

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
ODDELEK ZA FIZIKO

MODELSKA ANALIZA I
9. naloga: Stohastični populacijski modeli

Žiga Šinigoj, 28222025

Ljubljana, december 2022

1 Eksponentni model izumrtja populacije

Najprej obravnavam preprost eksponentni model izumrtja, kjer je dinamika stohastične narave

$$\Delta N = - P(\beta N \Delta t) , \quad (1)$$

kjer P označuje Poissonsko izžrebano število s povprečjem $\beta N \Delta t$. Povprečje Poissonove porazdelitve je odvisno od hitrosti umiranja β in časovnega koraka. Najprej sem pogledal časovni potek smrti pri različnih časovnih korakih in pri velikosti populacije $N = 25$, podoben je tudi potek za $N=250$. Poleg tega je prikazana še analitična rešitev (slika 1). Ker je simulacija stohastične narave se čas izumrtja razlikuje pri različnih simulacijah. Če je časovni korak prevelik se lahko zgodi, da v eni potezi izumre celotna populacija. Nekaj takih primerov (premic) je videti na sliki 1 s časovnim korakom $dt = 0.7$.

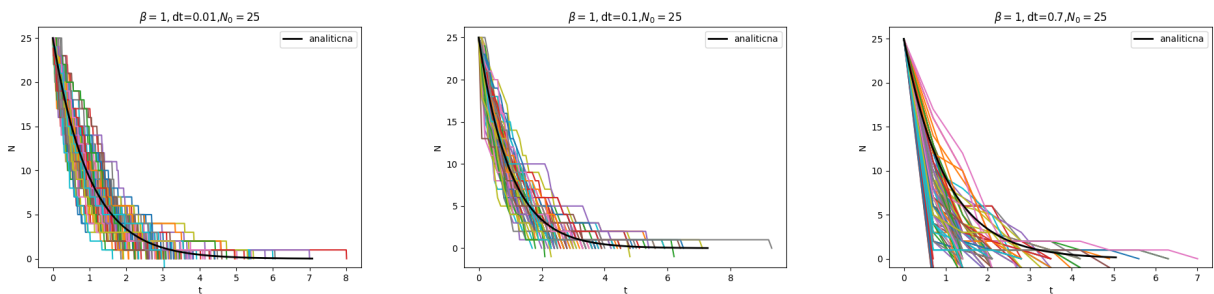
Model lahko izboljšam tako, da dodam še rojstva

$$\Delta N = - P(\beta_s N \Delta t) + P(\beta_r N \Delta t) , \quad (2)$$

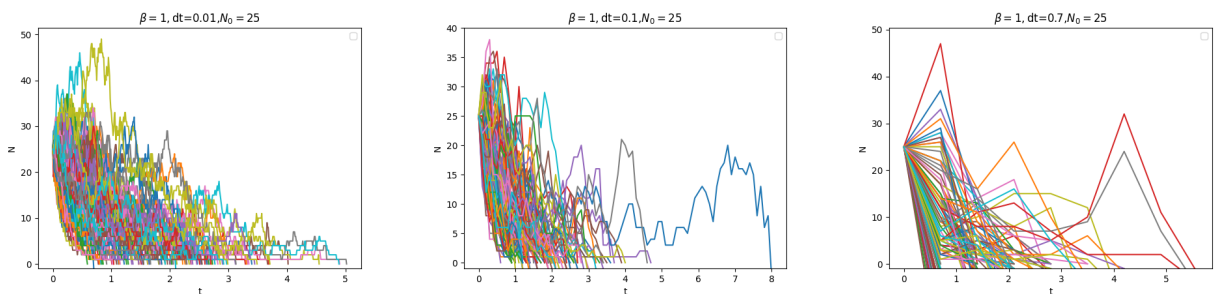
kjer sedaj v vsaki iteraciji izžrebam 2 naključni števili iz Poissonove porazdelitve z danima pričakovanima vrednostima. Naj za parametra β_i velja $\beta_s/\beta_r = 5/4$. Iz tega sledi

$$\Delta N = - P(5\beta N \Delta t) + P(4\beta N \Delta t) . \quad (3)$$

Dinamika in analitična rešitev je enaka kot v primeru smrti, vendar želimo sedaj obravnavati rojstvo in smrt kot 2 ločena dogodka. Primer dinamike pri začetni velikosti populacije $N_0 = 25$ prikazuje slika 2. V primeru rojstev lahko populacija poskoči čez začetno vrednost populacije.



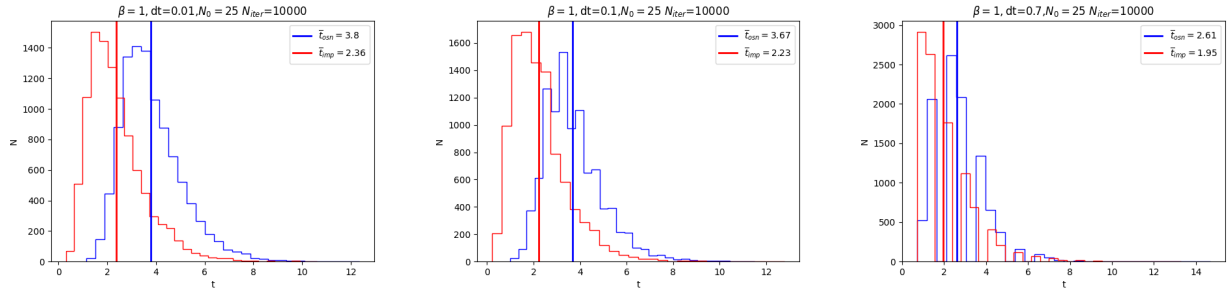
Slika 1: Časovna evolucija populacije za tri različne časovne korake v modelu, kjer ljudje samo umirajo. Število simuliranih evolucij:100



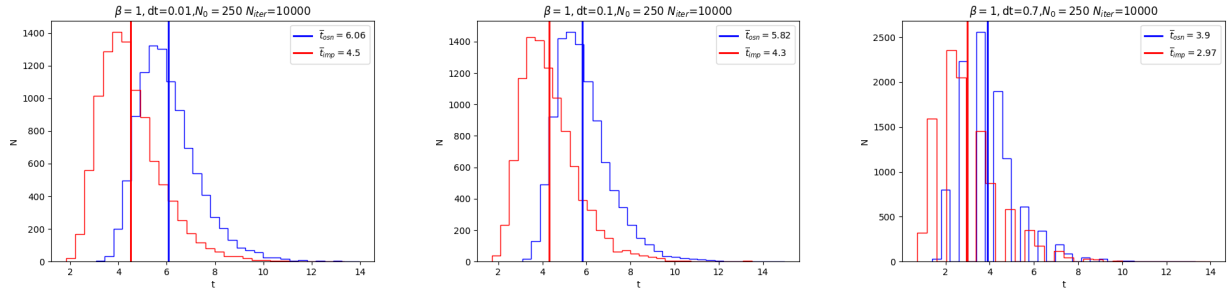
Slika 2: Časovna evolucija populacije za tri različne časovne korake v modelu, kjer se ljudje rojevajo in umirajo. Število simuliranih evolucij:100

Zanima nas statistika časov smrti populacije za 2 različni velikosti populacije $N_0 = 25$ in $N_0 = 250$. Statistiko za manjšo populacijo pri različnih časovnih korakih ($\beta = 1$) prikazuje slika 3 za oba modela. V modelu smrti (S) in modelu rojstva-smrti (RS) lahko vidim, da se z manjšanjem časovnega koraka večja povprečni čas izumrtja populacije. To je pričakovan rezultat, saj večji časovni korak pomeni večje spremembe populacije, kar v povprečju vodi do hitrejšega izumrtja. Če je korak velik, lahko v eni iteraciji izumre vsa populacija.

Kar je veliko presenečenje je to, da so povprečni časi RS modela manjši kot časi S modela. To je precej nenavadno, saj bi na prvi pogled trdil, da velja obratno. Kljub temu, da lahko imamo ob neničelnem času več populacije kot na začetku in da se populacija lahko večja, so povprečni časi manjši. Najverjetneje je razlog je v tem, da je raztresenost okrog analitične vrednosti dvakrat večja v primeru ko izžrebam 2 naključni števili (RS model) in ker so časi asimetrično porazdeljeni glede na povprečno vrednost, se pozna, če se raztresenost poveča za dvakrat v premiku povprečja. Statistiko za velikost populacije $N_0 = 250$ prikazuje slika 4. Za desetkrat večjo populacijo so povprečni časi izumrtja nekaj manj kot dvakrat večji. Pri večjih populacijah so tudi deleži umrlih večji. Razmik med povprečnim časom S in RS modela se bistveno spremeni samo tam, kjer je $dt = 0.7$, kar pa mislim da je posledica prevelikega časovnega koraka in posledično nenatančnosti modela. Za manjše korake zgleda, da je razmik približno konstanten.

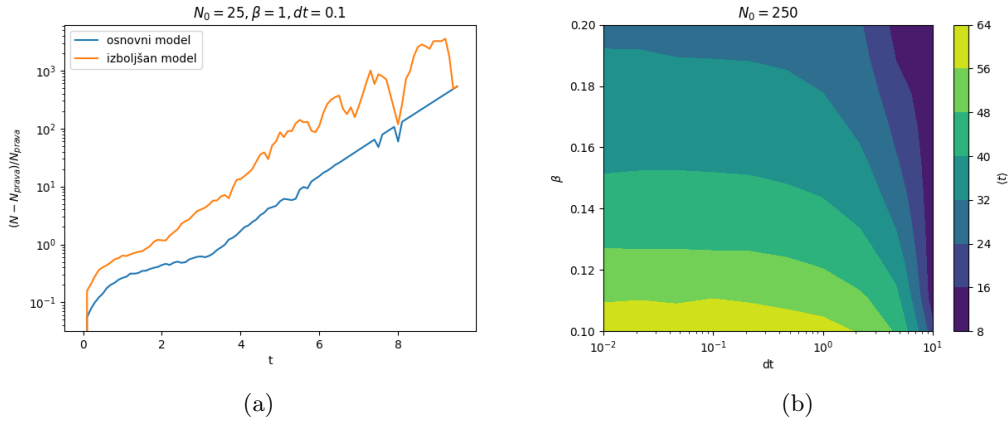


Slika 3: Statistika smrti populacije pri različnih velikostih časovnega koraka. N_{iter} označuje število dobljenih časov izumrtja.

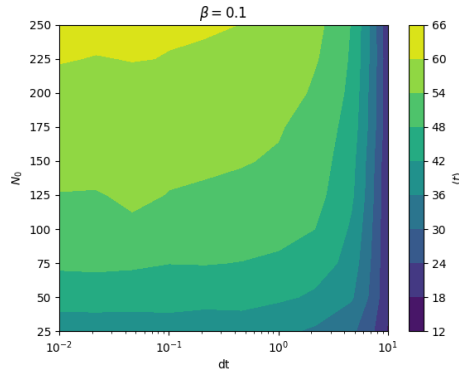


Slika 4: Statistika smrti populacije pri različnih velikostih časovnega koraka. N_{iter} označuje število dobljenih časov izumrtja.

Relativna odstopanja od analitične vrednosti modela prikazuje slika 5a, kjer sem povprečil odstopanja 100 simulacij od analitične rešitve. Vidim lahko, da je so odstopanja večja v RS modelu, kar je pričakovano, glede na to da dvakrat izžrebam naključno število. S časom se razpršenost oziroma varianca večja. Zanimivo je pogledati tudi odvisnost povprečnega časa izumrtja od parametra β in časovnega koraka dt (slika 5b). V parametru β nisem zasledil kakšne posebne odvisnosti. Z večanjem parametra pospešujem dinamiko populacije in z večjim parametrom dobim pričakovano manjši povprečni čas. Mogoče je opaziti odvisnost od časovnega koraka, za velike časovne korake je povprečni čas zelo odvisen od koraka. Če korak zmanjšam dobim natančnejšo rešitev, ki postane neodvisna od časovnega koraka-dobim stabilno rešitev. V ta namen sem v nalogi časovni korak 0.1 oziroma 0.01. Odvisnost povprečnega časa od velikosti populacije in časovnega koraka prikazuje slika 6. V velikosti populacije ni presenetljive odvisnosti, večja kot je populacija, večji je povprečni čas.



Slika 5: a) Relativno odstopanje povprečne vrednosti 100 simulacij od analitične rešitve za Rs model (izboljsan model) in S model (osnovni model) v odvisnosti od časa. b) Povprečni čas izumrtja v odvisnosti od parametra β in dt .



Slika 6: Povprečni čas izumrtja v odvisnosti od velikosti populacije in časovnega koraka.

1.1 Matrika prehodov

Nalogo je mogoče rešiti tudi z matriko prehodov. Stanje sistema opišem z vektorjem $\mathbf{x}(t)$, ki je velikosti 26 in 251 za model S in 51, 301 za model RS. Vsaka komponenta vektorja predstavlja število osebkov v populaciji (npr. $\mathbf{x} = (0, 0, 0, 0, 1)$ predstavlja 4 osebkke populacije). Rešitev ob naslednjem časovnem koraku dobimo kot

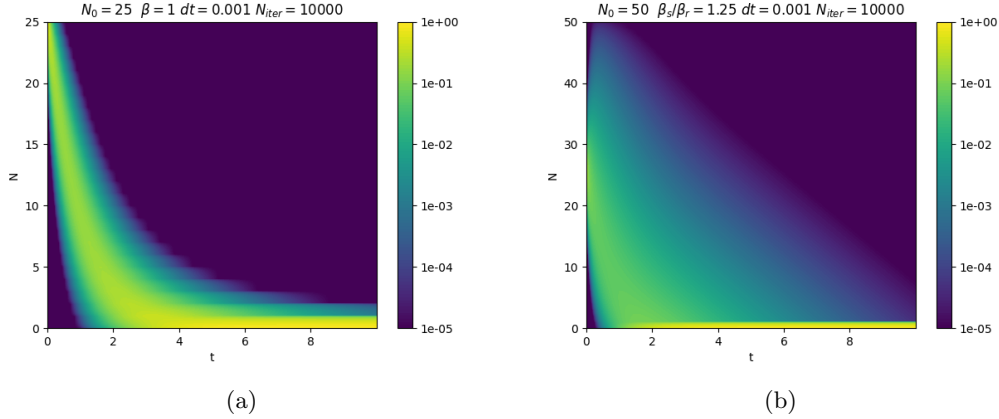
$$\mathbf{x}(t + \Delta t) = M\mathbf{x}(t), \quad (4)$$

kjer je M prehodna matrika. Prehodna matrika mora ohranjati verjetnost, torej mora biti unitarna. Prehodno matriko smo zapisali na predavanjih in je znana, zato jo ne bom pisal še enkrat. Pri obravnavi dinamike se pojavi vprašanje kakšen naj bo časovni korak, da bom ohranil verjetnost. Za unitarne matrike velja $\det(M) = |1|$, zato sem pogledal determinanto matrike M v odvisnosti od velikosti matrike in časovnega koraka (slika 9b). Za velike vrednosti časovnega koraka postane determinanta zelo velika, zato sem vse vrednosti večje od 1 in manjše od 10^{-5} nadomestil z vrednostjo 10^{-5} . Za matrike večje od 400 mora biti korak manjši od 10^{-6} , da ohranim verjetnost. Simulacijo sem zaustavil, ko je bila verjetnost, da je populacija izumrla 99.9 %. Za prvi del naloge, kjer je velikost vektorja 26 oz 51 lahko vzamem relativno velik časovni korak. Za majhne matrike lahko dobim rešitev ob danem času kot

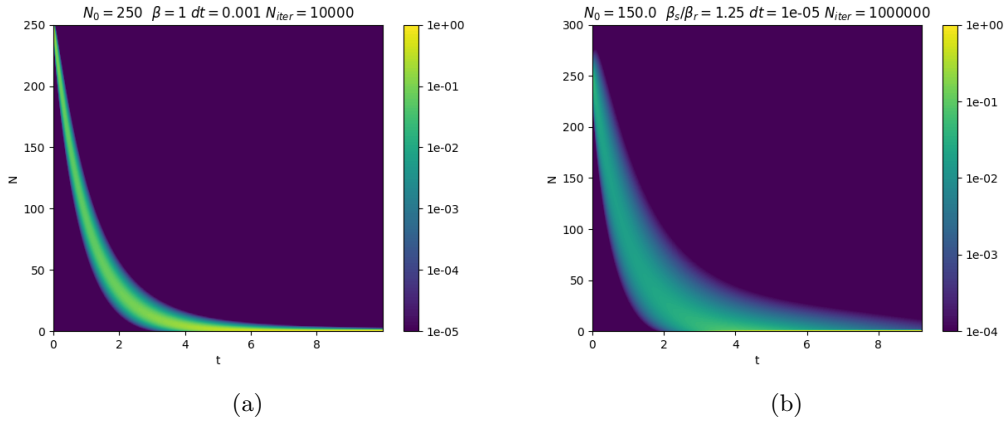
$$\mathbf{x}(t_n) = M^n \mathbf{x}(0), \quad (5)$$

ki je za majhne sisteme uporabna, če zagotovimo dovolj majhen časovni korak. Pri velikih matrikah pa lahko povzroči veliko težav. Če je na primer $\det(M) = 0.999$, bo na primer 1000 korak evolucije zahteval 1000 eksponent matrike, kar bo dalo napačen rezultat, saj bo tolikšno potenciranje povzročilo izgubo unitarnosti. V takih primerih je bolje reševati iterativno. Rešitev dinamike s prehodno matriko modela S in RS pri velikosti populacije $N_0 = 25$, prikazuje slika 7. Barve označujejo verjetnost, da je ob določenem času

določeno število ljudi v populaciji. Pri modelu RS lahko opazim večjo razpršenost, ki je posledica rojstev v populaciji. Rešitev pri večji populaciji prikazuje slika 8. V tem primeru je potrebno vzeti manjši korak, sploh takrat, ko uporabim večjo matriko (RS model). Porazdelitev je prav tako širša v primeru, ko upoštevam tudi rojstva. Pri modelu RS na prvi pogled zgleda, da se verjetnost ne ohranja, saj ob končnem času ni videti tankega rumenega pasa, ki je posledica premajhnega časovnega koraka. Če približam dinamiko ob končnih časih, lahko vidim, da se verjetnost dobro ohranja (slika 9a) in dobim tanek rumen pas.



Slika 7: a) Časovna evolucija verjetnostne porazdelitve populacije modela S. Barve predstavljajo verjetnost, da je ob danem času dano število osebkov. b) Časovna evolucija verjetnostne porazdelitve populacije modela RS. Barve predstavljajo verjetnost, da je ob danem času dano število osebkov. $N_0 = 25$ (napaka).



Slika 8: a) Časovna evolucija verjetnostne porazdelitve populacije modela S. Barve predstavljajo verjetnost, da je ob danem času dano število osebkov. b) Časovna evolucija verjetnostne porazdelitve populacije modela RS. Barve predstavljajo verjetnost, da je ob danem času dano število osebkov. $N_0 = 250$ (napaka).

Pogledal sem še povprečje in varianco porazdelitve za RS model kot funkciji časa. Iz numeričnih rešitev lahko momenta izračunam kot

$$\mu_i(t) = \sum_{n=0}^N n^i x_n(t), \quad i \in 1, 2, \quad (6)$$

kjer je n število osebkov in $x_n(t)$ verjetnost, da je v populaciji n osebkov. Varianco dobim kot

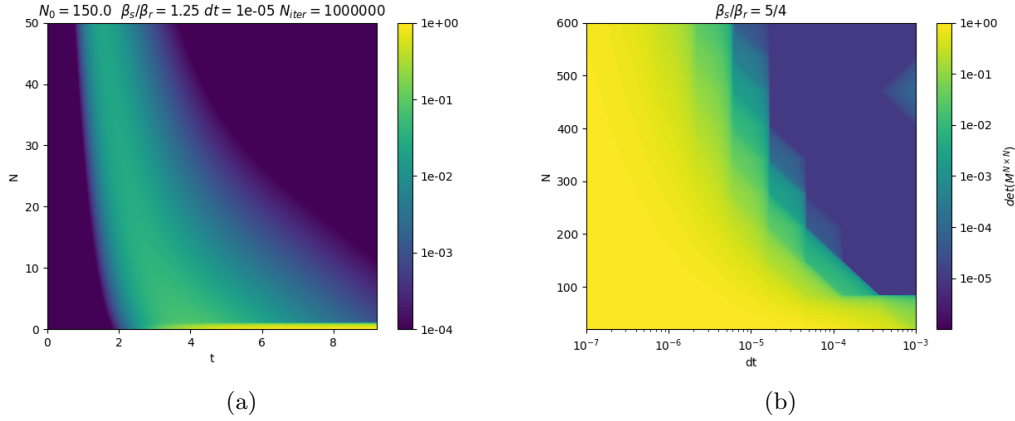
$$\sigma^2(t) = \mu_2(t) - \mu_1(t)^2 \quad (7)$$

Analitično rešitev za povprečje in varianco kot funkciji časa dobim kot

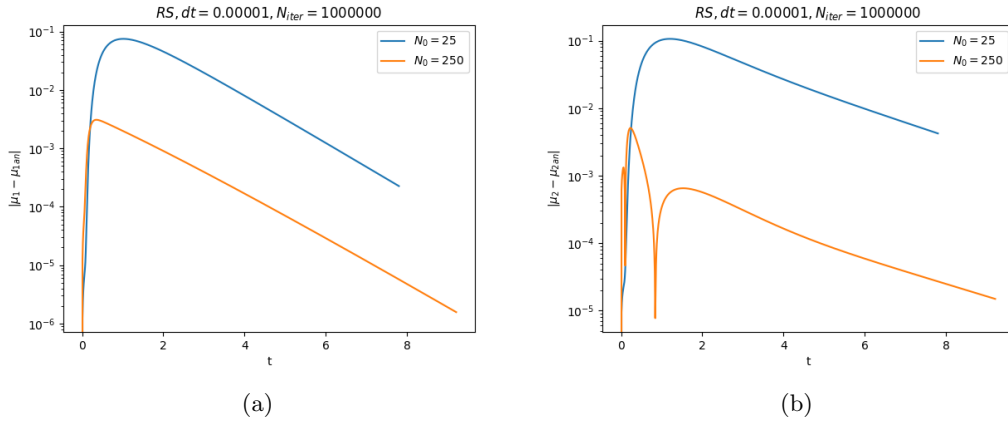
$$\mu_1(t) = N_0 e^{(\beta_r - \beta_s)t} \quad (8)$$

$$\sigma^2(t) = N_0 \frac{\beta_s + \beta_r}{\beta_r - \beta_s} e^{(\beta_r - \beta_s)t} \left(e^{(\beta_r - \beta_s)t} - 1 \right), \quad (9)$$

kjer je N_0 začetna velikost populacije. Odstopanja numeričnih rezultatov od analitičnih prikazuje slika 10. Večji vektor oz. populacija bolj natančno opiše analitično rešitev, saj bolj nadrobim interval ter bolje opišem dinamiko. Imata pa varianca in povprečje približno enak naklon padanja pri $N_0 = 25$ in $N_0 = 250$.



Slika 9: a) Časovna evolucija verjetnostne porazdelitve populacije modela S. Barve predstavljajo verjetnost, da je ob danem času dano število osebkov, približana rešitev. b) Determinanta prehodne matrike v odvisnosti od velikosti matrike in časovnega koraka. Vse vrednosti večje od 1 in manjše od 10^{-5} zavzamejo vrednost 10^{-5} .



Slika 10: a) Absolutna vrednost razlike povprečne vrednosti analitične in numerične rešitve v odvisnosti od časa za model RS. b) Absolutna vrednost razlike analitične in numerične variance v odvisnosti od časa za model RS. (Namesto μ_2 na grafu, bi morala biti varianca σ (napaka)).

2 Stohastični model zajci-lisice

Za model Lotka-Volterra veljajo enačbe

$$\dot{Z} = 5\alpha Z - 4\alpha Z - \frac{\alpha}{L_0} ZL \quad (10)$$

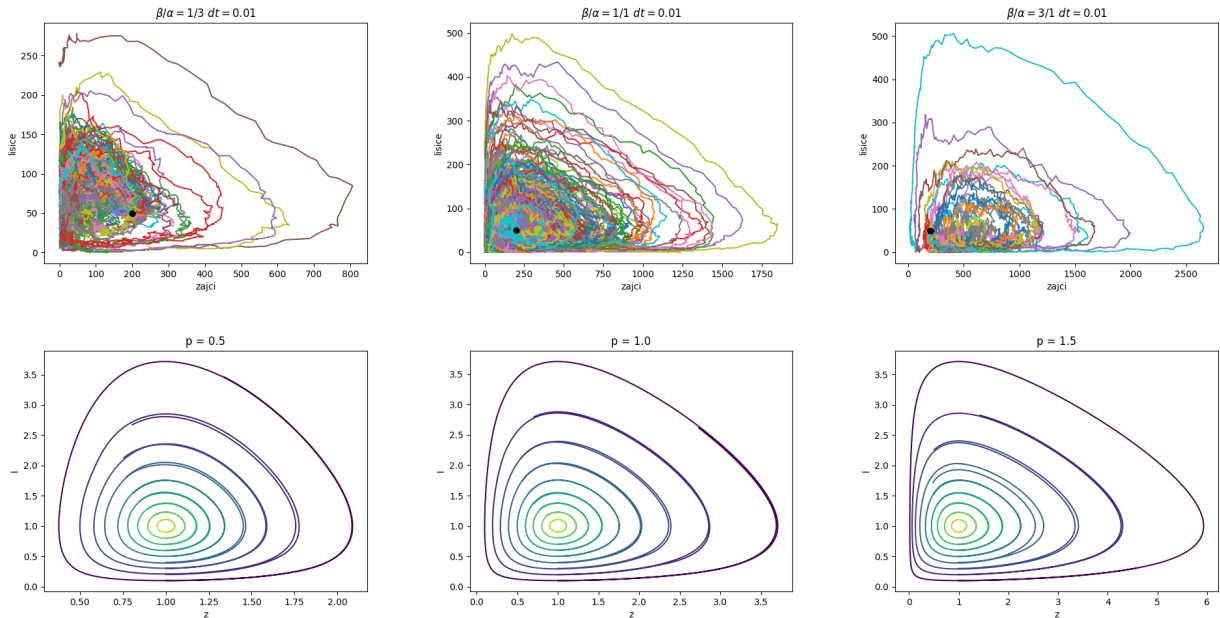
$$\dot{L} = 4\beta L - 5\beta L - \frac{\beta}{L_0} ZL, \quad (11)$$

kjer sta α in β prosta parametra. V stohastični obravnavi se enačbe prepišejo v

$$Z_{n+1} = Z_n + P(5\alpha Z_n \Delta t) - P(4\alpha Z_n \Delta t) - \frac{\alpha}{Z_0} Z_n L_n \quad (12)$$

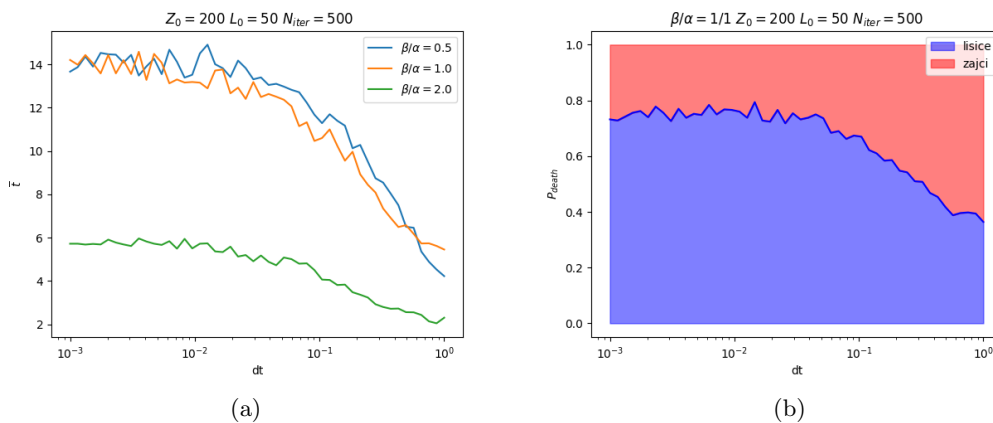
$$L_{n+1} = L_n + P(4\beta L_n \Delta t) - P(5\beta L_n \Delta t) - \frac{\beta}{L_0} Z_n L_n \quad (13)$$

, kjer sem v obeh primerih upošteval, da je razmerje med rodnošjo in smrtnostjo zajcev $5/4$ in obratno za lisice. Začetni pogoj je $L_0, Z_0 = (50, 200)$. Na sliki 11 je prikazanih nekaj stohastičnih potekov populacij za različna prosta parametra in poteki dobljeni iz reševanja diferencialnih enačb. Stohastična obravnava nam zašumi diferencialno rešitev. Tudi tukaj se postavi vprašanje kakšen naj bo časovni korak, da bom dobil dobre rešitve. S tem mislim rešitve, ki bodo neodvisne od časovnega koraka. V ta namen sem povprečil 500 rešitev za dani časovni korak (slika 12a). Za korak $dt = 0.01$ približno dosežem plato. Prav tako je



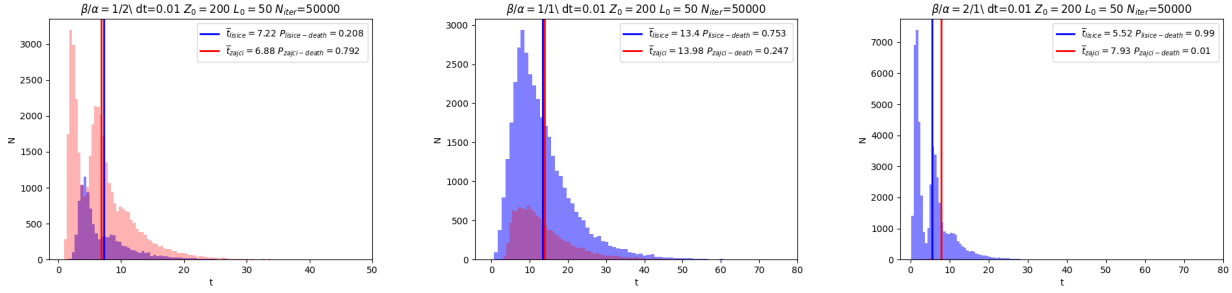
Slika 11: Fazni diagram LV v stohastični (zgoraj) in diferencialni (spodaj) obravnavi za različne vrednosti parametrov α in β . $p = \sqrt{\frac{\alpha}{\beta}}$.

pomembno da je tudi verjetnost za izumrtje dane populacije neodvisna od časovnega koraka (slika 12b). Tudi tukaj se izkaže, da je izbran časovni korak dovolj natančen. V nadaljnji obravnavi sem zato uporabil časovni korak $dt = 0.01$. Sedaj lahko naredim statistiko izumrtja zajcev in lisic, ter izračunam verjetnost da bo dana vrsta izumrla (slika 13). Kar se mi zdi presenetljivo je to, da ob spremembi enega izmed prostih parametrov vplivam na čas izumrtja druge populacije. Po eni strani je to razumljivo, saj sta populaciji sklopljeni, vendar statistika pokaže, da sta časa izumrtja skoraj enaka. V primeru ko je velika verjetnost 99 % za izumrtje lisic časa malce odstopata, vendar je to lahko posledica premajhne statistike zajcev (500 vzorcev).



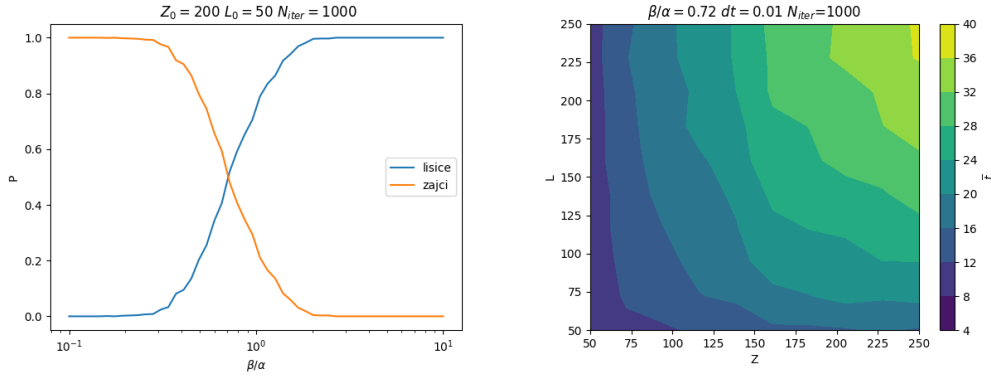
Slika 12: a) Povprečni čas izumrtja populacije v odvisnosti od časovnega koraka. b) Verjetnost za smrt populacije zajcev ali lisic v odvisnosti od časovnega koraka.

Zanimivo je tudi pogledati verjetnost za izumrtje dane vrste v odvisnosti od razmerja β/α (slika 14a). Enako



Slika 13: Statistika smrti lisic in zajcev pri različnih vrednostih α in β . $N_{iter} = 50000$ pove število simulacij. Če bi bila verjetnost 50 % za izumrtje zajcev ali lisic, bi potem vsaka vrsta imela na voljo statistiko 25000 simulacij.

verjetno izumrtje vrst ni pri razmerju 1, ampak je premaknjeno na vrednost okrog $\beta/\alpha = 0.72$. Najverjetneje zaradi nesimetričnih začetnih pogojev. Pogledal sem še povprečni čas izumrtja populacije v odvisnosti od začetnih populacij zajcev in lisic. Ker sem vzel razmerje parametrov tako, da je verjetnost 50 % za izumrtje populacije, bi pričakoval simetričen graf glede na simetralo 1. kvadranta. Mogoče so krive fluktuacije, da to ni popolnoma res.



Slika 14: a) Verjetnost za izumrtje populacije v odvisnosti od razmerja β/α . b) Povprečni čas izumrtja populacije v odvisnosti od začetne velikosti populacije zajcev in lisic. $N_{iter} = 1000$ pove število simulacij za vsako točko.

3 Stohastični model epidemije

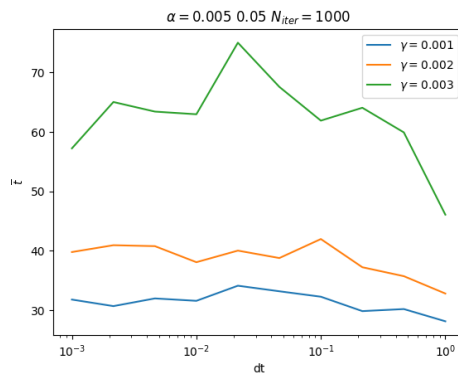
Stohastični model epidemije iz naloge 4. z dodatkom izgube imunosti opišejo enačbe

$$D_{n+1} = D_n - P(\alpha B_n D_n \Delta t) + P(\gamma I_n \Delta t) \quad (14)$$

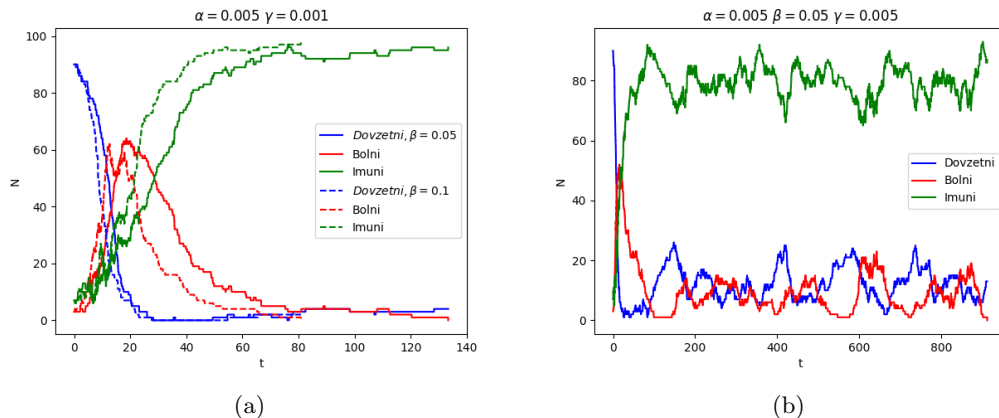
$$B_{n+1} = B_n + P(\alpha B_n D_n \Delta t) - P(\gamma I_n \Delta t) \quad (15)$$

$$I_{n+1} = N - B_n - D_n, \quad (16)$$

kjer je N začetno število populacije. Zadnja enačba izhaja iz ohranitve števila populacije. Prosti parametri α, β, γ povedo kako hitro prehajajo osebk med skupinami. Simulacijo končam, ko je $B=0$. Tako kot prej je najprej dobro pogledati čas eliminacije bolezni v odvisnosti od velikosti časovnega koraka (slika 15). V tem primeru dosežemo plato pri $dt = 0.1$ vendar se vrednosti precej spreminjajo za večje parametre γ . Poleg tega ima velik vpliv tudi začetno število posameznih skupin. Populacija na kateri sem obravnaval dinamiko je $D = 93$, $B = 3$, $I = 7$ osebkov. Potek epidemije prikazuje slika 16. Z večanjem parametra β bolni hitreje ozdravijo in dobimo premaknjen vrh k manjšim časom. Ob primerni nastavitvi hitrosti prehajanja iz skupine imunih v skupino dovzetnih lahko dosežem periodično obnašanje (slika 16b). V stohastični obravnavi lahko dosežem periodičnost brez prilagajanja enačb. V diferencialnem primeru je bilo potrebno

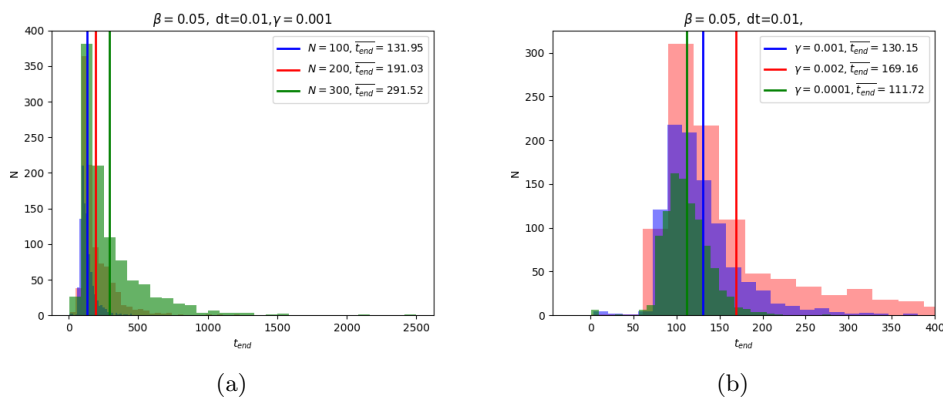


Slika 15: Povprečen čas eliminacije bolezni v odvisnosti od časovnega koraka pri treh različnih parametrih γ .



Slika 16: a) Simulacija epidemije v stohastičnem približku pri dveh različnih vrednostih parametra beta. b) Periodično obnašanje epidemije.

zapisati diferencialne enačbe z zakasnitvijo, ki so predstavljale dodatno stopnjo težavnosti problema. Za konec sem si pogledal še statistiko časov popolne eliminacije v odvisnosti od velikosti populacije in hitrosti izgubljanja imunosti (slika 17). Pričakovano se povprečni čas eliminacije bolezni večja z večanjem parametra γ , saj se hitra izguba imunosti odraža v periodičnosti epidemije. Prav tako se povprečni čas pričakovano večja z večanjem populacije. Kar se mi zdi zelo nenavadno in nisem našel napake v programu je to, da se širina stolpca histograma podvoji za vsak večji parameter γ oz. drugače, enako statistika velja za sosednja/e stolpce z večanjem γ . Za to ne poznam vzroka, mislim tudi, da ni preveč smiselno, vendar napake v programu nisem našel.



Slika 17: a) Statistika časov eliminacije bolezni pri treh velikostih populacije. b) Statistika časov eliminacije bolezni pri treh različnih γ .

4 Zaključek

Stohastična obravnava problemov ima prednosti in slabosti. Prednost je definitivno v preprosti implementaciji v primerjavi z drugimi metodami. Slabost je ta, da je potrebno paziti velikost časovnega koraka, saj lahko za napačen korak dobimo napačne rešitve.