Poglavlje 0

Uvod

Razne prirodne i društvene pojave odvijaju se na slučajan način. Teorija slučajnih procesa modelira evoluciju tih slučajnih pojava kroz vrijeme. Markovljevi lanci predstavljaju jedan od najvažnijih (a također i najjednostavnijh) modela slučajne evolucije. U ovom uvodnom poglavlju promotrit ćemo nekoliko primjera Markovljevih lanaca.

Primjer 0.1 Jednostavni proces grananja. Populacija nekih jedinki započinje u trenutku $n = 0$ s jednom jedinkom koju nazivamo predak. Predak čini nultu generaciju populacije. U trenutku $n = 1$ predak se podijeli na slučajan broj potomaka (ili daje, rađa, taj slučajan broj potomaka), a sam nestaje iz populacije (umire). Vjerojatnost da predak daje $k \geq 0$ potomaka je p_k, $0 \leq p_k \leq 1$. Ti potomci čine prvu generaciju. Svaki od tih potomaka podijeli se u trenutku $n = 2$, nezavisno od drugih jedinki u populaciji, na slučajan broj potomaka (a sam umire). Taj slučajan broj određen je istom vjerojatnosnom distribucijom $(p_k : k \geq 0)$. Proces se nastavlja u trenucima $n = 3$, $n = 4$, i tako dalje, sve dok u populaciji postoje jedinke.

Ovako opisan proces zove se *jednostavni proces grananja* ili *Galton-Watsonov proces*, te može služiti kao model rasta neke populacije na koju ne utječu vanjski efekti. Povijesno se proces grananja pojavio u posve drugom kontekstu, naime u proučavanju preživljavanja aristokratskih porodičnih imena u viktorijanskoj Engleskoj. Naime, primjećeno je da neka aristokratska prezimena odumiru (nestaju), što je dovelo do zabrinutosti u nekim krugovima. Francis Galton je 1873. godine postavio pitanje koliko muških potomaka je potrebno da porodično ime ne odumre u budućim generacijama. Odgovor je godinu dana kasnije dao Henry William Watson (otkud ime procesu).

Označimo sa X_n broj jedinki u n-toj generaciji, $n \geq 0$. Po pretpostavci je $X_0 = 1$. S druge strane, već X_1 je slučajan, te ga modeliramo slučajnom varijablom na nekom vjerojatnosnom prostoru. Očito su i X_n, $n \geq 2$, slučajne varijable i primaju vrijednosti u $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$. Familija $X = (X_n : n \geq 0)$ naziva se slučajan proces s prostorom stanja $\mathbb{Z}_+$. O čemu ovisi broj jedinki X_n u n-toj generaciji? Nakon kraćeg razmišljanja uvjerit ćete se da ovisi samo o broju jedinki X_{n-1} u $n-1$-voj generaciji (i očigledno o vjerojatnosnoj distribuciji potomaka $(p_k : k \geq 0)$), ali uvjetno na poznavanje X_{n-1}, ne ovisi o broju jedinki u generacijama koje prethode $n-1$-voj. Takvo svojstvo slučajnog proces X naziva se *Markovljevim svojstvom.*

Osnovno pitanje vezano uz jednostavni proces grananja je vjerojatnost izumiranja populacije, t.j., vjerojatnost da u nekoj generaciji više nema jedinki. Uočite da ako je vjerojatnost nula potomaka $p_0 > 0$, populacija izumire s pozitivnom vjerojatnošću već nakon nulte generacije (predak nema niti jednog potomka). Matematički, vjerojatnost izumiranja populacije pišemo kao $\mathbb{P}(\cup_{n=1}^{\infty}\{X_n = 0\})$, gdje $\{X_n = 0\}$ označava događaj da je populacija izumrla do (uključivo) n-te generacije. Primjetite da ako je $X_n = 0$, tada je i $X_m = 0$ za sve $m > n$.

Kod *binarnog grananja* broj potomaka je nula ili dva. Neka je $p_2 = p \in (0, 1)$ vjerojatnost dva potomka i $p_0 = q = 1 - p$ vjerojatnost nula potomaka. Izračunajte vjerojatnost izumiranja populacije.

Primjer 0.2 Slučajna šetnja. Slučajni šetač kreće se po cjelobrojnoj rešetci $\mathbb{Z}$ tako da se u svakom vremenskom trenutku $n \geq 1$ pomakne za jedno mjesto udesno ili za jedno mjesto ulijevo. Odabir smjera kretanja je slučajan. Označimo sa $p \in (0, 1)$ vjerojatnost koraka na desno, te sa $q = 1 - p$ vjerojatnost koraka na lijevo. Neka je X_n, $n \geq 0$, pozicija slučajnog šetača nakon n-koraka, te pretpostavimo da je u trenutku $n = 0$ šetač u ishodištu, $X_0 = 0$. Tada je $X = (X_n : n \geq 0)$ slučajni proces s prostorom stanja $\mathbb{Z}$ koji nazivamo *jednostavna slučajna šetnja*. Slučajna šetnja je simetrična ako je $p = q = 1/2$ (jednake vjerojatnosti koraka na lijevo i na desno). Uočite da, slično kao i u prethodnom primjeru, pozicija X_n ovisi samo o poziciji X_{n-1}, a ne i o prethodnim pozicijima koje je šetač posjetio. [1] Postavimo neka pitanja vezana za kretanje slučajnog šetača:

(i) Kolika je vjerojatnost da šetač ikad dođe u točku $+1$? A u točku $+10^{10^{10}}$?

(ii) Koliki je očekivani broj koraka do dolaska u točku $+1$?

(iii) Kolika je vjerojatnost da će se šetač ikada vratiti u početnu poziciju?

(iv) Koliki je očekivani broj posjeta stanju $+1$ prije prvog povratka u ishodište? A očekivani broj posjeta stanju $+10^{10^{10}}$ prije prvog povratka u ishodište?

Da li možete odgovoriti na ta pitanja za simetričnu jednostavnu slučajnu šetnju?

Primjer 0.3 Promotrimo slučajno kretanje skakača po šahovskoj ploči. U trenutku $n = 0$ skakač se nalazi u donjem lijevom kutu šahovske ploče. U trenutku $n = 1$ skakač skoči na jedno od mogućih mjesta s vjerojatnosti obrnuto proporcionalnoj broju mogućih mjesta, i tako dalje. Označimo sa X_n, $n \geq 0$, slučajnu poziciju skakača nakon n koraka. Tada je $X = (X_n : n \geq 0)$ slučajni proces sa prostorom stanja jednakim svim pozicijama šahovske ploče. Uočite da je opet pozicija X_n ovisna samo o X_{n-1}.

Označimo sa T (slučajno) vrijeme prvog povratka skakača u donji lijevi kut šahovske ploče. Da li možete izračunati distribuciju slučajne varijable T? Da li možete izračunati očekivanje $\mathbb{E}T$?

[1] Preciznije, uvjetno na poznavanje X_{n-1}, slučajna varijable X_n nezavisna je od $X_0, X_1, \ldots, X_{n-2}$.

Primjer 0.4 Osiguravajuća društva često u osiguranju motornih vozila primjenjuju bonus sustav (ili bonus-malus sustav). Ideja bonus sustava je da nagrađuje vozače koji nemaju (ili ne prijavljuju) štetu tokom prethodne godine manjom premijom u sljedećoj godini, dok vozači koji prijave štetu plaćaju veću premiju. Sustav bonusa obično se sastoji od nekoliko kategorija popusta. Na primjer, jednostavan sustava bonusa može imati šest kategorija popusta: u nultoj kategoriji vozač plaća punu premiju, u prvoj kategoriji ima 10% popusta na punu premiju, sve do pete kategorije u kojoj plaća 50% pune premije. Osim kategorija popusta, bonus sustav definiran je i pravilima prijelaza između kategorija. Pretpostavimo vrlo jednostavna pravila: ako tokom godine vozač ne prijavi štetu prelazi u sljedeću višu kategoriju popusta (osim ako je već na 50% popusta gdje u tom slučaju i ostaje). Vozač koji prijavi štetu tokom godine odmah pada u najnižu kategoriju popusta (t.j., sljedeće godine plaća punu premiju). Želimo modelirati kretanje nekog određenog vozača po kategorijama kroz više godina. Označimo prvo skup kategorija sa $S = \{0, 1, 2, 3, 4, 5\}$ gdje na primjer element 3 označava kategoriju s 30% popusta. Neka je nadalje $X_n \in S$ kategorija popusta u kojoj se vozač nalazi u n-toj godini, te pretpostavimo da je $X_0 = 0$ - vozač koji ulazi u sustav nema popust. Budući da je događaj da vozač unutar godine prijavi štetu slučajan, to je $X = (X_n : n \geq 0)$ slučajni proces. Odmah se vidi da kategorija popusta u kojoj se vozač nalazi u godini n ovisi samo o kategoriji popusta u godini $n - 1$ (i naravno o tome da li je vozač prijavio štetu), a ne ovisi o kategorijama popusta u kojima se vozač nalazio tokom bilo koje prethodne godine.[2] Dakle, X ima Markovljevo svojstvo.

Pretpostavimo (nerealistički) da je sustav bonusa zatvoren, t.j., da kroz niz godina vozači ne izlaze iz sustava, niti ne ulaze u sustav. Osiguravajuće društvo je zainteresirano za postotak vozača u pojedinoj kategoriji popusta nakon većeg broja godina. Označimo sa $p_{0j}^{(n)}$ vjerojatnost da se naš tipični vozač, koji kreće iz kategorije 0, nakon n godina nalazi u kategoriji $j \in S$. Ako postoji granična vrijednost $\pi_j = \lim_{n\to\infty} p_{0j}^{(n)}$, onda $p_{0j}^{(n)}$ možemo (dobro ili loše) aproksimirati s π_j. Vjerojatnosnu distribuciju $\pi = (\pi_j : j \in S)$ nazivamo graničnom distribucijom. Ako svi vozači u nultoj godini kreću iz kategorije 0, tada će ta granična vrijednost π_j predstavljati postotak ukupnog broja vozača koji se nalaze u kategoriji j.

Drugo zanimljivo pitanje za osiguravajuće društvo je da li postoji neka distribucija vozača po kategorijama tako da ako su u godini n vozači tako distribuirani, tada će u godini $n+1$ distribucija ostati nepromijenjena. Ako takva distribucija postoji, zvat ćemo je stacionarnom. Na nivou slučajnog procesa $(X_n : n \geq 0)$ (odnosno konkretnog vozača), to znači da je $\mathbb{P}(X_n = j) = \mathbb{P}(X_{n+1} = j)$, $j \in S$.

Jedan od važnih rezultata u teoriji Markovljevih lanaca govori o nužnim i dovoljnim uvjetima uz koje postoje granična, odnosno stacionarna, distribucija, te da su one jednake. Ti uvjeti su zadovoljeni za slučajni proces $(X_n : n \geq 0)$ vezan uz sustav bonusa.

[2] Opet, uvjetno na poznavanje kategorije u kojoj se vozač nalazi u godini $n - 1$.