{(n)} = n - t - Y_t^{(n)}$, u trenutku $T^{(n)}$ je $N_T^{(n)} = n - T^{(n)}$. Primijetimo da vrijedi $1 \leq T^{(n)} \leq n$ i $|C^{(n)}(v)| = T^{(n)}$ pa smo u tom trenutku pronašli veličinu komponente $C^{(n)}(v)$ grafa $G(n, p_n)$. Osim toga, proces grananja kod grafova jednak je algoritmu PPS do trenutka $T^{(n)}$, no za razliku od algoritma PPS, ovaj proces nastavlja se do trenutka n.

Zbog $N_t^{(n)} = N_{t-1}^{(n)} - Z_t^{(n)}$ slijedi da $N_t^{(n)}$ poprima vrijednosti iz skupa $\{0, 1, \ldots, N_{t-1}^{(n)}\}$. Za $k \in \{0, 1, \ldots, N_{t-1}^{(n)}\}$ zbog (4.3) vrijedi

$$\mathbb{P}[N_t^{(n)} = k] = \mathbb{P}[Z_t^{(n)} = N_{t-1}^{(n)} - k] = \binom{N_{t-1}^{(n)}}{N_{t-1}^{(n)} - k} p_n^{N_{t-1}^{(n)} - k} (1 - p_n)^k = \mathbb{P}[B(N_{t-1}^{(n)}, 1 - p_n) = k],$$

pa zaključujemo $N_t^{(n)} \sim N_{t-1}^{(n)} - B(N_{t-1}^{(n)}, p_n) \sim B(N_{t-1}^{(n)}, 1 - p_n)$. Osim toga, pokažimo da za $0 \leq t \leq n$ vrijedi

$$N_t^{(n)} \sim B\Big(n - 1, (1 - p_n)^t\Big). \tag{4.5}$$

Tvrdnja (4.5) dokazuje se matematičkom indukcijom. Primijetimo prvo da vrijedi $N_0^{(n)} = n - 1 \sim B(n - 1, 1)$. Pretpostavimo da je $N_k^{(n)} \sim B\Big(n - 1, (1 - p_n)^k\Big)$ za sve $k = 1, 2, \ldots, t - 1$. Budući da je $N_{t-1}^{(n)} \sim B\Big(n - 1, (1 - p_n)^{t-1}\Big)$, a $N_t^{(n)} \sim B\Big(N_{t-1}^{(n)}, 1 - p_n\Big)$, teorem 1.2.7 tada povlači $N_t^{(n)} \sim B\Big(n - 1, (1 - p_n)^t\Big)$. Pretpostavimo li da je $T^{(n)} = t$, tada nužno vrijedi $N_t^{(n)} = n - t$. Iz navedenih rasprava slijedi zanimljiv rezultat o veličini komponente slučajnog grafa određene zadanim vrhom v.

Propozicija 4.2.1. *Kod slučajnog grafa $G(n, p_n)$ vrijedi*

$$\mathbb{P}[|C^{(n)}(v)| = t] \leq \mathbb{P}\Big[B(n - 1, (1 - p_n)^t) = n - t\Big]$$

ili, ekvivalentno,

$$\mathbb{P}[|C^{(n)}(v)| = t] \leq \mathbb{P}\Big[B(n - 1, 1 - (1 - p_n)^t) = t - 1\Big]. \tag{4.6}$$

Dokaz. Dokaz slijedi iz sljedećih (ne)jednakosti.

$$\begin{aligned}
&\mathbb{P}[|C^{(n)}(v)| = t] \leq \mathbb{P}\Big[|C^{(n)}(v)| \leq t\Big] = \mathbb{P}[Y_t^{(n)} = 0] = \mathbb{P}[N_t^{(n)} = n - t] \\
&= \mathbb{P}\Big[B(n - 1, (1 - p_n)^t) = n - t\Big] \\
&= \binom{n - 1}{n - t} \Big[(1 - p_n)^t\Big]^{n-t} \Big[1 - (1 - p_n)^t\Big]^{t-1} \\
&= \binom{n - 1}{t - 1} \Big[(1 - p_n)^t\Big]^{n-t} \Big[1 - (1 - p_n)^t\Big]^{t-1} \\
&= \mathbb{P}\Big[B(n - 1, 1 - (1 - p_n)^t) = t - 1\Big].
\end{aligned}$$

$\square$

Napomena 4.2.2. *Kao što smo već naglasili, nama će od interesa biti tri različita procesa grananja: proces grananja kod grafova, binarni i Poissonov. Želja nam je analizirati proces grananja kod grafova, a već smo ranije spomenuli kako je taj proces sličan binomnom procesu grananja. Nadalje, zbog teorema 1.2.1 koji nam daje vezu binomne i Poissonove distribucije, možemo usporediti proces grananja kod grafova sa Poissonovim procesom. Uvest ćemo i nove oznake za ukupan broj jedinki nastalih u procesu grananja pa ćemo tako sa* T^{gr}_{n,p_n} *označavati ukupan broj jedinki nastalih u procesu grananja kod slučajnog grafa* $G(n, p_n)$*, sa* T^{bin}_{n,p_n} *ukupan broj jedinki nastalih u Galton-Watsonovom procesu gdje je* $Z \sim B(n, p_n)$*, a sa* T^{po}_c *ukupan broj jedinki nastalih u Galton-Watsonovom procesu uz* $Z \sim P(c)$*.*

Analiza i usporedba različitih procesa grananja

Teorem 4.2.3. *Promatrajmo Galton-Watsonov proces s distribucijom* $Z \sim P(c)$*, tj.* $Z_t \sim P(c)$*,* $t \geq 1$ *i označimo sa* T *vrijeme zaustavljanja. Za* $c \leq 1$*,* T *je konačan s vjerojatnosti* 1*, a za* $c > 1$*,* $T = \infty$ *s vjerojatnosti* $y = y(c)$*, gdje je* y *jedinstveni pozitivni realni broj koji zadovoljava jednadžbu*

$$e^{-cy} = 1 - y.$$

Dokaz. Dokažimo prvo slučaj kada je $c < 1$. Za $t < T$ vrijedi $Y_t > 0$, pa zbog rekurzije (4.4) zaključujemo $Y_0 + Z_1 + \cdots + Z_t - t > 0$, a budući da je $Y_0 = 1$, slijedi $Z_1 + \cdots + Z_t > t - 1$ ili, ekvivalentno, $Z_1 + \cdots + Z_t \geq t$. Za slučajne varijable Z_i, $i = 1, \ldots, t$ vrijedi $Z_i \sim P(c)$ pa iz teorema 1.2.6 slijedi da je suma $\sum_{i=1}^{t} Z_i \sim P(ct)$. Zbog pretpostavke $c < 1$ je $t - tc > 0$ pa iz Čebiševljeve nejednakosti slijedi

$$\begin{aligned}
\mathbb{P}[T > t] &= \mathbb{P}\Big[\sum_{i=1}^{t} Z_i \geq t\Big] = \mathbb{P}\Big[\sum_{i=1}^{t} Z_i - tc \geq t - tc\Big] \\
&\leq \mathbb{P}\Big[\Big|\sum_{i=1}^{t} Z_i - tc\Big| \geq \big|t - tc\big|\Big] \leq \frac{Var(\sum_{i=1}^{t} Z_i)}{(t - tc)^2} \\
&= \frac{c}{t(1-c)^2} \to 0,
\end{aligned}$$

za $t \to \infty$ pa zbog neprekidnosti vjerojatnosti u odnosu na padajući niz događaja $\{T > t\}$

$$\mathbb{P}[T = \infty] = \mathbb{P}\Big[\bigcap_{t=0}^{\infty}\{T > t\}\Big] = \lim_{t\to\infty} \mathbb{P}[T > t] = 0.$$

Zbog vjerojatnosti suprotnog događaja slijedi da je T konačan s vjerojatnosti jedan.

Neka je sada $c \geq 1$ i uvedimo oznaku $z = \mathbb{P}[T < \infty]$. Ako pretpostavimo da je jedinka od koje počinje proces grananja imala k potomaka, tada je vjerojatnost da proces završi u konačnom vremenu jednaka z^k, budući da svaki od potomka nezavisno započinje novi proces grananja i svaki od tih procesa također mora završiti u konačnom vremenu. Budući da je $Z_1 \sim P(c)$, slijedi

$$\begin{aligned} z = \mathbb{P}[T < \infty] &= \sum_{k=0}^{\infty} \mathbb{P}[T < \infty | Z_1 = k]\, \mathbb{P}[Z_1 = k] \\ &= \sum_{k=0}^{\infty} \mathbb{P}[Z_1 = k]\, z^k = \sum_{k=0}^{\infty} e^{-c} \frac{c^k z^k}{k!} = e^{c(z-1)}. \end{aligned} \tag{4.7}$$

Stavimo li $y = 1 - z$, iz (4.7) dobivamo jednadžbu $1 - y = e^{-cy}$, pri čemu je y vjerojatnost da je $T = \infty$. Kada je $c = 1$, imamo $1 - y = e^{-y}$, a budući da za $y > 0$ vrijedi $e^{-y} > 1 - y$, slijedi da je $y = 0$ jedino nenegativno rješenje jednadžbe. Zaključujemo kako je T konačan s vjerojatnosti jedan.

Ostalo je još dokazati tvrdnju teorema u slučaju $c > 1$. Za takav c definiramo funkciju $f(y) = 1 - y - e^{-cy}$. Tada je $f(0) = 0$, $f(1) = -e^{-c} < 0$ te $f'(0) = c - 1 > 0$, što znači da postoji $y \in (0, 1)$ takav da je $f(y) = 0$. Budući da je $f'' < 0$, funkcija je konkavna pa postoji točno jedan takav $y \in (0, 1)$. Dakle, funkcija $f(y)$ ima točno dvije nultočke na intervalu $[0, 1)$ i to $y_0 = 0$ i $y_0 \in (0, 1)$ pa zaključujemo da je $\mathbb{P}[T < \infty] = 1$ ili $\mathbb{P}[T < \infty] = 1 - y$. Pokažimo da vrijedi drugi slučaj. Pretpostavimo da je $Z_1 = k$. Tada je $Y_t = \sum_{i=1}^{t} Z_i - (t - 1) = \sum_{i=2}^{t} Z_i - t + k + 1 = \hat{Z}_t - t + k + 1$, gdje je $\hat{Z}_t = \sum_{i=2}^{t} Z_i$. Zbog teorema 1.2.6 je $\hat{Z}_t \sim P((t - 1)c)$ pa teorem 1.2.13 daje

$$\begin{aligned} \mathbb{P}[Y_{t+1} \leq 0 | Y_1 = k] &= \mathbb{P}[\hat{Z}_{t+1} \leq t - k | Y_1 = k] = \mathbb{P}\Big[\hat{Z}_{t+1} \leq ct(\frac{1}{c} - \frac{k}{ct})\Big| Y_1 = k\Big] \\ &= \mathbb{P}\Big[\hat{Z}_{t+1} \leq ct\Big(1 - (\frac{k}{ct} + \frac{c-1}{c})\Big)\Big| Y_1 = k\Big] \leq e^{-\frac{1}{2}\left(\frac{k}{ct} + \frac{c-1}{c}\right)^2 ct} \\ &= e^{-\frac{1}{2}\left(\frac{k^2}{ct} + \frac{2k(c-1)+(c-1)^2 t}{c}\right)} \leq e^{-\frac{k(c-1)}{c} - \frac{(c-1)^2 t}{2c}}, \end{aligned}$$

iz čega slijedi

$$L_k = \sum_{t \geq 2} \mathbb{P}[Y_t \leq 0 | Y_1 = k] \leq \sum_{t \geq 2} e^{-\frac{k(c-1)}{c} - \frac{(c-1)^2 (t-1)}{2c}} = e^{-\frac{k(c-1)}{c}} \sum_{t \geq 2} \left(\frac{1}{e}\right)^{\frac{(c-1)^2}{c}(t-1)} = e^{-\frac{k(c-1)}{c}} S,$$

gdje je $S = \sum_{t \geq 2} \left(\frac{1}{e}\right)^{\frac{(c-1)^2}{c}(t-1)} \in \mathbb{R}$. Odavde vidimo da zbog $\frac{c-1}{c} > 0$ vrijedi $L_k \to 0$ kada $k \to \infty$ pa za $\epsilon > 0$ postoji $k_0 \in \mathbb{N}$ tako da za sve $k \geq k_0$, uz uvjet $Y_1 = k$, vrijedi $L_k = \sum_{t \geq 2} \mathbb{P}[Y_t \leq 0 | Y_1 = k] < \epsilon$. Budući da je $Y_1 \sim P(c)$, vrijedi $\mathbb{P}[Y_1 > k] > 0$ za sve

$k \in \mathbb{N}$. Za $k \geq k_0$ je

$$\mathbb{P}[T = \infty | Y_1 = k] = \mathbb{P}\Big[\bigcap_{t=2}^{\infty} \{T > t\} \Big| Y_1 = k\Big] = 1 - \mathbb{P}\Big[\bigcup_{t=2}^{\infty} \{T \leq t\} \Big| Y_1 = k\Big]$$
$$\geq 1 - \sum_{t=2}^{\infty} \mathbb{P}[Y_t \leq 0 | Y_1 = k] > 1 - \epsilon$$

pa iz toga slijedi

$$\mathbb{P}[T = \infty] = \sum_{k=1}^{\infty} \mathbb{P}[T = \infty | Y_1 = k]\mathbb{P}[Y_1 = k] \geq \sum_{k=k_0}^{\infty} \mathbb{P}[T = \infty | Y_1 = k]\mathbb{P}[Y_1 = k] > 0.$$

Dakle, $\mathbb{P}[T = \infty] > 0$ pa mora vrijediti $\mathbb{P}[T < \infty] = 1 - y$ za $y \in (0, 1)$. □

Promatramo proces grananja kod grafova te sličnost tog procesa s Poissonovim procesom. Neka je $G(n, p_n)$ slučajni graf i $p_n = \frac{c}{n}$. Teorem 1.2.1 tada povlači da se $Z_1^{(n)} \sim B(n-1, \frac{c}{n})$ može aproksimirati s Poissonovom slučajnom varijablom s parametrom c kada $n \to \infty$. Slično vrijedi i za $Z_t^{(n)}$, $t \geq 2$ kad je $N_{t-1}^{(n)} \sim o(n)$, tj. kad je broj aktivnih i neaktivnih vrhova $o(n)$. Zato zaključujemo da postoje sličnosti i između algoritma PPS za pronalaženje komponente $C^{(n)}(v)$ i Poissonovog procesa sa parametrom c sve dok broj pronađenih vrhova u algoritmu PPS nije prevelik. Stoga će veza između ta dva procesa biti jača za $c < 1$, dok za $c > 1$ veza između algoritma PPS i Poissonovog processa s vremenom nestaje, što je i intuitivno jasno, budući da je Poissonov proces s parametrom $c > 1$ beskonačan s pozitivnom vjerojatnosti, dok je veličina komponente $C^{(n)}(v)$ sigurno konačna, štoviše, ne može biti veća od n.

Teorem 4.2.4. *Neka je $c > 0$, $k \in \mathbb{N}$ te v vrh u slučajnom grafu $G(n, \frac{c}{n})$. Tada vrijedi*

$$\lim_{n\to\infty} \mathbb{P}[|C^{(n)}(v)| = k] = \mathbb{P}[T_c^{po} = k]$$

Dokaz. Neka Z_t^{po}, Z_t^{gr} označavaju slučajne varijable Z_t u Poissonovom procesu s parametrom c i procesu grananja kod grafova s parametrima n i $p_n = \frac{c}{n}$. Označimo sa Γ sve uređene k-torke $z = (z_1, \ldots, z_k)$ nenegativnih cijelih brojeva tako da za rekurziju $y_0 = 1$, $y_t = y_{t-1} + z_t - 1$ vrijedi $y_t > 0$ za $t < k$ i $y_k = 0$. Tada je

$$\mathbb{P}[T_{n,p_n}^{gr} = k] = \sum_{z\in\Gamma} \mathbb{P}[Z_i^{gr} = z_i, 1 \leq i \leq k],$$
$$\mathbb{P}[T_c^{po} = k] = \sum_{z\in\Gamma} \mathbb{P}[Z_i^{po} = z_i, 1 \leq i \leq k]. \tag{4.8}$$

Za fiksirani $z \in \Gamma$ je, zbog nezavisnosti slučajnih varijabli $Z_t^{gr} \sim B(N_{t-1}^{gr}, p_n)$,

$$\mathbb{P}[Z_i^{gr} = z_i, 1 \leq i \leq k] = \prod_{i=1}^{k} \mathbb{P}[B(N_{n-1}^{gr}, p_n) = z_i]. \tag{4.9}$$

Fiksiramo li i, tada je $z_i = O(1)$ pa vrijedi $N_{i-1}^{gr} = n - O(1)$ i primjenom teorema 1.2.1 slijedi da $B(N_{i-1}^{gr}, p_n)$ možemo aproksimirati sa $P(c)$, tj. preciznije, zbog

$$(n - O(1))\frac{c}{n} \to c$$

je

$$\lim_{n\to\infty} \mathbb{P}[B(N_{i-1}^{gr}, p_n) = z_i] = \mathbb{P}[Z_i^{po} = z_i].$$

Budući da je produkt u (4.9) konačan, zbog nezavisnosti slučajnih varijabli Z_t^{po} je

$$\lim_{n\to\infty} \mathbb{P}[Z_i^{gr} = z_i, 1 \leq i \leq k] = \prod_{i=1}^{k} \lim_{n\to\infty} \mathbb{P}[B(N_{n-1}^{gr}, p_n) = z_i]$$
$$= \prod_{i=1}^{k} \mathbb{P}[Z_i^{po} = z_i] = \mathbb{P}[Z_i^{po} = z_i, 1 \leq i \leq k]$$

pa zbog konačnosti sume (4.8) dobivamo

$$\lim_{n\to\infty} \mathbb{P}[T_{n,p_n}^{gr} = k] = \mathbb{P}[T_c^{po} = k].$$

Zbog činjenice da algoritam PPS imitira proces grananja kod grafova do trenutka T_{n,p_n}^{gr}, vrijedi

$$\mathbb{P}[|C^{(n)}(v)| = k] = \mathbb{P}[T_{n,p_n}^{gr} = k],$$

čime smo dokazali tvrdnju teorema. □

Korolar 4.2.5. *Za $c > 0$, $k \in \mathbb{N}$ vrijedi*

$$\mathbb{P}[T_c^{po} = k] = \frac{e^{-ck}(ck)^{k-1}}{k!}.$$

Dokaz. Neka je $G(n, p_n)$ slučajni graf gdje je $p_n = \frac{c}{n}$ te neka je v proizvoljan vrh iz G. Budući da graf ima n vrhova, postoji $\binom{n-1}{k-1}$ izbora za skup vrhova S komponente $C(v)$ veličine k, a vjerojatnost da taj skup vrhova S ima više od $k - 1$ bridova je, zbog nezavisnosti, $p_n^k = O(n^{-k})$. Ukoliko slučajni graf $G(n, p_n)$ ima točno $k - 1$ bridova čiji su vrhovi u skupu S, oni nužno tvore stablo (inače $C^{(n)}(v)$ nije komponenta). Teorem 1.1.1 kaže da postoji k^{k-2} stabala s vrhovima u S, a vjerojatnost pojavljivanja takvog stabla je $p_n^{k-1}(1-p_n)^{\binom{k}{2}-k+1} \sim p_n^{k-1} = c^{k-1}n^{1-k}$, budući da se svaki od $k-1$ bridova pojavljuje (nezavisno) s vjerojatnosti p_n, dok se ni jedan od preostalih $\binom{k}{2}-k+1$ bridova ne smije pojaviti. Zaključujemo kako je vjerojatnost da podgraf grafa $G(n, p_n)$ induciran vrhovima skupa S čini povezan graf $\sim k^{k-2}c^{k-1}n^{1-k}$. Želimo li da taj podgraf čini povezanu komponentu, $C^{(n)}(v)$, ne smiju postojati bridovi između skupa S i njegovog komplementa na skupu vrhova grafa G. Vjerojatnost da je to ispunjeno je $(1-p_n)^{k(n-k)} = \left((1-\frac{c}{n})^{(n-k)}\right)^k \sim e^{-ck}$ jer postoji $k(n-k)$ bridova čiji je jedan kraj u skupu S, a drugi u komplementu skupa S i svaki takav brid neće se pojaviti u grafu G s vjerojatnosti $1-p_n$. Dolazimo do zaključka

$$\mathbb{P}[|C^{(n)}(v)| = k] \sim \binom{n-1}{k-1}k^{k-2}c^{k-1}n^{1-k}e^{-ck} = \frac{e^{-ck}(ck)^{k-1}}{k!}\frac{(n-1)!}{(n-k)!}n^{1-k} \to \frac{e^{-ck}(ck)^{k-1}}{k!}$$

kada $n \to \infty$ pa sada zbog teorema 4.2.4 slijedi tvrdnja korolara. □

Korolar 4.2.6. *Neka je $u > 0$. Za Poissonov proces $(Z_t, t \geq 0)$ s parametrom $c \neq 1$ vrijedi*

$$\mathbb{P}[T_c^{po} \geq u] < e^{-u(\alpha+o(1))}. \tag{4.10}$$

Dokaz. Iz korolara 4.2.5 i Stirlingove formule slijedi

$$\mathbb{P}[T_c^{po} = k] = \frac{e^{-ck}(ck)^{k-1}}{k!} \sim \frac{e^{-ck}(ck)^{k-1}}{\sqrt{2\pi k}(\frac{k}{e})^k} = \frac{1}{\sqrt{2\pi c}}k^{-\frac{3}{2}}e^{-k\alpha}, \tag{4.11}$$

za $\alpha = c - 1 - \ln c$. Za $c \neq 1$ je $e^{-\alpha} = ce^{1-c} < 1$ pa izraz u (4.11) teži prema nuli kada $k \to \infty$ i to eksponencijalnom brzinom. Dobivamo

$$\mathbb{P}[T_c^{po} \geq u] = \sum_{k \geq u} \mathbb{P}[T_c^{po} = k] = \sum_{k \geq u} \frac{1}{\sqrt{2\pi c}}k^{-\frac{3}{2}}e^{-k\alpha} < e^{-u(\alpha+o(1))}. \qquad \square$$

Navodimo još dva teorema koja pobliže opisuju Poissonov proces i čiji dokazi se mogu pronaći u knjizi Alon-Spencer [1].

Teorem 4.2.7. *Pretpostavimo $A \to \infty$ i $\epsilon \to 0^+$. Tada za Poissonov proces s parametrom $c = 1 - \epsilon$ vrijedi*

$$\mathbb{P}[T_{1-\epsilon}^{po} > A\epsilon^{-2}] < \epsilon e^{-(1+o(1))\frac{A}{2}} \tag{4.12}$$

kada $\epsilon \to 0^+$.

Teorem 4.2.8. *Za Poissonov proces s parametrom $c = 1 + \epsilon$ vrijedi*

$$\mathbb{P}[T^{po}_{1+\epsilon} = \infty] \sim 2\epsilon \tag{4.13}$$

kada $\epsilon \to 0^+$.

Osim s Poissonovim procesom, grananje kod grafova može se usporediti i s binarnim procesom grananja.

Propozicija 4.2.9. *Neka je $G(n, p_n)$ slučajni graf. Za $u > 0$ vrijedi*

$$\mathbb{P}[T^{gr}_{n,p_n} \geq u] \leq \mathbb{P}[T^{bin}_{n-1,p_n} \geq u].$$

Dokaz. Neka je v proizvoljan vrh iz $G(n, p_n)$. Tvrdnja propozicije slijedit će jednostavno iz modifikacije algoritma PPS. Promjena će se sastojati u tome da tijekom algoritma cijelo vrijeme imamo jednak broj neutralnih vrhova. Pretpostavimo da je vrh w izašao s početka slijeda Q i da, poštujući algoritam PPS, imamo $n-1-s$ neutralnih vrhova. U modificiranom algoritmu tim vrhovima dodajemo još s neutralnih vrhova w', tako da ukupno imamo $n-1$ neutralnih vrhova. Primijetimo da su w i w' susjedni s vjerojatnosi p_n, za svaki neutralni vrh w'. Ovim postupkom dobit ćemo komponentu veličine T^{bin}_{n-1,p_n}. Jasno je da je komponenta $C^{(n)}(v)$ grafa $G(n, p_n)$ podskup dobivene komponente. Odavde slijedi tvrdnja propozicije. □

Propozicija 4.2.10. *Neka je $G(n, p_n)$ slučajni graf. Za $u > 0$ vrijedi*

$$\mathbb{P}[T^{gr}_{n,p_n} \geq u] \leq \mathbb{P}[T^{bin}_{n-u,p_n} \geq u].$$

Napomena 4.2.11. *Primijetimo da za $n < u$ tvrdnja propozicije nema smisla. Budući da nas zanimaju asimptotski rezultati, pretpostavljamo da je $n > u$.*

Dokaz. Neka je v proizvoljan vrh iz $G(n, p_n)$. Uvodimo promjene u algoritam PPS. Prva promjena je da novi algoritam staje u trenutku kada dosegne veličinu u, tj. kada je zbroj aktivnih i neaktivnih vrhova jednak u. Ova promjena ne utječe na vjerojatnost pronalaženja barem u vrhova. Osim toga, u novom algoritmu, broj neutralnih vrhova bit će konstantan i jednak $n-u$. Zbog uvjeta zaustavljanja, u svakom trenutku će vrijediti $n-1-s \geq n-u$, gdje je $n-1-s$ broj neutranih vrhova koje bismo imali da provodimo nepromijenjen algoritam PPS pa od tih $n-1-s$ vrhova izaberemo bilo kojih $n-u$ vrhova i samo njih proglasimo neutralnima. Svaki put kada neki vrh w izlazi iz slijeda Q, provjeravamo povezanost w i w', gdje su w' novi neutralni vrhovi. Očito će tako dobiveni vrhovi biti podskup od $C^{(n)}(v)$ pa slijedi tvrdnja propozicije. □

4.3 Najveća komponenta slučajnog grafa

Slučajevi ispod kritičnog

Jako ispod kritičnog

Nakon analize procesa grananja iz prothodnog poglavlja možemo dokazati tvrdnje iz poglavlja 4.1. Podsjetimo se, kod slučaja jako ispod kritičnog imamo grubu parametrizaciju $p_n = \frac{c}{n}$, za $c < 1$.

Teorem 4.3.1. *Neka je $G(n, p_n)$ slučajan graf. U fazi jako ispod kritičnog je veličina najveće komponente $L_1^{(n)}$ grafa G reda* $\ln n$*.*

Dokaz. Neka je v proizvoljan vrh grafa $G(n, p_n)$. Iz propozicije 4.2.9 je

$$\mathbb{P}[|C^{(n)}(v)| \geq u] = \mathbb{P}[T^{gr}_{n,p_n} \geq u] \leq \mathbb{P}[T^{bin}_{n-1,p_n} \geq u],$$

pa zbog

$$(n-1)p_n = \frac{(n-1)c}{n} \to c \qquad \text{kada } n \to \infty$$

i teorema 1.2.1 slijedi

$$\mathbb{P}[|C^{(n)}(v)| \geq u] = (1 + o(1))\mathbb{P}[T^{po}_c \geq u].$$

Odavde, zbog korolara 4.2.6 vrijedi

$$\mathbb{P}[|C^{(n)}(v)| \geq u] \leq e^{-u(\alpha+o(1))}, \tag{4.14}$$

za $\alpha = c - 1 - \ln c > 0$. Budući da je $\alpha > 0$, desna strana u (4.14) teži prema nuli kada $u \to \infty$. Neka je K dovoljno velik takav da za $u = K \ln n$ vrijedi

$$\mathbb{P}[|C^{(n)}(v)| \geq u] < n^{-1.01}. \tag{4.15}$$

Budući da nejednakost (4.15) vrijedi za proizvoljan vrh v grafa G, a graf G ima n vrhova, slijedi da je vjerojatnost da za bilo koji vrh v grafa $G(n, p_n)$ vrijedi $|C^{(n)}(v)| \geq u$ manja od $nn^{-1.01}$, što teži prema 0 kada $n \to \infty$. Zaključujemo da je $L_1^{(n)} = O(\ln n)$ s velikom vjerojatnosti. □

Slabo ispod kritičnog

U ovom slučaju koristimo finu parametrizaciju pa pretpostavljamo da vrijedi $p_n = \frac{1-\epsilon}{n}$, za $\epsilon = \lambda n^{-\frac{1}{3}}$, $\epsilon = o(1)$ te $\lambda = \lambda(n) \to \infty$.

Teorem 4.3.2. *Neka je $G(n, p_n)$ slučajan graf. U fazi slabo ispod kritičnog je veličina najveće komponente $L_1^{(n)}$ grafa $G(n, p_n)$ reda $Kn^{\frac{2}{3}}\lambda^{-2}\ln\lambda$.*

Dokaz. Neka je v proizvoljan vrh grafa $G(n, p_n)$. Slično kao u dokazu prethodnog teorema,

$$\mathbb{P}[|C^{(n)}(v)| \geq u] = (1 + o(1))\mathbb{P}[T_{1-\epsilon}^{po} \geq u].$$

Za $K > 0$ stavimo $A = K\ln\lambda$ te $u = A\epsilon^{-2} = K\ln\lambda\epsilon^{-2} = Kn^{\frac{2}{3}}\lambda^{-2}\ln\lambda$ pa zbog pretpostavki $\epsilon = o(1)$ i $\lambda \to \infty$ vrijedi $\epsilon \to 0$, $A \to \infty$. Zadovoljene su pretpostavke teorema 4.2.7 pa

$$\mathbb{P}[T_{1-\epsilon}^{po} > A\epsilon^{-2}] < \epsilon e^{-(1+o(1))\frac{A}{2}},$$

tj. ako odaberemo dovoljno velik K,

$$\mathbb{P}[T_{1-\epsilon}^{po} \geq u] < \epsilon e^{-3.1\ln\lambda} = \epsilon\lambda^{-3.1}. \tag{4.16}$$

Za vrh v grafa $G(n, p_n)$ definiramo Bernoullijevu slučajnu varijablu $I_v^{(n)}$ koja detektira ima li komponenta $C^{(n)}(v)$ barem u vrhova te neka je

$$X^{(n)} = \sum_v I_v^{(n)}$$

slučajna varijabla koja broji vrhove v grafa $G(n, p_n)$ čije su komponente veličine barem u te $Y^{(n)}$ slučajna varijabla koja broji komponente grafa $G(n, p_n)$ veličine barem u. Iz definicije slučajne varijable $I_v^{(n)}$ slijedi $\mathbb{E}[I_v^{(n)}] = \mathbb{P}[T_{1-\epsilon}^{po} \geq u]$. Linearnost matematičkog očekivanja, nejednakost (4.16) i pretpostavka $\epsilon = \lambda n^{-\frac{1}{3}}$ povlače

$$\mathbb{E}[X^{(n)}] = \mathbb{E}[\sum_v I_v^{(n)}] = n\mathbb{E}[I_v^{(n)}] \leq n\epsilon\lambda^{-3.1} = n^{\frac{2}{3}}\lambda^{-2.1}.$$

Iz definicija slučajnih varijabli $X^{(n)}$ i $Y^{(n)}$ slijedi da je $X^{(n)} \geq uY^{(n)}$, tj. $Y^{(n)} \leq X^{(n)}u^{-1}$ pa

$$\mathbb{E}[Y^{(n)}] \leq u^{-1}\mathbb{E}[X^{(n)}] \leq (Kn^{\frac{2}{3}}\lambda^{-2}\ln\lambda)^{-1}n^{\frac{2}{3}}\lambda^{-2.1} = K^{-1}\lambda^{-0.1}(\ln\lambda)^{-1} \to 0,$$

budući da $\lambda \to \infty$ kada $n \to \infty$. Odavde zbog rasprave u poglavlju 3.1 slijedi $Y^{(n)} = 0$ s velikom vjerojatnosti što znači da je najveća komponenta $L_1^{(n)} = O(n^{\frac{2}{3}}\lambda^{-2}\ln\lambda)$ s velikom vjerojatnosti. □

Slučajevi iznad kritičnog

Kao i u slučaju ispod kritičnog, razlikovat ćemo dva manja slučaja—jako iznad kritičnog i slabo iznad kritičnog. Sličnost između slučajeva jako i slabo iznad kritičnog je u tome što oba sadrže dominantnu komponentu, koju ćemo u ovom poglavlju i precizno definirati. Sve ostale komponente su relativno male. Pokazat ćemo egzistenciju i jedinstvenost dominantne komponente.

Jako iznad kritičnog

Pretpostavimo da je $p_n = \frac{c}{n}$, za $c > 1$. Neka je $y = y(c)$ pozitivno realno rješenje jednadžbe (4.1). Neka je $\delta > 0$ proizvoljno mala konstanta te K proizvoljno velika konstanta. Definiramo $L_n^+ = (y + \delta)n$, $L_n^- = (y - \delta)n$ i $S_n = K \ln n$.

Definicija 4.3.3. *Neka je $G(n, p_n)$ slučajni graf i v vrh iz G. Za komponentu $C^{(n)}(v)$ grafa G kažemo da je* divovska *ako $L_n^- < |C^{(n)}(v)| < L_n^+$,* mala *ako $|C^{(n)}(v)| < S_n$, i* neobična *ako nije ni divovska ni mala.*

Lema 4.3.4. *Vjerojatnost da slučajni graf $G(n, p_n)$ ima neobičnu komponentu je $o(n^{-20})$.*

Dokaz. Neka je v proizvoljan vrh slučajnog grafa $G(n, p_n)$. Želimo li da komponenta $C^{(n)}(v)$ bude neobična, tada za $t = |C^{(n)}(v)|$ mora vrijediti $S_n < t < L_n^-$ ili $t > L_n^+$. Imamo ukupno $n - 2\delta n - K \ln n \sim n$ načina da odaberemo t i n načina da odaberemo vrh v. Pretpostavimo li da za dani vrh v i veličinu komponente t vrijedi

$$\mathbb{P}[|C^{(n)}(v)| = t] = q = o(n^{-22}), \tag{4.17}$$

tada slijedi da je vjerojatnost pojavljivanja neobične komponente u grafu $\sim n^2 q = o(n^{-20})$. Pokažimo sada da vrijedi (4.17). Iz propozicije 4.2.1 slijedi da je dovoljno dokazati da $\mathbb{P}\Big[B(n-1, 1-(1-p_n)^t) = t-1\Big] = o(n^{-22})$. Ovisno o t, dokaz se grana u dva slučaja.

Pretpostavimo da je $t = o(n)$. Tada slijedi

$$\frac{1 - (1 - \frac{c}{n})^t}{\frac{ct}{n}} = \frac{1 - \sum_{k=0}^{t} \binom{t}{k} \left(-\frac{c}{n}\right)^k}{\frac{ct}{n}} = 1 + o(1)$$

pa zaključujemo $1 - (1 - \frac{c}{n})^t \sim \frac{ct}{n}$. Tada iz teorema 1.2.10 slijedi

$$\begin{aligned}\mathbb{P}\Big[B(n-1, 1-(1-p_n)^t) = t-1\Big] &< \mathbb{P}\Big[B(n-1, 1-(1-p_n)^t) \le t\Big] \\ &\sim \mathbb{P}\Big[B(n-1, \frac{ct}{n}) \le t\Big].\end{aligned}$$

Kada $n \to \infty$, iz teorema 1.2.1 i Chernoffovih nejednakosti slijedi

$$\mathbb{P}\Big[B(n-1, \frac{ct}{n}) \le t\Big] \sim \mathbb{P}\Big[P(ct) \le t\Big] = \mathbb{P}\Big[P(ct) \le ct(1 - (1 - \frac{1}{c}))\Big] \le e^{-\frac{1}{2}(1-\frac{1}{c})^2 ct}$$

pa zaključujemo da $\mathbb{P}\Big[B(n-1, 1-(1-p_n)^t) = t-1\Big]$ teži u 0 eksponencijalnom brzinom kada $t \to \infty$. Budući da, zbog pretpostavke da je komponenta $C^{(n)}(v)$ neobična, mora vrijediti $t \ge K \ln n$, slijedi

$$e^{-\frac{1}{2}(1-\frac{1}{c})^2 ct} \le e^{-\frac{1}{2}(1-\frac{1}{c})^2 cK \ln n} = n^{-\frac{1}{2}(1-\frac{1}{c})^2 cK}$$

pa možemo odabrati K takav da promatrana vjerojatnost bude $o(n^{-22})$ čime je dokazan prvi slučaj.

Pretpostavimo sada da je $t \sim xn$, za fiksni x. Tada vrijedi

$$1 - (1 - \frac{c}{n})^t \sim 1 - (1 - \frac{c}{n})^{xn} = 1 - (1 - \frac{xc}{xn})^{xn} \to 1 - e^{-cx}$$

kada $n \to \infty$. Označimo $\alpha = 1 - e^{-cx}$. Analogno kao u prethodnom slučaju dobivamo

$$\mathbb{P}\Big[B(n-1, 1-(1-p_n)^t) = t - 1\Big] = \mathbb{P}\Big[P(n\alpha) \le t\Big]$$

pa koristeći Chernoffove nejednakosti

$$\mathbb{P}\Big[P(n\alpha) \le t\Big] = \mathbb{P}\Big[P(n\alpha) \le n\alpha(1 - (1 - \frac{t}{n\alpha}))\Big] < e^{-\frac{1}{2}(1-\frac{t}{n\alpha})^2 n\alpha}.$$

Zbog pretpostavke $t \sim nx$ slijedi $\frac{t}{n} \to x$ kada $n \to \infty$ pa kao i u prethodnom slučaju dobivamo $\mathbb{P}\Big[B(n-1, 1-(1-p_n)^t) = t-1\Big] = o(n^{-22})$.

Primijetimo na kraju da vjerojatnost da graf ima neobičnu komponentu možemo napraviti proizvoljno malom i to tako da namjestimo K da bude dovoljno velik. □

Teorem 4.3.5. *Neka je $G(n, p_n)$ slučajan graf. U fazi jako iznad kritičnog postoji točno jedna divovska komponenta.*

Dokaz. Pokazujemo prvo egzistenciju. Za prozivoljan vrh v grafa $G(n, p_n)$ izračunajmo vjerojatnost da komponenta $C^{(n)}(v)$ nije mala. Stavimo $\alpha = \mathbb{P}[|C^{(n)}(v)| \ge S_n]$. Propozicije 4.2.9 i 4.2.10 nam daju

$$\mathbb{P}[T^{bin}_{n-S_n, p_n} \ge S_n] \le \alpha \le \mathbb{P}[T^{bin}_{n-1, p_n} \ge S_n].$$

Zbog $(n - S_n)p_n \to c$ i $(n-1)p_n \to c$ kada $n \to \infty$, teorem 1.2.1 povlači

$$\begin{aligned} \mathbb{P}[T^{bin}_{n-S, p_n} \ge S] &\to \mathbb{P}[T^{po}_c \ge S] \qquad \text{kada } n \to \infty, \\ \mathbb{P}[T^{bin}_{n-1, p_n} \ge S] &\to \mathbb{P}[T^{po}_c \ge S] \qquad \text{kada } n \to \infty \end{aligned}$$

pa zbog teorema o sendviču vrijedi $\alpha \to \mathbb{P}[T^{po}_c \ge S]$. Označimo li $y = \mathbb{P}[T^{po}_c = \infty]$, teorem 4.2.3 povlači da je $y > 0$ i vrijedi $e^{-cy} = 1 - y$. Budući da je $c > 1$ fiksan, a $S_n \to \infty$, vrijedi

$$\mathbb{P}[T^{po}_c \ge S_n] \sim \mathbb{P}[T^{po}_c = \infty]$$

pa dobivamo $\alpha \sim y$. Zaključujemo da je vjerojatnost da $C^{(n)}(v)$ nije mala komponenta $\sim y$, dok lema 4.3.4 povlači da je vjerojatnost da je $C^{(n)}(v)$ neobična proizvoljno mala. Slijedi da je vjerojatnost da je $C^{(n)}(v)$ divovska komponenta $\sim y$. Budući da je vrh v bio proizvoljan, a imamo n mogućnosti izbora vrha v, zbog linearnosti matematičkog očekivanja je očekivani broj vrhova grafa $G(n, p_n)$ čija je komponenta $C^{(n)}(v)$ divovska $\sim ny$.

Preostalo je dokazati jedinstvenost. Pretpostavimo da $G(n, p_n)$ ima dvije divovske komponente. Neka su skupovi V_1 i V_2 vrhovi tih komponenti. Neka je $p'_n = n^{-\frac{2}{3}}$ (važno je da vrijedi $n^{-2} \ll p'_n \ll n^{-1}$) te neka su $G_1 \sim G(n, p'_n)$ i $G \sim G(n, p_n)$ na istom skupu vrhova veličine n. Za $G^+ = G_1 \cup G$ vrijedi $G^+ \sim G(n, p_n^+)$, gdje je $p_n^+ = p'_n + p_n - p'_n p_n$ jer se pojedini brid pojavljuje u $G_1 \cup G$ ako se ili pojavljuje u G_1, ili se ne pojavljuje u G_1, ali se pojavljuje u G. Vjerojatnost prvog događaja je p'_n, a vjerojatnost drugog (zbog nezavisnosti) iznosi $(1 - p'_n)p_n$. Zbog pretpostavke da su V_1 i V_2 skupovi vrhova divovskih komponenti, vrijedi $(y - \delta)n < |V_1| < (y + \delta)n$ i $(y - \delta)n < |V_2| < (y + \delta)n$ pa imamo $\Omega(n^2)$ izbora parova $\{v_1, v_2\}$ pri čemu je $v_1 \in V_1$ i $v_2 \in V_2$. Vjerojatnost da je barem jedan od parova $\{v_1, v_2\}$ povezan bridom u G_1 je, zbog $p'_n \gg n^{-2}$, jednaka $1 - o(1)$. Uključimo li taj brid u graf G^+, dobivamo komponentu $V_1 \cup V_2$ veličine barem $2(y - \delta)n$. Zbog $p'_n \ll n^{-1}$, $p_n = \frac{c}{n}$ je

$$\frac{p'_n + p_n - p'_n p_n}{p_n} \to 1$$

kada $n \to \infty$ pa zaključujemo $p_n^+ \sim p_n$. Lema 4.3.4 sada povlači da je vjerojatnost da graf G^+ ima komponentu te veličine $o(n^{-20})$ pa slijedi da je divovska komponenta jedinstvena.

□

Zbog jedinstvenosti divovske komponente i činjenice da je vjerojatnost postojanja neobične komponente zanemariva, slijedi da je $L_2^{(n)} \leq S_n = O(\ln n)$.

Slabo iznad kritičnog

Pretpostavimo da je $p_n = \frac{1+\epsilon}{n}$ za $\epsilon = \lambda n^{-\frac{1}{3}}$. Osim toga, pretpostavimo $\epsilon = o(1)$ i $\lambda = \lambda(n) \to \infty$. Pokazuje da se je analiza slučaja slabo iznad kritičnog teža kako $\lambda = \lambda(n)$ sporije raste prema beskonačnosti pa dodajemo još jedan uvjet $\lambda \gg \ln n$. Primijetimo da je $\epsilon^{-2} = \lambda^{-2} n^{\frac{2}{3}} \ll 2\lambda n^{\frac{2}{3}} = 2\epsilon n$. Kao i u prethodnom slučaju, pretpostavljamo da je $\delta > 0$ proizvoljno mala konstanta i $K > 0$ proizvoljno velika konstanta. Pretpostavljamo da je $S_n = K\epsilon^{-2}$, $L_n^- = (1 - \delta)2\epsilon n$ i $L_n^+ = (1 + \delta)2\epsilon n$. Neka je $G(n, p_n)$ slučajan graf i v proizvoljan vrh grafa $G(n, p_n)$.

Definicija 4.3.6. *Za komponentu $C^{(n)}(v)$ grafa $G(n, p_n)$ reći ćemo da je* mala *ako je $|C^{(n)}(v)| < S_n$,* dominantna *ako je $L_n^- < |C^{(n)}(v)| < L_n^+$ i* neobična*, inače.*

Napomena 4.3.7. *Razlog zašto u ovom slučaju najveću komponentu zovemo dominantnom, a ne divovskom kao prije, je taj što je izraz divovska rezerviran za komponentu reda n koja u slučaju slabo iznad kritičnog ne postoji.*

Lema 4.3.8. *Vjerojatnost da slučajni graf $G(n, p_n)$ ima neobičnu komponentu je $o(n^{-20})$.*

Dokaz ove leme je sličan dokazu leme 4.3.4 pa ga ovdje izostavljamo.

Teorem 4.3.9. *Neka je $G(n, p_n)$ slučajan graf. U fazi slabo iznad kritičnog postoji točno jedna dominantna komponenta.*

Dokaz. Pokazujemo egzistenciju. Za prozivoljan vrh v grafa $G(n, p_n)$ izračunajmo vjerojatnost da komponenta $C^{(n)}(v)$ nije mala. Zbog leme 4.3.8 to je jednako vjerojatnosti da je komponenta $C^{(n)}(v)$ dominantna. Stavimo $\alpha = \mathbb{P}[|C^{(n)}(v)| \geq S_n]$. Propozicije 4.2.9 i 4.2.10 nam daju

$$\mathbb{P}[T^{bin}_{n-S_n,p_n} \geq S_n] \leq \alpha \leq \mathbb{P}[T^{bin}_{n-1,p_n} \geq S_n]. \tag{4.18}$$

Zbog $(n-1)p_n \to 1+\epsilon$ kada $n \to \infty$, teorem 1.2.1 povlači

$$\mathbb{P}[T^{bin}_{n-1,p_n} \geq S_n] \to \mathbb{P}[T^{po}_{1+\epsilon} \geq S_n].$$

Iz definicije od S_n slijedi da $S_n \to \infty$, a budući da je $\epsilon = o(1)$, primijenjujemo teorem 4.2.8 pa vrijedi

$$\mathbb{P}[T^{po}_{1+\epsilon} \geq S_n] \sim \mathbb{P}[T^{po}_{1+\epsilon} = \infty] \sim 2\epsilon.$$

Pokušajmo sada aproksimirati $T^{bin}_{n-S_n,p_n}$ sa T^{po}. Zamijenimo li $n-1$ u prethodnom slučaju sa $n - S_n$, smanjili smo očekivanje za $(n-1)p_n - (n-S_n)p_n = (S_n - 1)p_n \sim S_n n^{-1}$. Sada pretpostavka $\lambda \gg \ln n$ te definicija od S_n i ϵ povlače

$$\frac{S_n n^{-1}}{\epsilon} = \frac{K\lambda^{-2} n^{\frac{2}{3}} n^{-1} \ln n}{\lambda n^{-\frac{1}{3}}} = K \frac{\ln n}{\lambda} \frac{1}{\lambda} \to 0$$

pa je $S_n n^{-1} = o(\epsilon)$. Zbog toga možemo $T^{bin}_{n-S_n,p_n}$ aproksimirati sa $T^{po}_{1+\epsilon-o(\epsilon)}$. Stavimo li $\epsilon' = \epsilon - o(\epsilon)$, onda $\epsilon \to 0$ povlači $\epsilon' \to 0$. Iz teorema 4.2.8 slijedi

$$\mathbb{P}[T^{po}_{1+\epsilon'} \geq S_n] \sim \mathbb{P}[T^{po}_{1+\epsilon'} = \infty] \sim 2\epsilon.$$

Zbog nejednakosti (4.18) i teorema o sendviču je $\alpha \sim 2\epsilon$ pa je komponenta $C^{(n)}(v)$ dominantna s vjerojatnosti $\sim 2\epsilon$. Budući da je vrh v bio proizvoljan, a imamo n mogućnosti izbora vrha v, zbog linearnosti matematičkog očekivanja je očekivani broj vrhova grafa $G(n, p_n)$ čija je komponenta $C^{(n)}(v)$ dominantna $\sim 2n\epsilon$. Dominantna komponenta je veličine između $(1-\delta)2n\epsilon$ i $(1+\delta)2n\epsilon$. Jedinstvenost se dokazuje analogno kao u dokazu teorema 4.3.5 uz odabir $p_1 = n^{-\frac{4}{3}}$. □

Dokazali smo da je $L_1^{(n)} \sim 2n\epsilon$, a budući da je druga najveća komponenta sigurno mala, vrijedi $L_2^{(n)} \leq K\epsilon^{-2} \ln n$.

Kritični slučaj

Pretpostavimo da je $p_n = \frac{1}{n} + \lambda n^{-\frac{4}{3}}$ pri čemu je $\lambda \in \mathbb{R}$ konstanta. Neka je $G(n, p_n)$ slučajni graf i $X^{(n)}$ slučajna varijabla koja broji komponente u $G(n, p_n)$ koje imaju strukturu stabla.

Teorem 4.3.10. *Neka su $a > 0$ i $b > 0$ takvi da $a < b$ te neka je $X^{(n)}$ slučajna varijabla koja za slučajni graf $G(n, p_n)$ broji komponente veličine između $an^{\frac{2}{3}}$ i $bn^{\frac{2}{3}}$ koje imaju strukturu stabla. U kritičnom slučaju za očekivanje od $X^{(n)}$ vrijedi*

$$\lim_{n\to\infty} \mathbb{E}[X^{(n)}] = \int_a^b e^{A(c)} c^{-\frac{5}{2}} (2\pi)^{-\frac{1}{2}} dc,$$

pri čemu je

$$A(c) = \frac{(\lambda - c)^3 - \lambda^3}{6}.$$

Dokaz. Neka je $c > 0$ i neka je $X^{(n)}$ slučajna varijabla koja broji komponente slučajnog grafa $G(n, p_n)$ veličine $cn^{-\frac{2}{3}} = k$ koje imaju strukturu stabla. Označimo sa $I_V^{(n)}$ Bernoullijevu slučajnu varijablu koja za k-člani skup vrhova V detektira čine li vrhovi V komponentu grafa $G(n, p_n)$ koja ima strukturu stabla. Tada je

$$E[I_V^{(n)}] = \mathbb{P}[I_V^{(n)}] = k^{k-2} p^{k-1} (1-p)^{k(n-k)+\binom{n}{2}-(k-1)}$$

jer u k-članom skupu vrhova možemo odabrati stablo na k^{k-2} načina (Cayleyeva formula). Svako stablo, da bi bilo povezano i bez ciklusa, mora imati točno $k - 1$ bridova, a ostali bridovi, kojih je $\binom{k}{2} - (k-1)$, ne smiju se pojavljivati. Osim toga, da bi to bila komponenta grafa $G(n, p_n)$, ni jedan od vrhova skupa V ne smije biti povezan ni s jednim vrhom izvan skupa V. Postoji $k(n - k)$ bridova koji povezuju vrh skupa V sa vrhom izvan skupa V. Linearnost matematičkog očekivanja povlači

$$\mathbb{E}[X^{(n)}] = \sum_V \mathbb{E}[I_V^{(n)}] = \binom{n}{k} \mathbb{E}[I_V^{(n)}] = \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)+\binom{n}{2}-(k-1)}. \tag{4.19}$$

Pojednostavimo formulu (4.19). Trik koji ćemo pri tome često koristiti je zapisivanje broja x kao $e^{\ln x}$. Osim toga, koristit ćemo

$$\ln(1 + x) = x - \frac{1}{2}x^2 + O(x^3) \tag{4.20}$$

za male vrijednosti x. Koristeći Stirlingovu formulu

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k!} \\ &= \frac{n^k}{k!} \prod_{i=1}^{k-1} (1 + \frac{i}{n}) \sim \frac{n^k e^k}{k^k \sqrt{2\pi k}} \prod_{i=1}^{k-1} (1 + \frac{i}{n}).\end{aligned} \tag{4.21}$$

Za $i = 0, 1, \dots k-1$, zbog (4.20) je

$$-\ln(1 - \frac{i}{n}) = \frac{i}{n} + \frac{i^2}{2n^2} + O(\frac{i^3}{n^3}).$$

Koristeći da je

$$\sum_{i=1}^{k-1} \frac{i}{n} = \frac{k^2}{2n} + o(1) \quad \text{i} \quad \sum_{i=1}^{k-1} \frac{i^2}{2n^2} = \frac{k^3}{6n^2} + o(1),$$

te definiciju od k, dobivamo

$$\sum_{i=1}^{k-1} -\ln(1 - \frac{i}{n}) = \frac{k^2}{2n} + \frac{k^3}{6n^2} + o(1) = \frac{k^2}{2n} + \frac{c^3}{6} + o(1).$$

Vratimo li se na (4.21), imamo

$$\binom{n}{k} \sim \frac{n^k e^k}{k^k \sqrt{2\pi k}} \prod_{i=1}^{k-1} (1 + \frac{i}{n}) = \frac{n^k e^k}{k^k \sqrt{2\pi k}} e^{-\frac{k^2}{2n} - \frac{c^3}{6} + o(1)}.$$

Za $p_n = \frac{1}{n} + \lambda n^{-\frac{4}{3}}$ je $p_n^{k-1} = n^{1-k}(1 + \lambda n^{-\frac{1}{3}})^{k-1}$ pa zbog (4.20) i $c = kn^{-\frac{2}{3}}$ imamo

$$\begin{aligned}(k-1)\ln(1 + \lambda n^{-\frac{1}{3}}) &= (k-1)(\lambda n^{-\frac{1}{3}} - \frac{1}{2}\lambda^2 n^{-\frac{2}{3}} + O(\lambda^3 n^{-1})) \\ &= k\lambda n^{-\frac{1}{3}} - \frac{1}{2}c\lambda^2 + o(1).\end{aligned}$$

Odavde slijedi

$$p_n^{k-1} = n^{1-k} e^{(k-1)\ln(1+\lambda n^{-\frac{1}{3}})} = n^{1-k} e^{k\lambda n^{-\frac{1}{3}} - \frac{c\lambda^2}{2} + o(1)},$$

a primijetimo i da vrijedi

$$(1 - p_n)^{k(n-k) + \binom{k}{2} - (k-1)} = e^{(k(n-k) + \binom{k}{2} - (k-1))\, \ln(1-p_n)}.$$

Prvo ćemo aproksimirati $\ln(1 - p_n)$

$$\ln(1 - p_n) = -p_n + O(n^{-2}) = -\frac{1}{n} - \frac{\lambda}{n^{\frac{4}{3}}} + O(n^{-2}),$$

a zatim i drugi član eksponenta (koristeći da je $k = cn^{\frac{2}{3}}$)

$$k(n-k) + \binom{k}{2} - (k-1) = k(n-k) + \frac{k(k-1)}{2} - (k-1) = kn - \frac{k^2}{2} + O(n^{\frac{2}{3}}).$$

Iz toga slijedi

$$\left(k(n-k)+\binom{k}{2}-(k-1)\right)\ln(1-p_n)=-k+\frac{k^2}{2n}-\frac{\lambda k}{n^{\frac{1}{3}}}+\frac{\lambda c^2}{2}+o(1)$$

pa je

$$(1-p_n)^{k(n-k)+\binom{k}{2}-(k-1)}=e^{-k+\frac{k^2}{2n}-\frac{\lambda k}{n^{\frac{1}{3}}}+\frac{\lambda c^2}{2}+o(1)}.$$

Konačno, dobivamo

$$\mathbb{E}[X^{(n)}]\sim\frac{n^k k^{k-2}}{k^k n^{k-1}\sqrt{2\pi k}}e^A=nk^{\frac{5}{2}}(2\pi)^{\frac{1}{2}}e^A,$$

za

$$A=k-\frac{k^2}{2n}-\frac{c^3}{6}+\frac{k\lambda}{n^{\frac{1}{3}}}-\frac{c\lambda^2}{2}-k+\frac{k^2}{2n}-\frac{\lambda k}{n^{\frac{1}{3}}}+\frac{\lambda c^2}{2}.$$

Iskoristimo li da je $k=cn^{\frac{2}{3}}$,

$$A=A(c)=-\frac{c^3}{6}+\frac{c\lambda^2}{2}-\frac{c^2\lambda}{2}=\frac{(\lambda-c)^3-\lambda^3}{6}$$

i

$$\mathbb{E}[X^{(n)}]\sim n^{-\frac{2}{3}}e^{A(c)}c^{-\frac{5}{2}}(2\pi)^{-\frac{1}{2}}. \tag{4.22}$$

Fiksirajmo vrijednosti λ, a i b te neka je $X^{(n)}$ slučajna varijabla koja broji komponente veličine između $an^{\frac{2}{3}}$ i $bn^{\frac{2}{3}}$. Tada je

$$\lim_{n\to\infty}E[X^{(n)}]=\int_a^b e^{A(c)}c^{-\frac{5}{2}}(2\pi)^{-\frac{1}{2}}\,dc. \qquad \square$$

Bibliografija

[1] N. Alon i J. H. Spencer, *The Probabilistic Method*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, 2008.

[2] B. Bollobás, *Random Graphs*, Cambridge studies in advanced mathematics 73, Cambridge University Press, 2001.

[3] P. Erdős i A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci **5** (1960), 17–61.

[4] R. Fagin, *Probabilities on finite models*, The Journal of Symbolic Logic **41** (1976), br. 1, 50–58.

[5] Glebskii, Kogan, Liagonkii i Talanov, *Range and degree of realizability of formulas in the restricted predicate calculus*, Cybernetics and Systems Analysis **5** (1969), br. 2, 142–154.

[6] G. Grimmett, *Probability on graphs*, Institute of Mathematical Statistics Textbooks, Cambridge University Press, 2010.

[7] S. Janson, T. Luczak i A.Rucinski, *Random graphs*, Wiley-Interscience series in discrete mathematics and optimization, John Wiley, 2000.

[8] N. Sarapa, *Teorija vjerojatnosti*, Udžbenici Sveučilišta u Zagrebu, Školska knjiga, 1992.

[9] S. Shelah i J. Spencer, *Zero-one laws for sparse random graphs*, Journal of the American Mathematical Society **1** (1988), br. 1, 97–115.

[10] P. W. Shor, *A new proof of Cayley's formula for counting labeled trees*, Journal of Combinatorial Theory, Series A **71** (1995), br. 1, 154–158.

Sažetak

U ovom radu proučava se tema slučajnih grafova, koja se može predstaviti kao spoj teorije vjerojatnosti i teorije grafova. Definira se slučajni graf kao vjerojatnosni prostor nad skupom svih grafova sa fiksiranim skupom vrhova. Također se proučavaju neka svojstva slučajnih grafova kao što je najveća klika u grafu, monotona svojstva i njihove prijelazne funkcije, zakoni nula-jedan i Erdős-Rényijeva promjena faze.

Summary

This thesis contributes to the topic of random graphs, which combines probability theory and graph theory. It gives the definition of random graph as a probability space over the set of graphs on a fixed vertex set. Also, it discusses some random graphs properties and phenomena like clique number, monotone properties and their threshold functions, zero-one laws and the Erdős-Rényi phase transition.

Životopis

Rođena sam 13. veljače 1992. godine u Zaboku, a odrasla u Krapini gdje pohađam i osnovnu školu do 2007. godine. Te godine upisujem Prirodoslovno-matematičku gimnaziju u Krapini koju završavam 2011. godine. Iste godine započinjem studij matematike na Prirodoslovno-matematičkom fakultetu u Zagrebu, a 2014. odlučujem se za upis Diplomskog studija Matematičke statistike.