Sveučilište J. J. Strossmayera u Osijeku
Odjel za matematiku

Antonija Živković

Markovljevi lanci u kreditnim rizicima

Diplomski rad

Osijek, 2017.

Sveučilište J. J. Strossmayera u Osijeku
Odjel za matematiku

Antonija Živković

Markovljevi lanci u kreditnim rizicima

Diplomski rad

mentor: izv. prof. dr. sc. Nenad Šuvak

Osijek, 2017.

Sadržaj

Uvod

Kreditne i financijske institucije odobravanjem kredita poduzećima i stanovništvu susreću se s velikim rizikom, ali ga prihvaćaju jer na taj način ostvaruju zaradu. Propadanje banaka najčešće je povezano s lošim kreditnim politikama. Kako bi se smanjio rizik, financijske i kreditne institucije koriste različite alate prilikom donošenja odluka o tome hoće li odobriti kredit ili ne. Tako razvijaju aplikativni skoring koji se koristi za nove klijente, te bihevioralni skoring za već postojeće klijente. Kod modeliranja aplikativnog i bihevioralnog skoringa koriste se različite statističke procedure kao što su logistička regresija, analiza preživljenja i Markovljevi lanci kojim ćemo se baviti u ovome radu.
Markovljevi lanci predstavljaju važnu ulogu u modeliranju kreditnog rejting sustava i kreditnog rizika. Naime, kreditni rizik se može definirati na sljedeće načine[1]:

- Posljedica ugovorene i/ili moguće transakcije između izdavatelja i uzimatelja sredstava, tj. varijacija mogućih povrata koji bi se mogli zaraditi zbog nepotpunog plaćanja glavnice i/ili kamata.
- Vjerojatnost da će izdavatelj vrijednosnog papira otići u default ako ne plati glavnicu i/ili kamate na vrijeme. To je kreditni rizik kod vlasničkih ulaganja.
- Vjerojatnost da dužnik neće otplatiti dug.

U radu ćemo se u primjerima usmjeriti uglavnom na kreditni rizik obveznica, preciznije korporativnih obveznica. Naime, obveznica je dužnički vrijednosni papir koji se izdaje sa ciljem prikupljanja financijskih sredstava s unaprijed definiranim rokom povrata. Korporativne obveznice izdaju poduzeća, bilo da su to javna poduzeća, transportna ili industrijska, banke ili druge financijske kompanije.
Default se prevodi kao status neispunjavanja obveza, a otići u default znači otići u status neispunjavanja obveza. Prema odluci o adekvatnosti kapitala koju je propisala HNB, status neispunjavanja obveza druge ugovorene strane nastaje kada je ispunjen barem jedan od sljedeća dva uvjeta:

1. kreditna institucija smatra da druga ugovorena strana neće u cijelosti otplatiti svoje kreditne obveze,
2. druga ugovorena strana kasni više od 90 dana po bilo kojoj materijalno značajnoj kreditnoj obvezi.

U prvom poglavlju rada biti će opisani Markovljevi lanci u diskretnom vremenu, te procjena prijelaznih vjerojatnosti u diskretnom vremenu. Model u diskretnom vremenu koji treba istaknuti je Mover-stayer model čije ćemo parametre procijeniti. Model se

[1] preuzeto iz [13]

definira pomoću dva Markovljeva lanca, od kojih jedan predstavlja tvrtke koje prelaze iz stanja u stanje, a drugi tvrtke koje konstantno borave u jednom od stanja Markovljevog lanca. Drugo poglavlje govori o Markovljevim lancima u neprekidnom vremenu, njihovoj konstrukciji pomoću skokova, te o generatoskoj matrici. Problem s kojim ćemo se susresti je procjena elemenata generatorske matrice prilikom povezivanja kreditnog rizika s Markovljevim lancima u neprekidnom vremenu. Zatim, kako bismo uvidjeli postoji li razlika između procijenjenih prijelaznih vjerojatnosti u diskretnom i neprekidnom vremenu napravit ćemo primjer na stvarnim podacima. Primjer će biti napravljen na podacima preuzetim od Moody's rejting agencije.
Kreditne rejting agencije procjenjuju kreditnu sposobnost raznih subjekata u gospodarskom sustavu. Agencije za kreditni rejting u praksi najčešće procjenjuju kreditnu sposobnost zajmoprimca, odnosno izdavatelja dužničkih vrijednosnih papira. U tom smislu rejting predstavlja mjeru kreditnog rizika i mogućnosti defaulta zajmoprimca. Agencije za kreditni rejting kao rezultat procjene dobivaju simbole prikazujući na taj način rizičnost izdavatelja vrijednosnih papira, a kreditni rejtinzi koji nastaju kao rezultat procjene rejting agencije razlikuju se s obzirom na rizičnost i kvalitetu na investicijski ili podinvesticijski stupanj. Tri su najveće kreditne rejting agencije koje simbolički nazivamo "Velika trojka", a to su Moody's, Standard & Poor's (S&P) i Fitch rejting. One kontroliraju oko 95% svjetskog tržišta.
Kako je kreditni rejting analogan bihevioralnom skoru, u trećem poglavlju rada napravit ćemo model Markovljevih lanaca na temelju bihevioralnog skora. Kako bismo modelirali bihevioralni skor pomoću Markovljevih lanaca, podijeliti ćemo ga u nekoliko različitih segmenata. Model će biti razvijan na podacima od glavne banke UK, te će biti razmatrane varijable kojima se model poboljšava.

1 Primjena Markovljevih lanaca u diskretnom vremenu u teoriji kreditnih rizika

1.1. Osnove teorije Markovljevih lanaca u diskretnom vremenu

Ovo poglavlje započet ćemo definicijom slučajnog procesa u diskretnom vremenu, te nastaviti definicijom Markovljevog lanca u diskretnom vremenu i osnovnih pojmova vezanih uz taj proces.

Definicija 1.1. *Neka je za svaki $n \in \mathbb{N}_0$, $X_n \colon \Omega \to K \subseteq R$ slučajna varijabla na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$. Familija slučajnih varijabli $(X_n, n \in N_0)$ na istom vjeorojatnosnom prostoru $(\Omega, \mathcal{F}, P)$ je slučajni proces u diskretnom vremenu s vrijednostima u K.*

Definicija 1.2. *Neka je K prebrojiv skup. Slučajni proces $X = (X_n, n \in \mathbb{N}_0)$ na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$ s vrijednostima u skupu K je Markovljev lanac ako vrijedi*

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \tag{1.1}$$

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

Svojstvo u relaciji (1.1) naziva se Markovljevo svojstvo.

Definicija 1.3. *Za $i, j \in K$ i $0 \leq s < t$, $s \in \mathbb{N}_0, t \in \mathbb{N}$ definiramo funkciju prijelaznih vjerojatnosti Markovljevog lanca na sljedeći način*

$$p(i, s; t, j) = P(X_t = j | X_s = i). \tag{1.2}$$

Prema prethodnoj definiciji, funkcija prijelaznih vjerojatnosti u jednom koraku dana je izrazom

$$p(i, n; n+1, j) = P(X_{n+1} = j | X_n = i), \tag{1.3}$$

$i, j \in K, n \in \mathbb{N}_0$.

Ako funkcija prijelaznih vjerojatnosti u jednom koraku ne ovisi o vremenu, odnosno, ako $\forall n, m \in \mathbb{N}$ vrijedi $p(i, n; n+1, j) = p(i, m; m+1, j)$ kažemo da se radi o homogenom Markovljevom lancu. Kako u ovom slučaju funkcija prijelaznih vjerojatnosti u jednom koraku ovisi samo u sadašnjem stanju i i budućem stanju j Markovljevog lanca pišemo $p(i, n; n+1, j) = p_{ij}, i, j \in K$. Ako je K konačan, te vjerojatnosti shematski zapisujemo u obliku matrice

$$\Pi = \begin{bmatrix} p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \cdots & p_{kk} \end{bmatrix}.$$

Matrica $\Pi = [p_{ij}]_{i,j\in K}$ naziva se matrica prijelaznih vjerojatnosti homogenog Markovljevog lanca. Ova matrica ima sljedeća svojstva:

- $0 \leq p_{ij} \leq 1, \quad \forall i,j \in K$
- $\sum_{j\in K} p_{ij} = \sum_{j\in K} P(X_{n+1} = j | X_n = i) = 1, \quad \forall i \in K,$

i takvu matricu nazivamo stohastička matrica.

Napomena 1.1. *Prijelazne vjerojatnosti Markovljevog lanca u više koraka ili n-koračne prijelazne vjerojatnosti su elementi matrice Π^n:*

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i).$$

Teorem 1.1. *Neka je $(X_n, n \in \mathbb{N}_0)$ slučajan proces s diskretnim skupom stanja K i konačnodimenzionalnim distribucijama $P(X_0 = i_0, \dots, X_n = i_n) = \lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}$, $\lambda_{i_0} = P(X_0 = i)$, gdje je $\lambda = (\lambda_i, i \in K)$ neka distribucija na K, a $\Pi = [p_{ij}]_{i,j\in K}$ neka stohastička matrica na K. Tada je taj proces X (λ, Π) Markovljev lanac.*

Teorem 1.2. *Prijelazne vjerojatnosti homogenog (λ, Π)-Markovljevog lanca zadovoljavaju Chapman-Kolmogorovljeve jednakosti:*

$$p_{ij}^{(m)} = \sum_{k\in K} p_{ik}^{(r)} p_{kj}^{(m-r)}, \quad \forall r \in \{1, \dots, m-1\}, i,j \in K. \tag{1.4}$$

Matrični zapis prethodne tvrdnje je

$$\Pi^m = \Pi^r \cdot \Pi^{n-r}.$$

1.2. Markovljevi lanci u diskretnom vremenu i kreditni rizik

Primjerice, razmotrimo N tvrtki čiji su prijelazi između različitih kategorija kreditnog rejtinga promatrani u diskretnim trenutcima $j \in \{0, \dots, J\}$. Pretpostavimo da su prijelazi između stanja nezavisni. Migracije tvrtki kroz vrijeme modeliramo slučajnim procesom (vektorom) $(X_j, j \in \{0, \dots, J\})$, gdje je X_j slučajna varijabla kojom je opisano stanje u kojem se tvrtka nalazi u trenutku j. Ako pretpostavimo da neka tvrtka prijeđe put $X_0, X_1, \dots, X_j$, vjerojatnost tog puta je $p_{X_0,X_1} p_{X_1,X_2} \cdots p_{X_{j-1},X_j}$. Prijelazne vjerojatnosi nam u praksi najčešće nisu poznate, te se suočavamo s problemom procjene prijelaznih vjerojatnosti. Skup stanja je skup rejting kategorija kojeg ćemo označiti s $K = \{1, \dots, D\}$. Uvedimo sljedeće oznake :

- $n_i(j)$ označava broj tvrtki u stanju i u trenutku j,
- $n_{ik}(j)$ označava broj tvrtki koje su prešle iz stanja i u trenutku $(j-1)$ u stanje k u trenutku j,

- $n_i^* = \sum_{j=1}^{J} n_i(j-1)$ označava ukupan broj posjeta stanju i do trenutka J,
- $n_{ik} = \sum_{j=1}^{J} n_{ik}(j)$ označava ukupan broj prijelaza iz stanja i u stanje k do trenutka J.

Ako pretpostavimo da je rejting migracija tvrtki nezavisan, dobivamo umnožak zasebnih vjerojatnosti, te funkcija vjerodostojnosti[2] ima sljedeći oblik:

$$\mathcal{L}(P) = \prod_{(i,k)} p_{ik}^{n_{ik}}$$

pri čemu je $p_{ik}^0 = 1$, a $P = [p_{ik}]_{i,k\in\{0,\ldots,J\}}$, gdje je p_{ik} vjerojatnost prijelaza iz stanja i u stanje k, a n_{ik} je ukupan broj takvih prijelaza.
Funkcija log-vjerodostojnosti[3]

$$\log \mathcal{L}(P) = \sum_{(i,k)} n_{ik} \log (p_{ik}).$$

Kada imamo restrikciju $\sum_{k=1}^{D} p_{ik} = 1$ za svaki i, možemo maksimizirati $\sum_{(i,k)} n_{ik} \log (p_{ik})$ u skladu s $\sum_{k=1}^{D} p_{ik} = 1$. Rješavanjem ovoga standardnim Lagrangeovim multiplikatorom[4] dolazimo do procjene prijelazne vjerojatnosti p_{ik}

$$\hat{p}_{ik} = \frac{n_{ik}}{n_i^*} \tag{1.5}$$

za sve $i, k \in K$.

Primjer 1.1. *Neka n-dimenzionalni slučajni vektor* $\mathbf{X}$ *ima multinomnu distribuciju. Tada je*

$$P(\mathbf{X} = \mathbf{x}; n; \mathbf{p}) = n! \prod_{k=1}^{K} \frac{p_k^{x_k}}{x_k!}$$

gdje je x_k *broj uspjeha k-te kategorije u n slučajnih pokušaja, a* p_k *je vjerojatnost uspjeha k-te kategorije. Primjetimo da je*

$$\sum_{k=1}^{K} x_k = n \quad i \quad \sum_{k=1}^{K} p_k = 1.$$

Pretpostavimo sada da imamo uzorak od N nezavisnih slučajnih vektora, $\mathbf{X_1}, \ldots, \mathbf{X_N}$*, s gornjom multinomnom distribucijom. Funkcija log-vjerodostojnosti je dana s*

$$\mathcal{L}(\mathbf{p}, n) = \sum_{i=1}^{N} \log P(\mathbf{x_i}; n; \mathbf{p})$$

[2]eng. likelihood function
[3]eng. log-likelihood function
[4]vidi [10] poglavlje 6.1.2

gdje je

$$\log P(\mathbf{x_i}; n; \mathbf{p}) = \log\left(\frac{n!}{\prod_k x_{ik}!}\right) + \sum_{k=1}^{K} x_{ik}\log(p_k)$$

$$\sum_{i=1}^{N}\log P(\mathbf{x_i}; n; \mathbf{p}) = C + \sum_{k=1}^{K} N_k\log(p_k)$$

pri čemu $N_k = \sum_{i=1}^{N} x_{ik}$ *označava ukupan broj uspjeha k-te kategorije u N uzoraka, a sa C smo označili vrijednosti koje ne ovise o parametrima. Za MLE od* $\mathbf{p}$*, uz pretpostavku da je n poznat, rješavamo optimizacijski problem:*

$$max_{\mathbf{p}}\mathcal{L}(\mathbf{p}, n) \quad tako \quad da \quad je \quad \sum_{k=1}^{K} p_k = 1.$$

Nadalje, računamo :

$$\frac{\partial\mathcal{L}(\mathbf{p}, n)}{\partial p_k} = \frac{N_k}{p_k} - \frac{N_K}{p_K} = 0.$$

Rješavanjem, uz $\sum_{k=1}^{K} p_k = 1$*, dobivamo MLE za* p_k*,*

$$\hat{p}_k = \frac{N_k}{nN}.$$

Primjetimo sličnost procjenitelja (1.5) s multinomnim procjeniteljem. Jasno je da dodavanjem opservacija tijekom vremena, procjenitelj implicitno koristi pretpostavku vremenske homogenosti.

1.3. Mover-stayer model

Mover-stayer model je slučajni proces $(Z_j, j \in \mathbb{N}_0)$ sa skupom stanja $K = \{1, \ldots, w\}$. Taj proces se definira korištenjem dva nezavisna Markovljeva lanca, jedan čija je matrica prijelaznih vjerojatnosti jednaka jediničnoj matrici I i drugi s matricom prijelaznih vjerojatnosti $M = [m_{ik}]_{i,k\in K}$. Neka su prijelazne vjerojatnosti procesa $(Z_j, j \in \mathbb{N}_0)$, gdje sa Z_j opisujemo stanje tvrtke u trenutku j, dane na sljedeći način:

$$P(Z_j = k|Z_0 = i) = \begin{cases} (1 - s_i)m_{ik}^j, & i \neq k \\ s_i + (1 - s_i)m_{ii}^j, & i = k \end{cases}, \tag{1.6}$$

pri čemu je $i, k \in K$, $M^j = [m_{ik}^j]_{i,k\in K}$, te je $0 \leq s_i \leq 1$ vjerojatnost onih tvrtki koje ostaju u stanju i. Ako sa S označimo dijagonalnu matricu čiji su elementi s_i, onda j-koračna prijelazna matrica Π^j s prijelaznim vjerojatnostima definiranim u 1.6 ima oblik:

$$\Pi^j = S + (I - S)M^j. \tag{1.7}$$

Nadalje, promotrimo problem procjene parametara Mover-stayer modela. Trebamo procijeniti matricu prijelaznih vjerojatnosti M i matricu S. Za procjenu provodimo proceduru MLE. Kako bismo zapisali funkciju vjerodostojnosti uvedimo dodatne oznake, uz one definirane u prethodnoj točki:

- $n_{i_0,\dots,i_j}$ označava broj tvrtki s nizom stanja $(i_0,\dots,i_j) \in K^{J+1}$, te uočimo $\sum_{i=1}^{w} n_i(j) = \sum_{(i_0,\dots,i_j)\in K^{J+1}} n_{i_0,\dots,i_j}$.
- n_i je broj tvrtki s nizom stanja $(i_0,\dots,i_j) \in K^{J+1}$ takvih da je $i_0 = i_1 = \cdots = i_l = i, i \in K$.

Ako pretpostavimo da su povijesna stanja tvrtki nezavisna, onda slučajne varijable $(n_{i_0,\dots,i_j}, (i_0,\dots,i_j) \in K^{J+1})$ imaju multinomnu distribuciju s multinomnim vjerojatnostima $(\pi_{i_0,\dots,i_j}, (i_0,\dots,i_j) \in K^{J+1})$ koje daju zajedničku distribuciju Mover-stayer modela.

$$\begin{aligned}\pi_{i_0,\dots,i_j} &= P(Z_0 = i_0, Z_1 = i_1, \dots, Z_J = i_J) \\ &= \begin{cases} s_{i_0}\eta_{i_0} + (1-s_{i_0})\eta_{i_0} m_{i_0 i_0}^J, & i_0 = i_1 = \cdots = i_J \\ (1-s_{i_0})\eta_{i_0} m_{i_0 i_1}\cdots m_{i_{J-1} i_J}, & \text{inače} \end{cases},\end{aligned}$$

gdje $(\eta_i, i \in K)$ označava početnu distribuciju procesa $(Z_j, j \in \mathbb{N}_0)$. Prema tome, funkcija vjerodostojnosti $\mathcal{L} = \mathcal{L}(\eta, M, S)$ ima sljedeći oblik:

$$\begin{aligned}\mathcal{L}(\eta, M, S) = [\pi_{i_0,\dots,i_j}]^{n_{i_0,\dots,i_j}} &= \prod_{i=1}^{w} [s_i\eta_i + (1-s_i)\eta_i m_{ii}^J]^{n_i} \\ &\cdot \prod_{(i_0,\dots,i_j)\in K^{J+1}, i_k \neq i_m} [(1-s_{i_0})\eta_{i_0} m_{i_0 i_1}\cdots m_{i_{J-1} i_J}]^{n_{i_0,\dots,i_j}}.\end{aligned}$$

Grupiranjem po i dobivamo

$$\begin{aligned}\mathcal{L}(\eta, M, S) &= \prod_{i=1}^{w} \eta_i^{n_i(0)} \prod_{i=1}^{w} [s_i + (1-s_i)m_{ii}^J]^{n_i} \\ &\cdot \prod_{i=1}^{w} (1-s_i)^{n_i(0)-n_i} \prod_{i=1}^{w} m_{ii}^{n_{ii}-Jn_i} \prod_{i\neq k}^{w} m_{ik}^{n_{ik}} \\ &= \prod_{i=1}^{w} \eta_i^{n_i(0)} \prod_{i=1}^{w} L_i,\end{aligned}$$

gdje je $L_i = [s_i + (1-s_i)m_{ii}^J]^{n_i}(1-s_i)^{n_i(0)-n_i} m_{ii}^{n_{ii}-Jn_i} \cdot \prod_{k=1, i\neq k}^{w} m_{ik}^{n_{ik}}, \quad i \in K$.

$$\begin{aligned}\log L_i &= n_i \log[s_i + (1-s_i)m_{ii}^J] + (n_i(0) - n_i)\log(1-s_i) \\ &+ (n_{ii} - Jn_i)\log(m_{ii}) + \sum_{k=1, i\neq k}^{w} n_{ik}\log(m_{ik}).\end{aligned} \tag{1.8}$$

Kako bismo našli kritične točke od L_i, s_i i m_{ij}, prvo stavimo

$$\frac{\partial \log L_i}{\partial s_i} = \frac{n_i(1-m_{ii}^J)}{s_i + (1-s_i)m_{ii}^J} - \frac{n_i(0) - n_i}{(1-s_i)} = 0.$$

Rješavajući po s_i dobivamo

$$s_i = \frac{n_i - n_i(0)m_{ii}^J}{n_i(0)(1-m_{ii}^J)}.$$

Ako uvrstimo ovaj izraz za s_i u (1.8) i sa C označimo vrijednosti koje ne ovise o parametrima, imamo sljedeće:

$$\log L_i = C - (n_i(0) - n_i)\log(1 - m_{ii}^J) + (n_{ii} - Jn_i)\log(m_{ii}) + \sum_{k=1, i\neq k}^{w} n_{ik}\log(m_{ik}).$$

Kako bismo olakšali notaciju, stavimo

$$g_i(m_{ii}) = (n_i(0) - n_i)\log(1 - m_{ii}^J) + (n_{ii} - Jn_i)\log(m_{ii}).$$

Sada $\log L_i$, za $i \neq w$, možemo zapisati kao

$$\log L_i = C - g_i(m_{ii}) + \sum_{k=1, i\neq k}^{w-1} n_{ik}\log(m_{ik}) + n_{ik}\log\left(1 - \sum_{k=1}^{w-1} m_{ik}\right). \tag{1.9}$$

Ako deriviramo (1.9) po $m_{i,w-1}$ imamo

$$\frac{\partial \log L_i}{\partial m_{i,w-1}} = \frac{n_{i,w-1}}{m_{i,w-1}} - \frac{n_{iw}}{1 - \sum_{k=1}^{w-1} m_{ik}} = 0$$

i rješavanjem po $m_{i,w-1}$ dobivamo

$$m_{i,w-1} = \frac{n_{i,w-1}\left(1 - \sum_{k=1}^{w-2} m_{ik}\right)}{\sum_{k=w-1}^{w} n_{ik}}.$$

Ako sada izraz za $m_{i,w-1}$ uvrstimo u (1.9) imamo

$$\log L_i = C - g_i(m_{ii}) + \sum_{k=1, i\neq k}^{w-2} n_{ik}\log(m_{ik}) + \sum_{k=w-1}^{w} n_{ik}\log\left(1 - \sum_{k=1}^{w-2} m_{ik}\right).$$

Nadalje, ponavljamo isti postupak za parametre $m_{i,w-2}, m_{i,w-3}, \ldots, m_{i,i+1}, m_{i,i-1}, \ldots, m_{i1}$ redom i dobivamo

$$\frac{\partial \log L_i}{\partial m_{ij}} = \frac{n_{ij}}{m_{ij}} - \frac{\sum_{k=j+1, k\neq i}^{w} n_{ik}}{1 - \sum_{k=1}^{j} m_{ik}} = 0, \quad j = w-1, w-2, \ldots, i+1, i-1, \ldots, 1.$$

Zatim riješimo po m_{ij}:

$$m_{ij} = \frac{n_{ij}\left(1 - m_{ii} - \sum_{k=1, k\neq i}^{j-1} m_{ik}\right)}{\sum_{k=j, k\neq i}^{w} n_{ik}}$$

i to uvrstimo u (1.9) pa imamo

$$\log L_i = C - g_i(m_{ii}) + \sum_{k=1, k\neq i}^{w} n_{ik}\log(m_{ik}) + \left(\sum_{k=j}^{w} n_{ik}\right)\log\left(1 - m_{ii} - \sum_{k=1, k\neq i}^{j-1} m_{ik}\right).$$

Nakon posljednjeg deriviranja (za $j = 1$ ili $j = 2$ ako je $i = 1$), dobivamo $\log L_i$ kao funkciju od m_{ii}:

$$\log L_i = C - g_i(m_{ii}) + \sum_{k=1,k\neq i}^{w} n_{ik} \log(1 - m_{ii}), \quad i = 1, 2, 3, \ldots, w-1. \tag{1.10}$$

Konačno, stavimo $\frac{\partial \log L_i}{\partial m_{ii}}, i \in K$ i primjetimo $\sum_{k=1,k\neq i}^{w} n_{ik} = n_i^* - n_{ii}$, te postavimo jednadžbu za $m_{ii}, i \in K$.

$$[n_i^* - Jn_i(0)]m_{ii}^{J+1} + [Jn_i(0) - n_{ii}]m_{ii}^{J} + [Jn_i - n_i^*]m_{ii} + n_{ii} - Jn_i = 0. \tag{1.11}$$

Sada iz ovoga možemo rekurzivno računati odgovarajuće $\hat{m}_{ij}$, za $j \neq i$:

$$\hat{m}_{ij} = \frac{n_{ij}(1 - \hat{m}_{ii} - \sum_{k=1,k\neq i}^{k-1} \hat{m}_{i,k})}{\sum_{k=1,k\neq i}^{k} n_{ik}} \quad i, j \in K, \tag{1.12}$$

počevši od najmanje j vrijednosti, odnosno, od $j = 1$ ili od $j = 2$ ako je $i = 1$. Odgovarajući $\hat{s}_i$ je tada:

$$\hat{s}_i = 1 - \frac{n_i(0) - n_i}{n_i(0)(1 - \hat{m}_{ii}^{J})} \quad i \in K. \tag{1.13}$$

Vrijednost $(n_i(0) - n_i)$ predstavlja broj tvrtki koje su startale u stanju i i napravile barem jedan prijelaz do trenutka J. Vrijednost $(n_i(0)(1-\hat{m}_{ii}^{J}))$ je očekivani broj takvih tvrtki, pod pretpostavkom da su sve $n_i(0)$ tvrtke koje se kreću. Razlika između 1 i kvocijenta broja tvrtki i očekivanog broja predstavlja procjenu udjela tvrtki koje ostaju, odnosno s_i.

Ako stavimo da $J \to \infty$, jednadžba (1.11) izgleda ovako:

$$[Jn_i - n_i^*]m_{ii} = Jn_i - n_{ii}, \quad i \in K, \tag{1.14}$$

a rješenje je

$$\bar{m}_{ii} = \frac{n_{ii} - Jn_i}{n_i^* - Jn_i}, \quad i \in K. \tag{1.15}$$

Vrijednost $(n_{ii} - Jn_i)$ predstavlja ukupan broj tvrtki koje su napravile barem J prijelaza tijekom vremena J i ostale u stanju i, a $(n_i^* - Jn_i)$ je broj tvrtki koje su posjetile stanje i i napravile barem J prijelaza do trenutka J.

Kada stavimo da $J \to \infty$ u (1.13) procjenitelj od s_i ima oblik:

$$\bar{s}_i = \frac{n_i}{n_i(0)}, \quad i \in K.. \tag{1.16}$$

Sada je procjenitelj $\bar{s}_i$ jednostavno udio onih tvrtki koje konstantno ostaju u svom početnom stanju i. Ovaj procjenitelj je intuitivno privlačan kada je J velik jer je tada razumno pretpostaviti da su one tvrtke koje konstantno stoje u svom početnom stanju tvrtke s_i.

1.3.1. Usporedba procjenitelja

Svrha usporedbe je prikazati razlike između procjenitelja, $\hat{m}_{ii}$ i $\bar{m}_{ii}$, te $\hat{s}_i$ i $\bar{s}_i$, prikazanih u prethodnoj točki. Ovaj primjer usporedbe preuzet je iz H. Frydman, *Maximum-likelihood estimation in the Mover-stayer model.* Naime, usporedba je napravljena na temelju podataka sastavljenih od 200 simuliranih realizacija i 3 stanja Mover-stayer modela s parametrima

$$M = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.2 & 0.7 & 0.1 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}, S = (0.7, 0.5, 0.1) \quad i \quad \eta = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

Primjerice, element 0.15 u matrici M predstavlja vjerojatnost da tvrtka koja se kreće prijeđe iz stanja 2 u stanje 1, dok element 0.5 vektora S predstavlja vjerojatnost da tvrtka koja stoji konstantno stoji u stanju 2 i dalje ostane u njemu. Elementi vektora η predstavljaju vjerojatnosti da se tvrtka nađe u nekom od 3 moguća stanja u trenutku 0. Svaka realizacija proteže se preko vremenskog intervala sastavljenog od 12 perioda. U tablicama 1.1 i 1.2 prikazani su rezultati procjene M i S dvjema metodama za 4 intervala različite duljine, a na temelju 200 simuliranih podataka. Primjetimo kako su oba procjenitelja od $m_{ij}, i \neq j$, izvedena iz rekurzivne formule koja kreće s procjenom m_{ii}. Stoga su razlike između procjenitelja od $m_{ij}, i \neq j$, posljedica razlika između procjenitelja od m_{ii}. Nadalje, kako bismo olakšali usporedbu, u tablici 1.1 je prikazana procjena od m_{ii} i s_i. Rezultati u Tablici 1.1 prikazuju kako je vrijednost $\hat{m}_{ii}$ i $\bar{m}_{ii}$,

Tablica 1.1: Usporedba procjenitelja od m_{ii} i s_i

		Vremenski interval			
Stvarna vrijednost	Procjenitelj	(0,3)	(0,6)	(0,9)	(0,12)
$m_{11} = 0.8$	$\hat{m}_{11}$	0.76	0.81	0.81	0.79
	$\bar{m}_{11}$	0.64	0.77	0.79	0.78
$s_1 = 0.7$	$\hat{s}_1$	0.63	0.60	0.58	0.62
	$\bar{s}_1$	0.80	0.71	0.64	0.64
$m_{22} = 0.7$	$\hat{m}_{22}$	0.73	0.72	0.71	0.71
	$\bar{m}_{22}$	0.64	0.69	0.70	0.71
$s_2 = 0.5$	$\hat{s}_2$	0.57	0.59	0.61	0.61
	$\bar{s}_2$	0.73	0.65	0.63	0.62
$m_{33} = 0.5$	$\hat{m}_{33}$	0.52	0.42	0.40	0.42
	$\bar{m}_{33}$	0.38	0.41	0.40	0.42
$s_3 = 0.1$	$\hat{s}_3$	0.06	0.17	0.18	0.18
	$\bar{s}_3$	0.19	0.18	0.18	0.18

te $\hat{s}_i$ i $\bar{s}_i$ u kraćim vremenskim intervalima jako različita. Što je inteval duži to su vrijednosti procjenitelja bliže, te skoro identične u najdužem vremenskom intervalu

$(0, 12)$. Možemo naslutiti, kada je J dovoljno velik, u ovom slučaju kada je jednak barem 9, koristimo procjenitelj $\bar{m}_{ii}$, a kada je manji koristimo $\hat{m}_{ii}$. Procjena cijele matrice M prikazana je u tablici 1.2. Iz nje izvodimo isti zaključak kao iz tablice 1.1 samo se ovdje radi o procjeni $m_{ij}, i \neq j$.

Tablica 1.2: Usporedba procjenitelja od M

Vremenski interval	$\hat{M}$			$\bar{M}$		
(0,3)	0.76	0.2	0.04	0.643	0.303	0.054
	0.18	0.73	0.09	0.24	0.64	0.12
	0.1	0.38	0.52	0.13	0.49	0.38
(0,6)	0.81	0.16	0.03	0.77	0.19	0.04
	0.2	0.72	0.08	0.21	0.69	0.09
	0.12	0.46	0.42	0.123	0.464	0.413
(0,9)	0.81	0.15	0.04	0.79	0.16	0.05
	0.21	0.71	0.08	0.21	0.7	0.09
	0.11	0.49	0.4	0.11	0.49	0.4
(0,12)	0.79	0.16	0.05	0.78	0.17	0.05
	0.2	0.71	0.09	0.2	0.71	0.09
	0.11	0.47	0.42	0.11	0.47	0.42

2 Primjena Markovljevih lanaca u neprekidnom vremenu u teoriji kreditnih rizika

2.1. Osnove teorije Markovljevih lanaca u neprekidnom vremenu

2.1.1. Definicija i osnovna svojstva

U ovome poglavlju ćemo proučavati Markovljeve u neprekidnom vremenu, tj. slučajne procese indeksirane vremenom $t \in [0, \infty)$. Započet ćemo formalnom definicijom Markovljevog lanca.

Definicija 2.1. *Markovljev lanac u neprekidnim vremenom s prebrojivim skupom stanja S je slučajni proces $X = (X_t, t \geq 0)$ definiran na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$ za koji vrijedi Markovljevo svojstvo, tj. $\forall n \geq 0$, za bilo koji izbor trenutaka $0 \leq t_1 \leq t_2 \leq \cdots \leq t_{n-1} \leq t_n$ i proizvoljna stanja $i_1, \ldots, i_{n-2}, i, j \in S$ vrijedi*

$$P(X_{t_n} = j | X_{t_1} = i_1, \ldots, X_{t_{n-2}} = i_{n-2}, X_{t_{n-1}} = i) = P(X_{t_n} = j | X_{t_{n-1}} = i). \tag{2.1}$$

Neka je za $i, j \in S$, te $0 \leq s < t$

$$P_{ij}(s, t) := P(X_t = j | X_s = i).$$

Funkcija $P_{ij}(s, t)$ zove se prijelazna vjerojatnost Markovljevog lanca X. Tada (2.1) zapisujemo kao

$$P(X_{t_n} = j | X_{t_1} = i_1, \ldots, X_{t_{n-2}} = i_{n-2}, X_{t_{n-1}} = i) = P_{ij}(t_{n-1}, t_n). \tag{2.2}$$

Definicija 2.2. *Markovljev lanac $X = (X_t, t \geq 0)$ je vremenski homogen ako prijelazne vjerojatnosti ovise samo o razlici vremenskih trenutaka, odnosno ako vrijedi*

$$P_{ij}(s, t) = P_{ij}(0, t - s), \quad \forall i, j \in S \quad i \quad \forall s, t \quad t.d. \quad 0 \leq s < t.$$

Uvedimo oznaku $P_{ij}(t - s)$ za $P_{ij}(s, t)$. i promotrimo uvjetne vjerojatnosti $P_i(A) := P(A | X_0 = i)$. Tada za $i, j \in S$ i $t \geq 0$ imamo

$$P_i(X_t = j) = P(X_t = j | X_0 = i) = P_{ij}(t). \tag{2.3}$$

Neka je za $i \in S$, $\lambda_i = P(X_0 = i)$. Tada je $\lambda = (\lambda_i, i \in S)$ početna distribucija Markovljevog lanca X.
Označimo

$$P(t) = (P_{ij}(t), i, j \in S), \quad t \geq 0. \tag{2.4}$$

Ako imamo konačan skup stanja onda se $P(t)$ shematski prikazuje kvadratnom matricom.

Propozicija 2.1. *Familija $(P(t), t \geq 0)$ je stohastička polugrupa, odnosno vrijedi sljedeće :*

1. $P(0) = I$, *gdje je* $I_{ij} = \delta_{ij}$,
2. $P(t)$ *je stohastička matrica,*
3. $P(t+s) = P(t)P(s)$ *(Chapman-Kolmogorovljeve jednakosti).*

Dokaz:

1. Vrijedi $P_{ij}(0) = P_i(X_0) = j = \delta_{ij}$.
2. Očito je $P_{ij} \leq 0$. Za $i \in S$,
$$\sum_{j \in S} P_{ij}(t) = \sum_{j \in S} P_i(X_t = j) = P_i(X_t \in S)$$
3. $P_{ij}(t+s) = P_i(X_{t+s} = j) = \sum_{k \in S} P_{ik}(t)P_{kj}(s) = (P(t)P(s))_{ij}$ što slijedi zbog (2.4).

■

Nadalje, pretpostavimo da Markovljev lanac X ima zdesna neprekidne trajektorije. Tada za $J_1 = inf\{t \geq 0 : X_t \neq X_0\}$ vrijedi $J_1 > 0$, pri čemu J_1 označava vrijeme skoka.

Definicija 2.3. *Vremena skokova Markovljevog lanca X definiramo na sljedeći način:*

$$J_0 = 0, \quad J_{n+1} = inf\{t \geq J_n : X_t \neq X_{J_n}\}$$

za $n = 0, 1, 2, 3, \dots$, gdje je $inf \emptyset = +\infty$.

Nadalje, ako je $J_n < \infty$, zbog neprekidnosti zdesna vrijedi $J_n < J_{n+1}$.

Definicija 2.4. *Vremena čekanja $(W_n, n \in \mathbb{N})$ od X definiramo na sljedeći način:*

$$W_n = \begin{cases} J_n - J_{n-1}, & J_{n-1} < \infty, \\ +\infty, & J_{n-1} = \infty. \end{cases}$$

Primjer 2.1. *Neka je $X = (X_t, t \geq 0)$ Markovljev lanac sa skupom stanja S takav da je $q : S \to (0, \infty)$, vrijednost joj označimo s q. Pripadajući lanac skokova je $Y = (Y_n : n \in \mathbb{N}_0)$ s prijelaznom matricom $\Pi = [\pi_{ij}]_{i,j \in S}$. Vremena skokova lanca X dana su sa*

$$J_n = \frac{1}{q} \sum_{j=1}^{n} E_j,$$

pri čemu je $(E_n, n \geq 1)$ niz nezavisnih eksponencijalnih slučajnih varijabli s parametrom 1. Tada vrijedi

$$X_t = Y_n, \quad za \quad J_n \leq t < J_{n+1}. \tag{2.5}$$

Neka je $N = (N_t, t \geq 0)$ *brojeći proces, gdje je*

$$N_t = \sum_{n=1}^{\infty} J_n 1_{[0,t]}, t \geq 0.$$

Uočimo da N_t *broji koliko se dogodilo skokova* J_n *do trenutka* $t \geq 0$, *odnosno broji koliko se nezavisnih eksponencijalnih vremena dogodilo do trenutka* $t \geq 0$. *Stoga je to Poissonov proces.*
Tada je $N_t = n$ *ako i samo ako vrijedi* $J_n \leq t < J_{n+1}$, *te uz (2.5) zaključujemo da je*

$$X_t = Y_{N_t}, t \geq 0.$$

Uočimo da su lanac Y *i Poissonov proces* N *nezavisni. Izračunajmo sada funkciju prijelaznih vjerojatnosti lanca* X.

$$\begin{aligned} P_{ij}(t) &= P_i(X_t = j) = P_i(Y_{N_t} = j) \\ &= \sum_{n=0}^{\infty} P_i(Y_n = j, N_t = n) \\ &= \sum_{n=0}^{\infty} P_i(Y_n = j) P_i(N_t = n) \\ &= \sum_{n=0}^{\infty} \pi_{ij}^n e^{-qt} \frac{(qt)^n}{n!} \end{aligned}$$

2.1.2. Konstrukcija Markovljevog lanca pomoću skokova

Pretpostavimo da je $Z =(Z_n, n \in \mathbb{N}_0)$ Markovljev lanac u diskretnom vremenu s prebrojivim skupom stanja S, s početnom distribucijom $\lambda = (\lambda_i, i \in S)$ i matricom prijelaznih vjerojatnosti $\Pi = [\pi_{ij}]_{i,j \in S}$ za koju vrijedi da je $\pi_{ii} = 0, \forall i \in S$. Neka je $(E_n, n \in \mathbb{N})$ niz nezavisnih eksponencijalnih slučajnih varijabli s parametrom 1 i neka su E_n nezavisne od Markovljevog lanca Z. Nadalje, neka je dana funkcija $q : S \to (0, \infty)$. Želimo konstruirati Markovljev lanac $X = (X_t, t \geq 0)$ u neprekidnom vremenu pomoću Markovljevog lanca Z i niza $(E_n, n \in \mathbb{N})$.
Stavimo $J_0 = 0$ te $W_1 = \frac{E_1}{q(Z_0)}$. Tada $\forall s \geq 0$ vrijedi

$$P(W_1 > s | Z_0 = i) = \frac{P(\frac{E_1}{q(Z_0)} > s, Z_0 = i)}{P(Z_0 = i)} = \frac{P(\frac{E_1}{q(i)} > s, Z_0 = i)}{P(Z_0 = i)}$$

$$= \frac{P(\frac{E_1}{q(i)} > s) P(Z_0) = i)}{P(Z_0 = i)} = e^{-q(i)s}.$$

Iz prethodne jednakosti vidimo da W_1, uvjetno na Z_0, ima eksponencijalnu distribuciju s parametrom $q(i)$. Neka je $J_1 = J_0 + W_1$ i $X_t = Z_0$, za $J_0 \leq t < J_1$, te neka je $W_2 = \frac{E_2}{q(Z_1)}$. Tada imamo

$$P(W_1 > s | Z_1 = i) = e^{-q(i)s}.$$

Dalje nastavljamo induktivno. Zatim pretpostavimo da smo konstruirali slučajne varijable $(J_m, 0 \leq m \leq n)$, $(W_m, 1 \leq m \leq n)$ te $(X_t, 0 \leq t < J_n)$. Neka je

$$W_{n+1} = \frac{E_{n+1}}{q(Z_n)}, \quad J_{n+1} = J_n + W_{n+1} \quad i \quad X_t = Z_n, \quad J_n \leq t < J_{n+1}.$$

Vidimo da vrijedi $J_n = \sum_{m=1}^{n} W_m, n \leq 1$ i

$$J_n - J_{n-1} = W_n = \frac{E_n}{q(Z_{n-1})}, \quad n \geq 1.$$

Označimo s $\nu := lim_{n \to \infty} J_n = \sum_{m=1}^{n} W_m \leq \infty$ vrijeme eksplozije slučajnog procesa $X = (X_t, t \geq 0)$. Slučajna varijabla X_t je definirana na intervalu $[0, \nu)$ i vrijedi

$$X_t = \sum_{n=0}^{\infty} Z_n \mathbf{1}_{[J_n, J_{n+1})}(t). \tag{2.6}$$

Nadalje, uvodimo dodatno stanje δ koje se često naziva groblje i stavimo $X_t = \delta$ za $t \leq \nu$.

Definicija 2.5. *Kažemo da je proces u neprekidnom vremenu $X = (X_t, 0 \leq t < \nu)$ regularan ako je $P_i(\nu = +\infty) = 1, \forall i \in S$. U tom je slučaju X_t definiran za sve $t \geq 0$.*

Propozicija 2.2. *Slučajne varijable $(W_n, n \in \mathbb{N})$ su uvjetno nezavisne i eksponencijalno distribuirane uz danu $\sigma - algebru$ $\mathcal{G} := \sigma(Z_n, n \in \mathbb{N}_0)$.*

Pogledajmo dovoljne uvjete uz koje je X regularan.

Propozicija 2.3. *Za svaki $i \in S$ vrijedi*

$$P_i(\nu < \infty) = P_i\left(\sum_{n=0}^{\infty} \frac{1}{q(Z_n)} < \infty\right). \tag{2.7}$$

Slučajni proces X je regularan ako i samo ako vrijedi

$$\sum_{n=0}^{\infty} \frac{1}{q(Z_n)} = +\infty \quad P_i - g.s., \quad \forall i \in S. \tag{2.8}$$

Dokaz:
Vrijedi $\nu = \sum_{n=1}^{\infty} W_n = \sum_{n=1}^{\infty} \frac{E_n}{q(Z_{n-1})}$. Uvjetno na $\mathcal{G} = \sigma(Z_n, n \geq 1)$, ν je zbroj nezavisnih eksponencijalnih slučajnih varijabli s parametrima $q(Z_{n-1})$. Nadalje, iskoristimo propoziciju:

Propozicija 2.4. *Neka je $(T_n, n \in \mathbb{N})$ niz nezavisnih eksponencijalnih slučajnih varijabli takvih da T_n ima parametar q_n. Tada je $\sum_{n=1}^{\infty} T_n < \infty$ g.s. ako i samo je $\sum_{n=1}^{\infty} \frac{1}{q_n} < \infty$.*

Prema navedenoj propoziciji slijedi :

$$P(\nu < \infty|\mathcal{G}) = \begin{cases} 1, & \sum_{n=1}^{\infty} \frac{1}{q(Z_{n-1})} < \infty, \\ 0, & \sum_{n=1}^{\infty} \frac{1}{q(Z_{n-1})} = \infty, \end{cases}$$

tj. $P(\nu < \infty|\mathcal{G}) = \mathbf{1}_{(\sum_{n=1}^{\infty} \frac{1}{q(Z_{n-1})} < \infty)}$ g.s. Računanjem uvjetnog očekivanja slijedi (2.6), dok je (2.7) neposredna posljedica. Ako je $c := \sup\limits_{i \in S} q(i) < \infty$ tada je

$$\sum_{n=1}^{\infty} \frac{1}{q(Z_{n-1})} \geq \sum_{n=1}^{\infty} \frac{1}{c} = +\infty$$

pa je X regularan. ∎

Napomena 2.1. *Specijalno, ako je* $\sup\limits_{i \in S} q(i) < \infty$*, onda je* X *regularan.* $\sup\limits_{i \in S} < \infty$ *je uvijek ispunjeno ako je* S *konačan.*

Ako je $q(i) \in (0, \infty)$ za svaki $i \in S$ onda se proces X naziva stabilnim. Ako bismo dopustili da je $q(i) = 0$ za neko stanje $i \in S$ vrijeme čekanja bi bilo beskonačno, odnosno uz $Z_{n-1} = i$ bismo imali $W_n = \frac{E_n}{q(Zn-1)} = +\infty$. Drugim riječima, stanje $i \in S$ bi bilo apsorbirajuće stanje procesa X. Kada bismo stavili $q(i) = +\infty$ za neki $i \in S$, uz dani $Z_{n-1} = i$ bismo imali $W_n = 0$. Takvo stanje i nazivamo trenutnim stanjem.

Definicija 2.6. *Vremena skokova procesa* X *definiramo na sljedeći način*

$$\widetilde{J_0} = 0,$$

$$\widetilde{J_{n+1}} := inf\{t \geq 0 : X_t \neq X_{\widetilde{J_n}}\}.$$

Očito je $\widetilde{J_n} = J_n$ i $X_{J_n} = Z_n$, odnosno, Markovljev lanac Z u diskretnom vremenu je proces skokova slučajnog procesa X u neprekidnom vremenu.

Teorem 2.1. *Neka je* X *regularan. Tada je* X *Markovljev lanac u neprekidnom vremenu.*

Dokaz:
Definirajmo prvo prijelaznu funkciju $P_{ij}(t)$ od X formulom $P_{ij}(t) = P_i(X_t = j), t \geq 0, i, j \in S$. Kako bismo dokazali da je X Markovljev, dovoljno je pokazati da konačno-dimenzionalne distribucije od X zadovoljavaju

$$P_i(X_{t_1} = j_1, \ldots, X_{t_n} = j_n) = P_{ij_1}(t_1)P_{j_1j_2}(t_2 - t_1) \cdots P_{j_{n-1}j_n}(t_n - t_{n-1})$$

$\forall n \geq 1$, sve $j_1, \ldots, j_n \in S$ i sve $0 < t_1 < \cdots < t_n$. U ovome dokazu zbog jednostavnosti ćemo se ograničiti na $n = 2$, te dokazati jednakost

$$P_i(X_{t_1} = j_1, X_{t_2} = j_2) = P_{ij_1}(t_1)P_{j_1j_2}(t_2 - t_1). \tag{2.9}$$

Za daljnju provedbu dokaza pogledati [16], Poglavlje 2, stranica 61, Teorem 2.20. ∎

2.1.3. Generatorska matrica

Na prijelaznu matricu Π i funkciju q koje su definirane u prethodnoj točki možemo gledati kao na lokalne karakteristike Markovljevog lanca: ako se X nalazi u stanju $i \in S$, tada u stanje $j \in S$, $j \neq i$, prelazi s vjerojatnošću π_{ij} nakon vremena čekanja koje je eksponencijalno s parametrom $q(i)$. Pitanje je, dakle, kako ti lokalni parametri određuju globalne karakteristike Markovljevog lanca, tj. vjerojatnosti $P_{ij}(t) = P_i(X_t = j)$. Radi jednostavnosti pretpostavit ćemo da je X regularan, te da je skup S konačan.

Definicija 2.7. *Generatorska matrica Markovljevog lanca $X = (X_t, t \geq 0)$ je matrica $Q = [q_{ij}]_{i,j \in S}$ definirana s*

$$q_{ij} = \begin{cases} -q(i), & j = i \\ q(i)\pi_{ij}, & j \neq i \end{cases} \tag{2.10}$$

Možemo primjetiti kako je $q_{ii} < 0 \quad i \quad q_{ij} \geq 0$ za $j \neq i$, te $\forall i \in S$ vrijedi

$$\sum_{j \in S} q_{ij} = -q(i) + \sum_{j \neq i} q(i)\pi_{ij} = 0.$$

Generatorska matrica Q je u potpunosti određena prijelaznom matricom Π i funkcijom q. Ako pretpostavimo obratno, da je dana matrica $Q = [q_{ij}]_{i,j \in S}$ takva da je $q_{ii} < 0 \quad i \quad \sum_{j \in S} q_{ij} = 0, \quad \forall i \in S$, te definiramo funkciju $q : S \to (0, \infty)$ i matricu $\Pi = [\pi_{ij}]_{i,j \in S}$ sa

$$q(i) = -q_{ii}, \quad \pi_{ii} = 0, \quad \pi_{ij} = -\frac{q_{ij}}{q_{ii}}.$$

Tada je očito da je $\pi_{ij} \geq 0 \quad i \quad \sum_{j \in S} \pi_{ij} = 1, \quad \forall i \in S$.

2.2. Markovljevi lanci u neprekidnom vremenu i kreditni rizik

Većina rejting sustava daje datum od kojeg se odvija rejting prijelaza kroz kategorije, a moguće je dobiti i druge podatke od rejting agencija. Korištenjem tih podataka značajno poboljšavamo procjenu prijelaznih vjerojatnosti tvrtki između različitih rejting kategorija.

Pretpostavimo da je promatran Markovljev lanac $X = (X_t, t \in [0, T])$ čiji skup stanja čine rejting kategorije. Simbolički, označimo takav skup stanja s $\{1, \ldots, D\}$, gdje je D stanje koje označava rejting kategoriju default. Kako bismo procijenili elemente generatorske matrice Q, pretpostavimo da nam je poznata početna distribucija tog Markovljevog lanca. Funkcija vjerodostojnosti ima sljedeći oblik:

$$\mathcal{L}(Q) = \prod_{i=1, i \neq j}^{D} (q_{ij})^{n_{ij}(T)} exp(-q_i Y_i(T)) \tag{2.11}$$

pri čemu je $Y_i(t) = \int_0^t 1_{\{X_s = i\}} ds$ ukupno vrijeme koje tvrtka provede u stanju i do trenutka t, a $n_{ij}(t)$ je definiran ranije u radu u Poglavlju 1.2.

Funkcija log-vjerodostojnosti je

$$\log \mathcal{L}(Q) = \sum_{i=1, i\neq j}^{D} \log{(q_{ij})} n_{ij}(T) - \sum_{i=1, i\neq j}^{D} q_{ij} Y_i(T). \tag{2.12}$$

Stoga je MLE za elemente generatorske matrice dan s

$$\hat{q}_{ij} = \frac{n_{ij}(T)}{Y_i(T)}. \tag{2.13}$$

Primjer 2.2. *Procjena prijelaznih vjerojatnosti na jednostavnom primjeru.*
Ilustrirajmo procjenitelj kroz jednostavan primjer na konkretnoj bazi podataka. Ovaj primjer objašnjava osnovnu ideju same procedure. Primjer je preuzet iz D. Lando*, Credit risk modeling: Theory and applications.*
Promotrimo rejting sustav sastavljen od dvije nondefault rating kategorije, A i B, i default kategorije D. Pretpostavimo da smo promatrali jednogodišnju povijest 20 tvrtki, od kojih 10 kreće iz kategorije A, a 10 iz kategorije B. Nadalje, pretpostavimo da tijekom godine, neka tvrtka iz kategorije A prelazi u rejting kategoriju B, tamo i ostaje do kraja godine. Zatim, pretpostavimo da je tijekom istog perioda neka tvrtka iz B kategorije napredovala, ostaje u rejting kategoriji A do kraja tog perioda, dok tvrtka koja kreće iz B i doživi default nakon 6 mjeseci, ostaje tamo preostali dio perioda. Za ilustraciju ovog primjera uzmimo sljedeće:

$$\hat{q}_{AB} = \frac{n_{AB}(1)}{Y_A(1)} = \frac{1}{\frac{119}{12}} = 0.10084.$$

Vrijeme koje tvrtka provede u stanju A do trenutka 1 *procjenjujemo s* $\frac{119}{12}$. *Postupajući na sličan način s ostalim unosima, pripazivši na to da je stanje D apsorbirajuće stanje, dobivamo procjenjenu generatorsku matricu*

$$\hat{Q} = \begin{pmatrix} -0.10084 & 0.10084 & 0 \\ 0.10435 & -0.20868 & 0.10435 \\ 0 & 0 & 0 \end{pmatrix}.$$

Odavde dobivamo MLE jednogodišnje prijelazne matrice kao

$$\widehat{P_1} = \begin{pmatrix} 0.9087 & 0.08657 & 0.00475 \\ 0.0895 & 0.81607 & 0.09434 \\ 0 & 0 & 1 \end{pmatrix}.$$

Vjerojatnost da jedna od 10 tvrtki čije je početno stanje A tokom jedne godine ostane u toj kategoriji je 0.9087. *Vidimo kako uobičajena metoda ne uviđa rizik od defaulta u kategoriji A zato što tvrtke ne odlaze u default direktno iz A. Primjetimo da metode u neprekidnom vremenu daju strogo pozitivne procjene defaulta iz rejting kategorije A tokom jedne godine unatoč tome da niti jedna tvrtka u uzorku ne odlazi u default direktno iz A. To se događa iz razloga što je vjerojatnost prijelaza u kategoriju B i vjerojatnost odlaska u default iz B pozitivna.*

2.3. Razlika između Markovljevih lanaca u diskretnom i neprekidnom vremenu: primjer na stvarnim podacima

U ovom poglavlju promotrimo kako djeluju metode u diskretnom, a kako metode u neprekidnom vremenu na primjeru preuzetom iz D. Lando, *Credit risk modeling: Theory and applications*. Primjer je napravljen na temelju podataka preuzetih od Moody's agencije. Moody's je jedna od tri najveće kreditne rejting agencije spomenute u uvodu rada. Rejting kategorije ove agencije prikazane su u Tablici 2.1.

Tablica 2.1: Rejting kategorije Moody's agencije

Rejting	**Opis značenja**
	Investicijski razred
Aaa	Najviša kvaliteta i najniži kreditni rizik
Aa1 Aa2 Aa3	Visoka kvaliteta i jako nizak rizik
A1 A2 A3	Visoka do srednja kvaliteta i nizak rizik
Baa1 Baa2 Baa3	Srednja kvaliteta s umjerenim rizikom
	Podinvesticijski razred
Ba1 Ba2 Ba3	Srednja-niža kvaliteta i zabilježen rizik
B1 B2 B3	Niža kvaliteta i visok kreditni rizik
Caa1 Caa2 Caa3	Niska kvaliteta i jako visok rizik
Ca	Jako blizu defaulta ili u defaultu, ali s mogućim oporavkom
D	Najniža kvaliteta, najčešće u defaultu i s malom vjerojatnošću oporavka

Promatrat ćemo podatke US korporacijskih izdavatelja u razdoblju od 1997. do 2001. Na temelju tih podataka, procjenjujemo generatorsku matricu korištenjem MLE procjenitelja (2.13) i dobivamo iz toga procjene jednogodišnjih prijelaznih vjerojatnosti na temelju procjenjene generatorske matrice. Zatim to usporedimo s multinomnim procjeniteljem (1.5) u diskretnom vremenu. Rezultate možemo vidjeti u Tablicama od 2.2 do 2.4. Naglasit ćemo dva promatranja.

Prvo, svi nul-elementi dobiveni multinomnom metodom su zamijenjeni elementima različitim od nule kada koristimo eksponent generatorske matrice. Primjerice, vjerojatnost prijelaza u jednoj godini s Aa na Ba ili niže je oko 3 bazna poena[5] u neprekidnom vremenu (vidi Tablicu 2.3), dok je u diskretnom vremenu procijenjeno s 0. Ali primjetimo da to nije jednoznačno. Naime, multinomna metoda daje veću vjerojatnost defaulta za tvrtke iz kategorije A unutar godine.
Drugo, promotrimo veliku razliku između jednogodišnje vjerojatnosti odlaska u default iz Caa. Ta razlika uzrokovana je povijesnim rejtingom. Multinomni procjenitelj uključit će samo prijelaze iz B kategorije u default i neće zabilježiti kratki boravak u Caa. Kratki boravak u ovoj niskoj kategoriji doprinosi većoj procjeni vjerojatnosti defaulta.

Tablica 2.2: Procijenjena generatorska matrica za US izdavatelje iz baze podataka od Moody's agencije za razdoblje od 01.01.1997. - 31.12.2001.

	Aaa	Aa	A	Baa	Ba	B	Caa	D
Aaa	-0.116959	0.096319	0.020640	0	0	0	0	0
Aa	0.014020	-0.116165	0.100142	0.002003	0	0	0	0
A	0.002781	0.042825	-0.148496	0.100666	0.002225	0	0	0
Baa	0.001090	0.002180	0.068131	-0.165150	0.085573	0.005996	0.001635	0.000545
Ba	0	0	0.006767	0.143236	-0.278577	0.121807	0.003384	0.003384
B	0	0.001023	0.003070	0.015349	0.065489	-0.317213	0.207724	0.024558
Caa	0	0	0	0	0	0.043140	-0.471457	0.428317
D	0	0	0	0	0	0	0	0

Tablica 2.3: Procijenjena matrica 1-godišnjih prijelaznih vjerojatnosti ($P = exp(Q^*)$) na temelju procijenjene generatorske matrice za razdoblje od 01.01.1997. - 31.12.2001.

	Aaa	Aa	A	Baa	Ba	B	Caa	D
Aaa	0.890255	0.086193	0.022366	0.001129	0.0000524	0.00000382	0.000000773	0.000000359
Aa	0.012615	0.829832	0.088104	0.006133	0.000286	0.0000234	0.00000495	0.00000236
A	0.002753	0.037812	0.866873	0.086560	0.005360	0.000493	0.000100	0.0000500
Baa	0.001052	0.003225	0.058826	0.855795	0.069110	0.008832	0.002029	0.001131
Ba	0.0000788	0.000365	0.009686	0.116332	0.764724	0.091264	0.011344	0.006206
B	0.0000189	0.000909	0.003175	0.015930	0.049378	0.734295	0.140670	0.055625
Caa	0.000000245	0.0000174	0.0000582	0.000294	0.001000	0.029196	0.627044	0.342391
D	0	0	0	0	0	0	0	1

Tablica 2.4: 1-godišnje prijelazne vjerojatnosti procijenjene multinomnom metodom za razdoblje od 01.01.1997. - 31.12.2001.

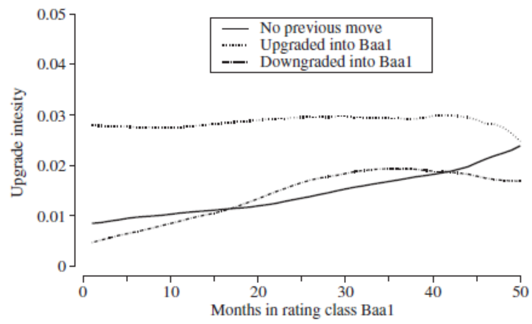
	Aaa	Aa	A	Baa	Ba	B	Caa	D
Aaa	0.841722	0.088608	0.012658	0	0	0	0	0
Aa	0.005874	0.888399	0.070485	0.002937	0	0	0	0
A	0.002165	0.028571	0.846502	0.063636	0.003896	0.001299	0.001299	0.000433
Baa	0.001596	0.001197	0.042282	0.861588	0.042282	0.008775	0.001596	0.001994
Ba	0	0	0.004061	0.104569	0.699492	0.061929	0.006091	0.006091
B	0	0.000804	0.000804	0.008039	0.040193	0.713826	0.097267	0.062701
Caa	0	0	0	0	0	0.029255	0.667553	0.210106

[5]Bazni poen je jedinica mjere kojom se iskazuju mali pomaci kamatne stope, deviznih tečajeva ili prinosa na obveznice. Jedan bazni poen je jedna stotinka jednog postotnog poena.

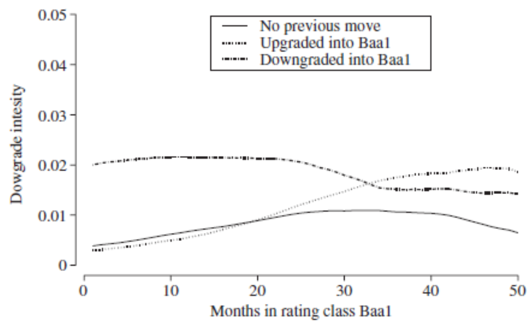
Generatorska matrica ima dodatnu značajku, čini svoje izračune korisnim voditeljima rizika. Prikazuje vrstu direktnog prijelaza koja se dogodila u određenom trenutku. Nule u generatorskoj matrici znače da se nije dogodio direktni prijelaz između promatranih stanja. Važno je napomenuti da su prijelazi preko više od dvije kategorije rijetki.

2.4. Markovljevo svojstvo u primjenama

Pretpostavka da vrijedi Markovljevo svojstvo često je neodrživa u praksi, tj. često nije primijenjiva. Naime, rejting agencije se suočavaju s višestrukim nazadovanjem i preokretom u rejtingu, te se tome žele oduprijeti. Međutim, to oboje ima tendenciju da stvara odstupanja procesa od Markovljevog slučaja. Zamislimo, primjerice, da se stanje neke tvrtke naglo pogorša. Opiranje višestrukom nazadovanju vodi do uzastopnog nazadovanja s kraćim vremenima zaustavljanja u posrednim stanjima nego karakteristične tvrtke u tim stanjima. Prosječan intenzitet tijekom razdoblja za US tvrtke ocijenjene od Moody-s agencije ilustrira efekte koji ne podržavaju Markovljevo svojstvo. Promotrimo Slike 2.1 i 2.2 koje prikazuju glatko uzdizanje i spuštanje intenziteta iz Baa1 tokom cijelog razdoblja. Intenziteti su procijenjeni kao funkcije vremena provedenog u stanju i podijeljeni prema tome je li Baa1 postignut kroz uspon ili pad.



Slika 2.1: Intenzitet napretka iz Baa1 kao funkcija vremena provedenog u stanju i za različite ulaske u razred.



Slika 2.2: Intenzitet nazadovanja iz Baa1 kao funkcija vremena provedenog u stanju i za različite ulaske u razred.

Kao što je prikazano grafovima, postoji značajan trenutni učinak i za nazadovanje i za napredak. Tvrtke koje su nazadovale u Baa1 imaju intenzitet napretka oko 6 puta manji od tvrtki koje su napredovale u ovaj razred. Ovaj učinak traje kroz jako dugo razdoblje. Slično, intenzitet nazadovanja je puno veći za tvrtke koje su nazadovale u Baa1 nego za tvrtke koje su napredovale u ovaj razred.
Inspirirani ovim grafovima, napravimo jednostavan primjer kako procjenitelj generatorske matrice može dati slabu procjenu vjerojatnosti ako su narušene Markovljeve pretpostavke. Razmotrimo uzorak od 100 tvrtki početno ocijenjenih s A i 100 tvrtki s početnom ocjenom B, gdje su A i B nondefault rating kategorije. Dodajmo još i

Tablica 2.5: Povijesni rating tvrtki čiji se rating mijenjao tokom razdoblja od 2 godine. Ostalih 90 A-rated i 85 B-rated tvrtki ostaje u svojim početnim kategorijama.

Tvrtke	Vrijeme	Novo stanje	Vrijeme	Novo stanje	Vrijeme
A1	1	B*	2	B	21
A2	3	B*	3	B	18
A3	5	B*	4	B	15
A4	7	B*	1	B	16
A5	9	B*	4	D	-
A6	11	B*	13	-	-
A7	13	B*	5	B	6
A8	15	B*	8	D	-
A9	17	B*	3	B	4
A10	19	B*	5	-	-
B1	3	D	-	-	-
B2	6	D	-	-	-
B3	9	D	-	-	-
B4	15	D	-	-	-
B5	18	D	-	-	-
B6	1	A	23	-	-
B7	3	A	21	-	-
B8	5	A	19	-	-
B9	6	A	18	-	-
B10	8	A	16	-	-
B11	9	A	15	-	-
B12	14	A	10	-	-
B13	15	A	9	-	-
B14	16	A	8	-	-
B15	18	A	6	-	-

"pobuđeno" stanje B* u koje tvrtke prelaze kada nazaduju iz stanja B. Ovo stanje smatramo stanjem koje ima veću vjerojatnost odlaska u default, a manju vjerojatnost napretka nego normalno stanje. Tvrtke se mogu, nakon određenog vremenskog perioda "normalizirati" i prijeći u B kategoriju. U stvarnom svijetu tražimo pokazatelja daljnjeg nazadovanja koji bi pomogao razlikovati kategorije B i B*. Ako prijelazi nisu zabilježeni, suočavamo se s problemom procjene. Za ovaj primjer usporedimo statističku analizu izvedenu na nerazlikovanju dvije B kategorije uz onu gdje su ova dva stanja zabilježena kao zasebna. Pretpostavimo da pratimo povijest tvrtki dvije godine. To su tvrtke čija je povijest rejting promjena prikazana u Tablici 2.5. Tvrtke A1-A10 su one čiji je početni rejting razred A, a B1-B15 čiji je početni razred B. Usporedimo sada jednogodišnje procjene prijelaznih vjerojatnosti na temelju neprekidnog procjenitelja s klasičnim multinomnim procjeniteljem. Najprije je procijenjena generatorska matrica, te su rezultati prikazani u Tablici 2.6. Zatim je uzet eksponent te matrice kako bi se

Tablica 2.6: Procjenitelj generatorske matrice bez stanja B*.

Stanje	A	B	D
A	-0.0499	0.0499	0
B	0.0519	-0.0882	0.0363
D	0	0	0

dobile neprekidne procjene jednogodišnjih prijelaznih vjerojatnosti. Rezultat ovoga je prikazan Tablicom 2.7. Primjetimo jako malu vjerojatnost defaulta iz A kategorije. Vjerojatnost da neka tvrtka ode u default iz rejting kategorije A unutar jedne godine iznosi 0.0009.

Tablica 2.7: 1-godišnje prijelazne vjerojatnosti dobivene eksponentom generatorske matrice.

Stanje	A	B	D
A	0.9525	0.0466	0.0009
B	0.048	0.918	0.0344
D	0	0	1

Nadalje, pogledajmo multinomni procjenitelj koji uzima stanja samo na početku i na kraju godine. Prikazan je u Tablici 2.8. Možemo vidjeti kako je ovdje vjerojatnost defaulta iza A puno veća, iznosi 0.005. Ovo nije uočeno generatorskom matricom iz razloga što kada tvrtka nazaduje u B ona postaje dio ukupne mase tvrtki iz rejting kategorije B.

Tablica 2.8: Procjena 1-godišnjih prijelaznih vjerojatnosti multinomnim procjeniteljem.

Stanje	A	B	D
A	0.95	0.045	0.005
B	0.0508	0.9184	0.0305
D	0	0	1

Ako bismo dodali stanje B* procjenitelj generatorske matrice bi bio u prednosti. Dobili bismo procjenitelja generatorske matrice prikazanog Tablicom 2.9 pomoću koje dobivamo procjene jednogodišnjih prijelaznih vjerojatnosti prikazane u Tablici 2.10. Zatim te rezultate usporedimo s multinomnim procjeniteljem iz Tablice 2.11. Usredotočimo se ponovno na vjerojatnost defaulta iz A kategorije, procjenitelj na temelju generatorske matrice je oko 8 puta veći nego bez stanja B* (procjena vjerojatnosti prijelaza iznosi 0.0071) i oko 60% veći nego procjenitelj u diskretnom vremenu koji uključuje stanje B*.

Tablica 2.9: Procjenitelj generatorske matrice uključujući stanje B*.

Stanje	A	B	B*	D
A	-0.0499	0.0499	0	0
B*	0	-2	1.5	0.5
B	0.0530	0	-0.0794	0.0265
D	0	0	0	0

Tablica 2.10: Procjena 1-godišnjih prijelaznih vjerojatnosti dobivena eksponentom generatorske matrice.

Stanje	A	B	B*	D
A	0.952	0.021	0.0202	0.0071
B*	0.0215	0.1356	0.6158	0.2271
B	0.0496	0.0007	0.924	0.0256
D	0	0	0	1

Tablica 2.11: Procjena 1-godišnjih prijelaznih vjerojatnosti dobivena multinomnom metodom u diskretnom vremenu.

Stanje	A	B	B*	D
A	0.95	0.015	0.03	0.005
B*	0	0.5	0	0.5
B	0.0513	0	0.92308	0.0256
D	0	0	0	1

Razumijevanje mehanike ovog primjera omogućava nam razumijevanje velike razlike između vjerojatnosti defaulta iz jako loše kategorije, primjerice Caa, i iz kategorije B procijenjene različitim metodama.

3 Markovljev model kreditnog rizika za potrošačke kredite

U praksi je puno manje zastupljena analiza kreditnog rizika potrošačkih kredita od analize kreditnog rizika korporativnih obveznica. Naime, kako je bihevioralni skor analogan rejtingu u korporativnom kreditnom riziku, u ovom poglavlju će biti napravljen model Markovljevih lanaca na temelju bihevioralnog skora. Takav model je privlačan zajmodavcima zato što oni računaju bihevioralni skor za sve svoje dužnike.
Izvodila su se mnoga istraživanja o modeliranju kreditnog rizika u portfeljima potrošačkih kredita. Primjerice, Musto and Souleles (2005.) koriste cijene kapitala kao analogiju za promjene u vrijednostima portfelja potrošačkih kredita. Koriste bihevioralni skor na način da uzimaju mjesečne promjene u bihevioralnim skorovima kao povrat na imovinu prilikom primjene svog modela. Andrade and Thomas (2007.) opisuju strukturni model za kreditni rizik u kome je bihevioralni skor zamjena za kreditnu sposobnost dužnika. Pomoću istraživanja jednog brazilskog kreditnog zavoda zaključili su kako je slučajna šetnja najbolji model za kreditnu sposobnost. Nadalje, Malik and Thomas (2010.) su razvili hazard model kojim se modelira vrijeme do defaulta za potrošačke kredite. Faktori rizika koje oni koriste prilikom razvijanja svog modela kreditnog rizika su bihevioralni skor, starost kredita i ekonomska varijabla. Thomas (2009. b) prikazuje modele za potrošački kreditni rizik i ističe analogiju s nekim već uspostavljenim modelima kreditnog rizika. Lando and Turnbull (1997.) donose pristup Markovljevih lanaca koji je postao popularan u modeliranju dinamike kreditnog rizika u portfeljima potrošačkih kredita. Ideja je bila opisati dinamiku rizika u terminima prijelaznih vjerojatnosti između različitih ocjena rejting agencija. Modeli Markovljevih lanaca su i ranije bili korišteni za potrošačke kredite, ali niti u jednom objavljenom članku za prostor stanja nije uzet bihevioralni skor niti je cilj bio modelirati procjene kreditnog rizika. Prvi primjer je bio model Markovljevih lanaca za potrošačke otplate koji je razvio Cyert (1962). Nadalje su Schneiderjans and Lock (1994) koristili model Markovljevih lanaca za modeliranje marketinških aspekata za upravljenje potrošačkih veza u bankovnom okruženju. Modeli Markovljevih lanaca koji se temelje na bihevioralnom skoru se ponekad koriste u industriji, ali uglavnom kao način procjene potrebnih zaliha.
U ovom poglavlju, napraviti ćemo model Markovljevih lanaca na temelju bihevioralnog skora. Najprije ćemo pogledati svojstva skora i Markovljevih lanaca, zatim ćemo opisati model Markovljevih lanaca na temelju potrošačkih kredita. Prijelazne vjerojatnosti Markovljevog lanca procijeniti ćemo logističkom regresijom.

3.1. Bihevioralni skor i Markovljevi lanci

Zajmodavci obnavljaju bihevioralni skor svakog mjeseca kako bi pristupili kreditnom riziku svakog zasebnog dužnika. Smatra se kako je skor vjerojatnost da dužnik bude

"Loš" i tako doživi default unutar određenog vremenskog perioda. Oni dužnici koji nisu svrstani u "Loše" klasificirani su kao "Dobri". U daljnjem tekstu "Loše" dužnike označavamo s B, a "Dobre" s G. Dužnik s karakteristikom x (koja može opisivati otplate dužnika, osobni dohodak, kamate i slično) ima skor $s(x)$. Skor možemo modelirati logističkom regresijom koja se može shvatiti kao linearna regresija u kojoj je ovisna varijabla nelinearna funkcija vjerojatnosti da će dužnik biti dobar. U tom slučaju jednadžba skora je sljedeća:

$$s(x) = w_0 + w_1x_1 + w_2x_2 + \cdots + w_mx_m, \tag{3.1}$$

gdje su $w_0, w_1, w_2, \ldots, w_m$ težine, a $x_1, \ldots, x_m$ su karakteristike dužnika. Skor definiran na ovaj način odgovara sljedećem:

$$s(x) = \ln\left(P(G|x)\right)P(B|x) = \ln\left\{(P_G/P_B)(P(x|G)/P(x|B))\right\}, \tag{3.2}$$

gdje je P_G udio G dužnika, a P_B je udio B dužnika.
Kako bismo pristupili bihevioralnom skoru dužnika u trenutku t, kojeg ćemo označiti s B_t, preko Markovljevih lanaca, moramo ga podijeliti u intervale od kojih će svaki predstavljati stanje Markovljevog lanca. Bihevioralni skor B_t je povezan s kreditnom sposobnošću dužnikove otplate kredita, U_t, koja također ovisi o duljini vremena otkada je podignut kredit i trenutnoj ekonomskoj situaciji. Model gradimo pod pretpostavkom da je B_t jedno od konačnog broja stanja, $\{s_0 = D, s_1, \ldots, s_n, C\}$, pri čemu $s_i, i > 0$ predstavlja i-ti interval bihevioralnog skora, $s_0 = D$ je stanje defaulta, a C je stanje u kojem se dužnik nalazi ako je otplatio ili zatvorio račun. C i D su apsorbirajuća stanja, te nema prijelaza iz tih stanja. Prisjetimo se kako nam Markovljevo svojstvo govori da stanje skora u neposrednoj budućnosti ovisi samo o sadašnjem stanju, a ne ovisi o prošlim stanjima. Neka je dan bihevioralni skor u stanju $s_i, i = 1, \ldots, n$ u trenutku $t-1$, onda varijablu U_t u trenutku t zapisujemo kao U_t^i. Za aktivni račun U_t^i se definira tako da je veza između B_t i U_t^i sljedeća

$$B_t = s_j \Leftrightarrow \mu_j^i \leq U_t^i \leq \mu_{j+1}^i, j = 0, \ldots, n,$$

gdje su μ_j^i vrijednosti sposobnosti otplate kredita koje odgovaraju krajnjim točkama intervala s_i. Pretpostavimo da je sposobnost otplate kredita, U_t^i, povezana s nekom varijablom x_{t-1} linearnom regresijom

$$U_t^i = -\beta_i x_{t-1} + \varepsilon_t^i, \tag{3.3}$$

pri čemu je β_i koeficijent, a ε_t^i je standardna greška. Prijelazne vjerojatnosti od B_t su u ovom slučaju dane na sljedeći način:

$$\begin{aligned}
P(B_t = D|B_{t-1} = s_i) &= logit(\mu_1^i + \beta_i x_{t-1}),\\
P(B_t = s_1|B_{t-1} = s_i) &= logit(\mu_2^i + \beta_i x_{t-1}) - logit(\mu_1^i + \beta_i x_{t-1}),\\
&\vdots\\
P(B_t = s_n|B_{t-1} = s_i) &= 1 - logit(\mu_n^i + \beta_i x_{t-1}).
\end{aligned} \tag{3.4}$$

Model kojem težimo je onaj u kojem pretpostavljamo da U_t^i ovisi o nekoj ekonomskoj varijabli, vremenu od otvaranja kredita i prethodnom bihevioralnom skoru. Pod tim pretpostavkama sposobnost otplate kredita u trenutku t, ako je u trenutku $t-1$ bila u stanju i, dana je s

$$U_t^i = -\sum_{k=2}^{K} a_{ik} State_{t-k} - b_i EcoVar_{t-1} - c_i MoB_{t-1} + \varepsilon_t^i, \qquad (3.5)$$

gdje $State_{t-k}$ označava stanje dužnika u trenutku $t-k$, $EcoVar_{t-1}$ je ekonomska varijabla u trenutku $t-1$, MoB_{t-1} označava vrijeme od podizanja kredita do trenutka $t-1$, a, b, i c koeficijenti, a ε_t^i je standardna greška. Sve dok U_t^i ovisi o i, stabilnost otplate kredita u trenutku t ovisi o stanju u trenutku $t-1$, pa tako i bihevioralni skor u trenutku t ovisi o stanju u trenutku $t-1$. Ako je $a_{ik} \neq 0$ onda stabilnost otplate kredita u trenutku t ovisi o stanju u trenutku $t-k$ pa bi Markovljev lanac odgovarajućeg bihevioralnog skora bio reda k. Markovljeve lance višeg reda u ovome radu nećemo razmatrati.

3.2. Izgradnja Markovljevog modela

Izgradnja Markovljevog modela u Malik M., Thomas L.C., *Modeling the credit risk of portfolios of consumer loans* napravljena je na klijentima glavne banke UK. Podaci su prikupljani od siječnja 2001. do prosinca 2005., a model gradimo na osnovu 50000 dužnika od siječnja 2001. do prosinca 2004. Baza podataka se sastoji od mjesečnog bihevioralnog skora dužnika, vremena kada je otvoren račun, vremena do defaulta ili do zatvaranja računa i određenih ekonomskih varijabli. Pretpostavljamo da je svaki dužnik koji kasni s plaćanjem rate kredita više od 90 dana otišao u default.
Nadalje, kako bismo analizirali promjene u distribuciji bihevioralnog skora, potrebno ga je najprije podijeliti u različite segmente. To radimo tako da promatramo decile distribucije skora svih dužnika. Dobijemo pet segmenata skora, $s_1 = [113, 680], s_2 = [681, 700], s_3 = [701, 715], s_4 = [716, 725]$ i s_5 =[726 i više]. Ovih pet segmenata čine skup stanja Markovljevog lanca, zajedno sa stanjima C i D. Kako je bihevioralni skor obnavljan svakog mjeseca, moguće je procijeniti 1-koračnu prijelaznu matricu. Budući da ima jako malo prijelaza u jednom mjesecu, umjesto 1-koračne, procijeniti ćemo 3-koračnu prijelaznu matricu. Kao što smo mogli i očekivati, dužnici koji se nalaze u najmanje rizičnom stanju (s_5) imaju najveću vjerojatnost ostanka u tom stanju u idućem kvartalu, a ta vjerojatnost iznosi 0.884. Također, najveća vjerojatnost defaulta je iz najrizičnijeg stanja (s_1), iznosi 0.067, a najmanja je iz stanja s_5.
Zatim, zanima nas koje ekonomske varijable imaju učinka na rizik od defaulta. Razmotriti ćemo sljedeće varijable:

- postotna promjena indeksa potrošačkih cijena tijekom 12 mjeseci,
- mjesečni prosjek stope kreditiranja,

Tablica 3.1: Procjena 3-koračne prijelazne matrice

Stanja	13-680	681-700	701-715	716-725	726-više	C	D
13-680	0.490	0.221	0.096	0.040	0.040	0.047	0.067
681-700	0.157	0.347	0.251	0.096	0.112	0.028	0.008
701-715	0.060	0.136	0.359	0.181	0.234	0.026	0.005
716-725	0.030	0.061	0.157	0.283	0.441	0.025	0.003
716-više	0.007	0.012	0.027	0.043	0.884	0.024	0.002

- godišnji povrat na FTSE 100,
- postotna promjena u BDP-u uspoređena s ekvivalentnim kvartalom u prethodnoj godini,
- stopa nezaposlenosti u UK,
- postotna promjena u internet kreditiranju tijekom 12 mjeseci.

Kako bismo odlučili koje od ovih varijabli uključujemo u model, napraviti ćemo korelacijsku matricu za svih 6 ekonomskih varijabli. Vidimo kako je kamatna stopa negativno

Tablica 3.2: Korelacijska matrica ekonomskih varijabli

Varijabla	Kamatna stopa	% promjena u CPI-u	% promjena u BDP-u	% promjena u net kreditiranju	Stopa nezaposlenosti	FTSE 100
Kamatna stopa	1	**-0.51**	**0.34**	0.14	0.01	**0.39**
% promjena u CPI-u	**-0.51**	1	-0.11	-0.23	**-0.45**	-0.09
% promjena u BDP-u	**0.34**	-0.11	1	**0.85**	**-0.71**	**0.87**
% promjena u net kreditiranju	0.14	-0.23	**0.85**	1	**-0.49**	**0.70**
Stopa nezaposlenosti	0.01	**-0.45**	**-0.71**	**-0.49**	1	**-0.73**
FTSE 100	**0.39**	-0.09	**0.87**	**0.70**	**-0.73**	1

korelirana sa postotnom promjenom u CPI-u, a pozitivno korelirana s postotnom promjenom BDP-a i FTSE 100. Slično, postotna promjena u net kreditiranju je negativno korelirana sa stopom nezaposlenosti, a pozitivno korelirana s BDP-om i FTSE 100. Kako bismo se uvjerili da ekonomske varijable imaju učinka na prijelazne vjerojatnosti tijekom vremena, procijenimo prijelaznu matricu tako da u model uključimo dvije ekonomske varijable, kamatnu stopu i net kreditiranje. U Tablici 3.3 su dane procijenjene prijelazne matrice, prva za vremenski period od siječnja do prosinca 2001., a druga od kolovoza 2003. do rujna 2004. Primjetimo kako su obje matrice slične prijelaznoj matrici iz Tablice 3.1. Također, primjetimo nekoliko značajnih razlika između prijelaznih vjerojatnosti ovih dviju matrica. Primjerice, dužnici koji su se našli u stanju

Tablica 3.3: Usporedba prijelaznih matrica za dva vremenska perioda

Sij - Pro 2001.							
	13-680	681-700	701-715	716-725	726-više	C	D
13-680	**0.5290**	0.2177	0.0924	0.0362	0.0367	**0.0331**	**0.0550**
681-700	**0.1780**	0.3556	**0.2386**	0.0951	0.1040	**0.0214**	0.0072
701-715	**0.0874**	**0.1484**	**0.3525**	0.1790	0.2272	0.0216	0.0029
716-725	**0.0328**	**0.0699**	0.1684	**0.2785**	**0.4264**	0.0212	0.0029
726-	**0.0072**	**0.0135**	**0.0286**	**0.0430**	**0.8839**	**0.0210**	**0.0028**
Kol 2003. - Ruj 2004.							
	13-680	681-700	701-715	716-725	726-više	C	D
13-680	**0.4624**	0.2268	0.0930	0.0403	0.0418	**0.0535**	**0.0822**
681-700	**0.1479**	0.3562	**0.2325**	0.0980	0.1099	**0.0274**	0.0082
701-715	**0.0542**	**0.1342**	**0.3730**	0.1820	0.2289	0.0233	0.0043
716-725	**0.0268**	**0.0563**	0.1617	**0.2934**	**0.4379**	0.0205	0.0033
726-više	**0.0062**	**0.0114**	**0.0265**	**0.0469**	**0.8880**	**0.0190**	**0.0019**

$s_1 = [13-680]$ tijekom prvog razdoblja imaju manju vjerojatnost defaulta od onih koji su se našli u tom stanju tijekom drugog vremenskog perioda. Nadalje, dužnici koji su se našli u stanju s_5 =[726 i više] imaju veću vjerojatnost otplate kredita u prvom razdoblju nego dužnici koji su se našli u tom stanju u drugom vremenskom razdoblju. Za testiranje razlika između odgovarajućih prijelaznih vjerojatnosti ove dvije matrice napravljen je Z-test o jednakosti proporcija. Podebljane prijelazne vjerojatnosti u Tablici 3.3 su one za koje je Z-test pokazao statistički značajnu razliku na razini značajnosti 0.05. Vidimo kako se 20 prijelaznih vjerojatnosti, od ukupno njih 35, razlikuje što nam govori da vremenski period ima značajan utjecaj.
Nadalje, razmotrimo povezanost vremena od otvaranja kredita i rizika od defaulta. U tu svrhu, podijelimo vrijeme od otvaranja kredita, MoB, u sljedećih 6 segmenata:

1. 0-6 mjeseci,
2. 7-12 mjeseci,
3. 13-18 mjeseci,
4. 19-24 mjeseca,
5. 37-48 mjeseci,
6. 48 i više.

Uključimo sada u modeliranje prijelaznih vjerojatnosti bihevioralnog varijablu MoB. U Tablici 3.4 imamo dvije matrice procijenjenih prijaznih vjerojatnosti. Prva, za one dužnike čiji je kredit u 2. segmentu MoB i druga za dužnike čiji je kredit u 6. segmentu MoB. I u ovim matricama prijelaznih vjerojatnosti kao i u prethodnim u koje smo uključili ekonomske varijable, vidimo da ima razlike. Razlike između odgovarajućih prijelaznih vjerojatnosti i ovdje su provjerene Z-testom. Podebljane prijelazne vjerojatnosti u Tablici 3.4 su one između kojih postoji statistički značajna razlika na

Tablica 3.4: Usporedba prijelaznih matrica za dužnike u različitim MoB

7-12 mjeseci MoB							
	13-680	681-700	701-715	716-725	726-više	C	D
13-680	**0.510**	**0.223**	**0.081**	**0.031**	**0.020**	**0.058**	**0.076**
681-700	**0.182**	**0.356**	**0.242**	**0.093**	**0.087**	**0.032**	0.008
701-715	**0.081**	**0.159**	**0.305**	**0.178**	0.256	0.027	**0.005**
716-725	**0.045**	**0.082**	0.147	**0.214**	**0.486**	**0.022**	0.003
726-više	**0.018**	**0.030**	**0.057**	**0.076**	**0.793**	0.023	0.002
48 i vise mjeseci MoB							
	13-680	681-700	701-715	716-725	726-više	C	D
13-680	**0.441**	**0.235**	**0.113**	**0.049**	**0.070**	**0.040**	**0.053**
681-700	**0.136**	**0.325**	**0.256**	**0.107**	**0.144**	**0.025**	0.006
701-715	**0.047**	**0.118**	**0.372**	**0.188**	0.248	0.025	**0.003**
716-725	**0.021**	**0.050**	0.149	**0.304**	**0.447**	**0.026**	0.003
726-više	**0.004**	**0.009**	**0.021**	**0.037**	**0.904**	0.024	0.002

razini značajnosti 0.05, a to je njih 27 od 35.
Možemo zaključiti kako su ekonomske varijable i MoB stvarno povezani s rizikom od defaulta u portfeljima potrošačkih kredita. Stoga, njihovim uključivanjem u model dobivamo bolje procjene prijelaznih vjerojatnosti. Međutim, kako bi takav Markovljev model bio dobar, on zahtijeva barem Markovljev lanac drugog reda, što je izvan granica ovog rada. Za detaljnije pogledati [6].

Literatura

[1] T. R. Fleming, D. P. Harrington, *Estimation for the discrete time nonhomogeneous Markov chains*, Stochastic processes and their applications 7 (1978.), 131-139

[2] H. Frydman, *Maximum-likelihood estimation in the mover-stayer model*, Journal of the American Statistical Association 387(1984), 632-638

[3] B.E. Hansen, *Econometrics*, University of Wisconsin, 2000.

[4] Y. Inamura, *Estimating continuous time transition matrices from discretely observed data*, Bank of Japan, 06-07

[5] D. Lando, *Credit risk modeling: Theory and applications*, Princeton University Press, 2004.

[6] Malik M., Thomas L.C., *Modeling the credit risk of portfolios of consumer loans*, Journal of the Operational Research Society, 61 (2010.), 411-420

[7] Moody's investor service, *Rating symbols and definitions*, 2016.

[8] J. E. Mosimann, *On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions*, Oxford University Press, 1962.

[9] J. R. Norris, *Markov chains*, Cambridge University Press, 1997.

[10] M. J. Osborne, *Mathematical methods for economic theory*, University of Toronto, 2016.

[11] A. Pavković, D. Vedriš, *Redefiniranje uloge agencija za kreditni rejting u suvremenom financijskom sustavu*, Ekonomski fakultet u Zagrebu, 2011.

[12] S. Spilerman, *Extensions of the Mover-Stayer Model*, American Journal of Sociology, 3 (1972.), 599-626

[13] N. Šarlija, *Klasična kreditna analiza*, materijali kolegija Upravljanje kreditnim rizicima, 2008.

[14] N. Šarlija, *Upotreba i primjena kredit skoring modela*, materijali kolegija Upravljanje kreditnim rizicima, 2008.

[15] L. C. Thomas, *A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers*, International Journal of Forecasting, 16 (2000.), 149-172

[16] Z. Vondraček, materijali kolegija *Slučajni procesi*, Sveučilište u Zagrebu, 2010.

Sažetak

Cilj ovog diplomskog rada je objasniti vezu između Markovljevih lanaca i kreditnog rizika. Kako za Markovljeve lance u diskretnom vremenu, tako i za one u neprekidnom potrebno je procijeniti prijelazne vjerojatnosti. Prijelazne vjerojatnosti su procijenjene različitim metodama. U neprekidnom vremenu je to metoda MLE, a u diskretnom radimo s multinomnim procjeniteljima. Prijelazi koji su promatrani su prijelazi tvrtki između različitih rejting kategorija, a najviše smo se fokusirali na vjerojatnost prijelaza iz određene rejting kategorije u kategoriju defaulta. Također su promatrane razlike između procjenitelja u diskretnom i u neprekidnom vremenu, te se boljim pokazao procjenitelj u neprekidnom vremenu. U diskretnom vremenu je razvijen Mover-stayer model koji uključuje dva Markovljeva lanca. Procijenili smo paramtere tog modela i usporedili ih na primjeru. Na kraju izveli smo još jedan Markovljev model u diskretnom vremenu, ali na temelju bihevioralnog skora. Kako bi takav model bio dobar, potrebno je uključiti Markovljeve lance višeg reda, ali to je izvan granica rada.

Ključne riječi: Markovljevi lanci, kreditni rizik, Mover-stayer model, MLE procjenitelj, kreditne rejting agencije, bihevioralni skoring.

Summary

The aim of this paper is to explain the relationship between Markov chains and credit risk. Transition probabilities need to be estimated for Markov chains in discrete time as well as in continuous time. Transition probabilities were estimated with two different methods, MLE method in continuous time, and multinomial estimator in discrete time. Observed transitions were transitions of companies between different rating categories and we were focused on the probability of transition from certain rating category to the category of default. We also observed differences between estimators in discrete and continuous time, and the better estimator was the one in continuous time. Mover-stayer was developed model in discrete time and it includes two Markov chains. We estimated parameters of the model and compare them in one example. Finally, we developed one more Markov model in discrete time, but this one is based on the behavioral score. In order for a model to be good, it should include second order Markov chains, but that is outside the boundaries of this work.

Key words: Markov chain, credit risk, Mover-stayer model, Maximum likelihood estimation, credit rating agencies, behaviour scoring.

Životopis

Rođena sam 17. listopada 1992. godine u Varaždinu. Od 1998. godine pohađala sam Osnovnu školu fra Ilije Starčevića u Tolisi. Tijekom osnovne škole sudjelovala sam na općinskim natjecanjima iz matematike. Zatim, upisala sam Opću gimnaziju u Orašju koju sam završila 2011. godine. Iste godine upisala sam Sveučilišni preddiplomski studij na Odjelu za matematiku kojeg završavam 2014. godine uz završni rad na temu *Školska kriptografija* pod mentorstvom izv. prof. dr. sc. Ivana Matića, te stječem naziv Sveučilišne prvostupnice matematike.
Od 2014. godine studentica sam Sveučilišnog diplomskog studija matematike, smjer Financijska matematika i statistika. Tijekom završne godine diplomskog studija imala sam stručnu praksu u tvrtki Žito d.o.o.