NYILATKOZAT

Név: Akály Alexandra Melinda

ELTE Természettudományi Kar, szak: Matematika BSc

NEPTUN azonosító: CRHE1S

Szakdolgozat címe:
A hitelnyújtás és a hitelkártyák optimális kérdései

A **szakdolgozat** szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló szellemi alkotásom, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2023.06.04.

Akály Alexa

a hallgató aláírása

Eötvös Loránd Tudományegyetem

Természettudományi Kar

Matematikai Intézet

A hitelnyújtás és a hitelkártyák optimalitási kérdései

— szakdolgozat —

Szerző:

Akály Alexandra Melinda

Alkalmazott matematikus szakirány

Témavezető:

Dr. Csiszár Villő

Budapest, 2023.

Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani a témavezetőmnek, Dr. Csiszár Villőnek, aki a rendszeres konzultációk során szakértelmével, észrevételeivel és világos magyarázatival hozzájárult a szakdolgozatom elkészítéséhez. Külön szeretném megköszönni neki a végtelen türelmét és megértését, amivel felém fordult.

Köszönöm továbbá a családomnak, hogy türelmesen megvárták, amíg visszatérek közéjük a mindennapokba. Köszönöm a barátaimnak a lelki támogatást, ami továbblendített a nehezebb pillanatokon.

Tartalomjegyzék

1. fejezet

Bevezetés

A hitel története egészen Kr.e. 3000-ig az ókori Mezopotámiáig vezethető vissza. [1] Természetesen eredetileg nem pénzt kölcsönöztek és fizettek vissza, hanem vetőmagot kölcsönöztek és a takarmányból fizették ki az adósságot. Törvényileg szabályozták ebben az időben, hogy a kamat értéke legfeljebb 33% lehet. Ha valaki nem tudta visszafizetni a felvett kölcsönt, a vissza nem fizetés okától függően vagy elengedték azt, vagy az adóst esetleg valamelyik családtagját 3 évre rabszolgásították. A középkorban vallási okok miatt a kamatos hitel nem volt elfogadott. Az 1700-as évek végén és az 1800-as évek elején jelentek meg a bankok Európában és Amerikában. Az első hitelek is megjelentek, de csak az 1870-es évektől lett gyakori a hitelfelfelvétel a munkás emberek körében.

A mai magyar társadalom meglehetősen negatívan áll a hitelekhez vagy legalábbis a hitelintézményekhez, ez részben biztosan köszönhető a sorra bedőlt devizahiteleknek. Azonban a hitel nem egy rossz lehetőség azoknak, akiknek pénzre van szükségük és máshogy nem tudják a jelenben előteremteni. Viszont a szemünk előtt kell tartani, hogy a bankok is vállalatok, felelősségük van az alkalmazottak felé és persze azok felé is, akik nem hitelt vesznek fel, hanem a pénzüket adják a banknak. Ezáltal a profit termelés a céljuk, pont ugyanúgy, mint bármely más vállalatnak. Azt is érdemes megemlíteni, ha valakinek nem nyújtanak hitelt, lehet épp jót tesznek a bankok. Mivel ha az illető mégse tudná fizetni a felvett kölcsönt a lakása/háza/autója lenne elárverezve, ami ugyan előnyösebb, mintha egy közeli családtagját rabszolgásítanák, azonban még mindig célszerű kerülni.

A szakdolgozatom két nagyobb részre osztható. Az első részben egyciklusos majd többciklusos modellel fogunk foglalkozni. Az egyciklusos modell, azt foglalja magában, hogy a bank egyszer meghozza a döntését, hogy nyújt hitelt vagy sem a leendő ügyfelének. Ehhez egy egyszerű Bernoulli próbát fogunk használni. A többciklusos modell alatt arra gondolok, amikor a bank rendszeresen, újra és újra hitelt nyújt az adott ügyfélnek. Legegyszerűbb példa a hitelkártya, ami hónapról, hónapra megújul. Magyarországon ugyan nem olyan elterjedt, de külföldön, főleg Amerikában szinte már nem is lehet nélküle élni. Amerika egyes államaiban, ha nem is mindegyikben, egy egyszerű házfelújítás során a kivitelezőnek is csak részletre lehet fizetni, amit a hitelkártya garantál a kivitelező szemében. A többciklusos modellnek két megoldási módszerét mutatja be a szakdolgozatom. Az első módszerben a Bayes becslés lesz segítségünkre. A másik módszerben dinamikus programozást fogunk használni. Ezt a módszert implementáltam, majd elemeztem a szakdolgozatomban.

A második részben azt fogom vizsgálni, hogy az emberi tulajdonságokból, hogy lehet modellt adni arra, hogy megéri-e hitelt nyújtani. Ez módszer $4+1$ lépésből áll. A kockázati index létérehozásából, ami az első négy lépés "$+1$" a használatából. A végén készítettem egy saját példát is a hazai lakáspiaci adatok alapján.

2. fejezet

Hitelnyújtási modellek

2.1. Egyciklusos modell

A legegyszerűbb modellel kezdjük az áttekintést. A következőkben azt fogjuk vizsgálni, amikor csak egyszer kell eldöntenünk, hogy egy ember hitelképes vagy sem. Például, ha valaki egyszer vesz egy házat, ahhoz szüksége van hitelre, utána viszont többé nem szorul banki segítségre. Tegyük fel, hogy egy termék ára R. A termék itt a hitel. C pedig a bank járulékos költsége az adott hitelnél. Ezt a módszercsoportot a kérdéskör feldolgozására az alábbi cikkből merítettem:[2]. Ennél a modellnél azzal a feltételezéssel élünk, hogy a hitelt az ügyfél teljesen visszafizette vagy egyáltalán nem fizetett vissza semmit. Erre tökéletesen tudjuk alkalmazni a Bernoulli próbát, hiszen, ez egy véletlenszerű kísérlet, aminek két kimenetele lehet. A két lehetőség a "siker" esetünkben az adhatunk hitelt, valószínűleg vissza fogják fizetni, illetve a "sikertelenség" azaz nem hitelképes az ügyfelünk a próba szerint. Legyen p annak a valószínűsége, hogy az ügyfél rendesen törleszt. A várható értékét nézzük annak, hogy mennyi pénzt szedtünk be. Ebből már ki tudjuk számolni, hogy mekkora a várható értéke a hozamnak (pR). Az a kérdés, hogy a várható hozam meghaladja-e a járulékos költséget (C).

$$pR \geq C \quad \text{vagy} \quad p \geq C/R$$

Például, ha egy termék 100 000 HUF forint, a járulékos költségünk 60 000 HUF forint. E szerint a modell szerint, akkor ajánljunk fel hitelt, amíg a visszafizetés valószínűsége nagyobb, mint 0.6

$$p \geq C/R = 60000/100000 = 0,6$$

2.2. Többciklusos modell

Eddig azt feltételeztük, hogy hitelt csak egyszer vesz fel az ember. Mostantól azokra az esetekre fókuszálunk, amikor egy ember többször vesz fel hitelt. Az egyszerűség kedvéért ezt hitelkártyával tudjuk a legjobban modellezni. Ugyanis minden hónapban a bank újra és újra felülvizsgálhatja, hogy az adott ügyfélnek folyósít-e hitelt vagy sem.

Tapasztalatok alapján, ha az ügyfél visszafizette a kapott hitelt, akkor a következő hónapban érdemes megnövelni annak a valószínűségét, hogy a bank visszakapja a pénzét (p). Ellenkező esetben, ha nem fizet, akkor pedig csökkenteni kell. Megeshet, hogy 0-ra csökkentjük, akkor egyértelműen nem fizetőképes. Például, ha $0,6$-os becsült beszedési valószínűséggel kezdünk és az ügyfél törleszt, akkor ezt az értéket érdemes növelnünk.

Bayes-becslést használni fogjuk a módszerben, ezért szeretném ezt levezetni előtte. A "recept" a következő:

Definíció. Legyen α és β két mérték az (X, M) mérhető téren. Az α abszolút folytonos a β-ra nézve, ha bármely $H \in M$ esetén, ha $\beta(H) = 0$, akkor $\alpha(H) = 0$.

- $\nu \in \Theta$, azaz a paraméter is egy valószínűségi változó. Véletlenül választjuk. $\Theta \subset \mathbb{R}^p$ nyílt
- $\nu \sim q(t)$ (sűrűségfüggvénye), $t \in \Theta$ azaz $P(\nu \in A) = \int\limits_A q(t)dt$, azaz ilyen eloszlás szerint nézzük.
- Majd a véletlen ν szerinti P_ν-ból veszünk mintát. $\mathbb{P}$ mérték $\mathcal{B}(\Theta \times \mathbb{R}^k)$-on, ahol k a minta elemszáma.
- Tegyük fel, hogy: P_ν abszolút folytonos R-re. $f_\nu(\underline{x}) = \dfrac{dP_\nu}{dR}(\underline{x}), x \in \mathbb{R}^k$.

- Azaz eddig van egy együttes eloszlás Θ-n és $\underline{X}$-en úgy, hogy a Θ-nak van a q eloszlása. Az $\underline{X}$-nek pedig ν paramétertől függő sűrűségfüggvénye $f_\nu(\underline{x})$.

Definíció (ν Bayes becslése)**.** $T := E_{\mathbb{P}}[\nu|\underline{X}] = E_{\mathbb{P}}[\nu|\sigma(\underline{X})]$ [3]

Ekkor:

$$\mathbb{P}((\nu, \underline{X}) \in A) = \int\limits_A \int q(t) f_t(\underline{x}) \quad dx \quad dt$$

- $q(t)$ az a priori eloszlásunk

Feltételes sűrűségfüggvényünk pedig:

$$f_{\nu|\underline{X}}(t|\underline{x}) = \frac{f_t(\underline{x}) \cdot q(t)}{f_{\underline{X}}(\underline{x})} = \frac{f_t(\underline{x}) \cdot q(t)}{\int\limits_\Theta f_t(\underline{x}) \cdot q(t) dt}$$

$$T = E_{\mathbb{P}}(\nu|\underline{x}) = \int\limits_\Theta t f_{\nu|\underline{X}}(t|\underline{x}) dt$$

$\underline{X}$ mintából $f_{\nu|\underline{X}}$ kiszámolható, ami az a posteriori sűrűségfüggvény ν eloszlás szerint. Onnan T integrálással adódik, hiszen ha sűrűségfüggvényt megszorozzuk azzal, ami szerint integrálunk, majd leintegráljuk akkor pont T-t kapjuk. Azaz megkapjuk a paraméter tetszőleges $g(\nu)$ függvényének a Bayes-becslését:

$$E_{\mathbb{P}}[g(\nu)|\underline{X}]$$

A Bayes-féle módszerrel egyesíteni tudjuk a megfigyelésből kapott adatokból és az a priori feltevésekből kapott információkat, ezáltal tudunk generálni egy a posteriori eredményt. A módszer hatékonyságát növelni tudjuk, hiszen az a priori becslésekről az a posteriori becslésekre történő áttérés tetszőlegesen sokszor ismételhető. Ezért ezt a módszert kitűnően tudjuk használni a begyűjtési valószínűség felülvizsgálatára. Minden hónap végeredménye egy Bernoulli-próba. A több eredmény együtt pedig egy binomiális eloszlás egy ismeretlen binomiális paraméterrel (t).

Erre úgy is tekinthetünk, hogy t – a modellünkben a visszafizetés valószínűsége – egy valószínűségi változó, amely Béta eloszlásnak megfelelően oszlik el. A paraméterei pedig (r, n), amik úgy vannak megválasztva, hogy tükrözze a korábbi jóslatunkat az adott ügyfélről. Amikor az (r, n) paramétereket választjuk ki fontos, hogy az apriori eloszlásnak megfelelően legyenek kiválasztva. Ha t várható értéke r/n és ha a két érték $r = 1, n = 2$ akkor az apriori eloszlás egyenletes, ahol $0 \leq t \leq 1$. Például ha $r = 19$ és $n = 20$, akkor a visszafizetés várható valószínűsége $19/20 = 0,95$ lesz, és az eloszlás unimodális lesz, ami azt jelenti, hogy egy lokális maximum helye van.

A visszafizetési valószínűség felülvizsgálata:
Az r és n paraméterekkel rendelkező Béta eloszlás sűrűségfüggvénye a következőképpen adható meg:

$$\begin{aligned} q(t|r,n) &= \frac{(n-1)!}{(r-1)!(n-r-1)!} t^{r-1}(1-t)^{n-r-1} \qquad 0 \leq t \leq 1, \quad n > r > 0 \\ &= \frac{\Gamma(n)}{\Gamma(r)\Gamma(n-r)} t^{r-1}(1-t)^{n-r-1} \end{aligned} \tag{2.1}$$

Ahol $\Gamma(n) = (n-1)!$
Ha pedig a megszokott Béta eloszlás szerint szeretnénk felírni, akkor:

$$f(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

Ahol: $x = t \quad \alpha = r \quad \beta = n - r \quad B = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

Mostantól én az első paraméterezéssel fogok dolgozni. Tegyük fel, hogy ezt az eloszlást (r, n paraméterekkel) a t apriori eloszlásának választottuk, ami az ügyfelünk ismeretlen visszafizetési valószínűsége. Ekkor a visszafizetés valószínűségének felülvizsgálata és újraszámolása az alábbiak szerint zajlik: Ha a hitelt n'-szer meghosszabbítja a bank (minden alkalommal konstans hitelösszeggel), és r'-nyi visszafizetés történik.

Bayes becslés itt:

- apriori eloszlás: $t \sim \beta(r, n)$ azaz a $q(t)$ a 2.1 egyenlet
- $f_t(r') = \binom{n'}{r'} \cdot t^{r'} \cdot (1-t)^{(n'-r')}$
- várható értéke: $E(t) = \dfrac{r}{n}$

Az a posteriori eloszlás számolása:

$$q^*(t|r') = \frac{f_t(r') \cdot q(t)}{f_{\underline{X}}(r')} = \frac{f_t(r') \cdot q(t)}{\int\limits_{\Theta} f_t(r') \cdot q(t)dt} = k(r') \cdot q(t) \cdot f_t(r')$$

Behelyettesítve:

$$q^*(t|r') = k_0(r') \cdot \left(\frac{(n-1)!}{(r-1)!(n-r-1)!} t^{r-1}(1-t)^{n-r-1} \right) \cdot \left(\binom{n'}{r'} \cdot t^{r'} \cdot (1-t)^{(n'-r')} \right)$$

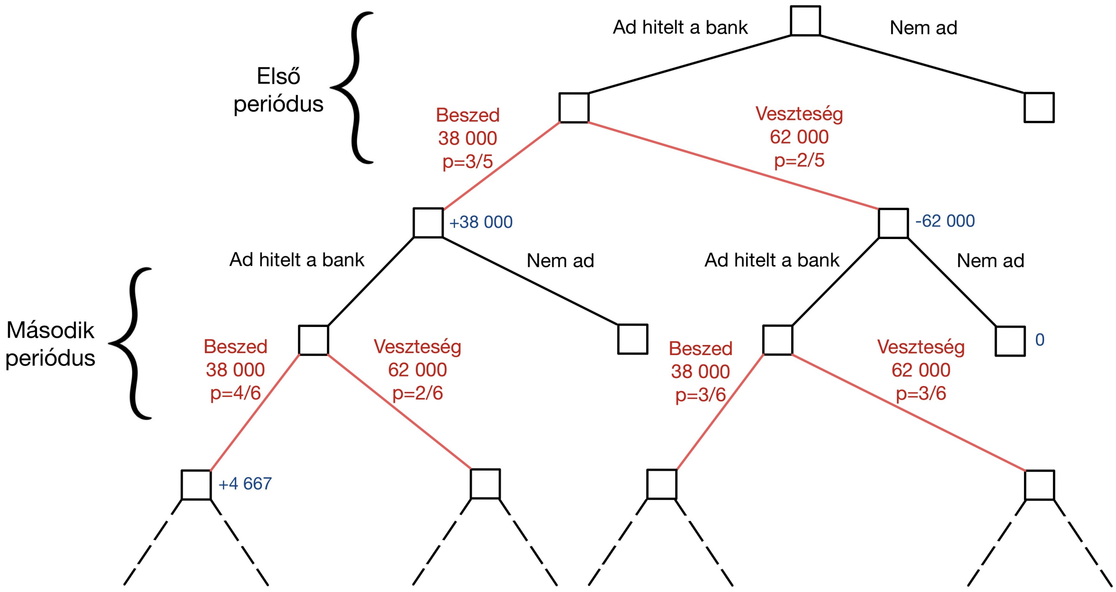
Ahol a $k_0(r')$ a nevező reciprokát jelenti. A $k_0(r')$ és $k(r')$ egy t-től független konstans, ami csak r'-től függ, ezért ezzel most nem foglalkozunk a becslés szempontjából.

$$\begin{aligned} q^*(t|r') &= k(r') \cdot t^{r-1} \cdot (1-t)^{n-r-1} \cdot t^{r'} \cdot (1-t)^{n'-r'} \\ &= k(r') \cdot t^{r+r'-1} \cdot (1-t)^{n+n'-(r+r')-1} \end{aligned}$$

Ebből következik, hogy a t a posteriori eloszlása: $\beta(r+r', n+n')$ azaz a r'' és n'' a következő :

$$r'' = r + r' \quad n'' = n + n'$$

Ekkor r'' és n'' a felülvizsgált Béta eloszlás paraméterei (a posteri eloszlása) és a t várható értéke most r''/n''. [4]



2.1. ábra. Döntési fa

A begyűjtési valószínűség felülvizsgálata:

Az előző példát nézzük meg ezen a modellen is. Most a Béta eloszlású p-nek ugye vannak paraméterei amik legyenek $r = 3, n = 5$. A visszafizetés várható valószínűsége $r/n = 3/5$. A hitel összege továbbra is legyen $R = 100\,000\,\mathrm{HUF}$ forint a járulékos költsége pedig $C = 62\,000\,\mathrm{HUF}$ forint (vagyis, ha nem fizet az ügyfél ennyi kára lesz a banknak.) Ha csak 1 ciklusra néznénk, akkor a korábbi képletet használva: $3/5 \cdot 100\,000 = 60\,000$. Vagyis ebben az esetben a banknak nem érné meg hitelt nyújtania, hiszen $2\,000\,\mathrm{HUF}$ kára keletkezne.

Most nézzük meg, hogy ha az ügyfél ugyanilyen paraméterekkel, de már két hónapra szeretne hitelt felvenni, az vajon megéri-e a banknak, ha szintén az ebben a részben részletezett modellt használjuk. Két esetet kell vizsgálnunk 1) ha nem fizet az ügyfél 2) ha fizet. Az első esetben 3/6-ra csökken annak a valószínűsége, hogy az ügyfél fizet a második hónapban, hiszen n nőtt r pedig nem mert nem történt visszafizetés. Ekkor a második hónapban várhatóan veszteséges lesz hitelt ajánlani, mert $3/6 \cdot 38\,000 + 3/6 \cdot (-62\,000) = -12\,000\,\mathrm{HUF}$. Azonban ha a második esetet vizsgáljuk, akkor 4/6-ra nő annak a valószínűsége, hogy visszakapja a bank a pénzét, mert r és n is nőtt. Ebben az esetben pedig várhatóan nyereséges lesz hitelt ajánlani, hiszen $4/6 \cdot 38\,000 + 2/6 \cdot (-62\,000) = 4\,667\,\mathrm{HUF}$. Így a hitelnyújtás maximális várható hozama két ciklus alatt:

$$3/5[38\,000 + 4\,667] + 2/5[-62\,000 + 0] = 800$$

Láthatjuk, hogy a banknak várhatóan nem 2 000 HUF kára hanem 800 HUF bevétele lesz. Ezért jobban megéri ugyanazt az ügyfelet alaposabban megvizsgálni.

Ha azt feltételezzük, hogy ha valaki nem fizeti vissza a tartozást utána soha nem kap hitelt, akkor kapunk egy összeget, ami a maximális várható érték alsó határa lesz. Itt egy végtelen sorozat összeget szeretnénk kapni, ami csak az r, n paraméterektől függ. Ennek a módszernek a hátránya, hogy egyáltalán nem könnyű minden lehetséges visszafizetés illetve nem visszafizetéshez kialakítani. Például ha csak 10 hónapig van valakinek hitelkártyája már az is 2^{10} opció lehet. Emiatt a következő részben ugyanezt a problémát dinamikus programozással fogom vizsgálni, hogy elkerüljük az exponenciálisan sok lehetőséget.

Most feltesszük, hogy kevésbé vagyunk biztosak abban, hogy kezdetben milyen várható értéke van annak, hogy az ügyfél visszafizeti a felvett kölcsönt. Ahelyett, hogy egy korábbi eloszlás paraméreivel kezdnénk, inkább kezdjük az ($r = 1,2 \;\; n = 2$) paraméterekkel. Mivel az r, n kisebb így nagyobb súllyal számít az, hogy a kölcsön visszafizetésre került-e. Azonban a visszafizetés kezdeti várhatóértéke továbbra is $0,6$, (mert $r/n = 1,2/2 = 0,6$). Most számoljuk ki ezekkel az adatokkal és a korábbi modellünkkel a várható hozamot.

$$(1,2/2)[38\,000 + (2,2/3)4\,667 + (0,8/3)(-62\,000)] + (0,8/2)[-62\,000 + 0] = 4800$$

Ebben az esetben jelentősen nagyobb nyereségre tett szert a cég, pedig a várható visszafizetési valószínűség az előző példával egyenlő a kezdetekben. A bankok gyakran küzdenek azzal a problémával, hogy kevés olyan információval rendelkeznek ami alapján "nyugodt szívvel" adhatnak hitelt. Ez a példa viszont jól mutatja, hogy megéri hitelt ajánlani, akár rizikósabbnak tűnő ügyfelek esetén is, mivel ha fizet akkor akár hosszú-hosszú évekre leszerződhetünk egy várhatóan megbízható ügyféllel.

Az eddig r, n paraméterek kiválasztása nehéz feladat lehet, és amint láttuk, a hitelnyújtás döntésében nagy szerepet játszanak. Ezen paraméterek érzékenysége az eljárás egyik fő gyengesége. A döntéshozónak át kell tekintenie a rendelkezésére álló információkat, és nem csak a begyűjtés várható értékét, hanem a relatív a prior tudást, ami ezt a valószínűséget adta. Feltételezzük, hogy egy adott ügyfél visszafizetésének várható értéke, előzetesen logikusan reprezentálhatóak egy Béta eloszlás r, n paramétereinek kiválasztásával.

Eddig az inflációt figyelmen kívül hagytuk, ami akkor sem szerencsés, ha az igen csekély, napjainkban viszont annál fontosabb, hogy ne maradjon ki a modellből. Az ügyfél által várható kifizetett összeg értékének a kiszámítása előtt minden jövőbeli i ciklusban meg kell szorozni a megfelelő diszkont tényezővel $1/(1+\rho)^i$, ahol ρ az adott periódus kamatlába. A diszkontálás hozzáadásával a következő időszakok veszítenek a fontosságukból, de természetesen továbbra sem elhanyagolhatóak.

2.3. Többciklusos modell dinamikus programozással

Mint már korábban említettem ennek a modellnek az előnye az előzővel szemben, hogy nem lesz exponenciálisan sok számolás, hanem a dinamikus programozás módszerével fogunk számolni. Így akár több száz illetve ezer hónapot adhatunk a számítógépnek, amivel kaphatunk egy-egy határértéket, ami alapján tudunk dönteni egy ügyfélről. További nagy előnye ennek a modellnek, hogy később bővíthető, így a hitelnyújtás további problémái is beleépíthetőek.

Tegyük fel, hogy továbbra is egy Béta eloszlású sűrűségfüggvény (r, n) paraméterekkel reprezentálja az a priori sűrűségfüggvényt t-n, vagyis a visszafizetés valószínűségét. Tekintettel arra, hogy ha az ügyfél fizet akkor a nyereség K_1, míg ha az adott hónapban nem fizet, akkor keletkezik a veszteség K_2, ahol $K_2 < 0$. Illetve még mindig ugyanazt a kérdést szeretnénk megválaszolni, csak egy kicsit árnyaltabban. Ha adott a bank hitelt, de az ügyfél nem fizetett, akkor mi alapján döntsön a bank, hogy a következő hónapban is nyújt-e hitelt az ügyfélnek vagy sem. Ez a kérdés pedig hónapról hónapra ismétlődik. A döntési probléma megoldása során továbbra is figyelembe kell vennünk a jelenlegi cselekmények jövőbeli hatását.

Legyen:

- (r, n) = állapotváltozó (kétdimenziós vektor, amely a jelenlegi Béta a prior eloszlás paramétereit reprezentálja az i-edik periódusban/hónapban)
- $f_i(r, n)$ = maximális diszkontált várható kifizetés az i-edik állapotból a következőbe. (r, n) adott.
- $a = 1/(1 + \rho)$ diszkont faktor
- $D_i(r, n)$ = az i-edik ciklusban végrehajtandó optimális kimenet, ahol
  - $D_i(r, n) = 1$ Jóváírjuk a hitelt
  - $D_i(r, n) = 0$ Nem adunk hitelt

Ekkor a következő általános ismétlődő képlet érvényesül:

$$f_i(r,n) = max\{ \quad (r/n)[K_1 + af_{i+1}(r+1, n+1)] + [1 - (r/n)][K_2 + af_{i+1}(r, n+1)] \quad ; \quad 0 \quad \} \tag{2.2}$$

Az első tag első felében a beszedés várható valószínűségét (r/n) megszorozzuk az ügyfél által visszafizetett összeggel, ami törlesztés esetén ennyi lesz:

$$[K_1 + af_{i+1}(r + 1, n + 1)]$$

A K_1 az egy periódus múlva fennálló helyzetből származó bevétel összege egy periódussal visszafelé diszkontálva, illetve az azonnali nyeresége a banknak. Az első tag második felében a beszedés várható hiányának valószínűségét $1 - (r/n)$ szorozzuk meg az ügyfél által nem visszafizetett összeggel.

$$[K_2 + af_{i+1}(r, n + 1)]$$

A K_2 összeg elvesztése a banknak csak később fog bekövetkezni ebben az esetben. Ezt reprezentálja a $f_{i+1}(r, n+1)$, amit ebben az esetben egy hónappal visszafelé kell diszkontálni. Ha a banknak nem származik nyeresége, akkor:

$$f_{i+1}(r, n + 1) = 0$$

lesz.

Egy konkrét példán bemutatom a modell működését:

A fenti 2.2 egyenlet megoldásához először feltételezzük, hogy egy tetszőleges távoli periódusban a folyamat leáll. Az algoritmus értékét ebben a sokadik ciklusban 0-ra állítunk. Ez a megközelítés, azért működik, mert a diszkontráta a jövőbeli értékek fontosságát folyamatosan csökkenti az idő előrehaladtával. A tetszőleges távoli ciklus pedig legyen $I = 40$.

$$f_I(r,n) = 0 \quad \forall \quad r\text{-re és } n\text{-re} \tag{2.3}$$

Most a a 2.2 egyenlet rekurzívan megoldható, hátulról kezdve $i = 39$. Nézzük az első pár iterációt.

$$f_{39}(r,n) = max\{\ (r/n)[K_1 + af_{40}(r+1,n+1)] + [1-(r/n)][K_2 + af_{40}(r,n+1)]\ ; \quad 0\ \}$$
$$\text{Mivel } f_{40}(r,n) = 0 \text{ ezért,}$$
$$f_{39}(r,n) = max\{\ (r/n)[K_1] + [1-(r/n)[K_2]\ ; \quad 0\ \} \tag{2.4}$$

Legyen B_{39} egyenlő azzal, amikor nullszaldós a banknak a hitelnyújtás. Az r/n hányadosunk pedig legyen annyi, hogy a bank legyen közömbös azzal kapcsolatban, hogy nyújt hitelt vagy elutasítja a hitelnyújtási kérelmet. A megoldás, amelynél mindkét lehetőség egyenlő:

$$\begin{gathered} B_{39}K_1 + (1-B_{39})K_2 = 0 \\ \text{Átrendezve pedig:} \\ B_{39} = \frac{-K_2}{(K_1 - K_2)} \end{gathered} \tag{2.5}$$

Ekkor már a 2.4 egyenletet egyszerűbben is fel tudjuk írni:

$$f_{39}(r,n) = \begin{cases} 0 & \text{ha}(r/n) \leq B_{39} \\ (r/n)[K_1] + [1-(r/n)[K_2] & \text{ha}(r/n) > B_{39} \end{cases} \tag{2.6}$$

Azaz az optimális döntés a 39. hónapban (amit jelöljön D_{39}) a következőképp írhatjuk fel: Adjon hitelt a bank ($D_{39} = 1$), ha (r/n) meghaladja B_{39}-et, amennyiben viszont kisebb vagy egyenlő ne nyújtson hitelt a bank ($D_{39} = 0$).

A következő iteráció felírása pedig, így néz ki:

$$
\begin{aligned}
&\qquad\qquad\qquad\quad \text{Visszafizette az ügyfél}\\
f_{38}(r,n) = max\{\ &(r/n)[K_1 + af_{39}(r+1,n+1)]\\
&\qquad\qquad\qquad\qquad\qquad \text{nem fizetett} \qquad\qquad\quad \text{ne nyújtsunk hitelt}\\
&+[1-(r/n)][K_2 + af_{39}(r,n+1)]\ ; \quad 0\ \}
\end{aligned}
\tag{2.7}
$$

ahol az f_{39} értékeit a 2.6 egyenletből kapjuk.

Az r és n lehetséges értékeinek csinálunk egy táblázatot, mondjuk 0-tól 100-ig. A táblázatba pedig f_{39} értékeit foglaljuk. Majd megoldjuk a 2.7-es egyenletet az f_{38}-ra.

Miután megoldottuk az $f_{38}(r,n)$ függvényt is táblázatban rögzítjük a kapott eredményeket. Végül ezt addig folytatjuk visszafelé, amíg eljutunk $f_1(r,n)$-hez és $D_1(r,n)$-hez. A végső táblázatba pedig már csak az optimális döntés fog szerepelni, azaz ha $D_1(r,n) = 1$ akkor egy 1-es, azaz megéri hitelt nyújtani, ha pedig $D_1(r,n) = 0$ akkor egy 0-ás, akkor optimálisabb elvileg elutasítani a hitel kérelmet a banknak. Szerintem egy érdekes és vicces megjegyzés, hogy az 1970-es években erre a módszerre úgy hivatkoznak, hogy jelentősen igénybe veszi a számítógépek kapacitását, ezért kifejezetten sok idő, mire kiszámolja a gép, de lehetséges. (Az én számítógépemnek ezekkel a kondíciókkal a számolás átlagosan 1 milliszekundumnyi időt vett igénybe úgy, hogy duplán futtattam.)

Hogy a korábbi példáinkhoz hűek maradjunk az alábbi értékekkel futtatott eredményt fogom bemutatni:

- $R = 40$ felsőkorlát r-re
- $N = 40$ felsőkorlát n-re
- $a = 1/(1+\rho) = 1/(1,05)$ diszkont faktor
- $I = 40$ ennyi ciklust fut a programunk
- $K_1 = 38\,000\,\text{HUF}$ Nyereség
- $K_2 = -62\,000\,\text{HUF}$ Veszteség

Ahol ρ a várható infláció. Ez az infláció az Európai Unióban reális, illetve egy kicsit túlzó, azzal együtt is, hogy az elmúlt időben előfordult, már a duplája is. Statisztikát erről [5] [6] itt található.

Ennek a modellnek két nagy előnye az előzővel szemben, az egyik az, hogy nem exponenciálisan sok időbe telik, mire eredményre jut. Előre kiszámolja melyik ügyfél mekkora kockázatot jelent, a rendelkezésre álló információk alapján.

X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
0	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
0	1	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
0	0	1	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
0	0	1	1	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
0	0	0	1	1	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X
0	0	0	0	1	1	1	X	X	X	X	X	X	X	X	X	X	X	X	X
0	0	0	0	1	1	1	1	X	X	X	X	X	X	X	X	X	X	X	X
0	0	0	0	0	1	1	1	1	X	X	X	X	X	X	X	X	X	X	X
0	0	0	0	0	1	1	1	1	1	X	X	X	X	X	X	X	X	X	X
0	0	0	0	0	0	1	1	1	1	1	X	X	X	X	X	X	X	X	X
0	0	0	0	0	0	0	1	1	1	1	1	X	X	X	X	X	X	X	X
0	0	0	0	0	0	0	1	1	1	1	1	1	X	X	X	X	X	X	X
0	0	0	0	0	0	0	0	1	1	1	1	1	1	X	X	X	X	X	X
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	X	X	X	X	X
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	X	X	X	X
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	X	X	X
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	X	X
0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	X

2.1. táblázat

Ugyan a programunk $R = 40$ és $N = 40$-ig futott, azonban én ebből levágtam egy 20×20 résztáblázatot, mert ennél nagyobb táblázattal nem tud dolgozni a latex. A 2.1 táblázatban az oszlopok rendre $r = 1, 2, \dots, 20$. A sorok pedig rendre $n = 1, 2, \dots, 20$. Annak nincs értelme, hogy r egyenlő vagy nagyobb, mint n, ezért a táblázatban az ennek megfelelő cellákba X-et írtunk. Abban az esetben ha $r = n$ azt jelentené szemléletesen, hogy a visszafizetés minden ciklusban biztos megtörténne a jövőben. A táblázatban, ahol 1-et írtunk megéri hitelt nyújtani, ahova 0-át ott nem éri meg. A nyereség összegét a következő alfejezetben fogom vizsgálni.

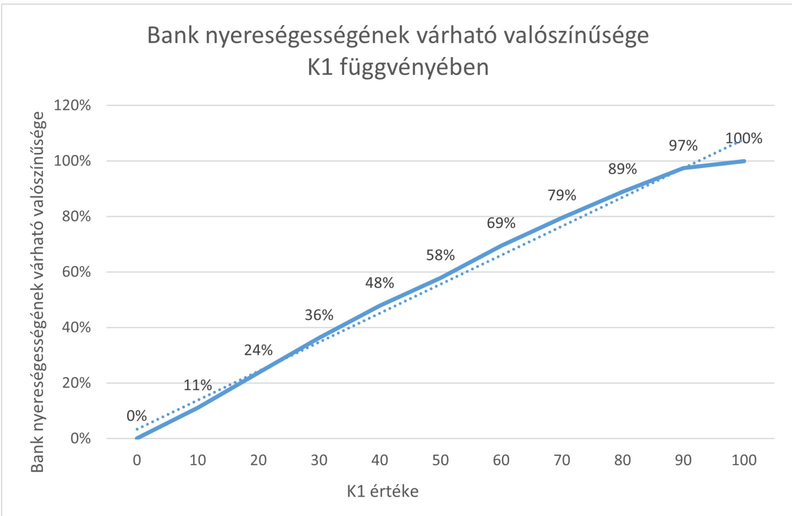
2.3.1. Saját elemzés

Ebben az alfejezetben az imént tárgyalt modellt, illetve az azokból kapott eredményeket fogom bemutatni. Először arról fogok értekezni, hogy a különböző változók, hogy befolyásolták a kapott eredményt.

Amennyiben az **infláció** mértékét 5%-ról 10%-ra növeltük. Azaz a diszkont faktort $a_1 = \frac{1}{1{,}05}$-ről $a_2 = \frac{1}{1{,}1}$-re csökkentettük akkor a várható nyereség 25%-kal csökkent, azaz a kapott eredményt jelentősen befolyásolta. A következő paraméterekkel futatott esetben:

- $R, N, I = 40$
- K_1, K_2 értékétől nem függ, azonban a_1 és a_2 estén is ugyanannyi kell legyen mindkét érték.

A hitel összege a továbbiakban legyen $R = 100\,000\,\text{HUF}$, a diszkont tényező pedig $a = \frac{1}{1{,}05}$. Azonban most a nyereség-veszteség arányának és a várható beszedés valószínűségnek a kapcsolatát vizsgáltam $R, N, I = 40$ esetén. Ebben a bekezdésben az R két külön dolgot jelöl, a hitel összegét illetve a r felső korlátját. A 2.2 ábrán azt szerettem volna megmutatni, hogy minél nagyobb $K1$ és ezzel párhuzamosan, minél kisebb abszolút értékben a K_2, annál nagyobb valószínűséggel lesz nyereséges a hitelnyújtás.



2.2. ábra. Nyereség értékének hatása a várható beszedésre

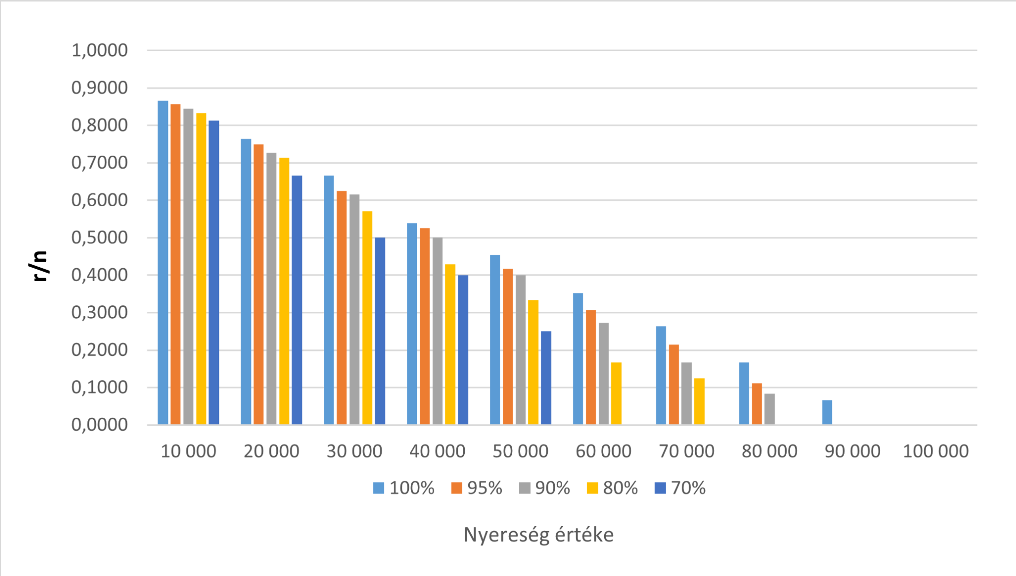
Arra is kíváncsi voltam, hogy ez a módszer, "húz-e" a felé, hogy nyújtsunk vagy sem hitelt, ezért az alábbi paraméterekkel futtattam le a programot.

- $R, N, I = 40$
- $a = 1/(1,05)$
- $K_1 = K_2 = 50\,000\,\mathrm{HUF}$

Itt azt tartottam fontosnak, hogy a nyereség és a veszteség értéke megegyezzen. Mint a 2.3 ábrán látszik ilyen paraméterek mellett, a banknak inkább megéri hitelt nyújtania. Ha több iterációt végzünk, az arány akkor még inkább az ügyfélnek kedvező irányba tolódik el.

2.3. ábra. K1=K2

A 2.1 táblázatban is jól látszott, hogy a főátlóhoz "tapadnak" az egyesek. Ezért azt vizsgáltam meg, hogy mekkora r/n hányados fölött mekkora valószínűséggel éri meg hitelt nyújtani, ha 40 ciklust futott a programunk. $(R, N, I = 40)$. A vizsgálat közben, azt is fontosnak tartottam, hogy a nyereség-veszteség mértéke változhat, ezért $10\,000\,\mathrm{HUF}$-os léptékkel növeltem a bevétel értékét, amíg a hitel összegét rögzítettem $100\,000\,\mathrm{HUF}$-ban. Nézzük is meg egy konkrét példán. Ha $K_1 = 30\,000\,\mathrm{HUF}$ és 90%-ig szeretnénk biztosak lenni abban, hogy megéri hitelt nyújtani, akkor a szürke oszlop alapján az r/n hányadosunk legalább $0,61$ kell, hogy legyen. A 2.4 grafikon, ezeknek a méréseknek az eredményeit mutatja.



2.4. ábra. Az r/n hányados kapcsolata a várhatóan jó hitelek számával

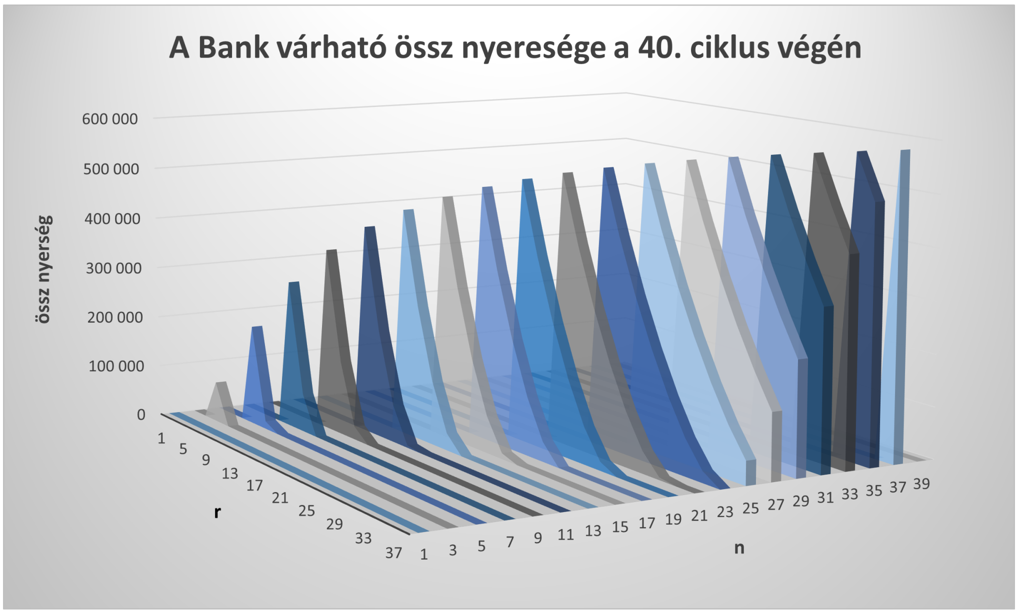
A 2.2 HOL látható táblázat az eredeti példánk, azaz az alábbi paraméterekkel futtatott program eredményének egy kis szelete:

- $R, N, I = 40$
- $a = 1/(1,05)$
- $K_1 = 38\,000\,\mathrm{HUF}$
- $K_2 = -62\,000\,\mathrm{HUF}$

Ahol a 2.2 táblázatban 0 van, az azt jelöli, hogy nem éri meg hitelt nyújtani. Megfigyelhető, hogy ahol az r/n hányados kevesebb egy soron illetve oszlopon belül, ott a várható nyereség is kevesebb. Azaz a számok balról jobbra illetve lentről felfele növekvő sorrendben vannak. Ezzel magyarázható, hogy a 2.5 grafikonon olyan derékszögű "háromszögek" vannak, amiknek az "átfogója" lépcsőzetes. Ahogy nő r és n egyre hamarabb "lecsapódik" az "átfogó", ennek az az oka, hogy a 40×40 táblázatunk is nagyon hasonlít a 20×20 táblázatunkra (2.1). Olyan szempontból, hogy oszloponként először nő a nullától különböző értékek száma, aztán elkezd csökkeni. Így például a 2.5 ábránkon $r = 37$ és $n = 37$ oszlopot nézzük olyan, mintha a 2.1 táblázatban az $r = 19$ és $n = 20$ cellát néznénk, azaz már csak egy érték szerepel.

r/n	6	7	8	9	10
7	426 323	X	X	X	X
8	245 107	456 733	X	X	X
9	119 579	288 073	480 917	X	X
10	41 087	162 269	324 556	500 502	X
11	3 242	74 182	201 816	355 500	516 630
12	0	20 924	109 033	237 546	381 837
13	0	0	44 837	143 610	269 345
14	0	0	8 670	72 802	176 459
15	0	0	0	25 245	102 385

2.2. táblázat



2.5. ábra

3. fejezet

Ügyfelek tulajdonságainak kockázatmérése

A következő fejezet az alábbi cikkek feldolgozásával készült: [7] [8]. Szeretném kiemelni, hogy ezek a cikkek 1940 − 60-as évek módszereit mutatják be. A mai világban már nem így vagy nem ezeket a megoldásokat használják a bankok. Azonban most is fontos, hogy milyen pénzügyi előmenete van a hitelkérőnek és milyen ingatlanokkal vagy ingóságokkal rendelkezik.

A továbbiakban a hitelnyújtás illetve nem nyújtás kérdéskörét onnan fogom megközelíteni, hogy az emberek tulajdonságai, neme, életkora, eddigi tartozásai, munkahelyei száma, ingóságai illetve ingatlanjainak értéke miként befolyásolja azt, hogy ő várhatóan egy nyereséget vagy veszteséget okozó ember lesz a cég számára.

Ez a módszer a hitelnyújtás problémájának arra a részére fókuszál, hogy megbecsüljük a valószínűségét annak, hogy bizonyos tulajdonságokkal rendelkező ember fizetőképes vagy sem. A statisztikai és logikai problémák miatt, itt nem egyből fogunk kapni egy konkrét valószínűséget, mint eddig hanem osztályozzuk az embert a tulajdonságai együttese alapján. Az osztályzatot, amit kapott pedig átkonvertáljuk egy becsléssé, ami egy valószínűségi érték lesz, hogy várhatóan mekkora eséllyel nem fizeti vissza az ügyfél a hitelt.

Eddig a jó és a rossz hitelt az alapján különböztettem meg, hogy az ügyfél teljesen visszafizette-e a tartozását vagy sem. Azonban a valóságban nem ennyire fekete-fehér, hiszen ha egy több éves törlesztés alatt 1-2 részletet kihagyott, az még a banknak lehet nyereséges. Elméletben az a hitel nyereséges, ahol a bank bevétele egy adott hitel esetén fedezi az összes kiadást, azon hitel esetében. Azonban ezt a valóságban egyáltalán nem egyszerű megmondani. A bankok emiatt gyakran több csoportja osztják az embereket. A, B és C csoportba. Az A csoportban azok a hitelek vannak, amik mindig késedelem nélkül lettek fizetve. A B csoportban azok, amiket törlesztettek, de előfordult, hogy késve lettek visszafizetve. A C-ben pedig azok, amik egyértelműen veszteségesek, például bírósági perrel fejeződtek be. A rossz kontraktus alatt azt értem a következőkben, hogy ahhoz az emberhez tartozó hitel, aki részben vagy egészben nem fizeti vissza a tartozását. Amíg a jó kontraktus alatt pedig azt értem, hogy a hitelező teljesen visszafizeti a tartozását. A módszer az alábbi 4 lépésből áll:

1. Összehasonlítjuk a jó és a rossz ügyfelek profilját és olyan jellemzőket keresünk, amik a rossz ügyfelekre jellemzőbbek.
2. A rossz kontraktus valószínűségének kiszámítása.
3. A rossz kontraktus valószínűségéből kockázati index kidolgozása, amely a számlák összehasonlítására és osztályozására használható.
4. A kockázati index értékelése és a kockázati indexhez értékéhez rendelt rossz kontraktus valószínűségének a kiszámítása.

3.1. Megkülönböztető jellemzők keresése

A kockázathoz kapcsolódó jelentős jellemzők keresését korlátozza a rendelkezésre álló információ halmaz és nehezíti, bonyolítja a tulajdonságok nagy száma. Továbbá nehezíti, hogy mely tulajdonság kombinációk relevánsak és melyek nem. Ebben az alfejezetben nem igazi "keresés" fog zajlani, hanem igazolni fogjuk, hogy egy-egy tulajdonság gyakrabban fordul elő jó illetve rossz hitelek esetén.

Nézzük meg egy konkrét példán. Kaptunk 100 jó és 100 rossz hitelfelvevőről különböző adatokat. A hitelezők fontos tulajdonságnak tartják, hogy egy lehetséges ügyfélnek mennyire stabil a munkahelye. Ezt abból következtetik ki, hogy hány éve dolgozik ugyanott. Azt szeretnénk megvizsgálni, hogy igazolják-e a minták azt, hogy egy stabil(abb) munkahellyel rendelkező egyén tényleg gyakrabban lesz jó vagy esetleg rossz ügyfél, mint aki rövidebb ideje dolgozik ugyanott a hitel felvétel pillanatában. A 200 hitelből, csak egynél nincs információnk erről az adatról. A számadatok jelentős részét a cikkből emeltem át.

Az egyik lehetséges módszer, hogy a két mintának az átlagát kiszámoljuk. Az eredmény a jó hitelek esetében $10,76$ év volt a rosszban pedig $7,16$. Itt szeretném megjegyezni, hogy a példánk 60-'as évekbeli, amikor az emberek ritkábban változtattak munkahelyet, mint napjainkban. Illetve az akkori szakértők ezt kimagaslónak találták, azonban a példánk szempontjából egyik sem releváns. Ebből arra következtethetünk, hogy a múltban a hosszú távú munkahellyel rendelkező ügyfelek kevésbé voltak kockázatosak.

A másik módszerben a jó és a rossz hitelek százalékos arányát vizsgáljuk az ügyfelek bizonyos csoportja szerint. A 3.1 táblázatban az első oszlop az adott munkahelyi évek számát mutatja, a második és harmadik oszlop rendre a rossz, illetve jó kölcsönök százalékos eloszlását. Az utolsó oszlop pedig a rossz és jó hitelek arányát mutatja. Nézzünk egy konkrét példát a táblázatból, a jó hitelekkel rendelkező ügyfelek $45,5$%-a 10 vagy annál több ideje dolgozik a jelenlegi munkahelyén. Ameddig a rossz hitelekkel rendelkező ügyfeleknek csupán 22%-a. A jó hitelek esetében csak 99 emberről volt információnk, azért vannak tört számok.

Ha összesítenénk az intervallumokat, akkor a rossz és jó hitelek aránya 1 lenne. Szóval az átlagos arány az az 1. Ha 1-nél nagyobb arány, akkor az átlagosnál nagyobb kockázatot jelent a banknak. Ha 1-nél kisebb az arány akkor pedig az átlagosnál jobbat. A példánkban ez azt jelenti, hogy $0 - 10$ évig az átlagosnál nagyobb a kockázat, hiszen a rossz hitelek aránya a jó hitelekhez képest rendre $1.4, 1.6, 1.4$. Amíg a 10 év vagy annál magasabb foglalkoztatottság egy helyen kevésbé kockázatos. Azonban elég kicsi mintával dolgozunk a példánkban, ezért javasolt több ezres vagy tízezres mintával dolgozni, hogy pontosabb képet kapjunk az átlagról.

Jelenlegi munkahelyi évek	Rossz kölcsönök %	Jó kölcsönök %	Rossz/Jó hitel
0-3	30	22,2	1,4
3-6	30	19,2	1,6
6-10	18	13,1	1,4
10 vagy annál több	22	45,5	0,5

3.1. táblázat

3.2. A rossz számla valószínűségének kiszámítása

Az egyes ügyféli tulajdonságokhoz kapcsolódó rossz kontraktus valószínűsége kiszámítható példaadatokból vagy a tulajdonságoknak a megkülönböztetésére használt módszerek eredményeinek eloszlásából.

Az előző alfejezettel ellentétben itt nem egyezik meg a jó hitelek száma a rossz hitelek számával. Itt veszünk egy bank eddigi adatait és ebből fogunk számolni. A mintánk nagysága: (107460) A szignifikáns különbségekre vonatkozó tesztekben használt eloszlások feltételes valószínűséggé alakíthatóak a jó és rossz kontraktus besorolására. Ezt úgy tehetjük meg, hogy az egyes tulajdonságokkal rendelkező jó kontraktusok számát elosztjuk az összes jó kontraktusok számával, majd ezt megismételjük a rossz kontraktusokkal is. Ezekből a feltételes valószínűségekből és a Bayes-tétel alapján, meg tudjuk határozni az a posteriori valószínűségeket.

Ahol legyen:

- R = rossz kontraktusok
- J = jó kontraktusok
- A_i = egy diszkriminatív jellemző az i osztályú tulajdonság szerint. Osztály alatt itt a tulajdonságok csoportját értem. Például, milyen régóta dolgozik az ügyfél egy munkahelyen, vagy hány éves, milyen a lakóhelye. (Saját ház, bérelt ház/lakás/szoba vagy egyéb)

$$\begin{aligned} P(R|A_i) &= \frac{P(R)P(A_i|R)}{P(R)P(A_i|R) + P(J)P(A_i|J)} \\ &= \frac{P(R, A_i)}{P(A_i)} \end{aligned} \tag{3.1}$$

3.3. Kockázati index kialakítása

A cél egy adott kontraktus rosszá válásának valószínűségének becslése. Az adott ember, akinek a szerződéséről szó van minden jellemzőjének az értékét külön-külön, és a jellemzők összes lehetséges kombinációját össze kell vetni az adott kontraktus várható minősítésével, azaz jó vagy rossz szerződésről beszélünk. Elméletileg a

Lakás típus	Rossz hitel	Összes hitel	Rossz hitel valószínűsége
Saját ház	564	81 985	0,0069
Bérelt ház	70	3153	0,0222
Bérelt lakás	331	10141	0,0326
Bérelt szoba	105	1436	0,0731
Egyéb	127	8956	0,0142

	Feltételes valószínűség		Együttes valószínűség		
Lakás típusa (A)	$P(A_i\|J)$	$P(A_i\|R)$	$P(J, A_i)$	$P(R, A_i)$	$P(A_i) = P(J, A_i) + P(R, A_i)$
Saját ház (1)	0,779	0,471	0,7700	0,0054	0,7754
Bérelt ház (2)	0,029	0,058	0,0287	0,0007	0,0294
Bérelt lakás (3)	0,094	0,277	0,0929	0,0032	0,0961
Bérelt szoba (4)	0,013	0,088	0,0129	0,0010	0,0139
Egyéb (5)	0,085	0,106	0,0840	0,0012	0,0852
	1,000	1,000	0,9885	0,0115	1,0000

legjobbnak tűnik, hogy az összes lehetséges kombinációnak az a posteriori valószínűségét kiszámoljuk a 3.2 szakaszban bemutatott módszerrel. Két gyakorlati nehézség azonban felmerül(t). Az első, hogy általában nincs annyi rossz konstrukcióról adatbázis, amiből megbízhatóan ki lehetne számolni a tulajdonságok kombinációnak valószínűségét. Legalábbis 1964-ben nem volt elég adat, sajnos nem találtam egyértelmű választ, hogy ma már van-e vagy sem. Valószínűleg már van, de kérdéses, hogy van-e olyan személy vagy vállalat, aki hozzáfér ezekhez. A második probléma, hogy a sok ember osztályozása bonyolult és időigényes lenne és a kidolgozott valószínűségi táblázatokat igényelne a tulajdonságok összes lehetséges kombinációjához. Azonban ezek az a problémák nem feltétlenül kritikusak, ha kevés változót használunk.

A különböző tulajdonságok valószínűségei logikai szabályokkal nem kombinálhatóak, mert nem tehetünk igazolható feltételezéseket a különböző tulajdonságok egymáshoz való viszonyáról. Statisztikailag nyilvánvaló, hogy nem függetlenek. Ha mégis kombinálnánk őket az bizonyos esetekben kioltanák a másik jó vagy rossz hatását vagy bizonyos esetekben túlzóan eltolódna egyik vagy másik irányba a hatásuk. Ezért közelítsük meg a problémát úgy, ha bármely két szerződés tulajdonságai megegyeznek egy kivételével. Akkor az a tulajdonság kockázatosabb, amelyik a rossz kontraktushoz tartozik, mint a másik. Ha ezt az érvelést kibővítjük hihetőnek tűnik, hogy az egyéni tulajdonságok valószínűségeinek összege alapul szolgálhat a kontraktusok kockázat szerinti ragsorolásához. Az így kapott összeget szorozva 1000-rel megkapjuk a kockázati indexet, amit így írhatunk fel:

$$R = (pa_i + pb_i + pc_i + \ldots + pn_i) \cdot 1000 \tag{3.2}$$

Ahol legyen:

- R Kockázati index
- p Az egyes tulajdonságok valószínűsége, amit a 3.1 képlet segítségével tudunk kiszámolni
- $a, b, c, \ldots, n$ Az egyes tulajdonságok az i-edik csoportban [1]

Amennyiben bebizonyítható, hogy az index hatékony, akkor számos előnye van, ami miatt érdemes használni. Például könnyű felépíteni a rendszert a használatához és sokféle tulajdonság beépítésére alkalmas. Alkalmazható olyan tulajdonságokra is, ami a szokottól eltérő lehet, mondjuk városrészek vagy irányítószámok alapján. Ez érdekes lehet, amikor a lehetséges ügyfelek egy jelentős része olyan helyen dolgozik, ahol köztudott, hogy az emberek %-ban kifejezhető hányada egy cégnél dolgozik ilyen például Zürichben a Google, bár várhatóan a Google nem fog csődöt jelenteni. Magyarország szegényebb régióiban viszont vannak települések, ahol egy gyár foglalkoztatja az ott élő népesség magas hányadát, ezekről a gyárakról viszont sajnos kevésbé biztos, hogy nem mennek csődbe vagy dönthetnek úgy, hogy a bizonytalan gazdasági helyzet vagy más okok miatt nem települnek ki az országunkból. Ez a módszer nagyon egyszerű megoldást nyújt az egyéni kontraktusok osztályozására. A lényeges tulajdonságokhoz tartozó értékeket összeadjuk, ezzel megkapva az adott ügyfél kontraktusának kockázati értékét.

[1] itt csoport alatt azt értjük például, hogy milyen házban lakik az ügyfél, milyen régóta dolgozik ugyanott stb.

3.4. Kockázati index értékelése

A kockázati index hatékonyságát azzal lehet mérni, hogy a kockázati indexet csoportosítjuk 10-esével, 20-asával, 50-esével, vagy ahogy szeretnénk. A csoportokban pedig összehasonlítjuk az összes és a rossz szerződések arányát. Ha a rossz hitel valószínűsége együtt emelkedik a kockázati index értékével, mint a 3.2 táblázatban, akkor jól működik. Ez azt is jelenti, hogy a rossz hitelek gyakrabban előfordulnak ott, ahol a kockázati index magasabb. A példánkban is megjelenik. A rossz hitelek harmada (416 db) 250 vagy a fölötti kockázati index tartományban van, ameddig a jó kontraktusoknak csak 3% esik ugyanebbe az indexérték csoportba. Ha a 300+ tartományt vizsgáljuk, akkor azt látjuk, hogy a rossz hitelek 12%-a van ott, amíg a jóknak csak $0,2\%$-a.

Kockázati index	**Összes hitel**	**Rossz hitel**	**Rossz hitel valószínűsége**
110-119	1754	2	0,0011
120-139	28756	97	0,0034
140-159	35748	152	0,0043
160-179	18620	161	0,0086
180-199	8875	117	0,0132
200-229	6605	160	0,0242
230-259	4109	184	0,0448
260-269	749	48	0,0641
270-279	766	65	0,0849
280-299	304	94	0,3092
300+	364	154	0,4231

3.2. táblázat

A példánkban az alábbi 13 tulajdonságokat vettük figyelembe:

- Telefon
- Lakhatás típusa
- Bankszámla és annak használati szokásai
- Kölcsön célja
- Kölcsön futamideje

- Utolsó tartozás óta eltelt idő
- Jelenlegi munkahelyen eltöltött idő
- Ügyfél neme és párkapcsolati státusza
- Lakhely kerülete
- Felvenni kívánt kölcsön időtartama
- Bevételének értéke
- Ügyfél kora
- Hány fős a család

3.5. Kockázati index használata

Ez az alfejezet már nem tartozik bele a módszer 4 lépésbe. Itt arról lesz szó, hogy lehet használni a kockázati indexet a gyakorlatban.

Amennyiben a kockázati indexet arra szeretnénk használni, hogy elutasítsunk lehetséges szerződéseket, akkor fontos, hogy a lehetséges jó számlák aránya a rossz számlákhoz képest a lehető legkisebb legyen az index magasabb tartományaiban. Ezt a küszöbindexet a példánkban válasszuk $0,25$-nek. A 3.2 táblázatunk negyedik oszlopában kell nézni. Azoknál a kockázati indexeknél, amik ennél magasabbak az összes lehetséges hitelt visszautasítanánk, akkor legfeljebb 3 várható jó hitel veszne el minden lehetséges rossz hitel után. Ezt a mértéket nevezzük hatékonysági rátának (E).

$$E = \frac{\text{rossz hitel valószínűsége} > 0,25}{\text{összes rossz hitel}} \tag{3.3}$$

Ugyan a legtöbb bank kifejezetten elégedett lenne, az előző bekezdésben ismertetett koncepcióval, azonban mégsem utasítanának, vagy kötnének szerződést csupán egy szám alapján. Némi ellenőrzés után, ha biztosan kizárták a csalást a kapott indexértéket inkább útmutatónak használják. Ez alapján mérlegelik, hogy megkötik-e szerződést vagy sem, esetleg tovább vizsgálják.

A legjobb vágási pont megtalálása viszonylag egyszerű. Vágási pont alatt itt azt értem, hogy milyen indexértéknél húzzunk meg egy képzeletbeli vonalat, ami alatt megéri hitelt nyújtani, de a fölött már nem. A rossz hitelek várható vesztesége ne haladja meg a jó hitelek várható nyereségét. A rossz hitelek veszteségének várható-értéke megbecsülhető úgy, hogy a rossz hitelek átlagos veszteségét (L) megszorozzuk a rossz hitelek valószínűségével p. A jó szerződések várható nyeresége pedig úgy becsülhető meg, hogy a jó hitelek átlagos nyereségét (C) megszorozzuk a jó hitelek valószínűségével g. A vágási pont az, amikor a rossz hitelek várható vesztesége (pL) megegyezik a jó hitelek várható bevételével (gC). Mivel a g egyenlő $1 - p$-vel, ezért a vágási pont a következőképp adható meg:

$$\begin{aligned} pL &= (1-p)C \\ pL + pC &= C \\ p &= C/(L+C). \end{aligned}$$

Nézzünk két gyakorlati példát. **Az első példában** az $1960 - 70$-es évek Amerikájából hoztam, ahol a várható nyereség jelentősen kisebb, mint a várható veszteség. Egy hitel átlagos megtérülése, az összes költség levonása után $C = 20\,\mathrm{USD}$. A veszteséges hitelek esetén, ahol van behajtási vagy per költség ott $L = 400\,\mathrm{USD}$, ekkor:

$$\begin{aligned} p &= \frac{20}{(20+400)} \\ p &= 0,0476 \end{aligned}$$

ez a korábbi példánkban azt jelenti, ha rossz hitel valószínűsége meghaladja ezt a számot, akkor veszteséges lenne. Ha nem haladja meg akkor pedig nyereséges. Ha ez alapján a valószínűség alapján szeretnénk megmondani, hogy milyen kockázati index értéknél húzzuk meg a határt akkor a 3.2 táblázat alapján látjuk, hogy 260 pontnál érdemes. Ez azt is jelenti, hogy a bank akár 15 jó számláról hajlandó lemondani, hogy egy rosszat elkerüljön. Mivel akár 749 számláról képes lemondani, azaz 699 jó szerződésről, hogy 48 rossz kontraktust elkerüljön. Természetesen, ha nézzük, a magasabb tartományokban lévő hiteleket is akkor arányaiban sokkal kevesebbről mond le.

A második példában Ezt a példát én hoztam létre a központi statisztikai hivatal adatai alapján. 2021 Magyarország lakáshiteleire voltam kíváncsi. [9]. Ebben a példánkban a várható nyereség még mindig sokkal kisebb, mint a várható veszteség. Azonban itt már "csak" a fele, szemben az előző példával, ahol $\frac{1}{20}$-a volt. A bankok által nyújtott lakossági lakáshitelek száma 334 472 darab volt. A hitelek összértéke pedig 2 688 milliárd forint volt. Ez azt jelenti, hogy átlagosan 8 036 547 forint volt egy-egy felvett hitel.

Pontos profitot nem tudunk meghatározni, mert bankonként, futamidőnként különböző. Az átlagos futamidő 16 év [10]. Amennyiben átlagos futamidővel számolunk, infláció nélkül, akkor a legkedvezőbb konstrukció, hogy 15 000 000 forintot kell visszafizetnünk a banknak. Ekkor átlagosan 5 millió forint körüli nyeresége van a banknak hitelenként, hiszen ki kell fizetnie az alkalmazottakat, irodát és még sok minden mást. Diszkontálva [11] ezt az összeget 5%-os inflációval 2 290 557 forint (C) lesz a tényleges nyeresége. A veszteség, úgy sem kicsi, hogy van fedezet vagy jelzálog az ilyen hiteleknél, hiszen fizetni kell az ügyintézőket/az aukciót stb., stb. Különböző kérelmekkel, bírósági pereskedésekkel, pedig az is lehet, hogy soha nem lát a bank egy fillért sem, de ugyanúgy van kiadása ezekkel. Ezért az eljárás költsége 4 000 000 forint (L) körül lehet. (Erre sajnos nem találtam megbízható forrást, ezért ezt a számot úgy választottam, hogy a példánkban látványos legyen, a kockázati index megválasztása.) Ezekkel az adatokkal, ha behelyettesítünk a képletbe

$$p = \frac{2\,290\,557}{(4\,000\,000 + 2\,290\,557)}$$
$$p = 0,3641$$

kapjuk. Ebben példánkban ez azt jelenti, ha rossz hitel valószínűsége nagyobb, mint $0,3641$ akkor várhatóan veszteséget jelentene a banknak. Ha nem haladja meg akkor pedig nyereséget. Ha ez alapján szeretnénk meghúzni a kockázati index határát a 3.2 táblázat alapján látjuk, hogy 300+ pontnál érdemes. Ez azt is jelenti, hogy a bank $1,68$ jó számláról hajlandó lemondani, hogy egy rosszat elkerüljön.

Mint láthattuk minél kisebb a szorzó a várható nyereség és várható veszteség között, annál kevesebb lehetséges jó számláról "kell" arányaiban lemondania a hitelintézetnek ahhoz, hogy várhatóan ne legyen veszteséges. Ez a kockázati indexek vonatkozásában pedig annyit tesz, hogy ahogy csökken a várható veszteség, nyereség hányados, annál magasabb kockázati indexnél éri el a bank, hogy várhatóan nem lesz veszteséges.

4. fejezet

Lezárás

A szakdolgozatomban bemutattam modelleket arra, hogy milyen módszerek segítik/segítették a bankokat abban, hogy kibe éri meg befektetniük a várható nyereség reményében. Természetesen nagyon sok másik módszer is létezik, amivel én nem foglalkoztam a diplomamunkám keretein belül. A gyakorlatban gyakran különböző módszereket együtt használnak, akár úgy, hogy több módszert kombinálnak vagy egymás ellenőrzésére használják őket. Kombinálni lehet őket úgy is, hogy az egy egyszerűbb gyorsabb modellt szűrőként használnak, egy lassabb, bonyolultabb modellt, pedig csak a kérdéses ügyfeleken alkalmazzák. Továbbá úgy is lehet kombinálni őket, hogy egymásba ágyaznak több módszert.

A hitelnyújtás problémája folyamatos foglalkoztatja a bankok világát, így folyamatosan fejlesztik a módszereiket. Több éve már a mesterséges intelligenciát is bevonták, a hatékonyabb eredmények reményében.

Hivatkozások

1. Kapralos, C. *A brief history of lending through the ages* `https://www.koho.ca/learn/history-of-lending/`.
2. Bierman, H. & Hausman, W. H. The Credit Granting Decision. *Management Science* 16. köt., B519–B532. `https://EconPapers.repec.org/RePEc:inm:ormnsc:v:16:y:1970:i:8:p:b519-b532` (1970).
3. Borovkov, A. A. *Matematikai Statisztika* (Typotex Kiadó, 1999).
4. Raiffa, H. & Schlaifer, R. *Applied Statistical Decision Theory* 50–51. old. `https://books.google.hu/books?id=dX2BxgEACAAJ` (Division of Research, Graduate School of Business Administration, Harvard University, 1961).
5. KSH. *A fogyasztóiár-index fogyasztási főcsoportok szerint, a nyugdíjas fogyasztóiár-index és a maginfláció* `https://www.ksh.hu/stadat_files/ara/hu/ara0002.html`. [Online; megtekintve a cikk: 2023-04-12]. 2022.
6. ECB. *Inflation rate* `https://sdw.ecb.europa.eu/`. [Online; megtekintve a cikk: 2023-04-12]. 2023.
7. Smith, P. F. Measuring Risk on Consumer Instalment Credit. *Management Science* 11. köt., 327–340. old. `https://EconPapers.repec.org/RePEc:inm:ormnsc:v:11:y:1964:i:2:p:327-340` (1964).
8. Durand, D. *Risk Elements in Consumer Instalment Financing, Technical Edition* 22–43. old. `http://www.nber.org/chapters/c12950` (NBER, 1941).
9. KSH. *Lakossági lakáshitelezés, 2021* `https://www.ksh.hu/docs/hun/xftp/idoszaki/lakashitel/20214/index.html`. [Online; megtekintve a cikk: 2023-04-12]. 2022.

10. KSH. *Folyósított lakáscélú hitelek, 2021* https://www.ksh.hu/stadat_files/lak/hu/lak0016.html. 2022.

11. Elemzésközpont. *Diszkontált CashFlow: Mit jelent? Hogyan működnek a DCF modellek* https://elemzeskozpont.hu/diszkontalt-cashflow-mit-jelent-hogyan-mukodnek-dcf-modellek. [Online; megtekintve a cikk: 2022-11-06]. 2022.