

UNIVERZA V LJUBLJANI
EKONOMSKA FAKULTETA

DIPLOMSKO DELO

**NAPOVEDOVANJE STEČAJEV PODJETIJ Z LOGIT
MODELOM IN DISKRIMINANTNO ANALIZO**

Ljubljana, januar 2008

MATJAŽ ŠIRCA

IZJAVA

Študent **Matjaž Širca** izjavljam, da sem avtor tega diplomskega dela, ki sem ga napisal pod mentorstvom **doc. dr. Igorja Mastena**, in dovolim objavo diplomskega dela na fakultetnih spletnih straneh.

V Ljubljani, dne 30.01.2008

Podpis: _____

KAZALO

1	UVOD	1
2	PODJETJE, STEČAJ, VPLIV NA BANKO IN BANČNI SISTEM	2
2.1	Stečaj podjetja	2
2.2	Vpliv stečaja podjetja na banko	4
2.3	Finančni kazalniki kot sredstvo manjšanja negotovosti za banko.....	6
3	DISKRIMINANTNA ANALIZA	7
3.1	Metodološka logika	8
3.2	Predpostavka 1: Normalno porazdeljevanje neodvisnih spremenljivk v skupinah...	10
3.3	Predpostavka 2: Enakost varianc neodvisnih spremenljivk v skupinah.....	11
4	LOGISTIČNA REGRESIJA	12
4.1	Metodološka logika	13
5	OPREDELITEV FINANČNIH KAZALNIKOV V ANALIZI	14
5.1	Lastnosti podatkov in primernost za uporabo diskriminantne analize	15
5.2	Podatkovna zasnova in prilagoditve začetne baze	15
5.2.1	Uporabljene spremenljivke v analizi	16
5.3	Metoda izbora najbolj primernih kazalnikov za analizo	17
5.3.1	1. korak-Faktorska analiza	17
5.3.2	2. korak-Logistična regresija in Analiza glavnih komponent	18
5.4	Končni izbor najboljše kombinacije spremenljivk.....	20
6	REZULTATI NA PODLAGI DISKRIMINANTNE ANALIZE	23
6.1	Ocena diskriminantnega modela in razvrščanje podjetij.....	23
6.2	Metode napovedovanja z ocenjenim modelom.....	25
6.3	Napovedovanje z ocenjenim modelom	26
6.3.1	Analiza območja prekrivanja	28
6.4	Napovedovanje s podatki dve leti pred stečajem	30
7	REZULTATI NA PODLAGI LOGISTIČNE REGRESIJE	31
7.1	Ocena logit modela in razvrščanje podjetij	32
7.2	Napovedovanje verjetnosti stečaja z ocenjenim modelom	33
7.2.1	Analiza območja prekrivanja	35
7.3	Napovedovanje s podatki dve leti pred stečajem	35

8	SKLEPNI DEL-KLJUČNE UGOTOVITVE TER PRIMERJAVE OBEH ORODIJ	36
8.1	Diskriminantna analiza.....	36
8.2	Logistična regresija	39
8.3	Primerjava obeh orodij ter končni komentarji.....	41
9	POMEMBNO, VENDAR NEVKLJUČENO V TEJ NALOZI	43
LITERATURA		45
VIRI		46

KAZALO TABEL

Tabela 1: Rezultati testov izbora optimalne kombinacije spremenljivk	21
Tabela 2: Razvrščanje podjetij pri različnih strukturah vzorca.....	24
Tabela 3: Opisne statistike za izbrane spremenljivke ter izbrani vzorec	25
Tabela 4: Napovedovanje razvrstitve z ocenjenim modelom	28
Tabela 5: Preizkus značilnosti napovedi z ocenjenim modelom.....	28
Tabela 6: Intervali zaupanja pri napačno razvrščenih podjetjih.....	29
Tabela 7: Ocenjeni logit model logistične regresije.....	33
Tabela 8: Kvaliteta razvrstitve podjetij v odvisnosti od točke reza	33
Tabela 9: Napovedovanje razvrstitve z ocenjenim modelom	34
Tabela 10: Preizkus značilnosti napovedi z ocenjenim modelom.....	34
Tabela 11: Intervali zaupanja pri napačno razvrščenih podjetjih.....	35
Tabela 12: Napovedovanje razvrstitve z ocenjenim modelom	36

KAZALO SLIK

Slika 1: Apriorna verjetnost kot dejavnik kvalitete razvrstitve pri diskriminantni analizi	37
Slika 2: Točka reza kot dejavnik kvalitete razvrstitve pri logistični regresiji.....	40

1 UVOD

Diskriminantna analiza in logistična regresija sta dve orodji multivariatne analize. Sodita v skupino t.i. metod odvisnosti, kjer poskušamo ugotoviti, ali posamezna oz. kombinacija neodvisnih (pojasnjevalnih) spremenljivk vpliva na eno oz. kombinacijo odvisnih (pojasnjenih) spremenljivk. Namen obeh orodij je razvrščanje statističnih enot na podlagi lastnosti (zakonitosti), ki določeno populacijo razdelijo v različne skupine oz. kategorije. Navadno nimamo podatkov za vse enote populacije, zato se moramo zadovoljiti z določenim vzorcem in na podlagi tega oceniti parametre modela. Poskušamo ugotoviti, ali je vpliv neodvisnih spremenljivk na odvisno dovolj velik oz. statistično značilen, da bi lahko na podlagi vzorčnih ocen parametrov sklepali o zakonitostih celotne populacije. Poskušamo torej oceniti takšen model, pri katerem bo pravilnost razvrstitve enot v skupine čim višja, tako na izbranem vzorcu kot na celotni populaciji. Pri tem predstavlja pomemben del naloge izbira vzorca na katerem ocenjujemo model in izbira neodvisnih spremenljivk, ki jih bomo uporabili pri ocenjevanju modela. Diskriminantna analiza in logistična regresija sicer sodita v enako skupino – metod odvisnosti, vendar sta metodološko gledano povsem različni.

Diskriminantna analiza je podvržena bistveno več predpostavkam kot logistična regresija. Kot največkrat kršene predpostavke diskriminantne analize se navajajo normalna porazdelitev ter enakost varianc neodvisnih spremenljivk v skupinah. Ker imamo za analizo velikokrat na razpolago le podatke za vzorec in ne za celotno populacijo, že v samem začetku naletimo na težavo pri uporabi tega orodja. To je pogosto pri majhnih vzorcih, kjer so porazdelitve često nenormalne in pri finančnih kazalnikih kot v primeru te naloge.

Na drugi strani logistična regresija ne predpostavlja normalnih porazdelitev in enakih varianc neodvisnih spremenljivk v skupinah. Zato se ob kršitvah teh dveh predpostavk zdi boljša izbira kot diskriminantna analiza. Na podlagi rezultatov analiz opravljenih v tej nalogi bomo videli, da temu ni vedno tako.

Namen naloge je primerjati diskriminantno analizo in logistično regresijo ter na podlagi rezultatov podati strnjene sklepe o tem, kdaj je bolje uporabiti katero izmed njiju. Hkrati je namen naloge dati analitiku nekaj napotkov pri soočanju s težavami, ki se pojavljajo pri analizah, ko imamo opravka z dvema skupinama enot, dihotomno odvisno spremenljivko na eni strani ter s kombinacijo numeričnih zveznih neodvisnih spremenljivk na drugi strani. S podatki podatkovne baze Ajpes za bankrotirana in nebankrotirana slovenska podjetja v obdobju 1994 – 2004, poskušam na podlagi vzorčnih ocen parametrov podati napoved razvrstitve oz. verjetnosti stečaja z uporabo kombinacije finančnih kazalnikov. Z vsakim izmed orodij poskušam napovedati pripadnost podjetij eni izmed dveh skupin ter na podlagi napak napovedi pri določenih okoliščinah podati sklepe o smiselnosti uporabe vsakega orodja.

Literatura, ki sem jo prebral za namen pisanja te naloge, se ne ukvarja posebej z neposredno primerjavo teh dveh orodij ter s presojanjem smiselnosti uporabe vsakega izmed njiju.

Večinoma je literatura omejena zgolj na uporabo enega izmed teh dveh orodij, nameni analiz pa so povečini bistveno drugačni – niso primerjalne narave. S to nalogo sem hkrati hotel pokazati, da ni splošnega pravila glede uporabe diskriminantne analize oz. logistične regresije. Ob določenih okoliščinah sta orodji, kljub kršenju osnovnih predpostavk, lahko tudi enakovredni in je zato z vidika kvalitete rezultatov vseeno katero orodje uporabimo. Pri tem je potrebno upoštevati namen analize. Spet ob drugačnih pogojih ima eno orodje prednosti pred drugim. Vsekakor je veliko odvisno od subjektivne presoje ter predvsem od poznavanja tako metodologij obeh orodij, kot tudi porazdelitvenih lastnosti neodvisnih spremenljivk podatkovne baze, s katero imamo opravka. Struktura naloge sledi strukturi, kot sem jo zasledil v strokovni literaturi.

Naloga je v grobem razdeljena v dva večja vsebinska sklopa. Prvi je posvečen diskriminantni analizi, drugi pa logistični regresiji. V vsakem sklopu sem najprej z izbranim vzorcem ocenil model, nato pa na podlagi vzorčnih ocen parametrov podal napoved na ostalem delu populacije. Poleg tega sem se v vsakem sklopu nekoliko posvetil tudi pojasnjevanju sistematičnih napak ocenjenega modela.

Zaradi omejenosti diskriminantne analize z omenjenima predpostavkama, sem se odločil za izbran potek analize v tej nalogi, izbiro vzorca ter neodvisnih spremenljivk. Pri tem sem kot referenčno orodje uporabljal diskriminantno analizo. Na podlagi teh rezultatov sem presojal o nadaljnem poteku analize. V drugem sklopu sem opravil analizo z logistično regresijo. Pri tem se nisem spuščal v novo iskanje optimalne kombinacije neodvisnih spremenljivk, pač pa sem uporabil iste spremenljivke in vzorce kot pri diskriminantni analizi. S tem sem hotel omogočiti primerljivost dobljenih rezultatov, na podlagi katerih sem tudi napravil končne sklepe.

2 PODJETJE, STEČAJ, VPLIV NA BANKO IN BANČNI SISTEM

2.1 Stečaj podjetja

Podjetje je lahko pravna ali fizična oseba, ki se ukvarja s pridobitno ali nepridobitno dejavnostjo. V tej nalogi bodo predmet analize podjetja, ki imajo profitni motiv. Ločimo jih glede na vrsto dejavnosti, ki jo opravljajo. V kontekstu te naloge, bi omenil razcepitev dejavnosti v smislu toka sredstev, ki je potreben za to, da neko podjetje lahko kontinuirano / normalno posluje (ang. *going concern*). V najbolj splošnem smislu s tega vidika ločimo podjetja na finančna in nefinančna. V današnjih, večinoma kapitalističnih sistemih, so raznovrstna podjetja povečini medsebojno povezana / odvisna. Da lahko podjetje nemoteno opravlja svojo pridobitno dejavnost potrebuje sredstva. Sredstva za potrebe poslovanja se lahko zagotavljajo iz denarnih tokov iz poslovanja, včasih pa ti denarni tokovi niso dovolj visoki in podjetje lahko v tem primeru poseže po posojilih, ki jih lahko dobi tudi pri bankah.

Kot sem že omenil opravljajo podjetja pridobitno dejavnost. Pridobitna je zato, ker je cilj poslovanja podjetja profit. To pomeni, da mora potem, ko od vseh prihodkov odštejemo vse odhodke ter davek na dobiček, ob uspešnem poslovanju podjetju ostati pozitivna številka, ki jo imenujemo (čisti) dobiček. Seveda je lahko ta številka tudi negativna. V tem primeru govorimo o izgubi. Dobiček se lahko porabi za različne namene. Pri mladih podjetjih gre dobiček praviloma za nove investicije za rast podjetja. Pri starejših podjetjih gre za dividende, če gre za delniške družbe, ali pa se dobički zadržijo in prenesejo v naslednja poslovna leta kot rezerva za pokrivanje potencialnih prihodnjih izgub ali novih investicijskih projektov. Pri generiranju profita mora podjetje opraviti veliko transakcij in kaj lahko se zgodi, da opravi blagovno transakcijo (proda blago svojemu kupcu), povratna denarna transakcija pa ni izvršena, ali pa plačilo ni izvršeno v dogovorjenem roku. Podjetje je obvezano poravnati svoje obveznosti ob dogovorjenih rokih (dobavitelji, itd.), zato je nespoštovanje rokov ali celo neplačilo s strani kupcev moteče za normalno poslovanje, tako podjetja kot njegove okolice. Z besedo normalno pri tem mislim na podjetje, ki posluje in pri tem ustvarja zadosten denarni tok, da je sposobno poravnati svoje obveznosti ob dogovorjenih rokih. Na kratki rok je likvidno. Hkrati je struktura poslovanja takšna, da je to v trenutku omogočeno tudi na dolgi rok – je solventno. Denimo, da podjetje v trenutku posluje v poslovnem ciklu, ki je ravno v konjunkturi. Najame veliko posojil, ker ocenjuje, da jih bo sposobno v prihodnosti odplačevati. Zaradi optimizma ter relativno nizke obrestne mere (gospodarstvo je v konjunkturi), se bo povečalo investicijsko povpraševanje. Pri relativno nizkih obrestnih merah investitorji menijo, da se bodo te povišala ter jim zagotovile dovolj visoke (pričakovane) donose na njihove investicije (Ribnikar, 2003, str. 34, 65). Zaradi tega se večina denarja, ki pride v gospodarstvo iz strani Centralne Banke aktivira – se porabi za investicije.

Danes so gospodarstva mednarodno povezana, zato obstaja na mednarodnih trgih relativno visoka mera negotovosti glede možnih scenarijev v prihodnosti. Ob recesiji na trgu / v gospodarstvu ali ob nenadnih šokih, kar je lahko posledica raznoraznih dejavnikov kot so splošna recesija, vojne, slabe razmere v določeni panogi, porast cen energentov (nafta), itd., podjetjem nepričakovano upadejo prihodki ali / in porastejo stroški. Posojila oz. investicije pa so planirala ob povsem drugače predpostavljenih scenarijih. Če so takšna podjetja visoko zadolžena, potem obstaja kredibilna grožnja lastnikom dolga, da ne bodo dobili povrnjenih svojih terjatev. Kako se to prenese na finančne trge? Če neko podjetje doživlja recesijo zaradi splošnega upada povpraševanja v panogi ali morda na globalnem trgu, potem ne bo moglo poravnati svojih obveznosti dobaviteljem. Če so le-ti zadolženi pri bankah, so lahko banke, ki takšna podjetja financirajo, potencialno ogrožene, da ne bodo dobile povrnjenih svojih terjatev. S tem primerom sem hotel prikazati kako vpliva povezanost (nefinančnih) podjetij na negotovost na finančnih trgih. Ko podjetje ni več sposobno poravnati svojih obveznosti, je to signal vsem upnikom, da njihove terjatve niso več tako varne kot so bile do tedaj, npr. v trenutku ob odobritvi posojila (blagovnega ali finančnega).

Nezmožnosti podjetja, da poravnava svoje obveznosti na kratki rok, pravimo nelikvidnost. Sama nelikvidnost še ne pomeni nujno, da bo šlo podjetje v stečaj. Glede na oceno stanja na

trgu se lahko dolg rekonstruira oz. refinancira. Če ima podjetje v prihodnosti dobre izglede v očeh posojilodajalca (banke), je torej na dobri poti da preživi kratkoročne (negativne) poslovne šoke. Če pa se izkaže, da tudi na dolgi rok ne bo uspelo poravnati svojih obveznosti, pravimo da je nesolventno. V tem primeru se upniki praviloma odločijo za izterjavo dela dolga, do katerega so upravičeni na podlagi določil iz posojilnih pogodb. Pravimo, da podjetje preneha opravljati svojo dejavnost. Stečaj je način prenehanja gospodarskih subjektov. Stečajni postopek se opravi dolžniku, ki je dalj časa plačilno nesposoben ali prezadolžen ter v drugih primerih, določenih z Zakonom o prisilni poravnavi, stečaju in likvidaciji. V splošnem obstajata dve možnosti vračila terjatev:

- Likvidacija, pri kateri se premoženje podjetja dejansko odproda. Iz dobljene denarne mase se izplačajo deleži upnikom po določenem vrstnem redu, ki sledi iz vrste posameznega vrednostnega papirja (obveznica, delnica, itd.);
- Sanacija, pri kateri se najprej zamenjajo vodstvene strukture v obstoječem podjetju in se določi stečajni upravitelj, ki vodi proces sanacije. V tem procesu se premoženje podjetja ohrani in podjetje ne preneha obstajati, temveč se reorganizira v drugačno pravno-organizacijsko obliko ali pa spremeni vrsto poslovanja, itd.

Da bi upniki prejeli plačilo, morajo osebi, ki je splošno znana kot upravitelj ali likvidacijski izvršitelj in je odgovorna za reorganizacijo ali likvidacijo dolžnikovega premoženja, dokazati svojo terjatev.

2.2 Vpliv stečaja podjetja na banko

Poleg nefinančnih obstajajo tudi podjetja, katerih pridobitna dejavnosti je prav zagotavljanje finančnih sredstev za nefinančna podjetja. Takšno podjetje, ki se ukvarja s finančnim posredništvom, je tudi banka. Banke so depozitni / denarni finančni posredniki, ki prerazporejajo sredstva od suficitnih gospodarskih celic k deficitnim gospodarskim celicam (Ribnikar, 2003, str. 165). To pomeni, da suficitne gospodarske celice, za katere velja, da so prihranki višji od investicij, vlagajo svoje prihranke v banko. Banka nato te prihranke posreduje naprej kot posojila deficitnim gospodarskim celicam (navadno podjetja), za katere velja, da imajo investicije višje od prihrankov. Banke zagotavljajo finančna sredstva za podjetja po določeni ceni, ki je obrestna mera. Podjetja lahko v splošnem posegajo po dveh virih sredstev: prvi je trg kapitala, delniški in /ali obvezniški; drugi pa je bančni trg kreditov. Čeprav v splošnem danes velja, da predvsem večja podjetja, delniške družbe, ki kotirajo na borzah, posegajo po trgu kapitala kot viru sredstev, pa takšna podjetja posegajo tudi po bančnih posojilih. Sam se bom usmeril na banke kot finančne posrednike / posojilodajalce in na podjetja, tako velika kot manjša, kot posojilojemalce. Nefinančno podjetje, ki se ukvarja s pridobitno dejavnostjo, prodaja svoje proizvode kupcem, ki so lahko spet podjetja ali pa končni potrošniki in v zameno zahteva plačilo. Prav tako tudi banka zahteva plačilo za dano posojilo. Vračilo sestavlja glavnica + obresti na izposojeni znesek / glavnico. Ko banka proda podjetju svoj produkt, posojilo, seveda pri tem prevzema tveganje, ki je povezano z

nevračilom posojila. V najbolj skrajni obliki je tveganje povezano z možnostjo, da bo podjetje nezmožno odplačevati posojilo, ki je ga najelo in banka ne bo dobila povrnjenega zneska, ki ga je posodila. Nevračilo je lahko le delno ali pa v višini celotnega posojenega zneska. Če stvar postavim nekoliko bolj v jezik financ oz. v logiko Mertonovega modela (Altman, Resti, Sironi, 2003, str. 3), je dano posojilo za banko prodajna opcija, katere izklicna cena je nominalna vrednost posojila (glavnica). Ko banka opcijo proda, obstajata dve možnosti:

- dejanska prodajna cena je enaka izklicni ceni na opciji. V tem primeru banka ne izgubi ničesar, v bistvu zasluži z obrestmi. Njen profit je enak višini obresti na dano posojilo;
- dejanska cena v času prodaje opcije je nižja kot je izklicna cena, vrednost podjetja je padla. Če ob dospetju posojila banka ne more prodati svoje opcije po izklicni ceni temveč po nižji, ali če je dejanska cena enaka 0, potem je banka dejansko naredila izgubo za razliko med izklicno ceno in dejansko prodajno ceno. Izguba je za banko lahko različno velika, je pa definirana na intervalu $[0, G]$, kjer je G posojena glavnica oz. izklicna cena opcije.

Če se odmaknem od razmišljanja o opciji in spet preidem na vpliv stečaja podjetja na bančne odločitve pri dajanju posojil, potem iz zgoraj napisanega sledi, da se bo banka poskušala maksimalno zavarovati pred tveganjem, da bi naredila z danim posojilom izgubo. Z vidika banke, če gledamo maksimizacijo profita ali minimizacijo tveganja, je zaželeno oz. obvezno, da se zavaruje pred možnostjo nevračila posojila s strani svojih posojilojemalcev. Eno odpisano posojilo še ne bo znatno ogrozilo obstoja banke, vendar iz systemskega vidika to pomeni neracionalno prerazdelitev dela prihrankov, ki so prešli iz banke na neko podjetje oz. vse, ki so z njim povezani. Če bi vsaka banka izgubila po eno ali nekaj posojil, bi to pomenilo iz vidika monetarnega sistema veliko ekonomsko škodo. Resnost problema se pokaže v drugi luči, če pomislimo, da bi banka izgubila polovico svojih posojil. Bančni komitenti, ki so v banko pred časom prinesli svoje prihranke, kar je omogočilo izdajanje posojil podjetjem s strani banke, bi hoteli dvigniti vse oz. relativno velik del svojih prihrankov. Banka, ki je izgubila svoja sredstva (odpisana dana posojila podjetjem) ne bi mogla več zagotavljati vračila prihrankov svojim komitentom, zato lahko nastopi splošna bančna kriza, ki se rodi iz nezaupanja v bančni sistem. Banka je pomembna za obstoj finančnega sistema zaradi svoje funkcije prerazporejanja sredstev / prihrankov, zato morajo biti njena sredstva racionalno in varno naložena. To pomeni, da mora banka dajati posojila le tistim podjetjem, za katera z dovolj visoko verjetnostjo ve, da bodo posojilo sposobna tudi vrniti. Res je, da mora podjetje v trenutku, ko najame posojilo izpolnjevati določene pogoje: garancija, zastavitev premoženja, itd. Vse to povečuje verjetnost, da bo posojilo vrnjeno, vsekakor pa to ni zadostni pogoj za odločitev banke pri dajanju posojil. Ključno za banko je, ali bo lahko podjetje poslovalo tudi v prihodnosti, kar je pogoj za vračilo posojila, ki se bo zagotovilo iz prihodnjih denarnih tokov, ki naj bi jih generiralo poslujoče podjetje. Analiza in napovedovanje prihodnjega poslovanja podjetja je postala izredno uporabno orodje, predvsem z razvojem statističnih metod v kombinaciji s finančnimi kazalniki izkaza stanja, izkaza poslovnega izida ter izkaza denarnih tokov. Poslovanje podjetja se v veliki meri odraža v njegovih finančnih

kazalnikih, če imamo opravka s kolikor toliko mednarodno primerljivim računovodskim sistemom kot velja v razvitih gospodarstvih. Finančni kazalniki so postali sredstvo za analiziranje in napovedovanje stečaja že veliko preden so se razvile metode multivariatne analize, vendar je bila takrat metodološka zasnova še zelo trivialna. Ni se upoštevalo medsebojne povezanosti (korelacij) spremenljivk, pa tudi računalniški sistemi ter programska orodja za obdelave velikih podatkovnih baz še niso bili tako razviti oz. dostopni, da bi lahko omogočali obsežnejše izvajanje statističnih metod kot jih poznamo danes.

2.3 Finančni kazalniki kot sredstvo manjšanja negotovosti za banko

S koncem industrijske revolucije in začetkom informacijske se je število podjetij na trgih, s čimer mislim tako mednarodno povezane kot tudi nacionalne trge, začelo skokovito povečevati. To je posledično pomenilo povečano povpraševanje po bančnih posojilih. Ker pa so banke pri dajanju posojil načeloma omejene s količino prihrankov, ki jih dobijo od svojih komitentov, morajo te prihranke razporejati racionalno. To pomeni, dajati posojila le tistim podjetjem, za katera pričakujejo, da jih bodo tudi sposobna odplačati. Finančni kazalniki so postali široko uporabni zaradi količine informacije, ki jo nudijo glede poslovanja podjetja. Seveda je bila na začetku analiza zgolj na univariatni ravni, brez kompleksnih statističnih metod in brez upoštevanja korelacij med spremenljivkami. To pomeni, da so iz vzorcev gibanja posameznih finančnih kazalnikov ugotavljali, v kakšni smeri se podjetje giblje oz. kakšen je trend napovedan na podlagi podatkov za poljubno število let iz preteklosti. Že v letu 1848 so začeli uporabljati analizo na podlagi finančnih kazalnikov za določanje kreditne bonitete posameznega podjetja (McGraw – Hill, 1961). V letu 1935 so raziskave pokazale, da podjetja, ki so pozneje bankrotirala, kažejo že v času pred bankrotom značilno / opazno različne vzorce gibanja finančnih kazalnikov, kot podjetja, za katera se je izkazalo, da niso bankrotirala (R. F. Smith in A. H. Winakor, 1935). Poznejše študije so pokazale, da se začnejo znaki stečaja kazati že v petem letu pred dejanskim stečajem. To je bilo potrjeno na določenih vzorcih bankrotiranih in nebankrotiranih podjetij (Altman, 1968). Študije so pokazale, da so finančni kazalniki lahko dobri indikatorji za napovedovanje stečaja podjetja.

V splošnem lahko finančne kazalnike razdelimo v naslednje skupine: kazalniki likvidnosti, kazalniki dobičkonosnosti, kazalniki učinkovitosti in kazalniki zadolženosti. Ni čisto določeno, kako je z vrstnim redom pomembnosti: kazalniki katere skupine bolj pravilno napovedo poslovanje (stečaj) podjetja. Pri pravilnosti napovedovanja na tej točki mislim predvsem kako zgodovinska dinamika gibanja določenega kazalnika uspe napovedati prihodnje gibanje vrednosti le-tega. Na podlagi preteklih podatkov je vsaka transformacija v prihodnost dosežena z določeno mero verjetnosti, ki pa je vedno manjša od 1. Enostavnost je bila pomemben razlog, da so se začeli finančni kazalniki široko uporabljati. Izkazalo se je, da je mogoče z manj kot 10 kazalniki zadovoljivo določiti stanje podjetja in tudi predvideti, sicer 'ad hoc', njegovo poslovanje v prihodnosti. Vendar je prav v tej preprostosti finančnih kazalnikov tudi zanka, zaradi katere so postajali sčasoma vse manj uporabni (Altman, 1968, str. 591). Tako je npr. banka izgubila posel in ni dala posojila kakšnemu podjetju, za katerega

se je na podlagi (univariatnih) analiz kazalnikov izkazalo, da je njegovo poslovanje v prihodnosti ogroženo – podjetje je bilo navidezno preveč tvegano. Nasprotno se je dogajalo, da je imelo neko drugo podjetje glede na analizo posameznih kazalnikov odlične izglede za prihodnost in mu je banka odobrila posojilo – podjetje je bilo navidezno relativno netvegano – kasneje pa se je izkazalo, da je šlo v stečaj in banka je bila ob del, ali celo ob vse svoje terjatve. Zaradi analiziranja zgolj na univariatni ravni je prihajalo do primerov, ko je bilo neko podjetje npr. trenutno nelikvidno, čeprav to sploh ne pomeni, da bo v prihodnosti nesolventno in zato ni dobilo odobrenega posojila. Dogajalo se je tudi nasprotno. Takšni primeri so pomenili nevzpodbudno prihodnost za finančne kazalnike kot sredstva za napovedovanje prihodnjega poslovanja podjetij za potrebe bank, da bi zagotovile kar najbolj racionalno reorganizacijo svojih sredstev (posojil) (Altman, 1968, str. 589, 591). Z razvojem statističnih metod multivariatne analize, ki upoštevajo tudi korelacije med spremenljivkami, se je pomen finančnih kazalnikov kot sredstva za napovedovanje prihodnjega poslovanja podjetij močno povečal. V naslednjih poglavjih sta predstavljeni dve orodji multivariatne analize, s katerima sem v tej nalogi opravil analizo napovedovanja stečajev podjetij glede na izbrane finančne kazalnike ter izbrane vzorce. Ti dve orodji sta: diskriminantna analiza (ang. *discriminant analysis*) in logistična regresija (ang. *logistic regression*).

3 DISKRIMINANTNA ANALIZA

Čeprav so se analize napovedovanja stečajev podjetij na podlagi (univariatnih) metod opisanih v prejšnjem poglavju izkazale kot dobre in dejansko uporabne, so njihove implikacije danes nekoliko vprašljive. Podjetja so imela slabe izglede na podlagi kazalnikov (npr. profitabilnosti) in so zaradi tega kazala znake slabega stanja. Na drugi strani so imela takšna podjetja dobre izglede na podlagi ostalih kazalnikov, npr. likvidnosti. Stvar je postala nekoliko nejasna in ni bilo določenega odgovora, katerim sklepom dati večjo težo. Še vedno je bilo precej negotovosti kljub uporabi statističnih metod. Naj takšno podjetje dejansko dobi posojilo od banke ali naj ga banka zavrne. Nejasnosti zgoraj opisanih rezultatov so med drugim tudi posledica nepoznavanja povezanosti med posameznimi skupinami kazalnikov in tudi med kazalniki samimi. Univariatna analiza je postajala vse bolj nezadostna za odločitve bank, ali naj se določenemu podjetju odobri posojilo ali ne. Problem univariatne analize je v tem, da analizira posamezni kazalnik izolirano od ostalih. Dodatno je nemogoče določiti ponder posameznemu kazalniku. Že mogoče, da ima podjetje dobre izglede na podlagi kazalnikov likvidnosti in slabe na podlagi dobičkonosnosti, ampak kakšno težo dati posameznemu kazalniku. Vsem enako? Bilo bi mogoče, vendar zelo verjetno neoptimalno, kar bi vodilo v napačne sklepe. Lahko bi relativno veliko zgrešili, če bi 'ad hoc' dali vsem kazalnikom enako težo. Iz problematike tega konteksta se je rodila ideja, kako določiti neko mero, ki bo določena na podlagi (linearne) kombinacije določenih finančnih kazalnikov. Diskriminantna analiza (in logistična regresija) se ukvarjata prav z določanjem linearne kombinacije, z upoštevanjem korelacij med spremenljivkami, ki za vsako enoto (podjetje) priredi določeno vrednost odvisne spremenljivke, ki določa stopnjo tveganosti podjetja.

3.1 Metodološka logika

Podobno kot pri linearni regresiji poskušamo tudi pri diskriminantni analizi s pomočjo nabora neodvisnih spremenljivk (prediktorjev) napovedati vrednosti določene odvisne spremenljivke. Osnovna razlika med regresijo in diskriminantno analizo je v odvisni spremenljivki. Za razliko od (linearnih) regresijskih modelov, ki uporabljajo zvezne numerične odvisne spremenljivke, se pri diskriminantni analizi omejimo zgolj na nezvezne (diskretne) numerične odvisne spremenljivke. Pravimo jim tudi kategorične, ker ponavadi vsaka vrednost opredeljuje določeno kategorijo / skupino. Kategorična spremenljivka ima lahko najmanj dve vrednosti. V primeru te naloge imam opravka z dvema skupinama enot. Vsaki skupini sem pripisal določeno vrednost odvisne spremenljivke, poimenoval sem jo *stečaj*. Vrednost 0 za spremenljivko *stečaj* uporabljam za skupino normalno poslujočih (nebankrotiranih) podjetij, vrednost 1 pa za skupino podjetij, ki so šla v stečaj. Osnovna ideja diskriminantne analize je določiti vrednostno os, ki bo razlikovala med skupinami, hkrati pa bo čimmanj mešanja med enotami posameznih skupin, da bi lahko zagotovili kar najvišjo možno učinkovitost modela – diskriminantne funkcije. Proces razvrščanja v tej nalogi poteka s pomočjo Fisherjevih diskriminantnih funkcij. Simbolični zapis Fisherjeve diskriminantne funkcije za skupino n je naslednji:

$$Z_n = \alpha + \beta_1 * X_{1,i} + \beta_2 * X_{2,i} + \dots + \beta_p * X_{p,i}. \quad (1)$$

Z_n je vrednost Fisherjeve diskriminantne funkcije za skupino n (v tej nalogi n zavzame dve vrednosti, in sicer 0 in 1), $X_{p,i}$ pa so neodvisne spremenljivke za enoto i . Enote, ki jih razvrščamo z diskriminantno analizo, so dejansko že razvrščene v skupine (nebankrotirana in bankrotirana podjetja), kar temelji na zgodovinskih podatkih. Vendar ta delitev na prvi pogled glede na vrednosti neodvisnih spremenljivk (finančnih kazalnikov), ni jasno razvidna. Če bi sklepali na podlagi prvega pogleda (ad-hoc), bi lahko relativno veliko zgrešili pri uvrstitvi podjetij v skupine. Banke podjetij ne razvrščajo glede kreditne bonitete ad-hoc, ampak poteka razvrščanje na podlagi sofisticiranih metodologij, tudi diskriminantne analize. Vsaka enota ima svoje attribute oz. nabor neodvisnih spremenljivk, na podlagi katerih poteka razvrščanje. β_i so koeficienti diskriminantne funkcije, ki določajo smer in velikost vpliva vsake neodvisne spremenljivke v enačbi 1.

Fisherjevih diskriminantnih funkcij je lahko toliko kot je skupin, zato sem v tej nalogi ocenil dve diskriminantni funkciji. Razvrščanje podjetij poteka po naslednjem postopku: vsakemu podjetju se priredi vrednost Z za vsako izmed ocenjenih fisherjevih diskriminantnih funkcij, nato pa se podjetje razvrsti v tisto skupino, v kateri ima spremenljivka Z najvišjo vrednost. Statistični paketi ponavadi razvrstijo podjetja na podlagi točke reza (ang. *cut-off point*). Program izračuna eno funkcijo, ki ji pravimo kanonična diskriminantna funkcija, njene vrednosti se porazdeljujejo na premici. Glede na ocenjene Fisherjeve diskriminantne funkcije se nato določi točka reza, ki razdeli premico na dva podprostor: v enem so razvrščena podjetja skupine 0, v drugem pa podjetja skupine 1. Vrednosti odvisne spremenljivke Z so tako porazdeljene na premici, kar je tudi glavna prednost diskriminantne analize, ki očitno

omogoča poenostavitev, v tem primeru, na le eno dimenzijo, enorazsežni prostor – premico. Vsaka enota ima določeno vrednost spremenljivke Z , ki je odvisna le in zgolj od vrednosti neodvisnih spremenljivk, v primeru te naloge, od vrednosti finančnih kazalnikov, ki sem jih vključil v analizo. Z njimi želim opisati podjetje glede na položaj na premici, ki jo določajo vrednosti diskriminantnih funkcij. Če smo izbrali prave neodvisne spremenljivke (kako je s izborom je pojasnjeno kasneje), bi se morale na vsaki strani premice oblikovati (relativno) homogene skupine (Rovan, 2007, str. 7). Na eni strani se bodo gostile enote, ki so bankrotirala, na drugi strani pa nebankrotirana podjetja. Nekje med tema skupinama je mejno področje, kjer se nahajajo enote, ki so napačno razvrščene z ocenjenim modelom. Prednost diskriminantne analize je sedaj že očitna. Na podlagi enostavnega kriterija omogoča razvrščanje podjetij glede tveganost poslovanja z uporabo kombinacije finančnih kazalnikov. Če bi bil model idealen, bi na koncu morala biti vsa podjetja na podlagi rezultatov diskriminantne analize razvrščena enako kot so bila razvrščena že na podlagi zgodovinskih podatkov. Vsa bankrotirana podjetja v eni skupini in vsa nebankrotirana v drugi skupini, brez mešanja med skupinami (napačnega razvrščanja).

Omeniti velja tri temeljne predpostavke, ki morajo biti izpolnjene zato, da so podatki sploh primerni za diskriminantno analizo. Za relativno majhne vzorce so te predpostavke naslednje:

- 1) neodvisne spremenljivke se znotraj skupin porazdeljujejo normalno;
- 2) enakost varianc neodvisnih spremenljivk znotraj skupin;
- 3) neodvisnost vrednosti spremenljivk znotraj skupin.

Predpostavke so lahko kršene, vendar pod pogojem da delamo z dovolj velikimi vzorci ($n > 100$). Da bi lahko pojasnil potrebnost zgoraj naštetih predpostavk, moram najprej predstaviti nekaj osnovnih pojmov. *Centroid* je vektor, ki ga simbolično zapišemo kot:

$$\mu'_g = [\mu_{1g}, \mu_{2g}, \dots, \mu_{pg}] \quad \text{in} \quad \mu_{pg} = \frac{1}{n} \sum_{n_g=1}^{n_g} y_{ng} \quad (3)$$

Indeks g predstavlja skupino, indeks p pa predstavlja spremenljivko, na katero se komponenta centroida nanaša. Centroid μ'_g torej predstavlja točko v prostoru, ki je določena z vrednostmi aritmetičnih sredin za vsako izmed neodvisnih spremenljivk v skupini g . Metode razvrščanja, ki se uporabljajo pri diskriminantni analizi, temeljijo na osnovi *razdalje* posameznih enot od neke določene točke v prostoru, ponavadi od centroida posamezne skupine. Osnovni kriterij je tisti, ki razvršča enote po evklidski razdalji od centroida. Vektorsko lahko evklidsko razdaljo med poljubnima točkama zapišem kot:

$$d_{AB}^2 = \sum_{i=1}^n (X_{iA} - X_{iB})^2 = [X_A - X_B]^* [X_A - X_B] \quad (4)$$

Indeks i predstavlja rezsežnost prostora, v katerem se nahaja točka (enota). Izraz na skrajno desni enačbe 4 se nanaša na matrično obliko zapisa evklidske razdalje. Za praktično uporabo ima metoda razvrščanja na osnovi evklidske razdalje pomankljivost, saj ne upošteva korelacij med spremenljivkami. Ko imamo opravka z medsebojno koreliranimi spremenljivkami, je

potrebno uvesti pojem variančno-kovariančna matrika (v nadaljevanju kovariančna matrika), ki jo simbolično zapišemo kot:

$$\Sigma = \begin{bmatrix} Var_{X_1} & Cov_{X_2X_1} & Cov_{X_3X_1} \\ Cov_{X_1X_2} & Var_{X_2} & Cov_{X_3X_2} \\ Cov_{X_1X_3} & Cov_{X_2X_3} & \dots Var_{X_n} \end{bmatrix} \quad (5)$$

X_n predstavlja neodvisno spremenljivko in $n \in \{1 \dots p\}$. Na glavni diagonali kovariančne matrike se nahajajo elementi, ki predstavljajo variance posameznih neodvisnih spremenljivk. Izven glavne diagonale so kovariance med spremenljivkami. V praksi se zaradi enostavnosti uporablja le spodnja / zgornja tirkotna kovariančna matrika, npr. $Cov_{X_1X_2} = Cov_{X_2X_1}$ in so zato zrcalni izvendiaagonalni elementi odveč. Če upoštevamo še korelacijo med spremenljivkami, kar sem shematično predstavil z matriko 5, lahko zapišem razdaljo med dvema enotama, ki sta medsebojno korelirani v spremenljivkah, z naslednjo formulo:

$$\Delta_{AB}^2 = [X_A - X_B]' \Sigma [X_A - X_B]. \quad (6)$$

kjer Σ predstavlja kovariančno matriko za nabor neodvisnih spremenljivk. Tako določeni razdalji (6) pravimo Mahalanobis razdalja, ki upošteva tudi korelacije. Sledi podrobnejši opis predpostavk, na katerih temelji diskriminantna analiza.

3.2 Predpostavka 1: Normalno porazdeljevanje neodvisnih spremenljivk v skupinah

V prejšnjem razdelku sem omenil, da se uporablja razdalja (evklidska, mahalnobis ali kakšna druga vrsta razdalje) kot kriterij za razvrščanje enot v skupine. Na tem mestu je potrebno dodati, da se poleg tega kriterija pri razvrščanju enot uporablja še drugi kriterij, in sicer relativna frekvenca. Huberty (1994, str. 46) postavlja naslednji kriterij za razvrstitev določene enote v eno izmed skupin. Razvrstimo enoto A v skupino g , če ima vektor opazovanja X enote A , $X'_A [X_{A1}, X_{A2}, \dots, X_{Ap}]$ v skupini g večjo relativno frekvenco kot v katerikoli skupini; g' , če velja $g \neq g'$. Simbolično lahko to zapišemo kot:

$$X_A \in g \text{ če } P(X_A | g) > P(X_A | g'). \quad (7)$$

$P(\cdot)$ je relativna frekvenca ali verjetnost, da ima vektor opazovanja enote A dejansko vrednost blizu X . To pomeni, da bo enota A razvrščena v tisto skupino, katere bolj verjeten predstavnik je. Verjetnosti $P(X | g)$ pravimo inverzna verjetnost (ang. *inverse probability*). V tem primeru dejansko variramo skupine g pri fiksirani vrednosti vektorja X in ko preletimo vse skupine, enota z vektorjem X pade v tisto, za katero je največja inverzna verjetnost. Verjetnost, definirano kot $P(g | X)$, imenujemo posteriorna verjetnost (ang. *posterior probability*) in ta je tista, ki nas neposredno zanima pri napovedovanju z diskriminantno analizo. Posteriorna verjetnost za enoto A z vektorjem opazovanja X je definirana kot:

$$P(g | X) = \frac{P(X | g)}{\sum_{g'=1}^k P(X | g')}. \quad (8)$$

Posteriorna verjetnost, da enota z vektorjem opazovanja X pade v skupino g , je torej inverzna verjetnost za to enoto v skupini g deljena z vsoto vseh inverznih verjetnosti za to enoto po vseh skupinah. Logično sledi, da bo enota z vektorjem X razvrščena v skupino g , če velja $P(g | X) > P(g' | X)$ če velja $g \neq g'$ in $g' \in \{1 \dots k\}$.

3.3 Predpostavka 2: Enakost varianc neodvisnih spremenljivk v skupinah

Ta predpostavka je potrebna zaradi zapisa enačbe posteriorne verjetnosti pod 9. Enakost varianc v skupinah pomeni enakost kovariančnih matrik, kar se nanaša na metodo izračuna¹. Ločimo: linearne forme in kvadratne forme. Linearne forme se od kvadratnih razlikujejo po elementih, ki jih upoštevajo pri izračunu posteriornih verjetnosti za posamezno enoto. Zaradi enakosti kovariančnih matrik linearna forma pri izračunu ne upošteva kovariančne matrike pri izračunu posteriornih verjetnosti (Huberty, 1994, str. 57). Denimo, da imamo opravka z deležem (%) enot celotne populacije, recimo temu deležu q , vse enote iz tega deleža pa pripadajo skupini g . Če na slepo izberemo eno enoto iz celotne populacije je verjetnost, da bo enota iz skupine g enaka ravno q . Tej verjetnosti pravimo apriorna verjetnost (ang. *prior probability*). Predpostavka o enakosti varianc je lahko kršena, vendar je priporočljivo, da operiramo z dovolj velikimi vzroci po skupinah. Če razpolagamo z enako velikimi skupinami, to pomeni enake q za vse skupine. To vpliva na kvaliteto izračuna. Če predpostavimo neenakost kovariančnih matrik in neenake q , se enačba za izračun posteriorne verjetnosti glasi (Huberty, 1994, str. 56):

$$P(g | X) = \frac{q_g * |S_g|^{-1/2} * \exp(-\frac{1}{2} D_{ug}^2)}{\sum_{g'=1}^k q_{g'} * |S_{g'}|^{-1/2} * \exp(-\frac{1}{2} D_{ug'}^2)} \quad (9)$$

kjer je q je apriorna verjetnost za enote skupine g , $|S_g|$ je determinanta kovariančne matrike za skupino g , izraz v eksponentu pa predstavlja kvadrat razdalje enote u od centroida skupine g . Iz sestave enačbe 9 se vidi, zakaj je ob kršitvi predpostavke o enakosti varianc v skupinah pomembna velikost skupin. Možno je, da izberemo takšne apriorne verjetnosti, ki bodo razlike med števcem in imenovalcem povečale / zmanjšale. Posledično bi lahko zagrešili pristranskost pri izračunu posteriorne verjetnosti za določeno enoto, kar lahko pomeni tudi napačno razvrstitev. Kljub temu je pri izbiri q veliko odvisno od subjektivne presoje.

Tretja predpostavka diskriminantne analize se nanaša na odnose med spremenljivkami. Te morajo biti medsebojno linearno neodvisne. To pomeni, da nobena spremenljivka ne sme biti linearna kombinacija preostalih (prepoved multikolinearnosti).

1) Varianca v tem primeru pomeni *generalizirano* varianco, ki je definirana kot determinanta kovariančne matrike S_g . Če je $p=1$, potem je $\det(S_g) = \sigma^2$, kjer je p število neodvisnih spremenljivk. V multivariatni analizi, kjer je $p > 1$, se pri determinanti odštejejo produkti stranskih diagonal, kar pomeni izločitev kovariabilnosti.

Za trenutek bi se še ustavil pri dveh oblikah oz. formah za izračunavanje posteriornih verjetnosti za posamezno enoto. Fisherjeva linearna forma, ki upošteva enake variance med skupinami, ima naslednjo obliko:

$$P(g | X_u) = \frac{q_g * \exp(-\frac{1}{2} D_{ug}^2)}{\sum_{g'=1}^k q_{g'} * \exp(-\frac{1}{2} D_{ug'}^2)}. \quad (10)$$

Smithova kvadratna forma za izračun posteriorne verjetnosti ima obliko, ki sem jo zapisal pod 9. Ker sem v tej nalogi uporabil velike vzorce predpostavljam, da bo tudi ob morebitni kršitvi katere izmed predpostavk, ki veljajo za diskriminantno analizo, analiza mogoča s Fisherjevo linearno formo. Za vsako izmed skupin lahko izračunamo po eno Fisherjevo diskriminantno funkcijo. V primeru te naloge imam dve skupini, zato sem ocenil dve Fisherjevi diskriminantni funkciji, ki sem jih izračunal v programskem paketu SPSS na podlagi izbranih finančnih kazalnikov. Vrednosti odvisne spremenljivke Fisherjeve diskriminantne funkcije za skupino nebankrotiranih podjetij sem označil z $Z0$, za skupino bankrotiranih pa z $Z1$. Če velja $Z0 > Z1$, potem bo podjetje na podlagi ocenjenih funkcij razvrščeno v skupino 0. Če pa velja $Z0 < Z1$, potem bo podjetje označeno kot zelo tvegano in bo zato razvrščeno v skupino 1. Navadno podamo poleg Fisherjevih diskriminantnih funkcij še zapis kanonične diskriminantne funkcije, ki kaže, kako so posamezne spremenljivke povezane z odvisno, v primeru te naloge dihotočno, spremenljivko.

4 LOGISTIČNA REGRESIJA

Logistično regresijo uporabljamo v primerih, ko nas zanima dejanska verjetnost pripadnosti neke enote eni ali drugi skupini. Logistična regresija spada v družino verjetnostnih modelov, v katerih napovedujemo verjetnost dogodka. Najenostavnejši predstavnik te družine je linearni verjetnostni model. Dogodek je definiran kot neko stanje enote. V principu to pomeni pripadnost neki skupini z določeno lastnostjo. Največkrat tako stanje označimo z 1, nasprotno stanje pa z 0. Obstajajo tudi drugi načini kodiranja (pripisovanja) vrednosti spremenljivkam za določene dogodke. V linearnem verjetnostnem modelu se verjetnost dogodka linearno povečuje na nekem intervalu neodvisnih spremenljivk. To je ponazorjeno na Sliki 1 v prilogah. Linearni verjetnostni model je v bistvu regresijska analiza, izvedena s pomočjo metode najmanjših kvadratov odklonov (ang. *ordinary least squares method*). Model zapišemo kot verjetnost, da ima odvisna spremenljivka vrednost 1 (dogodek se zgodi):

$$P(Y_i = 1) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + e_i = \sum_{j=1}^k b_j X_{ij} + e_i. \quad (11)$$

V primeru te naloge velja: $i \in \{1, 2, \dots, n_k\}$ in $k \in \{0, 1\}$. Pomankljivost linearnega verjetnostnega modela je v tem, da lahko neodvisna spremenljivka zavzame le vrednosti na nekem (omejenem) intervalu. Zunaj tega intervala je verjetnost dogodka 0 oziroma 1. Pri logističnem modelu razširimo interval na celotno os neodvisne spremenljivke. Grafično lahko model predstavimo, kot je razvidno na Sliki 2 (priloge).

4.1 Metodološka logika

Model je definiran z naslednjo enačbo: $P(X_i = 1) = \frac{e^Z}{1 + e^Z}$; pri tem pa je $Z = \sum_{j=1}^k \beta_j X_{ij}$, kar je

dejanska verjetnost, da se bo dogodek zgodil. Verjetnost, da do dogodka ne bo prišlo, je podana kot $P(Y_i = 0) = 1 - P(Y_i = 1)$. Če namesto verjetnosti dogodka uporabljamo možnosti

(odds), lahko model zapišemo tudi v linearni obliki, in sicer kot $\log\left(\frac{P_i}{1 - P_i}\right) = \sum_{j=1}^k b_j X_{ij}$, čemur

pravimo *logit transformacija* ali kratko *logit*, ki je linearna v parametrih. Parametri b so ocenjeni po metodi največjega verjetja; logika metode je analogna metodi najmanjših kvadratov. Želimo oceniti tak b_j , da bo verjetnost, da smo s funkcijo izbrali pravi Y kar največja. Podobno kot pri diskriminantni analizi določimo točko reza diskriminantne funkcije, lahko pri logistični regresiji določimo verjetnost pri točki reza (ang. *cut off point*), ki razdeli populacijo v dve skupini. Vrednosti za točko reza lahko določimo podobno kot določamo apriorne verjetnosti pri diskriminantni analizi. Privzeta verjetnost pri točki reza je 0,5. Koeficiente interpretiramo glede na njihov vpliv na možnosti (Pahor, 2003, str. 2).

Če izhajamo iz gornjega linearnega zapisa modela sledi, da pri pozitivnih vrednostih koeficientov β_j povečanje vrednosti neodvisne spremenljivke poveča logaritem možnosti. Če se poveča logaritem, se povečajo možnosti, kar pomeni, da je verjetnost dogodka večja. Pozitivni β_j torej kažejo pozitiven vpliv spremenljivke na verjetnost dogodka, negativni pa negativen. Kakovost modela lahko ocenimo z več statistikami. Problem je v tem, da nobena ni tako jasna in splošno sprejeta kot determinacijski koeficient pri regresijskih modelih. Ocenjujemo, kako dobro se ocenjen model prilega začetnim podatkom. Definiramo funkcijo verjetja (ang. *likelihood function*) kot verjetnost, da bo napovedan rezultat enak dejanskemu stanju. Vrednost funkcije verjetja pri popolnem prileganju je 1. Običajno je funkcija verjetja podana v obliki $-2 * \log_likelihood$. Nižja vrednost pomeni boljši model; pri idealnem prileganju bi bila vrednost tega izraza 0. Drug kazalec, ki kaže kakovost modela je definiran podobno kot vsota kvadratov odklonov. Zapišemo ga kot:

$$Z^2 = \sum_{i=1}^n \frac{(Y_i - P_i)^2}{P_i(1 - P_i)}$$

Poznamo pa tudi dva približka gornje statistike. Prvi je:

$$Cox \& Snell _ R^2 = 1 - \left(\frac{L(0)}{L(B)} \right)^{\frac{2}{n}}$$

kjer je $L(0)$ vrednost funkcije verjetja s samo konstanto in $L(B)$ vrednost funkcije verjetja ocenjenega modela. Drugi popravek je:

$$Nagel_ker_keR^2 = \frac{R^2}{R_{MAX}^2}; R_{MAX}^2 = 1 - (L(0))^2.$$

Logistična regresija ne predpostavlja normalno porazdeljenih neodvisnih spremenljivk. To pomeni, da je v primerih, ko spremenljivke niso normalno porazdeljene ali ko imamo opravka

z naborom mešanih, številskih in nominalnih neodvisnih spremenljivk, logistična regresija v večini primerov boljša izbira kot diskriminantna analiza. V sklepnem delu analize sem se posvetil neposredni primerjavi rezultatov dobljenih z diskriminantno analizo in logistično regresijo, zato v nadaljevanju opisujem še nekaj lastnosti logistične regresije, v zvezi s kvaliteto rezultatov.

Ugotavljam, kako je z variiranjem pravilnosti razvrstitve v odvisnosti od spreminjanja parametrov v modelu. Pri parametrov mislim na točko reza (ang. *cut-off point*), ki mi razdeli prostor (premico), na kateri so porazdeljene verjetnosti stečaja za podjetja na dva (ni nujno) enaka dela. V tej nalogi sem zaradi primerjave rezultatov stremel k takšni točki reza, ki mi bo zagotavljala karseda enakomerno relativno kvaliteto pravilnosti razvrstitve v vsaki izmed obeh skupin. S tem v zvezi definiram dva pojma, ki se uporabljata kot merilo kvalitete razvrstitve pri logistični regresiji: občutljivost (ang. *sensitivity*) in specifičnost (ang. *specificity*). Občutljivost se nanaša na skupino enot, za katere se je dogodek dejansko zgodil (stečaj=1 v primeru te naloge) in je definirana kot razmerje med deležem pravilno razvrščenih enot in vsemi enotami s takšno lastnostjo glede na podatke. Specifičnost se nanaša na skupino enot, za katere se dogodek dejansko ni zgodil (stečaj=0 v primeru te naloge) in je definirana kot razmerje med deležem pravilno razvrščenih takšnih enot in vsemi enotami s takšno lastnostjo glede na podatke.

Literatura (Hosmer, Lemeshow, 2000, str. 161) pravi, da sta si občutljivost ter specifičnost v obratnem odnosu, v odvisnosti od točke reza. Ko se s to točko oddaljujemo od vrednosti 0 proti 1, se občutljivost zmanjšuje, medtem ko se specifičnost povečuje. V neki točki se izenačita, nato pa desno od te točke reza občutljivost postane enaka 1 (100%), specifičnost pa postane enaka 0. V tej nalogi sem poskušal najti tisti približek točke reza, pri kateri sem v točki, ko bosta občutljivost in specifičnost enaki oz. zelo podobni. Enako strategijo sem uporabil tudi pri diskriminantni analizi, le da sem tam variiral apriorne verjetnosti in nastavil razmerje tako, da mi omogoča karseda enakomerno pravilnost razvrščanja v obeh skupinah podjetij.

5 OPREDELITEV FINANČNIH KAZALNIKOV V ANALIZI

Finančni kazalniki so izvedena razmerja iz podatkov, ki so zavedeni v računovodskih izkazih podjetja. V tej nalogi služijo kot prediktorji / neodvisne spremenljivke (ang. *predictors*) za napovedovanje stečaja podjetij. Standardni finančni kazalniki, ki se pogosteje uporabljajo v analizah poslovanja podjetja, zavzemajo naslednje skupine, ki se nanašajo na posamezne komponente, s katerimi lahko ocenimo stanje podjetja v določenem trenutku:

- kazalniki likvidnosti;
- kazalniki učinkovitosti;
- kazalniki zadolženosti;
- kazalniki dobičkonosnosti.

5.1 Lastnosti podatkov in primernost za uporabo diskriminantne analize

Osnovna predpostavka, na kateri temelji diskriminantna analiza, je predpostavka o normalnosti porazdelitev neodvisnih spremenljivk – finančnih kazalnikov v primeru te naloge. Iz literature, ki se ukvarja s finančnimi analizami je razbrati, da se v večini ne porazdeljujejo normalno, ampak v porazdelitvah, ki so asimetrične v desno. Diskriminantna analiza je na takšne kršitve ob določenih okoliščinah močno občutljiva. Zulkarnain, Mahamad in Annuar (2006) predlagajo logaritemsko transformacijo spremenljivk (finančnih kazalnikov) za njihovo normalizacijo.

5.2 Podatkovna zasnova in prilagoditve začetne baze

Podaki so iz podatkovne baze Ajpes (Agencija Republike Slovenije za javnopravne evidence in storitve) in obsegajo obdobje 1994 – 2004. Panoge, v katere spadajo podjetja, so večinoma proizvodne in tudi predelovalne. Obilica podatkov mi omogoča, da sem podjetja z nanjkajočimi vrednostmi spremenljivk izločil iz analize. V literaturi se sicer pojavlja več možnosti namesto brisanja enot z manjkajočimi vrednostmi, vendar je to grajeno pod predpostavko, da imamo opravka z manjšimi podatkovnimi bazami, ko je izguba vsake informacije relativno visok strošek. V literaturi se redko uporabljajo kazalniki, ki so oddaljeni več kot tri leta od leta stečaja. V članku E. Altmana (1968, str. 604) je evidentno, da se po tretjem letu učinkovitost modela drastično poslabša in pade celo pod 50%, kar pove, da je bolje razvrščati podjetja na oko, slučajno, kot pa graditi model. Analizo sem začel z osemindvajset (28) finančnimi kazalniki. Specifikacije kazalnikov so podane v Tabeli 1 v prilogah. Finančni kazalniki so za obe skupini podjetij izrazito asimetrični v desno (Tabeli 2 in 3, priloge), pa tudi hipoteza o normalnosti je gladko zavrnjena pri točni stopnji značilnosti $p=0,000$ (Tabela 5 priloge, 3. in 5. kolona). Na samem začetku sem odstranil 1% gornjih vrednosti pri vsaki spremenljivki, ker imam vse spremenljivke izrazito asimetrične v desno. Nato sem oblikoval standardizirane vrednosti – z spremenljivko – po naslednjem obrazcu:

$$z_{k,n} = \frac{y_i - \bar{y}}{\sigma} \quad (12)$$

za vsako spremenljivko, tako za nebankrotirana kot bankrotirana podjetja. S tako oblikovano z spremenljivko sem nato izločil enote, ki imajo vrednosti standardizirane spremenljivke večje od 3 ali manjše od -3, kar pomeni, da sem se omejil na ± 3 standardne odklone standardizirane spremenljivke. Po opisanih prilagoditvah podatkovne baze, je bilo vključenih 16011 podjetij, od tega 602 podjetji iz skupine 1. Iz Tabele 4 v prilogah je razvidno, da se je po teh preureditvah asimetričnost spremenljivk močno znižala. Razlike vrednosti mere asimetričnosti (ang. *skewness*) pred odstranjevanjem osamelcev so značilne za vsako spremenljivko v obeh skupinah podjetij. Vendar kljub znižanju asimetričnosti, spremenljivke še vedno niso normalno porazdeljene, saj lahko hipotezo o normalnosti še vedno zavrnemo pri točni stopnji značilnosti $p=0,000$, za vse spremenljivke (Tabela 5, priloge). To je posledica dejstva, da imam v podatkovni bazi veliko spremenljivk, ki imajo relativno veliko število

podjetij z vrednostjo, ki je enaka 0 ali zelo blizu 0 (tudi negativne vrednosti). Zaradi nenormalnih porazdelitev, bi moral spremenljivke prilagoditi / normalizirati z logaritemsko transformacijo, vendar logaritem pri vrednosti 0 in pri negativnih vrednostih argumenta ni definiran in bi zato izgubil relativno veliko število originalnih podatkov. Poleg tega sem opravil diskriminantno analizo na velikih vzorcih ($n > 100$), tako da sem lahko spregledal kršitev predpostavke o normalnosti. Zato spremenljivk nisem logaritmiral. Analizo sem torej opravil na originalni (nelogaritmirani) bazi, ki sem jo prilagodil zgolj za osamelce / ekstremne vrednosti, na katere je diskriminantna analiza močno občutljiva. Podatke za bankrotirana podjetja sem variiral, in sicer za oddaljenost od trenutka stečaja. Pričakujem, da se bo napovedna moč obeh modelov (diskriminantna analiza in logistična regresija) postopoma slabšala z oddaljevanjem od leta stečaja, ker so značilnosti stečaja, ki se kažejo v vrednostih kazalnikov, že bolj zabrisane. To se kaže v manjših povprečnih razlikah kazalnikov, kar pomeni manjšo moč diskriminacije za neodvisne spremenljivke. Ne glede na rezultate, sem se odločil za uporabo podatkov do največ tri leta pred stečajem, za skupino 1.

5.3 Uporabljene spremenljivke v analizi

Pri diskriminantni analizi in tudi pri logistični regresiji igra ključno vlogo izbor neodvisnih spremenljivk, na podlagi katerih poteka razvrščanje in / ali napovedovanje skupine, v katero dejansko pripada analizirano podjetje. Če hočemo dobiti karseda učinkovit model, to je, čim višji delež pravilno razvrščenih enot v vsako izmed skupin, morajo neodvisne spremenljivke zadoščati določenim kriterijem. Obsežen del literature se ukvarja s študijami, ki so posvečene zgolj iskanju primerne kombinacije neodvisnih spremenljivk, ki bi zagotovile kar najvišjo stopnjo pravilnosti razvrščanja enot v skupine (Tabachnick in Fidell, 2000). Vendar do sedaj v literaturi še ni splošno sprejete metode, ki bi zadostovala, da bi na podlagi rezultatov določili pravo / optimalno kombinacijo spremenljivk z vidika čim večjega deleža pravilno razvrščenih enot, v vsako izmed skupin. V računalniških statističnih paketih so sicer že vgrajene metode avtomatičnega izbiranja spremenljivk, npr. *Stepwise* metoda v SPSS paketu. Problem te metode je v tem, da kot kriterij uvrstitve / izločitve spremenljivke uporablja maksimiranje diskriminantnega kriterija λ , v katerem uporablja kot argumenta variabilnost spremenljivke, in sicer tisto znotraj skupin in tisto, ki se pojavlja med skupinami. Varianca je lahko v določenih primerih tudi pristranska cenilka in sicer v primeru, ko je med podatki prisotna heteroskedastičnost. Problem je v tem, da se ti kriteriji do določene mere navezujejo na predpostavke (normalnost, itd.), ki pa so v realnosti redko povsem izpolnjene. Zaradi tega so takšna integrirana orodja uporabna v omejenih primerih, če imamo seveda v mislih tudi cilj, ki je izbor optimalne kombinacije spremenljivk za analizo.

Omenjeni razlogi posledično pomenijo, da *stepwise* metoda ne zagotavlja nujno najboljšega izbora spremenljivk, ampak samo ob določenih pogojih. Ob upoštevanju pravkar navedenih dejstev, sem se za izbor najbolj primernih spremenljivk odločil uporabiti metodo, ki je v onovi sestavljena iz dveh korakov, idejo pa sem črpal iz članka avtorjev Brezigar-Masten in Masten (2007).

Na samem začetku sem najprej izločil podjetja, ki se ukvarjajo s finančnimi posli. Na stečaj takšnih podjetij vplivajo drugačni dejavniki, kot na stečaj podjetij, ki jih analiziram v tej nalogi. V bistvu je posledica ostalih (nefinančnih) podjetij velikokrat tudi razlog za povečanje verjetnosti stečaja finančnih podjetij. Od vseh začetnih kazalnikov so trije (3) mere likvidnosti, enajst (11) jih spada v skupino učinkovitosti, sedem (7) v skupino dobičkonosnosti ter sedem (7) v skupino zadolženosti. Odločil sem se, da bom v končnem izboru imel največ pet (5) neodvisnih spremenljivk in najmanj štiri (4). Iz vsake skupine kazalnikov sem v model vključil po eno spremenljivko, da bi pokril kar najširše vsa področja, s katerimi merimo poslovanje podjetja in s tem tudi presojava o njegovi uspešnosti. Če bo katera izmed preostalih spremenljivk po izboru štirih najpomembnejših še dovolj dobra in bo tudi ustrezala pogojem, ki so potrebni predvsem za diskriminantno analizo, to je dovolj nizka korelacija z vsemi že obstoječimi spremenljivkama v modelu (poleg ostalih kriterijev, ki sem jih uporabil v metodi izbora), bom za dano področje uporabil še dodatno spremenljivko. Pri analizah sem si izmenično pomagal s programskimi paketi Stata, SPSS, Soritec ter Excel.

5.4 Metoda izbora najbolj primernih kazalnikov za analizo

Kot že navedeno, spremenljivk nisem normaliziral (logaritmiral), s čimer bi rad potrdil hipotezo, da je model ocenjen z diskriminantno analizo v takšnih okoliščinah (kršenje predpostavk), pri napovedovanju slabši od modela ocenjenega z logistično regresijo. Vseeno se zavedam, da predpostavke v končni fazi niso huje kršene, ker delam z velikimi vzorci, zato bodo majhne razlike med rezultati diskriminantne analize ter logistične regresije razumljivi. Po dvo-stopenjski metodi za določitev najbolj primernih spremenljivk za končni model sem uporabil enake rezultate / enak izbor spremenljivk iz obeh korakov tako za diskriminantno analizo kot tudi za logistično regresijo. V osnovi je opisani postopek za določitev najboljše kombinacije spremenljivk sestavljen iz Faktorske Analize, Logistične Regresije ter iz Analize Glavnih Komponent.

5.5 1. korak-Faktorska analiza

Prvi korak pri določevanju najboljše kombinacije spremenljivk za nadaljno analizo je faktorska analiza (ang. *Factor Analysis*), ki pomaga določiti osnovna / širša področja, ki jih definira več skupin spremenljivk v začetni množici. Gre za to, da poskušamo najti eno ali več prikritih lastnosti, ki jih neposredno ne moremo meriti, domnevamo pa, da prav ti faktorji povzročajo spremembe v spremenljivkah, ki jih lahko opazujemo. Ima pa faktorska analiza tudi svoje omejitve. Ker se lastna vrednost (ki določa kolikšen delež variabilnosti podatkov pojasni posamezni faktor) računa kot determinanta enačbe lastne vrednosti $AX = \lambda X$, izraz oz. matrika $A - \lambda$ ne sme vsebovati proporcionalne vrstice, kar pomeni, da neodvisne spremenljivke ne smejo biti medsebojno linearno odvisne. V nasprotnem primeru ni mogoče določiti faktorjev, ker ne bi bilo mogoče izračunati determinante omenjene matrike, kar je pogoj za določitev rešitve enačbe lastne vrednosti. V faktorsko analizo sem vključil največje možno število spremenljivk, ki mi je še omogočalo izračun faktorjev (matrika A je bila še

vedno pozitivno definitna). V končno analizo naj bi vključili le tiste faktorje, ki imajo lastno vrednost večjo ali enako 1. To pomeni, da vsak faktor pojasni vsaj toliko variabilnosti kot vsaka posamezna spremenljivka. Glavni cilj tega koraka metode izbora, je izločiti tiste faktorje ter posledično spremenljivke, ki ne uspejo pojasniti dovolj variabilnosti med podatki in jih zato lahko v nadaljni analizi zanemarim. Upoštevajoč pravkar navedeno, sem obdržal prvih deset (10) faktorjev, znotraj vsakega pa spremenljivke, ki imajo dovolj visoke vrednosti uteži, in sicer večje od 0,5 po absolutni vrednosti.

Iz Tabele 6 v prilogah sledi, da prvih deset faktorjev pojasni več kot 70% celotne variabilnosti podatkov. Iz Tabele 7 v prilogah je razvidno, katere spremenljivke znotraj faktorjev sem uporabil v nadaljni analizi (osenčeno z rdečo barvo). Ostali faktorji (spremenljivke) pojasnijo le okrog 30% variabilnosti podatkov. To pomeni, da vsaka od preostalih petnajst (15) spremenljivk v povprečju pojasni le 2% celotne variabilnosti podatkov, kar smatram kot zanemarljivo, zato sem te spremenljivke izločil iz nadaljne analize. Kot vodilo pri metodi izbora sem uporabil tudi grafično analizo, v kateri sem primerjal razlike aritmetičnih sredin za vsako izmed spremenljivk. V analizo sem zajel vse enote, ki so ostale v vzorcu po osnovnih transformacijah na spremenljivkah ($n_0=15409$, $n_1=602$; n_g =št. podjetij, $g=0,1$). Kot je razvidno iz Slike 3 v prilogah, so povprečne vrednosti za nekatere spremenljivke močno različne med skupinama 0 in 1, za nekatere spremenljivke pa so razlike zanemarljive. Razlike sem izračunal kot absolutno odstotno razliko med povprečnimi vrednostmi. Pri tem sem negativne razlike transformiral v pozitivne, kar mi v tem primeru ne pomeni izgube informacije. Na podlagi te slike pričakujem, da se v končni model ne bodo uvrstile spremenljivke $X12$ in $X21$, za katere znaša razlika med povprečnimi vrednostmi v skupinah manj kot 10% (nizka pa je tudi razlika med spremenljivkami $X1$, $X14$ ter $X20$), medtem ko za spremenljivki $X22$ in / ali $X26$ pričakujem, da se bosta uvrstili v končni model. V nadaljevanju sem opravil testiranja, Slika 3 pa mi je služila kot dodatni kriterij pri izboru neodvisnih spremenljivk.

5.6 2. korak-Logistična regresija in Analiza glavnih komponent

Namen prvega koraka je bila izločitev tistih spremenljivk, ki ne doprinašajo dovolj k diskriminaciji med podjetji. V drugem koraku sem se posvetil končnemu izboru in ob tem poskušal upoštevati istočasno več kriterijev. Z upoštevanem več kriterijev hkrati sem poskušal povečati robustnost končno izbrane kombinacije spremenljivk, saj je bila odločitev sprejeta ob upoštevanju izpolnjevanja čimvečjega števila omenjenih kriterijev, ki so medsebojno komplementarni z vidika diskriminacije med skupinama. Drugi korak je sestavljen iz logistične regresije ter iz metode glavnih komponent (ang. *Principal Component Analysis*). Logistična regresija je bistveno manj občutljiva na kršenje predpostavk pod 2, ki veljajo za diskriminantno analizo. Z logistično regresijsko funkcijo sem za vsako spremenljivko testiral na bivariatni ravni, kako dobro uspe razvrstiti podjetja v skupino 0 oz. 1. Z besedo 'dobro' mislim na število podjetij, ki so na podlagi posamezne spremenljivke pravilno razvrščena v vsako izmed skupin. Med pomožnimi testiranjmi sem opazil, da so rezultati močno občutljiv na variiranje števila podjetij v obeh skupinah. S spreminjanjem strukture vzorca (razmerja števila

podjetij v vsaki izmed skupin), sem opazil drastično poslabšanje deleža pravilno razvrščenih podjetij v skupini 1. Ker hočem oceniti model, ki bo učinkovit in hkrati karseda robusten / stabilen, sem se odločil, da v nadaljevanju opravim testiranja za končni izbor v 4 različnih scenarijih. Vsak scenarij predstavlja drugačno strukturo vzorca, pri enaki strukturi pa sem opravil 4 ponovna testiranja.

V prvem scenariju sem za vsak test uporabil razmerje števila podjetij v skupinah $r=50:50$ oz. simbolično $n_0=n_1$, kjer je $r=n_0:n_1$. V vsako skupino sem uvrstil po 300 podjetij, kar predstavlja približno 50% vseh podjetij v skupini 1. V drugem scenariju sem določil razmerje $r=70:30$ ($n_1=300$), tretji scenarij je upošteval razmerje $r=80:20$ ($n_1=300$), četrti pa razmerje $r=90:10$ ($n_1=300$). Po teh testiranjih, sem opravil analizo sprememb kvalitete razvrščanja v odvisnosti od spreminjanja razmerja r . Za vsako spremenljivko sem izračunal varianco ter povprečje odstotnega deleža pravilno razvrščenih podjetij v vsaki izmed skupin, upoštevajoč vse štiri scenarije. Manjša variabilnost ter višja povprečna vrednost pomenita boljšo diskriminantno moč spremenljivke. Rezultati, prikazani na Slikah 4 in 5 v prilogah kažejo, da se razlike v deležu pravilno razvrščenih podjetij med prvim ter vsakim od preostalih treh scenarijev ne spreminjajo enako. Za vsako spremenljivko so torej v vsaki skupini 3 razlike, cilj pa je, da so le-te čim bolj podobne, kar pomeni nizko občutljivost spremenljivke na spreminjanje strukture vzorca. Razlike med prvim in drugim scenarijem se za vse spremenljivke v povprečju v skupini 0 povečajo za 38 o.t., med prvim in tretjim scenarijem za dodatne 3 o.t., med prvim in četrtim pa se razlike znižajo za 5,4 o.t. V skupini 1 se razlike med prvim in ostalimi scenariji povečujejo v zaporedju 52, 3,5, ter 2,2 o.t.

Razlog za tako nesorazmerno povečevanje omenjenih razlik med obema skupinama je naslednji: s tem ko večam delež nebankrotiranih podjetij in znižujem delež bankrotiranih podjetij v celotnem vzorcu, so ocenjeni parametri dejansko bolj pod vplivom podatkov za nebankrotirana podjetja. Že osnovni model ni razvrstil vseh podjetij pravilno v obe skupini. Ko večam delež nebankrotiranih podjetij ter obenem nižam delež bankrotiranih, večam verjetnost, da se bo pojavilo nebankrotirano podjetje, ki ima podobne lastnosti kot bankrotirano, posledično pa tudi podjetja, ki so dejansko v skupini bankrotiranih, pridejo po lastnosti bližje (novi – večji) množici nebankrotiranih. S tem manjšamo verjetnost, da bomo ob ocenjenih vrednostih parametrov pravilno razvrstili nebankrotirano podjetje, kar velja tudi za bankrotirana podjetja. Meja med podjetji postane dejansko manj jasna. Tabele 8 in 9 v prilogah kažejo slabšanje rezultatov v primerjavi s prvim scenarijem ($r=50:50$).

Metoda glavnih komponent poskuša določiti tiste medsebojno linearno neodvisne komponente (osi), v tej nalogi bodo to maksimalno štiri, ki bodo zajele karseda velik del variiranja prvotnih podatkov. Večji del variiranja podatkov kot zajamejo glavne komponente, večja je kvaliteta končnega modela, kar pomeni boljšo napovedovalno moč modela. Če imamo n spremenljivk, pomeni, da lahko podatki variirajo v n dimenzijah. Procedura pri določevanju glavnih komponent je takšna, da prva glavna komponenta kaže v takšni smeri, da pojasni kar največji delež variiranja prvotnih podatkov, druga komponenta pojasni največji

delež preostalega variiranja in tako naprej. Za ilustrativni prikaz glej Sliko 6 v prilogah. Cilj je postaviti takšno prvo komponento (os), ki bo pojasnila kar največji del variiranja. Kot je razvidno iz Slike 6 (priloge), bo prva komponenta $X1'$ potekala v $X1$ - $X2$ ravnini in bo imela smer iz izhodišča pod kotom cca 30° . To je smer variiranja, ki vsebuje najvišji delež celotnega variiranja podatkov. Druga os, ki bo pravokotna na $X2'$, bo os $X3'$, bo pojasnila najvišji delež preostalega variiranja. Če bi hoteli pojasniti celotno variiranje podatkov, bi na Sliki 6 postavili še tretjo os, ki bi bila pravokotna na $X1'$ in $X3'$. V bistvu se vidi, da gre za rotacijo koordinatnega sistema, linearno preslikavo, v takšno smer, da bo vsaka izmed osi pojasnila kar največji delež variiranja. Z rotacijo je možna redukcija podatkov, saj obstaja v prvotnem sistemu (vedno) med podatki neka mera kovariabilnosti. Rotacija mora biti takšna, da se kovariabilnost podatkov v novem, zarotiranem, sistemu čim bolj žniža. Ker pa ostane celotna variabilnost enaka, logično sledi, da bodo nove, zarotirane (glavne) osi zmožne pojasniti večji del variabilnosti, kar sledi iz naslednjega obrazca za linearno transformacijo: $X1' = X1 * \cos \varphi + X3 * \sin \varphi$ in $X3' = -X1 * \sin \varphi + X3 * \cos \varphi$. Prva $X1'$ komponenta-os bo torej zmožna pojasniti $X3 * \sin \varphi$ več variabilnosti kot prvotna $X1$ komponenta, kjer je φ korelacijski koeficient med spremenljivkama $X1$ in $X3$ v prvotnem-nerotiranem sistemu. Zaradi tega sledi, da bo nova spremenljivka $X1'$ pojasnila več variabilnosti kot $X1$ sama. Če pa lahko vsaka komponenta, oz. vsaj nekaj novih pojasni več variabilnosti kot prvotne, potem sledi, da lahko pridemo do zadovoljive pojasnitve variabilnosti z bistveno manjšim številom spremenljivk (osi) kot pri podatkih v prvotnem (nerotiranem) sistemu.

Z metodo glavnih komponent sem neodvisne spremenljivke v drugem koraku metode izbora razvrstil v štiri vsebinsko homogena področja. Bartletov test sferičnosti (Tabela 10, priloge), ki testira hipotezo, da je korelacijska matirka enotska matrika, kaže značilne razlike ($p=0,000$), kar pomeni, da je ocenjen model primeren za nadaljno analizo oz. da je združevanje spremenljivk v homogene skupine smiselno. V isti tabeli tudi Kaiser-Meyer-Olkin-ova mera ustreznosti vzorca (0,582) potrди primernost vzorca za analizo s to metodo (če je nižja kot 0,500 vzorec ni primeren za analizo s to metodo). Kot je razvidno iz Tabele 11 v prilogah, prve štiri komponente pojasnijo skoraj 67% celotne variabilnosti vseh (13) spremenljivk vključenih v analizo. Za jasnejšo sliko sem uporabil pravokotno rotacijo Varimax. Po tej rotaciji je struktura po glavnih komponentah za spremenljivke takšna, da ima vsaka spremenljivka visoko vrednost uteži na eni sami komponenti. Za nadaljno analizo sem upošteval tiste spremenljivke, ki imajo največjo utež višjo od 0,5 po absolutni vrednosti.

5.7 Končni izbor najboljše kombinacije spremenljivk

V tem razdelku opisujem izbor spremenljivk, na podlagi kriterijev navedenih v prejšnjih razdelkih tega poglavja. Zaradi nazornosti prikazujem v Tabeli 1 (str. 21) rezultate opisane metode izbora. V prvi koloni so navedene spremenljivke, ki so bile predmet metode izbora (po opravljenem prvem koraku). Druga kolona (razdeljena v dve podkoloni) pomeni variabilnost deleža pravilno razvrščenih podjetij za vsako skupino, tretja kolona (prav tako razdeljena v dve podkoloni) pomeni povprečno vrednost deleža pravilno razvrščenih podjetij

v vsaki skupini (v odstotkih), oboje ob upoštevanju vseh 4 scenarijev. Četrta kolona kaže najvišjo utež, ki jo ima posamezna spremenljivka, glede na vse štiri glavne komponente (Tabela 12 priloge), peta kolona pa ponazarja vsebinski pomen komponente, na katero se nanaša najvišja utež spremenljivke v četrti koloni. V zadnji koloni je še vrednost F testa.

Tabela 1: Rezultati testov izbora optimalne kombinacije spremenljivk

	Variabilnost		Pravilno razvrščenih (v %)		Najvišja utež	Komponenta	F test
	0	1	0	1			
Sprem./Skupina	0	1	0	1	-	-	-
X1	456,9	609,8	85,9	17,6	0,690	Učinkovitost	44,2
X6	296,9	775,7	87,7	17,8	0,832	Likvidnost	151,7
X7	278,7	1381,1	88,5	27,3	0,944	Dobičkonosnost	251,9
X9	423,0	684,5	86,6	16,9	0,519	Učinkovitost	93,7
X11	157,9	901,6	86,5	30,0	0,797	Likvidnost	137,7
X12	544,2	549,2	84,7	16,9	0,775	Učinkovitost	2,2
X14	576,5	1016,6	88,1	19,3	-0,163	Zadolženost	277
X15	277,6	1383,6	88,5	27,3	0,931	Dobičkonosnost	214,1
X17	285,1	1392,3	88,5	27,3	0,894	Dobičkonosnost	197,8
X20	513,1	545,2	84,9	17,1	0,867	Učinkovitost	19,1
X21	488,1	516,1	84,8	16,9	0,685	Likvidnost	0,1
X22	136,1	377,7	89,8	46,4	0,963	Zadolženost	1206,2
X26	125,7	391,4	90,5	46,3	-0,945	Zadolženost	1062,4

Vir: Ajpes; Lastni izračuni.

V prvi glavni komponenti, ki ponazarja dobičkonosnost in pojasni 21,9% variabilnosti podatkov (Tabela 11, priloge), imajo dovolj visoko utež naslednje spremenljivke: X7 (ROA), X15 (dobiček / dolg) in X17 (dobičkonosnost prodaje). Najvišjo utež ima X7 (Tabela 1), ki je glede na analizo glavnih komponent najboljši predstavnik izmed vseh spremenljivk, ki spadajo v kategorijo dobičkonosnosti podjetja. Ostali spremenljivki iz te skupine imata rahlo nižje uteži. Glede na razlike v povprečnih vrednostih pravilno razvrščenih podjetij v vsaki izmed skupin (Slika 3, priloge) vidimo, da imajo vse tri spremenljivke zelo podobne vrednosti razlik med povprečji v skupinah 0 in 1, prav tako pa imajo vse tri visoke vrednosti F testa. Dejansko sta spremenljivki X15 in X17 ob upoštevanju vseh kriterijev v Tabeli 1 zelo podobni, ima pa X17 najnižje korelacije z ostalimi (končno) izbranimi spremenljivkami (Tabeli 13 in 14, priloge). Zaradi navedenih razlogov sem X17 izbral kot predstavnika dobičkonosnosti podjetja v nadaljnji analizi.

Druga glavna komponenta ponazarja učinkovitost podjetja pri gospodarjenju s sredstvi in pojasni 17,5% celotne variabilnosti podatkov. Določajo jo spremenljivke X1 (prihodki / sredstva), X9 (produktivnost), X12 (obrat terjatev do kupcev) in X20 (koeficient obračanja

obratnih sredstev). Najvišjo utež ima *X20*, vendar ima nizko vrednost F testa glede na ostale spremenljivke in je hkrati v povprečju med slabšimi pri razvrščanju podjetij v skupini 0, kjer variabilnost v povprečju glede na štiri scenarije med višjimi glede na ostale tri spremenljivke, zato sem jo izločil iz nadaljnje analize. Naslednja po velikosti uteži v analizi glavnih komponent je *X12*, vendar ima tudi ta spremenljivka prenizko vrednost F testa ($F=2,2$). Hkrati *X12* najbolj variira med scenariji pri razvrščanju v skupini 0, zato sem jo izločil iz nadaljnje analize. Tretja najmočnejša v drugi komponenti je spremenljivka *X1*. Ima dovolj visoko vrednost F testa ($F=44,2$) in ima tudi relativno nizko variabilnost pri razvrščanju glede na štiri scenarije pri analizi v drugem koraku v obeh skupinah. Glede na vse opisano, sem spremenljivko *X1* uporabil kot indikator učinkovitosti podjetja. *X9* sem izločil, ker ima najnižjo utež pri analizi glavnih komponent, poleg tega pa rezultati glede na vse scenarije v skupini 1 najbolj variirajo. Spremenljivka *X1* je bila tudi v modelu, ki ga je ocenil E. Altman (1968), ki pravi, da daje ta spremenljivka v bivariatni analizi slabše rezultate kot v multivariatni. Hkrati sem z dodatnimi testi preveril vpliv izločitve *X1* ter nadomestitve z ostalimi iz iste skupine. Izkaže se, da ta spremenljivka najboljše razvršča podjetja, poleg tega pa je tudi priporočena s strani omenjenega avtorja in je med najbolj popularnimi v literaturi pri uporabi v diskriminantni analizi s finančnimi kazalniki.

Tretja glavna komponenta ponazarja zadolženost podjetja in pojasni 14,1% variabilnosti podatkov. Predstavljata jo spremenljivki *X22* (delež kratkoročnih obveznosti v celotnih sredstvih) in *X26* (kapital / celotni dolg). Spremenljivka *X26* ima negativno predznačeno utež, kar je tudi smiselno, saj naraščanje kapitala negativno vpliva na delež dolga. Vrednost F testa je pri obeh zelo visoka. Razlog je tudi asimetrična porazdelitev obeh spremenljivk. Glede na razlike v povprečnih vrednostih (Slika 3 priloge) se je bolje odrezala spremenljivka *X26*, pri analizi z logistično regresijo pa sta se odrezali podobno dobro. Menim, da je *X26* boljša in sem jo tudi uvrstil v končni model. To so mi potrdila tudi dodatna testiranja z diskriminantno analizo v kombinacijami z ostalimi spremenljivkami.

Četrta glavna komponenta ponazarja likvidnost podjetja in pojasni 13,3% variabilnosti podatkov. V tej komponenti so najmočnejše zastopane spremenljivke *X6* (kratkoročni koeficient), *X11* (hitri koeficient) ter *X21* (delovni kapital / sredstva). Najvišjo utež ima v tej komponenti *X6*, ki se glede na Sliko 3 v prilogah izkaže malenkost slabše kot *X11*, hkrati pa *X11* tudi občutno manj variira pri rezultatih v skupini 0 ter v povprečju veliko bolje razvršča v skupini 1. Spremenljivka *X21* ima izredno nizko vrednost F testa ($F=0,1$). V skladu z navedenim sem izbral spremenljivko *X11* kot predstavnika likvidnosti podjetja. Po opravljeni dvo-stopenjski metodi izbora optimalne kombinacije sem v nadaljnji analizi uporabil naslednje neodvisne spremenljivke:

- *X1* (prihodki / sredstva); kot predstavnika učinkovitosti podjetja,
- *X11* (hitri koeficient); kot predstavnika likvidnosti podjetja,
- *X17* (dobičkonosnost prodaje); kot predstavnika dobičkonosnosti podjetja,
- *X26* (kapital / celotni dolg); kot predstavnika zadolženosti podjetja.

Kot že prej navedeno, sem za te spremenljivke na celotni bazi (v kateri sem upošteval podatke 1 leto pred stečajem) opravil še test medsebojne koreliranosti. Na kratko bi opisal, zakaj je to pomembno. Vsaka spremenljivka nosi s seboj neko količino informacije (npr. zadolženost). Ta informacija je pomembna za diskriminacijo med skupinama 0 in 1. Če je katera izmed ostalih spremenljivk v modelu statistično značilno pozitivno korelirana s to spremenljivko, potem nova spremenljivka ne bo bistveno doprinesla k diskriminaciji med enotami skupin 0 in 1, ker je večino vpliva pobrala že prvotna spremenljivka. Druga je odvečna in jo je bolje nadomestiti z neko tretjo spremenljivko, ki ni korelirana s prvotno (ali pa je negativno korelirana) in zato nosi s seboj neko novo informacijo (npr. uspešnost), ki lahko diskriminira med enotami skupin v novi smeri. Ta smer je bodisi pravokotna na prvo, kot α med spremenljivkama pa je lahko tudi večji kot 90° . Če označimo spremenljivki z $X1$ in $X2$, potem bi bili v primeru da je kot α med osjo $X1$ in linearnim trendom ($X2$ v odvisnosti od $X1$) enak 90° spremenljivki nekorelirani, ker bi bil produkt $X1 * \sin \alpha$, kar je enako spremembi $X2$ ob spremembi $X1$, enak 0. Enostavno povedano: skalarni produkt dveh pravokotnih vektorjev je enak 0. Če pa je kot α topi kot, potem je produkt $X1 * \sin \alpha$ manjši od 0 (funkcija \cos je v drugem kvadrantu negativna). To pomeni, da je vsota križnih produktov negativna, ali kovariabilnost je negativna, zato je tudi koeficient korelacije negativen. Iz Tabel 13 in 14 v prilogah je razvidno, da so spremenljivke ali šibko pozitivno korelirane, za obe skupini, ali pa so celo negativno korelirane. Najvišja pozitivna (značilna) korelacija znaša 0,331 v skupini 0. Na tej točki sem tudi preveril, ali je katera izmed vseh ostalih spremenljivk mogoče močno negativno korelirana / nekorelirana s štirimi, ki so že v modelu in se hkrati tudi dobro odreže glede na ostale kriterije. Vendar nisem našel nobenega novega kandidata, ki bi ustrezal vsem kriterijem in bi se ga splačalo vključiti v analizo. Za preizkus sem opravil tudi analizo s *stepwise* metodo v SPSS paketu, kot kriterij za izbor sem izbral 'Wilks Lambda'. Spodnjo mejo za vključitev sem podal 0,001 in 0,01 za izključitev. Program je v model uvrstil vseh 13 spremenljivk. Izboljšanje je bilo le malenkostno (v skupini 0 je bil rezultat celo za 2,2 o.t. slabši, v skupini 1 pa je bil rezultat boljši le za 4,7 o.t.).

6 REZULTATI NA PODLAGI DISKRIMINANTNE ANALIZE

6.1 Ocena diskriminantnega modela in razvrščanje podjetij

Diskriminantno analizo sem začel s podatki za eno leto pred stečajem. Delam pod predpostavko o nenormalnih porazdelitvah spremenljivk ter tudi pod predpostavko neenakih varianc med skupinama, kar mi pokažeta Levenov preizkus ter Box-ov test (Tabeli 15 in 16, priloge). Zaradi uporabe relativno velikih vzorcev mi to še vedno zagotavlja, da lahko opravim razvrstitev z diskriminantno analizo korektno / v skladu s predpostavkami. Oba omenjena preizkusa o enakosti varianc sta močno občutljiva na velikost vzorca, tako da lahko tudi glede na zavrnjeno ničelno hipotezo o enakosti varianc vseeno delam pod predpostavko o enakih variancah med skupinama. Kot že navedeno, sem variiral število let od leta stečaja za ista podjetja v skupini 1. Pričakujem, da se bodo rezultati slabšali z večanjem števila let od

stečaja. V prvi fazi diskriminantne analize sem na podlagi vzorca ocenil Fisherjevi diskriminantni funkciji – pojasnjevalni model (ang. *explanatory model*). V drugi fazi sem ocenjen model še testiral na še neuporabljenih podjetjih oz. z ocenjenima funkcijama sem napovedoval pripadnost podjetij skupinam (ang. *predictionary model*). Uporabil sem zunanjo metodo preverjanja / napovedovanja, o čemer podrobneje pišem v naslednjem razdelku. Preden sem ocenil model, sem preveril, kako rezultat variira ob variiranju strukture vzorca, to je, ob spreminjanju razmerja r ($n_0:n_1$). V ta namen sem analiziral štiri možne scenarije in izbral najugodnejšega. Delal sem pod predpostavko enakih stroškov napačne klasifikacije za obe skupini. Zaradi primerljivosti rezultatov sem želel dobiti čimbolj enake rezultate po skupinah. V Tabeli 2 prikazujem rezultate testiranja (ocenjevanja modela) v tem razdelku.

Tabela 2: Razvrščanje podjetij pri različnih strukturah vzorca

	r = 50:50		r = 70:30		r = 80:20		r = 95:5	
Skupina	0	1	0	1	0	1	0	1
Delež pravilno razvrščenih (v %)	57	97,3	69,3	92,3	79,3	80	97,5	43,3
Tehtano povprečje	77,2		76,2		79,4		94,8	

Vir: Ajpes; Lastni izračuni.

Najprej sem uporabil vzorec s strukturo $r=50:50$ (v vsaki skupini enako število podjetij), kar je pogosto uporabljeno razmerje v literaturi (Altman, 1968), ki obsega razvrščanje na podlagi diskriminantne analize, zatem pa še vzorec s strukturo $r=95:5$, kot drugo skrajnost, kar je hkrati tudi dober približek dejanskega razmerja med nebankrotiranimi in bankrotiranimi podjetji v godspodarstvih. S tem sem želel oceniti občutljivost rezultatov po skupinah, glede na strukturo vzorca. V skupini 0 imam na voljo 15409 podjetij, v skupini 1 pa 602 podjetji. Uporabil sem 300 podjetij iz skupine 1 (približno polovico), tako da moram za razmerje 95:5 v skupini 0 uporabiti 5700 podjetij. Iz Tabele 2 (2. in 4. kolona) je razvidno, da je kvaliteta razvrstitve po skupinah med navedenima scenarijema obratna. V scenariju z enakim številom podjetij v obeh skupinah je 43% enot v skupini 0 razvrščenih napačno ter le 2,7% v skupini 1. Za drugi vzorec ($r=95:5$) je v skupini 0 2,5% podjetij razvrščeno napačno, medtem ko ta delež v skupini 1 znaša 56,7%. Iz tega sledi, da je v prvem vzorcu 43% nebankrotiranih podjetij klasificiranih kot (navidezno) bankrotirana, kar predstavlja oportunitetno izgubo za banko v znesku obresti, ki bi jih dobila, če bi odobrila posojilo tudi tem podjetjem. V drugem, najbolj neuravnoteženem, a hkrati realnejšem vzorcu, je 56,7% bankrotiranih podjetij razvrščenih kot (navidezno) normalno poslujoča. To bi v primeru, da banka tem podjetjem odobri posojilo pomenilo izgubo dela ali celotnih terjatev za banko. Ta podjetja so bankrotirala in ne bi bila sposobna poravnati svojih obveznosti do bank, delno ali v najslabšem primeru v celoti. Banka bi torej izgubila del svojih sredstev. Zaradi pravkar opisanih rezultatov sem se odločil še za testiranje vmesnih razmerij, in sicer najprej razmerje $r=70:30$ ($n_1=300$), pri katerem sta deleža pravilno razvrščenih enot za skupno 0 oz. 1 enaka 69,3% oz. 92,3%, kar še vedno pomeni, da banka 30,7% podjetjem ne bo odobrila kredita. V realnosti se izkaže, da bi bila ta podjetja dejansko sposobna odplačevati svoje obveznosti v primeru odobritve in je banka naredila oportunitetno izgubo. 7,7% podjetjem bi odobrila kredite in s tem povzročila izgubo svojih

sredstev, ker so ta podjetja dejansko v skupini bankrotiranih. Za razmerje $r=80:20$ ($n_1=300$) sem dobil 79,3% pravilno razvrščenih v skupini 0 ter 80% v skupini 1. V tem primeru banka okrog 20% svojih terjatev dejansko postavi pod tveganje nevrčila. Enak delež posojil tudi ne odobri normalno poslujočim podjetjem ter naredi izgubo v višini obresti. V skladu s temi rezultati sem v nadaljevanju uporabil vzorec z razmerjem podjetij 80:20 ($n_1=300$). Hkrati bo enakomerna razvrstitev po obeh skupinah služila tudi za primerjavo z logistično regresijo. Pri ocenjevanju diskriminantne funkcije sem predpostavil enake apriorne verjetnosti ($q=0,5$). Glede na rezultate dodatnih testiranj (Tabela 19, priloge), se ob uporabi apriornih verjetnosti, ki odražajo relativne velikosti skupin (80:20), kakovost modela močno poslabša v skupini 1, in sicer za 69 o.t.. Glede na to, je izboljšanje v skupini 0 za 20,7 o.t. majhno. Navedeno je v skladu navedbami v literaturi (Huberty, 1994, str. 112).

Tabela 3: Opisne statistike za izbrane spremenljivke ter izbrani vzorec

Spremenljivka	Skupina 0			Skupina 1		
	Povprečje	Std. odklon	n	Povprečje	Std. odklon	n
Prihodki / Sredstva	1,669	1,122	1200	2,042	2,416	300
Hitri Koeficient	1,367	2,113	1200	0,454	0,363	300
Dobičkonosnost prodaje	0,021	0,033	1200	0,001	0,007	300
Kapital / Celotni Dolg	0,332	0,369	1200	-1,064	3,756	300

Vir: Ajpes; Lastni izračuni.

V Tabeli 3 sem prikazal nekaj opisnih statistik za spremenljivke v izbranem vzorcu, s katerim sem ocenil diskriminantni model. Za 20% podjetij, ki bi jim glede na rezultate modela banka odobrila kredit, čeprav jim ne bi smela, bi moral opraviti dodatne analize. Tako bi lahko izboljšal model, vendar tega nisem napravil, ker to presega vsebino te naloge, katere cilj je predvsem primerjati kvaliteto modela ob danih lastnostih spremenljivk in vzorca. Poleg tega me zanima predvsem napovedovalni del analize. Zanima me, kako dobro bo ocenjeni model uspel napovedati pripadnost podjetij, ki niso bila uporabljena pri ocenjevanju diskriminantnih enačb. Podrobneje o tem pišem v nadaljevanju in v sklepnem delu.

Po oceni diskriminantnega modela ima končna kanonična nestandardizirana diskriminantna funkcija naslednjo obliko:

$$Z = -0,585 + 0,01 * X_1 + 0,188 * X_{11} + 19,37 * X_{17} + 0,402 * X_{26}. \quad (13)$$

Kot je razvidno iz Tabele 20 v prilogah, znašata vrednosti ocenjene kanonične nestandardizirane diskriminantne funkcije v centroidih za posamezni skupini 0,22 za skupino 0 ter -0,89 za skupino 1.

6.2 Metode napovedovanja z ocenjenim modelom

Kot sem že v prejšnjem razdelku omenil, je diskriminantna analiza sestavljena iz dveh korakov. Prvi korak je ocena modela, drugi korak pa preverjanje oz. napovedovanje z

ocenjenim modelom. Ko so enkrat koeficienti diskriminantnih funkcij ocenjeni na podlagi vzorca, statistični paketi večinoma omogočajo tudi preverjanje robustnosti / stabilnosti modela. Bistvo te metode je preveriti, kako dobro uspejo ocenjene diskriminantne funkcije razvrstiti enote, če vstavimo vrednosti spremenljivk izbrane enote v ocenjene funkcije. Notranja metoda preverjanja (ang. *internal method*) se imenuje tako, ker so za testiranje kvalitete modela (napovedovanje) uporabljene enote, ki so bile dejansko že uporabljene pri ocenjevanju modela. Cilj je razvrstiti čim več enot pravilno v vsako izmed skupin oz. doseči čim optimalnejše klasifikacijsko pravilo / funkcijo. V primeru notranje metode sta lahko dobljeni stopnjo pravilnosti razvrstitve v vsako izmed skupin precenjeni z vidika deleža pravilno razvrščenih enot oz. podcenjeni z vidika napake pri razvrščanju. Tako dobljena ocena pravilnosti razvrstitve (ang. *Accuracy Ratio – AR*) ne more biti ocena dejanskega AR (Huberty, 1994, str. 86). Formiranje klasifikacijskega pravila temelji na čim večjem deležu pravilno razvrščenih enot / čim nižjem deležu napačno razvrščenih enot. V tem primeru je uporaba enote, ki je že sodelovala pri optimizaciji omenjenega kriterija (ocenjevanju modela), kot sredstva za preverjanje neoptimalna oz. pristranska. Ne moremo trditi, da je dejansko dobljen AR s takšnim preizkusom tudi resničen AR, ki velja za populacijo. Zaradi pravkar navedenega, sem se odločil za zunanjo metodo preverjanja kvalitete modela (ang. *external method*). Pri tej metodi gre za to, da razdelimo vzorec na dva podvzorca. Prvega uporabimo za ocenjevanje modela, nato pa na tem modelu apliciramo drugi podvzorec. Spomnimo se, da sem pri oceni modela uporabil polovico podjetij iz skupine 1 ($n_1=300$). Preostanek (302) sem jih v skladu s pravkar navedenim uporabil za preverjanje kakovosti modela (napovedovanje).

6.3 Napovedovanje z ocenjenim modelom

Pri oceni modela sem zapisal kanonično nestandardizirano diskriminantno funkcijo (13). Spodaj sem izpisal Fisherjevi diskriminantni funkciji za vsako izmed skupin, ocenjeni na podlagi podatkov eno leto pred stečajem za vzorec z razmerjem enot $r=80:20$. Ocenjeni funkciji imata naslednjo obliko:

$$Z_0 = -2,049 + 0,920 * X1 + 0,375 * X11 + 28,116 * X17 + 0,244 * X26 \quad (14)$$

$$Z_1 = -1,770 + 0,909 * X1 + 0,167 * X11 + 6,710 * X17 - 0,201 * X26 \quad (15)$$

s povprečjema 0,48 oz. 0,12. Predznaki koeficientov so smiselni. Ti dve funkciji sem uporabil pri napovedovalnem delu diskriminantne analize. Na tej stopnji je model še vedno pojasnjevalne narave. Tehtana povprečna stopnja pravilno razvrščenih podjetij za obe skupini je analogna mera koeficientu korelacije v linearnih regresijskih modelih, kjer je odvisna spremenljivka zvezna. V nadaljevanju preverjam kvaliteto ocenjenega modela. Model bo napovedovalne narave. Prvo vprašanje, ki si ga lahko na tej točki zastavim je, ali se podjetja iz obeh skupin glede na izračunane vrednosti Z spremenljivk res statistično značilno medsebojno razlikujejo glede na vrednosti izračunanih Fisherjevih funkcij oz. ali je nadaljnja analiza na podlagi ocenjenega modela sploh smiselna. V ta namen sem opravil analizo vrednosti Fisherjevih diskriminantnih funkcij za obe skupini s t-preizkusom. Rezultati so pričakovani, saj lahko v obeh primerih zavrnem ničelno hipotezo: $H_0 : Z_0 - Z_1 = 0$. Razliki

sta v obeh skupinah statistično značilni: $p_0 = 0,000$ in $p_1 = 0,001$, kar pomeni, da je nadaljnja (napovedovalna) analiza z ocenjenima funkcijama smiselna. Podrobnejši rezultati so prikazani v Tabeli 23 v prilogah. Ustreznost modela potrди tudi dejstvo, da je v skupini 0 razlika Z_0-Z_1 pozitivno značilno različna od 0 ($t=11,59$), medtem ko je v skupini 1 ta razlika negativno značilno različna od 0 ($t=-3,23$). To je v skladu z razlago vsebinskega pomena Fisherjevih diskriminantnih funkcij.

Za napovedovalni del diskriminantne analize sem glede na rezultate pravkar opravljene analize značilnosti razlik Z vrednosti med skupinama uporabil enačbi 14 in 15 kot kriterij za napovedovanje pripadnosti podjetij skupinam. Kot sem že opisal v gornjem poglavju in kot predlaga tudi Frank et al. (1965), sem za napovedovanje na podlagi ocenjenega modela kreiral nove vzorce, ki vsebujejo podjetja, ki še niso bila zajeta v pojasnjevalnem delu diskriminantne analize. S tem sem se (vsaj deloma) izognil prenosu napak v prvotnem vzorcu, ki so lahko posledica različnih dejavnikov (odklon pri slučajnem generiranju vzorca, itd.). Hkrati sem z upoštevanjem še neuporabljenih podjetij izločil tudi napako, ki je povezana z izborom pravilne kombinacije spremenljivk iz prvotnih 28 na končne 4. To je posledica dejstva, da sem analizo, na podlagi katere sem se odločil za končni izbor spremenljivk (logistična regresija), opravljal tudi na majših vzorcih. Dobri rezultati ocenjenega modela sami po sebi še ne pomenijo, da bo model enako učinkovito deloval tudi na celotni populaciji. Zato je uporaba podjetij, ki še niso bila vključena v pojasnjevalni del analize tudi preizkus možnih odklonov zaradi omenjenega procesa izbora spremenljivk oz. je preizkus robustnosti ocenjenega modela.

Testiral sem tri scenarije, in sicer najprej prvega, z enakimi pogoji kot v ocenjenem modelu, kar je $r=80:20$ ($n_1=300$), nato drugega, kot boljšo aproksimacijo dejanskega razmerja med nebankrotiranimi in bankrotiranimi podjetji. Pri slednjem sem uporabil vsa nebankrotirana in bankrotirana podjetja iz prilagojene baze. Pri tem sem izključil vsa tista podjetja (v vsaki izmed skupin), ki sem jih že uporabil pri ocenjevanju pojasnjevalnega modela. Ob upoštevanju pravkar navedenega dobim za drugi scenarij vzorec podjetij s strukturo $r=98:2$ ($n_1=302$). Skupina 0 vključuje 14209 podjetij, skupina 1 pa 302 podjetji. Obe funkciji sem uporabil v programskem paketu Stata in v tem paketu tudi opravil napovedovanje na podlagi ocenjenih Fisherjevih diskriminantnih funkcij. V dejanskem svetu se v gospodarstvih pojavljajo tudi podjetja, ki imajo lastnosti enot, ki smo jih za oceno modela izločili in tako prilagodili bazo. To so osamelci (ang. *outliers*). V napoved pri tretjem scenariju sem vključil tudi vsa podjetja, ki sem jih odstranil na samem začetku (gornjih 1% vrednosti ter ostale osamelce). To so tiste enote, za katere velja, da se glede na vrednosti spremenljivk, od aritmetične sredine standardizirane Z spremenljivke razlikujejo za več kot ± 3 standardne odklone. V tem scenariju dobim strukturo vzorca $r=97,5:2,5$. Rezultati za vsakega od treh ocenjenih scenarijev so prikazani v Tabeli 4 (str. 28). Tabela kaže, da se rezultati pri prvem scenariju ($r=80:20$) in pri tretjem scenariju ($r=97,5:2,5$) po posameznih skupinah znatno razlikujejo. To je bilo tudi pričakovano, saj je diskriminantna analiza močno občutljiva na ekstremne vrednosti (osamelce) atributov / neodvisnih spremenljivk.

Tabela 4: Napovedovanje razvrstitve z ocenjenim modelom

Skupina	r = 80:20		r = 98:2		r = 97,5:2,5	
	0	1	0	1	0	1
Delež pravilno razvrščenih (v %)	78,7	78,8	79,6	82,8	80,3	61,4
Število podjetij – pravilna napoved	-	-	11305	250	-	-
Število podjetij – napačna napoved	-	-	2904	52	-	-
Tehtano povprečje	78,7		79,7		79,8	

Vir: Ajpes; Lastni izračuni.

Prvi scenarij je dokaj nerealen, saj je v slovenskem gospodarstvu razmerje 80:20 bistveno prenizko, medtem ko je tretji neprimeren z vidika občutljivosti diskriminantne analize na ekstremne vrednosti ter zaradi tega suboptimalne rešitve. Drugi scenarij ($r=98:2$) ustreza največ kriterijem, tako s tehničnega kot iz vsebinskega vidika, zato sem ga uporabil v nadaljni analizi. V drugem scenariju je model je relativno najbolj učinkovit v skupini 1, kjer nepravilno razvrsti 17,2% bankrotiranih podjetij (Tabela 4). Temu deležu (napravilno) razvrščenih podjetij pravimo napaka I. vrste (ang. *type I error*). Napačno razvrščenih nebankrotiranih podjetij je 20,4%, čemur pravimo napaka II. vrste (ang. *type II error*). V nadaljevanju sem analiziral podjetja, ki spadajo v območje porazdelitev, ki jih označujemo z omenjenima tipoma napak, to je v območje prekrivanja. Poleg tega sem za napovedovalni model s t-preizkusom opravil analizo statistične značilnosti vrednosti razlik Fisherjevih diskriminantnih funkcij med skupinama. Rezultati, ki so prikazani v Tabeli 5 kažejo, da model statistično značilno pravilno napove pripadnost podjetij v skupine.

Tabela 5: Preizkus značilnosti napovedi z ocenjenim modelom

Spremenljivka	Skupina 0		Skupina 1	
	Povprečje	Std. odklon	Povprečje	Std. odklon
Z0	0,68	1,58	-0,46	2,37
Z1	0,02	1,04	0,25	2,37
$H_0: (Z0-Z1)=0$	vrednost t – preizkusa=41,52*		vrednost t – preizkusa=-3,69*	
$H_0: (Z0-Z1)=0$	p – vrednost=0,000		p – vrednost=0,000	

Vir: Ajpes; Lastni izračuni. *Predpostavljene neenake variance med skupinama.

Ustreznost modela potrди tudi dejstvo, da je v skupini nebankrotiranih podjetij razlika $Z0-Z1$ pozitivno značilno različna od 0. V skupini bankrotiranih podjetij je ta razlika negativno značilno različna od 0, kar je v skladu z vsebinsko razlago Fisherjevih funkcij. To pomeni da ima večina podjetij v skupini 0 vrednost $Z0$ višjo kot vrednost $Z1$, obratno pa je v skupini 1.

6.3.1 Analiza območja prekrivanja

Z vidika banke kot posojilodajalca je najbolj zanimivo območje prekrivanja (ang. *overlapping area*). Polprostora na premici, diskriminantni osi, ki sta levo in desno od tega področja, za banko predstavljata veliko večjo predvidljivost. To pomeni, da ležijo levo od območja

prekrivanja podjetja, ki imajo tipične lastnosti bankrotiranega podjetja in neodobritev posojila takšnemu podjetju za banko predstavlja racionalno odločitev. Glede na rezultate modela se, banka pri takšnih podjetjih odloča racionalno in posojila ne odobri. Prav tako velja za podjetja desno od območja prekrivanja, da imajo tipične lastnosti normalno poslujočega podjetja in odobritev posojila takšnemu podjetju predstavlja za banko relativno nizko stopnjo tveganja v primerjavi s potencialno izgubo terjatev za dani kredit, se pravi racionalno odločitev. Področje, ki se nahaja med tema dvema (območje prekrivanja) je tisto, v katerem so pomešana tako podjetja, ki sodijo v skupino bankrotiranih, kot tudi tista, ki sodijo v skupino nebankrotiranih. Glede na dejanske podatke je nemogoče razbrati, katera podjetja sodijo v katero skupino. Iz ocenjenega modela ne vemo natančno, zakaj so ta podjetja padla v to območje. Na podlagi finančnih kazalnikov teh podjetij je potrebno opraviti dodatna testiranja.

Poskušal sem določiti, katere spremenljivke so v največji meri povzročitelj posledice, da se podjetja nahajajo v območju prekrivanja. V Tabeli 24 (priloge) sem prikazal nekaj opisnih statistik za vsako izmed spremenljivk. V tem primeru so me zanimala le podjetja, ki so bila v vsako izmed skupin napačno razvrščena. To so tista, ki tvorijo napake obeh vrst (I in II). Tabela 6 kaže izračune 95% intervalov zaupanja. V vsaki skupini sem za napačno razvrščena podjetja ($n=98:2$) izračunal intervale zaupanja za vsako spremenljivko v ocenjenem modelu.

Tabela 6: Intervali zaupanja pri napačno razvrščenih podjetjih

	Skupina 0			Skupina 1		
	Pravilno	Nepravilno		Pravilno	Nepravilno	
Sprem.	Povprečje	95% int. zaup.	Št. pod.	Povprečje	95% int. zaup.	Št. pod.
X1	1,66	1,44 – 1,48	222	1,95	1,09 – 1,38	30
X11	1,64	0,46 – 0,48	171	0,36	0,80 – 0,92	14
X17	0,03	0,00058 – 0,00062	15	0,00003	0,004 – 0,008	8
X26	0,43	-0,003 – 0,01	1349	-1,36	0,36 – 0,42	8

Vir: Ajpes; Lastni izračuni.

Iztočnica za analizo območja prekrivanja je struktura diskriminantnih enačb, ki so linearne v parametrih in spremenljivkah. Po drugi strani pa nimamo nobenega zagotovila, da velja med odvisno in neodvisnimi spremenljivkami v ocenjenem modelu res linearen odnos. Na Slikah 8 in 9 v prilogah je evidentno, da gre v skupini 1 za nelinearen odnos med deležem podjetij in vrednostmi spremenljivk $X1$ in $X26$, ki napačno razvrstita največ podjetij (Tabela 6). Zato bi lahko sistematično napako modela pripisali tudi nelinearnemu odnosu med spremenljivkami. Omenil sem že, da so se od nekaj pojavljala podjetja, ki so glede na vrednost določene spremenljivke v trenutku, kazala tipične znake skorajšnjega stečaja, dejansko pa so uspela izboljšati svoje poslovanje in banka je bila zaradi neodobritve posojila takšnim podjetjem ob del zaslužka. Večji problem predstavljajo nepravilno razvrščena podjetja iz skupine 1. Glede na vrednosti določene spremenljivke v trenutku, so ta podjetja videti povsem brez težav. Banka se zgolj na podlagi analize takšnih spremenljivk, ali na podlagi preveč površinske analize, nepravilno odloči in takšnemu podjetju odobri posojilo. Pozneje, v primeru tega

razdelka čez eno leto, se seveda izkaže, da kazalniki podjetja niso odražali dejanskega stanja, ki bi omogočalo, da bi lahko to poravnavalo svoje obveznosti do banke, ki je takšnemu podjetju odobrila posojilo. V Tabeli 6 (str. 29) sem za vsako skupino v kolonah 'Št. pod.' navedel število napačno razvrščenih podjetij, katerih vrednosti spremenljivk padejo v izračunan interval zaupanja (izpisan levo od te kolone). V skupini 1 se najbolj odreže spremenljivka *X1*, ki razkrije kar 30 izmed 52 takšnih podjetij, *X11* razkrije 14 takšnih podjetij, *X17* 8 takšnih podjetij, *X26* pa prav tako 8 podjetij.

Iz Tabele 6 (str. 29) se tudi vidi, da 8 podjetij izkazuje svojo nepravilno razvrstitev z več kot 1 spremenljivko. Kot je razvidno, sem s pomožno analizo uspel odkriti vsa tista podjetja, ki predstavljajo za banko grožnjo pred izgubo dela sredstev (terjatev). V skupini 0 se najbolj odreže spremenljivka *X26*, ki razkrije kar 1349 podjetij izmed 2904 napačno razvrščenih. Druga po številu identificiranih napačno razvrščenih podjetij je *X1*, ki razkrije 222 takšnih podjetij. *X11* razkrije 171 takšnih podjetij, najslabše pa se v tem delu analize odreže spremenljivka *X17*, ki razkrije le 15 napačno razvrščenih podjetij skupine 0.

Spremenljivki *X1* in *X26* torej povečini največ podjetij razvrstita napačno. Glede na Sliki 8 in 9 v prilogah ter v glede na pravkar navedeno lahko utemeljeno domnevam, da obstoječa nelinearna povezava med neodvisnimi ter odvisno spremenljivko negativno vpliva na kvaliteto razvrstitve. Tega v tej nalogi nisem upošteval, vsekakor pa bi to lahko služilo kot dodaten kriterij že pri izboru optimalne kombinacije spremenljivk, kar bi pripomoglo k še večji robustnosti ocenjenega modela. Druga možnost bi bila sprememba specifikacije modela z vključitvijo neodvisnih spremenljivk nelinearnih redov, za katere je ugotovljen nelinearen odnos z odvisno. Dodatna razlaga je podana v sklepnem delu.

6.4 Napovedovanje s podatki dve leti pred stečajem

Do sedaj sem se ukvarjal z diskriminantno analizo s podatki za podjetja v skupini 1 eno leto preden so šla ta podjetja uradno v stečaj. Altman (1968) pravi, da začno podjetja kazati znake poslabševanja kazalnikov že 5 let preden dejansko gredo v stečaj. Eno leto pred stečajem so razlike med kazalniki za podjetja skupine 0 in podjetja skupine 1 največje. V skladu s to razlago se pričakuje, da bo z oddaljevanjem leta, iz katerega jemljemo podatke (za podjetja iz skupine 1), od leta, katerega so šla ta podjetja dejansko v stečaj, model vedno slabše razvrščal. Drugače povedano, vedno več bo mešanja med podjetji obeh skupin in zato se bo tudi delež napačno razvrščenih enot povečeval. Opravil sem analizo za ista podjetja iz skupine 1 (nekaj sem jih izgubil, ker niso bili razpoložljivi podatki), tokrat s podatki dve leti pred stečajem.

V skupini 0 sem uporabil ista podjetja, kot sem jih uporabljal že dosedaj v napovedovalnem delu diskriminantne analize. V Tabeli 25 (priloge) sem prikazal nekaj opisnih statistik za spremenljivke v vsaki izmed skupin. Kot v prejšnjem delu analize sem tudi tu preveril statistično značilnost modela. Kot je razvidno iz Tabele 26 v prilogah, diskriminantni enačbi

14 in 15 ne razvrščata več statistično značilno podjetij v skupino 1. Vrednost t-preizkusa povprečne vrednosti razlike spremenljivk $Z0$ in $Z1$ znaša 0,13 oz. p – vrednost znaša 0,89. To je hkrati tudi v skladu z mojimi pričakovanji o slabšanju kvalitete razvrstitve z oddaljevanjem od leta stečaja. V skladu z rezultati analize za dve leti pred stečajem, nisem opravil še analize 3 leta pred stečajem, saj mi že zadnji model ne razlikuje statistično značilno. S tem sem tudi zaključil diskriminantno analizo in prehajam na logistično regresijo, kjer sem uporabil iste podatke, vendar ob bistveno manj predpostavkah (omejitvah), kot pri diskriminantni analizi.

7 REZULTATI NA PODLAGI LOGISTIČNE REGRESIJE

V tem (drugem) sklopu naloge sem si zadal enak cilj kot v prvem. Znatno del analize do sedaj sem posvetil izboru najboljše možne kombinacije spremenljivk glede na delež pravilno razvrščenih podjetij v vsako izmed skupin. Med drugim sem v razdelku 5.6. uporabil logistično regresijo tudi kot kriterij pri izboru spremenljivk. Pojasnil sem že, da sem uporabil logistično regresijo v metodologiji izbora zato, ker ne predpostavlja normalno porazdeljenih neodvisnih spremenljivk in je zato pomemben del metode izbora spremenljivk za napoved z diskriminantno analizo. Glavni namen te naloge je interpretirati rezultate, dobljene na podlagi diskriminantne analize ter logistične regresije kot (delnih) substitutov v multivariatni analizi. To interpretacijo sem poskušal narediti čimbolj primerljivo. Del primerljivosti sem dosegel tudi z uporabo že kreiranih vzorcev, ki sem jih pred tem uporabil pri diskriminantni analizi.

Z uporabo istih vzorcev sem skušal omogočiti čimbolj enake pogoje za obe orodji z vidika porazdelitvenih / parametričnih lastnosti ter s tem omogočiti interpretacijo razlik v rezultatih, ki temeljijo v večji meri 1) na razlikah v metodoloških zasnovah obeh orodij in 2) na lastnostih, ki jih imajo vzorci / spremenljivke, ki sem jih uporabil v analizi. Poleg pojasnjevanja razlik v rezultatih, sem predstavil tudi nekatere druge značilnosti rezultatov dobljenih z logistično regresijo, pri katerih sem našel pomembne implikacije.

Ob pisanju te naloge sem večkrat prišel do spoznanja, da sta metodi v nekem smislu tudi komplementarni. Iz literature ni težko, razbrati zakaj sta ti dve metodi substituta (Huberty 1994 ter Hosmer et al. 2000). Vendar me je zanimalo, v kakšni povezavi bi lahko bili komplementarni. Glede na predpostavke, ki jih ima vsako izmed teh dveh orodij in glede lastnosti podatkov oz. finančnih kazalnikov, lahko vsekakor trdim, da obe orodji nista enako uporabni v določenih okoliščinah. Pri okoliščinah mislim na porazdelitvene značilnosti podatkov s katerimi razpolagamo. Kot sem že omenil v razdelku 5.2 imajo finančni kazalniki (povečini) skupno značilnost – so porazdeljeni asimetrično v desno (Zulkarnain et al. 2006). Na misel mi je prišla ideja, da bi poskušal obe metodi kombinirati skozi proces določanja najboljšega modela za napovedovanje dogodka (stečaja). V tej nalogi sem podal le okvirne ideje, vsekakor pa se nisem posvetil analizi metodologij, ki je v literaturi niti nisem zasledil. Skozi posamezne korake bi moral sam, kar pa bi bilo preobširno in zato presega vsebinski okvir naslova te diplomske naloge. V tem sklopu sem se najprej posvetil razvrščanju podjetij na podlagi logistične regresije ter napovedovanju stečaja posredno, na osnovi verjetnosti.

7.1 Ocena logit modela in razvrščanje podjetij

Podobno kot pri diskriminantni analizi sem tudi pri logistični regresiji najprej ocenil model – pojasnjevalna analiza – in preveril, kako je s kakovostjo razvrstitve. Poleg tega me je tudi zanimalo, kako je z občutljivostjo rezultatov na spreminjanje določenih parametrov. Da bi omogočil primerljivost interpretacije razlik v rezultatih, sem za oceno modela uporabil isti vzorec kot sem ga v razdelku 6.1 pri pojasnjevalnem delu diskriminantne analize. To je tisti, ki je določen z razmerjem števila podjetij med skupinama 0 in 1 $r=80:20$. Pri logistični regresiji v prvi vrsti ocenjujemo verjetnost nastanka dogodka (stečaja v tem primeru) kot je zapisano pod enačbo (11). Vrednosti verjetnosti stečaja za vsa podjetja v vzorcu so tako porazdeljene na premici. Poiskal sem tisto verjetnost, ki mi razdeli premico, na kateri so porazdeljene verjetnosti stečaja na takšna podprostor, da bo vsak podprostor vseboval kar največ pravilno razvrščenih podjetij. Iščem tisto verjetnost (približek) (ang. *cut-off point*), ki mi bo zagotavljala kar najboljše razvrščanje v vsaki izmed skupin. Privzeta vrednost za verjetnost pri točki reza je 0,5. Na koncu sem tudi preveril, kako se delež pravilno razvrščenih podjetij v vsako izmed skupin spreminja (slabša) z oddaljevanjem od leta stečaja.

Prvi del analize je narejen na podlagi podatkov eno leto pred stečajem (za skupino 1). Ravnal sem se po enakem vodilu kot pri diskriminantni analizi. Za nadaljno analizo sem uporabil model pri tistih parametrih (točka reza), pri katerih mi bo ta zagotavljal relativno najbolj enake rezultate glede pravilno razvrščenih enot v vsako izmed skupin. S tem sem hotel (kolikor točno se da) tudi omogočiti interpretacijo razlik, ki so posledica variiranja ključnih parametrov obeh orodij. Verjetnost, ki mi razdeli populacijo v dve podskupini, sem variiral na intervalu (0,5 – 0,1). Tako sem dobil različne rezultate v odvisnosti od te verjetnosti (v točki reza). Na Slikah 10 – 13 (priloge) sem prikazal verjetnosti stečaja v odvisnosti od vrednosti posameznih spremenljivk za ocenjeni model (16). Iz definicije logistične regresije sledi, da ob ustrezno izbranih pojasnjevalnih spremenljivkah, verjetnost dogodka (stečaja) narašča / pada nelinearno / eksponentno, s spreminjanjem vrednosti teh spremenljivk. Iz teh slik je razvidno, da za spremenljivke $X11$, $X17$ in $X26$ to velja, medtem ko za $X1$ ta odnos ne velja. To bi lahko povezali z nelinearno odvisnostjo deleža podjetij v stečaju od vrednosti spremenljivke $X1$ (Slika 8, priloge).

V Tabeli 7 (str. 33) so prikazani rezultati ocenjenega logističnega regresijskega modela, v Tabeli 27 (priloge) pa sem prikazal tudi izpisek razvrstitve iz paketa SPSS. Kvaliteta razvrščanja je zelo podobna kot pri diskriminantni analizi v fazi ocenjevanja modela. Ocenjeni model logistične regresije (logit transformacija) ima naslednjo obliko:

$$\log\left(\frac{P_i}{1-P_i}\right) = 0,312 - 0,138 * X1 - 1,118 * X11 - 118,365 * X17 - 1,732 * X26. \quad (16)$$

V Tabeli 7 (str. 33) je prikazan izpisek, kjer so izpisane tudi točne stopnje značilnosti koeficientov v modelu (16). Iz predzadnje kolone v tej tabeli sledi, da so vsi koeficienti v ocenjenem modelu, razen konstante, statistično značilno različni od nič. To pomeni, da lahko uporabim model 16 za napovedovanje verjetnosti stečaja v napovedovalnem delu analize. V

Tabela 8 prikazujem kvaliteto rezultatov razvrščanja. Vrednost Cox&Snell $R^2 = 0,315$ (Tabela 28, priloge) pove, da se model 16 srednje dobro prilaga podatkom.

Tabela 7: Ocenjeni logit model logistične regresije

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	x1	-,138	,060	5,311	1	,021	,872
	x11	-1,118	,204	30,131	1	,000	,327
	x17	-118,365	18,864	39,370	1	,000	,000
	x26	-1,732	,225	59,069	1	,000	,177
	Constant	,312	,171	3,337	1	,068	1,366

a. Variable(s) entered on step 1: x1, x11, x17, x26.

Vir: Ajpes; Lastni izračuni.

Iz tretje in četrte kolone v Tabele 8 sledi, da kvaliteta razvrščanja narašča v skupini 1, ko se s točko reza pomikamo k dejanskemu razmerju podjetij v obeh skupinah (80:20). Obratno pa kvaliteta narašča v skupini 0, ko se s točko reza pomikamo proti 'navidezno' uravnoveženima skupinama (točka reza=0,5). To pa je dejansko ravno nasprotno kot gibanje pravilnosti razvrstitve po skupinah pri diskriminantni analizi (ob variiranju apriornih verjetnosti). Primerjave rezultatov med obema orodjema sem pustil za sklepn del. Ocenjen logit model, pri katerem sem uporabil točko reza pri verjetnosti stečaja 0,25, je statistično značilen. Povprečna razlika verjetnosti stečaja med skupinama ($P_0 - P_1$) znaša -38,2 o.t., p – vrednost za to razliko znaša 0,000 (Tabela 29, priloge).

Tabela 8: Kvaliteta razvrstitve podjetij v odvisnosti od točke reza

Model	Točka reza	Delež pravilno razvrščenih (v %)		Tehtano povprečje
	Verjetnost	Skupina 0	Skupina 1	Odstotek
1	0,5	95,3	43,0	84,8
2	0,4	98,1	56,7	89,8
3	0,3	85,3	74,0	83,0
4	0,2	76,2	86,3	78,2
5	0,1	61,3	96,0	68,2
6	0,25	81,6	82,3	81,7

Model 6: $-2 * \log\text{likelihood} = 934,038$, Cox&Snell $R^2 = 0,315$.

Vir: Ajpes; Lastni izračuni.

7.2 Napovedovanje verjetnosti stečaja z ocenjenim modelom

Tako kot sem z izbranimi spremenljivkami napovedal pripadnost podjetij določeni skupini pri diskriminantni analizi, sem tudi pri logistični regresiji napovedal verjetnost stečaja oz. razvrstil še neuporabljen podjetja v skupine. Na podlagi te ocenjene verjetnosti sem preveril, kako dobro mi model 16 razvrsti podjetja v vsako od skupin. Pričakujem, da se bo logistična regresija odrezala bolje kot diskriminantna analiza. Obenem pa ob dejstvu, da pri

diskriminantni analizi nisem huje kršil predpostavk (zaradi uporabe velikih vzorcev), ne pričakujem večje razlike v kvaliteti napovedi. Predvsem pa sem že pri ocenjevanju logit modela nastavil parametre tako, da sem določil točko reza, pri kateri dobim kvalitativno zelo podobne rezultate kot pri diskriminantni analizi. Glede na pravkar navedeno, domnevam, da napovedana pravilnost razvrstitve na podlagi ocenjenega logit modela, ne bo bistveno odstopala od tiste pri diskriminantni analizi. Za napovedovanje sem uporabil isti vzorec podjetij, kot pri napovedovanju z diskriminantno analizo ($r=98:2$), v tem razdelku s podatki eno leto pred stečajem. Točko reza sem določil enako kot pri ocenjevanju modela na verjetnost stečaja 0,25. V Tabeli 9 so prikazani rezultati na podlagi napovedovanja z ocenjenim modelom 16.

Tabela 9: Napovedovanje razvrstitve z ocenjenim modelom

	r = 98:2	
	Skupina 0	Skupina 1
Delež pravilno razvrščenih (v %)	81,4	82,1
Število podjetij – pravilna napoved	11566	248
Število podjetij – napačna napoved	2643	54
Tehtano povprečje		81,4

Napovedi na podlagi modela (16).

Vir: Ajpes; Lastni izračuni.

Rezultati v gornji tabeli kažejo, da ocenjeni logit model (16) v obeh skupinah (približno) enako dobro razvršča podjetja. Rezultati so rahlo boljši kot pri napovedovanju z diskriminantno analizo v razdelku 6.3. Vednar je bilo (vsaj enak) rezultat kot pri diskriminantni analize realno pričakovati.

Tabela 10: Preizkus značilnosti napovedi z ocenjenim modelom

	Verjetnost	
	Povprečje	Std. odklon
Skupina 0	0,1206	0,1553
Skupina 1	0,5059	0,2712
Povprečna razlika (v o.t.)		-38,53
H0: $P_0 - P_1 = 0$	vrednost t preizkusa=-24,6*	
H0: $P_0 - P_1 = 0$	p – vrednost=0,000	

*Predpostavljene neenake variance med skupinama.

Vir: Ajpes; Lastni izračuni.

V Tabeli 10 se kaže, da glede na verjetnosti stečajev vsa podjetja res ne prihajajo iz iste skupine. Povprečna razlika $P_0 - P_1$ znaša -38,53 o.t. in je statistično značilna pri p-vrednosti 0,000. Višina izračunane povprečne razlike za verjetnosti stečaja med podjetji skupin 0 in 1 mi v nekem smislu pove, da sta skupini dokaj polarizirani, glede na vrednosti izbranih finančnih kazalnikov.

7.2.1 Analiza območja prekrivanja

Tudi za rezultate logistične regresije na podlagi napovedovanja eno leto pred stečajem sem ocenil 95% intervale zaupanja spremenljivk za podjetja, ki so bila napačno razvrščena v vsako izmed skupin. Zanima me predvsem, kako grešijo spremenljivke v primerjavi z analizo območja prekrivanja pri napovedovanju z diskriminantno analizo v razdelku 6.3.1. Intervali zaupanja so prikazani v spodnji tabeli. Podrobnejše statistike za nepravilno razvrščena podjetja v vsako izmed skupin so prikazane v Tabeli 30 v prilogah.

Tabela 11: Intervali zaupanja pri napačno razvrščenih podjetjih

	Skupina 0			Skupina 1		
	Pravilno	Nepravilno		Pravilno	Nepravilno	
Sprem.	Povprečje	95% int. zaup.	Št. pod.	Povprečje	95% int. zaup.	Št. pod.
X1	1,68	1,26 – 1,34	467	1,84	1,30 – 2,35	48
X11	1,63	0,43 – 0,45	147	0,37	0,790 – 0,995	26
X17	0,03	0,0003 – 0,0004	73	0,000002	0,0007 – 0,007	5
X26	0,42	-0,006 – 0,026	1575	-1,33	0,232 – 0,386	21

Vir: Ajpes; Lastni izračuni.

Tabela 11 kaže, da spremenljivke grešijo v enakem vrstnem redu (glede števila napačno razvrščenih podjetij) pri napovedovanju stečaja kot pri diskriminantni analizi (Tabela 6, str. 29). V skupini 0 se kot najbolj učinkovita pri zaznavanju števila napačno razvrščenih podjetij izkaže spremenljivka *X26*. Druga najbolj učinkovita je *X1*, tretja *X11*. Najmanj učinkovita je spremenljivka *X17*. Enak vzorec zaznavanja kot pri diskriminantni analizi velja v skupini 1.

7.3 Napovedovanje s podatki dve leti pred stečajem

V tem razdelku preveril, kako je s kakovostjo razvrščanja na podlagi ocenjenih verjetnosti za podatke v skupini 1 dve leti pred stečajem. Rezultati so prikazani v Tabeli 12 (str. 36). Po pričakovanjih se je logistična regresija odrezala bolje pri napovedovanju verjetnosti stečaja na podlagi podatkov skupine 1 dve leti pred stečajem v primerjavi z diskriminantno analizo. Model pravilno napove več kot polovico podjetij v skupino 1. Kot je razvidno iz Tabele 31 v prilogah, povprečna razlika verjetnosti stečaja $P0-P1$ znaša -18,26 o.t., kar predstavlja poslabšanje (znižanje) te razlike za 20,27 o.t. glede na napovedovanje s podatki eno leto pred stečajem.

Razlog za boljšo napovedno moč modela logistične regresije proti diskriminantni analizi je najbrž v dejstvu, da se podatki v povprečju manj razlikujejo, če uporabimo za skupino 1 podatke dve leti pred stečajem. Diskriminantna moč spremenljivk se manjša, zato se napake oz. kršitve predpostavk lahko še bolj poznajo na končnem rezultatu. To je po mojem mnenju tudi razlog za neznačilnost rezultatov pri napovedovanju na podlagi podatkov dve leti pred stečajem pri diskriminantni analizi. Večino razlage razlik sem obravnaval v sklepnem delu.

Tabela 12: Napovedovanje razvrstitve z ocenjenim modelom

	r = 98:2	
	Skupina 0	Skupina 1
Delež pravilno razvrščenih (v %)	81,4	56,5
Število podjetij – pravilna napoved	11566	161
Število podjetij – napačna napoved	2643	124
Tehtano povprečje		80,9

Napoved na podlagi modela (16).

Vir: Ajpes; Lastni izračuni.

8 SKLEPNI DEL-KLJUČNE UGOTOVITVE TER PRIMERJAVE OBEH ORODIJ

V tem (sklepnem) delu diplomske naloge sem najprej povzel rezultate diskriminantne analize ter logistične regresije in primerjal rezultate z ugotovitvami v literaturi, ki sem jo prebral na to tematiko. Pri tem sem najprej obravnaval vsako orodje ločeno, nato pa na podlagi primerljivih rezultatov primerjal obe orodji glede na vrednosti parametrov, ki sem jih v tej nalogi pri vsakem orodju variiral in mi zato v nekem smislu neposredno določajo kvaliteto rezultatov.

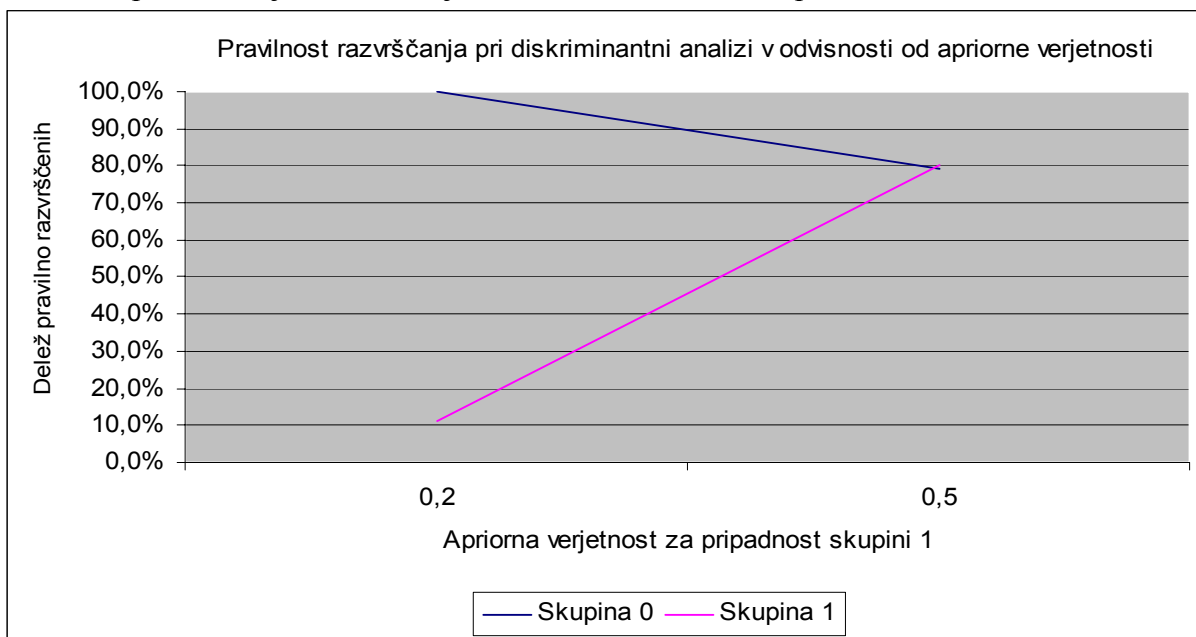
8.1 Diskriminantna analiza

Diskriminantno analizo sem opravil na podatkovni bazi, ki v osnovi ne ustreza prav nobeni izmed ključnih predpostavk (2) tega orodja. V izbrani literaturi nisem zasledil diskriminantne analize z nenormalno porazdeljenimi spremenljivkami, zato sem se odločil, da preverim, kako se to orodje obnese ob kršitvah predpostavk. Edina stvar, s katero sem kompenziral kršitve osnovnih predpostavk, je bila, da sem uporabljal relativno velike vzorce. Literatura večinoma navaja metode, kako se izogniti kršitvam s prilagoditvami / transformacijami spremenljivk, da bi se zagotovilo izpolnjevanje predpostavk. V tej nalogi transformacija, ki odgovarja porazdelitvenim karakteristikam finančnih kazalnikov (logaritemska transformacija), ne bi bila uporabna, ker bi s tem izgubil preveč podatkov, saj imam veliko podjetij z vrednostmi finančnih kazalnikov enako 0 ali celo negativno, kjer pa logaritem ni definiran.

Prva ugotovitev, do katere sem prišel pri diskriminantni analizi je, da ni enostavnega in splošnega pravila, po katerem bi lahko optimirali rezultate. Prva stvar, ki jo je potrebno narediti, je spoznati samo metodologijo. Druga pomembna stvar je dobro poznavanje karakteristike spremenljivk, s katerimi delamo. Preden pričnemo z metodološkim delom, je potrebno oceniti delež enot v populaciji, za katere na podlagi zgodovinskih podatkov velja, da se je dogodek (stečaj) zgodil / se ni zgodil – določiti je potrebno apriorne verjetnosti. Banka ima pri svojih odločitvah (v splošnem) dva možna scenarija, in sicer: lahko povečuje svoj profit ob neupoštevanju ostalih dejavnikov ali pa minimizira svoje tveganje. Ta dva kriterija / interesa sta v nekem smislu konfliktna. Odvisno je od odločitev bankirjev ter od zakonodaje,

kje se bo formirala točka, kjer bo obema kriterijema kar najbolj zadovoljeno. Baselski standardi do neke mere omogočajo bankam, da se izpostavljajo tveganjem, kar povečuje njihov profit, vendar so pri tem omejene. Prav zaradi te omejenosti je smiselno analizirati, kateri dejavniki pri diskriminantni analizi določajo pot do optimalne rešitve. V tem primeru sem to stanje določil kot čimbolj enak delež pravilno razvrščenih podjetij v vsaki izmed skupin (ob predpostavljenih enakih stroških napačne razvrstitve). Na Sliki 1 sem prikazal variiranje kakovosti razvrstitve po skupinah v odvisnosti od ključnega parametra, ki vstopa v enačbo posteriorne verjetnosti (9), to je apriorna verjetnost. Pogostosti pojava (stečaja) v Sloveniji sem (približno) prilagodil relativne velikosti skupin. Pri tem sem upošteval tudi omejitve, ki jih postavlja diskriminantna analiza ter seveda namen analize.

Slika 1: Apriorna verjetnost kot dejavnik kvalitete razvrstitve pri diskriminantni analizi



Vir: Ajpes, lastni izračuni

Če hočem napraviti sklepe glede variiranja kvalitete razvrstitve podjetij v odvisnosti od spreminjanja apriorne verjetnosti, je ključno opazovanje variiranja kvalitete rezultatov po posameznih skupinah. Glede na rezultate, ki jih prikazuje Slika 1, lahko sklepam:

- kadar delamo z neenako velikimi skupinami bodo rezultati v manjši skupini bolj občutljivi na variiranje apriorne verjetnosti;
- v večji skupini bodo rezultati bistveno manj variirali. Na tej točki velja omeniti, da odnos prikazan na Sliki 1 ni nujno linearen, kot je tu prikazano, kar pa ne vpliva ključno na interpretacijo dobljenih rezultatov;
- kadar imamo opravka s spremenljivkami, ki krišijo osnovne predpostavke diskriminantne analize (posebej predpostavko o enakosti varianc), je bolje izbrati enake apriorne verjetnosti za obe skupini ($q=0,5$), ne glede na relativno velikost skupin. Ta sklep je poleg Slike 1 tudi posledica sestave enačbe za izračun posteriorne verjetnosti (9), na podlagi katere poteka razvrščanje v skupine.

Iz zapisa enačbe je razvidno, da poleg razdalje, neenake variance med skupinami (Tabela 15, priloge) vplivajo na vrednost posteriorne verjetnosti. Pri različnih (večjih ali manjših) variancah v skupini 1 glede na skupino 0 (eno leto pred stečajem), bi izbor apriornih verjetnosti v skladu z relativnimi velikostmi skupin (ob neenako velikih skupinah) dejansko znižal vpliv vrednosti produkta $q_g * |S_g|^{-1/2} * \exp(-\frac{1}{2}D_{ug}^2)$, kar lahko zniža vpliv sicer majhne razdalje od centroida ter nizke kovariabilnosti – ki vplivata pozitivno na višino posteriorne verjetnosti v enačbi 9. To pomeni, da bo ob dani razdalji od centroida ter ob dani kovariabilnosti vpliv teh dveh parametrov na posteriorno verjetnost (preko nižjega produkta $q_g * |S_g|^{-1/2} * \exp(-\frac{1}{2}D_{ug}^2)$, dejansko zmanjšan, če za skupino 1 uporabimo $q = 0,2$ namesto $q = 0,5$). To nadalje pomeni nižjo posteriorno verjetnost za pripadnost skupini 1, kar večja verjetnost napačne razvrstitve.

V skladu s pravkar navedenim ter z načelom minimizacije tveganja (ob enakih stroških napačne razvrstitve), se je glede na Sliko 1 (str. 37) rezultat ob uporabi enakih apriornih verjetnosti močno izboljšal v skupini 1. Delež pravilno razvrščenih podjetij v to skupino se je povečal za več kot 6-krat, kar je nedvomno tako statistično kot tudi ekonomsko gledano pomembno povečanje. Na drugi strani je kvaliteta razvrščanja v skupini 0 veliko manj občutljiva na variiranje apriorne verjetnosti, saj se je delež pravilno razvrščenih podjetij ob uporabi enakih apriornih verjetnosti znižal le za 20,7%, glede na pravilnost razvrstitve pri uporabi apriornih verjetnosti, ki odsevajo relativne velikosti skupin. Ta ugotovitev je v skladu z ugotovitvami v literaturi (Huberty 1994, str. 112).

Ne glede na izbiro cilja analize (maksimizacija kvalitete rezultatov v vsaki od skupin ali neenaka maksimizacija kvalitete po posameznih skupinah) se mora analitik zavedati, da je ključna faza izbora najprimernejših spremenljivk, s katerimi bomo pojasnjevali odvisno spremenljivko. Od tega izbora so odvisni vsi rezultati, ki sem jih opisal v tem poglavju. Metoda, ki sem jo uporabil še zdaleč ni optimalna, je pa glede na možnosti, ki jih ponujajo statistični programski paketi daleč najprimernejša. Pri izboru spremenljivk je ključno zajeti karseda širok nabor dejavnikov, ki bi utegnili vplivati na kakovost rezultatov v poznejših fazah analize. Izbrana dvostopenjska metoda se temu dobro približa, kar se kaže tudi v stabilnosti rezultatov pri obeh uporabljenih orodjih. Dodatni testi robustnosti rezultatov na podlagi izvedene dvostopenjske metode bi bili prav tako priporočljivi, še posebej testi heteroskedastičnosti, ki jih v tej nalogi nisem opravljal. Če povzamem, bi glede uporabe diskriminantne analize v primeru dveh skupin rekel takole: Preden pričnemo z analizo, si je potrebno postaviti cilje, ki naj bi jih dosegli z analizo ali krajše povedano, določiti moramo namen rezultatov analize. V vsakem primeru je osnova:

- dobro poznavanje metodologije diskriminantne analize;
- dobro poznavanje (porazdelitvenih) karakteristik spremenljivk / baze s katero delamo.

To analitiku lahko veliko pove o primernosti podatkov glede na omejitve, ki jih postavlja diskriminantna analiza. Od tega je odvisno, na kakšen način se bomo opredelili glede izvedbe analize. Kadar imamo opravka s spremenljivkami, ki ne ustrezajo predpostavkam diskriminantne analize (2), se moramo predvsem zavedati, da je model pri pojasnjevalnem delu analize lahko prav tako dober kot logistična regresija, kar drži ob primerni metodologiji izbora spremenljivk ter uporabi dovolj velikih vzorcev. Pri napovedovanju na ostalem delu populacije se diskriminantna analiza odreže slabše kot logistična regresija, se pravi, da je manj zanesljiva / robustna. V takšnih primerih je potrebno spremenljivke dodatno 'normalizirati' ali celo opustiti diskriminantno analizo ter poseči po logistični regresiji. Ob predpostavljenih enakih stroških napačne razvrstitve je optimalna rešitev dokaj tehničnega značaja, apriorna verjetnost pa je tisti parameter, ki ga lahko variiramo ter s tem določamo kvaliteto razvrstitve.

8.2 Logistična regresija

Logistična regresija, tako ocenjevanje modela kot napovedovanje z ocenjenim modelom, je bila opravljena na istih vzorcih kot diskriminantna analiza. Specifika metodologije tega orodja je podvržena bistveno manj predpostavkam. Bistvena prednost logistične regresije pred diskriminantno analizo je ta, da ne predpostavlja normalnih porazdelitev ter enakih varianc neodvisnih spremenljivk po skupinah.

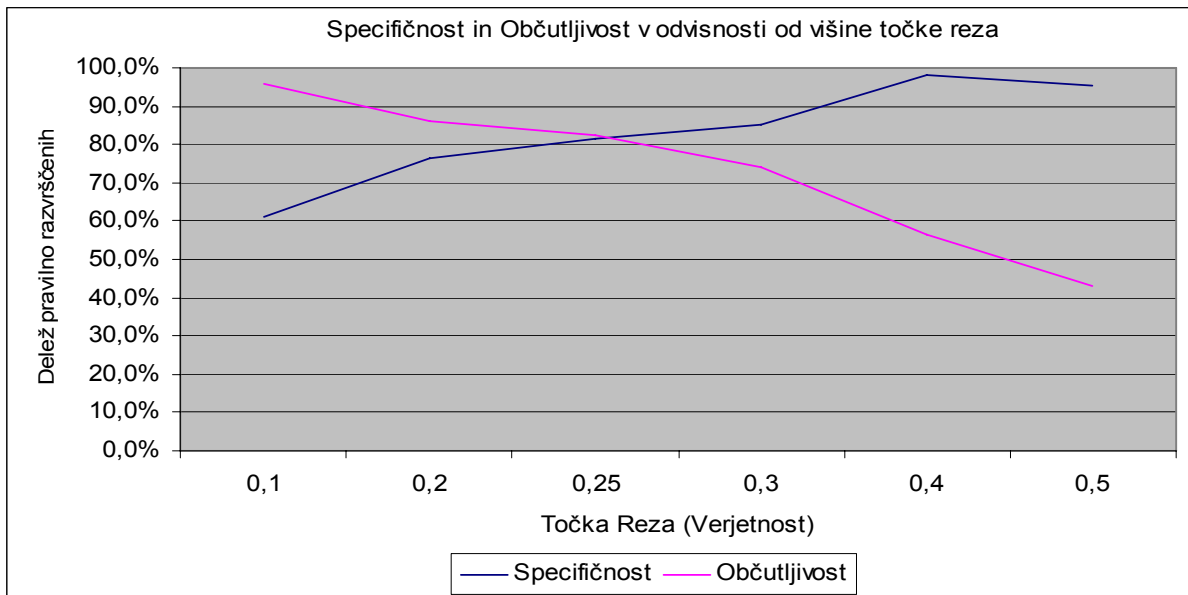
Točka reza je bil tisti parameter, ki sem ga uporabljal kot neodvisno spremenljivko pri določanju optimalne rešitve – čim bolj enakomernega pravičnega razvrščanja v vsaki izmed skupin, ob upoštevanju enakih stroškov napačne razvrstitve podjetij. Analogno kot pri diskriminantni analizi, prikazujem na Sliki 2 (str. 40) variiranje kakovosti razvrščanja (specifičnost ter občutljivost), tokrat v odvisnosti od verjetnosti v točki reza. Splošna zakonitost, ki je zavedena v literaturi (Hosmer, Lemeshow, 2000, str. 162), glede gibanja specifičnosti ter občutljivosti v odvisnosti od točke reza, se je potrdila tudi na vzorcu slovenskih podjetij v tej analizi. Iz Slike 2 (str. 40), se lahko razbere dvoje:

- z oddaljevanjem od referenčne točke reza (verjetnost stečaja 0,25), kjer se specifičnost in občutljivost (skoraj) izenačita, se kvaliteta razvrstitve spreminja; s pomikanjem od referenčne točke reza proti 0 se izboljšuje občutljivost (delež pravilno razvrščenih podjetij v skupini 1) in nasprotno se znižuje specifičnost (delež pravilno razvrščenih podjetij v skupini 0). Če pa se od referenčne točke reza pomikamo proti 0,5 se znižuje občutljivost ter zvišuje specifičnost,
- gibanje občutljivosti in specifičnosti se z oddaljevanjem od referenčne točke reza ne spreminja enakomerno (linearno) temveč nelinearno, praviloma po padajoči stopnji z oddaljevanjem od referenčne točke reza.

Iz Slike 2 (str. 40) pravkar navedeno sicer ni neposredno razvidno. Uporabil sem premajhen korak spreminjanja točke reza (0,10) ter samo 6 različnih nivojev. Praviloma je spreminjanje

rezultatov z oddaljevanjem od referenčne točke reza vedno manjše, sprememba pa se asimptotično približuje 0. To implicira, da kadar smo v točki, ko sta občutljivost in specifičnost izenačeni in seveda razumno visoki, smo ob predpostavljenih enakih stroških napačne razvrstitve podjetij v obeh skupinah v ravnovesnem položaju. V primeru te naloge predstavlja točka reza 0,25, ki sem jo uporabil pri ocenjevanju modela, približen delež bankrotiranih podjetij v vzorcu na katerem sem ocenil model.

Slika 2: Točka reza kot dejavnik kvalitete razvrstitve pri logistični regresiji



Vir: Ajpes, lastni izračuni

Logistična regresija ne predpostavlja normalnih porazdelitev ter enakih varianc, zato sklepam, da je potrebno v vsakem primeru pri logistični regresiji upoštevati relativne velikosti skupin pri določanju točke reza, če se želimo približati ravnovesnemu položaju (višja variabilnost podatkov v skupini 1). Če pa predpostavimo neenake stroške napačne razvrstitve, si lahko enostavno pomagamo s sklepi, ki sem jih izpeljal na podlagi rezultatov, ki jih prikazuje Slika 2 ter s primernim oddaljevanjem od referenčne točke reza določimo nov ravnovesni položaj. To ne predstavlja dejanskega (tehničnega) ravnovesja, vendar so napake v vsaki skupini neenako tehtane, tako da smo v končni fazi lahko v ravnovesju. Predvsem pri bankah, ki se odločajo o odobritvi posojil podjetjem na podlagi analize finančnih kazalnikov, je realno predpostavljati neenake stroške napačne razvrstitve. Če banka deluje racionalno in v skladu z vsemi standardi, ki jo tudi omejujejo glede količine danih posojil, potem je seveda smiselno predpostaviti višje stroške napačne razvrstitve bankrotiranih podjetij (napaka I. vrste). Ob tej predpostavki in ob pogoju ravnovesja se ne bi nahajali v presečišču krivulj specifičnosti in občutljivosti, temveč bi se nahajali levo od te točke. Podrobneje se v analizo takšnih situacij v tej nalogi nisem spuščal.

Naslednja ugotovitev in / ali priporočilo analitiku pri analiziranju z logistično regresijo, ki sledi iz moje analize, je pomembnost referenčne točke reza, kjer sta občutljivost in specifičnost izenačeni. Neodvisno od vrste predpostavljenih stroškov napačne razvrstitve,

podjetij rezultat pri referenčni točki še zdaleč ne pomeni, da smo s tem dosegli tudi zastavljeni cilj. Pomembno je, kako visoki sta občutljivost in specifičnost, kajti od tega je tudi odvisno, koliko zaloge smo pustili za optimizacijo modela. Pri predpostavljenih enakih stroških napačne klasifikacije, bi se vrednost (približno) 80% za občutljivost in specifičnost kot sem jih ocenil v tej nalogi, prav zagotovo dalo izboljšati. To pomeni, da se moramo vrniti na začetek, opraviti dodatne analize in morebiti določiti drugačno kombinacijo neodvisnih spremenljivk ter s tem izboljšati napovedovalno moč modela. Kot se je izkazalo pri napovedovanju z logistično regresijo (Tabela 9, str. 34), so rezultati doseženi ob uporabi kriterija enakosti specifičnosti ter občutljivosti, zelo robustni / stabilni, glede na rezultate pri pojasnjevalnem delu analize eno leto pred stečajem (Tabela 8, str. 33, model 6). To je lahko v pomoč in motivacijo, da posvetimo analizi izbora spremenljivk raje nekaj več kot manj časa. Od tega je močno odvisna napovedna moč ocenjenega modela.

Zadnja implikacija rezultatov logistične regresije (v primerjavi z diskriminantno analizo) pa je, da logistična regresija, v primeru nenormalnih porazdelitev veliko bolje zazna znake stečaja pri podjetjih tudi do dve leti pred stečajem (Tabela 12, str. 36). Diskriminantna analiza za podjetja, ki (potencialno) kažejo znake stečaja v roku, ki je daljši od enega leta (v primeru te naloge 2 leti), teh znakov ne uspe identificirati v dovolj veliki meri (Tabela 26, priloge). Uporaba diskriminantne analize v primeru, ko imamo opravka z nenormalno porazdeljenimi spremenljivkami, implicira, da se banka ne bi smela odločati za dajanje posojil, katerih rok zapadlosti je daljši od enega leta, saj bi v nasprotnem primeru, lahko relativno velik delež podjetij napačno razvrstila nepravilno. V roku, ki je daljši od enega leta in ob nepravilno določitvi ročnosti kreditov, bi lahko bila ob del svojih sredstev.

8.3 Primerjava obeh orodij ter končni komentarji

Večkrat sem že omenil, da tako diskriminantna analiza kot logistična regresija nista popolnoma komplementarni orodji. Tisti del vsakega izmed njiju, ki ju vsebinsko / metodološko ločuje, je najbolj pomemben, če hočem določiti smoternost uporabe enega ali drugega v posameznih okoliščinah. V tem kontekstu se mi zdi smiselno obe orodji soočiti z različnimi variantami, do katerih lahko pride analitik, ko skuša rešiti določen problem v zvezi z razvrščanjem v skupine in / ali napovedovanjem verjetnosti.

Preden se sploh odločamo, katero izmed orodij bomo uporabili, je potrebno definirati sam problem. Potrebno je vsebinsko določiti neodvisne spremenljivke, s katerimi bomo pojasnjevali odvisno spremenljivko. Pri tem mislim na odvisno spremenljivko, ki je diskretna ter dihotomna. S politomnimi spremenljivkami se v tej nalogi nisem ukvarjal. Ko smo spoznali osnovne značilnosti neodvisnih spremenljivk, si moramo postaviti vprašanje, kakšen je končni cilj analize. Ob tem je tudi zelo pomembno, da dobro opredelimo, kakšni so stroški napačne razvrstitve enot. Glavni namen te naloge je bil predvsem primerjati obe orodji ter podati ugotovitve glede primernosti uporabe vsakega. Ključne ugotovitve, na podlagi primerjave rezultatov obeh orodij sem strnil v 12 točkah (glej naslednjo stran).

1. Potrebno je spoznati metodologijo obeh orodij, saj lahko le tako vsako izmed njiju dovolj kvalitetno uporabimo ter dobimo dovolj kvalitetne rezultate, ne glede na omejitve, ki so posledica (porazdelitvenih) lastnosti spremenljivk, s katerimi opravljamo analizo ter omejitve / predpostavke, ki jih postavlja vsako orodje.
2. Za učinkovitost obeh orodij je ključno, da spoznamo osnovne karakteristike spremenljivk, predvsem njihove porazdelitvene lastnosti. Literatura iz področja analize porazdelitev je obsežna. Od tega sicer še ni odvisno, katero orodje bomo uporabili, je pa odvisno, kako se bomo opredelili glede načina analiziranja (velikosti vzorcev, itd.).
3. Preden pričnemo z analizo se moramo odločiti, ali je cilj naloge zgolj pojasnjevalna analiza (ocenitev modela) ali tudi napovedovalna analiza (test robustnosti / stabilnosti ocenjenega modela). Od tega je namreč odvisno, katero orodje bomo uporabili za doseg končnega cilja analize. Če se odločimo tudi za napovedovanje, priporočam napoved z uporabo zunanje metode preverjanja (glej str. 25).
4. Če je naš cilj zgolj pojasnjevalna analiza, je tudi v primeru, ko imamo opravka z nenormalno porazdeljenimi spremenljivkami kvaliteta rezultata neodvisna od izbire orodja, če le uporabljamo dovolj velike vzorce pri diskriminantni analizi.
5. Točka 4 drži, vendar je to pogojeno z metodo izbora optimalne kombinacije spremenljivk. Preden dokončno sklepamo o izboru orodja za analizo, je priporočljivo vložiti nekaj dodatnega časa v dodatne (globlje in širše) analize obnašanja spremenljivk ob čim več različnih scenarijih, ki lahko nastopijo v celotni populaciji. To je zelo koristno, še posebej če delamo tudi napovedovalno analizo na vzorcu enot izven vzorca, na katerem smo ocenili model (zunanja metoda preverjanja).
6. Potrebno je opredeliti, kakšni bodo stroški napačne razvrstitve enot v vsako izmed skupin. Ta predpostavka je pomembna, saj vpliva tako na izbor ključnih parametrov v obeh orodjih (apriorna verjetnost ter točka reza), kot tudi na način, po katerem bomo prišli do optimalnih rezultatov. Ob enakih stroških napačne klasifikacije je optimiranje rezultatov optimiranje z bolj tehničnega vidika uporabljenega orodja, medtem ko se mora ob neenakih stroških napačne klasifikacije analitik hkrati zanašati tudi na svoje izkušnje v skladu s ciljem, ki si ga je zastavil.
7. Če je naš cilj poleg pojasnjevanja tudi napovedovanje, je v primeru nenormalno porazdeljenih spremenljivk bolje uporabiti logistično regresijo. Glede na rezultate te naloge se je veliko bolje odrezala pri napovedovanju stečaja s podatki 2 leti pred stečajem. To v nekem smislu pomeni, da je logit model, ob danem vzorcu, ob danih spremenljivkah ter ob danih parametrih (apriorna verjetnost ter točka reza) boljše specificiran kot model diskriminantne analize.

8. Če se odločimo tudi za napovedovanje, je dobro, da čimbolje ocenimo pogostost dogodka za populacijo (apriorna verjetnost). Relativna velikost vzorcev, na podlagi katerih ocenimo model, namreč vpliva na ocenjene koeficiente oz. na diskriminacijsko moč modela. Ob tem se moramo zavedati, da se pogostost dogodka spreminja v odvisnosti od strukture podatkovne baze (ang. *survival bias*). Od vrste, velikosti, itd., podjetij (v primeru te naloge) je odvisno kakšno pogostost dogodka bomo določili.
9. Glede na točko 8, vedno preverimo porazdelitvene karakteristike neodvisnih spremenljivk. Če je mogoče uporabiti določeno transformacijo, s katero bi lahko spremenljivke normalizirali ter s tem zadostili kriterijem diskriminantne analize, ni izključeno, da se tudi pri napovedovanju diskriminantna analiza lahko odreže vsaj tako dobro kot logistična regresija.
10. Vezano na točko 9, preverimo tudi odnose med odvisno in vsako izmed neodvisnih spremenljivk, za vsako skupino. V primeru nelinearnih odnosov, je lahko koristno v model poleg linearnih, vključiti še nelinearne oblike izbranih spremenljivk. V nalogi sem na slikah 8 in 9 v prilogah prikazal nelinearen odnos v mojem modelu, česar nisem upošteval, saj cilj naloge ni bil optimizacija rezultatov temveč primerjava dveh orodij.
11. Ob dovolj kvalitetno izpeljani metodi izbora spremenljivk za končni model bo ne glede na to, ali uporabimo diskriminantno analizo ali logistično regresijo v obeh primerih vsaka spremenljivka grešila zelo podobno in po enakem vrstnem redu glede število napačno razvrščenih enot v obeh skupinah. To sledi iz Tabel 6 in 11 na straneh 29 oz. 35.
12. Glede na prejšnjo točko se zdi metoda izbora spremenljivk, kot sem jo uporabil v tej nalogi, s kombinacijo različnih orodij prava pot do izbora optimalne kombinacije spremenljivk, s katero bomo dosegli karseda kvalitetne ter tudi robustne rezultate.

9 POMEMBNO, VENDAR NEVKLJUČENO V TEJ NALOGI

Na več mestih v tej nalogi sem analizo dobesedno presekal, ker bi v nasprotnem primeru prešel vsebinski okvir, ki sem si ga zadal. Kljub temu se mi zdi smiselno omeniti nekaj stvari, ki so vredne posvetitve v nadaljnih metodoloških študijah diskriminantne analize in / ali logistične regresije. Sam sem jih opravil bolj površinsko oz. jih nisem opravil. Del, ki najbolj vpliva na končni rezultat, je metoda izbora spremenljivk. V tej nalogi sem opravil ta del relativno enostavno (v primerjavi z možnostmi, ki bi jih lahko še vključil), a kljub temu dobil relativno robustne / stabilne rezultate. Menim, da bi bilo smiselno metodo izbora tako razširiti, predvsem pa poglobiti ter s tem doseči še boljše rezultate. Od metode izbora je namreč odvisno vse nadaljnje optimiziranje pri ocenjevanju modelov.

Druga stvar, ki se mi je zdela zanimiva in sem jo tudi opravil relativno enostavno, je analiziranje napačno razvrščenih podjetij. V nekem smislu lahko tu iščemo faktorje, zaradi katerih ocenjen model (sistematično) greši. Prav gotovo leži del vzrokov tudi v nezadostno izvedeni metodi izbora optimalne kombinacije pojasnjevalnih spremenljivk in / ali neupoštevanje nelinearnih odnosov med odvisno in neodvisnimi spremenljivkami. S tega vidika je analiziranje napačno razvrščenih enot komplementaren del metodi izbora optimalne kombinacije spremenljivk in je izveden *ex-post* (po oceni modela). Iz tega izhaja, da bi bilo v analizo metode izbora optimalne kombinacije spremenljivk (po oceni modela) smiselno vključiti tudi metodo, s katero sem v tej nalogi identificiral napačno klasificirana podjetja. To je s pomočjo intervalov zaupanja. Lahko bi še našteval. Izhodišč za analizo odločitve ter optimiziranje je ogromno. V tej nalogi sem jih zajel res le peščico.

LITERATURA

1. Altman I. Edward: Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy : New York, The Journal of Finance – Vol. 23 (No. 4), 1968. 21 str.
2. Altman I. Edward, Resti Andrea, Sironi Andrea: Default Recovery Rates in Credit Risk Modeling: A Review of the Literature and Empirical Evidence, New York, 2003, 32 str.
3. Baum F. Christopher: An Introduction to Modern Econometrics Using Stata. Boston : Stata Corp LP, 2006. 340 str.
4. Brezigar Masten Arjana, Masten Igor: Comparison of Parametric, Semi – parametric and Non – parametric Methods in Bankruptcy Prediction. Ljubljana : Ekonomska Fakulteta, 2007. 23 str.
5. Frank R. E., Massy W. F. in Morrison G. D.: Bias in Multiple Discriminant Analysis. Journal of Marketing Research – Vol. 2, 1965. str. 250 – 258.
6. Hosmer W. David, Lemeshow Stanley: Applied Logistic Regression. Massachusetts : John Wiley & Sons, 2000. 375 str.
7. Huberty J. Carl: Applied Discriminant Analysis. Georgia : John Wiley & Sons, 1994. 466 str.
8. Ribnikar Ivan: Monetarna Ekonomija I. Ljubljana : Ekonomska fakulteta, 2003. 471 str.
9. Ribnikar Ivan: Monetarna Ekonomija III. Ljubljana : Ekonomska fakulteta, 2003. 216 str.
10. Rovan Jože: Statistika 3 – Diskrimintna Analiza, zapiski predavanj. Ljubljana, 2007. 13 str.
11. Smith R. F., Winakor A. H.: Changes in the Financial Structure of Unsuccessful Corporations. Illionis : Boreau of Business Research, 1935. str. 88.
12. Tabachnick G. Barbara: Fidell S. Linda: Using Multivariate Statistics. Boston : Allyn and Bacon, 2000. 1008 str.
13. Zulkarnain M. Sori, et al.: Some Basic Properties of Financial Ratios: Evidence from an Emerging Capital Market. Malaysia : International Research Journal of Finance and Economics, Issue 2 2006. 16. str.

VIRI

1. Agencija Republike Slovenije za Javnopravne Evidence in Storitve (AJPES) [URL: <http://www.ajpes.si/>], 2005.
2. Net MBA. Business Knowledge Center [URL: <http://www.netmba.com/finance/financial/ratios/>], 2007.
3. OSU Ecology. Ordination Methods for Ecologists [URL: <http://ordination.okstate.edu/PCA.htm>], 2006.
4. Pahor Marko: Statistika 3 – Logistična Regresija, zapiski predavanj. Ljubljana, [URL: ftp://ftp.ef.uni-lj.si/_dokumenti/predmeti/Logisticna_regresija_nov_nov.pdf], 2006. 7 str.
5. Stat Soft [URL: <http://www.statsoft.com/textbook/stdiscan.html>], 2008.

PRILOGE

Kazalo Prilog

Tabela 1: Specifikacije uporabljenih spremenljivk.....	3
Tabela 2: Pregled porazdelitev spremenljivk po prvem koraku s faktorsko analizo (po odstranitvi ekstremnih vrednosti-osamelcev).....	4
Tabela 3: Pregled porazdelitev spremenljivk po prvem koraku s faktorsko analizo (pred odstranitvijo osamelcev)	4
Tabela 4: Prikaz znižanja asimetričnosti po odstranitvi osamelcev	5
Tabela 5: Test univariatne normalnosti (brez osamelcev)	5
Tabela 6: Prvi korak metode izbora spremenljivk – glavni faktorji in pojasnitev odstotka variance v podatkih	6
Tabela 7: Faktorska matrika-prvi korak metode izbora optimalne kombinacije spremenljivk..	7
Tabela 8: Prikaz učinkovitosti spremenljivk pri univariatnem razvrščanju podjetij z logistično regresijo (2 korak metode izbora spremenljivk.....	8
Tabela 9: Prikaz učinkovitosti spremenljivk pri univariatnem razvrščanju podjetij z logistično regresijo (2 korak)	9
Tabela 10: Test ustreznosti podatkov za analizo glavnih komponent.....	10
Tabela 11: Tabela štirih glavnih komponent.....	10
Tabela 12: Štiri glavne komponente in spremenljivke s pripadajočimi utežmi	11
Tabela 13: Korelacije med izbranimi spremenljivkami za skupino 0.....	11
Tabela 14: Korelacije med izbranimi spremenljivkami za skupino 1	12
Tabela 15: Ne enakosti varianc med skupinama za izbrane spremenljivke	12
Tabela 16: Box-ov test enakosti varianc	12
Tabela 17: Test o enakosti varianc med skupinama.....	13
Tabela 18: Kazalniki kakovosti ocenjene diskriminantne funkcije	13
Tabela 19: Klasifikacija diskriminantnega modela; neenake apriorne verjetnosti ($r=80:20$) ..	13
Tabela 20: Vrednosti ocenjene diskriminantne funkcije pri centroidih	13
Tabela 21: Ocenjena standardizirana kanonična diskriminantna funkcija.....	14
Tabela 22: Klasifikacijska tabela ocenjenega diskriminantnega modela.....	14
Tabela 23: Preizkus značilnosti napovedi z ocenjenim modelom.....	14
Tabela 24: Primerjava parametrov za razvrščena podjetja.....	15
Tabela 25: Opisne statistike – 2 leti pred stečajem	15
Tabela 26: Preizkus značilnosti napovedi z ocenjenim modelom – 2 leti pred stečajem	15
Tabela 27: Klasifikacijska tabela ocenjenega logit modela	16
Tabela 28: Kazalniki kakovosti ocenjenega logit modela.....	16
Tabela 29: Preizkus značilnosti rezultatov ocenjenega logit modela.....	16
Tabela 30: Primerjava parametrov razvrščenih podjetij	16
Tabela 31: Preizkus značilnosti napovedanih rezultatov z logit modelom	17
Slika 1: Linearni verjetnostni model z diskretno neodvisno spremenljivko	17
Slika 2: Logistični regresijski model z zvezno neodvisno spremenljivko	17

Slika 3: Analiza razlik povprečnih vrednosti spremenljivk med skupinami.....	18
Slika 4: Spreminjanje povprečnih razlik v deležu pravilno razvrščenih podjetij med scenariji za skupino 0.....	18
Slika 5: Spreminjanje povprečnih razlik v deležu pravilno razvrščenih podjetij med scenariji za skupino 1.....	19
Slika 6: Podatki v treh dimenzijah in določanje glavnih komponent.....	19
Slika 7: Graf lastnih vrednosti za vsako imed (štirih) glavnih komponent.....	20
Slika 8: Nelinearna odvisnost deleža bankrotiranih podjetij od višine neodvisne spremenljivke X1(prihodki / sredstva) za obe skupini podjetij	20
Slika 9: Nelinearna odvisnost deleža bankrotiranih podjetij od višine neodvisne spremenljivke X26 (kapital / celotni dolg) za obe skupini podjetij	21
Slika 10: Verjetnost stečaja v odvisnosti od spremenljivke X1	22
Slika 11: Verjetnost stečaja v odvisnosti od spremenljivke X11	22
Slika 12: Verjetnost stečaja v odvisnosti od spremenljivke X17	23
Slika 13: Verjetnost stečaja v odvisnosti od spremenljivke X26	23

Tabela 1: Specifikacije uporabljenih spremenljivk

Spremenljivka	Specifikacija
X1	prihodki** / sredstva*
X2	dobiček / prihodki od prodaje (ROS)
X3	EBIT / celotna sredstva
X4	sprememba kapitala glede na prejšnje leto v %
X5	letna sprememba prihodkov od prodaje
X6	kratkoročni koeficient (kratkoročne terjatve / kratkoročne obveznosti)
X7	donos / celotna sredstva (ROA)
X8	zadržani dobički / celotna sredstva
X9	prihodki od prodaje / povprečno število zaposlenih (produktivnost)
X10	celotni dolg / kapital (debt-to-equity ratio)
X11	hitri koeficient (quick ratio)
X12	prihodki od prodaje / povprečna stanje terjatev v obdobju (obrat terjatev)
X13	donos / lastniški kapital (ROE)
X14	celotni dolg / celotna sredstva (debt ratio)
X15	dobiček / celotni dolg
X16	prihodki od prodaje / povprečna zaloga v obdobju (obračanje zalog)
X17	dobiček / prihodki od prodaje(dobičkonosnost prodaje)
X18	denarni tok / celotni dolg
X19	zaloga / kratkoročni kapital
X20	koeficient obračanja obratnih sredstev
X21	delovni kapital / celotna sredstva
X22	% kratkoročnih obveznosti v celotnih sredstvih
X23	% dolgoročnih obveznosti v celotnih sredstvih
X24	denarni tok / prihodke od prodaje
X25	obrat denarnih sredstev (prodaja/denar)
X26	kapital / celotni dolg
X27	sprememba prodaje v primerjavi z prejšnjim letom v %
X28	rast kratkoročnega dolga

*Celotna sredstva

**Prihodki od prodaje

Vir: Ajpes; Lastni preračuni.

Tabela 2: Pregled porazdelitev spremenljivk po prvem koraku s faktorsko analizo (po odstranitvi ekstremnih vrednosti – osamelcev)

Spremenljivka	Skupina 0		Skupina 1	
	Asimetričnost	Sploščenost	Asimetričnost	Sploščenost
X1	155,0864	30258,77	37,92804	1440,021
X6	384,5255	164559,1	31,95477	1056,774
X7	243,4822	69003,54	36,73411	1374,27
X9	58,25662	6351,834	6,51123	64,35802
X11	454,7285	234861	36,88767	1388,992
X12	78,33496	8931,311	31,3527	1092,667
X14	160,7267	32747,55	18,40431	372,8259
X15	303,4093	117877,7	37,96917	1444,055
X17	409,1694	199497,9	17,3356	334,7662
X20	174,2313	42596,21	12,92694	229,6334
X21	0,098274	1,769279	-0,1440162	1,522741
X22	205,5027	50516,54	20,10592	431,6668
X26	407,7861	197007,2	-23,74973	611,5551

Vir: Ajpes; Lastni izračuni.

Tabela 3: Pregled porazdelitev spremenljivk po prvem koraku s faktorsko analizo (pred odstranitvijo – osamelcev)

Spremenljivka	Skupina 0		Skupina 1	
	Asimetričnost	Sploščenost	Asimetričnost	Sploščenost
X1	1,077002	4,035391	2,833148	12,96677
X6	2,620925	12,78798	1,43161	4,939582
X7	1,824033	5,946309	9,433063	100,5549
X9	1,677954	5,565913	2,024864	6,914953
X11	4,154612	27,86993	1,289963	4,792073
X12	6,746075	62,14247	5,009038	39,95345
X14	-1,475832	14,02088	11,45203	169,8889
X15	2,290447	9,640901	9,521576	103,6187
X17	2,517373	10,39293	10,93168	143,9242
X20	1,30712	4,731864	1,135295	3,752703
X21	0,2540455	2,041811	0,2150118	1,664827
X22	0,7067868	3,914705	11,40918	155,9885
X26	15,97665	1612,304	-10,96782	144,2549

Vir: Ajpes; Lastni izračuni.

Tabela 4: Prikaz znižanja asimetričnosti po odstranitvi osamelcev**Paired Samples Test**

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 SK_0 - SK_1	3,23688	149,44092	9,30778	2,87638	3,59737	4,205	25	,000
Pair 2 KURT_0 - KUR	4668,65	8062,68824	4328,77	5157,99	4179,31	3,117	25	,005

Vir: Ajpes; Lastni izračuni.

Tabela 5: Test univariatne normalnosti (brez osamelcev)

Spremenljivka	Skupina 0			Skupina 1		
	Shapiro-Wilk W			Shapiro-Wilk W		
	Enot	W	Prob>z	Enot	W	Prob>z
X1	15409	0,92364	0,00000	602	0,75125	0,00000
X6	15409	0,72495	0,00000	602	0,86714	0,00000
X7	15409	0,79314	0,00000	602	0,48687	0,00000
X9	15409	0,82443	0,00000	602	0,74751	0,00000
X11	15409	0,50142	0,00000	602	0,89683	0,00000
X12	15409	0,78557	0,00000	602	0,73156	0,00000
X14	15409	0,79022	0,00000	602	0,25411	0,00000
X15	15409	0,73297	0,00000	602	0,49756	0,00000
X17	15409	0,71364	0,00000	602	0,44080	0,00000
X20	15409	0,89471	0,00000	602	0,79903	0,00000
X21	15409	0,96845	0,00000	602	0,94017	0,00000
X22	15409	0,96674	0,00000	602	0,24425	0,00000
X26	15409	0,69929	0,00000	602	0,23677	0,00000

Vir: Ajpes; Lastni izračuni.

Tabela 6: Prvi korak metode izbora spremenljivk – glavni faktorji in pojasnitev odstotka variance v podatkih

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,634	12,541	12,541	2,460	11,716	11,716	2,333	11,109	11,109
2	2,155	10,264	22,804	1,942	9,248	20,964	1,709	8,140	19,249
3	1,850	8,808	31,613	1,673	7,965	28,929	1,509	7,186	26,435
4	1,746	8,315	39,927	1,506	7,173	36,102	1,192	5,678	32,113
5	1,318	6,275	46,202	,931	4,431	40,533	,962	4,579	36,693
6	1,130	5,379	51,581	,413	1,967	42,500	,756	3,599	40,291
7	1,046	4,981	56,562	,361	1,721	44,221	,469	2,236	42,527
8	1,014	4,827	61,389	,326	1,550	45,771	,397	1,890	44,417
9	1,002	4,770	66,159	,228	1,086	46,857	,356	1,695	46,112
10	1,000	4,762	70,921	,138	,656	47,513	,294	1,400	47,513
11	,992	4,723	75,644						
12	,951	4,530	80,174						
13	,934	4,450	84,624						
14	,837	3,988	88,611						
15	,617	2,936	91,547						
16	,573	2,729	94,276						
17	,352	1,678	95,955						
18	,323	1,538	97,493						
19	,281	1,339	98,832						
20	,153	,727	99,559						
21	,093	,441	100,000						

Extraction Method: Principal Axis Factoring.

Vir: Aipes: Lastni izračuni.

Tabela 7: Faktorska matrika – prvi korak metode izbora optimalne kombinacije spremenljivk

Rotated Factor Matrix ^a

	Factor									
	1	2	3	4	5	6	7	8	9	10
Prinodki/sredstva	-.011	-.071	-.022	,287	,847	,225	,045	,024	,093	,033
ROA	,978	,027	,026	,021	,105	,069	,056	,015	,078	,051
Prilodki od prodaje/Popv. število zaposlenih (Produktivnost)	,082	,002	-.003	,075	,239	,140	,583	-.082	-.009	,063
Hitri Koefficient	,057	,054	,841	,015	-.040	,040	-.028	,009	,107	-.411
Prilodki/pter (Obrat terjavev do kupcev)	-.002	,014	-.064	,506	,149	-.096	-.102	,009	-.125	-.018
(Debt ratio)	-.085	-.027	,014	-.005	,010	,015	,016	,578	-.008	-.014
Dobitkonosnost prodaje	,754	,022	,038	-.053	-.175	-.060	,039	,009	,082	-.023
Koefficient obračanja zalog	,040	,001	,012	,091	-.018	,047	-.021	7,89E-005	,342	,002
Koefficient obračanja obratnih sredstev	,006	,014	-.042	,857	,299	-.128	,026	,014	,183	,185
Zaloga/Kratkoročni kapital	,002	,001	,003	,000	,002	-.001	,001	,000	,008	,001
Delovni Kapital/Celotna Sredstva	-.032	-.048	,167	-.199	,061	,653	,159	-.047	,339	-.051
rast kratkoročnega dolga rast prodaje (v%)	-.001	-.012	-.127	,030	-.037	,024	,030	-.004	,019	,167
% dolgoročnih obveznosti v celotnih sredstvih	-.026	-.030	,009	,001	-.084	-.434	,089	-.032	,078	,055
% kratkoročnih obveznosti v celotnih sredstvih	-.011	-.897	-.038	,017	,027	,053	,006	,063	-.046	-.042
Denarni tok/celotni dolg	-.009	,008	,029	,013	,076	,091	-.261	-.048	-.011	-.027
Kratak. Terjave/Kratak. Obveznosti (Kratkoročni koefficient)	,028	,031	,864	-.132	-.022	,112	-.035	,016	,148	,200
Sprememba kapitala glede na prejšnje leto v %	,015	,041	,012	,009	-.001	,034	,000	,001	-.040	,013
Sprememba prodaje v primerjavi z prejšnjim letom v %	-.010	-.014	-.012	,205	-.059	,044	,079	-.015	,074	,033
kapital/celotni dolg	,019	,942	,010	-.008	-.024	,047	-.016	,014	,002	,039
Dobitck/Sredstva	,886	,025	,012	,019	,092	,053	,017	-.212	,068	,050

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 11 iterations.

Vir: Aides: Lastni izračuni.

Tabela 8: Prikaz učinkovitosti spremenljivk pri univariatnem razvrščanju podjetij z logistično regresijo (2 korak metode izbora neodvisnih spremenljivk)

Scenarij 1: r = 50:50, n0 = n1 = 300; število pravilno razvrščenih v vsaki skupini (0,1)								
	Test1	Povpr.	Test2	Povpr.	Test3	Povpr.	Test4	Povpr.
Sprem.	% correct	%	% correct	%	% correct	%	% correct	%
X1	60,7; 34	47,3	58,7; 31,3	45	66,7; 33,3	49	66,6; 36	51,3
X6	50,7; 76	63,3	55,3; 78,7	67	56,7; 75,3	66	58; 72,7	65,3
X7	58,7; 94,7	76,7	59,3; 96	77,7	54; 96,7	75,3	57,3; 96	76,7
X9	48,7; 82	65,3	40; 74	57	44,7; 75,3	60	48; 77,3	62,7
X11	58,7; 79,3	69	66; 77,3	71,7	62,7; 76,7	69,7	68; 77,3	72,2
X12	36,7; 63,3	50	41,3; 70,7	56	38,7; 68,7	53,7	36; 66	51
X14	17,7; 96,7	57	92,7; 12,7	52,7	16,7; 95,3	56	93,3; 12	52,7
X15	58,7; 94,7	76,7	59,3; 95,3	77,3	53,3; 96	74,7	56,7; 95,3	76
X17	54,7; 94,7	74,7	60; 95,3	77,7	51,3; 94	72,2	56; 94,7	75,3
X20	65,3; 43,3	54,3	58; 32,3	45,3	66; 37,3	51,7	64; 35,3	49,7
X21	48,7; 52,7	50,7	50; 56	53	54; 50	52	45,3; 52	48,7
X22	75,3; 58	66,7	79,3; 62	70,7	74,7; 64,7	69,7	78,7; 68,7	73,3
X26	63,3; 61,3	67,3	76; 65,3	70,7	76,7; 70	73,3	76,7; 67,3	72
Scenarij 2: r = 70:30, n1 = 300; število pravilno razvrščenih v vsaki skupini (0,1)								
	Test1	Povpr.	Test2	Povpr.	Test3	Povpr.	Test4	Povpr.
Sprem.	% correct	%	% correct	%	% correct	%	% correct	%
X1	100; 2,7	70,8	100; 0	70	100; 0	70	100; 0	70
X6	90,9; 13,3	67,6	100; 0	70	100; 0	70	100; 0	70
X7	70,9; 94	77,8	100; 0	70	100; 0	70	100; 0	70
X9	100; 0	70	100; 0	70	100; 0	70	100; 0	70
X11	84,9; 48,7	74	87,7; 40	73,4	87,7; 40	73,4	90; 34,7	73,4
X12	100; 0	70	100; 0	70	100; 0	70	100; 0	70
X14	98,9; 3,3	70,2	99,4; 3,3	70,6	99,4; 3,3	70,6	98,9; 5,3	70,8
X15	71,1; 94	78	100; 0	70	100; 0	70	100; 0	70
X17	70,9; 94	77,8	100; 0	70	100; 0	70	100; 0	70
X20	100; 0,7	70,2	100; 1,3	70,4	100; 1,3	70,4	100; 0	70
X21	100; 0	70	100; 0	70	100; 0	70	100; 0	70
X22	96,6; 40,7	79,8	95,4; 47,3	81	95,4; 47,3	81	94,6; 45,3	79,8
X26	96,3; 42,7	80,2	96,6; 46	81,4	96,6; 46	81,4	96,6; 44	80,8

Vir: Ajpes; Lastni izračuni.

Tabela 9: Prikaz učinkovitosti spremenljivk pri univariatnem razvrščanju podjetij z logistično regresijo (2 korak metode izbora neodvisnih spremenljivk)

Scenarij 3: r = 80:20, n1 = 300; število pravilno razvrščenih v vsaki skupini (0,1)								
	Test1	Povpr.	Test2	Povpr.	Test3	Povpr.	Test4	Povpr.
Sprem.	% correct	%	% correct	%	% correct	%	% correct	%
X1	100; 5	81	100; 0	80	100; 7	81,4	100; 2	80,4
X6	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X7	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X9	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X11	99,3; 4	80,2	95,5; 29	82,2	96,3; 14	79,8	100; 0	80
X12	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X14	100; 1	80,2	100; 7	81,4	99; 8	80,8	99,8; 6	81,1
X15	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X17	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X20	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X21	100; 0	80	100; 0	80	100; 0	80	100; 0	80
X22	97,8; 37	85,6	98,8; 31	85,2	98,5; 47	88,2	97,8; 43	86,8
X26	98,5; 35	85,8	98,5; 31	85	99; 48	88,8	98; 40	84,4
Scenarij 4: r = 90:10, n1 = 300; število pravilno razvrščenih v vsaki skupini (0,1)								
	Test1	Povpr.	Test2	Povpr.	Test3	Povpr.	Test4	Povpr.
Sprem.	% correct	%	% correct	%	% correct	%	% correct	%
X1	100; 6	90,6	100; 0	90	100; 0	90	100; 0	90
X6	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X7	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X9	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X11	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X12	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X14	100; 0	90	100; 0	90	100; 2	90,2	100; 6	90,6
X15	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X17	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X20	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X21	100; 0	90	100; 0	90	100; 0	90	100; 0	90
X22	99,6; 26	92,2	99,3; 22	91,6	100; 26	92,6	100; 18	91,8
X26	99,6; 28	92,4	100; 28	92,8	99,3; 20	91,4	99,6; 16	91,2

Vir: Ajpes; Lastni izračuni.

Tabela 10: Test ustreznosti podatkov za analizo glavnih komponent**KMO and Bartlett's Test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,582
Bartlett's Test of Sphericity	Approx. Chi-Square	110372,5
	df	78
	Sig.	,000

Vir: Ajpes; Lastni izračuni.

Tabela 11: Tabela štirih glavnih komponent**Total Variance Explained**

Comp.	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,842	21,861	21,861	2,842	21,861	21,861	2,611	20,088	20,088
2	2,278	17,524	39,385	2,278	17,524	39,385	2,168	16,674	36,762
3	1,833	14,100	53,485	1,833	14,100	53,485	1,989	15,303	52,065
4	1,731	13,316	66,801	1,731	13,316	66,801	1,916	14,736	66,801
5	,997	7,670	74,472						
6	,957	7,365	81,837						
7	,758	5,832	87,669						
8	,584	4,493	92,162						
9	,388	2,984	95,145						
10	,250	1,925	97,071						
11	,234	1,799	98,870						
12	,090	,690	99,560						
13	,057	,440	100,000						

Extraction Method: Principal Component Analysis.

Vir: Ajpes; Lastni izračuni.

Tabela 12: Štiri glavne komponente in spremenljivke s pripadajočimi utežmi

Rotated Component Matrix ^a

	Component			
	1	2	3	4
Prihodki/sredstva	-,027	,690	,177	,198
Kratk. Terjatve/Kratk. Obveznosti (Kratkoročni koeficient)	,030	-,029	-,152	,832
ROA	,944	,163	-,014	,092
Prihodki od prodaje/Povp. število zaposlenih (Produktivnost)	,120	,519	,022	,027
Hitri Koeficient	,039	-,071	-,169	,797
Prihodki/pter (Obrat terjavev do kupcev) (Debt ratio)	,031	,775	-,106	-,192
Dobiček/Dolg	-,163	-,018	,129	,063
Dobičkonosnost prodaje	,931	,156	-,026	,067
Koeficient obračanja obratnih sredstev	,894	-,099	-,036	,010
Delovni Kapital/Celotna Sredstva	,043	,867	-,067	-,114
% kratkoročnih obveznosti v celotnih sredstvih	-,009	,038	,224	,685
kapital/celotni dolg	-,057	,025	,963	-,072
	,059	-,012	-,945	,079

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Vir: Ajpes; Lastni izračuni.

Tabela 13: Korelacije med izbranimi spremenljivkami za skupino 0

Correlations

		Prihodki/sredstva	Hitri Koeficient	Dobičkonosnost prodaje	kapital/celotni dolg
Prihodki/sredstva	Pearson Correlation	1	-,043	-,168**	-,035
	Sig. (2-tailed)		,140	,000	,221
	N	1200	1200	1200	1200
Hitri Koeficient	Pearson Correlation	-,043	1	,034	,331**
	Sig. (2-tailed)	,140		,242	,000
	N	1200	1200	1200	1200
Dobičkonosnost prodaje	Pearson Correlation	-,168**	,034	1	,180**
	Sig. (2-tailed)	,000	,242		,000
	N	1200	1200	1200	1200
kapital/celotni dolg	Pearson Correlation	-,035	,331**	,180**	1
	Sig. (2-tailed)	,221	,000	,000	
	N	1200	1200	1200	1200

** . Correlation is significant at the 0.01 level (2-tailed).

Vir: Ajpes; Lastni izračuni.

Tabela 14: Korelacije med izbranimi spremenljivkami za skupino 1

		Correlations			
		Prihodki/sredstva	Hitri Koeficient	Dobičkonosnost prodaje	kapital/celotni dolg
Prihodki/sredstva	Pearson Correlation	1	-,183**	-,061	-,318**
	Sig. (2-tailed)		,001	,296	,000
	N	300	300	300	300
Hitri Koeficient	Pearson Correlation	-,183**	1	,179**	,264**
	Sig. (2-tailed)	,001		,002	,000
	N	300	300	300	300
Dobičkonosnost prodaje	Pearson Correlation	-,061	,179**	1	,048
	Sig. (2-tailed)	,296	,002		,408
	N	300	300	300	300
kapital/celotni dolg	Pearson Correlation	-,318**	,264**	,048	1
	Sig. (2-tailed)	,000	,000	,408	
	N	300	300	300	300

** . Correlation is significant at the 0.01 level (2-tailed).

Vir: Ajpes; Lastni izračuni.

Tabela 15: Ne enakosti varianc med skupinama za izbrane spremenljivke

Test of Homogeneity of Variances				
	Levene Statistic	df1	df2	Sig.
Prihodki/sredstva	175,263	1	1498	,000
Hitri Koeficient	39,511	1	1498	,000
Dobičkonosnost prodaje	219,592	1	1498	,000
kapital/celotni dolg	203,214	1	1498	,000

Vir: Ajpes; Lastni izračuni.

Tabela 16: Box-ov test enakosti varianc

Test Results		
Box's M		5069,176
F	Approx.	504,363
	df1	10
	df2	1336024
	Sig.	,000

Tests null hypothesis of equal population covariance matrices.

Vir: Ajpes; Lastni izračuni.

Tabela 17: Test o enakosti varianc med skupinama

		ANOVA				
		Sum of Squares	df	Mean Square	F	Sig.
Prihodki/sredstva	Between Groups	46,791	1	46,791	31,536	,000
	Within Groups	8899,407	5998	1,484		
	Total	8946,198	5999			
Hitri Koeficient	Between Groups	262,631	1	262,631	66,206	,000
	Within Groups	23793,218	5998	3,967		
	Total	24055,850	5999			
Dobičkonosnost prodaje	Between Groups	,134	1	,134	101,0	,000
	Within Groups	7,939	5998	,001		
	Total	8,073	5999			
kapital/celotni dolg	Between Groups	560,018	1	560,018	630,0	,000
	Within Groups	5331,436	5998	,889		
	Total	5891,455	5999			

Vir: Ajpes; Lastni izračuni.

Tabela 18: Kazalniki kakovosti ocenjene diskriminantne funkcije

Wilks' Lambda				
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,836	268,678	4	,000

Vir: Ajpes; Lastni izračuni.

Tabela 19: Klasifikacija diskriminantnega modela; neenake apriorne verjetnosti (r=80:20)

Classification Results ^a					
		Predicted Group Membership		Total	
		,00	1,00		
Original	Count	,00	1200	0	1200
		1,00	267	33	300
	%	,00	100,0	,0	100,0
		1,00	89,0	11,0	100,0

a. 82,2% of original grouped cases correctly classified.

Vir: Ajpes; Lastni izračuni.

Tabela 20: Vrednosti ocenjene diskriminantne funkcije pri centroidih

Functions at Group Centroids

stecaj	Function
	1
,00	,222
1,00	-,887

Unstandardized canonical discriminant functions evaluated at group means

Vir: Ajpes; Lastni izračuni.

Tabela 21: Ocenjena standardizirana kanonična diskriminantna funkcija**Canonical Discriminant Function Coefficients**

	Function
	1
Prihodki/sredstva	,010
Hitri Koeficient	,188
Dobičkonosnost prodaje	19,317
kapital/celotni dolg	,402
(Constant)	-,585

Unstandardized coefficients

Vir: Ajpes; Lastni izračuni.

Tabela 22: Klasifikacijska tabela ocenjenega diskriminantnega modela**Classification Results^{b,c}**

	stecaj	Predicted Group Membership		Total	
		,00	1,00		
Original	Count	,00	952	248	1200
		1,00	60	240	300
	%	,00	79,3	20,7	100,0
		1,00	20,0	80,0	100,0
Cross-validated ^a	Count	,00	952	248	1200
		1,00	62	238	300
	%	,00	79,3	20,7	100,0
		1,00	20,7	79,3	100,0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 79,5% of original grouped cases correctly classified.

c. 79,3% of cross-validated grouped cases correctly classified.

Vir: Ajpes; Lastni izračuni.

Tabela 23: Preizkus značilnosti napovedi z ocenjenim modelom

Spremenljivka	Skupina 0		Skupina 1	
	Povprečje	Std. odklon	Povprečje	Std. odklon
Z0	0,66	1,51	-0,23	2,11
Z1	0,05	1,05	0,38	2,52
H0: (Z0-Z1)=0	vrednost t – preizkusa=11,59		vrednost t – preizkusa=-3,23	
H0: (Z0-Z1)=0	p – vrednost=0,000		p – vrednost=0,001	

Vir: Ajpes; Lastni izračuni. *Predpostavljene neenake variance med skupinama.

Tabela 24: Primerjava parametrov za razvrščena podjetja

	Skupina 0			Skupina 1		
	Pravilno	Nepravilno		Pravilno	Nepravilno	
Sprem.	Povprečje	Povprečje	Std. odklon	Povprečje	Povprečje	Std. odklon
X1	1,66	1,46	1,14	1,95	1,24	1,24
X11	1,64	0,47	0,29	0,36	0,86	0,50
X17	0,03	0,0006	0,001	0,00003	0,006	0,02
X26	0,43	0,003	0,41	-1,36	0,39	0,25

Vir: Ajpes; Lastni izračuni.

Tabela 25: Opisne statistike – 2 leti pred stečajem**Group Statistics**

stecaj		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
,00	Prihodki/sredstva	1,6100	1,10404	14209	14209,000
	Hitri Koeficient	1,4064	1,96340	14209	14209,000
	Dobičkonosnost prodaje	,0226	,03786	14209	14209,000
	kapital/celotni dolg	,3400	,40250	14209	14209,000
1,00	Prihodki/sredstva	2,2762	8,07874	277	277,000
	Hitri Koeficient	,8438	2,89626	277	277,000
	Dobičkonosnost prodaje	,0087	,03954	277	277,000
	kapital/celotni dolg	-,0414	1,21762	277	277,000
Total	Prihodki/sredstva	1,6227	1,56445	14486	14486,000
	Hitri Koeficient	1,3956	1,98670	14486	14486,000
	Dobičkonosnost prodaje	,0223	,03793	14486	14486,000
	kapital/celotni dolg	,3327	,43576	14486	14486,000

Vir: Ajpes; Lastni izračuni.

Tabela 26: Preizkus značilnosti napovedi z ocenjenim modelom – 2 leti pred stečajem

Spremenljivka	Skupina 0		Skupina 1	
	Povprečje	Std. odklon	Povprečje	Std. odklon
Z0	0,68	1,58	0,59	8,54
Z1	0,01	1,04	0,51	7,79
H0: (Z0-Z1)=0	vrednost t – preizkusa=41,78		vrednost t – preizkusa=0,13	
H0: (Z0-Z1)=0	p – vrednost=0,000		p – vrednost=0,8965	

Vir: Ajpes; Lastni izračuni. *Predpostavljene neenake variance med skupinama.

Tabela 27: Klasifikacijska tabela ocenjenega logit modela

Classification Table ^a

Observed		Predicted		
		stecaj		Percentage Correct
		,00	1,00	
Step 1	stecaj	,00	1,00	
	Overall Percentage			

a. The cut value is ,250

Vir: Ajpes; Lastni izračuni.

Tabela 28: Kazalniki kakovosti ocenjenega logit modela

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	934,038 ^a	,315	,498

a. Estimation terminated at iteration number 9 because parameter estimates changed by less than ,001.

Vir: Ajpes; Lastni izračuni.

Tabela 29: Preizkus značilnosti rezultatov ocenjenega logit modela

	Verjetnost	
	Povprečje	Std. odklon
Skupina 0	0,124	0,161
Skupina 1	0,506	0,282
Povprečna razlika (v o.t.)		-38,2
H0: P0-P1 = 0	vrednost t – preizkusa=-22,56*	
H0: P0-P1 = 0	p – vrednost=0,000	

Vir: Ajpes; Lastni izračuni. *Predpostavljene neenake variance med skupinama.

Tabela 30: Primerjava parametrov razvrščenih podjetij

Sprem.	Skupina 0			Skupina 1		
	Pravilno	Nepravilno		Pravilno	Nepravilno	
	Povprečje	Povprečje	Std. odklon	Povprečje	Povprečje	Std. odklon
X1	1,68	1,30	1,03	1,83	1,83	1,91
X11	1,63	0,44	0,27	0,37	0,89	0,37
X17	0,028	0,0004	0,001	0	0,005	0,01
X26	0,42	0,01	0,43	-1,33	0,31	0,28

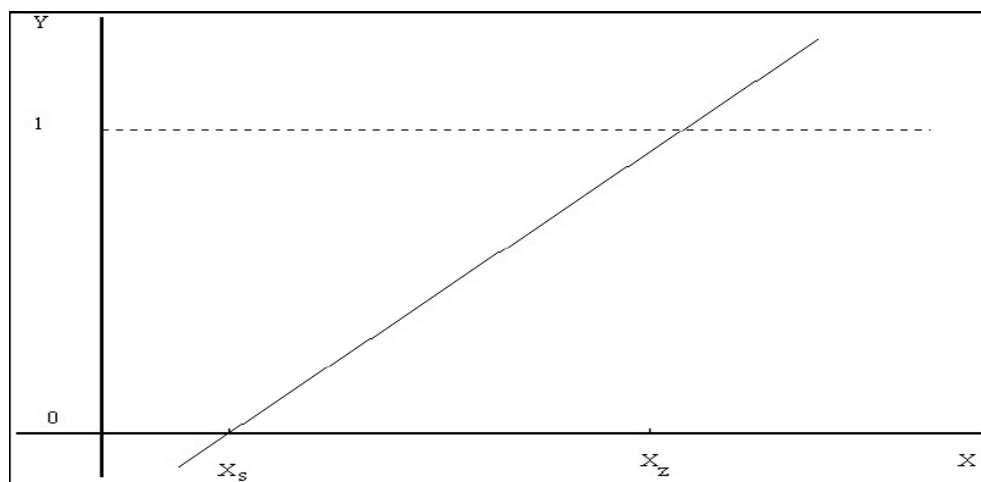
Vir: Ajpes; Lastni izračuni.

Tabela 31: Preizkus značilnosti napovedanih rezultatov z logit modelom

	Verjetnost	
	Povprečje	Std. odklon
Skupina 0	0,1206	0,1553
Skupina 1	0,3032	0,2231
Povprečna razlika (v o.t.)	-18,26	
H0: P0-P1 = 0	vrednost t – preizkusa=-13,55*	
H0: P0-P1 = 0	p – vrednost=0,000	

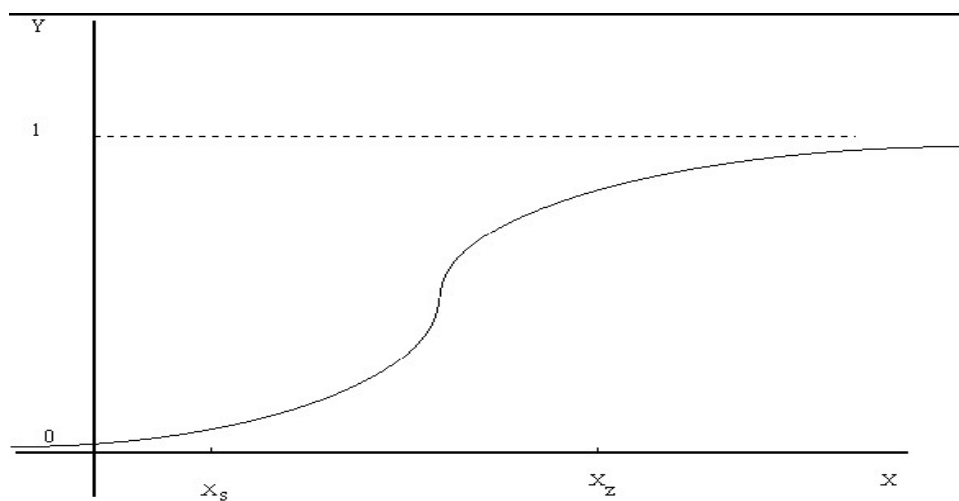
Vir: Ajpes; Lastni izračuni. *Predpostavljene neenake variance med skupinama.

Slika 1: Linearni verjetnostni model z diskretno neodvisno spremenljivko



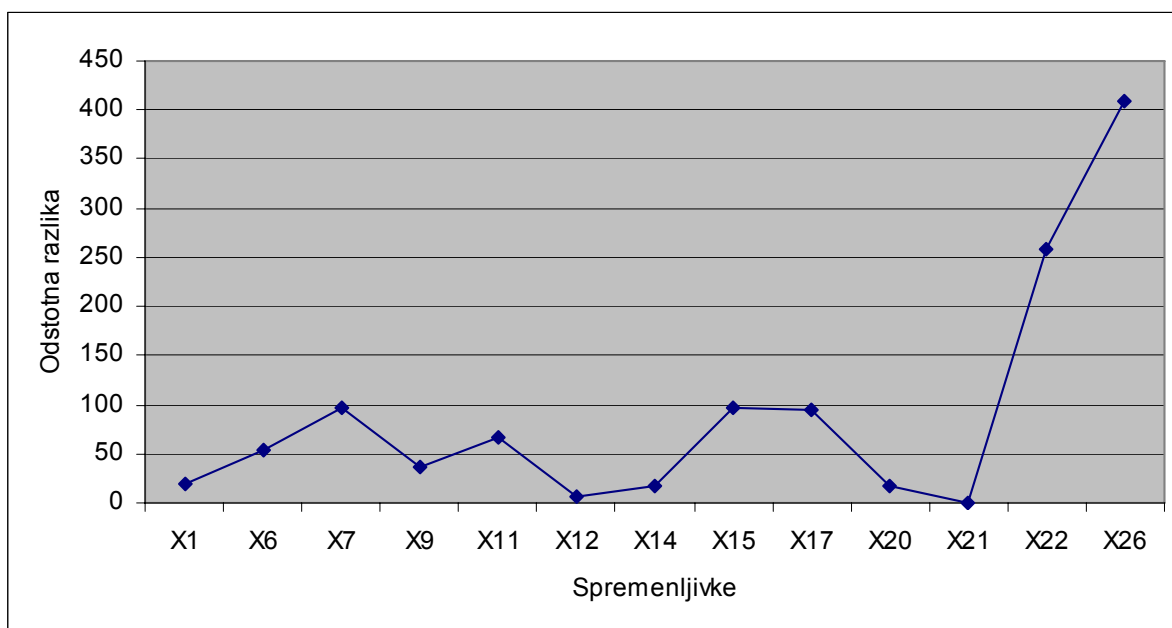
Vir: Pahor, 2003.

Slika 2: Logistični regresijski model z zvezno neodvisno spremenljivko



Vir: Pahor, 2003.

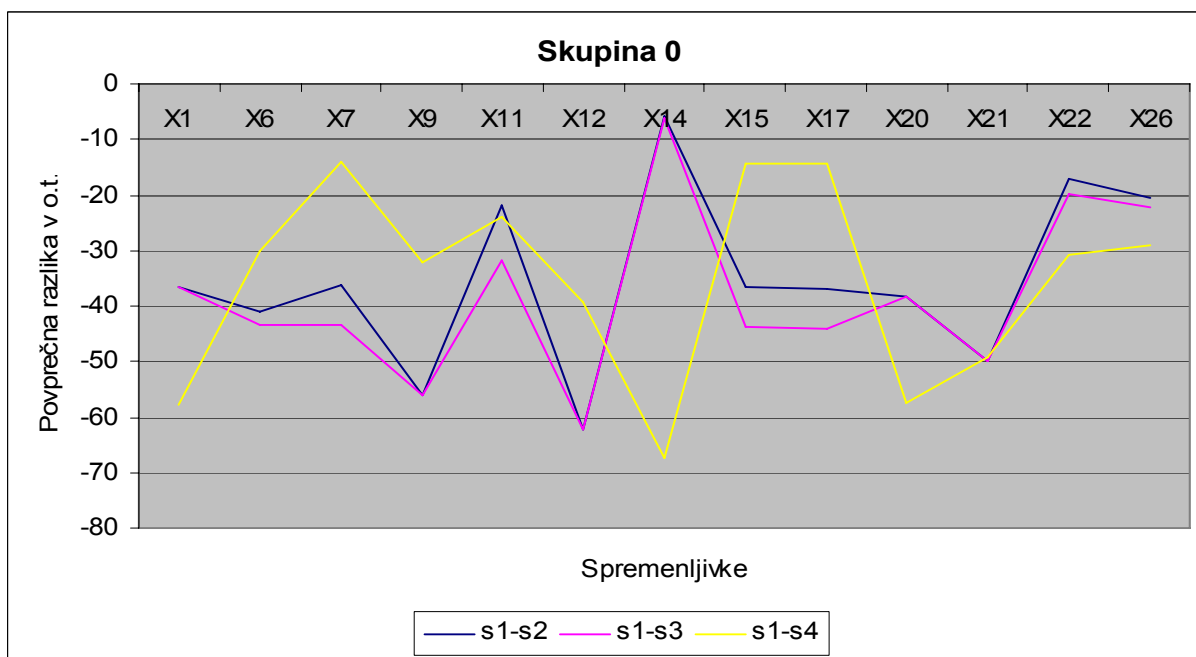
Slika 3: Analiza razlik povprečnih vrednosti spremenljivk med skupinami



Vir: Ajpes; Lastni izračuni.

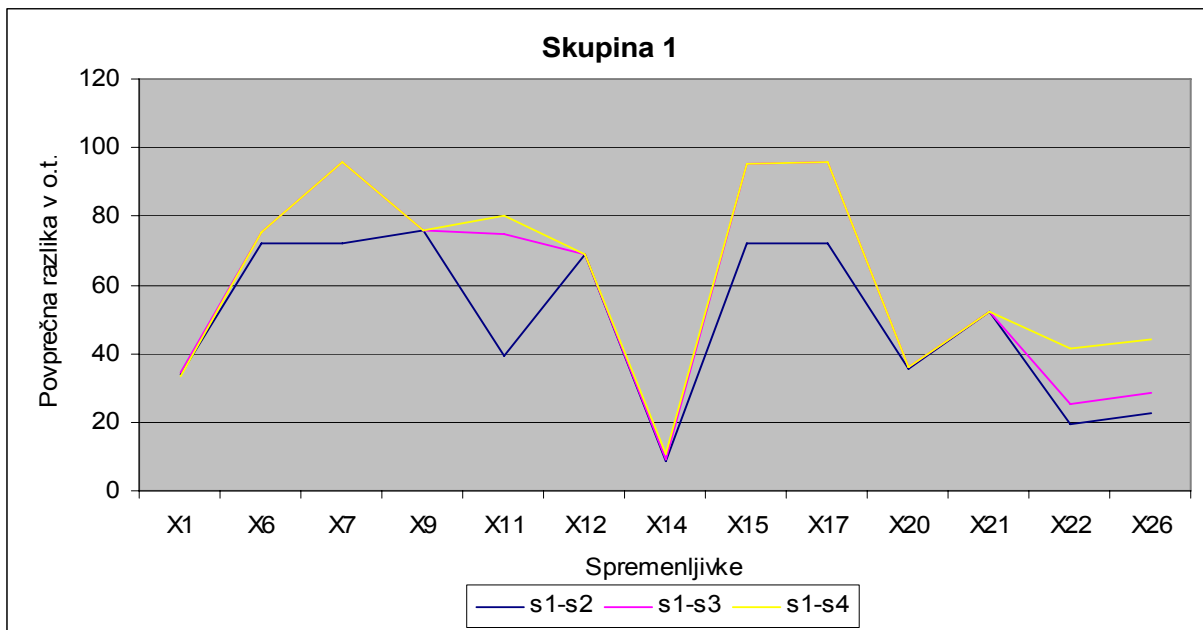
Slika 4: Spreminjanje povprečnih razlik v deležu pravilno razvrščenih podjetij med scenariji za skupino 0

Na slikah 4 in 5 prikazujem, kako se spreminja povprečen delež pravilno razvrščenih podjetij iz vseh štirih testiranj, v vsakem izmed štirih scenarijev (drugi korak izbora spremenljivk, z logistično regresijo). Vsaka črta na grafu, pomeni razliko med prvim ter enim izmed ostalih treh scenarijev. Pri tem oznake v legendi npr. 's1-s2', pomenijo prej omenjeno razliko med prvim in drugim scenarijem.



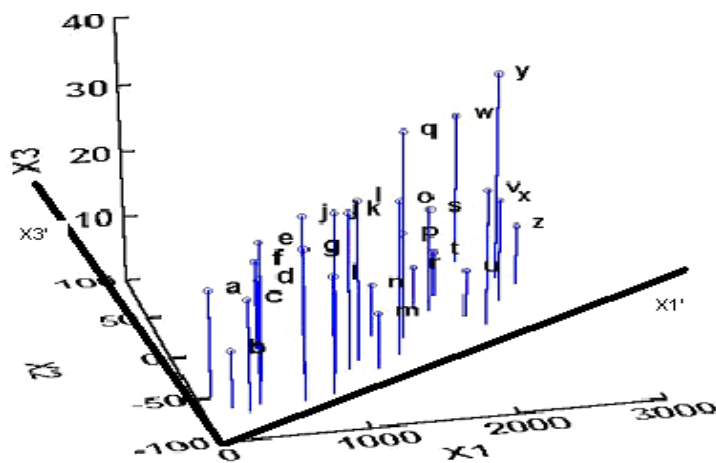
Vir: Ajpes; Lastni izračuni.

Slika 5: Spreminjanje povprečnih razlik v deležu pravilno razvrščenih podjetij med scenariji za skupino 1



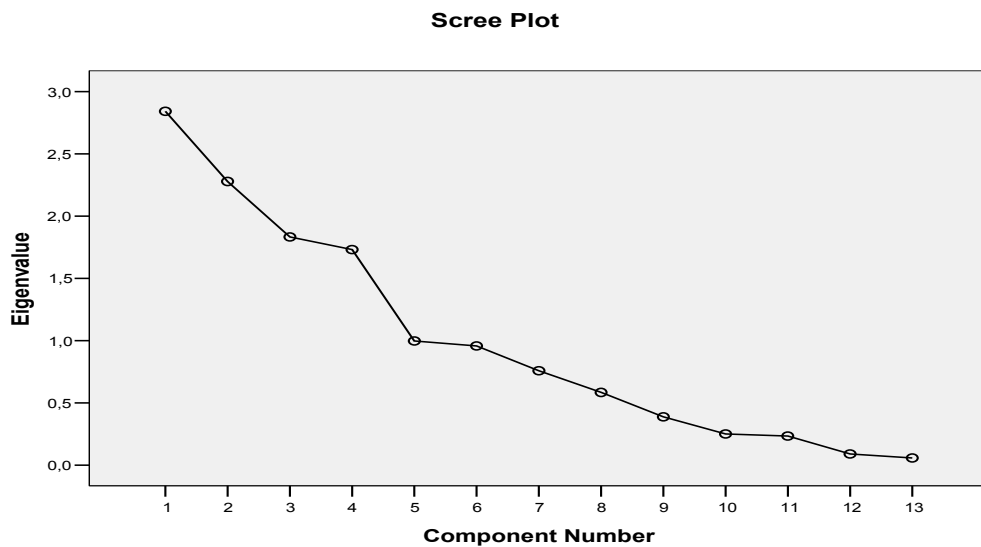
Vir: Ajpes; Lastni izračuni.

Slika 6: Podatki v treh dimenzijah in določanje glavnih komponent



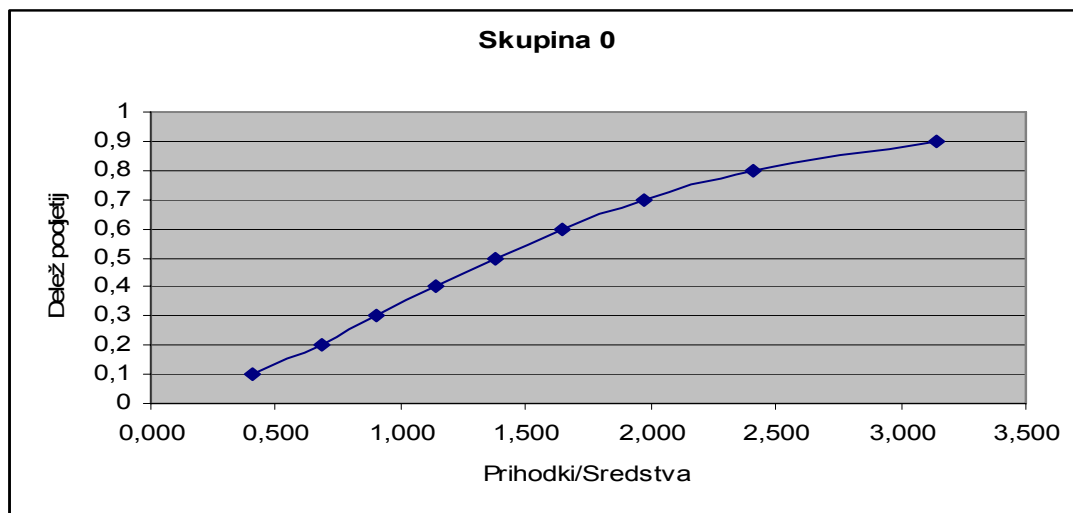
Vir: Ordination methods for Ecologists, 2007.

Slika 7: Graf lastnih vrednosti za vsako imed (štirih) glavnih komponent

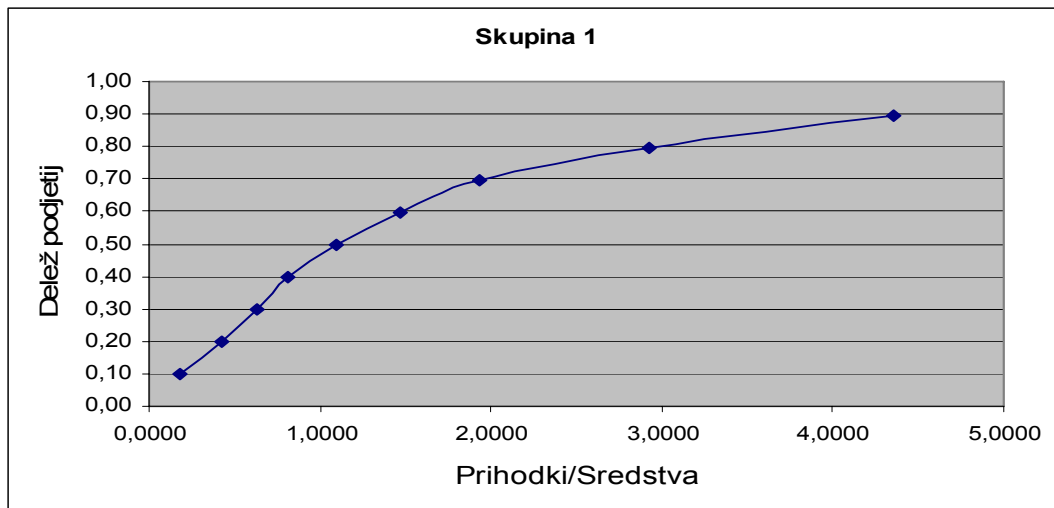


Vir: Ajpes; Lastni izračuni.

Slika 8: Nelinearna odvisnost deleža bankrotiranih podjetij od višine neodvisne spremenljivke X1(prihodki / sredstva) za obe skupini podjetij

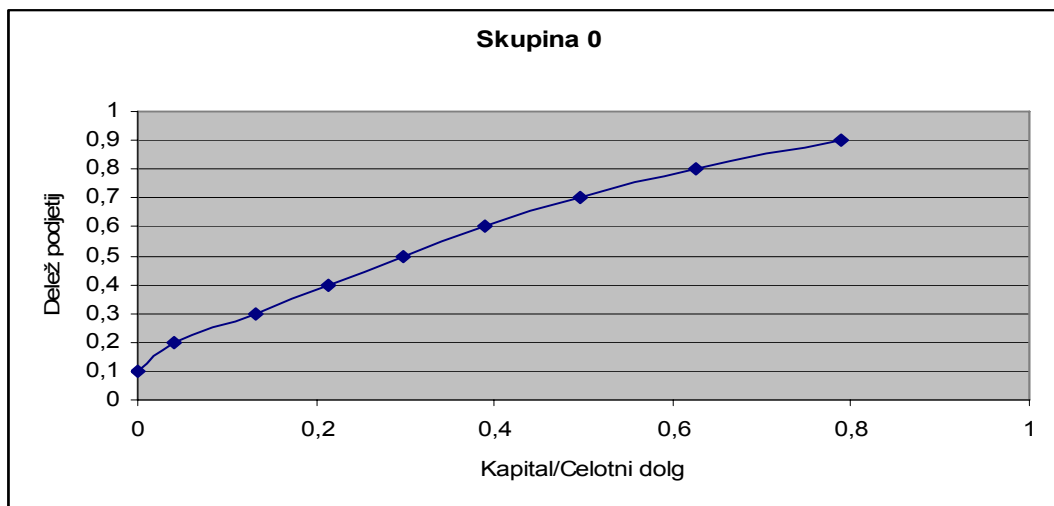


Vir: Ajpes; Lastni izračuni.

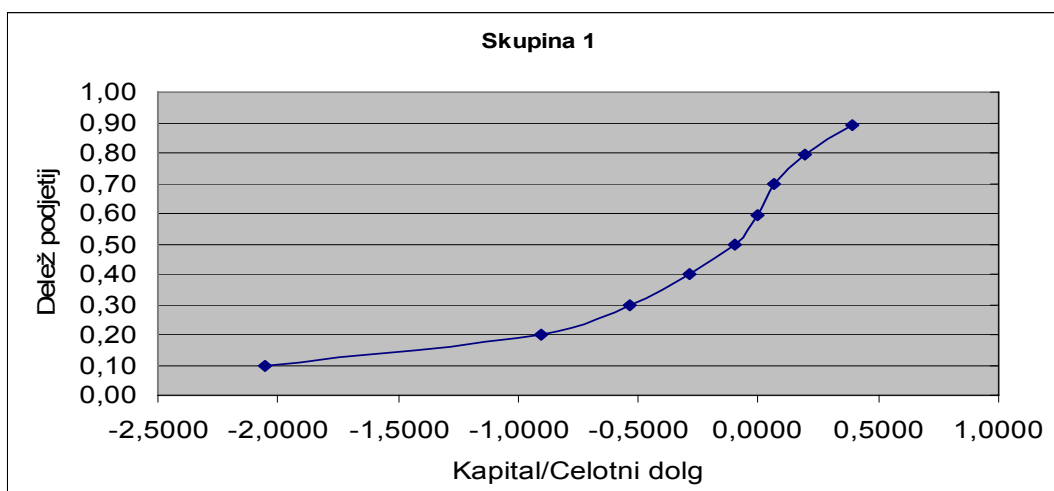


Vir: Ajpes; Lastni izračuni.

Slika 9: Nelinearna odvisnost deleža bankrotiranih podjetij od višine neodvisne spremenljivke X26 (kapital / celotni dolg) za obe skupini podjetij

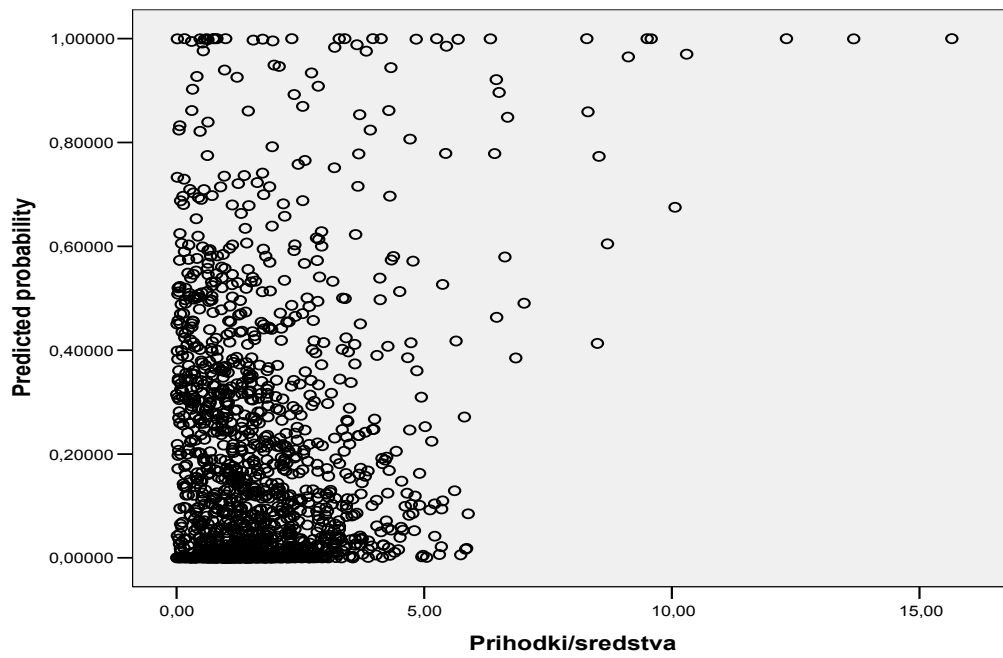


Vir: Ajpes; Lastni izračuni.



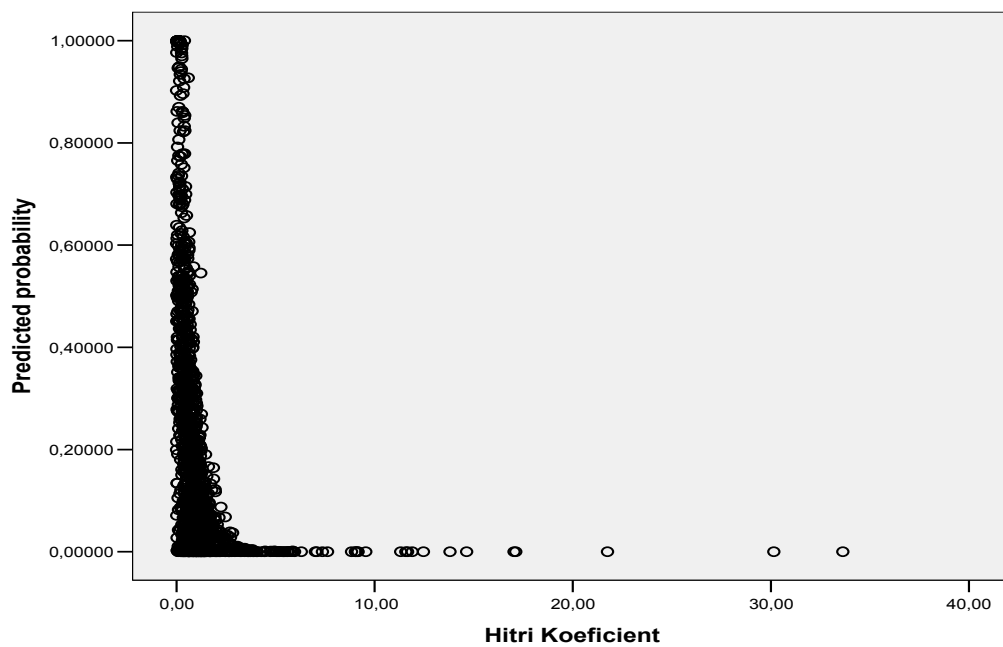
Vir: Ajpes; Lastni izračuni.

Slika 10: Verjetnost stečaja v odvisnosti od spremenljivke X1



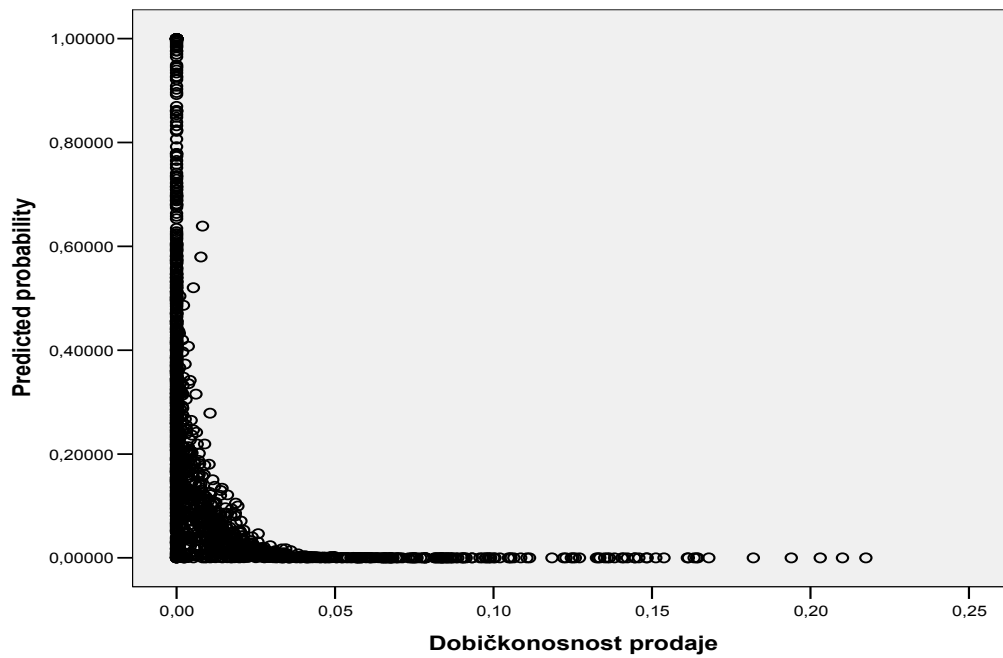
Vir: Ajpes; Lastni izračuni.

Slika 11: Verjetnost stečaja v odvisnosti od spremenljivke X11



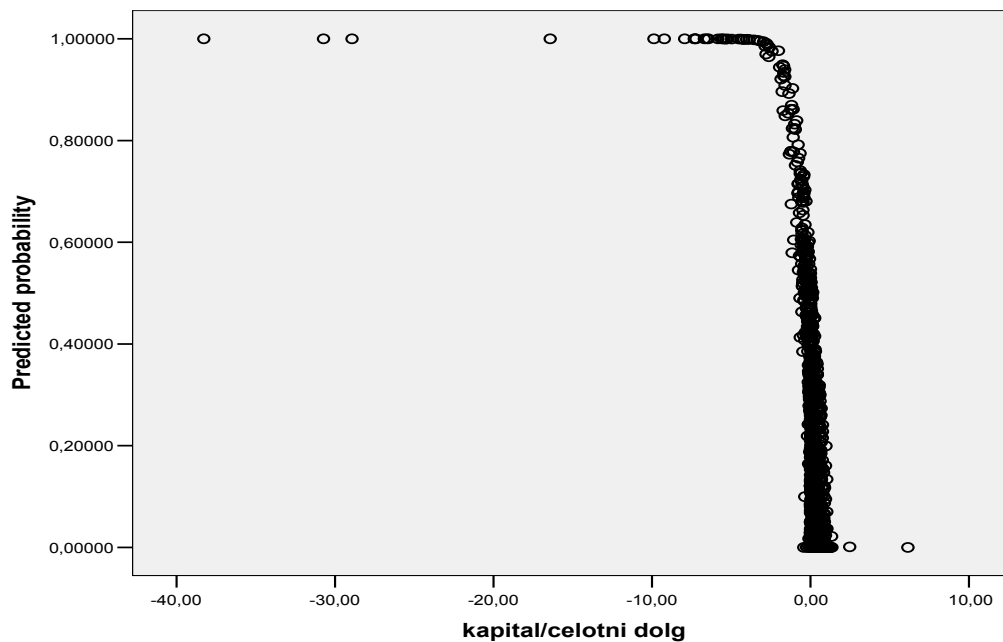
Vir: Ajpes; Lastni izračuni.

Slika 12: Verjetnost stečaja v odvisnosti od spremenljivke X17



Vir: Ajpes; Lastni izračuni.

Slika 13: Verjetnost stečaja v odvisnosti od spremenljivke X26



Vir: Ajpes; Lastni izračuni.