Galton-Watsonov proces grananja i Kestenovo stablo

Jurić, Luka

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Luka Jurić

GALTON-WATSONOV PROCES GRANANJA I KESTENOVO STABLO

Diplomski rad

Voditelj rada:
prof. dr. sc. Bojan Basrak

Zagreb, 2022

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. ____________________________

Zahvaljujem se mentoru prof. dr. sc. Bojanu Basraku na trudu, razumijevanju i nesebičnom dijeljenju znanja tijekom izrade ovog diplomskog rada. Zahvaljujem se profesorici Matematike iz srednje škole koja mi je pomogla otkriti afinitet prema matematici te bitno utjecala na moju odluku o upisivanju ovog studija. Na posljetku veliko hvala mojim roditeljima koji su uvijek u životu bili moj oslonac i sigurna luka.

Sadržaj

Uvod

Galton-Watsonov proces, koji se može smatrati prvim stohastičkim modelom rasta neke populacije, povijesno se pojavio kao posljedica proučavanja preživljavanja aristokratskih prezimena u viktorijanskoj Engleskoj. Tada je primijećeno kako određena aristokratska prezimena izumiru, što je dovelo do zabrinutosti plemićkih obitelji. Francis Galton 1873. godine postavlja pitanje u Educational Timesu o vjerojatnosti izumiranja plemićkih prezimena u Ujedinjenom Kraljevstvu. Galtonovo pitanje glasilo je:

> „*U velikoj naciji, od koje ćemo se baviti samo odraslim muškarcima, N brojčano, a svaki od njih nosi zasebno prezime, naseljavaju okrug. Njihovo pravilo razmnožavanja je takvo da, u svakoj generaciji, postotak od* p_0 *odraslih muškaraca nema mušku djecu koja dožive odraslu dob,* p_1 *ima jedno takvo muško dijete,* p_2 *ima dvoje i tako dalje do postotka* p_5 *koji imaju petero. Odredi koji će udio njihovih prezimena izumrijeti nakon r generacija te koliko će biti različitih prezimena koje nosi m osoba.*“

Matematičkim rječnikom rečeno, Francis Galton je pretpostavio da se sve jedinke razmnožavaju neovisno jedna o drugoj i da sve imaju istu distribuciju potomaka. Nakon što nije dobio konkretan odgovor na svoje pitanje, Francis Galton izravno je kontaktirao Henry William Watsona. Otud dolazi i ime procesa. Zajedno su radili na problemu te objavljuju članak godinu dana kasnije u kojem dokazuju da je vjerojatnost izumiranja fiksna točka funkcije izvodnice distribucije potomaka. U radu ćemo pokazati da je to točno, međutim oni su prebrzo zaključili da je ta vjerojatnost uvijek jednaka 1. Ponekad se u literaturi proces naziva i Bienaymé-Galton-Watsonov proces. Naime, francuski matematičar Irenée-Jules Bienaymé promatrao je sličan problem već 1845. godine i došao je do točnog odgovora.

U drugom poglavlju uvodimo skup diskretnih stabala kako bismo pomoću njega konstruirali Galton-Watsonovo stablo. Galton-Watsonova stabla ustvari opisuju genealogiju Galton-Watsonovog procesa. Vrlo lako možemo konstruirati Galton-Watsonov proces iz Galton-Watsonovog stabla, budući da se radi samo o broju jedinki u svakoj generaciji, odnosno razini stabla. Nakon toga navodimo rezultate o Galton-Watsonovim procesima koji su dokazani pomoću formalizma diskretnih stabala. Fokusirati ćemo se na vjerojatnost izumiranja i na opise procesa koji su uvjetovani izumiranjem ili preživljavanjem.

Poglavlje 1

Osnovne definicije

1.1 Funkcije izvodnice

Definicija 1.1.1. *Neka je X diskretna slučajna varijabla definirana na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, koja poprima vrijednosti u $\mathbb{N}_0$ te neka je $p_k := \mathbb{P}(X = k)$ za $k \in \mathbb{N}_0$. Funkciju izvodnicu od X definiramo kao*

$$G(s) := \sum_{k=0}^{\infty} p_k s^k,$$

za $s \in \mathbb{R}$ za koje je $\sum_{k=0}^{\infty} p_k |s|^k < \infty$.

Lako uočimo da za $-1 \leq s \leq 1$ red konvergira. Naime,

$$\left| \sum_{k=0}^{\infty} p_k s^k \right| \leq \sum_{k=0}^{\infty} p_k |s|^k \leq \sum_{k=0}^{\infty} p_k = 1.$$

Nadalje, $G(s) = \mathbb{E}\left[s^X\right]$. Sljedeću Lemu navodimo bez dokaza, vidi N.Sandrić i Z.Vondraček [6].

Lema 1.1.2. *Ako red potencija $\sum_{k=0}^{\infty} a_k s^k$ s nenegativnim koeficijentima apsolutno konvergira za $|s| < 1$, onda apsolutno konvergira i red $\sum_{k=1}^{\infty} k a_k s^{k-1}$ za $|s| < 1$. Ako je $f(s) := \sum_{k=0}^{\infty} a_k s^k$ za $|s| < 1$, onda je $f'(s) := \sum_{k=1}^{\infty} k a_k s^{k-1}$ za $|s| < 1$.*

Koristeći prethodnu Lemu uočavamo da za funkciju izvodnicu $G(s) = \sum_{k=0}^{\infty} p_k s^k$ za $|s| < 1$ je $G'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1}$ za $|s| < 1$. Indukcijom dobivamo da je

$$G^{(n)}(s) = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} p_k s^{k-n}, \ \ |s| < 1.$$

Posebno, $G^{(n)}(0) = n!p_n$, odnosno

$$p_n = \frac{G^{(n)}(0)}{n}, \quad \text{za } n \in \mathbb{N}_0.$$

Propozicija 1.1.3. *Ako su X i Y dvije diskretne slučajne varijable koje poprimaju vrijednosti u $\mathbb{N}_0$ takve da im se pripadne funkcije izvodnice, redom G_X i G_Y, podudaraju za sve $|s| < 1$, onda vrijedi $\mathbb{P}(X = k) = \mathbb{P}(Y = k)$ za sve $k \in \mathbb{N}_0$.*

Dokaz. Neka je $k \in \mathbb{N}_0$ proizvoljan. Imamo,

$$\mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!} = \frac{G_Y^{(k)}(0)}{k!} = \mathbb{P}(Y = k).$$

□

Sljedeću Lemu navodimo bez dokaza, za dokaz vidi N.Sandrić i Z.Vondraček [6].

Lema 1.1.4. *Neka je $\sum_{k=0}^{\infty} a_k s^k$ red potencija s nenegativnim koeficijentima. Tada, za svaki $s_0 > 0$ vrijedi*

$$\lim_{s \nearrow s_0} \sum_{k=0}^{\infty} a_k s^k = \sum_{k=0}^{\infty} a_k s_0^k.$$

Sada možemo pomoću funkcija izvodnica izračunati momente slučajnih varijabli. Za $0 \leq s < 1$ imamo

$$G'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1}.$$

Koristeći prethodnu Lemu, za $s = 1$ dobivamo

$$G'(1) := \lim_{s \nearrow 1} G'(s) = \sum_{k=0}^{\infty} k p_k = \mathbb{E}[X].$$

Sada će nam biti u cilju pokazati svojstvo koje ima funkcija izvodnica diskretne slučajne varijable koja je ustvari zbroj dvije nezavisne diskretne slučajne varijable. Neka su X i Y dvije nezavisne diskretne slučajne varijable koje poprimaju vrijednosti u N_0. Neka su f_X i f_Y te G_X i G_Y redom pripadne funkcije gustoće i funkcije izvodnice. Definiramo $Z := X+Y$. Neka su f_Z i G_Z redom pripadna funkcija gustoće i funkcija izvodnica. Iz nezavisnosti lako dobijemo da je

$$\mathbb{P}(Z = k) = \sum_{j=0}^{k} \mathbb{P}(X = j)\mathbb{P}(Y = k - j), \quad \text{za sve } k \in \mathbb{N}_0,$$

odnosno vrijedi

$$f_Z(k) = \sum_{j=0}^{k} f_X(j) f_Y(k-j), \text{ za sve } k \in \mathbb{N}_0.$$

Sljedećim Teoremom ćemo odrediti G_Z.

Teorem 1.1.5. *(a) Neka su X i Y dvije nezavisne diskretne slučajne varijable koje poprimaju vrijednosti u $\mathbb{N}_0$ te neka su G_X i G_Y redom pripadne funkcije izvodnice. Definiramo $Z := X + Y$ i neka je G_Z pripadna funkcija izvodnica. Tada vrijedi*

$$G_Z(s) = G_X(s) G_Y(s), \ \ |s| \leq 1.$$

(b) Neka su X_1, ..., X_n nezavisne diskretne slučajne varijable koje poprimaju vrijednosti u $\mathbb{N}_0$ te neka su G_{X_1}, ..., G_{X_n} redom pripadne funkcije izvodnice. Definiramo $S := X_1 + \cdots + X_n$ i neka je G_S pripadna funkcija izvodnica. Tada vrijedi

$$G_S(s) = G_{X_1}(s) \cdots G_{X_n}(s), \ \ |s| \leq 1.$$

Dokaz. (a) Kako je $G_Z(s) = \mathbb{E}\left[s^Z\right]$, imamo

$$G_Z(s) = \mathbb{E}\left[s^{X+Y}\right] = \mathbb{E}\left[s^X s^Y\right] = \mathbb{E}\left[s^X\right]\mathbb{E}\left[s^Y\right] = G_X(s) G_Y(s), \ \ |s| \leq 1,$$

gdje smo u trećoj jednakosti koristili nezavisnost slučajnih varijabli X i Y.

(b) Dokaz slijedi uz primjenu matematičke indukcije i to analogno kao u (a). □

Neka je $(X_n)_{n\in\mathbb{N}}$ niz nezavisnih i jednako distribuiranih diskretnih slučajnih varijabli koje poprimaju vrijednosti u $\mathbb{N}_0$ te neka je N diskretna slučajna varijabla koja također poprima vrijednosti u $\mathbb{N}_0$ i koja je nezavisna od niza $(X_n)_{n\in\mathbb{N}}$. Definiramo slučajnu sumu S kao

$$S(\omega) := \begin{cases} \sum_{i=1}^{N(\omega)} X_i(\omega), & N(\omega) \geq 1 \\ 0, & \text{inače.} \end{cases}$$

Definiramo još $S_n = X_1 + \cdots + X_n$ za $n \in \mathbb{N}$. Neka su G_{X_n}, za $n \in \mathbb{N}$, G_N i G_S funkcije izvodnice od, redom, $(X_n)_{n\in\mathbb{N}}$, N i S. Odredimo G_s. Za početak računamo

$$\begin{aligned}
\mathbb{P}(S = k) &= \sum_{n=0}^{\infty} \mathbb{P}(S = k, N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(\sum_{i=1}^{N} X_i = k, N = n\right) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_i = k, N = n\right) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(S_n = k)\,\mathbb{P}(N = n), \quad k \in \mathbb{N}_0.
\end{aligned}$$

Također, prethodno znamo da je $G_{S_n}(s) = G_{X_1}(s)^n$, za $|s| \leq 1$. Sada imamo,

$$\begin{aligned}
G_S(s) &= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \mathbb{P}(S_n = k)\,\mathbb{P}(N = n)s^k \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) \sum_{k=0}^{\infty} \mathbb{P}(S_n = k)\, s^k \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) G_{S_n}(s) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) G_{X_1}(s)^n \\
&= G_N\left(G_{X_1}(s)\right), \quad |s| \leq 1.
\end{aligned} \tag{1.1}$$

Sada možemo odrediti i očekivanje slučajne sume. Znamo da je $\mathbb{E}[S] = G'_S(1)$. Kako je

$$G'_S(s) = G'_N\left(G_{X_1}(s)\right) G'_{X_1}(s), \quad |s| < 1,$$

imamo

$$\mathbb{E}[S] = G'_S(1) = G'_N\left(G_{X_1}(1)\right) G'_{X_1}(1) = G'_N(1)\, G'_{X_1}(1) = \mathbb{E}[N]\,\mathbb{E}[X_1]. \tag{1.2}$$

1.2 Markovljevi lanci

Potpoglavlje započinjemo s formalnom definicijom slučajnog procesa, zatim nastavljamo s formalnom definicijom Markovljevog lanca na prebrojivom prostoru stanja.

Definicija 1.2.1. *Neka je S skup. Slučajan proces s diskretnim vremenom i prostorom stanja S je familija $X = (X_n : n \in \mathbb{N}_0)$ slučajnih varijabli definiranih na nekom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ s vrijednostima u S. Dakle, za svaki $n \in \mathbb{N}_0$, je $X_n : \Omega \to S$ slučajna varijabla.*

Definicija 1.2.2. *Neka je S prebrojiv skup. Slučajni proces $X = (X_n : n \in \mathbb{N}_0)$ definiran na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ s vrijednostima u skupu S je Markovljev lanac ako vrijedi*

$$\mathbb{P}\left(X_{n+1} = j \mid X_n = i,\ X_{n-1} = i_{n-1}, ..., X_0 = i_0\right) = \mathbb{P}\left(X_{n+1} = j \mid X_n = i\right), \tag{1.3}$$

za svaki $n \in \mathbb{N}_0$ i za sve $i_0, ..., i_{n-1},\ i,\ j \in S$ za koje su obje uvjetne vjerojatnosti dobro definirane.

Pretpostavimo da se nalazimo u vremenskom trenutku n. Tada vremena $0,\ 1, ...,\ n-1$ predstavljaju prošlost, a vrijeme $n+1$ neposrednu budućnost. Gornja relacija nam zapravo govori da je ponašanje Markovljevog lanca u neposrednoj budućnosti, uvjetno na sadašnjost i prošlost, jednako ponašanju Markovljevog lanca u neposrednoj budućnosti, uvjetno na samo sadašnjost. Relaciju 1.3 još nazivamo *Markovljevim svojstvom*. Iz *Markovljevog svojstva* jednostavno se pokaže da su ustvari budućnost i prošlost uvjetno nezavisne uz danu sadašnjost. Potrebni će nam biti samo *homogeni* Markovljevi lanci, odnosno Markovljevi lanci u kojima vrijednost desne strane iz 1.3 ne ovisi o vremenskom trenutku $n \in \mathbb{N}_0$.

Definicija 1.2.3. *Matrica $P = \left(p_{ij} : i,\ j \in S\right)$ naziva se stohastičkom matricom ako je $p_{ij} \geq 0$ za sve $i,\ j \in S$ te vrijedi*

$$\sum_{j \in S} p_{ij} = 1,$$

za sve $i \in S$.

Sada smo spremni definirati homogen Markovljev lanac.

Definicija 1.2.4. *Neka je $\Lambda = (\Lambda_i : i \in S)$ vjerojatnosna distribucija na S i neka je $P = \left(p_{ij} : i,\ j \in S\right)$ stohastička matrica. Slučajni proces $X = (X_n : n \in \mathbb{N}_0)$ definiran na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ s vrijednostima u skupu S je homogen Markovljev lanac s početnom distribucijom Λ i prijelaznom matricom P ako vrijedi:*

(i) $\mathbb{P}(X_0 = i) = \Lambda_i$, *za sve $i \in S$*

(ii) $\mathbb{P}\left(X_{n+1} = j \mid X_n = i,\ X_{n-1} = i_{n-1}, ..., X_0 = i_0\right) = p_{ij}$,
za svaki $n \in \mathbb{N}_0$ i za sve $i_0, ...,\ i_{n-1},\ i,\ j \in S$.

Uočimo, nije odmah jasno da svaki homogen Markovljev lanac zadovoljava Markovljevo svojstvo. Međutim, svaki homogen Markovljev lanac s početnom distribucijom Λ i prijelaznom matricom P zadovoljava jednadžbu

$$\mathbb{P}(X_0 = i_0,\ X_1 = i_1, ...,\ X_n = i_n) = \Lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n},$$

za sve $i_0,\ i_1, ...,\ i_n$. Sada se lako pokaže da je $\mathbb{P}(X_{n+1} = j \mid X_n = i) = p_{ij}$, tj. svaki homogen Markovljev lanac zadovoljava Markovljevo svojstvo. Sljedeći Teorem nam govori uz koje uvjete možemo znati da je neki proces Markovljev lanac. S njim ujedno i završavamo potpoglavlje o Markovljevim lancima.

Teorem 1.2.5. *Neka je $Y = (Y_n\ :\ n \in \mathbb{N}_0)$ niz nezavisnih jednako distribuiranih slučajnih elemenata u nekom prostoru E, $f\ :\ S \times E \to S$ funkcija te neka je X_0 neka slučajna varijabla u S, nezavisna od niza Y. Za $n \in \mathbb{N}$ definiramo*

$$X_n := f(X_{n-1},\ Y_n).$$

Tada je $X = (X_n\ :\ n \in \mathbb{N}_0)$ Markovljev lanac.

Dokaz. Računamo,

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = j \mid X_n = i,\ X_{n-1} = i_{n-1}, ...,\ X_0 = i_0) \\
&= \mathbb{P}(f(X_n,\ Y_{n+1}) = j \mid X_n = i,\ X_{n-1} = i_{n-1}, ...,\ X_0 = i_0) \\
&= \mathbb{P}(f(i,\ Y_{n+1}) = j \mid X_n = i,\ X_{n-1} = i_{n-1}, ...,\ X_0 = i_0) \\
&= \mathbb{P}(f(i,\ Y_{n+1}) = j),
\end{aligned}$$

gdje zadnji redak proizlazi iz činjenice da je slučajan element Y_{n+1} nezavisan s $X_0, X_1, ..., X_n$. Naime, $X_0, X_1, ..., X_n$ su funkcije slučajnih elemenata $Y_1,\ Y_2, ...,\ Y_n$ i slučajne varijable X_0, a oni su po pretpostavci nezavisni s Y_{n+1}. Analogno, $\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(f(i,\ Y_{n+1}) = j)$. Slijedi da je X Markovljev lanac s prijelaznim vjerojatnostima $p_{ij} = \mathbb{P}(f(i,\ Y_1) = j)$. □

1.3 Martingali

U ovom potpoglavlju uvodimo pojmove martingala, supermartingala i submartingala. Zatim ćemo navest Teorem o konvergenciji submartingala te Korolar za konvergenciju supermartingala.

Definicija 1.3.1. *Neka je $(\Omega, \mathcal{F})$ izmjeriv prostor.*

(i) Familija $\mathbb{F} = (\mathcal{F}_n\ :\ n \in \mathbb{N}_0)$ σ-podalgebri od $\mathcal{F}$ takvih da je $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ za svaki $n \in \mathbb{N}_0$ zove se filtracija.

(ii) *Slučajni proces $X = (X_n \ : \ n \in \mathbb{N}_0)$ zove se adaptiran s obzirom na filtraciju $\mathbb{F} = (\mathcal{F}_n \ : \ n \in \mathbb{N}_0)$ ako je za svaki $n \in \mathbb{N}_0$ slučajna varijabla X_n $\mathcal{F}_n$-izmjeriva.*

(iii) *Neka je $X = (X_n \ : \ n \in \mathbb{N}_0)$ slučajni proces. Za $n \in \mathbb{N}_0$ definiramo ${\mathcal{F}_n}^0 := \sigma(X_0, \ X_1, ..., \ X_n)$. Tada se filtracija $\mathbb{F}^0 = \left({\mathcal{F}_n}^0 \ : \ n \in \mathbb{N}_0\right)$ zove prirodna filtracija od X.*

Definicija 1.3.2. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor, $\mathbb{F} = (\mathcal{F}_n \ : \ n \in \mathbb{N}_0)$ filtracija te $X = (X_n \ : \ n \in \mathbb{N}_0)$ slučajni proces. Pretpostavimo da je X adaptiran s obzirom na $\mathbb{F}$ te da je $\mathbb{E}\left[|X_n|\right] < \infty$ za sve $n \in \mathbb{N}_0$.*

(i) *X se zove martingal, odnosno $(\mathbb{F}, \ \mathbb{P})$-martingal, ako vrijedi*

$$\mathbb{E}\left[X_{n+1} \,|\, \mathcal{F}_n\right] = X_n \ \text{g.s., za sve } n \in \mathbb{N}_0.$$

(ii) *X se zove supermartingal, odnosno $(\mathbb{F}, \ \mathbb{P})$-supermartingal, ako vrijedi*

$$\mathbb{E}\left[X_{n+1} \,|\, \mathcal{F}_n\right] \leq X_n \ \text{g.s., za sve } n \in \mathbb{N}_0.$$

(iii) *X se zove submartingal, odnosno $(\mathbb{F}, \ \mathbb{P})$-submartingal, ako vrijedi*

$$\mathbb{E}\left[X_{n+1} \,|\, \mathcal{F}_n\right] \geq X_n \ \text{g.s., za sve } n \in \mathbb{N}_0.$$

Sljedeći Teorem (o konvergenciji submartingala) navodimo bez dokaza, vidi Z.Vondraček [8].

Teorem 1.3.3. *Neka je $X = (X_n \ : \ n \in \mathbb{N}_0)$ submartingal takav da vrijedi*

$$\sup_{n \in \mathbb{N}_0} \mathbb{E}\left[{X_n}^+\right] < \infty.$$

Tada postoji $X_\infty = \lim_{n\to\infty} X_n$ g.s. te vrijedi $\mathbb{E}\left[|X_\infty|\right] < \infty$.

Korolar 1.3.4. *Neka je $X = (X_n \ : \ n \in \mathbb{N}_0)$ supermartingal takav da je $X_n \geq 0$ g.s. za sve $n \in \mathbb{N}_0$. Tada postoji $X_\infty = \lim_{n\to\infty} X_n$ g.s. te vrijedi $\mathbb{E}\left[X_\infty\right] \leq \mathbb{E}\left[X_0\right]$.*

Dokaz. Definiramo $Y_n := -X_n$, $n \in \mathbb{N}_0$. Tada je $Y = (Y_n \ : \ n \in \mathbb{N}_0)$ submartingal i vrijedi ${Y_n}^+ = 0$. Zato je $\sup_{n \in \mathbb{N}_0} \mathbb{E}\left[{Y_n}^+\right] < \infty$. Po prethodnom Teoremu postoji $Y_\infty = \lim_{n\to\infty} Y_n$ g.s. i $Y_\infty \leq 0$. Stavimo li $X_\infty = -Y_\infty$, slijedi da $X_n \to X_\infty$ g.s..Iz Fatouove Leme dobivamo da je $\mathbb{E}\left[X_\infty\right] \leq \liminf_{n \in \mathbb{N}_0} \mathbb{E}\left[X_n\right] \leq \mathbb{E}\left[X_0\right]$. □

Poglavlje 2

Galton-Watsonov proces i izumiranje

U ovom poglavlju ćemo uvesti pojam *Galton-Watsonovog stabla*, koje ćemo koristiti kao osnovni model razvoja populacije. Poglavlje započinjemo opisom *Galton-Watsonovog procesa* i njegovom definicijom.

2.1 Galton-Watsonov proces

Populacija nekih jedinki započinje s nultom generacijom, odnosno u trenutku $n = 0$, koja se sastoji od jedne jedinke i nazivamo ju predak. Predak u trenutku $n = 1$ daje slučajan broj potomaka $k \in \mathbb{N}_0$ po vjerojatnosnoj distribuciji $(p_k : k \in \mathbb{N}_0)$ i sam nestaje iz populacije. Njegovi potomci, točnije jedinke prve generacije, u trenutku $n = 2$ podijele se na slučajan broj potomaka, nezavisno od drugih jedinki populacije s istom vjerojatnosnom distribucijom $(p_k : k \in \mathbb{N}_0)$. Kao i njihov predak, u trenutku stvaranja potomaka te jedinke nestaju iz populacije. Dokle god postoje jedinke u populaciji, taj proces se analogno nastavlja u trenucima $n = 3, 4, 5, \ldots$. Opisani proces nazivamo *Galton-Watsonov proces*. *Galton-Watsonov proces* možemo matematički formalizirati na sljedeći način.

Definicija 2.1.1. *Neka je $(\gamma_{n,i} : n, i \in \mathbb{N})$ niz nezavisnih, jednako distribuiranih slučajnih varijabli s vrijednostima u $\mathbb{N}_0$, definiranih na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ te neka je $\mathbb{P}(\gamma_{n,i} = k) = p_k$, za sve $k \in \mathbb{N}_0$ i za sve $n, i \in \mathbb{N}$. Proces $Z = (Z_n : n \in \mathbb{N}_0)$ definiran s*

$$\begin{aligned} Z_0 &:= 1 \\ Z_1 &:= \gamma_{1,1} \\ Z_2 &:= \gamma_{2,1} + \gamma_{2,2} + \ldots + \gamma_{2,Z_1} \\ &\vdots \\ Z_n &:= \gamma_{n,1} + \gamma_{n,2} + \ldots + \gamma_{n,Z_{n-1}} , \end{aligned} \tag{2.1}$$

pri čemu $Z_n = 0$ povlači $Z_{n+k} = 0$, za sve $k \in \mathbb{N}$, nazivamo Galton-Watsonov proces.

Distribuciju $(p_k \ : \ k \in \mathbb{N}_0)$ iz Definicije 2.1.1 još nazivamo „*distribucijom potomaka*". Uočimo da *Galton-Watsonov proces* možemo zapisati na sljedeći način

$$Z_n = \begin{cases} 1 \ , & n = 0 \\ \sum_{i=1}^{Z_{n-1}} \gamma_{n,i} \ , & Z_{n-1} \neq 0 \\ 0 \ , & \text{inače.} \end{cases} \tag{2.2}$$

Ovakav zapis će nam biti koristan za daljnju analizu procesa jer lako uočavamo da je Z_n ustvari suma slučajnog broja nezavisnih i jednako distribuiranih slučajnih varijabli. Iz pretpostavke o jednakoj distribuciji broja potomaka za sve jedinke i (2.1), uočavamo da je distribucija *Galton-Watsonovog procesa* potpuno određena distribucijom slučajne varijable $\gamma_{n,i} \sim \gamma_{1,1} = Z_1$. Od sada na dalje označavamo njenu funkciju izvodnicu sa G i neka je G_n oznaka funkcije izvodnice slučajne varijable Z_n.

Propozicija 2.1.2. *Za sve $n \in \mathbb{N}$ vrijedi*

$$G_n = \underbrace{G \circ \cdots \circ G}_{n} \ .$$

Posebno, za sve $n \in \mathbb{N}$ i sve $s \in [-1, 1]$ vrijedi

$$G_n(s) = G_{n-1}(G(s)) = G(G_{n-1}(s)) \ .$$

Dokaz. Tvrdnju ćemo dokazati indukcijom. Za $n = 1$ tvrdnja trivijalno vrijedi.
Neka je $n \geq 2$ i pretpostavimo da tvrdnja vrijedi za $n - 1$. Slijedi da je

$$G_n = G_{Z_{n-1}} \circ G_{\gamma_{n,1}} = G_{n-1} \circ G = \underbrace{G \circ \cdots \circ G}_{n-1} \circ G = \underbrace{G \circ \cdots \circ G}_{n} = G \circ G_{n-1} \ .$$

U prvoj jednakosti smo koristili tvrdnju koju smo pokazali u 1.1, druga proizlazi iz činjenice da je $\gamma_{n,i} \sim Z_1$ te u trećoj i zadnjoj jednakosti koristimo pretpostavku indukcije. □

Označimo sa m očekivani broj potomaka jedne jedinke. Jasno nam je da zbog jednake distribucije broja potomaka, odnosno slučajnih varijabli $(\gamma_{n,i} \ : \ n, i \in \mathbb{N})$, vrijedi $m = \mathbb{E}[\,Z_1] = \mathbb{E}[\,\gamma_{n,i}]$, za sve $n, i \in \mathbb{N}$.

Lema 2.1.3. *Ako je $m < \infty$, tada za sve $n \in \mathbb{N}$ vrijedi*

$$\mathbb{E}[\,Z_n] = m^n \ .$$

Dokaz. Koristeći tvrdnju iz 1.2 slijedi

$$\mathbb{E}[\,Z_n] = \mathbb{E}[\,Z_{n-1}]\, m = \mathbb{E}[\,Z_{n-2}]\, m^2 = \cdots = \mathbb{E}[\,Z_0]\, m^n = m^n \ ,$$

za sve $n \in \mathbb{N}$. □

Posljedično možemo uočiti:

1. $\mathbb{E}[Z_n] \to 0$, za $m < 1$
2. $\mathbb{E}[Z_n] = 1$, za $m = 1$
3. $\mathbb{E}[Z_n] \to \infty$, za $m > 1$.

Sada ćemo prikazati jedan primjer *Galton-Watsonovog procesa*.

Primjer 2.1.4 (Nespolno razmnožavanje)**.** *Promatramo nespolnu reprodukciju zelenih algi, pri kojoj svaka plodna zelena alga ima dva do četiri potomka s jednakom vjerojatnošću te će svaki potomak biti ili sterilan ili plodan s jednakom vjerojatnošću. Na početku imamo samo jednu plodnu zelenu algu. Pretpostavimo da nas zanima očekivani broj ukupnih plodnih zelenih algi do dvadesete generacije.*

Označimo sa Z_n broj plodnih zelenih algi u n-toj generaciji a sa Z'_n označimo ukupan broj zelenih algi u n -toj generaciji. Računamo distribuciju za Z_1:

- $\mathbb{P}(Z_1 = 0) = \mathbb{P}(Z_1 = 0,\ Z'_1 \in \{2,3,4\}) = \mathbb{P}(Z_1 = 0,\ Z'_1 = 4) + \mathbb{P}(Z_1 = 0,\ Z'_1 = 3) + \mathbb{P}(Z_1 = 0,\ Z'_1 = 2) = \frac{1}{3}\cdot\left(\frac{1}{2}\right)^4 + \frac{1}{3}\cdot\left(\frac{1}{2}\right)^3 + \frac{1}{3}\cdot\left(\frac{1}{2}\right)^2 = \frac{7}{48}$
- $\mathbb{P}(Z_1 = 1) = \frac{1}{3}\cdot\left(\frac{1}{2}\right)^4\cdot\binom{4}{1} + \frac{1}{3}\cdot\left(\frac{1}{2}\right)^3\cdot\binom{3}{1} + \frac{1}{3}\cdot\left(\frac{1}{2}\right)^2\cdot\binom{2}{1} = \frac{3}{8}$
- $\mathbb{P}(Z_1 = 2) = \frac{1}{3}\cdot\left(\frac{1}{2}\right)^4\cdot\binom{4}{2} + \frac{1}{3}\cdot\left(\frac{1}{2}\right)^3\cdot\binom{3}{2} + \frac{1}{3}\cdot\left(\frac{1}{2}\right)^2 = \frac{1}{3}$
- $\mathbb{P}(Z_1 = 3) = \frac{1}{3}\cdot\left(\frac{1}{2}\right)^4\cdot\binom{4}{3} + \frac{1}{3}\cdot\left(\frac{1}{2}\right)^3 = \frac{1}{8}$
- $\mathbb{P}(Z_1 = 4) = \frac{1}{3}\cdot\left(\frac{1}{2}\right)^4 = \frac{1}{48}$.

Dakle,

$$\gamma_{n,i} \sim Z_1 \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ \frac{7}{48} & \frac{3}{8} & \frac{1}{3} & \frac{1}{8} & \frac{1}{48} \end{pmatrix} .$$

Kako je $m = \mathbb{E}[Z_1] = \frac{3}{2}$, koristeći Lemu 2.1.3 dobivamo

$$\mathbb{E}[Z_0 + Z_1 + \cdots + Z_{20}] = 1 + m + \cdots + m^{20} = \frac{1 - m^{21}}{1 - m} = \frac{1 - \left(\frac{3}{2}\right)^{21}}{1 - \frac{3}{2}} = 9973.77 .$$

Zaključujemo da je očekivani broj ukupnih plodnih zelenih algi do dvadesete generacije 9973.

Slijedi propozicija u kojoj ćemo pokazati bitno svojstvo *Galton-Watsonovog procesa*, a to je da on zadovoljava Markovljevo svojstvo.

Propozicija 2.1.5. *Galton-Watsonov proces $Z = (Z_n : n \in \mathbb{N}_0)$, definiran na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, je Markovljev lanac s vrijednostima u skupu $\mathbb{N}_0$.*

Dokaz. Neka je $n \in \mathbb{N}$ proizvoljan i neka su $i_0, i_1, \ldots, i_n \in \mathbb{N}_0$ proizvoljni. Računamo,

$$\begin{aligned}
&\mathbb{P}\left(Z_n = i_n \mid Z_{n-1} = i_{n-1}, \ldots, Z_1 = i_1, Z_0 = i_0\right) \\
&= \mathbb{P}\left(\sum_{j=1}^{i_{n-1}} \gamma_{n,j} = i_n \,\middle|\, Z_{n-1} = i_{n-1}, \ldots, Z_1 = i_1, Z_0 = 1\right) \\
&= \mathbb{P}\left(\sum_{j=1}^{i_{n-1}} \gamma_{n,j} = i_n\right) .
\end{aligned}$$

U prvoj jednakosti uočavamo da je uvjetna vjerojatnost dobro definirana samo za $i_0 = 1$ i koristimo 2.2, dok u zadnjoj jednakosti koristimo činjenicu da su slučajne varijable $\gamma_{n,1}, ..., \gamma_{n,i_{n-1}}$ nezavisne od slučajnih varijabli $Z_0, ..., Z_{n-1}$. Nezavisnost proizlazi iz činjenice da su slučajne varijable $Z_0, ..., Z_{n-1}$ funkcije familije slučajnih varijabli $\{\gamma_{t,j} : 1 \leq t \leq n-1, \ j \geq 1\}$. Analogno se pokazuje jednakost

$$\mathbb{P}\left(Z_n = i_n \mid Z_{n-1} = i_{n-1}\right) = \mathbb{P}\left(\sum_{j=1}^{i_{n-1}} \gamma_{n,j} = i_n\right) .$$

□

Možemo još izračunati prijelazne vjerojatnosti. Imamo,

$$p_{ij} = \mathbb{P}\left(Z_n = j \ , \ Z_{n-1} = i\right) = \mathbb{P}\left(\sum_{l=1}^{i} \gamma_{n,l} = j\right) = p_{j*i} ,$$

gdje je p_{j*i} oznaka za j-tu vrijednost i-te konvolucije *distribucije potomaka* $(p_k : k \in \mathbb{N}_0)$.

Za kraj potpoglavlja uvodimo jedan od bitnijih događaja kod analize *Galton-Watsonovog procesa*. To je pojam *izumiranja* populacije kojim ćemo se baviti u nastavku. Kažemo da je populacija izumrla u trenutku n ukoliko je $Z_n = 0$. Uočimo, događaj *izumiranja* Galton-Watsonovog procesa odgovara upravo događajima

$$\varepsilon := \bigcup_{n=1}^{\infty} \{Z_n = 0\} = \lim_{n \to +\infty} \{Z_n = 0\} , \tag{2.3}$$

a njegovu vjerojatnost ćemo označavati sa

$$\pi := \mathbb{P}(\varepsilon) . \tag{2.4}$$

U nastavku će nas zanimati kako izračunati π i pod kojim uvjetima je $\pi = 1$ ili $\pi < 1$.

2.2 Skup uređenih stabala

Nakon što smo uveli pojam Galton-Watsonovog procesa, prelazimo na Galton-Watsonovo stablo. Najprije moramo uvesti pojam skupa uređenih stabala i definirati stablo zajedno s njegovim podskupovima koji će nam biti od interesa. Za definiciju skupa uređenih stabala pridržavati ćemo se Neveuvog formalizma.

Neka je

$$\mathcal{U} := \bigcup_{n\geq 0}\{1, 2, ...\}^n,$$

skup koji se sastoji od konačnog niza prirodnih brojeva s konvencijom da je $\{1, 2, ...\}^0 = \{\varrho\}$. Čvorove stabala identificirati ćemo s elementima skupa $\mathcal{U}$. Za $n \in \mathbb{N}$ i $u_1, ..., u_n \in \mathbb{N}$, neka je $u = (u_1, u_2, ...u_n)$. Tada je očito $u \in \mathcal{U}$ te definiramo duljinu od u kao $|u| = n$, s konvencijom da je $|\varrho| = 0$. Ako su u i v dva niza iz $\mathcal{U}$, označavamo s uv operaciju *spajanja* dva niza, s konvencijom da je $uv = u$ ako je $v = \varrho$ te $uv = v$ ako je $u = \varrho$. Preciznije rečeno, ϱ je neutralni element operacije *spajanja*.

Definiramo parcijalni uređaj na skupu $\mathcal{U}$ kojeg označavamo sa $\preceq$ i nazivamo ga *genealoški poredak*. Dakle, za $v, u \in \mathcal{U}$ pišemo da je $v \preceq u$ ako i samo ako postoji $w \in \mathcal{U}$ takav da je $u = vw$. Kažemo da je v *predak* od u i pišemo $v \prec u$ ako je $v \preceq u$ i $v \neq u$. Uočimo da skup svih *predaka* od u možemo zapisati kao

$$A_u := \{v \in \mathcal{U}\ :\ v \prec u\}\,.$$

Nadalje, možemo definirati još skup $\bar{A}_u := A_u \cup \{u\}$.

Zadnji zajednički predak podskupa $s \subseteq \mathcal{U}$ (*eng. most recent common ancestor*), kojeg označavamo sa $MRCA(s)$, je jedinstven element v od $\bigcap_{u\in s} \bar{A}_u$ s maksimalnom duljinom $|v|$. Uvodimo još leksikografski uređaj na $\mathcal{U}$. Za $u, v \in \mathcal{U}$ je $u < v$ ako je $v \prec u$ ili je $v = wjv'$ i $u = wiu'$, gdje je $w = MRCA(\{v, u\})$ i $j < i$ za neke $i, j \in \mathbb{N}$. Slijedi definicija stabla.

Definicija 2.2.1. *Stablo $\boldsymbol{t}$ je podskup od $\mathcal{U}$ koji zadovoljava:*

(i) $\varrho \in \boldsymbol{t}$

(ii) Ako je $u \in \boldsymbol{t}$, onda je $A_u \subset \boldsymbol{t}$

(iii) Za svaki $u \in \boldsymbol{t}$, postoji $k_u(\boldsymbol{t}) \in \mathbb{N}_0 \cup \{+\infty\}$, takav da za svaki $i \in \mathbb{N}$ je $ui \in \boldsymbol{t}$ ako i samo ako $1 \leq i \leq k_u(\boldsymbol{t})$.

Iz definicije lako uočavamo da broj $k_u(\boldsymbol{t})$ ustvari predstavlja broj djece, odnosno potomaka, čvora u iz stabla $\boldsymbol{t}$. Dogovorno, $k_u(\boldsymbol{t}) = -1$ ako $u \notin \boldsymbol{t}$. Čvor ϱ nazivamo korijenom stabla $\boldsymbol{t}$.

Definicija 2.2.2. *Neka je $\boldsymbol{t}$ stablo. List stabla $\boldsymbol{t}$ je čvor $u \in \boldsymbol{t}$ takav da je $k_u(\boldsymbol{t}) = 0$.*

Naravno, stablo može imati više listova. Skup svih listova stabla $\boldsymbol{t}$ je skup $\mathcal{L}_0(\boldsymbol{t}) = \{u \in \boldsymbol{t} \ : \ k_u(\boldsymbol{t}) = 0\}$. Također, beskonačan čvor je čvor u za kojeg vrijedi $k_u(\boldsymbol{t}) = +\infty$.

Skup svih stabala označavamo sa $\mathbb{T}_\infty$, a sa $\mathbb{T}$ označavamo skup stabala koji ne sadrže niti jedan beskonačan čvor. Preciznije,

$$\mathbb{T} := \{\boldsymbol{t} \in \mathbb{T}_\infty \ : \ k_u(\boldsymbol{t}) < +\infty, \ \forall \ u \in \boldsymbol{t}\} \, .$$

Za $\boldsymbol{t} \in \mathbb{T}_\infty$, neka je $|\boldsymbol{t}| = \mathbf{card}\,(\boldsymbol{t})$. Uočimo da za stablo $\boldsymbol{t} \in \mathbb{T}_\infty$, vrijedi da je

$$\sum_{u \in \boldsymbol{t}} k_u(\boldsymbol{t}) = |\boldsymbol{t}| - 1. \tag{2.5}$$

Sada možemo uvesti prebrojiv podskup od $\mathbb{T}_\infty$

$$\mathbb{T}_0 := \{\boldsymbol{t} \in \mathbb{T}_\infty \ : \ |\boldsymbol{t}| < +\infty\} \, , \tag{2.6}$$

a nazivamo ga skupom konačnih stabala. Jasno nam je da je $\mathbb{T}_0$ podskup i od skupa $\mathbb{T}$. Definirajmo visinu i širinu stabla.

Definicija 2.2.3. *Neka je $\boldsymbol{t}$ stablo. Visinu stabla $\boldsymbol{t}$ definiramo sa*

$$H(\boldsymbol{t}) := \sup_{u \in \boldsymbol{t}} |u| \, ,$$

a širinu stabla $\boldsymbol{t}$ na razini $h \in \mathbb{N}_0$ definiramo sa

$$z_h(\boldsymbol{t}) := \mathbf{card}\,(\{u \in \boldsymbol{t} \ : \ |u| = h\}) \, .$$

Za $\boldsymbol{t} \in \mathbb{T}_\infty$, visina $H(\boldsymbol{t})$ i širina $z_h(\boldsymbol{t})$ na nekoj razini $h \in \mathbb{N}_0$ ne moraju biti konačne. Uočimo da za $\boldsymbol{t} \in \mathbb{T}$, vrijedi da je $z_h(\boldsymbol{t})$ konačno, za sve $h \in \mathbb{N}_0$. Za $\boldsymbol{t} \in \mathbb{T}_0$ je $H(\boldsymbol{t}) = \max_{u \in \boldsymbol{t}} |u|$ i također je konačno.

Neka je $h \in \mathbb{N}$, sa $\mathbb{T}^{(h)}$ označavamo podskup skupa stabala s visinom manjom od h. Naime,

$$\mathbb{T}^{(h)} := \{\boldsymbol{t} \in \mathbb{T} \ : \ H(\boldsymbol{t}) \leq h\} \, . \tag{2.7}$$

Možemo definirati i podstablo stabla $\boldsymbol{t}$ iznad čvora $u \in \boldsymbol{t}$ sa

$$\mathcal{S}_u(\boldsymbol{t}) := \{v \in \mathcal{U} \ : \ uv \in \boldsymbol{t}\} \, .$$

Kako bi uveli pojam stabla s jednom *beskonačnom granom*, prvo neka je $v = (v_k \ : \ k \in \mathbb{N}) \in (\mathbb{N})^{\mathbb{N}}$ i označimo $\bar{v_n} = (v_1, ..., v_n)$, s konvencijom da je $\bar{v_0} = \varrho$. Tada skup $\bar{\mathbf{v}} = \{\bar{v_n} \ : \ n \in \mathbb{N}_0\}$ predstavlja neku beskonačnu *granu* nekog stabla. Sa $\mathbb{T}_1$ označavamo skup stabala s jednom beskonačnom *granom* i možemo ga zapisati kao skup

$$\mathbb{T}_1 := \{\boldsymbol{t} \in \mathbb{T} \ : \ \exists! \ v \in (\mathbb{N})^{\mathbb{N}} \text{ takav da } \bar{\mathbf{v}} \subset \boldsymbol{t}\} \, .$$

Definicija 2.2.4. *Neka je $h \in \mathbb{N}_0$ proizvoljan. Definiramo funkciju restrikcije $r_h : \mathbb{T} \longmapsto \mathbb{T}^{(h)}$ sa*

$$r_h(\boldsymbol{t}) := \{u \in \boldsymbol{t} \ : \ |u| \leq h\}, \ \textit{za sve } \boldsymbol{t} \in \mathbb{T}\,.$$

Zapravo na $r_h(\boldsymbol{t})$ možemo gledati kao na podstablo stabla $\boldsymbol{t}$, koje smo dobili tako da smo odsjekli stablo $\boldsymbol{t}$ na visini h.

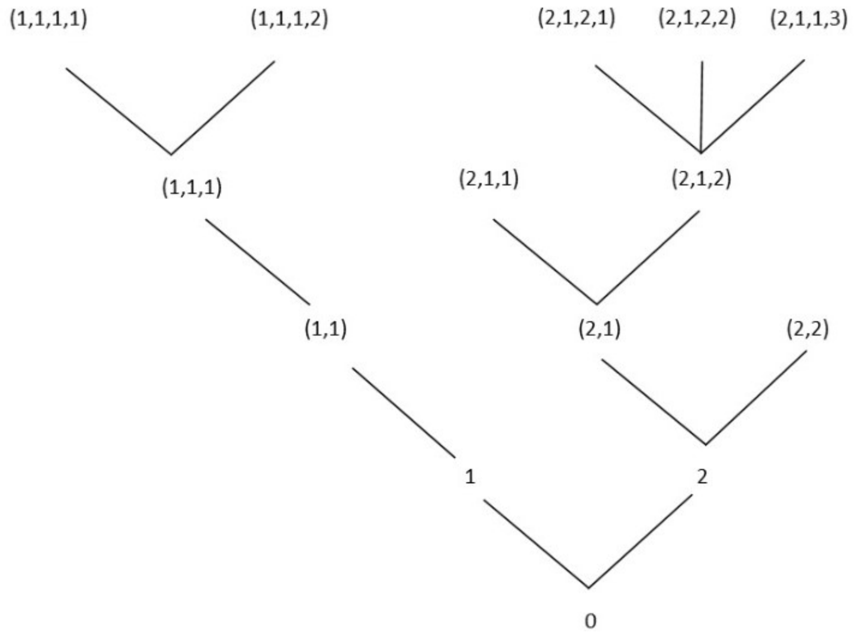
Na skupu $\mathbb{T}$ možemo definirati udaljenost za sve $\boldsymbol{t}, \boldsymbol{t'} \in \mathbb{T}$ sa

$$\delta(\boldsymbol{t}, \boldsymbol{t'}) := 2^{-\sup_{h\in\mathbb{N}_0}(r_h(\boldsymbol{t})=r_h(\boldsymbol{t'}))}.$$

Udaljenost δ je ustvari metrika na skupu $\mathbb{T}$, bolje rečeno $(\mathbb{T},\ \delta)$ je metrički prostor. Pokažimo svojstvo nejednakosti trokuta. Neka su $\boldsymbol{t}$, $\boldsymbol{t'}$ i $\boldsymbol{t''} \in \mathbb{T}$ proizvoljni. Pretpostavimo da je $\sup_{h\in\mathbb{N}_0}(r_h(\boldsymbol{t}) = r_h(\boldsymbol{t''})) \leq sup_{h\in\mathbb{N}_0}(r_h(\boldsymbol{t}) = r_h(\boldsymbol{t'}))$. Iz toga odmah slijedi da je $\delta(\boldsymbol{t}, \boldsymbol{t'}) \leq \delta(\boldsymbol{t}, \boldsymbol{t''})$. Suprotno, ako je $sup_{h\in\mathbb{N}_0}(r_h(\boldsymbol{t}) = r_h(\boldsymbol{t''})) > sup_{h\in\mathbb{N}_0}(r_h(\boldsymbol{t}) = r_h(\boldsymbol{t'}))$, onda mora biti $sup_{h\in\mathbb{N}_0}(r_h(\boldsymbol{t'}) = r_h(\boldsymbol{t''})) = sup_{h\in\mathbb{N}_0}(r_h(\boldsymbol{t}) = r_h(\boldsymbol{t'}))$ te sada dobivamo da je $\delta(\boldsymbol{t}, \boldsymbol{t'}) = \delta(\boldsymbol{t'}, \boldsymbol{t''})$. Zaključujemo da je $\delta(\boldsymbol{t}, \boldsymbol{t'}) \leq \max\left(\delta(\boldsymbol{t}, \boldsymbol{t'})\ ,\ \delta(\boldsymbol{t'}, \boldsymbol{t''})\right) \leq \delta(\boldsymbol{t}, \boldsymbol{t'}) + \delta(\boldsymbol{t'}, \boldsymbol{t''})$. Uočimo još da za proizvoljan $h \in \mathbb{N}_0$ vrijedi $\delta(r_h(\boldsymbol{t}), r_h(\boldsymbol{t'})) \leq \delta(\boldsymbol{t}, \boldsymbol{t'})$, to jest funkcije restrikcija $(r_h \ : \ h \in \mathbb{N}_0)$ su Lipschitz-neprekidne na metričkom prostoru $(\mathbb{T},\ \delta)$ pa samim time su i neprekidne.

Nakon što smo uveli metriku, sada možemo definirati i konvergenciju obzirom na tu metriku. Niz $(\boldsymbol{t}_n \ : \ n \in \mathbb{N}_0)$ stabala iz $\mathbb{T}$ konvergira prema stablu $\boldsymbol{t} \in \mathbb{T}$ obzirom na metriku δ ako i samo ako, za sve $h \in \mathbb{N}_0$, je $r_h(\boldsymbol{t}_n) = r_h(\boldsymbol{t})$ za dovoljno velike n. Završiti ćemo potpoglavlje s jednim primjerom konačnog stabla.

Primjer 2.2.5. *Slijedi nam grafički prikaz simulacije našeg Galton-Watsonovog procesa iz Primjera 2.1.4 do uključivo 4. generacije. Taj prikaz možemo definirati kao konačno stablo $\boldsymbol{t}_1 \in \mathbb{T}_0$.*



Slika 2.1: Konačno stablo $\boldsymbol{t}_1$

Odredimo visinu stabla, širinu stabla na razini 3, skup listova, broj djece čvora $(2,1)$*,* $|\boldsymbol{t}_1|$ *i zadnjeg zajedničkog pretka skupa* $\{(2,1,1,3),\ (2,1,1)\}$.

- *Visina stabla* $\boldsymbol{t}_1$ *je* $H(\boldsymbol{t}_1) = 4$.
- *širina na razini 3 je* $z_3(\boldsymbol{t}_1) = 3$.
- *skup listova je* $\mathcal{L}_0(\boldsymbol{t}_1) = \{(1,1,1,1),\ (1,1,1,2),\ (2,1,2,1),\ (2,1,2,2),\ (2,1,2,3)\}$.
- $k_{(2,1)}(\boldsymbol{t}_1) = 2$ *i* $|\boldsymbol{t}_1| = 14$.
- *MRCA* $(\{(2,1,1,3),\ (2,1,1)\}) = (2,1)$.

2.3 Galton-Watsonova stabla

Neka je $(p_k \,:\, k \in \mathbb{N}_0)$ vjerojatnosna distribucija na nekom skupu nenegativnih brojeva i X slučajna varijabla definirana na tom diskretnom vjerojatnosnom prostoru s vrijednostima u $\mathbb{N}_0$. Neka je $G_p(s) = \mathbb{E}[\, s^X]$, $s \in [0, 1]$, funkcija izvodnica slučajne varijable X. Sa $m(p)$ označavamo očekivanje vjerojatnosne distribucije $(p_k \,:\, k \in \mathbb{N}_0)$. Uočimo, $m(p) = G'_p(1) = \mathbb{E}[X]$ te od sada na dalje pišemo samo m za unaprijed danu diskretnu distribuciju.

Definicija 2.3.1 (Svojstvo grananja)**.** *Slučajno stablo τ, s vrijednostima u $\mathbb{T}$, ima svojstvo grananja ako za svaki $n \in \mathbb{N}$, uvjetno na $\{k_\varrho(\tau) = n\}$, podstabla $(\mathcal{S}_1(\tau),\ \mathcal{S}_2(\tau), ...,\ \mathcal{S}_n(\tau))$ su nezavisna i jednako distribuirana kao originalno slučajno stablo τ.*

Definicija 2.3.2 (Galton-Watsonovo stablo)**.** *Slučajno stablo τ, s vrijednostima u $\mathbb{T}$, je Galton-Watsonovo stablo s distribucijom potomaka $(p_k \,:\, k \in \mathbb{N}_0)$, ako zadovoljava svojstvo grananja i distribucija od $k_\varrho(\tau)$ je $(p_k \,:\, k \in \mathbb{N}_0)$.*

Bitna karakterizacija Galton-Watsonovog stabla je da τ je Galton-Watsonovo stablo s distribucijom $(p_k \,:\, k \in \mathbb{N}_0)$ ako i samo ako za svaki $h \in \mathbb{N}$ i $\boldsymbol{t} \in \mathbb{T}^{(h)}$ vrijedi

$$\mathbb{P}(r_h(\tau) = \boldsymbol{t}) = \prod_{u \in \boldsymbol{t},\ |u|<h} p_{k_u(\boldsymbol{t})}. \tag{2.8}$$

Štoviše, distribucija slučajnog stabla τ s vrijednostima u skupu $\mathbb{T}_0$ je dana sa

$$\mathbb{P}(\tau = \boldsymbol{t}) = \prod_{u \in \boldsymbol{t}} p_{k_u(\boldsymbol{t})}, \text{ za sve } \boldsymbol{t} \in \mathbb{T}_0. \tag{2.9}$$

Prisjetimo se Galton-Watsonovog procesa $Z = (Z_n \,:\, n \in \mathbb{N}_0)$ iz Definicije 2.1.1. Sljedeća Lema, koju navodimo bez dokaza, dati će nam poveznicu između Galton-Watsonovog procesa i Galton-Watsonovog stabla.

Lema 2.3.3. *Neka je τ Galton-Watsonovo stablo s distribucijom potomaka $(p_k \,:\, k \in \mathbb{N}_0)$. Proces širine stabla $(z_h(\tau) \,:\, h \in \mathbb{N}_0)$ je jednako distribuiran kao Galton-Watsonov proces $Z = (Z_n \,:\, n \in \mathbb{N}_0)$.*

Definicija 2.3.4. *Distribucija potomaka Galton-Watsonovog stabla je redom:*

(i) Kritična, ako je $m(p) = 1$

(ii) Sub-kritična, ako je $m(p) < 1$

(iii) Super-kritična, ako je $m(p) > 1$.

Ukoliko kažemo da je Galton-Watsonovo stablo kritično, sub-kritično ili super-kritično ustvari ćemo misliti na njegovu distribuciju potomaka. Prisjetimo se definicije izumiranja Galton-Watsonovog procesa i vjerojatnosti izumiranja iz 2.3 i 2.4. Sada ćemo uvesti analogne pojmove za Galton-Watsonovo stablo τ s distribucijom potomaka $(p_k \ : \ k \in \mathbb{N}_0)$.

Uočimo, događaj izumiranja Galton-Watsonovog stabla je jednak događaju

$$\varepsilon(\tau) = \{\tau \in \mathbb{T}_0\} , \tag{2.10}$$

te ćemo ga zbog prethodne Leme označavati jednako kao za Galton-Watsonov proces s ε. Također, zadržavamo oznaku π za vjerojatnost izumiranja. Kao što smo naveli na kraju potpoglavlja 2.1, krećemo s analizom vjerojatnosti izumiranja. Prikažimo sad koje distribucije potomaka nam neće biti od interesa i zašto.

Za distribuciju potomaka $(p_k \ : \ k \in \mathbb{N}_0)$ vrijedi:

- Ako je $p_0 = 0$, onda je $\pi = 0$ jer $\tau \notin \mathbb{T}_0$
- Ako je $p_0 = 1$, onda zbog $\tau = \{\varrho\}$ je $\pi = 1$
- Ako je $p_0 = 0$ i $p_1 = 1$, onda je $m = 1$ i $\tau = \bigcup_{n\in\mathbb{N}_0} \{1\}^n$, to jest τ je stablo s jednom *beskonačnom granom*, gdje je $\bar{v_n} = (v_1, ..., v_n) = (1, ..., 1)$. U ovom slučaju je trivijalno $\pi = 0$
- Ako je $0 < p_0 < 1$ i $p_0 + p_1 = 1$, onda je $H(\tau) + 1$ geometrijska slučajna varijabla s parametrom p_0 i $\tau = \bigcup_{0\le n\le H(\tau)} \{1\}^n$. Iz konačnosti geometrijske distribucije slijedi $\pi = 1$.

Od sada pretpostavljamo da naša distribucija potomaka $(p_k \ : \ k \in \mathbb{N}_0)$ zadovoljava sljedeće zahtjeve

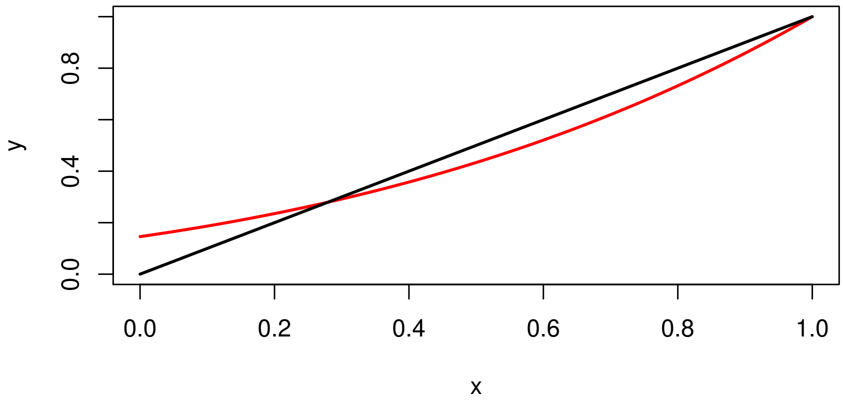
$$0 < p_0 < 1 \quad i \quad p_0 + p_1 < 1 . \tag{2.11}$$

Slijedi nekoliko tehničkih informacija vezanih uz funkciju izvodnicu $G_p(s)$ vjerojatnosne distribucije s gornjim svojstvom.

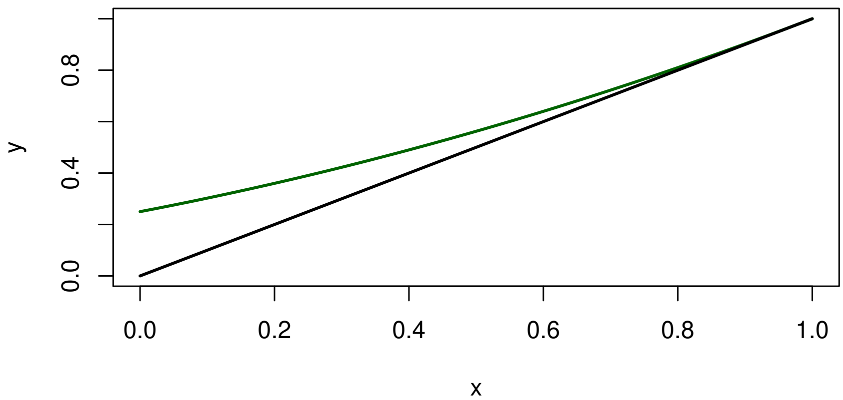
Napomena 2.3.5. *Kako vrijedi 2.11, zaključujemo da je sigurno $p_n > 0$ za neki $n \ge 2$. U tom slučaju je $G'_p(s) > 0$ i $G''_p(s) > 0$ za sve $s \in \langle 0 \ , \ 1\rangle$ pa slijedi da je G_p strogo rastuća i strogo konveksna na $[\,0\ ,\ 1]$.*
Vidimo da je $G_p(0) = p_0 > 0$ i $G_p(1) = 1$ i dolazimo do zaključka da jednadžba

$$G_p(s) = s,$$

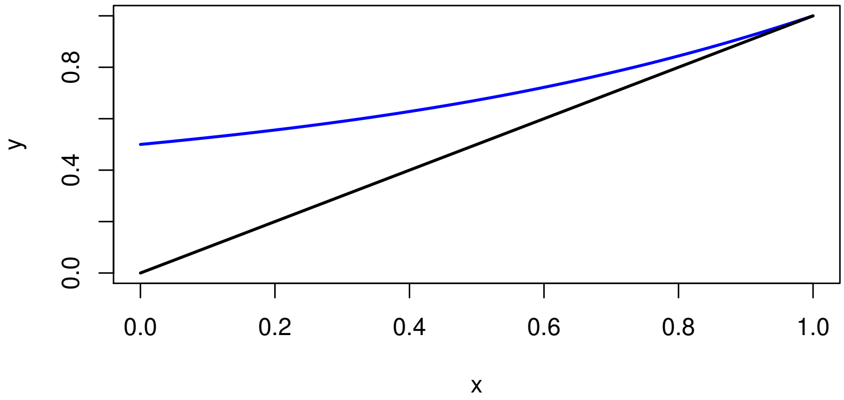
ima najviše dva moguća rješenja na intervalu $[\,0\ ,\ 1]$. Neka je q najmanje rješenje jednadžbe $G_p(s) = s$ na $[\,0\ ,\ 1]$. Uočimo da je $q = 1$ ako je $G'_p(1) \le 1$ i da tada imamo *samo jedno rješenje na intervalu* $[0, 1]$. *U tom slučaju se radi o kritičnoj ili sub-kritičnoj distribuciji potomaka.*



Slika 2.2: Prikaz funkcije izvodnice super-kritične distribucije potomaka



Slika 2.3: Prikaz funkcije izvodnice kritične distribucije potomaka



Slika 2.4: Prikaz funkcije izvodnice sub-kritične distribucije potomaka

Korolar 2.3.6. *Za kritično ili sub-kritično Galton-Watsonovo stablo τ s distribucijom potomaka $(p_k \; : \; k \in \mathbb{N}_0)$ koja zadovoljava 2.11, vrijedi izumiranje, odnosno $\mathbb{P}(\varepsilon) = 1$.*

Dokaz. Vrijedi

$$\begin{aligned}
\pi &= \mathbb{P}(\varepsilon(\tau)) \\
&= \sum_{n \in \mathbb{N}_0} \mathbb{P}\Big(\varepsilon(\mathcal{S}_1(\tau)), ..., \varepsilon(\mathcal{S}_n(\tau)) \mid k_\varrho(\tau) = n\Big) \, p_n \\
&= \sum_{n \in \mathbb{N}_0} \mathbb{P}(\varepsilon(\tau))^n \, p_n \\
&= G_p(\mathbb{P}(\varepsilon(\tau))) \\
&= G_p(\pi) \; ,
\end{aligned}$$

gdje smo u trećoj jednakosti koristili nezavisnost i jednaku distribuiranost od $(\mathcal{S}_1(\tau), ..., \mathcal{S}_n(\tau))$ sa τ uvjetno na $\{k_\varrho(\tau) = n\}$ koja proizlazi iz svojstva grananja stabla τ.
Kako je $\pi \in [0 \, , \, 1]$, slijedi da je π rješenje jednadžbe $G_p(s) = s$. Tvrdnja Korolara je ustvari neposredna posljedica prethodne Napomene. □

Preostaje nam još slučaj super-kritičnog Galton-Watsonovog stabla. Pretpostavimo da imamo super-kritičnu distribuciju potomaka. U tom slučaju znamo da je $0 < q < 1$, gdje je q najmanje rješenje jednadžbe iz Napomene 2.3.5. Iz prethodnog dokaza zaključujemo da je $\pi > 0$. Uvedimo novu vjerojatnosnu distribuciju $(\hat{p}_k \; : \; k \in \mathbb{N}_0)$ jednadžbom

$$\hat{p}_k = q^{k-1} p_k, \quad \text{za k } \in \mathbb{N}_0. \tag{2.12}$$

Pokažimo prvo da je $(\hat{p}_k \ : \ k \in \mathbb{N}_0)$ stvarno vjerojatnosna distribucija na skupu $\mathbb{N}_0$.

1. $q > 0$ i za sve $k \in \mathbb{N}_0$ je $p_k \geq 0 \ \Rightarrow \ \hat{p}_k \geq 0$, za sve $k \in \mathbb{N}_0$.
2. $\displaystyle\sum_{k \in \mathbb{N}_0} \hat{p}_k = \frac{1}{q} \sum_{k \in \mathbb{N}_0} q^k p_k = \frac{G_p(q)}{q} = 1.$

Uočimo da je $G_{\hat{p}}(s) = \frac{G_p(qs)}{q}$ i da vrijedi $G'_{\hat{p}}(1) = G'_p(q) < 1$. Dakle, ukoliko razmišljamo o $(\hat{p}_k \ : \ k \in \mathbb{N}_0)$ kao distribuciji potomaka i definiramo Galton-Watsonovo stablo $\hat{\tau}$ s tom distribucijom potomaka, dobivamo stablo sa sub-kritičnom distribucijom. Bitno je još uočiti da $(\hat{p}_k \ : \ k \in \mathbb{N}_0)$ zadovoljava 2.11. Sada znamo da je $\hat{\pi} = \mathbb{P}(\varepsilon(\hat{\tau})) = 1$.
Slijedi Lema koja će nam ponešto reći o distribuciji super-kritičnog Galton-Watsonovog stabla.

Lema 2.3.7. *Neka je τ Galton-Watsonovo stablo sa super-kritičnom distribucijom potomaka $(p_k \ : \ k \in \mathbb{N}_0)$ koja zadovoljava 2.11. Tada je $\pi = q$. Nadalje, uvjetno na događaj izumiranja, stablo τ je jednako distribuirano kao Galton-Watsonovo stablo $\hat{\tau}$ s distribucijom potomaka iz 2.12.*

Dokaz. Zbog toga što je $\hat{\pi} = \mathbb{P}(\varepsilon(\hat{\tau})) = 1$, jasno nam je da je $\hat{\tau} \in \mathbb{T}_0$.
Neka je $\boldsymbol{t} \in \mathbb{T}_0$ proizvoljno stablo. Koristeći 2.9 i 2.12 imamo

$$\begin{aligned} q\mathbb{P}(\hat{\tau} = \boldsymbol{t}) &= q \prod_{u \in \boldsymbol{t}} \hat{p}_{k_u(\boldsymbol{t})} \\ &= q \prod_{u \in \boldsymbol{t}} p_{k_u(\boldsymbol{t})} q^{k_u(\boldsymbol{t})-1} \\ &= q^{1+\sum_{u \in \boldsymbol{t}} (k_u(\boldsymbol{t})-1)} \prod_{u \in \boldsymbol{t}} p_{k_u(\boldsymbol{t})} \\ &= \mathbb{P}(\tau = \boldsymbol{t}), \end{aligned}$$

te u zadnjoj jednakosti koristimo činjenicu da je $\sum_{u \in \boldsymbol{t}} (k_u(\boldsymbol{t}) - 1) = |\boldsymbol{t}| - 1 - |\boldsymbol{t}|$. Sumiranjem gornje jednakosti po svim stablima $\boldsymbol{t} \in \mathbb{T}_0$, zaključujemo da za bilo koju nenegativnu funkciju $\mathcal{G}$ definiranu na skupu $\mathbb{T}_0$ vrijedi

$$q\mathbb{E}[\mathcal{G}(\hat{\tau})] = \mathbb{E}[\mathcal{G}(\tau)\mathbb{1}_{\{\tau \in \mathbb{T}_0\}}].$$

Za $\mathcal{G} = \mathbb{1}$, dobivamo da je $\mathbb{E}[\mathbb{1}_{\{\tau \in \mathbb{T}_0\}}] = \pi = q$. Sada imamo za proizvoljan $\boldsymbol{t} \in \mathbb{T}_0$

$$\mathbb{P}(\tau = \boldsymbol{t} \mid \varepsilon(\tau)) = \frac{\mathbb{P}(\tau = \boldsymbol{t}, \ \tau \in \mathbb{T}_0)}{\mathbb{P}(\tau \in \mathbb{T}_0)} = \mathbb{P}(\hat{\tau} = \boldsymbol{t}).$$

□

Slijedi Korolar koji je trivijalna posljedica prethodne Leme i Leme 2.3.3.

Korolar 2.3.8. *Neka je* $Z = (Z_n : n \in \mathbb{N}_0)$ *Galton-Watsonov proces sa super-kritičnom distribucijom potomaka koja zadovoljava jednadžbu 2.11. Uvjetno na događaj izumiranja* ε*,* Z *je jednako distribuiran kao Galton-Watsonov proces* $\hat{Z} = \left(\hat{Z}_n : n \in \mathbb{N}_0\right)$ *sa subkritičnom distribucijom potomaka* $(\hat{p}_k : k \in \mathbb{N}_0)$ *definiranom s 2.12.*

2.4 Super-kritična Galton-Watsonova stabla na događaju preživljavanja

Prethodno potpoglavlje smo završili s Lemom koja nam daje distribuciju Galton-Watsonovog stabla u slučaju izumiranja. Sada ćemo se baviti dekompozicijom Galton-Watsonovog stabla na događaju *preživljavanja*, odnosno na događaju ε^c. Neka je τ Galton-Watsonovo stablo sa super-kritičnom distribucijom potomaka $(p_k : k \in \mathbb{N}_0)$ koja zadovoljava jednadžbu 2.11. Događaj preživljavanja Galton-Watsonovog stabla τ možemo zapisati kao $\varepsilon^c = \{H(\tau) = +\infty\}$. Iz prethodnog potpoglavlja znamo da se ε^c događa s pozitivnom vjerojatnošću $1 - q$, gdje je q najmanje rješenje jednadžbe $G_p(s) = s$ na intervalu $[0, 1]$.

Za početak uvodimo svojstvo *preživljavanja* čvora nekog stabla iz $\mathbb{T}$. Za $u \in \boldsymbol{t}$ kažemo da je *preživio* u stablu $\boldsymbol{t} \in \mathbb{T}$ ako vrijedi da je $\mathbf{card}\left(\{v \in \boldsymbol{t} : u \prec v\}\right) = +\infty$. Ukoliko čvor $u \in \boldsymbol{t}$ ne zadovoljava navedeno svojstvo, kažemo da je *izumro*. Sada možemo definirati proces preživljavanja $z^s(\boldsymbol{t}) = \left(z_h^s(\boldsymbol{t}) : h \in \mathbb{N}_0\right)$

$$z_h^s(\boldsymbol{t}) := \mathbf{card}\left(\{u \in \boldsymbol{t} : |u| = h \text{ i } u \text{ je } \textit{preživio}\}\right).$$

Očito je da za svaki $h \in \mathbb{N}_0$ vrijedi $z_h^s(\boldsymbol{t}) \subseteq z_h(\boldsymbol{t})$. Navedimo još da je korijen ϱ stabla $\boldsymbol{t}$ *preživio* s vjerojatnošću $1 - q$. Sa S ćemo označavati broj djece korijena koja su *preživjela* te analogno sa E broj djece korijena koja nisu, to jest koja su *izumrla*. Djecu čvora u stabla $\boldsymbol{t}$ smatramo čvorovima oblika $\{ui \in \boldsymbol{t} : i \in \mathbb{N}\}$.

Za $r, l \in [0, 1]$ definiramo funkciju

$$S(r, l) = \mathbb{E}\left[r^S l^E \,\middle|\, \varepsilon^c\right].$$

Sljedeća Lema nam daje poveznicu između funkcije $S(r, l)$ i super-kritičnog Galton-Watsonovog stabla.

Lema 2.4.1. *Neka je* τ *Galton-Watsonovo stablo sa super-kritičnom distribucijom potomaka koja zadovoljava 2.11 te neka je* q *najmanje rješenje jednadžbe* $G_p(s) = s$ *na intervalu* $[0, 1]$*. Za* $r, l \in [0, 1]$ *vrijedi*

$$S(r, l) = \frac{G_p((1 - q)r + ql) - G_p(ql)}{1 - q}.$$

Dokaz. Računamo,

$$
\begin{aligned}
\mathbb{E}\left[r^S l^E \,\middle|\, \varepsilon^c\right] &= \frac{\mathbb{E}\left[r^S l^E \mathbb{1}_{\{S\geq 1\}}\right]}{q-1} \\
&= \frac{1}{q-1}\sum_{n\in\mathbb{N}} p_n \sum_{k=1}^{n} \binom{n}{k} (1-q)^k r^k q^{n-k} l^{n-k} \\
&= \frac{1}{q-1}\sum_{n\in\mathbb{N}} p_n \left[((1-q)r+ql)^n - (ql)^n\right] \\
&= \frac{G_p\left((1-q)r+ql\right) - G_p\left(ql\right)}{1-q}.
\end{aligned}
$$

U drugoj jednakosti smo koristili svojstvo grananja Galton-Watsonovog stabla te svojstvo da je Galton-Watsonovo stablo konačno s vjerojatnosti q, što ustvari odgovara vjerojatnosti pripadnosti čvora skupu E. □

Nakon što smo počeli dijeliti čvorove u skupove S i E, možemo uvesti pojam viševrsnog Galton-Watsonovog stabla. Viševrsno Galton-Watsonovo stablo $\hat{\tau}^s$ zadovoljava sljedeće:

(1) Čvorovi su ili tipa s ili tipa e, odnosno možemo na to gledati kao da je svaki čvor ili *preživio* ili je *izumro*

(2) Korijen $\hat{\varrho}^s$ stabla $\hat{\tau}^s$ je tipa s

(3) Čvor koji je tipa e "reproducira" samo čvorove tipa e i to uz sub-kritičnu distribuciji potomaka $(\hat{p}_k \ : \ k \in \mathbb{N}_0)$ danu u 2.12

(4) Čvor tipa s ima točno $S \geq 1$ čvorova djece tipa s i E čvorova djece tipa e. Funkcija izvodnica tako opisanog čvora je dana s $\mathbb{E}\left[r^S l^E\right] = S(r,l)$ iz prethodne Leme. Raspored između S djece tipa s te E djece tipa e je uniformno distribuiran između $\binom{S+E}{S}$ mogućih kombinacija.

Objasnimo detaljnije (4). Pretpostavimo da je čvor u tipa s. Vjerojatnost da on ima n djece, od kojih je $\{ui \ : \ i \in I\}$ tipa s, gdje je I neprazan podskup od $\{1, ..., n\}$, je jednaka

$$p_n(1-q)^{|I|-1} q^{n-|I|}.$$

Pokažimo da je to uistinu vjerojatnosna distribucija. Dakle,

$$
\begin{aligned}
&\sum_{n\in\mathbb{N}} \sum_{I\subset\{1,\dots,n\},\, I\neq\emptyset} p_n(1-q)^{|I|-1}q^{n-|I|}\\
&= \sum_{n\in\mathbb{N}} \sum_{k=1}^{n} p_n \binom{n}{k}(1-q)^{k-1}q^{n-k}\\
&= \sum_{n\in\mathbb{N}} p_n \frac{1-q^n}{1-q} = \frac{1}{1-q}\left(\sum_{n\in\mathbb{N}} p_n - \sum_{n\in\mathbb{N}} p_n q^n\right)\\
&= \frac{G_p(1)-G_p(q)}{1-q}\\
&= 1\ .
\end{aligned}
$$

Uočimo, uvjet (2) nam osigurava bar jednu *beskonačnu granu* zajedno s uvjetom (4). Sa τ^s označavamo slučajno stablo iz $\mathbb{T}$ koje definiramo kao $\hat{\tau}^s$ ukoliko zaboravljamo tipove.

Pokazali smo već distribuciju super-kritičnog Galton-Watsonovog stabla uvjetno na događaj izumiranja. Nakon što smo uveli pojam viševrsnog Galton-Watsonovog stabla τ^s, konačno možemo nešto reći o distribuciji super-kritičnog Galton-Watsonovog stabla uvjetno na događaj preživljavanja.

Lema 2.4.2. *Neka je τ Galton-Watsonovo stablo sa super-kritičnom distribucijom potomaka $(p_k\ :\ k\in\mathbb{N}_0)$ koja zadovoljava 2.11. Uvjetno na događaj preživljavanja ε^c, stablo τ je jednako distribuirano kao stablo τ^s.*

Dokaz. Za početak označimo skup čvorova stabla τ^s na visini h i tipa s sa

$$\mathcal{S}_h = \{u \in \tau^s\ :\ |u| = h \text{ i čvor } u \text{ je tipa } s\}.$$

Kako čvorovi tipa s mogu biti samo potomci čvorova tipa s te svaki čvor tipa s ima barem jednog potomka tipa s, zaključujemo da stablo $\hat{\tau}^s$ odrezano na razini h, odnosno $r_h(\hat{\tau}^s)$, je potpuno određeno skupovima $r_h(\tau^s)$ i $\mathcal{S}_h$.

Neka je $\boldsymbol{t} \in \mathbb{T}_0$ proizvoljan takav da je $H(\boldsymbol{t}) = h$ te $A \subset \{u \in \boldsymbol{t}\ :\ |u| = h\}$, $A \neq \emptyset$ i $n = z_h(\boldsymbol{t})$. Označimo sa $\mathcal{A} = \bigcup_{u\in A} A_u$ skup predaka od A. Uočimo, $\mathcal{A}^c$ je dan skupom $r_{h-1}(\boldsymbol{t}) \setminus A$. Za čvor $u \in \mathcal{A}$, s $k_u^s(\boldsymbol{t}, A)$ označavamo broj djece čvora u koja pripadaju skupu $\mathcal{A} \cup A$.

Računamo,

$$
\begin{aligned}
&\mathbb{P}\left(r_h(\tau^s) = \boldsymbol{t}\ ,\ \mathcal{S}_h = A\right) \\
&= \prod_{u\in\mathcal{A}^c} \hat{p}_{k_u(\boldsymbol{t})} \prod_{u\in\mathcal{A}} p_{k_u(\boldsymbol{t})}(1-q)^{k_u^s(\boldsymbol{t},A)-1} q^{k_u(\boldsymbol{t})-k_u^s(\boldsymbol{t},A)} \\
&= \left(\prod_{u\in\mathcal{A}\cup\mathcal{A}^c} p_{k_u(\boldsymbol{t})}\right)\left(\frac{1-q}{q}\right)^{\sum_{u\in\mathcal{A}}(k_u^s(\boldsymbol{t},A)-1)} q^{\sum_{u\in\mathcal{A}\cup\mathcal{A}^c}(k_u(\boldsymbol{t})-1)} \\
&= \mathbb{P}(r_h(\tau) = \boldsymbol{t})\left(\frac{1-q}{q}\right)^{|A|-1} q^{n-1} \\
&= \frac{1}{q-1}\mathbb{P}(r_h(\tau) = \boldsymbol{t})(1-q)^{|A|} q^{n-|A|},
\end{aligned}
$$

gdje smo u trećoj jednakosti dva puta koristili činjenicu da je $\sum_{u\in\boldsymbol{t}} k_u(\boldsymbol{t}) = |t| - 1$ i činjenicu da je $n = z_h(\boldsymbol{t})$. Sumirajući prethodnu jednakost po svim mogućim skupovima A dobivamo,

$$
\begin{aligned}
\mathbb{P}\left(r_h(\tau^s) = \boldsymbol{t}\right) &= \sum_A \frac{1}{q-1}\mathbb{P}(r_h(\tau) = \boldsymbol{t})(1-q)^{|A|} q^{n-|A|} \\
&= \frac{1}{q-1}\mathbb{P}(r_h(\tau) = \boldsymbol{t}) \sum_{k=1}^{n} \binom{n}{k}(1-q)^k q^{n-k} \\
&= \mathbb{P}(r_h(\tau) = \boldsymbol{t})\frac{1-q^n}{1-q}.
\end{aligned}
$$

S druge strane dobivamo,

$$
\begin{aligned}
\mathbb{P}\left(r_h(\tau) = \boldsymbol{t} \mid \varepsilon^c\right) &= \frac{\mathbb{P}(r_h(\tau) = \boldsymbol{t}) - \mathbb{P}(r_h(\tau) = \boldsymbol{t}\ ,\ \varepsilon)}{1-\mathbb{P}(\varepsilon)} \\
&= \frac{\mathbb{P}(r_h(\tau) = \boldsymbol{t}) - \mathbb{P}(r_h(\tau) = \boldsymbol{t})q^n}{1-q} \qquad = \mathbb{P}(r_h(\tau) = \boldsymbol{t})\frac{1-q^n}{1-q},
\end{aligned}
$$

gdje smo u drugoj jednakosti koristili svojstvo grananja na visini h. Uočimo da smo pokazali da je $\mathbb{P}\left(r_h(\tau) = \boldsymbol{t} \mid \varepsilon^c\right) = \mathbb{P}\left(r_h(\tau^s) = \boldsymbol{t}\right)$ za proizvoljan $\boldsymbol{t} \in \mathbb{T}_0$ i proizvoljnu visinu $h \in \mathbb{N}$. □

Iz definicije stabla τ^s, lako uočavamo da je proces preživljavanja $z^s(\tau) = \left(z_h^s(\tau)\ :\ h \in \mathbb{N}_0\right)$, uvjetno na događaj ε^c, Galton-Watsonov proces čija je distribucija potomaka $(\bar{p}_k\ :\ k \in \mathbb{N}_0)$ definirana funkcijom izvodnicom

$$
G_{\bar{p}}(r) = S(r,1) = \frac{G_p((1-q)r+q) - q}{1-q}.
$$

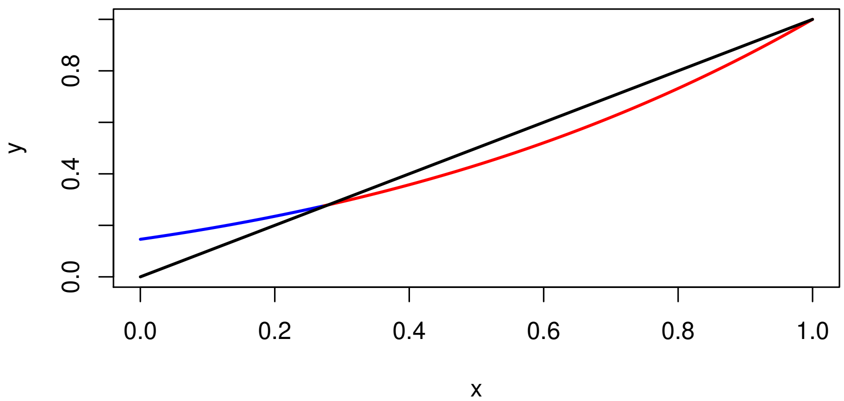
Također, zbog $G'_{\bar{p}}(r) = G'_p((1-q)r+q)$, slijedi da je očekivanje distribucije potomaka $(\bar{p}_k : k \in \mathbb{N}_0)$ jednako očekivanju distribucije potomaka $(p_k : k \in \mathbb{N}_0)$, jer je $G'_{\bar{p}}(1) = G'_p(1)$.

Nadalje, kako je $\bar{p}_0 = 0$, jer imamo $G_p(q) - q = q - q = 0$, zaključujemo kako je distribucija potomaka $(\bar{p}_k \,:\, k \in \mathbb{N}_0)$ super-kritična te samim time ne dolazi do izumiranja.

Ranije smo pokazali da Galton-Watsonovo stablo τ uvjetno na događaj izumiranja ε je Galton-Watsonovo stablo s distribucijom potomaka $(\hat{p}_k \,:\, k \in \mathbb{N}_0)$. Podsjetimo se, funkcija izvodnica distribucije $(\hat{p}_k \,:\, k \in \mathbb{N}_0)$ je $G_{\hat{p}}(s) = \frac{G_p(qs)}{q}$. Promatrajmo funkciju izvodnicu G_p za Galton-Watsonovo stablo sa super-kritičnom distribucijom potomaka $(p_k \,:\, k \in \mathbb{N}_0)$. Nju možemo dobiti pomoću vjerojatnosti izumiranja $\pi = q$ te funkcija izvodnica $G_{\hat{p}}$ i $G_{\bar{p}}$. Dakle,

$$G_p(s) = qG_{\hat{p}}\left(\frac{s}{q}\right)\mathbb{1}_{[0,\,q]}(s) + \left(q + (1-q)G_{\bar{p}}\left(\frac{s-q}{1-q}\right)\right)\mathbb{1}_{\langle q,\,1]}(s).$$

Slijedi grafički prikaz funkcije izvodnice G_p pomoću sub-kritične funkcije izvodnice $G_{\hat{p}}$ (označena plavom bojom) i super-kritične funkcije izvodnice $G_{\bar{p}}$ (označena crvenom bojom) pomnožene redom s q i $1 - q$.



Slika 2.5: Prikaz super-kritične funkcije izvodnice pomoću sub-kritične i super-kritične funkcije izvodnice

2.5 Kestenovo stablo

Nakon što smo definirali viševrsno Galton-Watsonovo stablo u prošlom potpoglavlju slijedi definicija Kestenovog stabla. Kestenovo stablo je još jedan tip viševrsnog Galton-Watsonovog stabla u kojem su svi čvorovi nezavisno distribuirani, međutim sama distribucija čvora ovisi o tipu čvora. Zbog tog svojstva Kestenovo stablo je očito viševrsno Galton-Watsonovo stablo. Prije nego što formalno definiramo Kestenovo stablo potrebna nam je definicija vjerojatnosne distribucije *pristrane po veličini*.

Definicija 2.5.1. *Neka je $(p_k \ : \ k \in \mathbb{N}_0)$ vjerojatnosna distribucija na skupu N_0 s konačnim očekivanjem $m < +\infty$. Pripadajuću vjerojatnosnu distribuciju koja je pristrana po veličini definiramo sa*

$$\check{p}_k := \frac{kp_k}{m}, \ \text{za sve } k \in \mathbb{N}_0.$$

Definicija 2.5.2 (Kestenovo stablo)**.** *Neka je $(p_k \ : \ k \in \mathbb{N}_0)$ distribucija potomaka koja zadovoljava 2.11 i ima konačno očekivanje $m < +\infty$. Kestenovo stablo koje povezujemo s distribucijom potomaka $(p_k \ : \ k \in \mathbb{N}_0)$ je viševrsno Galton-Watsonovo stablo $\hat{\tau}^*$ distribuirano na sljedeći način.*

(i) Čvorovi stabla su ili normalni ili posebni.

(ii) Korijen stabla je poseban čvor.

(iii) Normalan čvor reproducira samo normalne čvorove s obzirom na distribuciju potomaka $(p_k \ : \ k \in \mathbb{N}_0)$.

(iv) Poseban čvor reproducira čvorove s obzirom na distribuciju potomaka $(\check{p}_k \ : \ k \in \mathbb{N}_0)$, odnosno s obzirom na pripadajuću vjerojatnosnu distribuciju koja je pristrana po veličini vjerojatnosne distribucije $(p_k \ : \ k \in \mathbb{N}_0)$. Jedan od potomaka koji je uniformno izabran je poseban dok su ostali čvorovi potomci normalni.

Uočimo, poseban čvor uvijek ima bar jednog potomka jer je $\check{p}_0 = 0$. Također, s τ^* označavamo slučajno stablo s vrijednostima u $\mathbb{T}$ koje je definirano kao $\hat{\tau}^*$ kad zaboravljamo tipove. Slučajno stablo τ^* je iz $\mathbb{T}_1$ ukoliko je distribucija potomaka $(p_k \ : \ k \in \mathbb{N}_0)$ subkritična ili kritična. Slijedi Lema koja će nam dati poveznicu između Galton-Watsonovog stabla τ s distribucijom potomaka $(p_k \ : \ k \in \mathbb{N}_0)$ i τ^*.

Lema 2.5.3. *Neka je τ Galton-Watsonovo stablo s distribucijom potomaka $(p_k \ : \ k \in \mathbb{N}_0)$ koja zadovoljava 2.11 i konačnim očekivanjem $m < +\infty$ te neka je τ^* Kestenovo stablo koje povezujemo s tom distribucijom potomaka. Za sve $n \in \mathbb{N}_0$, $\boldsymbol{t} \in \mathbb{T}_0$ i $v \in \boldsymbol{t}$ takve da je $H(\boldsymbol{t}) = |v| = n$, vrijedi*

$$\mathbb{P}\left(r_n(\tau^*) = \boldsymbol{t}, \ v \text{ je poseban}\right) = \frac{1}{m^n}\mathbb{P}\left(r_n(\tau) = \boldsymbol{t}\right)$$

i

$$\mathbb{P}\left(r_n(\tau^*) = \boldsymbol{t}\right) = \frac{z_n(\boldsymbol{t})}{m^n}\mathbb{P}\left(r_n(\tau) = \boldsymbol{t}\right). \tag{2.13}$$

Dokaz. Najprije uočimo, ako je u poseban, vjerojatnost da čvor u ima k_u djece i da je ui poseban čvor za $1 \leq i \leq k_u$ je jednaka

$$\frac{k_u p_{k_u}}{m} \frac{1}{k_u} = \frac{p_{k_u}}{m}.$$

Neka su $n \in \mathbb{N}_0$, $\boldsymbol{t} \in \mathbb{T}^{(n)}$ i $v \in \boldsymbol{t}$ takvi da je $H(\boldsymbol{t}) = |v| = n$. Računamo,

$$\begin{aligned} \mathbb{P}\left(r_n(\tau^*) = \boldsymbol{t},\ v \text{ je poseban}\right) &= \prod_{u \in \boldsymbol{t} \setminus A_v,\, |u|<n} p_{k_u(\boldsymbol{t})} \prod_{u \in A_v} \frac{p_{k_u(\boldsymbol{t})}}{m} \\ &= \frac{1}{m^n} \mathbb{P}\left(r_n(\tau) = \boldsymbol{t}\right). \end{aligned}$$

Druga tvrdnja proizlazi iz činjenice da postoji samo jedan poseban čvor stabla $\boldsymbol{t}$ na razini n među $z_n(\boldsymbol{t})$ čvorova stabla $\boldsymbol{t}$ na razini n i prve jednadžbe koju smo dokazali. $\square$

Uvedimo sada pojam renormaliziranog Galton-Watsonovog procesa.

Definicija 2.5.4. *Neka je τ Galton-Watsonovo stablo s distribucijom potomaka $(p_k\ :\ k \in \mathbb{N}_0)$ kojoj je očekivanje jednako $m < +\infty$. Renormaliziran Galton-Watsonov proces stabla τ je proces $W = (W_n\ :\ n \in \mathbb{N}_0)$ takav da je*

$$W_n := \frac{z_n(\tau)}{m^n}, \ \textit{za sve } n \in \mathbb{N}_0.$$

Vidimo da je $W_0 = 1$. Također ćemo za proizvoljno slučajno stablo $\boldsymbol{t}$ s $W_n(\boldsymbol{t})$ označavati $\frac{z_n(\boldsymbol{t})}{m^n}$. Uvedimo još filtraciju $\mathbb{F} = (\mathcal{F}_n\ :\ n \in \mathbb{N}_0)$ koja je generirana s Galton-Watsonovim stablom τ, tako da je $\mathcal{F}_n = \sigma(r_n(\tau))$. Neka je $Z = (Z_n\ :\ n \in \mathbb{N}_0)$ Galton-Watsonov proces iz Leme 2.3.3. Možemo zaključiti dvije važne stvari.

- $W_n = \frac{Z_n}{m^n}$.
- $\sigma(r_n(\tau)) = \sigma(\{\gamma_{t,j}\ :\ 1 \leq t \leq n,\ j \geq 1\})$.

Kako je $Z_n = \displaystyle\sum_{i=1}^{Z_{n-1}} \gamma_{n,i}$, slijedi da je Z_n $\mathcal{F}_n$-izmjeriv. Sada možemo dokazati sljedeći korolar.

Korolar 2.5.5. *Neka je $(p_k\ :\ k \in \mathbb{N}_0)$ distribucija potomaka koja ima konačno očekivanje $m < +\infty$. Renormalizirani Galton-Watsonov proces $W = (W_n\ :\ n \in \mathbb{N}_0)$ je nenegativan martingal adaptiran obzirom na filtraciju $\mathbb{F} = (\mathcal{F}_n\ :\ n \in \mathbb{N}_0)$.*

Dokaz. Adaptiranost smo već pokazali, ona slijedi iz adaptiranosti od Z_n. Prvo, koristeći Teorem o uvjetnoj monotonoj konvergenciji, računamo

$$\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[\sum_{i=0}^{+\infty} Z_{n+1}\mathbb{1}_{\{Z_n=i\}} \,\middle|\, \mathcal{F}_n\right] = \sum_{i=0}^{+\infty} \mathbb{E}\left[Z_{n+1}\mathbb{1}_{\{Z_n=i\}} \mid \mathcal{F}_n\right].$$

Nadalje, imamo

$$\begin{aligned}
\mathbb{E}\left[Z_{n+1}\mathbb{1}_{\{Z_n=i\}} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[(\gamma_{n+1,1} + \cdots + \gamma_{n+1,Z_n})\,\mathbb{1}_{\{Z_n=i\}} \mid \mathcal{F}_n\right] \\
&= \mathbb{1}_{\{Z_n=i\}}\mathbb{E}\left[\gamma_{n+1,1} + \cdots + \gamma_{n+1,Z_n} \mid \mathcal{F}_n\right] \\
&= \mathbb{1}_{\{Z_n=i\}}\mathbb{E}\left[\gamma_{n+1,1} + \cdots + \gamma_{n+1,i}\right] \\
&= \mathbb{1}_{\{Z_n=i\}}mi,
\end{aligned}$$

gdje druga jednakost vrijedi zbog izmjerivosti Z_n sa $\mathcal{F}_n$ i treća jednakost zbog nezavisnosti $\gamma_{n+1,1}...\gamma_{n+1,Z_n}$ i familije $(\gamma_{t,j} \ : \ 1 \leq t \leq n, \ j \geq 1)$. Konačno,

$$\begin{aligned}
\mathbb{E}\left[W_{n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[\frac{Z_{n+1}}{m^{n+1}} \,\middle|\, \mathcal{F}_n\right] \\
&= \frac{1}{m^{n+1}}\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] \\
&= \frac{1}{m^{n+1}}\sum_{i=0}^{+\infty}\mathbb{E}\left[Z_{n+1}\mathbb{1}_{\{Z_n=i\}} \mid \mathcal{F}_n\right] \\
&= \frac{1}{m^{n+1}}\sum_{i=0}^{+\infty}\mathbb{1}_{\{Z_n=i\}}mi \\
&= \frac{mZ_n}{m^{n+1}} = W_n.
\end{aligned}$$

□

Lema 2.5.6. *Neka je $(p_k \ : k \in \mathbb{N}_0)$ distribucija potomaka Galton-Watsonovog stabla τ koja zadovoljava 2.11 i koja ima konačno očekivanje $m < +\infty$. Renormalizirani Galton-Watsonov proces $W = (W_n \ : \ n \in \mathbb{N}_0)$ konvergira prema slučajnoj varijabli $W_\infty = \lim_{n\to+\infty} W_n$ te vrijedi $\mathbb{E}[W_\infty] \leq 1$. Štoviše, vrijedi i $\mathbb{P}(W_\infty = 0) \in \{q, \ 1\}$.*

Dokaz. Prvu tvrdnju Leme dokazujemo pomoću prethodne Leme i Korolara 1.3.4. Naime, Renormalizirani Galton-Watsonov proces $W = (W_n \ : \ n \in \mathbb{N}_0)$ zadovoljava sve pretpostavke Korolara 1.3.4 te posljedično je $\mathbb{E}[W_\infty] \leq \mathbb{E}[W_0] = 1$.

Pokažimo sada da je $\mathbb{P}(W_\infty = 0) \in \{q, \ 1\}$. Dekompozicijom Galton-Watsonovog stabla

τ s obzirom na broj djece korijena dobivamo da je

$$W_n(\tau) = \frac{1}{m} \sum_{i=1}^{k_\varrho(\tau)} W_{n-1}\left(\mathcal{S}_i(\tau)\right).$$

Svojstvo grananja nam kaže da uvjetno na $k_\varrho(\tau)$, slučajna stabla $\mathcal{S}_1(\tau), ..., \mathcal{S}_{k_\varrho(\tau)}(\tau)$ su nezavisna i jednako distribuirana kao τ. Posljedično, kako proces $(W_n(\mathcal{S}_i(\tau)) \ : \ n \in \mathbb{N}_0)$ konvergira prema slučajnoj varijabli W^i, vrijedi da su $(W^i \ : \ i \in \mathbb{N})$ nezavisne nenegativne slučajne varijable jednako distribuirane kao W_∞ i nezavisne su od $k_\varrho(\tau)$. Puštanjem $n \to +\infty$, vrijedi

$$W_\infty = \frac{1}{m} \sum_{i=1}^{k_\varrho(\tau)} W^i.$$

Uočimo,

$$\mathbb{P}(W_\infty = 0) = \sum_{k \in \mathbb{N}_0} p_k \mathbb{P}(W^1 = 0, ..., W^k = 0) = G_p\left(\mathbb{P}(W_\infty = 0)\right),$$

čime smo pokazali da je $\mathbb{P}(W_\infty = 0)$ nenegativno rješenje jednadžbe $G_p(s) = s$, za $s \in [0, 1]$. Sada nam je jasno da je $\mathbb{P}(W_\infty = 0) \in \{q, 1\}$. □

Napomena 2.5.7. *Uočimo, $\varepsilon \subset \{W_\infty = 0\}$. Pretpostavimo da je $\mathbb{P}(W_\infty = 0) = q$. Kako je $\mathbb{P}(\varepsilon) = q$, uvjetno na događaj preživljavanja ε^c je $\lim_{n\to+\infty} W_n = \lim_{n\to+\infty} \frac{z_n(\tau)}{m^n} = W_\infty > 0$ $\mathbb{P}$-g.s..*

Prisjetimo se slučajne varijable X definirane na skupu $\mathbb{N}_0$, za koju vrijedi da je $\mathbb{P}(X = k) = p_k$, za sve $k \in \mathbb{N}_0$. Sada će nam biti cilj izračunati $\mathbb{P}(W_\infty = 0)$ te ujedno s tim završavamo ovo poglavlje. Prije toga ćemo navesti posljedicu Borel-Cantellijeve Leme koju ćemo koristiti u dokazu završnog Teorema.

Lema 2.5.8. *Neka su $L, L_1, L_2, ...$ nenegativne jednako distribuirane slučajne varijable. Tada vrijedi*

$$\limsup_{n\to+\infty} \frac{L_n}{n} = \begin{cases} 0 & , \text{ako } \mathbb{E}[L] < +\infty \\ +\infty & , \text{ako } \mathbb{E}[L] = +\infty. \end{cases}$$

Teorem 2.5.9. *Neka je $(p_k : k \in \mathbb{N}_0)$ super-kritična distribucija potomaka koja zadovoljava 2.11 i ima konačno očekivanje $m < +\infty$. Tada vrijedi*

(a) Ako je $\mathbb{E}\left[X \ln^+(X)\right] < +\infty$, onda je $\mathbb{P}(W_\infty = 0) = q$.

(b) Ako je $\mathbb{E}\left[X \ln^+(X)\right] = +\infty$, onda je $\mathbb{P}(W_\infty = 0) = 1$.

Dokaz. Neka su $\mathbb{P}$ i $\mathbb{P}^*$ redom distribucije, na skupu $\mathbb{T}$, Galton-Watsonovog stabla τ sa distribucijom potomaka $(p_k\ :\ k \in \mathbb{N}_0)$ i Kestenovog stabla τ^* kojeg povezujemo s tom distribucijom potomaka. Koristeći 2.13 iz Leme 2.5.3 dolazimo do zaključka da za sve $n \in \mathbb{N}_0$ vrijedi

$$d\mathbb{P}^*_{|\mathcal{F}_n} = W_n\, d\mathbb{P}_{|\mathcal{F}_n} \Rightarrow \frac{d\mathbb{P}^*_{|\mathcal{F}_n}}{d\mathbb{P}_{|\mathcal{F}_n}} = W_n.$$

Uz pomoć Teorema 4.3.5. iz R.Durrett [3], dobivamo da za svaki izmjerivi $A \subset \mathbb{T}$

$$\mathbb{P}^*(A) = \int_A W_\infty\, d\mathbb{P} + \mathbb{P}^*\left(A \cap \{W_\infty = +\infty\}\right).$$

Uzimajući $A = \Omega$, prethodna jednakost nam daje sljedeće.

- $\mathbb{E}\left[W_\infty\right] = 1 \Leftrightarrow \mathbb{P}^*\left(W_\infty = +\infty\right) = 0.$
- $\mathbb{P}\left(W_\infty = 0\right) = 1 \Leftrightarrow \mathbb{P}^*\left(W_\infty = +\infty\right) = 1.$

Dakle to nam omogućava da promatramo $\mathbb{P}^*$-ponašanje od W_∞ što će se pokazati puno jednostavnijim.

Neka je $(\check{p}_k\ :\ k \in \mathbb{N}_0)$ distribucija pristrana po veličini distribucije $(p_k\ :\ k \in \mathbb{N}_0)$ te neka je Y slučajna varijabla takva da je slučajna varijabla $Y + 1$ distribuirana s $(\check{p}_k\ :\ k \in \mathbb{N}_0)$ na skupu $\mathbb{N}_0$. Neka su $(\check{p}_k\ :\ k \in \mathbb{N}_0)$ nezavisne slučajne varijable distribuirane kao slučajna varijabla X. Također, neka su $(Y_n\ :\ n \in \mathbb{N})$ nezavisne slučajne varijable distribuirane kao slučajna varijabla X te nezavisne od $(\check{p}_k\ :\ k \in \mathbb{N}_0)$. Definiramo $Z_0^* = 0$ i za sve $n \in \mathbb{N}$ neka je

$$Z_n^* = Y_n + \sum_{i=1}^{Z_{n-1}^*} \alpha_{n,i},$$

uz konvenciju da je $\sum_\emptyset = 0$. Po konstrukciji nam je jasno da je $(Z_n^* + 1\ :\ n \in \mathbb{N}_0)$ jednako distribuiran kao $(z_n(\tau^*)\ :\ n \in \mathbb{N}_0)$. Jednostavnim računom vidimo da je $\mathbb{P}^*$-distribucija od $(W_n\ :\ n \in \mathbb{N}_0)$ jednaka distribuciji od $(W_n^* + m^{-n}\ :\ n \in \mathbb{N}_0)$, za $W_n^* = \frac{Z_n^*}{m^n}$. Naime, za proizvoljan $\boldsymbol{t} \in \mathbb{T}$ i $n \in \mathbb{N}_0$ imamo

$$\mathbb{P}^*(W_n = t) = \mathbb{P}\left(\frac{z_n(\tau^*)}{m^n} = \boldsymbol{t}\right) = \mathbb{P}\left(\frac{Z_n^* + 1}{m^n} = \boldsymbol{t}\right) = \mathbb{P}\left(W_n^* + m^{-n} = \boldsymbol{t}\right).$$

(I.) Pretpostavimo da je $\mathbb{E}\left[X \ln^+(X)\right] < +\infty$. Koristeći monotonost očekivanja i definicije distribucije pristrane po veličini odmah nam slijedi da je $\mathbb{E}\left[\ln^+(Y)\right] < +\infty$. Sada iz Leme 2.5.8 imamo da za proizvoljan $\epsilon > 0$ je $Y_n \leq e^{n\epsilon}$ za dovoljno velike n $\mathbb{P}$-g.s.. Označimo sa $\mathcal{Y} = \sigma((Y_n \,:\, n \in \mathbb{N}))$ te sa $(\mathcal{F}_n^* \,:\, n \in \mathbb{N}_0)$ filtraciju generiranu sa $(W_n^* \,:\, n \in \mathbb{N}_0)$. Sada dobivamo

$$\begin{aligned}
\mathbb{E}\left[W_n^* \,|\, \mathcal{Y}, \mathcal{F}_{n-1}^*\right] &= \frac{1}{m^n}\mathbb{E}\left[Z_n^* \,|\, \mathcal{Y}, \mathcal{F}_{n-1}^*\right] \\
&= \frac{1}{m^n}\mathbb{E}\left[Y_n \,|\, \mathcal{Y}, \mathcal{F}_{n-1}^*\right] + \frac{1}{m^n}\mathbb{E}\left[\sum_{i=1}^{Z_{n-1}^*} \alpha_{n,i} \,\middle|\, \mathcal{Y}, \mathcal{F}_{n-1}^*\right] \\
&= \frac{Y_n}{m_n} + \frac{Z_{n-1}^*}{m^{n-1}} \\
&\geq W_{n-1}^*.
\end{aligned}$$

Dolazimo do zaključka da je $(W_n^* \,:\, n \in \mathbb{N}_0)$ nenegativni submartingal u odnosu na $\mathbb{P}(\cdot|\mathcal{Y})$. Dodatno,

$$\begin{aligned}
\mathbb{E}\left[W_n^* \,|\, \mathcal{Y}\right] &= \frac{Y_n}{m^n} + \mathbb{E}\left[W_{n-1}^* \,|\, \mathcal{Y}\right] \\
&\;\;\vdots \\
&= \sum_{k=1}^{n} \frac{Y_k}{m^k} \\
&\leq \sum_{k=1}^{+\infty} \frac{Y_k}{m^k}.
\end{aligned}$$

Kao što smo ranije pokazali, za proizvoljan $\epsilon > 0$ i za dovoljno velike n je $Y_n \leq e^{n\epsilon}$ $\mathbb{P}-$ g.s. te zaključujemo da je $\mathbb{P}-$ g.s. $\sup_{n\in\mathbb{N}_0} \mathbb{E}\left[W_n^* \,|\, \mathcal{Y}\right] < +\infty$. Koristeći tu činjenicu i činjenicu da je $(W_n^* \,:\, n \in \mathbb{N}_0)$ nenegativni sub-martingal u odnosu na $\mathbb{P}(\cdot|\mathcal{Y})$, pomoću Teorema 1.3.3 znamo da postoji $W_\infty^* = \lim_{n\to+\infty} W_n^*$ $\mathbb{P}(\cdot \,|\, \mathcal{Y})-$ g.s. te vrijedi $\mathbb{E}\left[W_\infty^* \,|\, \mathcal{Y}\right] < +\infty$. Kako je $\mathbb{E}\left[W_\infty^*\right] < +\infty$, a $(W_n^* + m^{-n} \,:\, n \in \mathbb{N}_0)$ jednako distribuiran kao $(W_n \,:\, n \in \mathbb{N}_0)$ pod $\mathbb{P}^*$, vrijedi da je $\mathbb{E}^*\left[W_\infty\right] < +\infty \Rightarrow \mathbb{P}^*(W_\infty = +\infty) = 0 \Leftrightarrow \mathbb{E}\left[W_\infty\right] = 1$. Iz Leme 2.5.6 imamo da je $\mathbb{P}(W_\infty = 0) \in \{q,\, 1\}$ pa uz $\mathbb{E}\left[W_\infty\right] = 1$, lako dolazimo do zaključka da je $\mathbb{P}(W_\infty = 0) = q$.

(II.) Pretpostavimo da je $\mathbb{E}\left[X \ln^+(X)\right] = +\infty$. Analogno kao prije zaključujemo da je $\mathbb{E}\left[\ln^+(Y)\right] = +\infty$ te koristeći Lemu 2.5.8 znamo da za svaki $\epsilon > 0$ je $Y_n \geq e^{n\epsilon}$ za beskonačno mnogo $n-$ova. Sada, kako je $Z_n^* \geq Y_n$ i $(W_n^* + m^{-n} \,:\, n \in \mathbb{N}_0)$ jednako distribuiran kao $(W_n \,:\, n \in \mathbb{N}_0)$ pod $\mathbb{P}^*$, lako dolazimo do zaključka da je $W_n \geq e^{n\left(-\ln(m)+\frac{1}{\epsilon}\right)}$, za beskonačno mnogo n-ova $\mathbb{P}^*-$ g.s..Jasno nam je da kada uzimamo dovoljno male ϵ, imamo $\lim_{n\to+\infty} W_n = W_\infty = +\infty$ $\mathbb{P}^*-$ g.s. $\Rightarrow \mathbb{P}^*(W_\infty = +\infty) = 1 \Leftrightarrow \mathbb{P}(W_\infty = 0) = 1$. $\square$

Bibliografija

[1] Romain Abraham i Jean François Delmas, *An introduction to Galton-Watson trees and their local limits*, 2015, `https://arxiv.org/abs/1506.05571`.

[2] Gerold Alsmeyer, *Übungen zur Vorlesung Spezielle Stochastische Prozesse im WS 2010/2011i*, 2010, `https://www.uni-muenster.de/Stochastik/lehre/WS1011/SpezielleStochastischeProzesse/`.

[3] Rick Durrett, *Probability: theory and examples*, sv. 49, Cambridge university press, 2019.

[4] Russell Lyons, Robin Pemantle i Yuval Peres, *Conceptual Proofs of L Log L Criteria for Mean Behavior of Branching Processes*, The Annals of Probability **23** (1995), br. 3, 1125–1138, ISSN 00911798, `http://www.jstor.org/stable/2244865`, posjećena 2022-09-04.

[5] Jacques Neveu, *Arbres et processus de Galton-Watson*, Annales de l'IHP Probabilités et statistiques, sv. 22, 1986, str. 199–207.

[6] Nikola Sandrić i Zoran Vondraček, *Vjerojatnost–predavanja. 2019*.

[7] Zoran Vondraček, *Markovljevi lanci*, PMF-MO skripta (2008).

[8] Zoran Vondraček, *Slučajni procesi*, Prirodoslovno-matematički fakultet, Zagreb (2010).

Sažetak

Galton-Watsonov proces ili jednostavan proces grananja je prvi stohastički model razvoja populacije. Proces kreće od jedne jedinke te se svaka jedinka razmnožava jednako i nezavisno od ostalih s određenom distribucijom potomaka. U ovom diplomskom radu bavimo se Galton-Watsonovim stablima. Najprije uvodimo skup diskretnih stabala, definiciju stabla i određene podskupove stabla koji će nam biti od interesa kao što su: širina stabla na određenoj razini, listovi, podstablo stabla iznad određenog čvora i podstablo stabla s određenom visinom. Nakon toga smo spremni definirati Galton-Watsonovo stablo, koje ustvari predstavlja Galton-Watsonov proces kada promatramo širinu stabla na određenoj razini, odnosno broj jedinki iste generacije. S obzirom na distribuciju potomaka, preciznije njenom očekivanju m, razlikujemo kritična ($m = 1$), sub-kritična ($m < 1$) i super-kritična ($m > 1$) Galton-Watsonova stabla. Potvrđujemo da je vjerojatnost izumiranja najmanja fiksna točka funkcije izvodnice distribucije potomaka stabla na intervalu $[0,\ 1]$, međutim ona je jednaka 1 samo za kritična i sub-kritična stabla. Također, super-kritično stablo uvjetno na događaj izumiranja je distribuirano jednako kao sub-kritično stablo s određenom distribucijom potomaka. Na kraju pokazujemo da se na događaju preživljavanja veličina populacije Galton-Watsonovog procesa sa super-kritičnom distribucijom ponaša kao konačna slučajna konstanta pomnožena s m^n, u slučaju kada je $\mathbb{E}\left[X \ln^+(X)\right] < +\infty$, gdje je X slučajna varijabla s distribucijom koja je jednaka super-kritičnoj distribuciji potomaka procesa.

Summary

The Galton-Watson process or the simple branching process is the first stochastic model for population evolution. The process starts from one individual and each individual reproduces equally and independently from the others with a given offspring distribution. In this thesis we deal with Galton-Watson trees. Firstly, we introduce a set of discrete trees, a definition of a tree and certain subsets of a tree that will be of interest to us such as: tree width at a certain level, the set of leaves of a tree, a subtree above a certain node of a tree and a subtree below a certain height of a tree. Subsequently we are ready to define the Galton-Watson tree, which actually represents the Galton-Watson process when we observe the width of the tree at a certain level. In other words, we observe the number of individuals of the same generation. Considering the offspring distribution, more precisely its mean m, we distinguish critical ($m = 1$), sub-critical ($m < 1$) and super-critical ($m > 1$) Galton-Watson trees. We confirm that the probability of extinction is the smallest fixed point of the generating function of the tree's offspring distribution on the interval $[0, 1]$. However, it is equal to 1 only for critical and sub-critical trees. Also, the super-critical tree conditionally on the extinction event is equally distributed as a sub-critical tree with a special offspring distribution. Finally, we show that at the survival event the population size of the Galton-Watson process with super-critical distribution behaves as a finite random constant multiplied by m^n, when $\mathbb{E}\left[X \ln^+(X)\right] < +\infty$, where X is a random variable with a distribution equal to the super-critical offspring distribution of the process.

Životopis

Rođen sam 21.02.1999. godine u Zagrebu, a veliki dio djetinjstva proveo sam u Bratislavi, Slovačkoj. Izuzev 3. i 4. razreda osnovne škole koje sam pohađao u Zagrebu (OŠ Silvije Strahimir Kranjčević), osnovnu školu upisujem i završavam u Bratislavi u The British International School Bratislava. U ljeto 2013. godine upisujem XVI. Gimnaziju u Zagrebu, dvojezičan razred. Kao učenik jezične gimnazije afinitet prema matematici stječem uz pomoć profesorice matematike. Zbog toga po završetku srednjoškolskog obrazovanja 2017. godine upisujem preddiplomski sveučilišni studij Matematike na Prirodoslovno-Matematičkom fakultetu u Zagrebu. 2020. godine obrazovanje nastavljam na diplomskom sveučilišnom studiju Financijska i poslovna matematika.