

SZABADOSNÉ NÉMETH ZSUZSANNA–DÁVID LÁSZLÓ

A KIS- ÉS KÖZÉPVÁLLALATI SZEGMENS MULASZTÁSI VALÓSZÍNŰSÉGÉNEK ELŐREJELZÉSE MAGYARORSZÁGI KÖRNYEZETBEN*

A Bazel II fogalmai, elvárásai lassan beépülnek a bankok mindennapi gyakorlatába. Nagyon sok értekezés, tanulmány született már, amely összefoglalta és rendszerezte a Bazel II előírásait. Jelen tanulmányban a szerzők arra vállalkoznak, hogy bemutatják a magyarországi kis- és középvállalatok mulasztási valószínűségének előrejelzésére alkalmas PD-modell egy lehetséges megvalósítását.

Az első részben az elméleti kereteket, a fő irányvonalakat és a módszertan fejlődését mutatják be, míg a második részben egy konkrét modell fejlesztésének elméleti és gyakorlati problémáira térnek ki. Tapasztalataikat és következtetéseiket konkrét portfólión elvégzett elemzésből merítették, részeredményeket az adatok titkossága miatt nem közölnek.¹

* Lektorálta: Király Júlia, Nemzetközi Bankárképző Központ, vezérigazgató.

¹ Szeretnénk köszönetet mondani a Bankárképző munkatársainak, akik a Bazel II-höz kapcsolódó szakmai ismereteikre és a közgazdasági tudásukra támaszkodva folyamatosan segítséget nyújtottak a modellezés lépéseiben, biztosítva ezzel, hogy a kialakításra kerülő modell mind az ajánlásoknak, mind a matematikai-közgazdasági elvárásoknak megfeleljen. Az elemzés a Budapest Bank kis- és középvállalati portfólióján készült, szeretnénk megköszönni a Bank dolgozóinak a modellezésben való részvételüket és szakmai támogatásukat.

**A KÖZÉP-KELET-EURÓPAI PIACOK
HITELMODELLEZÉSI GYAKORLATÁNAK
RÖVID TÖRTÉNETE**

Míg a banki hitelezés története 5000 évre tekint vissza, addig az ügyfelekre számított „rating”² csupán 50 éves hagyományal rendelkezik. A diszkriminanciaanalízis statisztikai interpretációja Fisher nevéhez fűződik. A banki gyakorlatban a módszer alkalmazási lehetőségét Durand vetette fel 1941-ben. Ezután jelentős fejlődésen ment keresztül, a hitelezési portfólióban szereplő „jó” és a „rossz” ügyfelek hatékony megkülönböztetését segítő módszerek alkalmazására egyre újabb és újabb technikák jelentek meg a diszkriminanciaanalízis területén, amelyek átszivárogtak a hitelezési piacokra is.

A credit scoring gyakorlati alkalmazhatóságát az 1960-as évek során egyre szélesebb körben elterjedt hitelkártya-kibocsátás végképp bebizonyította. A credit score-t elsőként használó cégek tapasztalatai azt mutatták, hogy számos folyamatot lehet automatizálni, meggyorsítani és a „defaultosok”³ arányát akár 50%-kal lehetett csökkenteni. A credit scoring teljes kiteljesedését és általános elfogadottságát

az 1975–76-ban hozott Equal Credit Opportunity Act tette lehetővé.

Az 1980-as években – a korábban csak a hitelkártya-kibocsátásoknál használt – credit scoring rendszereket fokozatosan kezdték használni más termékekre is, például a fogyasztási hitelek vagy a direct marketing területén az 1990-es években. A scorecard-építés technikájában a húzóerő az 1980-as években a logisztikus regresszió és a lineáris programozás lett, míg az 1990-es években megjelentek a mesterséges intelligencia technikák, mint például a neurális hálók is.

Vállalati hitelkockázati modellek

A lakossági területen alkalmazott modellezési technikák fokozatosan kezdtek megjelenni a bankok vállalati hitelezési folyamataiban. Amíg a matematikai-statisztikai technikák tökéletesen megfelelték a lakossági „sokasághoz hasonlóságon” alapuló ügyfélminősítésének, addig a vállalati ügyfelek minősítését ezek a megoldások csak részben tudták kielégíteni. A hitelezési sajátosságok megkövetelték, hogy a vállalatok esetében egyedi, szakértői elemzésen alapuló szempontokat is beépítsenek a modellekbe. Fontos szempont tehát, hogy megtaláljuk a szegmensnek, a portfóliónak megfelelő hitelkockázati modellt. Hiszen egy nem megfelelő modell alkalmazása hibákat, kockázatokat visz a hitelezési folyamatba, aminek eredménybefolyásoló hatása lehet.

A hitelkockázati modellek fejlődésében három fő időszakot különböztethetünk meg.

- 2 A gyakorlatban az angol „rating” kifejezés a magyar „minősítés” szóval azonosítható, ugyanakkor tartalmában tágabb foglalat jelent, mert az egész mögöttes modellre, számolási logikára utal. Mivel ez a magyar megfelelőjéből hiányzik, illetve a magyar szakirodalomban is gyakran használják az eredeti fogalmat, így a továbbiakban mi is a „rating” kifejezést használjuk.
- 3 A „default” elnevezés a Bazel II előírásaiban szerepel, jelentése röviden: az ügyfél adósságát nem tudja időben rendezni, fizetési problémája keletkezik, azaz mulaszt. Részletes definíciót *A modellezés alapjául szolgáló adatbázis összeállítás*a részénél adunk.

Az *első generációs* modellek a 1960-as, 70-es évekre tehetőek. Jellemzőjük, hogy számos olyan feltételezésen alapultak, amelyek akkoriban kifinomultnak számítottak, de mára bebizonyosodott, hogy jelentősen csorbítják a modellek hatékonyságát. Általánosan elmondható azonban, hogy ezek a kutatások nagy előrelépést jelentettek a véletlen mozgások (random walk) modellezésében, hiszen ezek hatására a determinisztikus változók helyett megjelentek a sztochasztikus változók.

A modellalkotás legfontosabb képviselői:

- Black–Scholes: Opcióárazási modell [1973]
- az eredeti Merton-modell [1974]
- a kamatlábmozgások egyfaktoros modellje
- Jamshidian: kötvényopció-modell (1989)

A modellek egyik közös jellemzője, hogy valamennyi a cég eszközértékének a fejlődését feltételezi, illetve hogy a modellekben használt hozamgörbék azonosítják a kockázatkerülő befektetők hozamelvárásaival. A Black–Schole-féle opcióárazási modellben (BS-modell) azzal a kiindulási ponttal találkozunk, hogy a vételi vagy az eladási opció vásárlója biztosan fizetni fog, míg az opció alaptermékének az ára bizonytalanságot jelent.

Merton ismert tanulmányában a Black–Scholes-modellből kiindulva olyan kötvényeknek az árazását vezeti le, melyek működésében benne van a bedőlés kockázata (diszkont és kupon kötvényekre is).

A kutatók közül többen Nobel-díjat kaptak közgazdasági eredményeikért.

A *második generációs* modellek megkísérelték a hitelkockázatok területén alkalmazni az első generációs modelleket. A hitelkockázat sajátosságai azonban mindenképpen megkövetelték a korábbi modellek módosítását. Az egyik legismertebb második generációs modell a Merton által készített a „Kockázatos adósság”⁴ elemekre épülő modell. Kiinduló pontja, hogy a részvény nem más, mint egy Black–Scholes call opció a vállalat eszközeire. A modellben egyetlen kockázatos változóként a cég piaci értéke szerepel, a a BS-modellhez hasonlóan a kamatlábat ez a modell is konstansnak tekintette.

Egy következő példa lehet a második generációs modellekre a portfólió értékének Monte Carlo-szimulációja. Kiindulópont az első generációs modellek eszközérték- meghatározása egy default/non-default változó bekapcsolásával. Ezen adatok alapján szimulálták az adós hitelfizető képességét, illetve a teljes portfólió értékét.

A *harmadik generációs* modelleknek két közös jellemzőjük van. A korábbi egy véletlen változóval szemben legalább két véletlen változót tartalmaznak, illetve az ügyfél- defaulttal kapcsolatban határozott feltételezéssel élnek. Az 1990-es években nagyon sok modell született ebben a szelvényben Ide tartozott a

- Jarrow–Turnbull-modell, amelyben a default folyamat és a kamatláb szerepelnek mint véletlen változók (1995)
- Jarrow–Lando–Turnbull-modell: ebben a default valószínűségek számításánál megjelenik az átmenetmátrix (1998)

4 Merton Model of Risky Debt.

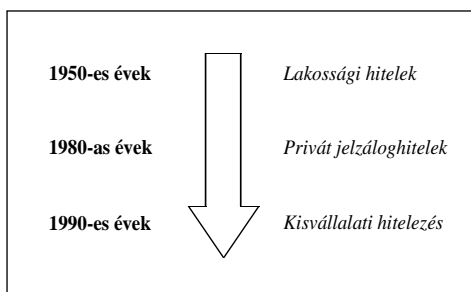
- Shimko–Tejima–van Deventer-modell: a Merton-modell módosítása, a konstans kamatlábakat véletlen változóra cserélték (1993)

A kelet-közép-európai (továbbiakban: KKE) hitelezési gyakorlat is ezen a fejlődési folyamaton ment keresztül, ám a fejlett amerikai bankokkal összehasonlítva számos különbséget fedezhetünk fel:

- a régió számos országára igaz, hogy a fenti fejlődési folyamat csak jóval később kezdődik el;
- általánosságban elmondható, hogy a KKE régió országainak egyikében sem működik olyan európai szintű rating cég, amely nemzetközi szinten is elismert minősítést publikál. Így a bankok az ezen intézmények által adott vállalati ratinget nem építhetik be saját belső adóminősítési folyamatukba;
- az előző tulajdonságok miatt a KKE régió bankjainak az amerikai gyakorlattól (best practice) eltérő szakmai bázisra kell építeniük a hitelezési folyamatot, rendszert.

Jelenleg a rating modellek felhasználási területe egyre gazdagabb, ezt a fejlődési folyamatot az 1. ábrán követhetjük nyomon.

1. ábra



Mivel a KKE bankok döntő többségükben fejlett anyabankok leányvállalatai, ezért nagyon gyakori megoldás, hogy a nyugaton már bevált módszert kívánják adaptálni az adott országban. Általában a nagy forgalommal és standardizált termékekkel jellemezhető lakossági és jelzáloghitelezés területén kisebb-nagyobb módosításokkal lehet alkalmazni ezeket a régióta bevált módszereket. Ám a kisvállalati hitelezés egyedi tulajdonságai miatt ez szinte lehetetlen. A vállalati szereplők működése nagyban függ az adott ország gazdasági helyzetétől, szektorális felépítésétől és a piac jellegétől. Ez nyilvánvalóan nemcsak a nyugat-európai és kelet-európai összehasonlításban áll fent, hanem a KKE régió egyes országai között is. A vállalati hitelezés mindig tartalmaz egyedi vonásokat, nem lehet standard termékekkel megcélózni az ügyfeleket. Minden vállalat saját egyedi igényeire kell kialakítani a hitelezési termékeket.

A HATÉKONY ADÓMINŐSÍTÉS

E rövid áttekintés után vizsgáljuk meg, hogy mi a célja egy rating modellnek, illetve mi az elvárásunk egy hatékony adóminősítő modellel szemben.

A rating alapvető célja a hitelezési folyamatban az, hogy feltárja az egyes adósok kockázatát, és így segítséget nyújtson a döntéshozóknak a hiteligeny elbírálásában. A Bazel II szabályainak bevezetésével azonban a ratinget már nemcsak a hitelezési folyamat döntéshozói használják, hanem szabályozói elvárás lett. A felügyelet a tőkekövetelményt a hitelintézet

belső rendszerei alapján határozza meg, amihez a rating szolgáltatja az információt az adósok kockázatáról. Így a ratingnek kettős szerepe lett: objektív mérce a döntéshozóknak és a tőkekövetelmény meghatározásának alapja. A következőkben nézzük meg, hogy ilyen keretek között hogyan is kell felépülnie és milyen alapvető kritériumokat kell teljesítenie egy hatékony rating rendszernek!

Egy hatékony adósminősítési modellnek ki kell elégítenie az alább felsorolt elvárásokat.

- Legyen széles körű, azaz legyen képes minősíteni minden múltbeli, jelenbeli és jövőbeli ügyfelet.
- Megbízhatónak kell lennie, nem függhet az adósminősítő személyétől, vagy az adósminősítés időpontjától.
- Legyen közgazdaságilag értelmes, mind külön-külön az egyes mutatók, mind a mutatók közötti tartalom tekintetében.
- Számolnia kell minden hozzáférhető információval.
- Ne legyen túlságosan bonyolult.
- Lehetőleg annyi változót tartalmazzon, amennyi szükséges, illetve lehetőleg annyira kevés változót tartalmazzon, amennyire lehetséges.
- Fel kell készíteni a rendszert, hogy az üzletkötők a szubjektív elemeket jobbra értékelhetik, mint az objektíveket, így torzíthatják a modell eredményét.
- A minősítő rendszernek annyi kategóriával kell rendelkeznie, amennyinél még megfelelően elválnak egymástól az ügyfelek.
- A hasonló kockázatú cégeknek azonos kategóriába kell kerülniük (homogén csoport alkotási képesség).
- Az előzetesen becsült (ex ante) PD értékek nem különbözhetnek jelentősen az utólag tapasztalt (ex post) PD értékektől, a modellezés célja ennek összehangba hozása (előrejelző képesség).
- A modell hatékonyan szeparálja a „rossz” és „jó” ügyfeleket. (Lásd alfa-béta hiba grafikon.)
- A szakértők által elfogadható legyen a modell.
- A modellben legyen kicsi a korreláció a pénzügyi mutatók között.
- A modellnek tartalmaznia kell az eltérő iparágak eltérő kockázatát.
- Növelje a modell a hitelezési folyamat hatékonyságát (ne kerüljön több időbe adott hitelkérelem elbírálása, mint korábban).
- A modellnek igazodnia kell a baseli előírásokhoz (basel-konformitás).
- A modellnek a felügyeleti szerv előírásaival is összhangban kell lennie (home-konformitás).
- Ugyanakkor a modell ne igazodjon teljes mértékben a kialakításkor jellemző portfólió sajátosságaihoz (ne legyen overfit), mert adott portfólió-átrendeződéskor nagyon félreviheti a minősítéseket. → Ezt a célt szolgálja többek között az out-of-sample teszt.

A fenti kritériumokkal lényegében összegyűjtöttük, hogy milyen „feladatok” kell teljesítenie a modellnek. Mielőtt azonban a konkrét modellfejlesztéshez hozzáférnénk vizsgáljuk meg, hogy milyen megközelítéseket alkalmazhatunk.

VÁLASZTHATÓ HITELMINŐSÍTÉSI MEGKÖZELÍTÉSEK

Két fő hitelminősítési filozófia létezik, a „*Gazdasági ciklus modell*” (angol elnevezés szerint: *Through the Cycle*) és az „*Azonnali minősítés*” (angol elnevezés szerint: *Point in Time Rating*).⁵ Mind a két technikának vannak olyan sajátosságai, amiket az elemezni kívánt portfólió ismeretében mérlegelnünk kell.

A *Gazdasági ciklus modell* hosszú időszakot – egy vagy több üzleti vagy akár gazdasági ciklust – vizsgál, hosszú távú minősítés megállapítására törekszik, és nem célja a vállalat pillanatnyi helyzetének, illetve közeljövőjének a bemutatása. Ilyet alkalmaznak a nagy minősítő ügynökségek, mint a Standard and Poor’s vagy a Moody’s. Számukra az a fontos, hogy a minősítések viszonylag stabil szinten mozogjanak, ne befolyásolják rövid távú – 1-2 éves – események.

Ezzel szemben az *Azonnali minősítés* az adós helyzetét és ennek legvalószínűbb változását próbálja előrejelezni egy meghatározott időszakra, például 1 évre. A minősítés azonnal megváltozik, amint az adós helyzete módosul, akár egy üzleti cikluson belül is. Ilyen például a KMV Merton típusú modellje, amit a bankok szívesen alkalmaznak belső hitelminősítő modellként.

A Bázeli elvárásaiban nincs egyértelműen előírás arra vonatkozóan, hogy melyik minősítési filozófiát kell alkalmazni a tőkekövetelmény számítása során. Több

kutató elemezte a két filozófia hatásait a tőkekövetelményre. Carey és Hrcay (2001) azt vizsgálták, hogy milyen hatása lesz, ha a bankok a külső minősítéseket alkalmazzák a belső elbírálási folyamatokban. Crouchy, Gallai és Mark (2003) viszont azt bizonyították, hogy az Azonnali minősítés az alkalmasabb a tőkekövetelmény számítására. Catarineu-Rabell (2003) az Azonnali minősítéssel kapcsolatban felvetette a tőkekövetelmény-változás pro-ciklikusságának problematikáját.

A tőkekövetelmény szempontjából a módszertan kiválasztásánál a generált mulasztási valószínűségek mellett az eszközök korrelációját is vizsgálnunk kell. A Gazdasági ciklus modell stabil mulasztási valószínűséget és magas korrelációt, míg az Azonnali minősítés időben változó mulasztási valószínűséget és alacsony korrelációt mutat. Ennek következtében a két módszertan különböző tőkekövetelményt határoz meg a hitelintézet számára.

Nézzük meg, hogyan alakulhat a tőkekövetelmény különböző gazdasági ciklusokban.

Recessziós időszakban az adósok késedelmei emelkednek, aminek következtében az Azonnali minősítési modell magasabb mulasztási valószínűséget generál, és egyben magasabb tőkekövetelményt határoz meg. Ezzel szemben a Gazdasági ciklus modell esetében az adósok késedelmének növekedése nem okoz növekedést a mulasztási valószínűségek becslésében, így a tőkekövetelmény változatlan marad.

Gazdasági fellendülés esetén épp az ellenkezője játszódik le. Ha feltételezzük, hogy a Bázeli II-ben leírt szabályozás eredeti célja az volt, hogy egyrészt átlátha-

⁵ A problémával a Bázeli Bizottság is foglalkozott, lásd Basel Committee on Banking Supervision (2000a,b): *Through the Cycle versus Point in Time Ratings*.

többá tegye a hitelezés kockázatait, másrészt a valós kockázatok alapján határozza meg a tőkekövetelményt, akkor teljesen természetes, hogy recesszióban növekszik a késedelembe eső adósok száma. Ugyanakkor biztosítani kell, hogy a két módszertan azonos tőkekövetelményt írjon elő, ne adjon lehetőséget szabályozói arbitrázsra. Megoldást jelenthet, ha az Azonnali minősítés tőkekövetelményének rövid távú ingadozását kompenzáljuk a módszertan alacsonyabb állóeszköz-korrelációja miatt. Vélelmezzük, hogy a végleges szabályozás nem teszi lehetővé, hogy a különböző módszertanok más tőkekövetelményt állapítsanak meg.

Összefoglalásként a két módszertan definíciójára vonatkozóan megállapíthatjuk, hogy míg az Azonnali minősítés módszertan annak a valószínűségét tükrözi, hogy az adós mulasztani fog-e például az elkövetkező 1 évben, addig a Gazdasági ciklus modell definíciója nem ennyire meghatározott. A tanulmányokból az is kitűnik, hogy a két modell eltérő információtartalommal dolgozik. A Gazdasági ciklus modell nem tartalmaz minden olyan információt, ami adott pillanatban az adósról fellelhető, helyette több gazdasági ciklusban összegyűjtött információra támaszkodik, és az ezen adatok alapján felépített modellből hosszú távú, stabil előrejelzést készít. Nem szabad figyelmen kívül hagynunk, hogy ez a modell egyes gazdasági időszakokban nagyon nehézkes lehet, például amikor egy cég vevőállománya visszaesik, nem olyan könnyű megmondani, hogy ez a ciklus miatt vagy a vállalat „immanens” tulajdonságai miatt következett be.

Ezek figyelembevételével, illetőleg azal a feltételezéssel, hogy a tőkekövetelmény- számítás szempontjából a két módszertan azonos elvárásokat számszerűsít, megállapítható, hogy a magyarországi gyakorlatban az Azonnali minősítési modellek alkalmazása javasolt.

VÁLASZTATÓ MODELLFEJLESZTÉSI MEGKÖZELÍTÉSEK

1. Szakértői modell

Amikor a szakértői rendszereket, döntési struktúrákat, hitelezési szabályokat, a vezetőség kockázatvállalási hajlandóságát egy konkrét matematikai modellbe kell beépíteni, akkor szokták a nagyobb részt egyéni tapasztalatokra, véleményekre épülő adósminősítési modellt használni (*human judgement*). Ebben az esetben általában egyedi elbírálást igénylő, nagy hitelösszegű ügyfelek hitelezéséről van szó. A termékek nem sztenderdek. A modell kialakításához nem áll rendelkezésre elegendő adat. A hazai gyakorlatban elsősorban a *nagyvállalati hitelezésnél* használják ezt a módszert.

Előnyök:

- nem szenved csorbát a közgazdasági tartalom,
- az ügyfél igényeihez mindig teljeskörűen lehet alkalmazkodni.

Hátrányok:

- lassú modellkialakítási procedúra,
- az eltérő vélemények és tapasztalatok összhangba hozása sokszor nem egyszerű,

- a döntési folyamat mindig időigényes marad, mert az elbírálás egyedileg történik (a modell csak kiinduló pont).

2. Statisztikai modell

A csak matematikai-ökonometriai alapon kialakított modellt olyan hitelezési folyamatban lehet használni, ahol a cél minél gyorsabb döntéshozatal megfelelő előrejelzési pontosság mellett. A hitelporfólióra elsősorban a nagy számú, kisebb hitelösszegű ügyfélállomány jellemző, amelyeket sztenderd termékekkel céloznak meg. A tisztán statisztikai modellek szükséges, de nem elégséges feltétele, hogy megfelelő mennyiségű és minőségű adat álljon rendelkezésre. A hazai gyakorlatban elsősorban a *lakossági hitelezésben* használatos.

Előnyök:

- viszonylag gyors modellkialakítás,
- rövid döntési ciklusidő.

Hátrányok:

- gyakran figyelmen kívül hagyja az egyes változók kiválasztásakor a közgazdasági értelmet,
- nincs lehetőség egyedi szakértői mérlegelésre.

3. Hibrid modell

A két előző modell vegyítése. Akkor használatos, ha bizonyos fokú automatizmust ugyan kell adni a hitelezési döntésnek, ám a statisztika mellett rendkívül fontos szerepe van a szakértői tapasztalatok beépítésének a modellezésbe. Nem szükséges a lakossági hitelezésben tapasztalható adatbővességgel rendelkezni. A szakértők eddigi

tapasztalatának és a statisztikai-ökonometriai módszereknek a harmonizációjáról van szó. A hazai gyakorlatban ezt a megoldást a *kis- és középvállalatok hitelezésében* (KKV szegmens) célszerű alkalmazni.

Előnyök:

- a modellben nem szenved csorbát a közgazdasági tartalom,
- a modell kialakításában a szakértők is részt vesznek, így nagyobb az elkötelezettségük,
- rövid a döntési ciklusidő, hiszen az egyedi elbírálási feltételeket is beépítettük a modellbe.

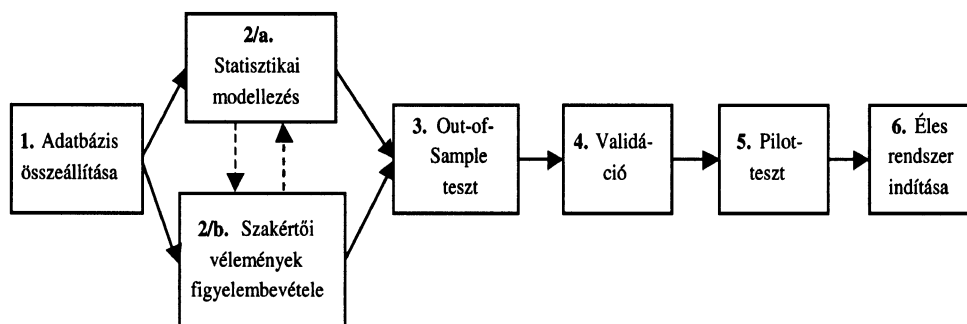
Hátrányok:

- a modell kialakítása időigényes,
- nehéz elnyerni a szakértők elkötelezettségét.

A kis- és középvállalati szegmens jellemzőit tekintve sorolható a lakossági portfólióba, hiszen a cégek mérete és az egy adóssal szembeni kitettség a Nyugat-Európai szinten belefér a lakossági definícióba. Hiba lenne azonban azt gondolni, hogy Magyarországon a KKV szegmensben egy statisztikai modell mindenre úgy tud választ adni, mint a lakossági szegmensben. Így a lakossági portfóliónál használt módszerek jó kiindulópontot nyújtanak, de a banki üzemben dolgozók véleménye és piaci ismeretei nélkülözhetetlenek egy jó modell sikeres bevezetéséhez. Ha a másik oldalról vizsgáljuk a döntéshozatalt, a szakértők véleménye önmagában szintén nem elegendő, hiszen az ügyfeleket ugyan egyedileg bírálják el, de erősen támaszkodnak a statisztikai alapon kialakított homogén csoportokra.

Erre a szegmensre a fent bemutatott hibrid modellt lehet alkalmazni, amely

2. ábra



részben statisztikai, részben szakértői elemekre épít. A modell előnyei ötvözik mind a lakossági, mind a szakértői modell pozitívumait, hátrányai pedig kezelhetők.

A MODELLEZÉS FŐBB LÉPÉSEINEK BEMUTATÁSA

A 2. ábrán követhetjük nyomon a mulasztási valószínűség előrejelző modell készítését (továbbiakban: PD-modellezés⁶).

Alapvetően öt nagy szakaszra bontható a folyamat.

1. A modellezés alapjául szolgáló adatbázis összeállítása

A lakossági scorecard-építési gyakorlatban gyakran használt megoldás, hogy a nagy adathalmazból megfelelő módszerrel vett mintán történik a modellfejlesztés. Default arány tekintetében általában 50-50% vagy 25-75%-os összetételű mintavételt szoktak alkalmazni, a kialakított

⁶ A PD-modell a Bazel II szabályaiban szereplő „Probability of Default”-ot, a mulasztási valószínűséget próbálja előrejelezni.

modellt pedig a maradék adathalmazon tesztelik vissza. Mivel a vállalati hitelezésben sokkal kevesebb ügyfél kerül a portfólióba, ezért az esetek többségében a mintavételezést nem lehet megoldani. Különösen igaz ez a hazai hitelezési gyakorlatban. A mintavételezés helyett a teljes portfólión történő elemzés jelenthet alternatív megoldást,⁷ azzal együtt, hogy a modell visszatesztelése a banki portfóliótól független mintán történhet. Az adatbázis összeállítása az alábbi alfolyamatokra bontható:

a) *A portfólióban szereplő minősítésekre vonatkozó egyedi, pénzügyi adatok és default-események összegyűjtése*

Célszerű több évre visszamenőleg elkészíteni az alapadatbázist. Az alapadatokat és a pénzügyi adatok mellett javasolt minden olyan egyedi információ összegyűjtése, amiket az alkalmazott hitelezési folyamatban bekérnek.

⁷ Az adatbázis összeállítása az elegendő darabszám eléréséhez több év adataiból történhet, így egy adott ügyfél többször is szerepelhet az adatbázisban. Mivel eredeti célunk egy „application” típusú scoring rendszer kidolgozása, így azok az ügyfelek, akik több even keresztül szerepelnek a portfólióban, minősítések egymástól független eseménynek tekinthetők.

A „default” definíciójának meghatározása kritikus pont lehet.

Bázell II előírások alapján „default”-nak nevezzük, ha

- az ügyfél 90 napos késedelembe esik és/vagy
- úgy néz ki, hogy az ügyfél nem tudja visszafizetni az adósságait (ilyen lehet: csődbe megy, adósságátütemezést kér, fizetési nehézségei jelentkeznek stb).

A Bázell II előírása egy minimum elvárás, de használhatunk szigorúbb „default” definíciót is. A szigorúbb „default” – például 30 napos késedelem – meghatározásának oka lehet, hogy növelni szeretnénk a mintában a bedőlt események számát. A mintaszám növelése egyben a fejlesztendő modell megbízhatóságát is növeli, hiszen a statisztikai modellek erősségének egyik alapja a minta elemszáma. A szigorúbb definíció ugyanakkor hátrányt is jelenthet, hiszen magasabb kockázatot mutat, így magasabb tőkekövetelményt eredményez. Ez a probléma azonban kiküszöbölhető, ha a fejlesztésnél a szigorúbb definíciót alkalmazzuk, viszont a tőkekövetelmény meghatározásánál csak a minimum elvárásnak megfelelőt.

b) Adatgyűjtés az elutasított ügyfelekről (reject inference kezelése)

Nagy probléma a scorecard és a rating modellek fejlesztésében, hogy amennyiben a meglévő portfólió adataiból indulunk ki, csak az elfogadott ügyfelek adataiból modellezhetünk. Még ha gyűjtjük is az elutasított ügyfelekről az információt, akkor sem tudjuk egyér-

telműen eldönteni, hogy ők vajon „rosszak” vagy „jóak” lettek volna. Ez a probléma a modellezésnek nyilván bizonyos korlátot teremt. Számos technika született, mely megpróbálja kezelni ezt a problémát. A problémának és megoldási módszereinek összefoglaló neve: reject inference.⁸ Manapság rengeteg vita születik arról, hogy milyen mértékű és jellegű hibát követhetünk el a modellezés során, ha nem megfelelően kezeljük a reject inference-t.

Elméletileg az egyik legtisztább módja a reject inference megszüntetésének, hogy beengedünk mindenkit, és majd a portfólióban eldől, hogy ki lesz a „rossz” és ki a „jó”. Ez természetesen teljesen eliminálja a scorecardok kockázatmérés és -kezelés tulajdonságát. A gyakorlatban ez nem egy járható út. A szakirodalomban alapvetően 5 fajta reject inference kezelési lehetőséget említhetünk.

1. Az elutasítottakat „rossznak” tekinteni

Ez az egyik legegyszerűbb mód, amely ugyanolyan mértékben követ el hibát, mint amennyire egyszerű. Nagy gondot okoz, hogy ezzel a módszerrel valamilyen szinten összekeveredik az alfa és béta hiba. Adott bank hitelezési folyamatában elfogadott „default” definíció is csorbát szenved.

⁸ A „reject inference” problémával a szakirodalom nagyon részletesen foglalkozik, ennek a tanulmánynak a keretében csupán felvillantjuk, hogy ez milyen kérdéseket vet fel a KKV modellezésben, de nem célunk a teljes irodalom bemutatása.

2. Extrapoláció

Hend és Henley (1993) a reject inference kutatásai során rámutattak arra, hogy alapvetően két szituáció létezik az X_{old} és X_{new} viszonylatában. X_{old} a jelenlegi scorecard alapját képező sokaság (tehát csak az elfogadottakat tartalmazó sokaság), X_{new} pedig az újonnan fejlesztendő scorecard alapjául szolgáló sokaság. Amennyiben $X_{old} \in X_{new}$, akkor azokról az ügyfelekről, melyeket nem fogadott el a régi scorecard, nem tudjuk a „rossz”/„jó” megoszlást. Éppen ezért egyfajta megoldás lehet, hogy az elfogadottak adatbázisán (X_{old}) kifejlesztjük a scorecardot, és ezt valamilyen módon extrapoláljuk a visszautasított ügyfelek adataira.

3. Augmentáció

Abban az esetben, ha X_{old} nem részalmazza X_{new} -nek, akkor a helyzet az előzőhöz képest némileg bonyolultabb. Hsia 1978-ban vázolta fel a probléma egy lehetséges megoldási módját. Először az elfogadottakra kell építeni egy modellt, mely megmondja, hogy kik a „rosszak” és „jók”: $P(G|x, A)$, ahol x a magyarázó változók. Ezután az előzővel hasonló technikával egy olyan modellt kell készíteni, mely meghatározza, kiket fogadtunk el vagy utasítottunk el: $P(A|x) = P(A|s(x)) = P(A|s)$, ahol s az elfogadás/visszautasítás scoreja. A továbbiakban azzal a feltételezéssel élünk, hogy $P(G|s, R) = P(G|s, A)$. Ennek tudatában újrasúlyozzuk a scorecard-fejlesztés alapjául szolgáló mintát. A legnagyobb

hiányossága ennek a módszernek, hogy az elfogadottak és a visszautasítottak körében hasonló „jó/rossz” arányt feltételezünk.

4. Eloszlások keverése

Abból a megközelítésből indul ki, hogy feltételezi, hogy a minta két eloszlásból tevődik össze: „jók” és „rosszak” eloszlása.

$$p(x) = p(x|G)p_G + p(x|B)p_B.$$

A $p(x|G)$ és $p(x|B)$ végső paramétereit külön-külön becsülik egy iteratív eljárás segítségével.

5. Három csoport megközelítés

A mintát itt 3 csoportra bontjuk: „jók”, „rosszak” és „visszautasítottak”. A problémát az jelenti, hogy ebből a 3 csoportból végül 2 csoportot kell készítenünk, hiszen a hitelezési döntésbe a „visszautasított” státus nem kerülhet be változóként.

Általános megegyezés alapján racionális feltételezni, hogy

$$p(G|s, R) < p(G|s, A).$$

Bár a kis- és középvállalati szegmens hiteligenyét tekintve hasonlít a lakossági szegmensre, hitelezési folyamata inkább a vállalati hitelezési folyamat sajátosságait viseli magán. Ez pedig azt jelenti, hogy igazi elutasított ügyfelet (leszámítva a jogi problémákkal, a csalás gyanúval, tiltott tevékenységgel rendelkezőket stb.) nem találunk, így klasszikus reject inference-szel nem számolhatunk. Az elutasítás helyett a hitelezés három dimenzió mentén mozog, úgymint termék, biztosíték és ár. Ténylegesen az ügyfél dönt, hogy a kínált terméket, adott feltételekkel és kínált áron elfogadja vagy sem.

c) *Adatok tisztítása*

A hiányzó adatokat, illetve a „rossz” adatokat mindenképpen kezelni kell. Célszerű minden esetben ellenőrizni a mérleg és eredménykimutatás nevezetes egyenlőségeit. A majdani pénzügyi mutatók számításánál a nevezőkben szereplő pénzügyi adatokat úgy kell módosítani, hogy ne alakulhasson ki nullával való osztás.⁹

d) *A pénzügyi mutatók kiszámítása*

Általában érdemes az összes mutatót kiszámítani (ez akár 100-200 mutató is lehetséges), mely a rendelkezésre álló pénzügyi adatokból kvantifikálható. Az egyes pénzügyi indikátorokat mindenképpen csoportosítanunk kell közgazdasági értelmük és jellegük alapján.¹⁰

e) *Az „outlierek” kezelése*

A legáltalánosabban elfogadott módszer a kiugró értékek kezelésére, hogy az egyes mutatóknál a felső (1-5-ödik) percentiliséknél nagyobb értékeket a felső percentilissel, az alsó (1-5-ödik) percentiliséknél kisebb értéket pedig az alsó percentilissel helyettesítjük.

Egy másik módszert a logit vagy probit átalakítás jelenti. Egy extrém csúcsos, leptokurtikus eloszlással rendelkező megfigyeléshalmaz esetében a logit és a probit transzformáció kiugró értékektől mentes eloszlást hoz létre. A szakirodalomban leggyakrabban használt

logit-transzformáció a következő alakot ölti:

$$x_i^{transz} = \frac{e^{x_i}}{1 + e^{x_i}}$$

Az átalakítás egy 0 és 1 közötti számot eredményez.

2. Statisztikai modellezés a szakértői vélemények figyelembevételével

A modellezés ebben az esetben egy iteratív (visszacsatolásokban gazdag) folyamat. A szakértői tapasztalatokat, véleményeket a statisztikai-ökonometriai módszertan támogatja. A modellezés alapjául szolgáló változók körét a szakértők bevonásával lehet csak kialakítani. A statisztikai modellezés számos módszer alapján történhet. Három főbb modellcsaládot említhetünk meg. *Lineáris módszerek* (pl.: lineáris regresszió, logisztikus regresszió, diszkriminancia analízis stb.); *nem-lineáris módszerek* (pl.: neurális hálók, kernel becslések, közeli szomszédok módszere stb.); *egyéb eszközök* (pl.: döntési fák, lineáris programozás, Bayes hálók stb.). A gyakorlatban az egyik leggyakrabban alkalmazott módszer a PROBIT:

$$g(p) = \Phi^{-1}(p) = \beta_0 + \beta^1 x + \varepsilon$$

- p : annak a valószínűsége, hogy az adott cég egy éven belül „default” állapotba kerül (az eredményváltozó bináris jellegű: konvencionálisan 1-et vesz fel, ha „defaultos” egy ügyfél, és 0-t, ha nem)

⁹ Nézzünk egy példát! A kamatfedezeti mutatóknál jelentkezhet a nullával való osztás problematikája, megoldás lehet, ha helyettesíthetjük a 0 Ft-ot 1 Ft-tal. A helyettesítést úgy kell meghatározni, hogy ne csorbuljon az adott mutató **közgazdasági értelme**.

¹⁰ A későbbi modellezésben az egyedi elemzéseknél az lesz a cél, hogy minden egyes mutatócsoportból a leghatékonyabb mutatókat válasszuk ki.

- $\Phi^{-1}(\cdot)$: a standard normális eloszlás eloszlásfüggvényének inverze
- β_0 : a regresszió konstans változó-független kezdő paramétere
- β' : a változó-függő paraméterek által alkotott vektor
- x : a magyarázó változók által alkotott vektor (az esetek döntő többségében az előzőekben említett pénzügyi indikátorok)
- ε : a regresszió során megfigyelhető hibavektor. Feltételezzük, hogy ezek standard normális eloszlást követnek.

A paraméterek becslésére különböző ökonometriai módszerek állnak rendelkezésre, a becslés nagyon gyakran maximum-likelihood becslési algoritmussal történik. Ám a likelihood-függvénynek nem minden esetben van véges megoldása. A függvény megoldhatósága függ annak a multidimenzionális térnek a szerkezetétől, melyet az alapadatbázis változói alkotnak. 3 alapvető szerkezetet definiálhatunk ebben a térben:

Legyen Y_i az eredményváltozó, x_i a magyarázó változók által alkotott vektor, b pedig a paraméterek által alkotott vektor.

1. Teljes szeparáció: létezik olyan b vektor, mely minden egyes megfigyelést tökéletesen el tud különíteni egymástól.

$$\begin{cases} b'x_i > 0 & Y_i = 1 \\ b'x_i < 0 & Y_i = 2 \end{cases}$$

Teljes szeparáció esetén nincs egyedi megoldás (b -re vonatkozó becslések száma végtelen). Ekkor a log-likelihood függvény hamar nullához kon-

vergál \rightarrow nincs megoldása az optimalizációnak.

2. Kvázi-teljes szeparáció: a megfigyeléseket nem lehet teljes egészében elkülöníteni egymástól, de létezik olyan b vektor, amelyre:

$$\begin{cases} b'x_i \geq 0 & Y_i = 1 \\ b'x_i \leq 0 & Y_i = 2 \end{cases}$$

Ebben az esetben sem véges a paraméterbecslések száma. Ekkor – hasonlóan a teljes szeparációhoz – a szórásmatrix nem korlátos, és a likelihood függvény egy nemzérő konstanshoz konvergál.

3. Átfedés: amennyiben a megfigyeléseket annyira sem lehet elkülöníteni egymástól, mint a kvázi-teljes szeparációnál, akkor átfedésről beszélünk. Ebben az esetben létezik egyedi b becslés.

A teljes és kvázi-teljes szeparáció tipikusan kis minták esetén előforduló probléma.

Amennyiben becslésre kerültek a paraméterek, akkor a PD a következő képlet alapján határozható meg (ez az előző egyenlet inverzének a felírása):

$$y = \beta_0 + \beta'x$$

$$PD = \int_{-\infty}^{y(x)} \frac{1}{\sqrt{2\pi}} e^{-\frac{y(x)}{2}} dy(x)$$

A fent bemutatott modell esetében az alábbi feltételezésekkel élünk:

– a regressziós hibavektorok standard normális eloszlást követnek

- a magyarázó változók egymástól függetlenek¹¹
- a magyarázó változók olyan teret alkotnak, melyben nincs sem teljes, sem kvázi-teljes szeparáció. Természetesen az lenne a megfelelő, hogy teljes legyen a szeparáció az adott sokaságban, a probléma viszont az, hogy teljes szeparáció a gyakorlatban nem létezik, illetve ha van, akkor a szeparáló síkok száma nem véges, és az optimalizáció nem vezet megoldásra.

Ebben a szakaszban az alábbi alfolyamatok jelennek meg:

a) *A modellezés alapjául szolgáló mutatók kiválasztása*

A korábban kialakított mutatócsoportokban minden egyes mutatóra célszerű ábrázolni, hogy milyen függvény-szerű összefüggés figyelhető meg a mutató értéke és a mintabeli „default-ráta”¹² között. Azokat érdemes a modellezésbe bevonni, melyeknél ez a kapcsolat monoton. Természetesen az összefüggés irányának összhangban kell lennie a mutató közgazdasági értelmével. Elemezni kell, hogy önmagában az egyes mutatók milyen mértékben határozzák meg a default tényét.¹³ A mutatók kiválasztásánál pedig figyelembe kell venni, hogy a lehető legki-

sebb korreláció legyen közöttük. Általában a kiválasztás során követendő szabály lehet, hogy mutatócsoportonként 2-3 mutatót válogassunk be a regresszió alapját képező inputhalmazba.

b) *Regresszió futtatása*

A inputmutatókból történő végső mutatókiválasztás alapvetően 3 módon történhet:

➤ Forward-módszer:

Minden egyes lépésben egy új mutató kerül kiválasztásra. A folyamat addig ismétlődik, míg minden kiválasztott mutató teljesíti a szignifikanciakritériumot. Amely mutató be lett választva, az a későbbi lépésekben nem kerül szelektálásra.

➤ Backward-módszer:

Az összes mutató kiválasztásra kerül a 0-adik lépésben. Minden egyes lépésben egy mutató kerül elhagyásra. A folyamat addig megy, míg minden kiválasztott mutató teljesíti a szignifikanciakritériumot. Amely mutató el lett távolítva az a későbbi lépésben már nem kerül vissza.

➤ Stepwise-módszer:

Az előző kettő vegyítése. Minden egyes lépésben mutatók kerülnek bevonásra és szelektálásra. A folyamat addig ismétlődik, míg a kiválasztott mutatók mindegyike teljesíti a szignifikancia-kritériumot. Amely mutató kiválasztásra került, az a későbbi lépésekben még kiszelektálható, illetve amely mutató kiszelektálódott, az a későbbi lépésekben újra bevonható.

c) *A modellezésben megfigyelhető multidimenzióális outlierek szűrése és a paraméterbecslés újbóli elvégzése*

11 Ezt a korábbi lépéseknél elvégzett korrelációs számítás biztosítja, pl.: mutatók számításánál elkerülhetetlen, hogy a bizonyos mutató a többi mutató lineáris kombinációja.

12 Ez nem más, mint az adott kategóriában a „rossz ügyfelek” száma elosztva az összes ügyfél számával.

13 Az egyes mutatók default előrejelzőképességét több információs mérőszámmal is kifejezhetjük, leggyakrabban használt például a GINI és KS görbe, de lehet más információs mutatókat is számítani.

Léteznek olyan megfigyelések, melyek „kiugró” jellege csak szofisztikáltabb mérésekkel azonosítható. Amennyiben ezek bennmaradnak a paraméterbecslésben, akkor instabillá tehetik a modellt. Ezeket mindenképpen ki kell szűrni, és a paraméterbecslést újra el kell végezni. Számos módszer lehet segítségünkre. Az egyik leggyakrabban alkalmazott eljárás, hogy kiszámítjuk a paramétereket az adott megfigyeléssel és az adott megfigyelés elhagyásával. Amennyiben a két eredmény jelentősen eltér egymástól, akkor valószínűsíthető, hogy az adott megfigyelés „outlier”.¹⁴ Egy másik lehetséges megoldás a lineáris regresszió elméleténél használatos Cook-távolság ötletének felhasználása a logisztikus regressziós modellben. A hányados képlete:

$$C_j = \frac{\chi_j^2 h_{jj}}{(1 - h_{jj})^2}$$

$$\chi_j = \frac{\sqrt{w_j}(r_j - n_j \hat{p}_j)}{\sqrt{n_j \hat{p}_j \hat{q}_j}}$$

$$h_{jj} = w_j n_j \hat{p}_j \hat{q}_j (1, x'_j) \hat{V}_b \begin{pmatrix} 1 \\ x_j \end{pmatrix}$$

- w_j : a j -edik megfigyelés teljes súlya
- r_j : amennyiben a j -edik megfigyelés „defaultos”, akkor $r_j = 1$, egyébként $r_j = 0$
- n_j : a „kísérletek” száma, bináris eredményváltozó esetén $n_j = 1$
- \hat{p}_j : a j -dik megfigyelés becsült PD-je
- $\hat{q}_j = 1 - \hat{p}_j$

¹⁴ Általában a sokaság kevesebb mint 1%-ának kiszűréséről van szó.

A C_j -t általában egy számláló index függvényében szokták ábrázolni. A kiugró értékeket a görbéből már könnyen le lehet olvasni.

d) *A modell szeparációs erejének mérése*

A leggyakrabban alkalmazott módszer a GINI-görbe felrajzolása, az abból számítható GINI-indikátor számítása, illetve a Kolmogorov–Smirnov (KS) statisztika számítása.

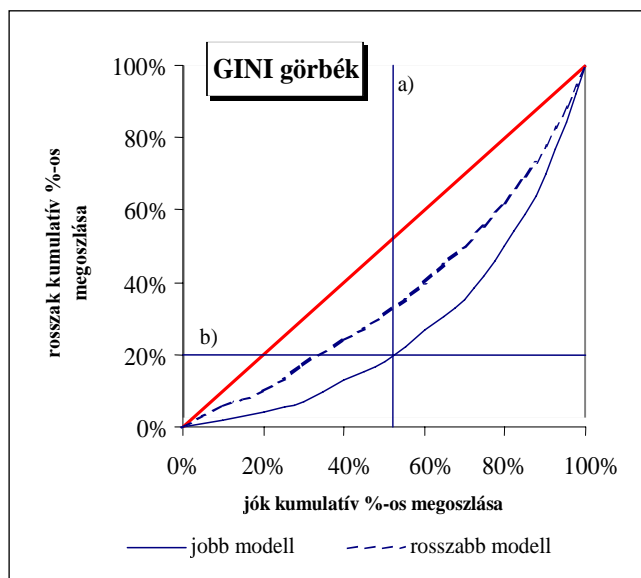
A GINI-görbe esetében a „rossz” ügyfelek kumulatív %-os megoszlását ábrázoljuk a „jó” ügyfelek kumulatív %-os megoszlásának függvényében. Minél távolabb van a görbe az átlótól, annál jobb a modell szeparációs ereje. A GINI-indikátor pedig azt mondja meg, hogy hogyan aránylik a GINI-görbe és az átló közötti terület a 0,5-hez. A GINI-mutató alapvetően csak relatív összehasonlításban értelmezhető, ám a nemzetközi gyakorlatban a legalább 45–50%-os GINI értéket tartják elfogadhatónak.

A 3. ábrán egy erősebb és egy gyengébb teljesítményű modellt láthatunk.

A görbének alapvetően kétfajta értelmezése lehet:

- Adott mennyiség mellett magasabb minőség: ahhoz, hogy adott mennyiségű jó ügyfelet lehessen beengedni a portfólióba, a folytonos vonallal jellemzett jobb modell esetében kevesebb rossz ügyfelet kell beengedni (x-tengelymetszet, lásd a) egyenes)
- Adott minőség mellett magasabb mennyiség: amennyiben ugyanannyi rossz ügyfelet engedünk be, akkor a folytonos vonallal jellemzett jobb modell esetében lehetőség van

3. ábra



nagyobb mennyiségű jó ügyfél egyidejű beengedésére (y-tengelymet-szet, lásd b) egyenes)

A KS-görbe esetében (lásd 4. ábra) mind a „jó”, mind a „rossz” ügyfelek kumulatív %-os megoszlását ábrázoljuk a PD (vagy PD-kategóriák) függvényében. A KS-mutató a két görbe közötti távolság maximuma.

$$KS = \max_s |P_G(s) - P_B(s)|$$

- s : a számított score, kategória vagy individuális PD
- $P_G(s)$: a „jó” ügyfelek kumulatív száma megoszlása (valószínűsége) az s függvényében
- $P_B(s)$: a „rossz” ügyfelek kumulatív száma megoszlása (valószínűsége) az s függvényében

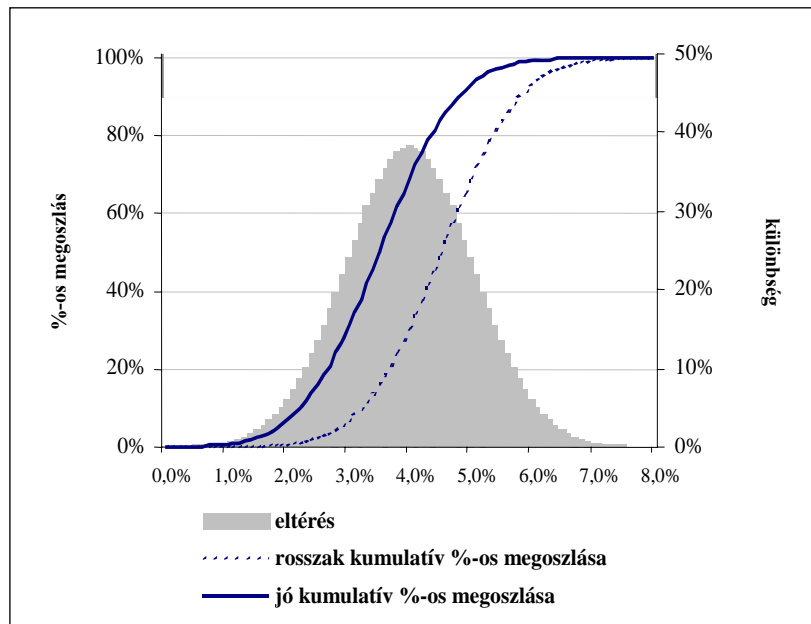
A KS-értéket – a GINI-hez hasonlóan – nem lehet alapesetben abszolút értelemben felhasználni, csupán relatív összehasonlításokban lehet értelmezni, ennek ellenére a nemzetközi gyakorlatban egy jó modelltől legalább 40%-os értéket várnak el.

Minél nagyobb a különbség a folytonos és a szaggatott vonal között, annál hatékonyabban tudja a modell elválasztani egymástól a „jó” ügyfeleket a „rosszak”-tól.

e) *A PD-k kategóriákká alakítása (klaszifikáció)*

A hitelezési folyamatban az egyedi PD-k segítségével homogén csoportokra bontjuk a portfóliót. A kategorizálásnál a következőket kell szem előtt tartani:

4. ábra



➤ Monotonitás:

Figyelni kell arra, hogy a tényleges default-ráta az egyes kategóriákban monoton növekvő legyen. Akkor tükrözi jól a kategorizálás a tényleges bedőlés valószínűségét, ha a magasabb kategóriákba arányaiban több defaultos kerül.

➤ Jó illeszkedés:

Annak is teljesülnie kell, hogy a függvény ne csak monoton növekvő legyen, hanem az egyes kategóriákban a tényleges rossz/összes arány összhangban legyen az adott kategóriába kerülő minősítések individuális PD-inek átlagával.

➤ Magyarázó erő:

Mivel folytonos függvényből készítettünk diszkrét értékeket, ezért a kate-

gorizálás után a modell magyarázó ereje nagy valószínűséggel csökkeni fog. Törekedni kell arra, hogy lehetőleg tartsuk meg a folytonos modellnél fennálló GINI és KS indikátorok értékét (vagy legalábbis ne rontsuk túlságosan).

➤ Kicsi koncentrátság:

Úgy célszerű megválasztani a PD-határokat, hogy a mintaeloszlás ne legyen túl koncentrált. Ne legyen az egyik kategóriában sokkal több megfigyelés, mint a többiben.

f) A modell szakértői felülvizsgálata

A kész modellel szemben fontos elvárás, hogy közgazdaságilag értelmes legyen. Ez azt jelenti, hogy a modellben használt változóknak önmagukban is közgazdasági tartalommal kell bírniuk,

és a szakértők véleménye szerint is magyarázó szerepet kell betölteniük az adós kockázatának előrejelzésében. Alapvető célunk, hogy a modell minél pontosabban jelezze előre a default valószínűségét, így a lehető legmagasabb GINI-t szeretnénk elérni, ugyanakkor kompromisszumot kell kötni a közgazdasági tartalom és a GINI növelése között. Az esetek nagy százalékában fel kell adnunk GINI-maximalizálási törekvésünket és a szakértők által elfogadható mutatók kombinációja mellett kell a lehető legmagasabb GINI-t elérni!

g) A szakértői konzultációból levonható következtetések ismeretében visszacsatolás a b. ponthoz

Általában 5-6 körös iteráció után már kialakul a szakértők által is elfogadható tartalmú és erejű modell. Amikor kialakul a végleges modell, akkor lépünk a 3-as pontra!

3. Out-of-sample (továbbiakban: OOS) adatbázison való tesztelés

A modell erejének visszamérését egy független, de a modellezés alapjául szolgáló mintához hasonló összetételű mintán kell elvégezni. Az adatbázis összeállításának folyamata teljes mértékben megegyezik az 1. pontban leírtakkal:

a) Vállalatok egyedi, pénzügyi adatainak, valamint default-események összegyűjtése

Az OOS adatbázis felállítása során problémát jelenthet a default definíció. A hitelintézet sajátos hitelezési folya-

matában alkalmazott default meghatározás eltérhet a piacon nyilvánosan hozzáférhető default információktól. Ennek eredményeként nem tudunk ugyanolyan definíciót alkalmazni az OOS adatbázison, mint a minta adatbázison. Ez azt jelenti, hogy kérdésessé válik az eredmény értelmezése, hiszen két eltérő dolgot szeretnénk összehasonlítani. Ez nem feltétlen jelent problémát, de fontos, hogy ha ilyet alkalmazunk, az eredmény értelmezésénél és a következtetések levonásánál figyelni kell.

b) Adatok tisztítása

c) A pénzügyi mutatók kiszámítása

d) Az „outlierek” kezelése

Valószínűsíthetően nagy adathalmazt kapunk. Éppen ezért mintavételi eljárásra van szükség. Mivel a reprezentativitáshoz rengeteg dimenzió alapján kellene a mintavételezést elvégezni, ezért az egyik leggyorsabb megoldás, hogy egy bináris PROBIT vagy LOGIT modellt szerkesztünk annak eldöntésére, hogy az out-of-sample megfigyelései milyen valószínűséggel elemei az eredeti adatbázisnak. A független változók körébe érdemes mind pénzügyi, mind nem pénzügyi változókat bevonni. A regressziós paraméterek megbecslése után az eredményváltozót csökkenő sorrendbe rendezve könnyedén ki lehet választani a teszteléshez szükséges elemszámú megfigyelést.

Általában az out-of-sample mintán futtatott GINI- és KS-értékek elmaradnak a fejlesztési mintán számítottaktól. Általánosan elfogadott hüvelykujj-szabály, hogy akkor megfelelő a modellezés, ha a GINI értéke az out-of-sample teszten 10–15%-nál többel nem csökken.

4. Validáció

Nemcsak a hitelintézet belső érdeke, hanem már szabályozói elvárás is a PD-modell tesztelése és folyamatos (éves) monitorálása. A hitelintézetnek ezt saját magának kell megtennie, a Felügyelet csupán a tesztelés eredményét vizsgálja, nem ő végzi a tesztet. A bank választhat, hogy belső erőforrását vagy külső független szakértőket vesz igénybe a teszt elvégzéséhez.

5. Pilot teszt

Még mielőtt a modellt éles működésbe állítanánk, ki kell alakítani az IT-fejlesztők segítségével egy tesztkörnyezetet, amelyen az elemzők, kockázatkezelők vizsgálhatják a PD-modell működését. Folyamatosan elemezni kell a visszajelzéseket, melyeket adott esetben be kell építeni a modellbe.

ÖSSZEGZÉS

A tanulmány célja az volt, hogy a magyarországi KKV szegmensben bemutassa az előrejelző-PD-modell egy lehetsé-

ges megvalósítását. A hitelintézetek portfóliójuk méretét, összetételét, adatbázisuk nagyságát, hitelezési folyamatuk sajátosságait figyelembe véve eldönthetik, hogy milyen módszerrel kívánja előrejelezni az egyes ügyfelek bedőlési valószínűségét. Rámutattunk arra, hogy bár a közép-kelet-európai térségben a kis- és középvállalati szegmens hiteligenyét, méretét és egyéb sajátosságait tekintve nagyon közel áll a lakossági portfólióhoz, mégis, az ügyfelek hatékony klasszifikálásához a lakossági hitelezésben alkalmazott módszerek felhasználásán túlmenően további kritikus szempontokat is figyelembe kell venni. Mindenképpen törekedni kell arra, hogy minden számszerűsíthető információt felhasználjunk az ügyfél elemzéséhez, de ezen túl külön hangsúlyt kapnak a szubjektív, egyedi, nehezen-kvantifikálható adatok. Fontos, hogy a számszerűsíthető információk mögötti közgazdasági értelem is teljesüljön a modellépítés folyamán. Ennek eléréséhez a tanulmányban említett hibrid modell a leginkább alkalmas, ugyanis a szakértői véleményeknek és a statisztikai-ökonometriai technikáknak szinergikus kapcsolatrendszerét kell kiépíteni.

IRODALOM

- Basel Committee on Banking Supervision [2000a]. Range of practice on Banks' internal ratings systems. *Working Paper, January* www.bis.org
- Basel Committee on Banking Supervision [2000b]. Credit ratings and complementary sources of credit quality information. *Working Paper, No.3, August* www.bis.org
- Basel Committee on Banking Supervision [2004]. Implementation of Basel II: Practical Considerations. *Bank for International Settlements, July* <www.bis.org>
- Basel Committee on Banking Supervision [2004]. International Convergence of Capital Measurement and Capital Standards. A revised framework, *Bank for International Settlements, June* www.bis.org
- CAIRE, D. [2004]: Building Credit Scorecards For Small Business Lending in Developing Markets. CFA, Banknook Consulting [2004]
- Crouhy, M.–Galai, D.–Mark, R. [2000]: A Comparative Analysis of Current Credit Risk Models. *Journal of Banking & Finance* 24 (2000) 59–117.

- FALKENSTEIN, E. –BORAL, A. [2000]: RiskCalc™ For Private Companies: Moody's Default Model. Moody's Investors Service, Global Credit Research [2000]
- HAND, D. J. [2001]: Modelling Consumer Credit Risk. IMA Journal of Management Mathematics (2001) 12., 139–155.
- KRALMEN–WEBER [2001]: Generally Accepted Rating Principles. (A primer. Journal of Banking and Finance 25. 1–2. pp. 3–35.)
- MERTON, R.C. [1974]: On the Pricing of Corporate Debt: the Risk Structure of Interest Rates. The Journal of Finance, vol. 29. (May,1974.) 449–470.
- RÖSCH, D. [2005]. An empirical comparison of default risk forecasts from alternative credit rating philosophies. International Journal of Forecasting 21 (2005) 37–51.
- THOMAS, L. C.–CROOK, J.–EDELMA, D.–THOMAS. L., [2002]: Credit Scoring And Its Application: Siam, 2002.
- VAN DEVENTER, D.–IMAI, K. [2003]: Credit Risk Models & the Basel Accords John Wiley & Sons (Asia) Pte Ltd.