EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

BUDAPESTI CORVINUS EGYETEM
KÖZGAZDASÁGTUDOMÁNYI KAR

SZTOCHASZTIKUS MÓDSZEREK A NEM-ÉLETBIZTOSÍTÁSOK TARTALÉKOLÁSÁBAN

MSc szakdolgozat

Írta: Orbán Barbara
Biztosítási és Pénzügyi Matematika MSc, aktuárius szakirány

Témavezető:
Antalffy-Németh Gabriella
aktuárius, Generali Biztosító Zrt.

Budapest, 2014

KÖSZÖNETNYILVÁNÍTÁS

Ezúton szeretnék köszönetet mondani témavezetőmnek, Antalffy-Németh Gabriellának, aki hasznos útmutatásaival és precíz észrevételeivel segítette a munkámat. Neki köszönhetem, hogy megismerkedhettem a téma gyakorlati alkalmazásával, és ezzel jelentősen hozzájárult a szakmai fejlődésemhez.

Köszönettel tartozok továbbá a Generali Biztosító Zrt. Aktuáriusi Osztályának, amiért lehetővé tették számomra az alkalmazás során felhasznált ResQ program használatát.

Végül szeretném megköszönni a családomnak és a barátaimnak, hogy az egyetemi évek alatt végig támogattak és bíztattak.

TARTALOMJEGYZÉK

1. FEJEZET

BEVEZETÉS

A nem-életbiztosítással foglalkozó aktuáriusok egyik legfontosabb feladata a kártartalékok képzése és ellenőrzése, valamint hogy legjobb becslést (*„best estimate”*) adjanak a szerződésállományon eddig bekövetkezett összes kár összes jövőbeli kifizetésére. A jelenleg érvényben lévő EU-s rezsim, a Szolvencia I. rendszer tartalékolási elvei számos szempontból kritizálhatóak. Ez a rendszer többek között nem kockázatérzékeny, nincs tekintettel az alakuló IFRS-re, valamint az alapjául szolgáló modell napjainkban korlátozottan tekinthető érvényesnek. Mindezek versenyhátrányt jelentenek a biztosítók világversenyében, illetve a biztosító tőkéjének nem megfelelő hasznosulásához vezethetnek. A hibák kiküszöbölésére szolgál a napjainkban bevezetés alatt álló Szolvencia II. rendszer, amely az EU-ban letelepedett biztosítókra és viszontbiztosítókra vonatkozik. Az új, a banki Bázel II. rendszerhez hasonló hárompilléres rendszer bevezetése az aktuáriusokat további feladatok elé állította. A legfontosabb újítás a szavatoló tőkeszükséglet számításában jelenik meg. A Szolvencia II. rendszer előírja, hogy a szavatoló tőkének 1 éves időhorizontra vonatkozóan 99,5%-os valószínűséggel elegendőnek kell lennie minden felmerülő kötelezettség teljesítésére. Ennek a szavatoló tőkének az egyik eleme a tartalékolási kockázat, amely a best estimate becslésének pontosságához kapcsolódik. Mivel az összes kifizetés mértékére nem tudunk pontos előrejelzéseket adni, ezért felhasználjuk, hogy a felmerülő károk mögött valamilyen eloszlás áll, és ezen eloszlás alapján adjuk meg a becslési hibát és az ebből adódó tőkeszükségletet. Bár a determinisztikus tartalékolási technikákkal a legjobb becslés számítható ki, nem mondanak túl sokat a becslés pontosságáról.

Az elmúlt három évtizedben kapott nagyobb figyelmet az aktuáriusok körében a sztochasztikus tartalékolás témája, illetve ezen újszerű modellek és a klasszikus lánc-létra módszer közötti összefüggések tanulmányozása. A sztochasztikus kártartalékolási módszerek célja, hogy pontosan ugyanazt a becslést adják a tartalékok mértékére, mint a lánc-létra módszer, azonban a becslések pontosságát is megkapjuk belőlük. Ezen modellek statisztikai eszközökkel vizsgálják az illesztett modellek „jóságát”. Fontos eszközük a reziduálisok elemzése, mivel a reziduálisok rávilágítanak az illesztett modell eltéréseire az eredeti adatoktól.

A szakdolgozat célja, hogy bemutassunk néhány sztochasztikus módszert, és rávilágítsunk a köztük lévő összefüggésekre, illetve ezek közül néhányat a gyakorlatban is alkalmazunk. Fontos hangsúlyozni, hogy a bemutatott modellek közül néhány jobb illeszkedést mutat, ha a kárkifizetésekre, vagy a bekövetkezett károk számára alkalmazzuk őket, mintha az ún. *„incurred”* háromszöggel dolgoznánk, amely a kifizetésen felül a tételes függőkár tartalékot is tartalmazza. Ennek oka, hogy ez utóbbi esetben felülbecslés fordulhat elő a tételes függőkár tartalék megállapításánál, ami negatív inkrementális értékekhez vezethet. A dolgozatban *pénzügyi évként* hivatkozunk a kárbekövetkezéstől számított j-edik évre, ez a felhasznált angol nyelvű szakirodalmak *„development year”* vagy *„financial year”* kifejezésének felel meg. A téma kiindulópontjául az *Általános biztosítás II.* kurzuson megismert lánc-létra módszer fog szolgálni.

2. FEJEZET

AGGREGÁLT KÁROK MODELLEZÉSE

Tegyük fel, hogy rendelkezésünkre állnak a bekövetkezett károk nagyságára vonatkozó adatok, és ezek háromszög alakba rendezettek. Az alábbi jelölést alkalmazzuk:

$$\{X_{i,j}: i = 1, \dots, n; j = 1, \dots, n - i + 1\},$$

ahol i a kárbekövetkezés évét jelenti, a j index pedig a kárbekövetkezéstől számított j. pénzügyi évre utal. Az $X_{i,j}$ *inkrementális kárnagyságokat* tartalmazó felső háromszögmátrixot fogjuk *kifutási háromszögnek* nevezni. A *kummulált károkat* a következőképpen definiáljuk:

$$C_{i,j} = \sum_{k=1}^{j} X_{i,k}$$

Bevezetjük az úgynevezett *növekedési faktorokat*: $\{\lambda_j: j = 2, \dots, n\}$. Tudjuk, hogy lánc-létra módszer a növekedési faktorokra a következő becsléseket adja:

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} C_{i,j}}{\sum_{i=1}^{n-j+1} C_{i,j-1}}.$$

A kummulált kárráfordítás jövőbeli értékére az alábbi formulák alkalmazásával kapunk becslést:

$$\hat{C}_{i,n-i+2} = C_{i,n-i+1}\hat{\lambda}_{n-i+2}$$

$$\hat{C}_{i,k} = \hat{C}_{i,k-1}\hat{\lambda}_k, \quad k = n - i + 3, n - i + 4, \dots, n.$$

Ezek a becsült értékek lesznek a kummulált kifutási háromszögünkhöz tartozó alsó háromszögmátrix elemei. Ebből könnyen kiszámítható a megfelelő inkrementális alsó háromszögmátrix. A továbbiakban *tartalék* alatt az inkrementális alsó háromszögmátrix elemeinek összegét értjük.

Ha a lánc-létra módszert a fenti formában alkalmaztuk, akkor az úgynevezett *„ultimate"*, vagyis végső kárnagyságokat becsültük meg. Ez alatt most azt értjük, hogy csak az adott kárévtől számított n. évre vonatkozólag lesznek becsléseink, és nem foglalkozunk a károk eloszlásának későbbi viselkedésével („*tail factor*"). Ez azért veszélyes, mert előfordulhat, hogy nem a várakozásainknak megfelelően alakulnak a károk és lehet, hogy ekkor nem lesz elegendő tartaléka a biztosítónak. Ez adódhat például az ultimate kárnagyságok alulbecsléséből. A károk változékonyságának mérésére az *előrejelzési hibát* alkalmazzuk, amely definíció szerint a tartalék eloszlásának szórása.

Ahhoz, hogy meghatározzuk az előrejelzés hibáját, feltevéseket kell alkalmaznunk. Ha az a cél, hogy a lánc-létra módszer becsléseihez hasonló eredményekre jussunk egy sztochasztikus módszer segítségével, ezt kétféleképpen érhetjük el: vagy az eloszlást határozzuk meg az adatokból, vagy csak az első két momentumot. A továbbiakban az első módszerre példaként bemutatjuk az *ODP* („*over-dispersed Poisson"*) és a *negatív binomiális modellt*, a második módszert pedig a *Mack-modellen* keresztül szemléltetjük.

2.1 ODP MODELL

Az ODP-eloszlás annyiban különbözik a Poisson-eloszlástól, hogy az előbbi szórásnégyzete nem egyenlő a várható értékével, hanem egy arányszám szorzóval megkapható belőle. Erre utal az elnevezésben a „túlszórtság". Az ODP modell feltételezése, hogy az inkrementális $X_{i,j}$ értékek független, ODP-eloszlású véletlen változók, az alábbi tulajdonságokkal:

$$E\left[X_{i,j}\right] = m_{i,j} = x_i y_j \quad \text{és} \quad D^2\left[X_{i,j}\right] = \phi x_i y_j,$$

ahol

$$\sum_{k=1}^{n} y_k = 1\,.$$

Itt x_i az i. kárbekövetkezési évre vonatkozó végső kár várható értéke *(„expected ultimate claims")*, y_j pedig a j. pénzügyi évre vonatkozó arány, melynek segítségével megkapjuk az adott évi kárráfordítás nagyságát az ultimate kárnagyságból. A „túlszórtságra" az ismeretlen ϕ paraméter utal, melyet az adatokból becslünk. Azzal, hogy a varianciában megengedjük ezt a ϕ szorzót, a paraméterbecsléseken nem változtatunk, viszont a becslés standard hibájára hatással van (ld. [9]). Mivel az y_j megjelenik a szórásnégyzetben, automatikusan feltettük, hogy értékei csak pozitívak lehetnek. Ezért az inkrementális károk összege a j. oszlopban is pozitív lesz, ami a modell egyik korlátja. Néhány negatív érték előfordulhat a mátrixban, az oszlopösszegeknek azonban pozitívnak kell lennie.

2.2 NEGATÍV BINOMIÁLIS MODELL

Ez a modell hasonló az előzőekben bemutatott ODP modellhez, de a paraméterei jobban hasonlítanak a lánc-létra módszer növekedési faktorához. A negatív binomiális modell kifejezhető inkrementális és kummulált kárnagyságokkal is. A $X_{i,j}$ inkrementális kárnagyságok eloszlása „túlszórt" negatív binomiális, ha az alábbiak teljesülnek:

$$E\left[X_{i,j}\right] = (\lambda_j - 1)C_{i,j-1} \quad \text{és} \quad D^2\left[X_{i,j}\right] = \phi\lambda_j\left(\lambda_j - 1\right)C_{i,j-1},$$

ahol λ_j a lánc-létra módszernél bemutatott növekedési faktor. A ϕ paraméter most is ismeretlen. Nyilvánvaló a $C_{i,j} = C_{i,j-1} + X_{i,j}$ összefüggés, és feltesszük, hogy ebben a rekurzív megközelítésben a $C_{i,j-1}$ ismert. Tehát látszik, hogy ez a modell felírható kummulált kárnagyságok segítségével is. Ekkor a $C_{i,j}$-k eloszlása szintén túlszórt negatív binomiális lesz, az alábbi várható értékkel és szórásnégyzettel:

$$E\left[C_{i,j}\right] = \lambda_j C_{i,j-1} \quad \text{és} \quad D^2\left[C_{i,j}\right] = \phi\lambda_j\left(\lambda_j - 1\right)C_{i,j-1}.$$

Ez a modell nem alkalmazható negatív inkrementális értékek esetén. Ilyen esetekben a negatív binomiális modell normális eloszlással való közelítése használható. Erről a módszerről a [3] tanulmányban olvashatunk bővebben.

2.3 MACK MODELLJE

A Mack által konstruált modell az egyik legkorábbi sztochasztikus modell, alkalmazásával ugyanazokat a becsléseket kapjuk, mint a lánc-létra módszerrel. Alkalmazásának előnye, hogy a károk eloszlására vonatkozólag kevesebb feltételezéssel él, mivel csak az első két momentumot használja. A kummulált károkra a következők teljesülnek:

$$E\left[C_{i,j}\right] = \lambda_j C_{i,j-1} \quad \text{és} \quad D^2\left[C_{i,j}\right] = \sigma_j^2 C_{i,j-1}.$$

Az ismeretlen λ_j és σ_j^2 paraméterekre vonatkozó becslések:

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} w_{i,j} f_{i,j}}{\sum_{i=1}^{n-j+1} w_{i,j}}, \quad \text{ahol } w_{i,j} = C_{i,j-1} \text{ és } f_{i,j} = \frac{C_{i,j}}{C_{i,j-1}},$$

illetve

$$\hat{\sigma}_j^2 = \frac{1}{n-j} \sum_{i=1}^{n-j+1} w_{i,j} (f_{i,j} - \hat{\lambda}_j)^2.$$

Látható, hogy a növekedési faktorra vonatkozó becslés ugyanaz, mint a standard súlyozású lánc-létra módszer esetében. Mack bebizonyította, hogy ez torzítatlan becslést ad a növekedési faktorokra. Valójában az évenkénti növekedési faktorok súlyozatlan átlaga is torzítatlan becslést ad, azonban érdemesebb a súlyozott átlagukat használni, mivel annak kisebb a varianciája. Sőt, Mack azt is belátta, hogy pontosan a súlyozott átlag adja a minimum varianciájú torzítatlan becslést.

A továbbiakban *reziduális* alatt a kummulált károk tényleges növekedési aránya - tehát az $f_{i,j}$ – és a becsült növekedési faktorok különbségét értjük. A σ_j^2 variancia komponenst a súlyozott reziduálisok átlagával becsüljük, de úgy, hogy leosztjuk a reziduálisok száma mínusz eggyel. Így torzítatlan becslést kapunk erre a komponensre. A variancia komponenst ugyan nem használjuk a növekedési faktorok becslésekor, de szükséges a jövőbeli kifizetések előrejelzési hibájának kiszámításához.

Ez a modell eloszlásmentesnek tekinthető, mivel a kárnagyságok teljes eloszlásáról nem mondunk semmit. Ez leegyszerűsíti a vizsgálatokat, mivel csak az első két momentummal kell foglalkozni. További feltételezéseket akkor szükséges tenni, ha az eredményeket dinamikus pénzügyi elemzéshez használjuk fel, mivel ekkor a szükséges tartalék eloszlását is szimulálni kell.

3. FEJEZET

ELŐREJELZÉSEK ÉS HIBÁIK

Az eddigiek alapján azt mondhatjuk, hogy a kártartalékok képzése egy előrejelzési folyamat: rendelkezésünkre állnak az adatok, amelyek alapján megpróbáljuk előre jelezni a jövőbeli kárkifizetések nagyságát. Az előző fejezetben különböző modellek segítségével próbáltunk előrejelzéseket adni. A bevezetőben említettük, hogy az előrejelzés hibája nagy jelentőségű, ezért nem hagyhatjuk figyelmen kívül. Ebben a fejezetben az eddig bemutatott modellek hibáival fogunk foglalkozni.

Legyen y egy véletlen változó, melynek várható értékét jelölje $\hat{y}$. A *becslés átlagos négyzetes hibáját* (*MSEP - mean squared error of prediciton*) a következőképpen definiáljuk:

$$E[(y-\hat{y})^2] = E[((y-E[y]) - (\hat{y} - E[y]))^2]$$

Vizsgáljuk meg ezt az egyenletet. Az utolsó várható értékes tagban helyettesítsünk y helyére $\hat{y}$-ot, így az alábbi közelítést kapjuk:

$$E[(y-\hat{y})^2] \approx E[(y-E[y])^2] - 2E[(y-E[y])(\hat{y}-E[\hat{y}])] + E[(\hat{y}-E[\hat{y}])^2].$$

Tegyük fel, hogy a jövőbeli megfigyelések függetlenek a korábbi megfigyelésektől, így:

$$E[(y-\hat{y})^2] \approx E[(y-E[y])^2] + E[(\hat{y}-E[\hat{y}])^2].$$

Ez pontosan azt jelenti, hogy az előrejelzés szórásnégyzete felírható az eredeti y változó szórásnégyzetének és a becslés szórásnégyzetének összegeként.

Fontos megjegyezni, hogy az előrejelzés hibája nem egyezik meg a standard hibával. Az utóbbi pontosan a becslés szórásnégyzetének négyzetgyöke, az előrejelzés hibája pedig figyelembe veszi az előrejelzés változékonyságát, a paraméterek becsléséből esetlegesen adódó eltéréseket, illetve az adatok szórását is.

3.1 AZ ODP LÁNC-LÉTRA MODELL

Az ODP lánc-létra modellt általánosan az alábbi tulajdonságokkal írhatjuk fel:

$$E\left[X_{i,j}\right] = m_{i,j} \text{ és } D^2\left[X_{i,j}\right] = \phi m_{i,j},$$

ahol az $m_{i,j}$ paraméterre kétféle struktúrát javaslunk:

$$m_{i,j} = x_i y_j \tag{3.1}$$

vagy

$$\log\left(m_{i,j}\right) = c + \alpha_i + \beta_j. \tag{3.2}$$

Ha a (3.1) alakot használjuk a modell felírásához, akkor egy paramétereiben nem-lineáris modellt kapunk, ekkor pedig szükséges a paraméterbecsléseket kiszámolnunk. Erre a legegyszerűbb módszer a maximum likelihood becslés lenne, azonban ez hosszadalmas számolást igényel. A (3.2) struktúra egy általánosított lineáris modellt (GLM) határoz meg, amelyben az $X_{i,j}$ kárnagyságok Poisson-eloszlású véletlen változók, a $\eta_{i,j} = c + \alpha_i + \beta_j$ logaritmikus linkfüggvénnyel. Itt a „túlszórtság" akkor kap szerepet, amikor a ϕ ismeretlen skálaparamétert szeretnénk becsülni. A logaritmikus linkfüggvény alkalmazásával a modellünk lineáris lesz a paramétereiben. A paraméterbecsléseket a gyakorlatban GLM-illesztő statisztikai programcsomagok segítségével kaphatjuk meg. Ezek a programcsomagok is maximum likelihood becsléseket használnak.

A jövőbeli kárnagyságok becslését ezután úgy kaphatjuk meg, ha a paraméterbecsléseket visszahelyettesítjük a (3.2) egyenletbe, és alkalmazzuk rá az exponenciális függvényt:

$$\hat{X}_{i,j} = \hat{m}_{i,j} = \exp\left(\hat{\eta}_{i,j}\right). \tag{3.3}$$

A teljes tartalékra vonatkozó becslést értelemszerűen az $\hat{X}_{i,j}$-k soronkénti és oszloponkénti összegzéséből kapjuk meg.

Kíváncsiak vagyunk az előrejelzés hibájára is. Először vizsgáljuk meg egyetlen inkrementális $X_{i,j}$ kárnagyság előrejelzésének hibáját:

$$MSEP[\hat{X}_{i,j}] = E\left[\left(X_{i,j} - \hat{X}_{i,j}\right)^2\right] \approx D^2[X_{i,j}] + D^2[\hat{X}_{i,j}].$$

A modell általános felírásánál láttuk:

$$D^2[X_{i,j}] = \phi m_{i,j}. \tag{3.4}$$

Ezt összevetve a (3.3) egyenlettel, és Taylor-sorba fejtést alkalmazva [8] alapján a következőt kapjuk:

$$D^2[\hat{X}_{i,j}] \approx \left|\frac{\partial m_{i,j}}{\partial \eta_{i,j}}\right|^2 D^2[\hat{\eta}_{i,j}]. \tag{3.5}$$

A (3.4) és (3.5) egyenleteket összevetve adódik a közelítés a kárnagyság előrejelzésének átlagos négyzetes hibájára:

$$MSEP[\hat{X}_{i,j}] \approx \phi \hat{m}_{i,j} + \hat{m}_{i,j}^2 D^2[\hat{\eta}_{i,j}]. \tag{3.6}$$

A (3.6) egyenlet utolsó komponense a logaritmikus linkfüggvény szórásnégyzete, ezt általában közvetlenül megadják a statisztikai szoftverek, és általa az előrejelzés hibája egyszerűen számolható.

A kárévekre vonatkozó tartalékbecslések előrejelzési hibája és a teljes tartalék becslése is kiszámolható, de ez jóval több erőfeszítést igényel. Ekkor az előrejelzett értékek összegének szórásnégyzetét kell tekinteni, így az értékek közötti kovarianciákat is ki kell számolni. [8] alapján csak a becslés szórásnégyzetét szükséges megvizsgálni.

Vezessük be a Δ jelölést a becsült károkat tartalmazó kifutási háromszögre. Így az i. kárévre vonatkozólag a tartalék becslése:

$$\hat{X}_{i,+} = \sum_{j \in \Delta_i} \hat{X}_{i,j}.$$

[8] alapján adott kárév becsült tartalékának négyzetes előrejelzési hibája:

$$MSEP[\hat{X}_{i,+}] \approx \sum_{j \in \Delta_i} \phi \hat{m}_{i,j} + \sum_{j \in \Delta_i} \hat{m}_{i,j}^2 D^2[\hat{\eta}_{i,j}] + 2 \sum_{\substack{j_1, j_2 \in \Delta_i \\ j_2 > j_1}} \hat{m}_{i,j_1} \hat{m}_{i,j_2} Cov[\hat{\eta}_{i,j_1}, \hat{\eta}_{i,j_2}],$$

a teljes tartalék becslése:

$$\hat{X}_{+,+} = \sum_{i,j\epsilon\Delta} \hat{X}_{i,j},$$

a teljes tartalék négyzetes előrejelzési hibája pedig:

$$MSEP\left[\hat{X}_{+,+}\right] \approx \sum_{i,j\in\Delta} \phi\widehat{m}_{i,j} + \sum_{i,j\in\Delta} \widehat{m}_{i,j}^2 D^2\left[\hat{\eta}_{i,j}\right] + 2 \sum_{\substack{i_1 j_1\in\Delta \\ i_2 j_2\in\Delta \\ i_1 j_1\neq i_2 j_2}} \widehat{m}_{i_1 j_1} \widehat{m}_{i_2 j_2} Cov\left[\hat{\eta}_{i_1 j_1}, \hat{\eta}_{i_2 j_2}\right].$$

3.2 Rekurzív modellek hibája

Mielőtt a 2.2 fejezetben bemutatott negatív binomiális modell folyamat és előrejelzési hibáját kiszámítjuk, vizsgáljuk meg általánosan a rekurzív modellek esetét. Kezdjük a folyamat hibájával. Először a kétlépéses előrejelzést vizsgáljuk, majd a háromlépésest, végül rekurzívan megkapjuk a k-lépéses előrejelzésre vonatkozó képletet. Az egyszerűség kedvéért csak a kifutási háromszög $i.$ sorát tekintjük. A várható értékekre és a szórásnégyzetekre vagyunk kíváncsiak.
Kétlépéses előrejelzés:

$$\begin{aligned}
E\left[C_{j+1}|X_1, X_2, \dots, X_{j-1}\right] &= E\left[E\left[C_{j+1}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right] \\
&= E\left[\lambda_j C_j|X_1, X_2, \dots, X_{j-1}\right] \\
&= \lambda_{j+1}\lambda_j C_{j-1} \\
D^2\left[C_{j+1}|X_1, X_2, \dots, X_{j-1}\right] &= E\left[D^2\left[C_{j+1}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right] \\
&\quad + D^2\left[E\left[C_{j+1}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right] \\
= E\left[D^2\left[C_{j+1}|X_1, X_2, \dots, X_j\right]|X_1, X_2, \dots, X_{j-1}\right] &+ D^2\left[\lambda_{j+1} C_j|X_1, X_2, \dots, X_{j-1}\right] \\
= E\left[D^2\left[C_{j+1}|X_1, X_2, \dots, X_j\right]|X_1, X_2, \dots, X_{j-1}\right] &+ \lambda_{j+1} D^2\left[C_j|X_1, X_2, \dots, X_{j-1}\right].
\end{aligned}$$

Az utolsó kifejezés már csak az egy lépéssel későbbi szórásnégyzetet használja, így bevezethetjük a rekurziót. Ebből adódik a háromlépéses előrejelzésre:

$$E\left[C_{j+2}|X_1, X_2, \dots, X_{j-1}\right] = E\left[E\left[C_{j+2}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right]$$

$$= E\left[\lambda_{j+2}\lambda_{j+1}C_j|X_1, X_2, \dots, X_{j-1}\right]$$
$$= \lambda_{j+2}\lambda_{j+1}\lambda_j C_{j-1}$$

$$D^2\left[C_{j+2}|X_1, X_2, \dots, X_{j-1}\right] = E\left[D^2\left[C_{j+2}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right]$$
$$+D^2\left[E\left[C_{j+2}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right]$$
$$= E\left[D^2\left[C_{j+1}|X_1, X_2, \dots, X_j\right]|X_1, X_2, \dots, X_{j-1}\right] + \lambda_{j+2}^2\lambda_{j+1}^2 D^2\left[C_j|X_1, X_2, \dots, X_{j-1}\right].$$

Itt már csak az egy- illetve két lépéssel későbbi szórásnégyzetet használjuk. Hasonlóan írható fel a formula a k-lépéses esetre:

$$E\left[C_{j+k-1}|X_1, X_2, \dots, X_{j-1}\right] = E\left[E\left[C_{j+k-1}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right]$$
$$= E\left[\lambda_{j+k-1}\lambda_{j+k-2} \dots \lambda_{j+1}C_j|X_1, X_2, \dots, X_{j-1}\right]$$
$$= \lambda_{j+k-1}\lambda_{j+k-2} \dots \lambda_{j+1}C_{j-1}$$

$$D^2\left[C_{j+k-1}|X_1, X_2, \dots, X_{j-1}\right] = E\left[D^2\left[C_{j+k-1}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right]$$
$$+D^2\left[E\left[C_{j+k-1}|X_1, X_2, \dots, X_{j-1}, X_j\right]|X_1, X_2, \dots, X_{j-1}\right]$$
$$= E\left[D^2\left[C_{j+k-2}|X_1, X_2, \dots, X_j\right]|X_1, X_2, \dots, X_{j-1}\right]$$
$$+ \lambda_{j+k-1}^2\lambda_{j+k-2}^2 \dots \lambda_{j+1}^2 D^2\left[C_j|X_1, X_2, \dots, X_{j-1}\right].$$

A következőkben a rekurzív modellek becslési hibáját szeretnénk megállapítani. Az egyszerűség kedvéért ismét csak a kifutási háromszög i. sorát tekintjük. A $D^2\left[\hat{C}_n|C_{n-i+1}\right]$ kifejezést szeretnénk meghatározni, ami az adott kárévre vett összes becsült kummulált kár nagyságának az utolsó ismert kummulált kárra vett feltételes szórásnégyzete. Ez természetesen ugyanaz, mintha az inkrementális becsült értékek összegének feltételes szórásnégyzetét néznénk, mindkét esetben ugyanazt kapjuk a becslési hibára:

$$D^2\left[\hat{C}_n|C_{n-i+1}\right] = D^2\left[\hat{\lambda}_{n-i+2} \dots \hat{\lambda}_n C_{n-i+1}|C_{n-i+1}\right] = C_{n-i+1}^2 D^2\left[\hat{\lambda}_{n-i+2} \dots \hat{\lambda}_n|C_{n-i+1}\right].$$

Tehát a becsült növekedési faktorok szorzatának szórásnégyzetére vagyunk kíváncsiak, amit rekurzívan megkaphatunk úgy, hogy az utolsó ismert kummulált károk négyzetével megszorozzuk a becsült növekedési faktorok szorzatának szórásnégyzetét. A második sorra a becslés szórásnégyzete egyszerűen $C_{2,n-1}^2 D^2[\hat{\lambda}_n]$. A harmadik sorra pedig $C_{3,n-2}^2 D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n]$, ahol

$$D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n] = \left(E[\hat{\lambda}_{n-1}]\right)^2 D^2[\hat{\lambda}_n] + \left(E[\hat{\lambda}_n]\right)^2 D^2[\hat{\lambda}_{n-1}] + D^2[\hat{\lambda}_n] D^2[\hat{\lambda}_{n-1}].$$

Itt felhasználtuk azt a feltételezést, hogy a növekedési faktorok becsült értékei egymástól függetlenek, vagy legalább korrelálatlanok. A gyakorlatban a korrelálatlanság könnyen megmutatható a növekedési faktorok kovariancia-mátrixának segítségével, mivel a főátló kivételével minden elem nulla lesz. A modell illesztésekor megkapjuk a növekedési faktorok becsléseinek szórásnégyzetét, és a megfigyelt értékekkel helyettesítjük a fenti várható értékeket:

$$D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n] \approx \left(\hat{\lambda}_{n-1}\right)^2 D^2[\hat{\lambda}_n] + \left(\hat{\lambda}_n\right)^2 D^2[\hat{\lambda}_{n-1}] + D^2[\hat{\lambda}_n] D^2[\hat{\lambda}_{n-1}], \quad (3.7)$$

a következő sorra pedig felírható:

$$D^2[\hat{\lambda}_{n-2}(\hat{\lambda}_{n-1}\hat{\lambda}_n)] = \left(E[\hat{\lambda}_{n-2}]\right)^2 D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n] + \left(E[\hat{\lambda}_{n-1}\hat{\lambda}_n]\right)^2 D^2[\hat{\lambda}_{n-2}] + D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n] D^2[\hat{\lambda}_{n-2}]. \quad (3.8)$$

Az (3.7)-ben kiszámolt $D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n]$ értéket behelyettesítjük a (3.8) egyenletbe, és így rekurzívan megoldható a feladat. A rekurzió minden egyes lépésében egy plusz növekedési faktor lép be. A teljes tartalékra vonatkozó becslési variancia [8] alapján a következő:

$$D^2[\hat{R}_+] \approx \sum_{i=2}^{n} D^2[\hat{C}_{i,n}] + 2\sum_{\substack{i=2 \\ j>i}}^{n} Cov[\hat{C}_{i,n}, \hat{C}_{j,n}],$$

ahol $\hat{R}_+$ jelölje a teljes becsült tartalékot. Tehát ki kell számolni még a kovarianciás kifejezéseket. Az $i = 2$ esetben:

$$Cov[\hat{C}_{2,n}, \hat{C}_{3,n}] = Cov[C_{2,n-1}\hat{\lambda}_n, C_{3,n-2}\hat{\lambda}_{n-1}\hat{\lambda}_n] = C_{2,n-1}C_{3,n-2}\hat{\lambda}_{n-1}D^2[\hat{\lambda}_n],$$

ha a növekedési faktorok függetlenek. Hasonlóan:

$$Cov[\hat{C}_{2,n}, \hat{C}_{4,n}] = Cov[C_{2,n-1}\hat{\lambda}_n, C_{4,n-3}\hat{\lambda}_{n-2}\hat{\lambda}_{n-1}\hat{\lambda}_n] = C_{2,n-1}C_{4,n-3}\hat{\lambda}_{n-2}\hat{\lambda}_{n-1}D^2[\hat{\lambda}_n],$$

és így tovább ki kell számolni minden kettőnél nagyobb indexű sor kovarianciáját a második sorral.

Legyen $i = 3$, ekkor:

$$\begin{aligned} Cov[\hat{C}_{3,n}, \hat{C}_{4,n}] &= Cov[C_{3,n-2}\hat{\lambda}_{n-1}\hat{\lambda}_n, C_{4,n-3}\hat{\lambda}_{n-2}\hat{\lambda}_{n-1}\hat{\lambda}_n] \\ &= C_{3,n-2}C_{4,n-3}\hat{\lambda}_{n-2}D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n]. \end{aligned}$$

A $D^2[\hat{\lambda}_{n-1}\hat{\lambda}_n]$ kifejezés helyére beírjuk a (3.7) kifejezést, és hasonló módon kiszámoljuk minden háromnál nagyobb indexű sor kovarianciáját a harmadik sorral. Ezt az eljárást folytatjuk a $n-1$. sorig, ahol:

$$\begin{aligned} Cov[\hat{C}_{n-1,n}, \hat{C}_{n,n}] &= Cov[C_{n-1,2}\hat{\lambda}_3\hat{\lambda}_4 \dots \hat{\lambda}_n, C_{n,1}\hat{\lambda}_2\hat{\lambda}_3 \dots \hat{\lambda}_n] \\ &= C_{n-1,2}C_{n,1}\hat{\lambda}_2 D^2[\hat{\lambda}_3\hat{\lambda}_4 \dots \hat{\lambda}_n] \end{aligned}$$

és a növekedési faktorok szorzatának szórásnégyzetébe mindig behelyettesítjük az előzőleg kapott értéket.

3.3 A NEGATÍV BINOMIÁLIS LÁNC-LÉTRA MODELL

Korábban láttuk, hogy a negatív binomiális modell illesztése során inkrementális, illetve kummulált adatokat is használhatunk. Mindkét módszerrel ugyanazokat az illesztett értéket kapjuk, mivel a negatív binomiális modell csak a pénzügyi éveket használja paraméterként a számolások során – ellenben az ODP modellben a kárbekövetkezési és a kifizetési évek is paraméterként szerepelnek. A negatív binomiális modell inkrementális adatokra felírva a következő tulajdonságokkal rendelkezik:

$$E[X_{i,j}] = (\lambda_j - 1)C_{i,j-1} \quad \text{és} \quad D^2[X_{i,j}] = \phi\lambda_j(\lambda_j - 1)C_{i,j-1},$$

ahol a $C_{i,j}$ értékeket ismertnek tekintjük. A fenti várható értéket jelölje $m_{i,j}$. Ekkor:

$$\log(m_{i,j}) = \log(\lambda_j - 1) + \log(C_{i,j-1}), \tag{3.7}$$

továbbá felírható a következő:

$$\log(\lambda_j - 1) = c + \alpha_{j-1}, \quad \text{ahol } \alpha_1 = 0, j \geq 2. \tag{3.8}$$

A (3.8) egyenletet behelyettesítve a (3.7)-be a következő összefüggést kapjuk:

$$\log(m_{i,j}) = c + \alpha_{j-1} + \log(C_{i,j-1}).$$

Erre az átalakításra azért volt szükség, mert így egy általánosított lineáris modellt kapunk logaritmikus linkfüggvénnyel és negatív binomiális hibastruktúrával. Akárcsak az ODP modell esetében, most is megkaphatjuk a paraméterek maximum likelihood becslését statisztikai programcsomagok segítségével. A növekedési faktorok a (3.8) egyenlet alapján becsülhetőek, közelítő standard hibájuk pedig a következő egyenletből kapható meg:

$$D^2[\hat{\lambda}_j] = D^2[\hat{\lambda}_j - 1] \approx \exp(\hat{c} + \hat{\alpha}_{j-1})^2 D^2[\hat{c} + \hat{\alpha}_{j-1}], \qquad j \geq 2.$$

Most nézzük meg a folyamat varianciájára és a becslés varianciájára adódó becsléséket. Jelölje U_i az i. kárbekövetkezési évre vonatakozó „ultimate” károkat, azaz:

$$U_i = C_{i,n.}$$

Az i. kárbekövetkezési évre vonatakozó tartalék becslése:

$$R_i = U_i - C_{i,n-i+1},$$

ahol $C_{i,n-i+1}$ ismert. Így felírhatóak a következők:

$$D^2[R_i] = D^2[U_i] = D^2[C_{i,n}],$$

illetve

$$D^2[\hat{R}_i] = D^2[\hat{U}_i] = D^2[\hat{C}_{i,n}].$$

Egy adott kárbekövetkezési évre vonatkozó folyamat és becslés varianciája megbecsülhető a $D^2[C_{i,n}]$ és $D^2[\hat{C}_{i,n}]$ segítségével. Először nézzük meg a folyamat varianciáját. A negatív binomiális modell bevezetésénél láttuk, hogy az egylépéses előrejelzés szórásnégyzete a következőképpen adódik:

$$D^2[C_j | X_1, X_2, \dots, X_{j-1}] = \phi\lambda_j(\lambda_j - 1)C_{j-1}.$$

Az előzőekben bemutatott általános eset alapján a kétlépéses előrejelzés szórásnégyzete:

$$\begin{aligned} D^2[C_{j+1} | X_1, X_2, \dots, X_{j-1}] &= E[\phi\lambda_{j+1}(\lambda_{j+1} - 1)C_j | X_1, X_2, \dots, X_{j-1}] + \lambda_{j+1}^2\phi\lambda_j(\lambda_j - 1)C_{j-1} \\ &= \phi\lambda_{j+1}(\lambda_{j+1} - 1)\lambda_j C_{j-1} + \lambda_{j+1}^2\phi\lambda_j(\lambda_j - 1)C_{j-1} \\ &= \phi\lambda_j\lambda_{j+1}(\lambda_j\lambda_{j+1} - 1)C_{j-1}, \end{aligned}$$

a háromlépéses előrejelzés szórásnégyzete:

$$\begin{aligned} D^2[C_{j+2} | X_1, X_2, \dots, X_{j-1}] &= E[\phi\lambda_{j+1}\lambda_{j+2}(\lambda_{j+1}\lambda_{j+2} - 1)C_j] + \lambda_{j+2}^2\lambda_{j+1}^2\phi\lambda_j(\lambda_j - 1)C_{j-1} \\ &= \phi\lambda_{j+1}\lambda_{j+2}(\lambda_{j+1}\lambda_{j+2} - 1)\lambda_j C_{j-1} + \lambda_{j+2}^2\lambda_{j+1}^2\phi\lambda_j(\lambda_j - 1)C_{j-1} \\ &= \phi\lambda_j\lambda_{j+1}\lambda_{j+2}(\lambda_j\lambda_{j+1}\lambda_{j+2} - 1)C_{j-1}. \end{aligned}$$

Ezek alapján a k-lépéses előrejelzés szórásnégyzete:

$$D^2[C_{j+k-1} | X_1, X_2, \dots, X_{j-1}] = \phi\lambda_j\lambda_{j+1} \dots \lambda_{j+k-1}(\lambda_j\lambda_{j+1} \dots \lambda_{j+k-1} - 1)C_{j-1}.$$

A becslés varianciájára a következő adódik:

$$D^2[\hat{C}_{i,n}] \approx D^2\left[C_{i,n-i+1} \prod_{k=n-i+2}^{n} \hat{\lambda}_k\right] = C_{i,n-i+1}^2 D^2\left[\prod_{k=n-i+2}^{n} \hat{\lambda}_k\right].$$

Ezután vizsgáljuk meg a teljes tartalékra vonatkozó folyamat és becslés varianciáját. Legyen $R_+ = \sum_{i=2}^{n} R_i$. A teljes tartalékra vonatkozó folyamat varianciája a különböző kárbekövetkezési évekre vonatkozó tartalékok folyamat varianciáinak összege, feltéve, hogy a különböző évekre vonatkozó tartalékok függetlenek egymástól. A teljes tartalékra vonatkozó becslés varianciája a 3.2 fejezet alapján:

$$D^2[\hat{R}_+] \approx \sum_{i=2}^{n} D^2[\hat{C}_{i,n}] + 2\sum_{\substack{i=2 \\ j>i}}^{n} Cov[\hat{C}_{i,n}, \hat{C}_{j,n}].$$

A [2] cikkben kapott eredmények alapján, ha összehasonlítjuk a negatív binomiális modell eredményeit az ODP modell alkalmazásával kapott eredményekkel, akkor megállapíthatjuk, hogy a tartalékra vonatkozó becslések a két módszerrel megegyeznek, illetve az előrejelzések hibája is nagyon közeli, a negatív binomiális modell esetében valamivel nagyobb. A folyamat varianciájának és a becslés varianciájának összege lényegileg megegyezik a két esetben. A becslés varianciája az ODP modell esetében lesz a nagyobb, ennek a modellre vonatkozó nagyobb paraméterszám az oka. Elmondható, hogy a két modell csak abban különbözik, hogy hogyan paraméterezzük őket. Tehát összességében mindegy, hogy melyik modellt illesztjük az adatokra, a kapott eredmények ugyanazok lesznek.

3.4 MACK MODELLJE

A 2.3 fejezetnek megfelelően ez a modell a kummulált kárnagyságokra vonatkozóan a következőket mondja:

$$E[C_{i,j}] = \lambda_j C_{i,j-1} \quad \text{és} \quad D^2[C_{i,j}] = \sigma_j^2 C_{i,j-1}.$$

Mack levezette a kárbekövetkezési évek tartalékára és a teljes tartalékra vonatkozó előrejelzési hibát, csak a kummulált kárnagyságok, a növekedési faktorok és a varianciakomponens felhasználásával (ld. [6]). A kárévenkénti tartalék folyamat varianciájára a következőt kapta:

$$D^2[R_i] \approx \hat{C}_{i,n}^2 \sum_{k=n-i+1}^{n-1} \frac{\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2 \hat{C}_{i,k}},$$

az előrejelzés varianciájára pedig:

$$D^2\left[\hat{R}_i\right] \approx \hat{C}_{i,n}^2 \sum_{k=n-i+1}^{n-1} \frac{\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2 \sum_{q=1}^{n-k} C_{q,k}},$$

Ezek összege az átlagos négyzetes előrejelzési hiba az $i.$ kárévre:

$$MSEP\left[\hat{R}_i\right] \approx \hat{C}_{i,n}^2 \sum_{k=n-i+1}^{n-1} \frac{\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2}\left(\frac{1}{\hat{C}_{i,k}} + \frac{1}{\sum_{q=1}^{n-k} C_{q,k}}\right).$$

Kárévenként összegezve pedig adódik a teljes tartalékra vonatkozó átlagos négyzetes előrejelzési hiba:

$$MSEP\left[\hat{R}_+\right] = \sum_{i=2}^{n} \left\{ MSEP\left[\hat{R}_i\right] + \hat{C}_{i,n}\left(\sum_{q=i+1}^{n} \hat{C}_{q,n}\right) \sum_{k=n-i+1}^{n-1} \frac{2\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2 \sum_{q=1}^{n-k} C_{q,k}} \right\},$$

ahol a korábbiakhoz hasonlóan

$$\hat{R}_+ = \sum_{i=2}^{n} \hat{R}_i.$$

Bár ez a formula ijesztően néz ki, a gyakorlatban könnyedén és gyorsan alkalmazható.

4. FEJEZET

A BOOTSTRAP ELJÁRÁS

Az előző fejezetek során az volt a célunk, hogy becslést kapjunk a best estimate-re és az előrejelzési hibára. Az előrejelzési hiba fontos mérőszáma a tartalékbecslések pontosságának, azonban kiszámításával csak a tartalék teljes prediktív eloszlásának második momentumáról kapunk információt. A célunk viszont az, hogy a teljes prediktív eloszlást ismerjük, hiszen ez alapján ki tudjuk számolni a szükséges kvantiliseket. Amíg nem ismerjük az adatok eloszlását, feltételezésekkel kell élnünk – például Mack feltette, hogy a tartalékok mértéke közelítőleg normális vagy lognormális eloszlás szerint alakul. A számítógépek fejlődésével és a szimulációs technikák népszerűvé válásával azonban ma már egy sokkal jobb módszer is alkalmazható a tartalékszámításban: a bootstrap eljárás.

A bootstrap egy egyszerűnek tűnő szimulációs eljárás, melynek célja hogy egy adathalmazból információkat nyerjünk ki. A gyakorlatban az eljárást aktuáriusi programcsomagok – például ResQ – segítségével futtathatjuk, azonban ezeknél a programoknál csak az input és output adatokat látjuk, azt viszont nem, hogy mi is történik a „fekete dobozban". Ennek a fejezetnek az a célja, hogy megmutassuk a bootstrap matematikai hátterét, illetve annak alkalmazhatóságát.

Az eljárás lényege, hogy a rendelkezésünkre álló megfigyelt adatok halmazába pszeudo-adatok helyettesítésével megfelelően nagy mennyiségű új mintát hozzunk létre, melyekben az adatok eloszlása megegyezik az eredeti adathalmaz eloszlásával. Ezután statisztikai módszerekkel ebből a nagyobb adathalmazból már több lényeges információt tudunk kiszűrni. Egy egyszerű példa, hogyha több pszeudo-adathalmazt hozunk létre, majd kiszámoljuk ezen adathalmazok várható értékét egyenként, illetve az így kapott értékek szórását, akkor becslést kapunk a standard hibára. England és Verrall a bootstrap eljárás segítségével adta meg a lánc-létra módszerrel becsült tartalékok becslési hibáját (ld. [9]). Később England kiterjesztette a módszert úgy, hogy bootstrap segítségével szimulálja már a folyamat hibáját is.

Az eljárás alkalmazása során feltesszük, hogy a vizsgált adatok függetlenek és azonos eloszlásúak, és az új pszeudo-minta létrehozásához magából az eredeti adathalmazból visszatevéses mintavétellel veszünk mintaelemeket. Megjegyezzük, hogy regressziós típusú feladatokban általában azt szokás feltenni, hogy az adatok függetlenek, de nem azonos eloszlásból származnak. Ezért az ilyen feladatokban a bootstrap eljárást a reziduálisokra hajtjuk végre az adatok helyett, mivel a reziduálisok általában függetlenek és azonos eloszlásúak, vagy egyszerű transzformációk segítségével azzá alakíthatóak. Tehát fontos, hogy legyen egy egységes reziduális definíciónk az ilyen esetekre. Az ODP lánc-létra modell esetében a Pearson-reziduálisokra fogjuk végrehajtani a bootstrap eljárást. A kárbekövetkezési és pénzügyi évre vonatkozó indexeket elhagyva a Pearson-reziduálisokat a következőképpen definiáljuk:

$$r_p = \frac{X - \widehat{m}}{\sqrt{\widehat{m}}}, \tag{4.1}$$

ahol $\widehat{m}$ az ODP lánc-létra módszernél illesztett inkrementális kárnagyság. Az eljárás során a reziduálisokból visszatevéses mintavétellel véletlenszerűen veszünk mintaelemeket, és új mintát hozunk létre, úgy, hogy erre az adathalmazra alkalmazzuk a (4.1) képletet. Az új r_p^* Person-reziduálisok és az illesztett $\widehat{m}$ kárnagyság segítségével felírható a bootstrappel kapott X^* inkrementális kárnagyság:

$$X^* = r_p^* \sqrt{\widehat{m}} + \widehat{m}.$$

4.1 AZ ALGORITMUS

Az alábbi lépések segítségével implementáljuk a bootstrap eljárás algoritmusát:

1. A kummulált adatokból kiszámoljuk a standard lánc-létra módszer növekedési faktorait.
2. Fordított rekurzióval kiszámoljuk a kummulált illesztett értékeket egy múltbeli háromszögre. Kiindulásként az utolsó megfigyelt kummulált értékeket használjuk, és legyen $\hat{C}_{i,n-i+1} = C_{i,n-i+1}$, illetve $\hat{C}_{i,k-1} = \hat{C}_{i,k} \lambda_k^{-1}$.
3. A kummulált illesztett értékekből kiszámoljuk ezen múltbeli háromszög illesztett $\widehat{m}_{i,j}$ inkrementális értékeit.
4. Kiszámítjuk a „skálázatlan” Pearson-reziduálisokat a múltbeli háromszögre:

$$r_{i,j}^{(P)} = \frac{X_{i,j} - \widehat{m}_{i,j}}{\sqrt{\widehat{m}_{i,j}}} \tag{4.2}$$

5. A ϕ Pearson-skálaparaméter a Pearson-reziudálisok négyzetösszege leosztva a szabadságfokkal, ahol a szabadságfok a megfigyelések száma mínusz a becsült paraméterek száma, azaz:

$$\phi = \frac{\sum_{i,jn-i+1}\left(r_{i,j}^{(P)}\right)^2}{\frac{1}{2}n(n+1) - 2n + 1}.$$

6. Kiszámoljuk a módosított reziduálisokat, így kiküszöbölhetjük az esetleges torzításokat:

$$r_{i,j}^{adj} = \sqrt{\frac{n}{\frac{1}{2}n(n+1) - 2n + 1}}\, r_{i,j}^{(P)}.$$

7. Elég nagyszámú N-re ismételjük az alább iterációt (például $N = 10000$):
   a) létrehozunk egy új „múltbeli” kifutási háromszöget a reziduálisokból, a módosított reziduálisokból való visszatevéses mintavétellel;
   b) ezen háromszög minden elemére alkalmazva (4.2)-t létrehozzuk a pszeudo-inkrementális adathalmazt;
   c) létrehozzuk a megfelelő pszeudo-kummulált adathalmazt;
   d) ezekre a pszeudo-kummulált adatokra illesztjük a standard lánc-létra modellt;
   e) majd létrehozzuk a jövőbeli kummulált kifizetésekre vonatkozó kifutási háromszöget;
   f) ebből differenciálással megkapjuk a jövőbeli inkrementális kifizetésekre vonatkozó kifutási háromszöget – ennek értékeit fogjuk várható értékként használni az eloszlásból való szimulálások során;
   g) ezen háromszög minden elemére szimuláljuk a kifizetést az eloszlásból („*process distribution*”) – ami ez esetben az ODP modell, - feltéve, hogy az eloszlás várható értéke az előző pontban megkapott $\widetilde{m}_{i,j}$ érték, szórásnégyzete pedig $\phi\widetilde{m}_{i,j}$;

h) kárbekövetkezési évenként összegezve ezeket a szimulált kifizetéseket megkapjuk az adott kárbekövetkezési évre vonatkozó tartalék becslését, a szimulált értékek teljes összege pedig a teljes tartalék becslését adja;

i) mentsük el az eredményeket, majd kezdjük elölről az iterációt.

Ezek az elmentett eredmények határozzák meg a prediktív eloszlást. A módszer megfelelőségét például tesztheljük úgy, hogy az elmentett értékek átlagát összehasonlítjuk a standard lánc-létra módszerből adódó tartalékbecslésekkel. Az elmentett eredmények szórása becslést ad az előrejelzés hibájára.

A bootstrap eljárást egyszerűen és gyorsan implementálhatjuk az ODP lánc-létra modellre, mivel ekkor a sima lánc-létra módszert alkalmazhatjuk minden egyes iterációban. További jól definiált modellekre is alkalmazható a bootstrap, azonban általában sokkal bonyolultabb az eljárást implementálni. Például, ha a Mack modellre szeretnénk végrehajtani az eljárást, akkor a skálázott Pearson-reziduálisokkal kell dolgoznunk. Ezeket a következőképpen definiáljuk:

$$r_{i,j} = \frac{\sqrt{w_{i,j}}(f_{i,j} - \lambda_j)}{\sigma_j}$$

Ebben az esetben lánc-létra módszerrel újrabecsüljük a növekedési faktorokat, és ezekkel a szimulált növekedési faktorokkal szorozzuk az ismert kummulált károkat, így kapjuk meg a következő periódusra vonatkozó kummulált károk becsült értékét. A szimuláció során ezt az eljárást kell ismételni.

4.2 SCHNIEPER MODELL

A Schnieper-modell célja, hogy jobb becslést adjon a fennálló kötelezettségek mértékére azokban az esetekben, amikor a rendelkezésre álló káradataink eléggé volatilisek. A problémát a modell úgy közelíti meg, hogy különválasztja egymástól az *IBNR* és *IBNER* károkat, és mindegyikre külön képez tartalékot. Emlékeztetőül, IBNR (*„Incurred But Not Reported”*) károknak neveztük az olyan károkat, amelyek már bekövetkeztek, de még nem jelentették be őket a biztosítónak (például a mérlegforduló napjáig). Az IBNER károk (*„Incurred But Not Enough Reserved”*) pedig azok, amelyek már bekövetkeztek, be is jelentették őket, de még nem áll rendelkezésünkre róluk elég információ, így előfordulhat, hogy nem képeztünk a fedezetükre elégséges tartalékot. Sok ilyen típusú kár negatív bonyolítási eredményhez vezethet. Ezen két kártípus szétválasztásának az az előnye, hogy könnyebb a két kárkifutási háromszöget külön kezelni, illetve így pontosabb becsléseket kapunk a tartalék nagyságára vonatkozóan. Az előzőekben tárgyalt bootstrap eljárással ellentétben ez egy rekurzív módszer a tartalékképzésre.

A modell alapfeltevése, hogy a két kifutási háromszögben az adatok egymástól függetlenek. Az eddigiekhez hasonlóan az adataink háromszög alakba rendezettek, ahol az $i.$ sorindex a kárbekövetkezési évre utal, a $j.$ oszlopindex pedig az adott kárbekövetkezési évtől számított j. pénzügyi évet jelenti. A háromszögekben a kummulált bekövetkezett („*incurred*") károk nagyságai állnak rendelkezésünkre. Vezessük be a következő jelöléseket: $-C_{i,j}$ a „régi" károk, azaz a kummulált felmerült és bejelentett károk változása a korábbi periódusokban (tehát az IBNER károk), $N_{i,j}$ pedig az „új" károkat jelöli, azaz a $j.$ pénzügyi évben bejelentett károk nagysága. Feltesszük, hogy az inkrementális bekövetkezett károk a „régi" és az „új" inkrementális bekövetkezett károk összege, azaz:

$$X_{i,j} - X_{i,j-1} = C_{i,j} + N_{i,j},$$

a kummulált károkra felírva:

$$X_{i,j} = X_{i,j-1} + C_{i,j} + N_{i,j}.$$

Schnieper azt is felteszi, hogy minden kárbekövetkezési évre ismert az E_i kitettség nagysága. Kitettség alatt a feladattól függően érthetünk biztosítási összeget, állomány darabszámot, megszolgált díjat, stb. Az adott kárévtől számított $n.$ pénzügyi évhez tartozó kummulált károkra a „ultimate" károkként fogunk hivatkozni, az $n.$ év utáni előrejelzésekkel nem foglalkozunk. H_k-val jelöljük a $k.$ pénzügyi évig rendelkezésünkre álló információk összességét, F_k-val pedig az adott kárévtől számított $k.$ periódusig ismert információkat, vagyis:

$$H_k = \{N_{i,j}, C_{i,j} : 1 \leq i,j \leq n; i + j - 1 \leq k\},$$
$$F_k = \{N_{i,j}, C_{i,j} : 1 \leq i,j \leq n; j \leq k\}.$$

A következő feltételeket alkalmazzuk az általános modellre:

1. feltétel: a λ_j és δ_j konstansok léteznek, úgy, hogy E_i nagysága ismert és a következők teljesülnek:

$$E\left[N_{i,j}|H_{i+j-2}\right] = E_i\lambda_j,\ \ 1 \leq i,j \leq n,$$
$$E\left[C_{i,j}|H_{i+j-2}\right] = X_{i,j-1}\delta_j,\ \ 1 \leq i \leq n, 2 \leq j \leq n.$$

2. feltétel: a σ_j^2és τ_j^2 konstansok léteznek és a következők teljesülnek:

$$D^2\left[N_{i,j}|H_{i+j-2}\right] = E_i\sigma_j^2,\qquad 1 \leq i,j \leq n,$$
$$D^2\left[C_{i,j}|H_{i+j-2}\right] = X_{i,j-1}\tau_j^2,\ \ 1 \leq i \leq n, 2 \leq j \leq n.$$

3. feltétel: a kárévek egymástól függetlenek, azaz $\left\{N_{1,j}, C_{1,j} : 1 \leq j \leq n\right\}, \dots, \left\{N_{n,j}, C_{n,j} : 1 \leq j \leq n\right\}$ függetlenek.
4. feltétel: a kárévek között korrelálatlanság áll fenn, azaz $\left\{N_{i,j}|H_{i+j-2} : 1 \leq j \leq n\right\}$ és $\left\{C_{i,j}|H_{i+j-2} : 1 \leq j \leq n, 2 \leq j \leq n\right\}$ korrelálatlanok.

Ezen feltételek mellett a paraméterek becslésére [5] alapján a következők adódnak:

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} N_{i,j}}{\sum_{i=1}^{n-j+1} E_i},\ \ 1 \leq j \leq n,$$

$$\hat{\delta}_j = \frac{\sum_{i=1}^{n-j+1} C_{i,j}}{\sum_{i=1}^{n-j+1} X_{i,j-1}},\ \ 2 \leq j \leq n,$$

$$\hat{\sigma}_j^2 = \frac{1}{n-j}\sum_{i=1}^{n-j+1}\frac{1}{E_i}\left(N_{i,j} - \hat{\lambda}_j E_i\right)^2,\qquad 1 \leq j \leq n-1,$$

$$\hat{\tau}_j^2 = \frac{1}{n-j}\sum_{i=1}^{n-j+1}\frac{1}{X_{i,j-1}}\left(C_{i,j} - \hat{\delta}_j X_{i,j-1}\right)^2,\ \ 2 \leq j \leq n-1.$$

A bootstrap eljárás alkalmazásánál ezeket a paraméterbecsléseket használjuk. A bekövetkezett károk becslésére a következő igaz:

$$\hat{X}_{i,j} = \hat{E}\left[X_{i,j}|H_n\right] = \left(1-\hat{\delta}_j\right)\hat{X}_{i,j-1}, \;\; j\epsilon\{n-i+2, n-i+3, \dots, n\}$$

$$\hat{E}\left[X_{i,n-i+1}|H_n\right] = X_{i,n-i+1}$$

és a rekurziót $\hat{X}_{i,n-i+1} = X_{n-i+1}$-ből indítjuk.

4.3 BOOTSTRAP ALKALMAZÁSA A SCHNIEPER MODELLRE

A fent ismertett modell érdekes példa a bootstrap alkalmazásának szempontjából, mivel két különálló kifutási háromszögre kell egymástól függetlenül új mintát létrehozni. Ez teljesen más, mintha csak egyetlen kifutási háromszög áll rendelkezésünkre a kártartalék megállapításához, és arra végezzük el a bootstrap eljárást. A következőkben erre a speciális esetre mutatjuk be az algoritmust.

Ebben az esetben is szükséges feltétel, hogy az adatok függetlenek és azonos eloszlásúak legyenek. Mivel a rendelkezésre álló adataink általában nem ilyenek, most is a reziduálisokat fogjuk használni. Pontosabban, mivel a Schnieper-modell rekurzív modelleken alapul, az $\frac{N_{i,j}}{E_i}$ és $\frac{C_{i,j}}{X_{i,j-1}}$ arányok reziduálisaival fogunk számolni. Ezen arányok reziduálisainak kiszámításához szükségünk van a várható értékükre és a varianciájukra. Ezek az előzőek alapján az alábbiak:

$$E\left[\frac{N_{i,j}}{E_i}\middle| H_{i+j-2}\right] = \lambda_j$$

és

$$E\left[\frac{C_{i,j}}{X_{i,j-1}}\middle| H_{i+j-2}\right] = \delta_j,$$

illetve

$$D^2\left[\frac{N_{i,j}}{E_i}\middle| H_{i+j-2}\right] = \frac{\sigma_j^2}{E_i}$$

és

$$D^2\left[\frac{C_{i,j}}{X_{i,j-1}}\middle| H_{i+j-2}\right] = \frac{\tau_j^2}{X_{i,j-1}}.$$

Az előző fejezethez hasonlóan, például a becslési hiba bootstrap becslését úgy kapjuk meg, hogy elegendően nagyszámú - általában 1000, vagy akár 10000 - darab új adathalmazt generálunk az eredeti adatokból, és mindegyik új adathalmazra kiszámoljuk a tartalék becslését. Célunk az MSEP becslése a jövőbeli megfigyelésekre – tehát az $\{X_{i,m}: n-i+1 < m \leq n\}$ értékekre - vonatkozóan:

$$MSEP[\hat{X}_{i,m}|H_n] = E\left[\left(X_{i,m} - \hat{X}_{i,m}\right)^2\middle| H_n\right]$$

$$= E\left[\left(\left(X_{i,m} - E[X_{i,m}|H_n]\right) - \left(\hat{X}_{i,m} - E[X_{i,m}|H_n]\right)\right)^2\middle| H_n\right]$$

$$= E\left[\left(X_{i,m} - E[X_{i,m}|H_n]\right)^2\middle| H_n\right] + \left(\hat{X}_{i,m} - E[X_{i,m}|H_n]\right)^2$$

$= $ *folyamat varianciája + becslési hiba.*

Ahhoz, hogy a folyamat varianciáját megkapjuk, szükség van egy plusz szimulációra minden bootstrap lépés után. Legyenek $f_{i,j} = \frac{N_{i,j}}{E_i}$ és $g_{i,j} = \frac{C_{i,j}}{X_{i,j-1}}$. A skálázott Pearson-reziduálisokat a két kifutási háromszögre a következőképpen határozzuk meg:

$$r_{PS}\left(f_{i,j}, \hat{\lambda}_j, E_i, \hat{\sigma}_j\right) = \frac{\sqrt{E_i}\left(f_{i,j} - \hat{\lambda}_j\right)}{\hat{\sigma}_j}$$

és

$$r_{PS}\left(g_{i,j}, \hat{\delta}_j, X_{i,j-1}, \hat{\tau}_j\right) = \frac{\sqrt{X_{i,j-1}}\left(g_{i,j} - \hat{\delta}_j\right)}{\hat{\tau}_j}.$$

A bootstrap becsléshez most is szükséges korrekciót alkalmazni a rezduálisokra, az esetleges torzítások elkerülése miatt, így a módosított reziduálisok:

$$r_{i,j} = \sqrt{\frac{n-j}{n-j+1}} r_{PS}\left(f_{i,j}, \hat{\lambda}_j, E_i, \hat{\sigma}_j\right)$$

és

$$s_{i,j} = \sqrt{\frac{n-j}{n-j+1}} r_{PS}\left(g_{i,j}, \hat{\delta}_j, X_{i,j-1}, \hat{\tau}_j\right).$$

Ezeket a módosított reziduálisokat használjuk az új, reziduálisokból álló Bootstrap minták létrehozásához. A pszeudo-adatokból álló kifutási háromszögeket ezután a reziduális definíció alapján behelyettesítéssel kapjuk meg:

$$f_{i,j}^B = r_{i,j}^B \frac{\hat{\sigma}_j}{\sqrt{E_i}} + \hat{\lambda}_j,$$

illetve

$$g_{i,j}^B = s_{i,j}^B \frac{\hat{\tau}_j}{\sqrt{X_{i,j-1}}} + \hat{\delta}_j,$$

ahol $r_{i,j}^B$ és $s_{i,j}^B$ az új, reziduálisokból álló Bootstrap minták elemei, $i = 1,2,\dots,n$ és $j = 1,2,\dots,n-i+1$.

Minden egyes bootstrap mintára a paraméterek bootstrap becslését az egyéni növekedési faktorok súlyozott átlagából kapjuk meg:

$$\lambda_j^B = \frac{\sum_{i=1}^{n-j+1} f_{i,j}^B E_i}{\sum_{i=1}^{n-j+1} E_i}$$

és

$$\delta_j^B = \frac{\sum_{i=1}^{n-j+1} g_{i,j}^B X_{i,j-1}}{\sum_{i=1}^{n-j+1} X_{i,j-1}}.$$

Vegyük észre, hogy az $X_{i,j-1}$ megfigyelések és az E_i kitettség most súlyokként viselkednek, és a súlyoknál nem szabad bootstrappel kapott értékeket használni. A kárévenkénti és a teljes tartalékra vonatkozó bootstrap becslést az λ_j^B és δ_j^B paraméterek segítéségével kapjuk meg. Az [5] tanulmány az alábbi formulát adja a bekövetkezett károk becslésére:

$$\hat{X}_{i,j} = \left(1 - \delta_j^B\right)\hat{X}_{i,j-1} + E_i\lambda_j^B,$$

ahol $j \in \{n-i+2, n-i+3, \dots, n\}$ és a kezdeti érték $\hat{X}_{i,n-i+1} = X_{i,n-i+1}$.

A bootstrap eljárás csak a modell becslési hibáját adja meg. Azonban a tartalékolási folyamat során a becslési hibára és a tartalék eloszlásának becslésére is szükségünk van. Ahhoz, hogy ezeket megkapjuk, szükségünk van a folyamat varianciájára. Ezt a folyamat eloszlásából kaphatjuk meg. Mivel ehhez csak az első két momentumra van szükségünk, a legegyszerűbb, ha feltesszük, hogy $N_{i,j}$ és $C_{i,j}$ normális eloszlásúak - de például használhatnánk az ODP modellt is. Így mindkét háromszögre megkapjuk a szimulált értékeket az inkrementális kárnagyságokra, a megfelelő feltételes eloszlások felhasználásával:

$$\left.\frac{N_{i,j}}{E_i}\right| H_{i+j-2} \sim N\left(\lambda_j^B, \frac{\sigma_j^2}{E_i}\right),$$

illetve

$$\left.\frac{C_{i,j}}{X_{i,j-1}}\right| H_{i+j-2} \sim N\left(\delta_j^B, \frac{\tau_j^2}{X_{i,j-1}}\right).$$

Ezeknek a szimulált értékeknek a segítségével kapunk becslést a fennálló kötelezettségek mértékére. A folyamat varianciáját az eloszlás varianciájából kapjuk meg szimuláció segítségével. A teljes eloszlás becslését ezekből a szimulált értékekből tudjuk kiszámítani a bootstrap eljárás segítségével.

Összefoglalva tehát, a bootstrap eljárást a következőképpen implementáljuk a Schnieper-modellre:

1. Kiszámítjuk az $f_{i,j}$ és $g_{i,j}$ hányadosokat és azok varianciáit az IBNR és IBNER kifutási háromszögekre. A σ_j^2 és τ_j^2 varianciák az eljárás alatt végig változatlanok maradnak, mivel nem számoljuk újra az értéküket a bootstrap mintákból.
2. Kiszámítjuk az $r_{PS}\left(f_{i,j}, \hat{\lambda}_j, E_i, \hat{\sigma}_j\right)$ és $r_{PS}\left(g_{i,j}, \hat{\delta}_j, X_{i,j-1}, \hat{\tau}_j\right)$ skálázott Pearson-reziduálisokat.

3. A bootstrap torzításának kiküszöbölésére megszorozzuk a skálázott Pearson-reziduálisok értékét a $\sqrt{\frac{n-j}{n-j+1}}$ tényezővel.

   Az iterációt N-szer hajtjuk végre, ahol N megfelelően nagy, például $N \geq 1000$.
4. Legyen $B = 1$.
5. Véletlenszerűen, visszatevéses mintavétellel válasszunk ki elemeket a reziduálisokat tartalmazó kifutási háromszögekből, ezek halmazát jelölje $R = \{r_{i,j}, i = 1, \dots, n; j = 1, \dots, n-i+1\}$ és $S = \{s_{i,j}, i = 1, \dots, n; j = 1, \dots, n-i+1\}$. Jelölje a bootstrapből kapott reziduálisokat $r_{i,j}^B$ és $s_{i,j}^B$. Így létrehoztunk két pszeudo mintát az igazi IBNR és IBNER károkra vonatkozó Pearson-reziduálisokból. Jelölje ezt a két mintát $R^B = \{r_{i,j}^B, i = 1, \dots, n; j = 1, \dots, n-i+1\}$ és $S^B = \{s_{i,j}^B, i = 1, \dots, n; j = 1, \dots, n-i+1\}$.
6. Kiszámítjuk $f_{i,j}^B$-t és $g_{i,j}^B$-t.
7. Kiszámítjuk az $N_{i,j}$-vel és $C_{i,j}$-vel súlyozott átlagos bootstrap növekedési faktorokat az eredeti IBNR és IBNER értékekre. Ezeket jelölje λ_j^B és δ_j^B.
8. Szimuláljuk a kifutási háromszögekhez tartozó alsó háromszögmátrixok minden egyes elemének értékét, azaz a jövőbeli kárkifizetéseket. Ehhez a folyamat eloszlását használjuk fel, ez előző pontban kiszámolt λ_j^B és δ_j^B várható értékekkel, azaz:

$$\left.\frac{N_{i,j}}{E_i}\right| H_{i+j-2} \sim N\left(\lambda_j^B, \frac{\sigma_j^2}{E_i}\right)$$

   az eredeti IBNR károkra vonatkozóan, illetve:

$$\left.\frac{C_{i,j}}{X_{i,j-1}}\right| H_{i+j-2} \sim N\left(\delta_j^B, \frac{\tau_j^2}{X_{i,j-1}}\right)$$

   teljesül a jövőbeli IBNER károkra.
9. A tartalékok becslését az előző pontban kiszámolt szimulált inkrementális kárnagyságokból $(-C_{i,j} + N_{i,j})$ állapíthatjuk meg.
10. Tároljuk az így kapott eredményeket, legyen $B = B + 1$, és az 5. ponttól kezdjük el az iterációt. Az iteráció akkor ér véget, ha $B = N$.

5. FEJEZET

SZTOCHASZTIKUS MÓDSZEREK A GYAKORLATBAN

5.1 ALKALMAZÁS

Ebben a fejezetben az előzőekben ismertetett módszerek közül néhányat a gyakorlatban is bemutatunk és a kapott eredményeket összehasonlítjuk.

A kifutási háromszög, amellyel dolgozunk (*ld. 1. Melléklet*), egy fiktív biztosító egyik gépjármű ágazatának kummulált anyagi kárkifizetéseit tartalmazza, 1 forintban megadva. Tehát a személyi sérülésekből adódó károkkal most nem foglalkozunk. Feltehetjük, hogy az adataink a nagy károktól már meg vannak tisztítva. Célunk, hogy a rendelkezésre álló adatok alapján legjobb becslést adjunk az ágazat 2013. év végi tartalékára, meghatározzuk a káreloszlást, és a tartalékolási kockázatot. A számítások a Towers Watson *ResQ* tartalék-modellező szoftverével készültek.

Először a standard lánc-létra módszerből adódó becslést szeretnénk megkapni. A növekedési faktorokat (*ld. 2. Melléklet*) a kummulált kárkifizetések alapján megadja a program, de könnyedén ki is számolhatóak a tanult képlet szerint. Látszik, hogy az utolsó három évben a növekedési faktorok értéke kiugró az előző értékekhez képest. Ez például adódhat az ágazat portfóliójának csökkenéséből. Ha a növekedési faktorok között kiugró értéket tapasztalunk, érdemes megvizsgálni, hogy az miből adódhat, mivel előfordulhat, hogy az nagyban torzítja a becslést. A szoftver lehetőséget ad arra, hogy ilyen problémás esetekben figyelmen kívül hagyjunk bizonyos kiugró faktorokat a becslésből. A *2. Melléklet* alsó táblázatában kiválasztottuk, hogy milyen „mértékben” vesszük figyelembe az egyes pénzügyi évekhez tartozó növekedési faktorokat. Választhatjuk az utolsó 1,2,…,5 év növekedési faktorainak átlagát, a lánc-létra módszer súlyozását, az adott oszlophoz tartozó legkisebb növekedési faktort, stb. Ezek között az aktuáriusok általában a múltbeli tapasztalatok alapján döntenek, vagy ha attól eltérő módszert alkalmaznak, akkor azt alaposan meg kell indokolni. Ebben a példában a standard lánc-létra módszert választottuk (ld. *„Volume – all”).* Fontos megjegyezni, hogy ha valamely növekedési faktor értéke kisebb, mint 1, akkor az ODP módszer az adott kifutási háromszögre nem alkalmazható. Ilyen előfordulhat például akkor, ha kárkifizetések helyett az „incurred” adatokból szeretnénk számolni. A továbbiakban az ODP és a Mack modellt szeretnénk összehasonlítani a standard lánc-létrát használva kiindulásként.

A standard lánclétra növekedési faktorainak kiszámítása után a ResQ-ban lehetőség van görbét illeszteni a kapott értékekre, illetve farok eloszlást meghatározni *(ld. 3. Melléklet)*, amennyiben az adott ágazat hosszabb kifutású mint a tapasztalatunk. Egyszerűség kedvéért a farokeloszlással most nem foglalkozunk, és nem illesztünk görbét, hanem az eredeti lánclétra növekedési faktorokat használjuk a további számítások során. A lánc-létra módszerből kapott eredmények az 4. *Mellékletben* találhatóak. A táblázat 1. oszlopa (*„Latest”*) az összes eddig ismert kifizetést tartalmazza az ágazat állományára vonatkozóan. Megkaptuk a kárkifutási háromszög adatai alapján szükséges tartalék nagyságát, a farokeloszlás figyelmen kívül hagyásával. A legjobb becslést a *„Total Reserve”* oszlopban találjuk.

Most vizsgáljuk meg, hogy hogyan módosulnak az eredményeink, ha az ODP, illetve a Mack modellre alkalmazzuk a bootstrap eljárást. A bootstrap segítségével megkapjuk a teljes prediktív eloszlást, ezáltal a tartalékolási kockázat nagyságát. A bootstrap során a *ResQ* 10.000 szimulációt hajt végre az adatokon – a szimulációk száma tetszőlegesen állítható -, a 4.1 fejezetben ismertetett algoritmus alapján. Mindkét modellre futtattuk az eljárást, a kapott eredmények az *5. Mellékletben* találhatóak. A bootstrap alapján várható tartalék nagysága az *„Expected Reserve”* oszlopban található. Az eredmények alapján elmondható, hogy a két esetben a szükséges tartalék nagysága viszonylag közeli, és természetesen a standard lánc-létra módszerből kapott tartalékszükséglethez is közel vannak. Mindkét szimulációs eljárás során nagyobb tartalékszükségletet kaptunk, mint a lánc-létra módszer esetében. A Mack módszer becslése adja a standard lánc-létra becsléséhez közelebbi eredményt. A bootstrappel az előrejelzési hibák nagyságát is megkaptuk, ezek is közelinek mondhatóak, a Mack-módszer esetében 1,5%-al nagyobb.

A szimulációból megkaptuk az eloszlást is, és a hozzá tartozó kvantilisek nagyságát. A kvantilisek közül számunkra a 0,5%-os érdekes, mivel annak az ellentettje adja a 99,5%-os VaR-t, tehát azt az értéket, amit akkor kell tartalékolnunk, ha egy éves időhorizonton 99,5%-s valószínűséggel el akarjuk kerülni a csődbemenést. A *6. Mellékletben* tehát a Szolvencia II. szerinti tartalékolási kockázathoz kapcsolódó szavatolótőke szükségletének ellentettjeit láthatjuk a két módszer esetében egy fiktív biztosító KGFB anyagi káraira számolva.

5.2 PROBLÉMÁK

Bár a sztochasztikus tartalékolási módszerek elmélete és gyakorlata az elmúlt évtizedek során rengeteget fejlődött, a gyakorlati alkalmazások során még mindig felmerülnek problémás helyzetek és megoldatlan kérdések. Ezek között vannak olyanok, amelyek a tartalékoló szoftverek technikai korlátaiból adódnak, és olyanok is, amelyekre még nem született meg a megfelelő matematikai megoldás. Most néhány ilyen problémára mutatunk rá, amelyekkel az aktuáriusok gyakran szembesülnek a tartalékolási feladatok során.

Az alkalmazás során megjegyeztük, hogy az egyszerűség kedvéért a károk farokeloszlását figyelmen kívül hagyjuk, azonban a hosszú kifutású ágazatoknál ez nem szerencsés választás. Viszont ha az ún. *„tail factort"* is figyelembe vesszük, akkor a bootstrap eljárás során a visszatevéses mintavételnél ebből a farokeloszlásból is veszünk adatokat. Ez pedig azért probléma, mert nagy előrejelzési hibát okozhat, ami a tőkeszükséglet felülbecsléséhez vezethet.

A példában láthattuk, hogy a növekedési faktorok értéke az utolsó 3 évre kiugró, és ezt a portfólió változásával magyaráztuk. Ekkor értelemszerűen célszerűbb lenne az utolsó 3 évre vonatkozóan más módszert alkalmazni, mint az előzőekre, vagy a legjobb becslést akár több módszer átlagával vagy súlyozott átlagával meghatározni. Azonban a bootstrap eljárást több módszer súlyozott átlagára nem tudjuk végrehajtani. Ennek a problémának a kiküszöbölésére a ResQ ugyan kínál egy megoldást (*„Target Reserve"*), amelyben a végső tartalék értékéhez lehet igazítani az eloszlás várható értékét, de ez nagyobb eltérések esetén nem alkalmazható.

Fontos megemlíteni az infláció kérdését, mivel a jelenlegi gazdasági környezetben és az évről évre sokat változó jövőbeli inflációs feltételezések miatt ezzel is érdemes lenne számolni, azonban az infláció alkalmazhatósága a sztochasztikus tartalékolási technikák során sokat vitatott kérdés.

Egy másik kérdés a nagy károk kezelése. Bár az alkalmazás során feltettük, hogy az adataink a nagy károktól megtisztítottak, a valóságban a best estimate megállapítása során a nagy károkkal, illetve a nagy károkhoz kapcsolódó tartalékolási kockázattal is számolni kell. Itt felmerül egy újabb probléma, mivel a nagy károkra, jellegükből adódóan rájuk a fenti módszer nem alkalmazható.

6. FEJEZET

ÖSSZEFOGLALÁS

Ezen szakdolgozat két fő célja a sztochasztikus kártartalékolási technikák elméleti és gyakorlati bemutatása volt. Mivel ez a téma elég sokszínű, így csak azokat a módszereket emeltük ki, amelyek vagy elméleti szempontból érdekesek, vagy a gyakorlatban jól alkalmazhatóak. A témának még nincsen magyar szakirodalma, ezért sok helyen az angol kifejezéseket is szerepeltettük az egyértelműség érdekében.

A 2. fejezetben az ODP, a negatív binomiális és a Mack modell elméletét ismertettük. Ezen modellek célja, hogy a lánc-létra módszer eredményeihez hasonló becslést adjanak a kártartalék nagyságára, de kevesebb feltételezés mellett. A 3. fejezetben az így becsült tartalék négyzetes előrejelzési hibájával és a becslési folyamatok hibájával foglalkoztunk. A rekurzív modellekre vonatkozó folyamat és becslési hibákat általánosan is ismetettük. Összehasonlítottuk az ODP és a negatív binomiális modellt, és megállapítottuk, hogy a két modell esetében az eredmények lényegében ugyanazok, csak a paraméterezésükben különböznek.

A 4. fejezet célja a bootstrap eljárás bemutatása volt. Ez a számítógépen gyorsan futtatható modern eljárás azért hasznos, mert az eddig bemutatott módszerek előrejelzési hibája csak a teljes tartalék prediktív eloszlásának második momentumáról adott információt. A bootstrap segítségével azonban megkapjuk a teljes eloszlást, amiből megállapíthatjuk a kvantiliseket, és így a 99,5%-os VaR értékét, ami a Szolvencia II. szerinti szavatoló tőkeszükséglete. Ismertettük a bootstrap eljárás feltételeit, és az algoritmust, valamint annak implementálását az ODP és a Mack modellekre. Bemutattuk a Schnieper-modellt, ami megkülönbözteti egymástól az IBNR és IBNER károkat. Ez egy érdekes modell a bootstrap szempontjából, mivel az algoritmus a két kifutási háromszöget egymástól függetlenül kezeli.

Az 5. fejezetben néhány eddig ismertetett módszert egy fiktív biztosító egyik gépjármű ágazatának kummulált anyagi kárkifizetéseire alkalmaztunk a ResQ tartalékoló program segítségével. Ismertettük a program használatának főbb lépéseit, illetve, hogy melyek azok a legfontosabb dolgok, amelyekre a beállítások során figyelni kell. Célunk az ODP és a Mack modell összehasonlítása volt a lánc-létra módszerrel, ezekre futtattuk az eljárást. A kapott eredményeket a mellékletekben szemléltetjük. Megkaptuk mindegyik esetre a legjobb becslést és az előrejelzési hibát. Elmondható, hogy mindkét bootstrap eljárással a tartalék nagyságára hasonló eredményt kaptunk, mint a lánc-létra módszerrel, és a két bootstrap előrejelzési hibái is közeliek lettek. A szimulációkból megkaptuk a teljes prediktív eloszlást, ezáltal a kvantiliseket, amelyekből megállapítható a 99,5%-os VaR értéke.

Végezetül összefoglaltuk azokat a főbb problémákat, amelyek a bemutatott módszerek gyakorlati alkalmazása során felléphetnek. Ezek alapján elmondható, hogy bár a sztochasztikus tartalékolási módszerek rengeteg szempontból hasznosak, és többet tudnak a determinisztikus technikáknál, még sok fejlesztésre szorulnak, mind az elméleti, mind a technikai megvalósítás szempontjából.

Accident year	0	1	2	3	4	5	6	7	8	9	10	11	12	13
2000	231 263 678	330 245 990	341 991 212	346 954 166	349 622 841	350 645 404	351 012 585	351 210 639	351 567 669	351 742 934	351 757 398	351 794 715	351 794 715	351 926 452
2001	246 204 962	345 518 869	359 789 523	365 505 438	368 559 002	369 839 031	370 404 001	371 579 676	372 403 500	372 483 587	372 509 639	372 568 503	372 601 559	
2002	235 166 557	337 630 703	349 696 889	356 410 992	359 158 029	360 159 992	360 724 044	361 029 312	361 096 285	361 176 032	361 179 167	361 420 331		
2003	235 673 804	340 291 019	350 988 876	354 053 927	356 702 086	357 319 225	358 335 086	358 564 282	358 679 388	358 694 044	358 991 673			
2004	211 947 450	297 433 177	307 743 045	311 350 235	312 428 238	313 819 264	314 692 360	315 057 734	315 110 282	315 220 395				
2005	178 042 348	260 575 058	270 116 197	273 077 794	274 926 895	275 963 715	276 574 311	277 009 648	277 027 792					
2006	163 466 376	234 813 783	242 226 082	245 559 983	247 304 215	248 712 465	249 197 882	250 030 366						
2007	143 923 921	204 349 040	211 536 398	217 680 992	219 338 511	220 919 981	221 804 914							
2008	126 919 443	188 757 536	195 186 037	199 202 565	202 092 405	203 551 598								
2009	117 351 427	165 445 570	174 083 671	177 637 557	178 801 134									
2010	113 795 290	172 692 660	182 177 897	185 451 279										
2011	70 736 916	103 805 139	109 935 837											
2012	48 887 183	82 766 045												
2013	47 735 289													
Total	**2 171 114 644**	**3 064 324 588**	**3 095 471 663**	**3 032 884 928**	**2 868 933 357**	**2 700 930 674**	**2 502 745 182**	**2 284 481 657**	**2 035 884 916**	**1 759 316 992**	**1 444 437 877**	**1 085 783 549**	**724 396 273**	**351 926 452**

1. Melléklet – Kummulált kárkifutási háromszög.

Development Factor Method - Ratios & Average Selection

Origin Period	1	2	3	4	5	6	7	8	9	10	11	12	13
Ratios													
2000	1.42801	1.03557	1.01451	1.00769	1.00292	1.00105	1.00056	1.00102	1.00050	1.00004	1.00011	1.00000	1.00037
2001	1.40338	1.04130	1.01589	1.00835	1.00347	1.00153	1.00317	1.00222	1.00022	1.00007	1.00016	1.00009	
2002	1.43571	1.03574	1.01920	1.00771	1.00279	1.00157	1.00085	1.00019	1.00022	1.00001	1.00067		
2003	1.44391	1.03144	1.00873	1.00748	1.00173	1.00284	1.00064	1.00032	1.00004	1.00083			
2004	1.40333	1.03466	1.01172	1.00346	1.00445	1.00278	1.00116	1.00017	1.00035				
2005	1.46356	1.03662	1.01096	1.00677	1.00377	1.00221	1.00157	1.00007					
2006	1.43647	1.03157	1.01376	1.00710	1.00569	1.00195	1.00334						
2007	1.41984	1.03517	1.02905	1.00761	1.00721	1.00401							
2008	1.48722	1.03406	1.02058	1.01451	1.00722								
2009	1.40983	1.05221	1.02041	1.00655									
2010	1.51757	1.05493	1.01797										
2011	1.46748	1.05906											
2012	1.69300												
2013													
Averages													
1: Volume - 1	1.69300	1.05906	1.01797	1.00655	1.00722	1.00401	1.00334	1.00007	1.00035	1.00083	1.00067	1.00009	1.00037
2: Volume - 2	1.55965	1.05648	1.01916	1.01076	1.00722	1.00292	1.00241	1.00012	1.00019	1.00042	1.00041	1.00005	1.00037
3: Volume - 3	1.53913	1.05488	1.01966	1.00961	1.00665	1.00266	1.00194	1.00020	1.00020	1.00030	1.00031	1.00005	1.00037
4: Volume - 4	1.49588	1.04865	1.02227	1.00887	1.00581	1.00269	1.00155	1.00019	1.00020	1.00024	1.00031	1.00005	1.00037
5: Volume - 5	1.49358	1.04535	1.02022	1.00836	1.00547	1.00273	1.00139	1.00064	1.00026	1.00024	1.00031	1.00005	1.00037
6: Volume - all	1.44314	1.03821	1.01586	1.00755	1.00401	1.00215	1.00155	1.00070	1.00026	1.00024	1.00031	1.00005	1.00037
7: Vol + 0,9 - all	1.45508	1.03997	1.01658	1.00771	1.00441	1.00230	1.00162	1.00061	1.00025	1.00027	1.00033	1.00005	1.00037
8: Simple - all	1.46225	1.04019	1.01662	1.00772	1.00436	1.00224	1.00161	1.00066	1.00026	1.00024	1.00031	1.00004	1.00037
9: Lowest - all	1.40333	1.03144	1.00873	1.00346	1.00173	1.00105	1.00056	1.00007	1.00004	1.00001	1.00011	1.00000	1.00037
10: Highest - all	1.69300	1.05906	1.02905	1.01451	1.00722	1.00401	1.00334	1.00222	1.00050	1.00083	1.00067	1.00009	1.00037
11: User Entry	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

2. Melléklet – Növekedési faktorok a standard lánc-létra módszerből.

Development Factor Method - Curves Data

Future Dev. Periods: 1

	Initial Selection (1)	Include	Exp. Decay (2)	Inverse Power (3)	Power (4)	Weibull (5)	User Entry (6)	Selected Value
1	1.44314	Yes	1.07079	1.30681	1.06900	1.23967	1.00000	1.44314
2	1.03821	Yes	1.03893	1.05914	1.03764	1.07318	1.00000	1.03821
3	1.01586	Yes	1.02141	1.01839	1.02068	1.02870	1.00000	1.01586
4	1.00755	Yes	1.01178	1.00743	1.01140	1.01258	1.00000	1.00755
5	1.00401	Yes	1.00648	1.00355	1.00630	1.00589	1.00000	1.00401
6	1.00215	Yes	1.00356	1.00190	1.00348	1.00289	1.00000	1.00215
7	1.00155	Yes	1.00196	1.00110	1.00193	1.00147	1.00000	1.00155
8	1.00070	Yes	1.00108	1.00068	1.00107	1.00077	1.00000	1.00070
9	1.00026	Yes	1.00059	1.00045	1.00059	1.00041	1.00000	1.00026
10	1.00024	Yes	1.00033	1.00030	1.00033	1.00022	1.00000	1.00024
11	1.00031	Yes	1.00018	1.00021	1.00018	1.00013	1.00000	1.00031
12	1.00005	Yes	1.00010	1.00015	1.00010	1.00007	1.00000	1.00005
13	1.00037	Yes	1.00005	1.00011	1.00006	1.00004	1.00000	1.00037
Ult	1.00000		1.00003	1.00009	1.00003	1.00002	1.00000	1.00000
Fit			OK	OK	OK	OK		
R-squared %			86.94%	94.81%	87.45%	92.83%		
Tail Pattern							X	
14			1.00003	1.00009	1.00003	1.00002	1.00000	

3. Melléklet – Illeszthető görbék.

Development Factor Method - Results Summary

Origin Period	Latest	Triangle Reserve	Tail Reserve	Total Reserve	Ultimate Values
2000	351 926 452	0	0	0	351 926 452
2001	372 601 559	139 529	0	139 529	372 741 087
2002	361 420 331	151 841	0	151 841	361 572 172
2003	358 991 673	262 438	0	262 438	359 254 111
2004	315 220 395	304 989	0	304 989	315 525 384
2005	277 027 792	340 538	0	340 538	277 368 330
2006	250 030 366	483 757	0	483 757	250 514 124
2007	221 804 914	774 188	0	774 188	222 579 102
2008	203 551 598	1 149 373	0	1 149 373	204 700 971
2009	178 801 134	1 731 395	0	1 731 395	180 532 530
2010	185 451 279	3 209 610	0	3 209 610	188 660 889
2011	109 935 837	3 676 365	0	3 676 365	113 612 202
2012	82 766 045	6 035 676	0	6 035 676	88 801 721
2013	47 735 289	26 176 874	0	26 176 874	73 912 164
Total	**3 317 264 664**	**44 436 574**	**0**	**44 436 574**	**3 361 701 238**

4. Melléklet – DFM eredmények.

Bootstrap Method - Results Summary
ODP

Accident Year	Latest	Expected Reserve	Prediction Error	Prediction Error%	Expected Ultimate	DFM Reserve	Reserve Difference
2000	351 926 452	0	0	0.00%	351 926 452	0	0
2001	372 601 559	138 384	78 382	56.64%	372 739 942	139 529	-1 145
2002	361 420 331	149 737	78 628	52.51%	361 570 068	151 841	-2 104
2003	358 991 673	259 711	145 569	56.05%	359 251 384	262 438	-2 727
2004	315 220 395	302 983	205 877	67.95%	315 523 378	304 989	-2 006
2005	277 027 792	337 488	197 042	58.38%	277 365 280	340 538	-3 049
2006	250 030 366	482 310	337 472	69.97%	250 512 676	483 757	-1 447
2007	221 804 914	765 824	473 603	61.84%	222 570 738	774 188	-8 364
2008	203 551 598	1 149 839	527 726	45.90%	204 701 437	1 149 373	465
2009	178 801 134	1 720 828	685 530	39.84%	180 521 963	1 731 395	-10 567
2010	185 451 279	3 213 682	951 831	29.62%	188 664 961	3 209 610	4 072
2011	109 935 837	3 684 735	1 284 818	34.87%	113 620 573	3 676 365	8 370
2012	82 766 045	6 016 253	1 686 976	28.04%	88 782 298	6 035 676	-19 423
2013	47 735 289	26 222 538	4 109 010	15.67%	73 957 827	26 176 874	45 664
Total	**3 317 264 664**	**44 444 312**	**5 021 721**	**11.30%**	**3 361 708 976**	**44 436 574**	**7 739**

Mack

Accident Year	Latest	Expected Reserve	Prediction Error	Prediction Error%	Expected Ultimate	DFM Reserve	Reserve Difference
2000	351 926 452	0	0	0.00%	351 926 452	0	0
2001	372 601 559	139 341	32 610	23.40%	372 740 900	139 529	- 188
2002	361 420 331	152 777	42 698	27.95%	361 573 109	151 841	936
2003	358 991 673	262 090	135 054	51.53%	359 253 763	262 438	- 348
2004	315 220 395	301 649	190 262	63.07%	315 522 044	304 989	-3 340
2005	277 027 792	339 224	186 356	54.94%	277 367 016	340 538	-1 314
2006	250 030 366	485 561	318 004	65.49%	250 515 927	483 757	1 803
2007	221 804 914	773 995	440 975	56.97%	222 578 909	774 188	- 193
2008	203 551 598	1 151 505	483 023	41.95%	204 703 103	1 149 373	2 131
2009	178 801 134	1 742 860	628 434	36.06%	180 543 994	1 731 395	11 464
2010	185 451 279	3 202 112	859 988	26.86%	188 653 391	3 209 610	-7 498
2011	109 935 837	3 668 802	1 167 574	31.82%	113 604 639	3 676 365	-7 563
2012	82 766 045	6 029 585	1 577 336	26.16%	88 795 630	6 035 676	-6 091
2013	47 735 289	26 190 817	5 070 944	19.36%	73 926 107	26 176 874	13 943
Total	**3 317 264 664**	**44 440 318**	**5 679 188**	**12.78%**	**3 361 704 982**	**44 436 574**	**3 744**

5. Melléklet – Bootstrap eredmények (ODP, Mack).

Bootstrap Run-off Result - Percentile Run-off Results
ODP

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Mean	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std. Dev.	0	78 382	45 395	128 441	160 113	75 342	283 421	358 142	267 517	484 737	650 375	1 065 639	1 273 318	3 928 482
Min.	0	-363 223	-213 403	-661 409	-880 628	-359 965	-1 465 197	-1 449 271	-1 159 514	-1 851 791	-2 783 057	-3 750 064	-4 388 726	-17 191 320
0.5%	0	-257 060	-146 016	-425 406	-543 788	-225 292	-947 963	-1 075 843	-769 145	-1 407 096	-1 832 296	-2 934 622	-3 469 088	-11 306 273
													Total	-12 606 325

Mack

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Mean	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std. Dev.	0	32 610	32 339	129 488	150 864	67 672	265 744	330 171	238 937	438 396	586 551	973 549	1 185 769	4 883 708
Min.	0	-121 066	-135 340	-519 058	-571 258	-253 551	-981 199	-1 243 300	-845 249	-1 412 080	-1 977 364	-3 148 115	-3 566 405	-15 140 611
0.5%	0	-91 916	-89 204	-356 508	-415 342	-183 991	-739 760	-898 230	-659 863	-1 191 238	-1 632 594	-2 675 179	-3 253 407	-13 390 913
													Total	-14 633 582

6. Melléklet – 0,5%-os kvantilisek.

IRODALOMJEGYZÉK

[1] *Általános biztosítás II.* kurzus előadásjegyzete, ELTE, 2013-2014/1. félév

[2] ENGLAND, P. D., VERRALL, R. J.: *Analytic and bootstrap estimates of prediction errors in claims reserving*, Insurance: Mathematics and Economics 25. (1999) 281-293.

[3] ENGLAND, P. D., VERRALL, R. J.: *Stochastic claims reserving in general insurance*, British Actuarial Journal 8/3. (2002) 443-518.

[4] INTERNATIONAL ACTUARIAL ASSOCIATION: *Stochastic modeling – Theory and reality from an actuarial perspective* (2010)

[5] LIU, H., VERRALL, R. J.: *A bootstrap estimate of the predictive* d*istribution of outstanding claims for the Schnieper model*, ASTIN Bulletin 39/2. (2009) 677-689.

[6] MACK, T.: *Distribution-free calculation of the standard error of chain-ladder reserve estimates*, ASTIN Bulletin 23. (1993) 213-225.

[7] PETERS, G., WÜTHRICH, M., SHEVCHENKO, P.: *Chain-ladder method: Bayesian bootstrap versus classical bootstrap, Research report (2009), ETH Zurich*.

[8] RENSHAW, A. E.: *Claims reserving by joint modelling*, Actuarial Research Paper 72. (1994), Department of Actuarial Science and Statistics, City University, London

[9] RENSHAW, A. E., VERRALL, R. J.: *A stochastic model underlying the chain-ladder technique*, B.A.J. 4. (1998) 903-923.