Istraživanja I-37

Procjena matrica kreditnih migracija pomoću agregatnih podataka – bajesovski pristup

Davor Kunovac

Zagreb, ožujak 2012.

HNB

ISTRAŽIVANJA I-37

IZDAVAČ
Hrvatska narodna banka
Direkcija za izdavačku djelatnost
Trg hrvatskih velikana 3, 10002 Zagreb
Telefon centrale: 01/4564-555
Telefon: 01/4565-006
Telefaks: 01/4564-687

WEB-ADRESA
www.hnb.hr

GLAVNI UREDNIK
Evan Kraft

UREDNIŠTVO
Ljubinko Jankov
Gordi Sušić
Maroje Lang
Boris Vujčić

UREDNICA
Romana Sinković

PREVODITELJICA
Tina Antonini

LEKTORICA
Sanda Uzun-Ikić

DIZAJNER
Vjekoslav Gjergja

GRAFIČKA UREDNICA
Gordana Bauk

TISAK
Printera d.o.o.

Tiskano u 400 primjeraka

ISSN 1332-1900 (tisak)
ISSN 1334-0077 (online)

HNB

HRVATSKA NARODNA BANKA

ISTRAŽIVANJA I-37

Procjena matrica kreditnih migracija pomoću agregatnih podataka – bajesovski pristup

Davor Kunovac

Zagreb, ožujak 2012.

Sažetak

U ovom se radu proučava Gibbsov *sampler*, konstruiran u Rodriguez-Yam *et al.* (2004.), koji se može upotrebljavati za procjenu parametara restringiranoga regresijskog modela. Rad ima dva cilja: prikazati efikasnu metodologiju procjene kojom se dosad nije koristilo u literaturi o kreditnom riziku i pomoću nje procijeniti migracijsku matricu koja odgovara dinamici kredita sektoru poduzeća u Hrvatskoj kada su dostupni samo agregatni (nadzorni) podaci o iznosu bankovnih kredita po rizičnim skupinama određenima prema očekivanoj razini naplativosti. Rezultati analize upućuju na to da bajesovski procjenitelj upotrijebljen u ovom radu ima neke važne komparativne prednosti u odnosu na procjenitelje koji su se prije rabili u literaturi.

Ključne riječi:

matrice kreditnih migracija, bajesovsko zaključivanje, restringirani regresijski model, odrezani normalni vektor

JEL klasifikacija:

JEL: C13, C50

Zahvaljujem Gordonu Kempu, Bojanu Basraku i Evanu Kraftu, koji su svojim komentarima prethodnih verzija rada pridonijeli njegovu poboljšanju. Također zahvaljujem Ivanu Huljaku na pomoći oko podataka.

Sadržaj

1. Uvod i motivacija

1.1. Preliminarna rasprava

Ovaj rad ima dva cilja. Prvi je prikazati i ispitati efikasnu metodologiju procjene parametara restringiranoga regresijskog modela kojom se dosad u literaturi o kreditnom riziku nije koristilo. Drugi je cilj pomoću te metodologije procijeniti migracijsku matricu koja odgovara dinamici kredita sektoru poduzeća u Hrvatskoj kad su dostupni samo agregatni (nadzorni) podaci o bankovnim kreditima u svakoj skupini određenoj prema očekivanoj razini naplativosti. Empirijska analiza u ovome se radu stoga bavi modeliranjem migracije bankovnih kredita među različitim rizičnim skupinama. Osnovno pitanje ove analize stoga glasi: Kolika je vjerojatnost da će dobar kredit postati loš (ili nenaplativ) i obrnuto?

Migracijske matrice obično se upotrebljavaju za opisivanje dinamike kreditnog portfelja tijekom vremena i mogu biti od koristi za prognoziranje kvalitete kredita. U slučaju kad su dostupni podaci o prijelazu pojedinačnih kredita iz jedne rizične skupine u drugu, vjerojatnosti prijelaza jednostavno se izračunavaju metodom maksimalne vjerodostojnosti iz relativnih frekvencija dolazaka i odlazaka za svaku skupinu. Međutim, uglavnom su dostupne samo agregatne vremenske serije o iznosima kredita u svakoj kvalitativnoj skupini, koje obično prikuplja regulatorno tijelo. U tom se slučaju vjerojatnosti prijelaza mogu promatrati kao vjerojatnosti prijelaza Markovljeva lanca prvoga reda s diskretnim vremenom i diskretnim prostorom stanja. Problem procjene vjerojatnosti prijelaza tako je ekvivalentan problemu procjene parametara restringiranoga regresijskog modela (Kelton, 1981.; Lee *et al.*, 1965.; Rodriguez-Yam *et al.*, 2004.). No, literatura u kojoj se pri procjeni migracijskih matrica uzimaju u obzir ovi minimalni zahtjevi za podatke nije opsežna. Primjerice, Jones (2005.) procjenjuje matrice prijelaza za portfelj bankovnih kredita u SAD-u pomoću ICLS procjenitelja (engl. *Inequality Constrained Least Squares Estimator*). Zaključivanje otežavaju nepostojanje relevantne asimptotske teorije za ovaj procjenitelj i pripadajuće česte granične procjene. Primjerice, vjerojatnosti prijelaza često se procjenjuju kao nule, što implicira ekstremne zaključke, koji katkad zavaravaju i nije ih jednostavno tumačiti. Kako bi izbjegao probleme s ICLS-om, Christodoulakis (2006.) proučava istu skupinu podataka, ali rabi drugačiju strategiju procjene. Vjerojatnosti prijelaza procjenjuju se Monte Carlo integracijom, kako predlažu Kloek i van Dijk (1978.) te Geweke (1986.). Empirijski rezultati koje navodi Christodoulakis (2006.) u nekim su slučajevima bliski procjenama Jonesa (2005.), ali pokazuju i neke statistički značajne razlike kada je riječ o procijenjenim vjerojatnostima prijelaza. Međutim, simuliranje koje se zasniva na Monte Carlo integraciji često je sporo i nedjelotvorno. Christodoulakis (2006.) ne navodi performanse svoga *samplera*, nego samo navodi da se zaključivanje zasniva na velikom broju ponavljanja (100.000).

U ovom se radu procjenjuju vjerojatnosti prijelaza za tipičan kredit sektoru poduzeća u Hrvatskoj. Rad se u određenoj mjeri nadovezuje na radove Jonesa (2005.) i Christodoulakisa (2006.), no u njemu se vjerojatnosti prijelaza između skupina kredita procjenjuju pomoću alternativne, bajesovske strategije.

1.2. Restringirani regresijski model

Pod pretpostavkom da tranzicija tipičnoga kredita prati Markovljev lanac prvoga reda, pripadajuća se matrica prijelaza procjenjuje restringiranim regresijskim modelom. Stoga ćemo sada detaljnije razmotriti procjenu ovoga modela.

Ekonometričari se u praksi često susreću s problemom procjenjivanja modela linearne regresije, za koji je poznato da neki od njegovih parametara pripadaju a priori danim intervalima. Taj je model u potpunosti definiran skupom nejednakosti koje trebaju poštovati parametri klasičnoga regresijskog modela, a naziva se restringirani regresijski model. Apriorno znanje o tim parametrima uglavnom se oslanja na uvjete koje postavlja ekonomska ili statistička teorija. Alternativno, nejednakosti čine samo ad hoc ograničenja, koja se s obzirom na prirodu relevantnog problema smatraju očitima.

U skupinu ograničenja koja se najčešće susreću u praksi spadaju ograničenja predznaka regresijskih parametara. U teoriji se često pretpostavljaju smjerovi marginalnih efekata pojedinačnih regresora u modelu. Međutim, predznaci procijenjenih koeficijenata ne moraju se nužno slagati s onima koji su predviđeni u teoriji. Na primjer, nedosljednost predznaka često se pojavljuje zbog visoko koreliranih regresora koji katkad pomiješaju pojedinačne efekte u modelu. Kako navodi Geweke (1986.), istraživači ovaj praktičan problem rješavaju različitim pristupima. Najčešće se regresori koji proizvode pogrešan predznak jednostavno izostavljaju iz specifikacije i navode se samo rezultati zasnovani na specifikacijama sa svim očekivanim predznacima. No takva praksa može rezultirati pristranim i neefikasnim procjenama parametara (Lovell i Prescott, 1970.). S druge strane, po konstrukciji nije uvijek moguće jednostavno izostaviti neku varijablu iz linearnog modela, osobito kad osnovni teoretski model potječe iz teorije vjerojatnosti. Radi ilustracije, odavno se zna da se procjena vjerojatnosti prijelaza diskretnoga Markovljeva lanca jednostavno svodi na problem procjene restringiranoga regresijskog modela (Lee *et al.*, 1965.; MacRae, 1977.). Dakako, elementi u svakom retku matrice prijelaza moraju biti nenegativni, manji od jedan ili jedan, i trebaju imati sumu jedan. Isti skup ograničenja mora se stoga pojaviti i u pripadajućoj regresiji.

Umjesto da se traži skup eksplanatornih varijabli koji naposljetku daje parametre s očekivanim predznakom, u alternativnom se pristupu razmatra procjena ograničenog modela u kojemu su svi parametri prisiljeni proizvesti očekivane predznake ili, u općenitijem smislu, zadovoljiti dana ograničenja nejednakosti. Na tragu te ideje Judge i Takayama (1966.) rješavaju kvadratični program pomoću Dantzig-Cottleova algoritma i minimiziraju sumu kvadrata reziduala linearnog modela uz ograničenja nejednakosti na parametrima. Na taj se način dobiva ICLS procjenitelj. S ovim se pristupom povezuje nekoliko većih problema. Asimptotska svojstva ICLS procjenitelja jasna su i dobro potkrijepljena samo za nekoliko jednostavnih slučajeva i stoga nisu od velike koristi u primjeni. Standardni test značajnosti koeficijenata zasnovan na Studentovoj distribuciji može upućivati na dvosmislene zaključke (Lovell i Prescott, 1970.), a momenti procjenitelja teško se izvode (Judge i Takayama, 1966.). S praktičnijeg stajališta, ne uključuju svi standardni ekonometrijski softverski paketi procedure za izračun ICLS procjena. Naposljetku, rješenje dobiveno kvadratičnim programiranjem često se nalazi na granici područja određenog danim nejednakostima i stoga se često smatra neinformativnim.

Kako bi se zaobišli nedostaci ICLS procjenitelja, problemu se pristupilo i iz bajesovske perspektive. U rane teoretske radove spadaju O'Hagan (1973.), Davis (1978.) te Chamberlain i Leamer (1976.). Međutim, bajesovski su procjenitelji postali relevantni za praktičare tek nakon popularizacije računski zahtjevnih Monte Carlo metoda u osamdesetim i devedesetim godinama. Nadovezujući se na Kloeka i van Dijka (1978.), Geweke (1986.) simulira Monte Carlo integracijom uzorak iz aposteriorne distribucije regresijskih parametara ograničenih nejednakošću. Međutim, pokazalo se da je ovaj pristup nedjelotvoran i spor do granice neprovedivosti u slučaju kad prostor parametra definira malo područje i/ili kad je broj parametara koji se procjenjuju velik (Geweke, 1996.). Gibbsov *sampler* namijenjen simulaciji iz aposteriorne distribucije za regresijske parametre ograničene nejednakošću prvi je predstavio Geweke (1996.). Upotrijebljene apriorne distribucije bile su difuzne apriorne distribucije koje se obično upotrebljavaju u linearnim bajesovskim modelima, ali su sada ograničene kako bi odražavale eventualno apriorno znanje o parametrima. Aposteriorna distribucija za koeficijente u ovom je slučaju odrezana višedimenzionalna normalna distribucija i problem uzorkovanja iz aposteriorne distribucije svodi se na uzorkovanje iz odrezane normalne distribucije. Međutim, uzorkovanje iz višedimenzionalnih odrezanih distribucija nije jednostavno, osim ako je riječ o niskim dimenzijama. Geweke se u tu svrhu koristi metodologijom rada Hajivassiliou i McFadden (1990.), Hajivassiliou, McFadden i Ruud (1996.) te Geweke (1991.). Premda je Gewekeov *sampler* pogodan za rješavanje širokog spektra problema, nedostaci su i dalje brojni. Prije svega, broj ograničenja koja se mogu nametnuti regresijskim koeficijentima ne smije premašivati ukupan broj parametara koji se procjenjuju. Nadalje, ograničenja trebaju biti linearno nezavisna. Na primjer, ti nedostaci onemogućuju primjenu ovog *samplera* u procjeni vjerojatnosti prijelaza Markovljeva lanca u okviru restringiranoga regresijskog modela. Osim toga, Gewekeov *sampler* proizvodi visoko autokorelirane realizacije i stoga neefikasno istražuje aposteriornu distribuciju. Naravno, aposteriorne relizacije iz Gibbsova *samplera* realizacije su Markovljeva lanca i, kao takve, po definiciji korelirane. Ipak, problem je ozbiljan i uzrokuje sporu konvergenciju prema aposteriornoj distribuciji.

Naposljetku, Gibbsov sampler koji su konstruirali Rodriguez-Yam, Davis i Scharf (2004.) prevladava probleme koji karakteriziraju prethodne specifikacije. Autori pokazuju da njihov *sampler* može inkorporirati više ograničenja od Gewekeova (1996.), da proizvodi slabije korelirane realizacije i konvergira prema aposteriornoj distribuciji brže od alternativnih specifikacija. Zbog toga se u ovom radu detaljnije proučava *sampler* Rodriguez-Yama, Davisa i Scharfa (2004.).

U ovoj se analizi strategija procjene oslanja na rad Rodriguez-Yama *et al.* (2004.). Vjerojatnosti migracije među različitim skupinama za kreditni portfelj hrvatskih banaka djelotvorno se procjenjuju pomoću Gibbsova *samplera*, gdje se konvergencija bajesovskih procjenitelja postiže nakon manje od 10.000 ponavljanja. Nadalje, procijenjene vjerojatnosti nalaze se unutar područja ograničenog danim nejednakostima i stoga ih je lako interpretirati.

Ostatak rada strukturiran je na sljedeći način. U sljedećem se odjeljku pokazuje kako se može procijeniti matrica kreditnih migracija pomoću agregatnih podataka. Nakon toga, u trećem se odjeljku izlažu tehničke pojedinosti o *sampleru* na koji se oslanja u cijelome radu. Kako bi se mogla procijeniti pouzdanost procjenitelja koji se upotrebljava u ovom radu, u četvrtom se odjeljku dodatno uspoređuju njegove osobine s onima OLS i ICLS procjenitelja u jednostavnom Monte Carlo eksperimentu. U petom se odjeljku navode rezultati procjene, a u posljednjem zaključak.

2. Procjena matrica kreditnih migracija pomoću agregatnih podataka

U ovom se odjeljku prikazuje kako se i pod kojim uvjetima procjena vjerojatnosti prijelaza Markovljeva lanca svodi na procjenu restringiranoga regresijskog modela.

U sklopu procesa upravljanja rizikom financijske institucije često konstruiraju jednostavne ekonometrijske modele kako bi sažele prethodno kretanje različitih financijskih pokazatelja. U slučaju kada su ti modeli korisni za prognoziranje, oni pružaju tim institucijama i njihovim regulatorima važan input pri upravljanju kreditnim rizikom.

Osnovni je sastavni element mnogih modela kreditnog rizika migracijska matrica kreditnog portfelja. U njoj su dane vjerojatnosti migracije kredita u portfelju među različitim skupinama koje su određene očekivanom razinom naplativosti kredita. Primjerice, prirodno je da upravitelje rizicima zanima procjena vjerojatnosti da će dobar (naplativ) kredit postati loš (nenaplativ) ili obrnuto. U slučaju kad mehanizam na kojemu se zasniva migracija među skupinama u sljedećim razdobljima ostane sličan, migracijske matrice postaju važno sredstvo za predviđanje bankovnih gubitaka.

2.1. Matrica prijelaza Markovljeva lanca u regresijskom okviru

Često se pretpostavlja da prijelaz kredita između različitih skupina prati diskretan Markovljev lanac. U slučaju kad su dostupni samo agregatni podaci o udjelima kredita u svakoj skupini, vjerojatnosti prijelaza ne mogu se izravno izračunati i problem procjene migracijske matrice rješava se u okviru restringiranoga regresijskog modela (Lee *et al.*, 1965.; Rodriguez-Yam *et al.,* 2004.), kako slijedi.

Neka $S = \{S_1, \ldots, S_k\}$ označuje diskretan prostor stanja (skupine kredita prema naplativosti) vremenski homogenoga[1] Markovljeva lanca $X = \{X_0, X_1, \ldots\}$. Problem koji nas ovdje zanima jest procjena vjerojatnosti prijelaza između stanja koje sažeto prikazuje $k \times k$ matrica P:

$$P = [p_{ij}]_{i,j=1}^{k} = [P(X_t = S_j \mid X_{t-1} = S_i)]_{i,j=1}^{k}.$$

Strategija procjene ovisi o prirodi raspoloživih podataka. U slučajevima kad se realizacije lanca (tj. krediti po skupinama) promatraju izravno, vjerojatnosti prijelaza jednostavno se dobiju procjenom maksimalne vjerodostojnosti kao (vidi Lee *et al.*, 1968.):

$$P_{ML} = [\hat{p}_{ij}]_{i,j=1}^{k} = \left[\frac{n_{ij}}{\sum_{j=1}^{k} n_{ij}} \right]_{i,j=1}^{k},$$

gdje n_{ij} označuje broj opaženih migracija iz stanja i u stanje j. Ovaj pristup nazivamo procjenom pomoću *mikropodataka*. Suprotno tomu, često se rezultati lanca ne mogu promatrati izravno, nego samo kao agregatni broj jedinica (u našem slučaju kredita) u svakom stanju lanca u vremenu. Pretpostavlja se da su te jedinice međusobno nezavisne i da prate isti Markovljev proces s matricom prijelaza P. Individualni prijelazi u ovom se slučaju s *makropodacima* ne promatraju, pa se strategija procjene razlikuje od prethodnog slučaja.

U tu svrhu, odredimo najprije zajedničku vjerojatnost za dvije uzastopne realizacije lanca

$$P(X_{t-1} = S_i, X_t = S_j) = P(X_{t-1} = S_i)P(X_t = S_j \mid X_{t-1} = S_i) = P(X_{t-1} = S_i)p_{ij} \tag{1}$$

Bezuvjetna vjerojatnost da će lanac u vremenu t biti u stanju j sada je:

$$P(X_t = S_j) = \sum_{i=1}^{k} P(X_{t-1} = S_i, X_t = S_j) = \sum_{i=1}^{k} P(X_{t-1} = S_i)p_{ij} \tag{2}$$

U praksi, vjerojatnosti $P(X_t = S_j)$ procjenjuju se kao omjeri jedinica u stanju j u vremenu t uz oznaku y_{jt}. Ti se omjeri mogu jednostavno dobiti iz agregatne skupine podataka (makropodaci), čak iako se realizacije pripadajućeg lanca ne opažaju. Kao posljedica toga, uzorački pandan od (2) definira sljedeći restringirani regresijski model:

$$y_{jt} = \sum_{i=1}^{k} y_{it-1}p_{ij} + \varepsilon_{jt} \tag{3}$$

uz

$$\sum_{j=1}^{k} p_{ij} = 1 \text{ za sve } i \tag{4}$$

$$p_{ij} \geq 0 \text{ za sve } i, j, \tag{5}$$

gdje su $i = 1,...,$ stanja Markovljeva lanca (odnosno moguće skupine kredita), a y_{jt} je omjer kredita skupine u vremenu t. Osim toga, pretpostavka postojanja jednoga apsorpcijskog (terminalno) stanja uvodi dodatnih k restrikcija:

1 Vremenska homogenost lanca odražava činjenicu da se vjerojatnosti tranzicije ne mijenjaju u vremenu.

$$p_{kk} = 1 \text{ i } p_{k1} = p_{k2} = \ldots = p_{kk-1} = 0 \tag{6}$$

Ograničenja jednakosti uključena su u specifikaciju na standardan način. Uključivanje (4) u (3) i pretpostavka jednoga apsorpcijskog stanja (6) daju *reducirani model* koji se može procjenjivati pomoću ICLS procjenitelja. Međutim, ograničenja jednakosti ovdje nameću ograničenja parametara između jednadžbi (3), pa se zato regresije za sve j trebaju procjenjivati istodobno. Za naše potrebe dovoljno je ilustrirati postupak za lanac s četiri stanja (odnosno $k = 4$) i jednim apsorpcijskim stanjem.

Razmotrimo skup jednadžbi (3) uz uvjet (6), ali sada prikazanih ovako:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \Delta y_4 \end{pmatrix} = \begin{pmatrix} X & 0 & 0 & 0 \\ 0 & X & 0 & 0 \\ 0 & 0 & X & 0 \\ 0 & 0 & 0 & X \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} + \varepsilon,$$

gdje je $y_j = (y_{j2}, \ldots, y_{jT})^\tau$, X zajednička matrica regresora iz (3), p_j označuje stupce tranzicijske matrice, Δ je standardni operator diferenciranja, a $\varepsilon \sim N(0, \sigma^2 I)$ slučajna greška.[2]

Kako bismo nastavili s procjenom, promotrimo stupčastu reprezentaciju ove složene matrice:

$$X_1 = \begin{pmatrix} X \\ 0 \\ 0 \\ 0 \end{pmatrix}, X_2 = \begin{pmatrix} 0 \\ X \\ 0 \\ 0 \end{pmatrix}, X_3 = \begin{pmatrix} 0 \\ 0 \\ X \\ 0 \end{pmatrix} \text{ i } X_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ X \end{pmatrix}$$

Sada sve i (4) možemo napisati kao:

$$p_{i4} = 1 - p_{i1} - p_{i2} - p_{i3} \tag{7}$$

ili $p_4 = \mathbf{1} - p_1 - p_2 - p_3$ u matričnom zapisu, što daje:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \Delta y_4 \end{pmatrix} = (X_1 X_2 X_3 X_4) \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \mathbf{1} - p_1 - p_2 - p_3 \end{pmatrix} + \varepsilon = X_1 p_1 + X_2 p_2 + X_3 p_3 + X_4 (\mathbf{1} - p_1 - p_2 - p_3) + \varepsilon.$$

Naposljetku, imamo

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \Delta y_4 \end{pmatrix} - X_4 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = (X_1 - X_4) p_1 + (X_2 - X_4) p_2 + (X_3 - X_4) p_3 + \varepsilon \tag{8}$$

što je jednadžba koja se može procijeniti u okviru restringiranoga regresijskog modela uz:

$$p_{ij} \geq 0, \text{ za i} = 1,2,3 \text{ i } j = 1,2,3 \tag{9}$$

$$p_{i1} + p_{i2} + p_{i3} \leq 1, \text{ za } i = 1,2,3. \tag{10}$$

2 Činjenica da baratamo podacima o omjerima podrazumijeva *istodobnu* korelaciju različitu od nule između slučajnih grešaka između jednadžbi. Doista, zbrajanje obiju strana u $y_{jt} = \sum_{i=1}^{K} y_{it-1} p_{ij} + \varepsilon_{jt}$ za $j = 1, 2, 3, 4$ uzimajući u obzir ograničenja vjerojatnosti prijelaza kao i $\sum_{j=1}^{4} y_{jt} = 1$ povlači $\sum_{j=1}^{4} \varepsilon_{jt} = 0$ za sve t. Zbog toga najmanji kvadrati daju *nepristrane i konzistentne* (Lee, Judge i Zellner, 1970.), ali *neefikasne* procjene vjerojatnosti prijelaza. Zato MacRae (1977.) predlaže iterativni GLS procjenitelj. Naposljetku, Kelton (1981.) pokazuje ekvivalentnost GLS i OLS procjenitelja uz relativno slabe zahtjeve. Primarna je svrha ovoga primjera usporediti procjene zasnovane na *sampleru* iz Rodriguez-Yam, Davis i Scharf (2004.) s konvencionalnijim rješenjima dobivenima pomoću OLS i ICLS procjenitelja, pa se zato, unatoč neoptimalnosti izbora, pretpostavlja da je kovarijacijska matrica dijagonalna, odnosno $\varepsilon \sim N(0, \sigma^2 I)$.

Kako navode Rodriguez-Yam *et al.* (2004.), broj restrikcija sada je veći od broja parametara koji se procjenjuju, pa stoga alternativni Gibbsov *sampler* (Geweke, 1996.) nije odgovarajuća metoda procjene.

3. Bajesovska procjena restringiranoga regresijskog modela

U ovom se odjeljku opisuje bajesovski procjenitelj za linearni restringirani regresijski model koji predlažu Rodriguez-Yam, Davis i Scharf (2004.). Ove procjene, zasnovane na aposteriornim realizacijama simuliranima pomoću Gibbsova *samplera*, prevladavaju najvažnije probleme prethodnih specifikacija. Autori pokazuju da se u njihov *sampler* može jednostavno uključiti velik broj potencijalno kolinearnih ograničenja, da proizvodi slabije korelirane realizacije i da brže konvergira prema aposteriornoj distribuciji nego konkurentne bajesovske specifikacije. Nasuprot tome, čini se da je klasičan pristup problematičan zbog nepostojanja relevantne asimptotske teorije za ICLS procjenitelje. Štoviše, ovako proizvedena granična rješenja katkad navode na pogrešne zaključke.

U bajesovskom se modelu normalne linearne regresije obično pretpostavlja normalna apriorna distribucija za regresijske koeficijente, tako da je, zbog oblika funkcije vjerodostojnosti, aposteriorna distribucija također normalna. To vrijedi i za ograničenu regresiju, ali uz jednu važnu razliku. Prethodno znanje o koeficijentima ugrađeno je u model pomoću odrezane normalne distribucije kao apriorna distribucija, pa je zbog toga aposteriorna distribucija ponovo iz odrezane normalne distribucije. Za razliku od neograničenog slučaja, teoretske momente ograničene Gaussove distribucije uglavnom nije jednostavno izračunati, pa se radi simulacije uzorka iz aposteriorne distribucije primjenjuju MCMC metode (engl. *Markov Chain Monte Carlo*). Ipak, kao što je već rečeno, ni simulacija iz odrezane višedimenzionalne distribucije nije lak zadatak i procjenitelj za restringirani regresijski model koji su prikazali Rodriguez-Yam, Davis i Scharf (2004.) zapravo je primjena efikasnoga Gibbsova *samplera* iz odrezane normalne distribucije.

U sljedećem se odjeljku najprije opisuje Gibbsov *sampler* za odrezanu normalnu distribuciju, a zatim se pokazuje kako se on prirodno uklapa u okvir restringiranoga regresijskog modela.

3.1. Simulacija uzorka iz odrezane normalne distribucije

Odrezani normalni vektor formalno je definiran sljedećom definicijom (Rodriguez-Yam, Davis i Scharf, 2004.).

Definicija 1. (Odrezani normalni vektor) Ako je R podskup od $\mathbb{R}^k$ s pozitivnom Lebesgueovom mjerom, slučajni k-vektor Y nazivamo odrezanim normalnim vektorom i pišemo $Y \sim N_R(\mu, \Sigma)$ ako je njegova funkcija vjerojatnosti gustoće proporcionalna $f(x;\mu,\Sigma) I_R(x)$, gdje je $f(x;\mu,\Sigma)$ k-dimenzionalna normalna gustoća s vektorom očekivanja μ i kovarijacijskom matricom Σ, a $I_R(x)$ funkcija je indikatora za R.

Gibbsov algoritam za odrezanu normalnu distribuciju oslanja se na neka dobra obilježja koja karakteriziraju linearne transformacije i uvjetne distribucije neograničenoga normalnog vektora, koje je ograničeni vektor također naslijedio. Ti standardni rezultati omogućuju da linearna transformacija ograničenoga normalnog vektora također bude normalno distribuirana, a zajamčena je i normalnost jednodimenzionalnih uvjetnih distribucija. Važnost ovih rezultata proizlazi iz činjenice da se realizacije Gibbsova *samplera* oslanjanju na simulaciju iz uvjetnih distribucija. Formalno se ova dva svojstva mogu navesti u obliku dviju propozicija (vidi Dodatak).

Ovdje je krajnji cilj simulirati slučajni uzorak iz vektora

$$X \sim N_T(\mu, \sigma^2\Sigma), \text{ gdje je } T = \left\{x \in \mathbb{R}^k : Bx \leq b\right\}.$$

Za razliku od sličnih aplikacija, ovdje je dopušteno da matrica ima kolinearne retke, odnosno ograničenja nejednakosti ne moraju biti linearno nezavisna. Neka je kvadratna matrica punog ranga, takva da je

$$A\Sigma A^{\tau} = I_k \tag{11}$$

i neka je $Z = AX$ linearna transformacija od X Propozicije 1 (vidi Dodatak) i (11) podrazumijevaju distribuciju od:

$$Z \sim N_S(A\mu, \sigma^2 A\Sigma A^{\tau}) = (2) = N_S(A\mu, \sigma^2 I_k), \text{ gdje } S = \{Ax : x \in \mathbb{R}^k \ \& \ Bx \leq b\}. \quad (12))$$

Područje S može se napisati kao

$$S = \{z \in \mathbb{R}^k : Dz \leq b\}, \text{ gdje je } D = BA^{-1}. \tag{13}$$

Svrha transformacije vektora X u (12) jest pojednostavljenje kovarijacijske matrice. Zbog toga su sve jednodimenzionalne uvjetne distribucije potrebne za jednu rundu Gibbsova *samplera* definirane u cijelosti ako su poznati μ a i σ. Potrebno je još detaljnije upoznati se s jednodimenzionalnim nejednakosnim skupovima i pripadajućim jednodimenzionalnim uvjetnim distribucijama vektora Z. U tu svrhu, definirajmo:

$$\alpha = A\mu, Z_{-j} = (Z_j, \ldots, Z_{j-1}, Z_{j+1}, \ldots, Z_k) \text{ and } z_{-j} = (z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_k),$$

pri čemu Z_{-j} označuje vektor Z s isključenom j-tom komponentom, a z_{-j} je realni vektor u $\mathbb{R}^k$. Iz propozicije 2 slijedi da je:

$$Z_j | (Z_{-j} = z_{-j}) \sim N_{S_j}(\alpha_j, \sigma^2), \tag{14}$$

gdje je

$$S_j = \{z_j \in \mathbb{R} : z \in S\} = \{z_j \in \mathbb{R} : Dz \leq b\}.$$

Neka je D = $(d_1,...,d_k)$ stupčana reprezentacija od D , a D_{-j} matrica kad se stupac j ukloni iz D. Sada se skup S_j može napisati kao

$$S_j = \{z_j \in \mathbb{R} : d_j z_j \leq b - D_{-j} z_{-j}\}. \tag{15}$$

Svi su objekti D_{-j}, z_{-j}, d_j i b poznati i, stoga, zbog konveksnosti, granice S_j mogu se odrediti iz skupa linearnih nejednakosti $d_j z_j \leq b - D_{-j} z_{-j}$. Nakon izvođenja granica od S_j u (15) jednostavno se izvode simulacije iz jednodimenzionalnih uvjetnih distribucija $Z_j | (Z_{-j} = z_{-j})$ danih u (14). Premda algoritam omogućuje realizaciju iz transformiranog vektora $Z = AX$, prethodno množenje ovoga matricom A^{-1} daje realizaciju iz X. Cijela procedura jednostavno se provodi slijedeći korake algoritma:

Algoritam (Gibbsov *sampler* za odrezani višedimenzionalni normalni vektor)

1. Inicijalizirajmo lanac $z_0 \in S$. Pretpostavimo da u koraku t imamo $z^{(t)} = (z_j^{(t)}, \ldots, z_k^{(t)})$
2. Simulirajmo $z_j^{(t+1)}$ iz $f(z_j | z_1^{(t+1)}, \ldots, z_{j-1}^{(t+1)}, z_{j+1}^{(t)}, \ldots, z_k^{(t)})$, za j = 1,...,k
3. Poništimo linearnu transformaciju: $X^{(t+1)} = A^{-1} z^{(t+1)}$.

3.2. Gibbsov *sampler* za restringirani regresijski model

U ovom se odjeljku uvodi Gibbsov *sampler* namijenjen simuliranju iz aposteriorne distribucije parametara restringiranoga regresijskog modela. Simuliranje koeficijenata regresije iz aposteriorne distribucije svodi se na simuliranje iz odrezane normalne distribucije. U tu se svrhu primjenjuje Gibbsova shema opisana u prethodnom odjeljku.

Osnovni model koji razmatramo jest k-dimenzionalna regresija, gdje koeficijenti regresije trebaju zadovoljiti skup linearnih nejednakosti. U obliku matrice taj *restringirani regresijski model* definiramo na sljedeći način:

$$Y = X\beta + \varepsilon, \text{ tako da } B\beta \leq b,$$

pri čemu je $\varepsilon \sim N(0, \sigma^2 I_n)$ standardni bijeli šum, B označuje poznatu $l \times k$ matricu restrikcija, pri čemu l može biti veći od k, a b je poznati k-vektor. Nejednakosti definiraju skup:

$$T = \{\beta \in \mathbb{R}^k : B\beta \leq b\}.$$

Skup parametara koji se procjenjuje uključuje regresijske koeficijente β i varijancu šuma σ. Oni zajednički tvore $k+1$-dimenzionalni vektor:

$$\theta = (\beta, \sigma).$$

Strategija procjene je bajesovska, pa je stoga model u cijelosti definiran ako je poznata apriorna distribucija za θ. Standardan odabir za bajesovski linearni model jest postaviti normalnu apriornu distribuciju za β i *inverznu-gama*[3] apriornu distribuciju za σ.

$$\beta \sim N_T(\mu_0, \sigma_0^2 (X^\tau X)^{-1}) \tag{16}$$

$$\sigma^2 \sim IG(\nu, \lambda), \tag{17}$$

gdje σ_0, ν i λ označuju poznate skalare, a μ_0 poznati vektor. Pretpostavimo također da su β i σ nezavisni. Prethodno znanje o koeficijentima ugrađeno je u model korištenjem odrezane normalne distribucije kao apriorne distribucije, a, slijedom toga, aposteriorna distribucija također je odrezana normalna. Za razliku od neograničenog slučaja, momenti iz ograničene Gaussove distribucije najčešće se ne izvode jednostavno, pa se zato MCMC metode primjenjuju za uzorkovanje iz aposteriorne distribucije.

Kako bismo izveli aposteriornu distribuciju za $\theta = (\beta, \sigma)$, koristimo se Bayesovim teoremom:

$$f(\beta, \sigma^2 | y) = \frac{f(y | \beta, \sigma^2) f(\beta, \sigma^2)}{f(y)} \propto f(y | \beta, \sigma^2) f(\beta, \sigma^2) = L(\beta, \sigma^2 | y) f(\beta) f(\sigma^2). \tag{18}$$

gdje je $L(\beta, \sigma^2 | y)$ standardna Gaussova funkcija vjerodostojnosti:

$$L(\beta, \sigma^2 | y) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-(y - X\beta)^\tau (y - X\beta)/2\sigma^2\}. \tag{19}$$

Sada se pomoću izraza (16), (17), (18) i (19) mogu izvesti sljedeće aposteriorne distribucije:

$$\beta | (\sigma^2, y) \sim N_T(\mu_1, \Sigma_1) \tag{20}$$

3 Za slučajnu varijablu X kaže se da je inverzno-gama distribuirana s parametrima $X \sim IG(a,b)$ i ako je $\frac{1}{X}$ i gama distribuirana parametrima a i b.

$$\sigma^{-2}|(\beta,y) \sim (SS(\beta) + 2\lambda^{-1})^{-1}\chi^2_{n+2\nu}, \tag{21}$$

gdje je:

$$\mu_1 = \gamma\hat{\beta} + (1-\gamma)\mu_0 \tag{22}$$

$$\Sigma_1 = \sigma^2\gamma(X^\tau X)^{-1} \tag{23}$$

$$SS(\beta) = (y - X\beta)^\tau(y - X\beta) \tag{24}$$

$$\gamma = \sigma_0^2/(\sigma_0^2 + \sigma^2) \tag{25}$$

$$\hat{\beta} = (X^\tau X)^{-1}X^\tau y. \tag{26}$$

Dok je normalna aposteriorna distribucija za β standardan rezultat u literaturi, $\chi^2_{n+2\nu}$ aposteriorna distribucija za varijancu nije očit rezultat (vidi Dodatak).

Vektor $\mu_1 = \gamma\hat{\beta} + (1-\gamma)\mu_0$ konveksna je kombinacija apriornog parametra μ_0 i OLS procjenitelja $\hat{\beta}$. Koliki se ponder dodjeljuje pojedinačnom vektoru μ_0 ili $\hat{\beta}$ ovisi o nesigurnosti u vezi s apriornom distribucijom koja se odražava u $\gamma = \sigma_0^2/(\sigma_0^2 + \sigma^2)$. Na primjer, ako postoji visoka razina sigurnosti o apriornoj distribuciji, varijanca apriorne distribucije bit će manja, odnosno konstanta σ_0^2 bit će mala. Kao posljedica toga, γ će biti bliži nuli, a konveksna kombinacija μ_1 stoga je bliža μ_0.

U slučaju kad OLS procjenitelj zadovoljava sva ograničenja, jednak je ICLS procjenitelju. Štoviše, ako je određeno da je apriorni parametar μ_0 ICLS, odnosno sada OLS, slijedi da je $\mu_1 = \gamma\hat{\beta} + (1-\gamma)\mu_0 = \hat{\beta}$ a aposteriorna distribucija dobiva oblik $\beta|(\sigma^2,y) \sim N_T(\hat{\beta},\Sigma_1)$. Kao posljedica toga, bajesovski procjenitelj jednim brojem bit će bliže OLS procjenitelju.

Kako bismo simulirali realizaciju aposteriorne distribucije, u svakom krugu algoritma trebaju se simulirati k-vektor iz odrezane višedimenzionalne distribucije i jedna vrijednost iz $\chi^2_{n+2\nu}$ distribucije. U tu svrhu pratimo proceduru iz prethodnog odjeljka. Neka je A regularna matrica takva da je $A(X^\tau X)^{-1}A^\tau = I_k$. Nadalje, transformirajmo vektor β pomoću prikladne linearne transformacije

$$\eta = A\beta.$$

Sada na osnovi propozicije 2 i prethodne rasprave slijedi da je

$$\eta|(\sigma^2,y) \sim N_S(A\mu_1, \sigma^2\gamma A(X^\tau X)^{-1}A^\tau) = N_S(A\mu_1, \sigma^2\gamma I_k)$$

gdje je

$$S = \{\eta \in \mathbb{R}^k : D\eta \leq b\}, \text{ gdje je } D = BA^{-1}. \tag{27}$$

Gibbsov *sampler* sada je definiran pomoću sljedećeg algoritma.
Algoritam (Gibbsov *sampler* za restringirani regresijski model)

1. Inicijalizirajmo lanac $\theta^{(0)} = (\eta_1^{(0)},\ldots,\eta_k^{(0)},\sigma^{2(0)})$. Pretpostavimo da je $\theta^{(t)}$ posljednja simulirana vrijednost.

2. Vrijednost $\theta^{(t+1)} = (\eta_1^{(t+1)},\ldots,\eta_k^{(t+1)},\sigma^{2(t+1)})$
   - simulirajmo $\eta_1^{(t+1)}$ from $f(\eta_1|\eta_2^{(t)},\eta_3^{(t)},\ldots,\eta_k^{(t)},\sigma^{2(t)},y)$,
   - simulirajmo $\eta_2^{(t+1)}$ from $f(\eta_2|\eta_1^{(t+1)},\eta_3^{(t)},\ldots,\eta_k^{(t)},\sigma^{2(t)},y)$

     $\vdots$
   - simulirajmo $\sigma^{2(t+1)}$ from $f(\sigma^{2(t)})|\eta_1^{(t+1)},\eta_2^{(t+1)},\ldots,\eta_k^{(t+1)},y)$.

3. Poništimo linearnu transformaciju: $\beta^{(t+2)} = A^{-1}\eta^{(t+1)}$
4. Ponovimo postupak za $\theta^{(t+2)}$, $\theta^{(t+3)}$,...

Ukratko, detaljno smo opisali efikasnu metodu simuliranja iz apriorne distribucije parametara restringiranoga regresijskog modela. Simulacija je zasnovana na Gibbsovu *sampleru* koji generira Markovljev lanac koji u distribuciji konvergira prema aposteriornoj distribuciji, pa stoga simulirana sekvenca aproksimira slučajni uzorak iz aposteriorne distribucije.

4. Procjena pomoću mikropodataka u odnosu na procjenu pomoću makropodataka – simulirani primjer

Često su na raspolaganju samo agregatni podaci ili makropodaci o udjelima kredita u svakoj od rizičnih skupina. U tom se slučaju, zbog činjenice da procijenjene vjerojatnosti jesu migracijske vjerojatnosti pojedinačnih, nezavisnih mikrojedinica, uključenih u makrouzorak, postavlja pitanje pouzdanosti procijenjenih parametara. U praksi, kad su dostupni mikropodaci, vjerojatnosti prijelaza trebaju se uvijek procjenjivati pomoću maksimalne vjerodostojnosti. S druge strane, kad su na raspolaganju samo agregatni podaci, teško je odrediti u kojoj mjeri procjene dobivene regresijskim metodama aproksimiraju strukturne vjerojatnosti prijelaza na kojima se zasniva dinamika mikrojedinica. Kako bismo to ispitali, izvodi se jednostavan Monte Carlo eksperiment.

Pomoću proizvoljne matrice prijelaza:

$$P = \begin{pmatrix} 0,90 & 0,05 & 0,05 \\ 0,2 & 0,4 & 0,4 \\ 0,03 & 0,02 & 0,95 \end{pmatrix},$$

simulira se pet nezavisnih lanaca duljine n.[4] Nakon toga iz simuliranih se mikropodataka izvodi vremenska serija udjela u svakom stanju kako bi se usporedile procjene vjerojatnosti prijelaza zasnovane na makropodacima i mikropodacima. U Tablici 1. uspoređene su procjene dobivene različitim pristupima. Naposljetku, kako bi se ilustrirala osjetljivost procjena u odnosu na duljinu upotrijebljenih serija, eksperiment se izvodi i za $n = 100$ i za $n = 10.000$.

Prema očekivanjima, s obzirom na velik broj simuliranih realizacija lanca, tj. $n = 10000$, sve procjene vrlo dobro aproksimiraju prave vjerojatnosti prijelaza. Štoviše, sve metode procjene proizvode vjerojatnosti koje su jednake, do drugoga decimalnog mjesta, a regresijske metode zasnovane na makropodacima jednako su dobre kao procjenitelj maksimalne vjerodostojnosti zasnovan na mikropodacima. Međutim, za naše su potrebe zanimljiviji rezultati dobiveni za $n = 100$, što je broj opservacija koji se očekuje u praksi. Za razliku od slučaja velikog uzorka, većina procjena ovdje snažnije odstupa od pravih vjerojatnosti. Metodom maksimalne vjerodostojnosti, u kojoj se rabe mikropodaci, općenito se dobivaju najpreciznije procjene. Što se tiče procjenitelja regresije, OLS procjenitelj proizveo je jednu negativnu procjenu, pa se zato ista vjerojatnost procjenjuje kao nula pomoću ICLS procjenitelja. Bajesovski prosječni aposteriorni procjenitelj ovdje proizvodi pozitivnu vrijednost. S druge strane, OLS i ICLS procjenitelji proizvode bolje procjene za p_{22} u usporedbi s bajesovskim procjeniteljem. Osim toga, sve regresijske metode proizvode slične procjene za preostale parametre.

4 U slučaju kad OLS procjenitelj zadovoljava sva ograničenja nejednakosti, ekvivalentan je ICLS procjenitelju i veoma sličan bajesovskim procjenama. Svrha je ovoga primjera ilustrirati rezultate različitih procjenitelja u slučaju kad OLS procjenitelj nije u skladu s nametnutim ograničenjima. Stoga su prve tri simulirane serije koje rezultiraju OLS procjeniteljem sa zadovoljenim svim ograničenjima odbačene i ovdje se navodi samo četvrti pokušaj koji je proizveo *problematične* OLS procjene.

Tablica 1. Vjerojatnosti prijelaza

		N = 10000			N = 100		
		j = 1	*j* = 2	*j* = 3	*j* = 1	*j* = 2	*j* = 3
MLE		0,90	0,05	0,05	0,90	0,06	0,04
OLS		0,90	0,05	0,05	0,87	0,05	0,08
ICLS	*i* = 1	0,90	0,05	0,05	0,87	0,05	0,08
Bayes		0,90	0,05	0,05	0,87	0,05	0,08
MLE		0,21	0,39	0,40	0,05	0,38	0,57
OLS		0,21	0,39	0,40	–0,10	0,39	0,71
ICLS	*i* = 2	0,21	0,39	0,40	0,00	0,34	0,66
Bayes		0,21	0,39	0,40	0,05	0,31	0,64
MLE		0,03	0,02	0,95	0,03	0,01	0,96
OLS		0,03	0,02	0,95	0,06	0,02	0,92
ICLS	*i* = 3	0,03	0,02	0,95	0,05	0,02	0,93
Bayes		0,03	0,02	0,95	0,05	0,02	0,93

Kako bismo prosudili sposobnost pojedinačnih procjenitelja da aproksimiraju pravu tranzicijsku matricu, treba primijeniti odgovarajuću mjeru udaljenosti između matrica. U tu se svrhu može upotrijebiti standardna matrična 2-norma. Prema tom kriteriju udaljenosti, procjenitelj maksimalne vjerojatnosti matrice p najbliži je pravoj matrici, s procijenjenom udaljenošću od 0,23. Nakon toga slijedi bajesovski procjenitelj s udaljenošću od 0,3, a ICLS i OLS procjenitelji malo su lošiji, s udaljenostima od 0,34 i 0,44.

5. Rezultati

5.1. Vjerojatnosti prijelaza za kreditni portfelj hrvatskih banaka

U ovom se odjeljku prikazuju bajesovske procjene matrice prijelaza na kojoj se zasniva dinamika kredita danih domaćim poduzećima u razdoblju od drugog tromjesečja 1999. do drugog tromjesečja 2011. Svojstva bajesovskih procjena na kraju se uspoređuju s onima konvencionalnijih procjena dobivenih pomoću OLS i ICLS procjenitelja. Dostupni su samo podaci o agregatnim veličinama u svakoj rizičnoj skupini. Tromjesečne serije agregatnih kredita svrstane su u četiri široke kategorije rizika, ovisno o mogućnosti naplate, odnosno o očekivanim budućim novčanim tokovima (HNB, 2009.; Cerovac i Ivičić, 2009.):

A – plasmani za koje nisu uočeni dokazi o umanjenju vrijednosti

B – plasmani za koje su uočeni dokazi o umanjenju vrijednosti, tj. djelomično naplativi plasmani

C – plasmani za koje su uočeni dokazi o umanjenju vrijednosti, koja je jednaka njihovoj knjigovodstvenoj vrijednosti, odnosno potpuno nenaplativi plasmani

D – kumulativni otpisi kredita. Ovo je definirano kao apsorpcijska rizična kategorija[5].

Dužnikovo pravodobno ispunjavanje obveza prema kreditnoj instituciji važan je kriterij integriran u navedenu klasifikacijsku shemu koji podrazumijeva spuštanje iz kategorije A u nižu kategoriju B ako dužnik ima

5 Analiza je provedena i za lanac s tri stanja, s tim da su izostavljena apsorpcijska stanja. Zbog niskog udjela kredita u kategoriji *otpisa* u našem uzorku, rezultati se u tom slučaju nisu znatnije promijenili. Drugo se ispitivanje robustnosti odnosi na stacionarnost osnovnoga Markovljeva lanca. Metodologija klasifikacije znatno se promijenila u 2003. Zbog toga je analiza također provedena a da se nije uzelo u obzir razdoblje prije prekida. Glavni rezultati ponovo su se pokazali robustnima u odnosu na promjenu uzorka.

nepodmirene obveze duže od 90 dana do 365 dana, a iz kategorije B u kategoriju C ako dužnik ima nepodmirene obveze duže od 365 dana. Za detaljniji prikaz klasifikacije kredita vidi HNB (2009.).

Migracije između klasa A, B, C i D modelirane su kao staza Markovljeva lanca s diskretim vremenom i četiri stanja, pomoću podataka o udjelima kredita u svakom stanju, kako je prikazano u prethodnom odjeljku. Vjerojatnosti prijelaza u lancu, odnosno migracijske vjerojatnosti pojedinih kredita procjenjuju se pomoću realizacija koje proizvodi Gibbsov *sampler* u okviru restringiranoga regresijskog modela.

Da bi se *sampler* inicijalizirao, treba specificirati parametre apriorne distribucije i početnu vrijednost lanca. Kao i Rodriguez-Yam *et al.* (2004.), definiram apriorni parametar μ_0 kao ICSL procjenitelj, dobiven kvadratičnim programiranjem.[6] Nadalje, σ_0 treba biti dovoljno velika, ovdje je 100, kako bi odražavala značajnu nesigurnost u vezi s apriornom distribucijom koeficijenata. Naposljetku, parametri apriorne distribucije za varijancu šuma trebali bi također odražavati nedostatak prethodnih informacija i postavljeni su na[7] $\nu = \lambda = 0{,}01$. Nakon određivanja apriornih distribucija, generira se 10.000 realizacija iz aposteriornih uvjetnih distribucija za šest regresijskih koeficijenata i varijancu šuma. Realizacije za preostale koeficijente izračunavaju se izravno iz jednakosti (6). Bajesovski procjenitelj koji se ovdje rabi standardan je aposteriorni procjenitelj, odnosno prosjek uzorka realizacija aposteriorne distribucije.

5.2. Rezultati procjene

U Tablici 2. uspoređuju se procjene vjerojatnosti prijelaza za kredite koje hrvatske banke daju sektoru poduzeća, dobivene pomoću OLS procjenitelja (odnosno, zanemarivanjem ograničenja), ICLS procjenitelja (dobivenoga kvadratičnim programiranjem) i bajesovskog procjenitelja.

Prema očekivanjima, procjene metodom OLS ne zadovoljavaju ograničenja nejednakosti i neke su procjene veće od jedinice ili manje od nule.[8] S druge strane, po konstrukciji, i ICLS i bajesovski procjenitelj zadovoljavaju sva ograničenja. Međutim, među njima postoje neke razlike. Primjerice, ICLS procjenitelj često daje procjene koje su blizu granice danih skupova. Procijenjene vjerojatnosti tako otkrivaju nepoželjnu osobinu procjenitelja ICLS, a ta je da često daje rubna rješenja koja ne pružaju dovoljno informacija. Osim toga,

Tablica 2. Vjerojatnosti prijelaza u jednom tromjesečju, OLS, ICLS i bajesovske procjene (u postocima)

		j = *A*	*j* = *B*	*j* = *C*	*j* = *D*
OLS		100,04	0,05	–0,08	–0,01
ICLS	*i* = *A*	99,74	0,23	0,03	0
Bayes		**99,45**	**0,28**	**0,19**	**0,08**
OLS		–16,04	112,42	5,13	–1,51
ICLS	*i* = *B*	0,00	96,55	3,45	0
Bayes		**2,87**	**91,7**	**4,43**	**1**
OLS		34,18	–26,41	88,94	3,29
ICLS	*i* = *C*	10,69	0,00	89,31	0
Bayes		**9,74**	**6,05**	**82,16**	**2,05**
OLS		0	0	0	100
ICLS	*i* = *D*	0	0	0	100
Bayes		**0**	**0**	**0**	**100**

6 Vidi Byrd *et al.* (1995.) te Coleman i Li (1996.).

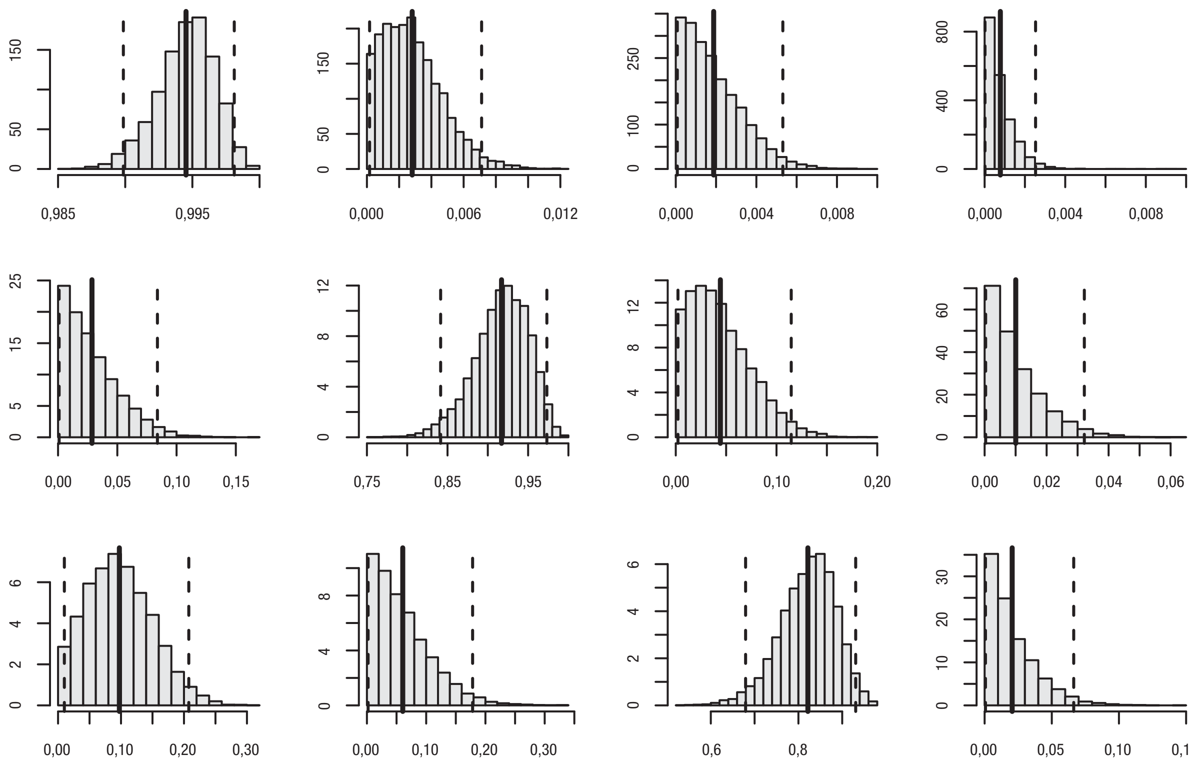
7 Može se pokazati da inverzna gama apriorna distribucija, $IG(\varepsilon,\varepsilon)$ za varijancu σ^2 konvergira prema uniformnoj apriornoj distribuciji za log $\log\sigma^2$ kad $\varepsilon \to 0$. Stoga postavljanje ε na malu vrijednost pridodaje sličnu gustoću svim potencijalnim vrijednostima apriorne distribucije, što odražava nesigurnost apriorne distribucije u vezi s varijancom.

nepostojanje relevantne asimptotske teorije za procjenitelj ICLS otežava ocjenu vjerodostojnosti parametara.

Nasuprot tome, bajesovski su procjenitelji prihvatljivi i lako ih je tumačiti. Krediti u kategoriji A najčešće ostaju u istoj rizičnoj skupini i u sljedećem tromjesečju, s visokom vjerojatnošću od 0,99. Dinamika kredita iz skupina B i C u posljednjih nekoliko godina nije pokazala takvu postojanost. Primjerice, krediti u kategoriji B nastavljaju se spuštati u niži rang C s vjerojatnošću manjom od 5%. Vjerojatnosti za premještanje u viši rang prilično su niske, manje od 3%. S druge strane, ICLS procjenitelj predviđa nultu vjerojatnost da će se kredit B u potpunosti oporaviti i ponovo postati A. Krediti u najnižoj skupini C otpisuju se s vjerojatnošću od 2%, a procijenjene vjerojatnosti prijelaza upućuju na to da rizična skupina C nije konačno (apsorpcijsko) stanje tipičnoga kredita. Ponovo napominjemo da ICLS procjenitelj predviđa nultu vjerojatnost da će kredit C biti otpisan.

Slika 1. prikazuje procijenjene gustoće za uvjetne aposteriorne distribucije regresijskih koeficijenata zajedno s bajesovskim intervalima vjerodostojnosti od 95%. U klasičnom su pristupu granice intervala pouzdanosti, a ne parametri, slučajne veličine. Zbog toga se ne može zaključiti da postoji, primjerice, vjerojatnost od 0,95 da parametar upada u interval pouzdanosti od 0,95. Ispravan je zaključak da interval pouzdanosti uključuje danu vrijednost s vjerojatnošću od 0,95, prije nego što se uopće vide podaci. Nasuprot tome, bajesovski intervali vjerodostojnosti omogućuju konstruiranje takvih intervala da je vjerojatnost da će (slučajni) parametar

Slika 1. Aposteriorni histogrami uvjetnih distribucija za regresijske koeficijente. Grafikon u redu i i stupcu j znači gustoću od $p_{ij} = P(y_t = j \mid y_{t-1} = i)$. Vertikalne crte jesu procjene 0,025, i 0,975 kvantila i srednju vrijednost aposteriorne distribucije.



Izvor: Izračun autora

8 Reci matrice prijelaza po definiciji se zbrajaju u jedinicu.

upasti u ovaj interval u 0,95. Procijenjeni histogrami ilustriraju učinak odsijecanja, a neke od uvjetnih distribucija karakteriziraju izrazito asimetrične empirijske distribucije.

Da bi se ispitala brzina konvergencije našeg *samplera* prema stacionarnoj distribuciji, korisno je također usporediti pomične prosjeke realizacija aposteriorne distribucije za vjerojatnosti prijelaza sa srednjim vrijednostima izračunatima u cijelom uzorku (vidi Dodatak).

6. Zaključak

Ovaj se rad bavi Gibbsovim *samplerom* za parametre restringiranoga regresijskog modela koji predlažu Rodriguez-Yam *et al.* (2004.). Metodologija je primijenjena na problem procjene vjerojatnosti migracije između različitih rizičnih skupina za kredite u portfelju hrvatskih banaka. Analiza se provodi za kredite sektoru poduzeća. Pretpostavlja se da svaka jedinica u portfelju prati diskretan Markovljev lanac. U tom se slučaju migracijske vjerojatnosti mogu procijeniti u okviru restringiranoga regresijskog modela, iako su dostupni samo podaci o udjelima kredita u svakoj rizičnoj skupini. Najvažniji nalazi ovoga rada mogu se sažeto prikazati kako slijedi.

Bajesovske su procjene zasnovane na Gibbsovu *sampleru* prihvatljive i jednostavno ih je tumačiti. Najkvalitetniji krediti uglavnom ostaju u istoj klasi u idućem tromjesečju, s vjerojatnošću od 0,99. Drugim riječima, rizična skupina A bila je veoma bliska apsorpcijskom stanju za tipičan kredit u Hrvatskoj u proteklom desetljeću. Nasuprot tome, krediti u kategorijama B i C u posljednjih deset godina nisu pokazali takvu perzistentnost. Primjerice, krediti iz skupine B prelaze u nižu kategoriju C s vjerojatnošću manjom od 5%. Vjerojatnosti prelaska u višu kategoriju prilično su niske, manje od 3%. Krediti s najnižim rejtingom C otpisuju se s vjerojatnošću od 2%, a procijenjene vjerojatnosti prijelaza upućuju na to da rizična skupina C nije konačno (apsorpcijsko) stanje tipičnoga kredita.

Za razliku od bajesovskih procjena, čini se da su standardno korištene procjene najmanjih kvadrata problematičnije. Procjene metodom OLS ne zadovoljavaju ograničenja nejednakosti i stoga se ne mogu tumačiti kao vjerojatnosti prijelaza. S druge strane, ICLS procjene svojom konstrukcijom zadovoljavaju sva ograničenja, ali sklone su tome da proizvode granična rješenja zbog kojih ih je teško tumačiti i mogu navesti na pomalo ekstremne zaključke. Primjerice, ograničeni procjenitelj predviđa nultu vjerojatnost da će se kredit B u potpunosti oporaviti i postati kredit A. Slično tome, ograničeni procjenitelj predviđa nultu vjerojatnost da će kredit C biti otpisan. Premda mikropodaci o pojedinačnim kreditima nisu dostupni, pa se ove tranzicije ne mogu izravno promatrati, takvi su slučajevi zabilježeni u prethodnom desetljeću.

Ova se analiza koristi nadzornim, agregatnim podacima koji se odnose samo na udjele kredita u svakoj skupini, pa je zato teško procijeniti u kojoj mjeri dobivene procjene aproksimiraju strukturne vjerojatnosti prijelaza. Kako bi se ovo detaljnije ispitalo, razmatra se ilustrativan primjer zasnovan na simuliranim putanjama mikrojedinica kredita. Ako vremenska dimenzija nije veoma velika ($n = 100$), regresijske procjene zasnovane na agregatnim podacima mogu znatno odstupati od pravih vjerojatnosti. Međutim, općenito i bajesovske procjene i procjene dobivene ICLS metodom daju dobru približnu vrijednost pravih prijelaza. Prema očekivanjima, procjene maksimalne vjerodostojnosti zasnovane na mikropodacima daju procjene s najvećom preciznošću i uvijek bi se trebale upotrebljavati kad su mikropodaci na raspolaganju. Kad se uzima matrična norma kao mjera za udaljenost, čini se da su bajesovske procjene malo bliže pravim parametrima nego procjene dobivene ICLS procjeniteljima. OLS, ponovimo, daje loše rezultate.

Osnovno obilježje primijenjenog *samplera* jest to da uzorak aposteriorne distribucije ima brzoopadajuću funkciju autokorelacije (Dodatak). Zbog toga prvi momenti aposteriorne distribucije vrlo brzo konvergiraju, a simulirane realizacije efikasno aproksimiraju aposteriornu distribuciju. Konvergencija se najčešće postiže nakon samo otprilike 1000 ponavljanja, što je znatno brže nego u alternativnim specifikacijama, kada je potrebno na desetke ili stotine tisuća ponavljanja. Nadalje, u korištenu specifikaciju jednostavno ugrađujemo velik brojpotencijalno kolinearnih restrikcija. To obilježje omogućuje procjenu vjerojatnosti prijelaza diskretnoga Markovljeva lanca i ne karakterizira alternativni Gibbsov *sampler* za ograničenu regresiju (Geweke, 1996.).

Ova je analiza dodatno potvrdila sva dobra obilježja *samplera* koji predlažu Rodriguez-Yam *et al.* (2004.). Dobiveni rezultati upućuju na to da bi se procjene koje se zasnivaju na ovom *sampleru* trebale uzimati u obzir pri procjeni Markovljevih matrica prijelaza u praksi.

Literatura

Byrd, R. H., Lu, P., Nocedal, J. i Zhu, C. (1995.): A limited memory algorithm for bound constrained optimization, SIAM J. Scientific Computing, 16, str. 1190 – 1208

Casella, G. i George E. I. (1992.): Explaining the Gibbs sampler, The American Statistician, 46 (3), str. 167 – 174

Cerovac, S. i Ivičić, L. (2009.): Credit Risk Assessment of Corporate Sector in Croatia, rad predstavljen na 15. Dubrovačkoj ekonomskoj konferenciji

Chamberlain, G. i Leamer, E. E. (1976.): Matrix weighted averages and posterior bounds, Journal of the Royal Statistical Society, serija B, str. 73 – 84.

Christodoulakis, G. A. (2006.): Markovian Credit Risk Transition Probabilities under Non-Negativity Constraints for the US Portfolio 1984-2004, Manchester Business School, Working Paper Series

Coleman, T. F. i Li, Y. (1996.): A Reflective Newton Method for Minimizing a Quadratic Function Subject to Bounds on Some of the Variables, SIAM Journal on Optimization, svezak 6, br. 4, str. 1040 – 1058.

Davis, W. W. (1978.): Bayesian analysis of the linear model subject to linear inequality constraints, Journal of the American Statistical Association, str. 573 – 579.

Geman, S. i Geman, D. (1984.): Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6, str. 721 – 741

Feller, W. (1971.): An Introduction to Probability Theory and Its Applications, svezak 2, 2. izdanje, Wiley.

Geweke, J. (1986.): Exact Inference in the Inequality Constrained Normal Linear Regression Model, Journal of Applied Econometrics, John Wiley & Sons, Ltd., svezak 1(2), str. 127 – 141

Geweke, J. (1991.): Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints and the Evaluation of Constraint Probabilities, rad je bio pripremljen za prezentaciju na Computing Science and Statistics: the Twenty-Third Symposium on the Interface, Seattle, 22. – 24. travnja 1991.

Geweke, J. (1996.): Bayesian Inference for Linear Models Subject to Linear Inequality Constraints, in Modeling and Prediction: Honouring Seymour Geisser, uredili Johnson, W. O., Lee, J. C. i Zellner, A., New York, Springer, str. 248 – 263

Geyer, C., internetska stranica: www.stat.umn.edu/126charlie/

Hajivassiliou, V. A. i McFadden, D. (1990.): The Method of Simulated Scores for the Estimation of LDV Models with an Application to External Debt Crisis, Cowles Foundation Discussion Papers 967, Cowles Foundation, Yale University

Hajivassiliou, V. A., McFadden, D. i Ruud, P. (1996.): Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results, Journal of Econometrics, Elsevier, Vol. 72(1-2), str. 85 – 134.

HNB (2009.): Odluka o klasifikaciji plasmana i izvanbilančnih obveza kreditnih institucija

Judge, G. G.i Takayama, T. (1966.): Inequality Restrictions In Regression Analysis, Journal of the American Statistical Association, 61, str. 166 – 181

Jones, M. T. (2005.): Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk, IMF Working Paper WP/05/219.

Kelton, C. M. L. (1981.): Estimation of Time-Independent Markov Processes with Aggregate Data: A Comparison of Techniques, Econometrica, svezak 49, br 2, str. 517 – 518.

Kloek, T. i Van Dijk, H. K. (1978.): Bayesian estimates of equation system parameters: an application of integration by Monte Carlo, Econometrica, 46, str. 1 – 19.

Lee, T. C., Judge, G. i Zellner, A. (1968.): Maximum Likelihood and Bayesian Estimation of Transition Probabilities, Journal of the American Statistical Association, svezak 63, br. 324, str. 1162 – 1179.

Lee, T. C., Judge, G. G. i Takayama, T. (1965.): On estimating the transition probabilities of a Markov process, Journal of Farm Economics, svezak 47, br. 3, str. 742 – 762.

Lovell, M. C.i Prescott, E. (1970.): Multiple Regression with inequality constraints: Pretesting Bias, Hypothesis Testing, and Efficiency, Journal of the American Statistical Association, 65, str. 913 – 925.

MacRae, E. C. (1977.): Estimation of Time-Varying Markov Processes with Aggregate Data, Econometrica, svezak 45(1), str. 183 – 198.

Metropolis, N. i Ulam, S. (1949.): The Monte Carlo Method, Journal of the American Statistical Association, svezak 44, br 247 (Sep., 1949), str. 335 – 341

O'Hagan, A. (1973.): Bayes estimation of a convex quadrature, Biometrika, 60, str. 565 – 571

Robert, C. P. i Casella, G. (2004.): Monte Carlo Statistical Methods, Springer-Verlag, New York.

Rodriguez-Yam, G., Davis, R. i Scharf, L. (2004.): Efficient Gibbs sampling of truncated multivariate Normal with application to constrained linear regression, Working paper, Colorado State University, Fort Collins, CO.

Dodatak

Aposteriorna distribucija za varijancu šuma

Kako bismo pokazali (21), napomenimo najprije da za apriornu distribuciju od σ^2 vrijedi:

$$f(\sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{\nu+1} e^{-\frac{1}{\lambda\sigma^2}}.$$

Iz Bayesova teorema slijedi:

$$f(\sigma^2 | \beta, y) \propto L(\sigma^2 | \beta, y) f(\sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2}SS(\beta)} \left(\frac{1}{\sigma^2}\right)^{\nu+1} e^{-\frac{1}{\lambda\sigma^2}} = \left(\frac{1}{\sigma^2}\right)^{\frac{n+2\nu}{2}+1} e^{-\frac{\lambda SS(\beta)+2}{2\lambda\sigma^2}},$$

a uspoređujući ovo s inverznom gama gustoćom dobivamo:

$$\sigma^2 \Big| (\beta, y) \sim IG\left(\frac{n+2\nu}{2}, \frac{2\lambda}{\lambda SS(\beta)+2}\right),$$

ili

$$\sigma^{-2} \Big| (\beta, y) \sim G\left(\frac{n+2\nu}{2}, \frac{2\lambda}{\lambda SS(\beta)+2}\right).$$

Nadalje, karakteristična funkcija gama gustoće dana je izrazom:

$$\varphi_{G\left(\frac{n+2\nu}{2}, \frac{2\lambda}{\lambda SS(\beta)+2}\right)}(t) = \left(1 - \frac{2\lambda}{\lambda SS(\beta)+2} it\right)^{-\frac{n+2\nu}{2}}.$$

Prisjetimo li se da je gama distribucija $G(n/2,2)$ po definiciji jednaka $\chi^2(n)$ distribuciji, odmah slijedi:

$$\varphi_{\frac{\chi^2(n+2\nu)}{SS(\beta)+2\lambda^{-1}}}(t) = \varphi_{\chi^2(n+2\nu)}\left(\frac{t}{SS(\beta)+2\lambda^{-1}}\right) = \left(1 - \frac{2}{SS(\beta)+2\lambda^{-1}} it\right)^{-\frac{n+2\nu}{2}} =$$
$$= \left(1 - \frac{2\lambda}{\lambda SS(\beta)+2} it\right)^{-\frac{n+2\nu}{2}} = \varphi_{G\left(\frac{n+2\nu}{2}, \frac{2\lambda}{\lambda SS(\beta)+2}\right)}(t).$$

Naposljetku, po Teoremu jedinstvenosti za karakteristične funkcije (Feller, 1971.), slučajne varijable $\frac{\chi^2(n+2\nu)}{SS(\beta)+2\lambda^{-1}}$ i $G\left(\frac{n+2\nu}{2}, \frac{2\lambda}{\lambda SS(\beta)+2}\right)$ imaju jednaku distribuciju i (21) vrijedi. Na sličan se način može prikazati (20).

Svojstva odrezane normalne distribucije

Propozicija 1. *Neka je $X \sim N_R(\mu, \Sigma)$ gdje je $R \in \mathbb{R}^k$ pozitivne Lebersqueove mjere i Σ pozitivno definitna matrica. Neka je $Y := AX$, gdje je A $r \times k$ matrica uz $r \leq k$. Tada je $Y{:}N_T(A\mu, A\Sigma A^\tau)$, gdje je $T := \{Ax : x \in R\}$.*

Propozicija 2. *Particionirajmo X, μ, Σ ina sljedeći način*

$$X = \begin{pmatrix} X_1 \\ X_k \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_k \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_1 \\ \Sigma_1^\tau & \sigma_{kk} \end{pmatrix}$$

pri čemu je X_k jednodimenzionalna komponenta od X. Tada jednodimenzionalne uvjetne varijable imaju sljedeće distribucije:

$$X_k | (X_1 = x_1) \sim N_{R_k}(\mu_k^*, \sigma_{kk}^*),$$

gdje je

$$\mu_k^* = \mu_k + \Sigma_1^\tau \Sigma_{11}^{-1}(x_1 - \mu_1)$$

$$\sigma_{kk}^* = \sigma_{kk} - \Sigma_1^\tau \Sigma_{11}^{-1} \Sigma_1,$$

za novi ograničeni skup:

$$R_k := \{x_k \in R : (x_1, x_k) \in R\}.$$

Bajesovska statistika

Parametri modela obično se procjenjuju *metodama maksimalne vjerodostojnosti* i *metodom momenata*. U ovom klasičnom pristupu, parametar se smatra fiksnom (determinističkom), ali nepoznatom varijablom. Prethodna uvjerenja ili informacija o parametrima koji se procjenjuju nisu relevantni za proces procjene. Nasuprot tome, bajesovsko zaključivanje zasniva se na drugačijoj ideji. Ovdje se parametar smatra slučajnom varijablom s dvije relevantne pripadajuće distribucije. Prva, *apriorna distribucija*, odražava naše znanje o parametru čak i prije nego što se prikupe podaci, a druga, *aposteriorna distribucija,* uvjetovana je promatranim podacima. Apriorna i aposteriorna distribucija povezane su preko Bayesova teorema, kako slijedi:

$$f(\beta|\mathcal{D}) = \text{Bayesov teorem} = \frac{f(\mathcal{D}|\beta)f(\beta)}{f(\mathcal{D})} \propto f(\mathcal{D}|\beta)f(\beta), \quad 28)$$

gdje β označava (slučajan) parametar, $\mathcal{D}$ su podaci, $f(\beta)$ je apriorna distribucija, $f(\beta|\mathcal{D})$ je aposteriorna, a $f(\mathcal{D}|\beta) = L(\beta|\mathcal{D})$ je funkcija vjerodostojnosti modela. Za razliku od standardnih strategija procjene, bajesovska statistika o relevantnom parametru rabi i prethodno znanje (apriorna distribucija) i informacije o uzorku (funkcija vjerodostojnosti). Međutim, iako se mogućnost uključivanja prethodnih uvjerenja o parametru u proces procjenjivanja često smatra važnom prednošću u odnosu na standardne tehnike procjene, to je također najčešća meta kritike bajesovskih modela. Činjenica da pojedinci a priori posjeduju različita stajališta o relevantnim parametrima može dovesti do njihove znatno drugačije specifikacije aposteriorne distribucije, pri čemu se onda postavlja pitanje dosljednosti pristupa.

Bajesovski model često se definira pomoću aposteriorne distribucije koja pripada istoj familiji kao i apriorna distribucija. Dobro poznati primjer je *konjugirana normalna gama apriorna distribucija* u normalnom linearnom regresijskom modelu. Drugim riječima, ako se normalna apriorna distribucija pretpostavi za regresijske koeficijente, a gama apriorna za varijancu šuma, marginalne aposteriorne distribucije za koeficijente i varijancu također posjeduje normalnu i gama distribuciju. Ovo se odmah može vidjeti iz (1) i iz oblika funkcije vjerodostojnosti Gaussova linearnog modela. U tom slučaju trebaju nam samo teorijski momenti aposteriorne distribucije kako bismo u potpunosti odredili distribuciju parametra. Premda cijela aposteriorna distribucija sadržava informacije o parametru koji se procjenjuje, to je procjena jednim brojem koja je najčešće zanimljiva za aplikacije. Za procjenu jednim brojem obično se odabire jednostavan *procjenitelj očekivanja aposteriorne distribucije.*

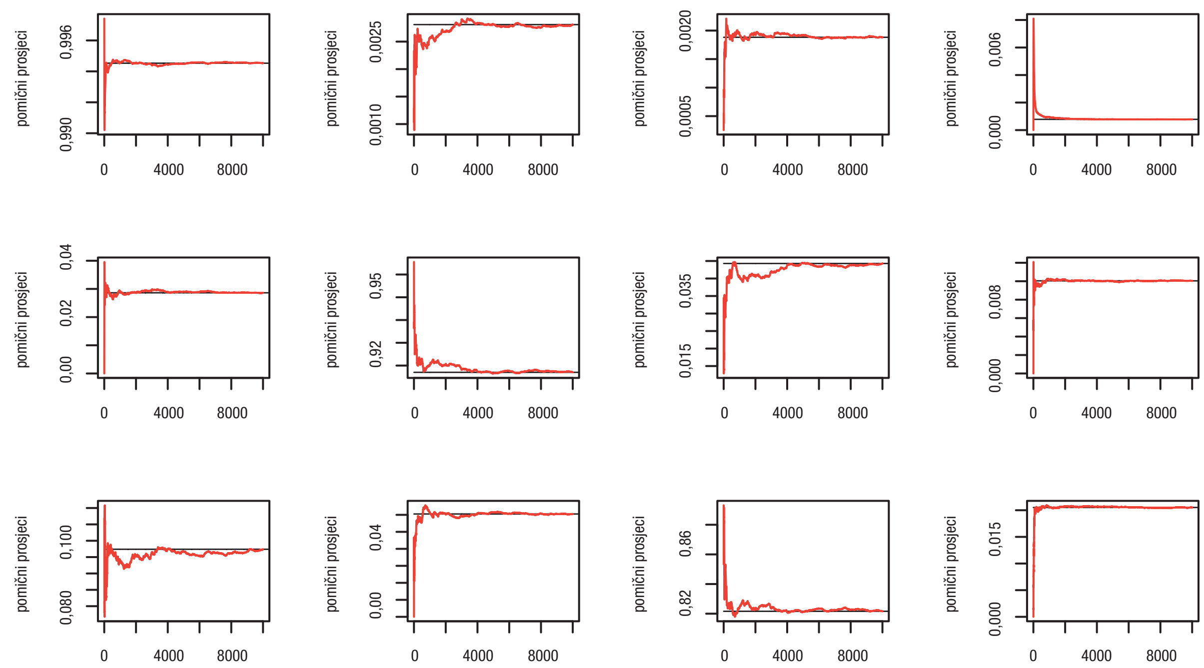
Međutim, u mnogim se slučajevima aposteriorna distribucija jednostavno izračunava samo do na konstantu i $f(\beta|\mathcal{D}) \propto f(\mathcal{D}|\beta)f(\beta)$ je ponekad najbolja dostupna informacija o formi aposteriorne distribucije. Razlog tome je najčešće činjenica da je normalizirajuću konstantu veoma teško dobiti analitički pomoću visokodimenzionalne integracije. Osim toga, aposteriorna distribucija ne treba pripadati nijednoj od standardnih distribucija vjerojatnosti i analiza aposteriorne distribucije postaje otežana. To je tradicionalno najvažniji nedostatak mnogih bajesovskih modela i njihovih aplikacija. MCMC metode pokazale su se veoma djelotvornima u rješavanju ovog problema korištenjem simulacija.

MCMC metode

MCMC metode zasnivaju se na činjenici da je katkad moguće simulirati slučajan uzorak iz distribucije koja može biti poznata samo do na normalizirajuću konstantu. Osnovna je ideja MCMC metoda simulirati putanju Markovljeva lanca kojemu je granična distribucija upravo tražena distribucija. U smislu izraza (1) dovoljno je simulirati lanac koji *konvergira* prema $f(\beta|\mathcal{D})$. Na taj način, pod relativno blagim uvjetima (detaljnije vidi Robert i Casella, 2004.), dovoljno duga simulirana staza Markovljeva lanca može se smatrati uzorkom iz te distribucije. Premda egzaktno zaključivanje o aposteriornoj distribuciji možda nije moguće, MCMC metodama često se može simulirati uzorak koji aproksimira uzorak iz tražene distribucije. Važno je napomenuti da je uzorak generiran nekom od MCMC metoda realizacija Markovljeva lanca i kao takav ne može se smatrati serijom nezavisnih realizacija iz granične distribucije. Ipak, *ergodični Markovljev lanac*, uz neke dodatne pretpostavke, nasljeđuje sve dobre osobine iz *Zakona velikih brojeva* (ZVB) i *Centralnoga graničnog teorema* (CGT) i stoga statistike od interesa dobro konvergiraju.

MCMC metode nastale su iz Metropolisova algoritma (Metropolis i Ulam, 1949.) kad je skupina fizičara pokušala izračunati komplicirane integrale pomoću simulacija. Nadovezujući se na taj algoritam, konstruiraju se dvije najvažnije MCMC metode, *Metropolis-Hastings algoritam* i, naposljetku, njegov poseban slučaj, Gibbsov *sampler* (vidi Geman i Geman, 1984.).

Slika 2. Pomični prosjeci uvjetnih distribucija za regresijske koeficijente. Za svaku seriju srednje vrijednosti izračunate u cijelom uzorku prikazane su punom linijom.



Izvor: Izračun autora

Dijagnostika konvergencije i uvjetne aposteriorne distribucije

Pouzdanost bajesovskih procjenitelja umnogome ovisi o dvama činiteljima. Kao prvo, disperzija aposteriorne distribucije važna je jer naznačuje preciznost procjena. Što je manja varijanca, to smo sigurniji u bajesovske procjene. Kao drugo, bajesovski se procjenitelji oslanjaju na Gibbsove realizacije pod pretpostavkom da je pripadajući lanac doista konvergirao prema aposteriornoj distribuciji. To je najlakše provjeriti grafom pomičnih prosjeka.

Kako bismo provjerili brzinu konvergencije našega *samplera* prema stacionarnoj distribuciji, na Slici 2. uspoređujemo pomične prosjeke realizacija aposteriornih distribucija za vjerojatnosti prijelaza sa srednjim vrijednostima izračunatima na cijelom uzorku. Za sve parametre pomični se prosjeci konvergiraju prema ukupnom prosjeku već nakon nekoliko stotina realizacija *samplera*. Vidi se kako prosječne vrijednosti uvelike ovise o početnom stanju lanca.[9] Stoga se često takozvani *burn-in* dio uzorka koji se sastoji od početnih realizacija izostavlja iz daljnjih analiza. Jasno, odbacivanje nekih ponavljanja na početku izvođenja algoritma samo je drugi način da se započne lanac, pomoću točke s većom vjerojatnošću pojavljivanja u aposteriornoj distribuciji. Međutim, odbačene vrijednosti također su mogući ishodi za aposteriornu distribuciju, ali uz vrlo male vjerojatnosti u ovako kratkom roku. Umjesto da jednostavno odbacimo početne opservacije lanca, u ovom se radu razmatraju nešto duže MCMC sekvencije i početne realizacije ne utječu na statistike uzorka iz aposteriorne distribucije. Inicijalne opservacije stoga se ne izostavljaju iz analize.[10]

9 Kako bismo provjerili osjetljivost statistika aposteriorne distribucije na razne specifikacije početnih stanja, *sampler* je pokrenut s različitih točaka realizacija aposteriorne distribucije. Statistike su konvergirale brzo za različite početne točke.

10 U zanimljivoj raspravi Charles Geyer (Geyerova internetska stranica) objašnjava zašto i kako *burn-in* nije nužan. *Burn-in je samo jedna metoda, i to ne osobito dobra, nalaženja dobrog polazišta. Naziv burn-in dolazi iz elektronike. Mnoge elektroničke komponente brzo propadaju. One koje ne propadaju su pouzdaniji podskup. Zato se burn-in obavlja u tvornici kako bi se najgori primjerci eliminirali. Markoljevi lanci ne funkcioniraju na isti način Neuspjeh (nekonvergencija) Markovljeva lanca razlikuje se od propadanja elektroničkih komponenti. Prvi se može ispraviti ako duže traje, ali pokvareni tranzistor ne može se popraviti. Zato je burn-in loš pojam u metodi MCMC, ali nije samo ta riječ pogrešna, nego je cijeli koncept neuvjerljiv.*

Do sada objavljena Istraživanja

Broj	Datum	Naslov	Autor(i)
I-1	studeni 1999.	Je li neslužbeno gospodarstvo izvor korupcije?	Michael Faulend i Vedran Šošić
I-2	ožujak 2000.	Visoka razina cijena u Hrvatskoj – neki uzroci i posljedice	Danijel Nestić
I-3	svibanj 2000.	Statističko evidentiranje pozicije putovanja – turizam u platnoj bilanci Republike Hrvatske	Davor Galinec
I-4	lipanj 2000.	Hrvatska u drugoj fazi tranzicije 1994. – 1999.	Velimir Šonje i Boris Vujčić
I-5	lipanj 2000.	Mjerenje sličnosti gospodarskih kretanja u Srednjoj Europi: povezanost poslovnih ciklusa Njemačke, Mađarske, Češke i Hrvatske	Velimir Šonje i Igeta Vrbanc
I-6	rujan 2000.	Tečaj i proizvodnja nakon velike ekonomske krize i tijekom tranzicijskog razdoblja u Srednjoj Europi	Velimir Šonje
I-7	rujan 2000.	OLS model fizičkih pokazatelja inozemnoga turističkog prometa na hrvatskom tržištu	Tihomir Stučka
I-8	prosinac 2000.	Je li Srednja Europa optimalno valutno područje?	Alen Belullo, Velimir Šonje i Igeta Vrbanc
I-9	svibanj 2001.	Nelikvidnost: razotkrivanje tajne	Velimir Šonje, Michael Faulend i Vedran Šošić
I-10	rujan 2001.	Analiza pristupa Republike Hrvatske Svjetskoj trgovinskoj organizaciji upotrebom računalnog modela opće ravnoteže	Jasminka Šohinger, Davor Galinec i Glenn W. Harrison
I-11	travanj 2002.	Usporedba dvaju ekonometrijskih modela (OLS i SUR) za prognoziranje dolazaka turista u Hrvatsku	Tihomir Stučka
I-12	veljača 2003.	Strane banke u Hrvatskoj: iz druge perspektive	Evan Kraft
I-13	veljača 2004.	Valutna kriza: teorija i praksa s primjenom na Hrvatsku	Ivo Krznar
I-14	lipanj 2004.	Privatizacija, ulazak stranih banaka i efikasnost banaka u Hrvatskoj: analiza stohastičke granice fleksibilne Fourierove funkcije troška	Evan Kraft, Richard Hofler i James Payne
I-15	rujan 2004.	Konvergencija razina cijena: Hrvatska, tranzicijske zemlje i EU	Danijel Nestić
I-16	rujan 2004.	Novi kompozitni indikatori za hrvatsko gospodarstvo: prilog razvoju domaćeg sustava cikličkih indikatora	Saša Cerovac
I-17	siječanj 2006.	Anketa pouzdanja potrošača u Hrvatskoj	Maja Bukovšak
I-18	listopad 2006.	Kratkoročno prognoziranje inflacije u Hrvatskoj korištenjem sezonskih ARIMA procesa	Andreja Pufnik i Davor Kunovac
I-19	svibanj 2007.	Kolika je konkurencija u hrvatskom bankarskom sektoru?	Evan Kraft
I-20	lipanj 2008.	Primjena hedonističke metode za izračunavanje indeksa cijena nekretnina u Hrvatskoj	Davor Kunovac, Enes Đozović, Gorana Lukinić, Andreja Pufnik
I-21	srpanj 2008.	Modeliranje gotovog novca izvan banaka u Hrvatskoj	Maroje Lang, Davor Kunovac, Silvio Basač, Željka Štaudinger
I-22	listopad 2008.	Međunarodni poslovni ciklusi u uvjetima nesavršenosti na tržištu dobara i faktora proizvodnje	Ivo Krznar
I-23	siječanj 2009.	Rizik bankovne zaraze u Hrvatskoj	Marko Krznar
I-24	kolovoz 2009.	Optimalne međunarodne pričuve HNB-a s endogenom vjerojatnošću krize	Ana Maria Čeh i Ivo Krznar
I-25	veljača 2010.	Utjecaj financijske krize i reakcija monetarne politike u Hrvatskoj	Nikola Bokan, Lovorka Grgurić, Ivo Krznar, Maroje Lang
I-26	veljača 2010.	Priljev kapitala i učinkovitost sterilizacije – ocjena koeficijenta sterilizacije i ofset koeficijenta	Igor Ljubaj, Ana Martinis, Marko Mrkalj
I-27	travanj 2010.	Postojanost navika i međunarodne korelacije	Alexandre Dmitriev i Ivo Krznar
I-28	studeni 2010.	Utjecaj vanjskih šokova na domaću inflaciju i BDP	Ivo Krznar i Davor Kunovac
I-29	prosinac 2010.	Dohodovna i cjenovna elastičnost hrvatske robne razmjene – analiza panel-podataka	Vida Bobić
I-30	siječanj 2011.	Model neravnoteže na tržištu kredita i razdoblje kreditnog loma	Ana Maria Čeh, Mirna Dumičić, Ivo Krznar
I-31	travanj 2011.	Analiza kretanja domaće stope inflacije i Phillipsova krivulja	Ivo Krznar
I-32	svibanj 2011.	Identifikacija razdoblja recesija i ekspanzija u Hrvatskoj	Ivo Krznar
I-33	listopad 2011.	Globalna kriza i kreditna euroizacija u Hrvatskoj	Tomislav Galac
I-34	studeni 2011.	Središnja banka kao krizni menadžer u Hrvatskoj – analiza hipotetičnih scenarija	Tomislav Galac
I-35	siječanj 2012.	Ocjena utjecaja monetarne politike na kredite stanovništvu i poduzećima: FAVEC pristup	Igor Ljubaj
I-36	ožujak 2012.	Jesu li neke banke blaže od drugih u primjeni pravila klasifikacije plasmana	Tomislav Ridzak

Upute autorima

Hrvatska narodna banka objavljuje u svojim povremenim publikacijama Istraživanja, Pregledi i Tehničke bilješke znanstvene i stručne radove zaposlenika Banke i vanjskih suradnika.

Prispjeli radovi podliježu postupku recenzije i klasifikacije koji provodi Komisija za klasifikaciju i vrednovanje radova. Autori se u roku od najviše dva mjeseca od primitka njihova rada obavještavaju o odluci o prihvaćanju ili odbijanju članka za objavljivanje.

Radovi se primaju i objavljuju na hrvatskom i/ili na engleskom jeziku.

Radovi predloženi za objavljivanje moraju ispunjavati sljedeće uvjete.

Tekstovi moraju biti dostavljeni elektroničkom poštom ili optičkim medijima (CD, DVD), a mediju treba priložiti i ispis na papiru. Zapis treba biti u formatu Microsoft Word.

Na prvoj stranici rada obvezno je navesti naslov rada, ime i prezime autora, akademske titule, naziv ustanove u kojoj je autor zaposlen, suradnike te potpunu adresu na koju će se autoru slati primjerci za korekturu.

Dodatne informacije, primjerice zahvale i priznanja, poželjno je uključiti u tekst na kraju uvodnog dijela.

Na drugoj stranici svaki rad mora sadržavati sažetak i ključne riječi. Sažetak mora biti jasan, deskriptivan, pisan u trećem licu i ne dulji od 250 riječi (najviše 1500 znakova). Ispod sažetka treba navesti do 5 ključnih pojmova.

Tekst treba biti otipkan s proredom, na stranici formata A4. Tekst se ne smije oblikovati, dopušteno je samo podebljavanje (bold) i kurziviranje (italic) dijelova teksta. Naslove je potrebno numerirati i odvojiti dvostrukim proredom od teksta, ali bez formatiranja.

Tablice, slike i grafikoni koji su sastavni dio rada, moraju biti pregledni, te moraju sadržavati broj, naslov, mjerne jedinice, legendu, izvor podataka te bilješke. Bilješke koje se odnose na tablice, slike ili grafikone treba obilježiti malim slovima (a, b, c…) i ispisati ih odmah ispod. Ako se posebno dostavljaju (tablice, slike i grafikoni), potrebno je označiti mjesta u tekstu gdje dolaze. Numeracija mora biti u skladu s njihovim slijedom u tekstu te se na njih treba referirati prema numeraciji. Ako su već umetnuti u tekst iz nekih drugih programa, onda je potrebno dostaviti i te datoteke u formatu Excel (grafikoni moraju imati pripadajuće serije podataka).

Ilustracije trebaju biti u standardnom formatu EPS ili TIFF s opisima u Helvetici (Arial, Swiss) veličine 8 točaka. Skenirane ilustracije trebaju biti rezolucije 300 dpi za sivu skalu ili ilustraciju u punoj boji i 600 dpi za lineart (nacrti, dijagrami, sheme).

Formule moraju biti napisane čitljivo. Indeksi i eksponenti moraju biti jasni. Značenja simbola moraju se objasniti odmah nakon jednadžbe u kojoj se prvi put upotrebljavaju. Jednadžbe na koje se autor poziva u tekstu potrebno je obilježiti serijskim brojevima u zagradi uz desnu marginu.

Bilješke na dnu stranice treba označiti arapskim brojkama podignutima iznad teksta. Trebaju biti što kraće i pisane slovima manjima od slova kojima je pisan tekst.

Popis literature dolazi na kraju rada, a u njega ulaze djela navedena u tekstu. Literatura treba biti navedena abecednim redom prezimena autora, a podaci o djelu moraju sadržavati i podatke o izdavaču, mjesto i godinu izdavanja.

Uredništvo zadržava pravo da autoru vrati na ponovni pregled prihvaćeni rad i ilustracije koje ne zadovoljavaju navedene upute.

Pozivamo zainteresirane autore koji žele objaviti svoje radove da ih pošalju na adresu Direkcije za izdavačku djelatnost, prema navedenim uputama.

Hrvatska narodna banka izdaje sljedeće publikacije:

Godišnje izvješće Hrvatske narodne banke

Redovita godišnja publikacija koja sadržava godišnji pregled novčanih i općih ekonomskih kretanja te pregled statistike.

Polugodišnje izvješće Hrvatske narodne banke

Redovita polugodišnja publikacija koja sadržava polugodišnji pregled novčanih i općih ekonomskih kretanja te pregled statistike.

Tromjesečno izvješće Hrvatske narodne banke

Redovita tromjesečna publikacija koja sadržava tromjesečni pregled novčanih i općih ekonomskih kretanja.

Bilten o bankama

Redovita publikacija koja sadržava pregled i podatke o bankama.

Bilten Hrvatske narodne banke

Redovita mjesečna publikacija koja sadržava mjesečni pregled novčanih i općih ekonomskih kretanja te pregled monetarne statistike.

Istraživanja Hrvatske narodne banke

Povremena publikacija u kojoj se objavljuju kraći znanstveni radovi zaposlenika Banke i vanjskih suradnika.

Pregledi Hrvatske narodne banke

Povremena publikacija u kojoj se objavljuju stručni radovi zaposlenika Banke i vanjskih suradnika.

Tehničke bilješke

Povremena publikacija u kojoj se objavljuju informativni radovi zaposlenika Banke i vanjskih suradnika.

Hrvatska narodna banka izdaje i druge publikacije: numizmatička izdanja, brošure, publikacije na drugim medijima (CD-ROM, DVD), knjige, monografije i radove od posebnog interesa za Banku, zbornike radova s konferencija kojih je organizator ili suorganizator Banka, edukativne materijale i druga slična izdanja.

ISSN 1332-1900 (tisak) • ISSN 1334-0077 (online)