Diplomamunka

Nem-élet tartalékolási módszerek megbízhatóságáról

Írta: Szabó István

matematikus szak

Témavezető:

Arató Miklós, egyetemi docens
Valószínűségelméleti és Statisztika Tanszék

Eötvös Loránd Tudományegyetem
Természettudományi Kar
2011

Tartalomjegyzék

1. Klasszikus módszerek **5**
1.1. Kifutási háromszögek 5
1.2. Láncszemhányados módszerek, a lánclétra módszer 6
1.3. Bornhuetter-Ferguson módszer 8

2. A Mack-modell **10**
2.1. Klasszikus módszerek keverése 11
2.2. Optimális súly - a Mack-modell 12

3. Gamma eloszlású kárhányad **19**
3.1. Kiinduló feltételek teljesülése 19
3.2. Az elméleti optimális súly 22
3.3. Mack-modell a gyakorlatban 26
3.3.1. A becslés lépései 26
3.3.2. Az elértségi szintekről 32
3.3.3. Eredmények 34
3.3.4. Randomizált eredmények 39
3.3.5. Összefoglalás 45

Bevezető

Egy biztosítási szerződéskötés során a biztosító díj ellenében kockázatot vállal át a szerződőtől. Ha a biztosítási esemény bekövetkezik, fontos, hogy tudja teljesíteni azt, amiben a két fél megállapodott, hiszen ezért a biztonságért köt biztosítást a szerződő. A különböző jövőbeli kifizetésekre a jelenben a biztosítónak tartalékot kell képeznie a már beszedett díjakból, több okból is. Egy hosszabb távú szerződésnél a díj arányosan lett megállapítva, a jövőben várható kifizetések egy részét a már beszedett díjak fedezik. Egy rövid távú szerződés esetén is gyakran előfordul, hogy a kifizetés valamiért késik: késhet a bejelentés, elhúzódhat a kárrendezés. A biztosítástechnikai tartalék a jövőbeli kifizetések és a jövőbeli (díj)bevételek különbségére képzett tartalék. A biztosítónak azért kell ezzel rendelkeznie, hogy később rendezni tudja a kárt.

A tartalékok másik igen fontos szerepe a biztosító mérlegének helyes kimutatásában van. Ha a díjat beszedtük, de teljesítenünk csak a jövőben kell, akkor ha nem képzünk tartalékot, amit a kötelezettségek közé állítunk, esetlegesen túlzottan pozitív eredményt mutatunk ki. Másrészről, a tartalékképzéssel csökkenthető az eredmény, ezzel adózás szempontjából jól jár a biztosító. Emiatt a két ok miatt a tartalékképzés fő szabályai törvényileg meghatározottak, de természetesen többféle képzési módszer is megfelelő, amelyek eltérő tartalékszükségletet jelezhetnek előre.

A tartalékokat csoportosíthatjuk díjtartalékokra és kártartalékokra. A díjtartalékok esetében azok a károk, melyekre a tartalékot képezzük, még nem következtek be, a kártartalékok esetében a károk már bekövetkeztek, csak később kerülnek rendezésre. A kártartalékok között foglal helyet a függőkárok tartaléka, ez a nem-élet ági tartalékok közül a legnagyobb nagyságú (lásd: [3] 5.4.1. ábra), Magyarországon két részre bontva kell megképezni: tételes függőkár tartalék és IBNR tartalék (incurred but not reported - bekövetkezett, de még nem bejelentett). Felelősségbiztosításnál például gondolhatunk arra, hogy egy balesetből származó vagyoni kárt a biztosító már megtérített, de előfordulhat, hogy a baleset következtében hónapok, évek múlva jelentkezik valamiféle balesettel összefüggő állapotromlás a biztosítottnál. A szükséges tartalékok minél pontosabb becslése tehát alapvető fontosságú a biztosító életében. A gyakorlatban egyenlőre még jobban jelenlévő hagyományos és a sztohasztikus módszerek egy átfogó leírása található [5]-ben. A sztohasztikus feltevéseket használó kártartalékolási módszerekről egy rendszerező, sok gyakorlati példát is bemutató cikk [6].

A dolgozatomban kifutási háromszögön alapuló IBNR tartalékolási módszereket fogok vizsgálni. Az első fejezetben [3]-at követve (és némiképp [4]-et) bemutatom a klasszikus módszereket. A második fejezetben Mack [1] cikkében található optimális keverésre vonatkozó elméleti eredményeket mutatom

be. Ez az elméleti eredmény adja a vizsgálódásunk további tárgyát. A 3. fejezetben belátom, hogy Gamma eloszlású kárhányadot feltételezve teljesülnek a Mack-modell feltételei, majd meghatározom az elméleti optimális súlyt. 3.3-ban a Mack-modell gyakorlati megvalósítását vizsgálom: kifutási háromszöget generálok Gamma eloszlásból, a szükséges paramétereket megbecsülöm, majd a két klasszikus (Bornhuetter-Ferguson, Chain-Ladder) és a Mack-modellből származó módszer által adott tartalékbecsléseket összehasonlítom. Vizsgálom, hogy a korábbi, már teljes kifutással rendelkező évek száma és az adataink bizonytalansága (szórása) milyen hatással van az eredményekre. A becslések elkészítésére megírt program és a számítások saját eredmények.

Köszönetnyilvánítás

Szeretném megköszönni témavezetőmnek, Arató Miklósnak, hogy az egész év során rendszeresen időt szakított rám, témajavaslata, szakmai útmutatása és kérdéseimre adott válaszai nagyon sokat segítettek abban, hogy ez a dolgozat elkészülhessen.

1. Klasszikus módszerek

Ebben a fejezetben definiálom a kifutási háromszög fogalmát, majd röviden bemutatom a biztosítási tartalékolásban elterjedt, kifutási háromszög vizsgálatán alapuló klasszikus tartalékolási módszereket. A fejezet megírásához [3]-at vettem alapul.

1.1. Kifutási háromszögek

A kifutási háromszög egy összesített kártáblázat. Az alapfelbontás szerint a kár bekövetkezésének vagy a kár bejelentésének éve alapján csoportosítjuk a károkat, a lebonyolítás szerint pedig a bejelentés, kifizetés vagy a lezárás alapján csoportosítunk. Mindezt egy olyan mátrixszal ábrázoljuk, ahol a mátrix sorai felelnek meg az alapfelbontási, a mátrix oszlopai pedig a lebonyolítási periódusoknak.

Alapfelbontás éve	Lebonyolítás éve				
	1	2	...	t-1	t
1	$X_{1,1}$	$X_{1,2}$	...	$X_{1,t-1}$	$X_{1,t}$
2	$X_{2,1}$	$X_{2,2}$	...	$X_{2,t-1}$	
...	...	...	...		
...	...	...			
t	$X_{t,1}$				

1. ábra. Kifutási háromszög

Ez alapján az $X_{i,j}$ mátrixelemek jelenthetik a következőket:

- Az i. évben bekövetkezett, $(j-1)$ év késéssel bejelentett/kifizetett/lezárt károk száma/összege.
- Az i. évben bejelentett, $(j-1)$ év késéssel kifizetett/lezárt károk száma/összege.
- Az i. évben bekövetkezett, *legfeljebb* $(j-1)$ év késéssel bejelentett/kifizett/lezárt károk száma/összege.
- Az i. évben bejelentett, *legfeljebb* $(j-1)$ év késéssel kifizetett/lezárt károk száma/összege.

Az utobbi két esetben **kumulált háromszög**ről beszélünk, a mátrixelem a lebonyolításban őt megelőző évek adatait is tartalmazza. A függőkár tartalék meghatározása éppen a mátrix hiányzó elemeinek minél pontosabb becslését kívánja.

A fejezet további részében $X_{i,j}$ jelentése az i. alapfelbontási peridódusban (évben) bekövetkezett, legfeljebb $(j-1)$ év késéssel kifizetett károk összege, vagyis: a kifutási háromszög a kumulált kárkifizetéseket tartalmazza.

Egy kifutási háromszöget **teljesnek** nevezünk, ha annak a valószínűsége, hogy a kumulált károk t év után tovább nőnek, közel van a 0-hoz, ellenkező esetben a tartalék pontos becsléséhez szükség lehet egy úgynevezett végkifutási faktor becslésére is, ami azt mutatja meg, hogy a t. lebonyolítási év *után* a teljes károk mekkora részére számíthatunk - ezt általában a korábbi évek tapasztalaiból határozzuk meg. A már teljes kifutással rendelkező adatokat jelöljük $X_{i,t+}$-al. Ha feltesszük, hogy a t. év után már nem történik kárkifizetés, akkor $X_{i,t+} = X_{i,t}$

A most bemutatásra kerülő klasszikus módszerek azzal a közös feltevéssel élnek, hogy a j. év utáni **elértségi szintek**, tehát az $X_{i,j}/X_{i,t+}$ hányadosok nem függenek (lényegesen) i-től, vagyis attól, hogy melyik bekövetkezési év adatait vizsgáljuk.

1.2. Láncszemhányados módszerek, a lánclétra módszer

Feltételezzük, hogy a $c_j(i) = \dfrac{X_{i,j+1}}{X_{i,j}}$ **növekedési faktorok** nem függenek erősen i-től és körülbelül c_j-vel egyenlők (ezt úgy is tekinthetjük, hogy $\forall\, j > 0$ lebonyolítási évhez létezik egy olyan c_j valódi növekedési faktor, amely egy valószínűségi változó, és amelynek $c_j(i)$ indexek az egyes realizációi. Ezekről a valószínűségi változókról feltesszük, hogy függetlenek, vagyis azt, hogy az adott növekedési faktor a lebonyolítási évre jellemző és nem függ a múltbeli növekedésektől. Másképp: az számít, hogy hány év késésnél tartunk, ez határozza meg a még várt növekedést, nem a korábbi növekedések).

- A $c_t = \dfrac{X_{i,t+}}{X_{i,t}}$ faktor becsléséhez használjuk a korábbi, már teljes kifutásűnak minősíthető évek adatait.
- A többi c_j faktort a ténylegesen megfigyelt $c_j(i)$ hányadosok valamilyen függvényével becsüljük.

A végső kárkifizetések becslése és a szükséges tartalékok(R_i)[1] meghatározása:

$$
\begin{aligned}
\hat{X}_{1,t+} &= c_t X_{1,t}\,, \qquad R_1 = \hat{X}_{1,t+} - X_{1,t} = (c_t - 1)X_{1,t}\,,\\
\hat{X}_{2,t+} &= c_t c_{t-1} X_{2,t-1}\,, \qquad R_2 = \hat{X}_{2,t+} - X_{2,t-1} = (c_t c_{t-1} - 1)X_{2,t-1}\,,\\
&\vdots\\
\hat{X}_{i,t+} &= c_t c_{t-1} \dots c_{t-i+1} X_{i,t-i+1}\,, \qquad (1.1)\\
R_i &= \hat{X}_{i,t+} - X_{i,t-i+1} = (c_t c_{t-1} \dots c_{t-i+1} - 1)X_{i,t-i+1}\,, \qquad (1.2)\\
&\vdots\\
\hat{X}_{t,t+} &= c_t c_{t-1} \dots c_1 X_{t,1}\,,\\
R_t &= \hat{X}_{t,t+} - X_{t,1} = (c_t c_{t-1} \dots c_1 - 1)X_{t,1}.
\end{aligned}
$$

Az alkalmazott függvény megválasztása a láncszemhányados módszer különböző változataihoz vezet:

1.változat c_j faktorokat a megfigyelt hányadosok átlagával határozzuk meg.

$$c_j = \frac{c_j(1) + c_j(2) + \dots + c_j(t-j)}{t-j} \quad j = 1, 2, \dots, t-1.$$

2.változat: c_j faktorokat a megfigyelt hányadosok *rész*átlagával határozzuk meg. A részátlag a nagyság szerint rendezett $c_j(i)$-ből a néhány legkisebb és néhány legnagyobb (pl. a legkisebb és a legnagyobb) értékek elhagyásával kapott értékek átlaga, ezzel a technikával kiszűrhetjük az ún. outlier-eket.

3.változat: c_j faktorokat a megfigyelt hányadosok maximumával becsüljük.

$$c_j = \max(c_j(1), c_j(2), \dots, c_j(t-j)) \quad j = 1, 2, \dots, t-1.$$

Ebből a módszerből egy óvatos, konzervatív becslés származik.

4.változat c_j faktorokat a megfigyelt hányadosok súlyozott átlagával határozzuk meg.

$$c_j = \frac{a_{1,j}c_j(1) + a_{2,j}c_j(2) + \dots + a_{t-j,j}c_j(t-j)}{a_{1,j} + a_{2,j} + \dots + a_{t-j,j}} \quad j = 1, 2, \dots, t-1.$$

[1] a jelölés R betűje az angol reserve - tartalék szóból

1.1 Definíció: A lánclétra-módszer esetén $a_{i,j} = X_{i,j}$, vagyis a c_j faktorok becslése:

$$c_j = \frac{X_{1,j}c_j(1) + X_{2,j}c_j(2) + \cdots + X_{t-j,j}c_j(t-j)}{X_{1,j} + X_{2,j} + \cdots + X_{t-j,j}} = \\ = \frac{X_{1,j+1} + X_{2,j+1} + \cdots + X_{t-j,j+1}}{X_{1,j} + X_{2,j} + \cdots + X_{t-j,j}}. \tag{1.3}$$

A lánclétra-módszerrel [2] meghatározott tartalékot ezentúl $R_{CL,k}$-val jelöljük.

1.3. Bornhuetter-Ferguson módszer

1.2 Definíció: Legyen

$$p_k = \frac{1}{c_t}\frac{1}{c_{t-1}}\cdots\frac{1}{c_k} = \frac{1}{\prod_{j=k}^{t} c_j} \quad \forall\, k = 1, \ldots, t. \tag{1.4}$$

Az ilyen módon definiált faktorokat a k. évre jellemző **elértségi szint**eknek nevezzük. Bevezetjük még a $q_k = 1 - p_k$ jelölést.

A definícióból rögtön következik, hogy p_k elértségi szintek megmutatják, hogy az adott lebonyolítási év után a várt teljes kár hányad részénél tartunk, összhangban azzal, ahogy eddig a $X_{i,j}/X_{i,t+}$ hányadosokra használtuk az elértségi szint megnevezést. A növekedési faktorokból 1.4 alapján az elértségi szintek számolhatóak, és megfordítva, az elértségi szintekből is számolhatóak a növekedési faktorok:

$$c_k = \frac{p_{k+1}}{p_k} \quad \forall k = 1, \ldots, t-1 \quad \text{és} \quad c_t = \frac{1}{p_t}. \tag{1.5}$$

Tehát a növekedési faktorok minden becslése becslést szolgáltat az elértségi szintekre is, valamint ugyanez igaz fordítva is.

1.3 Definíció: A Bornhuetter-Ferguson módszer a k. év tartalékszükségletéhez először megbecsüli a várt végső kárt (ezt nevezzük a végső kár előzetes becslésének): $\hat{X}_{i,t+} = P_k \cdot \hat{\kappa}_k$ ahol P_k a k. évhez tartozó díjelőírás (ezt ismerjük), $\hat{\kappa}_k$ pedig a várt végső kárhányad becslése. Ezután a k. év tartalékszükségletét a következő módon állapítjuk meg:

$$R_{BF,k} = R_k = \hat{X}_{i,t+} \cdot q_k. \tag{1.6}$$

[2] angolul chain-ladder, használni fogjuk a CL jelölést

[7]-ben egy összefoglaló írás olvasható a Bornhuetter-Ferguson módszer és más ismert módszerek (például: Cape-Cod módszer, additív-módszer) közös gyökeréről és mindezen módszerek egy általános kereteken belüli felírásáról.

1.4 Megjegyzés: A későbbiek miatt fontos kihangsúlyoznunk, hogy amíg a lánclétra módszernél 1.2-ből könnyen láthatóan a tartalékot a k. lebonyolítási évig **már megismert** károkból számoljuk, addig a Bornhuetter-Ferguson módszer esetében a várt végső kár egy előzetes, például a korábbi évek kártapasztalataiból származó becsléséből indulunk ki!

2. A Mack-modell

Ebben a fejezetben főleg Mack [1] cikkét követem.

A következőkben egy alapfelbontási periódussal foglalkozunk, most egy bekövetkezési évre koncentrálunk. X_i-vel jelöljük tehát a legfeljebb $(i-1)$ év késéssel kifizett károk összegét. Továbbá feltesszük, hogy a t. év után már nem történik kárkifizetés, vagyis az eddigi jelöléseinket használva, feltesszük, hogy:

$$c_t = \frac{X_{t+}}{X_t} = 1, \quad \text{vagy ezzel ekvivalensen} \quad p_t = 1.$$

Ismertnek vesszük az elértségi szinteket is:
Legyenek $0 < p_1 < p_2 < \cdots < p_t = 1$ rendre az i. lebonyolítási évre jellemző elértségi szintek, vagyis

$$p_i = E\left(\frac{X_i}{X_t}\right) = E\left(\frac{1}{\prod_{j=k}^{t} c_j}\right).$$ [3]

Az X_t végső kárra bevezetjük az $X_t = U$ jelölést.

Jelölje egy M módszerrel meghatározott tartalékbecslés nagyságát R_M , a végső kár becslését U_M. Ha a k. lebonyolítási év káradatai ismertek, akkor értelemszerűen fennállnak

$$U_M = X_k + R_M \quad \text{és} \quad R_M = U_M - X_k.$$

Jelölje U_0 a végső kár egy olyan előzetes becslését (pl. a korábbi évek tapasztalatait felhasználva), ami a vizsgált év már meglévő kártapasztalatát nem veszi figyelembe.
Ezzel a jelöléssel 1.3 definíció alapján a Bornhuetter-Ferguson tartalék valamely $k < n$ év után:

$$R_{BF} = q_k U_0\,, \quad \text{ahol} \quad q_k = 1 - p_k \quad \text{és} \quad U_0 = P\hat{\kappa}, \tag{2.1}$$

P az adott évre vonatkozó díjelőírás, $\hat{\kappa}$ pedig a végső kárhányad becslése.
A lánclétra-módszer 1.1 egyenlőség alapján a végső kárra a következő becslést adja:

$$U_{CL} = \frac{X_k}{p_k} \Rightarrow R_{CL} = U_{CL} - X_k = (1 - p_k)U_{CL} = q_k U_{CL}. \tag{2.2}$$

[3] korábbi megjegyzésünk szerint a növekedési faktorokra már úgy gondolunk, mint valószínűségi változókra

Ahogy 1.4 megjegyzésben már megjegyeztük, 2.1-ből jól látszik, hogy R_{BF} csak az előzetes becslésre - U_0 - épít, addig 1.2-ben U_{CL} X_k-tól függ, így R_{CL} a vizsgált év már meglévő kártapasztalatait - X_k - veszi figyelembe. Megtehetjük azonban, hogy mindkét információt bevonjuk a becslésünkbe, egy kevert tartalékolási módszer esetleg sikeresebb lehet.

2.1. Klasszikus módszerek keverése

U_0 kezdeti kárbecslés helyett nézzük a következőt:

$$U_c = cU_{CL} + (1-c)U_0\,, \quad \text{ahol} \quad 0 \le c \le 1. \tag{2.3}$$

Ebben megjelenik mind az előzetes, korábbi információn alapuló becslés (U_0), mind pedig a már meglévő kártapasztalatunk U_{CL}-en keresztül. Vegyük észre, hogy minél több lebonyolítási éven vagyunk túl, annál több információt tartalmaz a már meglévő kártapasztalat, így azt a súlyozásnál érdemes egyre nagyobb súllyal figyelembe venni.

G. Benktander ezért [2]-ben azt javasolja, legyen $c = p_k$.

2.1 Definíció: Az ilyen módon kapott módszert **Benktander-módszer**nek nevezzük, a módszerből nyerhető tartalékbecslés

$$R_{GB} = q_k U_{p_k} = q_k(\underbrace{p_k U_{CL}}_{X_k} + \underbrace{(1-p_k)U_0}_{q_k U_0 = R_{BF}}) = q_k(X_k + R_{BF}) = q_k U_{BF}. \tag{2.4}$$

Vagyis $R_{GB} = q_k U_{BF}$. Erre úgy is nézhetünk, hogy a Bornhuetter-Ferguson módszert alkalmazzuk, annyi különbséggel, hogy az U_0 előzetes kárbecslés helyett a Bornhuetter-Ferguson módszer által szolgáltatott U_{BF} becslésből indulunk ki. Emiatt hívják ezt a módszert **iterált Bornhuetter-Ferguson módszer**nek is. A GB módszer végső kárbecslése:

$$\begin{aligned} U_{GB} = X_k + R_{GB} &= U_{CL}p_k + q_k(p_k U_{CL} + q_k U_0) = \\ &= (1-q_k^2)U_{CL} + q_k^2 U_0 = U_{1-q_k^2}. \end{aligned} \tag{2.5}$$

Mi történik, ha ezt tovább iteráljuk? Erről szól a következő, indukcióval egyszerűen belátható állítás.

2.2 Állítás: Az

$$R^{(m)} = q_k U^{(m)}, \quad U^{(m+1)} = X_k + R^{(m)}$$

iteráció tetszőleges $U_0 = U^{(0)}$ kezdőértékre a következőt adja:

$$\begin{aligned} U^{(m)} &= (1-q_k^m)U_{CL} + q_k^m U_0\,, \\ R^{(m)} &= (1-q_k^m)R_{CL} + q_k^m R_{BF}. \end{aligned}$$

BIZONYÍTÁS: Tegyük fel, hogy az állítást már tudjuk $(m-1)$-re, belátjuk m-re is.

$$\begin{aligned} U^{(m)} &= X_k + R^{(m-1)} = U_{CL}p_k + q_k U^{(m-1)} \stackrel{1}{=} \\ &= U_{CL}p_k + q_k[(1-q_k^{m-1})U_{CL} + q_k^{m-1}U_0] = \\ &= (p_k + q_k - q_k^m)U_{CL} + q_k^m U_0 = (1-q_k^m)U_{CL} + q_k^m U_0\,, \end{aligned}$$

1-nél használtuk az indukciós feltevést. A másik állítás teljesen hasonlóan igazolható. ■

$k < n$ esetén $0 < p_k < 1 \;\Rightarrow 0 < q_k < 1$, amiből

$$\lim_{m\to\infty} q_k^m = 0\,, \quad \text{tehát} \lim_{m\to\infty} R^{(m)} = R_{CL}\,,$$

vagyis az iteráció m $\to \infty$ esetén a CL módszerhez tart.

Ráadásul az iteráció meglehetősen gyorsan tart a CL módszerhez, még nem túl magas elértségi szintek esetén is, ezt mutatja a következő táblázat:

	$(1-q_k^m)$ értéke, ha m értéke				
p_k	3	4	5	6	8
0,4	0,784	0,870	0,922	0,953	0,983
0,5	0,875	0,938	0,969	0,984	0,996
0,6	0,936	0,974	0,990	0,999	
0,7	0,973	0,992	0,998	0,999	
0,8	0,992	0,998	0,999		

1. táblázat. A CL módszer súlya az iteráció lépései során

2.2. Optimális súly - a Mack-modell

Ebben az alfejezetben Mack [1]-beli eredményeit ismertetem egy saját megjegyzéssel kiegészítve.

2.3 Definíció: Az R tartalékszükséglet egy M módszerrel való R_M becslésének átlagos négyzetes eltérésén a következőt értjük:

$$E(R_M - R)^2 = \text{mse}(R_M).\ ^4 \tag{2.6}$$

Legyen 2.3-hoz hasonlóan

$$R_c = cR_{CL} + (1-c)R_{BF}, \quad \text{ahol} \quad 0 \leq c \leq 1,$$

vagyis az alkalmazott M módszerünk az, hogy valamilyen c súllyal keverjük a CL és a BF módszert.

$$R_c = c(R_{CL} - R_{BF}) + R_{BF}.$$

Ez a kifejezés c-ben lineáris, így mse(R_c) c-nek másodfokú függvénye, vagyis létezik c-ben minimuma. Célünk ezen optimális c súly meghatározása.

A levezetéshez feltesszük, hogy U_0 előzetes becslés független X_k-től (vagyis nem használja az adott évi kártapasztalatot), független R-től és U-tól, valamint fennáll, hogy E(U_0)=E(U), vagyis az előzetes becslésre használt módszer várható értékben "pontos" a végső kárra. U_0 szórását jelöljük Var(U_0)-al. Ezen feltevések mellett igaz a következő

2.4 Tétel: Az mse(R_c)-t minimalizáló optimális c^* súly

$$c^* = \frac{p_k}{q_k} \cdot \frac{Cov(X_k, R) + p_k\, q_k\, Var(U_0)}{Var(X_k) + p_k^2\, Var(U_0)}. \tag{2.7}$$

Bizonyítás:

$$\begin{aligned} E(R_c - R)^2 &= E[c(R_{CL} - R_{BF}) + R_{BF} - R]^2 = \\ &= c^2 E(R_{CL} - R_{BF})^2 - 2cE[(R_{CL} - R_{BF})(R - R_{BF})] + \\ &+ E(R_{BF} - R)^2. \end{aligned}$$

Ezt c szerint deriválva :

$$\frac{\partial}{\partial c} E(R - R_c)^2 = 2cE(R_{CL} - R_{BF})^2 - 2E[(R_{CL} - R_{BF})(R - R_{BF})].$$

[4] mse - mean squared error

Ezt 0-ra rendezve az optimális c^* súly:

$$c^* = \frac{E[(R_{CL} - R_{BF})(R - R_{BF})]}{E(R_{CL} - R_{BF})^2} = \frac{E[\left(\frac{q_k}{p_k}X_k - q_kU_0\right)(R - q_kU_0)]}{E(\frac{q_k}{p_k}X_k - q_kU_0)^2} =$$

$$= \frac{\frac{q_k}{p_k} \cdot E[(X_k - p_kU_0)(R - q_kU_0)]}{\left(\frac{q_k}{p_k}\right)^2 E(X_k - p_kU_0)^2} = \frac{p_k}{q_k} \cdot \frac{E[(X_k - p_kU_0)(R - q_kU_0)]}{E(X_k - p_kU_0)^2} =$$

$$= \frac{p_k}{q_k} \cdot \frac{Cov[(X_k - p_kU_0), (R - q_kU_0)]}{Var(X_k - p_kU_0)} = \frac{p_k}{q_k} \cdot \frac{Cov(X_k, R) + p_kq_kVar(U_0)}{Var(X_k) + p_k^2Var(U_0)}.$$

Az utolsó előtti lépésben azt használtuk ki, hogy

$$E(X_k - p_kU_0) = E(X_k) - p_kE(U_0) = X_k - p_kE(U) = 0$$

p_k definíciója szerint, és $E(R - q_kU_0) = 0$ hasonlóan.

Az utolsó lépésben pedig felhasználtuk, hogy U_0 független R-től és X_k-tól, így $Cov(U_0, R) = Cov(U_0, X_k) = 0$ és $Var(X_k - p_kU_0)$ is felírható a független valószínűségi változók varianciáinak összegeként. ■

Ahhoz, hogy becslést tudjunk adni c^*-ra, a 2.7-ben szereplő $\mathrm{Var}(X_k)$-ról és $\mathrm{Cov}(X_k, R)$-ről kell tudnunk mondani valamit. Az X_k/U feltételes eloszlására élünk feltevéssel, az első feltevés lényegében ekvivalens azzal, ahogy p_k-ra eddig gondoltunk, a második feltétel pedig a feltételes eloszlás szórására követel meg egy paraméteres alakot (a későbbiekben ezekre fogunk úgy hivatkozni, mint a *Mack-modell kiinduló feltételei*):

$$E\left(\frac{X_k}{U} \,|\, U\right) = p_k\,, \tag{2.8}$$

$$Var\left(\frac{X_k}{U} \,|\, U\right) = p_kq_k\beta^2(U). \tag{2.9}$$

Bevezetjük még az $\alpha^2(U) = U^2 \cdot \beta^2(U)$ jelölést.
2.8-ból és 2.9-ből egyszerűen kapjuk a következőket:

$$E(X_k \,|\, U) = p_kU, \tag{2.10}$$

$$Var(X_k \,|\, U) = p_kq_k\alpha^2(U). \tag{2.11}$$

Várható értéket véve 2.10-ben:

$$E(E(X_k \,|\, U)) = E(X_k) = p_kE(U). \tag{2.12}$$

2.11-re felírva a teljes szórásnégyzet tételét, Var(X_k)-ra adódik:

$$\begin{aligned} Var(X_k) &= E(Var(X_k \,|\, U)) + Var(E(X_k \,|\, U)) = \\ &= p_k q_k E(\alpha^2(U)) + p_k^2 Var(U) = \\ &= p_k E(\alpha^2(U)) + p_k^2 \left(Var(U) - E(\alpha^2(U))\right). \end{aligned} \tag{2.13}$$

Cov(X_k, R)-re:

$$Cov(X_k, R) = Cov(X_k, U - X_k) = Cov(X_k, U) - Var(X_k). \tag{2.14}$$

$$Cov(X_k, U) = Cov(E(X_k \,|\, U), U) = p_k Var(U), \tag{2.15}$$

ezt és 2.13-at beírva 2.14-be:

$$\begin{aligned} Cov(X_k, R) &= p_k Var(U) - \left[p_k E(\alpha^2(U)) + p_k^2 \left(Var(U) - E(\alpha^2(U))\right)\right] = \\ &= (p_k - p_k^2) Var(U) - p_k q_k E(\alpha^2(U)) = \\ &= p_k(1 - p_k) Var(U) - p_k q_k E(\alpha^2(U)) = \\ &= p_k q_k \left(Var(U) - E(\alpha^2(U))\right). \end{aligned} \tag{2.16}$$

Írjuk be Var(X_k)-ra és Cov(X_k, R)-re kapott 2.13-as és 2.16-os kifejezéseket az optimális c^*-ot megadó 2.7-es formulánkba, ebből közvetlen adódik a

2.5 Tétel: 2.8 és 2.9 feltevésekkel élve a $mse(R_c)$-t minimalizáló optimális c^* súly:

$$c^* = \frac{p_k}{p_k + t}, \quad \text{ahol} \quad t = \frac{E(\alpha^2(U))}{Var(U_0) + Var(U) - E(\alpha^2(U))}. \tag{2.17}$$

Megjegyezzük még, hogy:

$$E(R) = E(U - X_k) = E(U) - E(X_k) = q_k E(U).$$

A korábbiak felhasználásával pedig

$$\begin{aligned} Var(R) &= Var(U) - 2Cov(X_k, U) + Var(X_k) = \\ &= Var(U)(1 - 2p_k + p_k^2) + p_k q_k E(\alpha^2(U)) = \\ &= q_k^2 Var(U) + p_k q_k E(\alpha^2(U)) = \\ &= q_k E(\alpha^2(U)) + q_k^2 \left(Var(U) - E(\alpha^2(U))\right). \end{aligned}$$

Mennyi az átlagos négyzetes eltérés minimumát adó $mse(R_{c^*})$? Általában fogjuk felírni $mse(R_c)$-t minden $0 \leq c \leq 1$ -re:

$$\begin{aligned} mse(R_c) &= E(cR_{CL} + (1-c)R_{BF} - R)^2 = \\ &= E[c(R_{CL} - R) + (1-c)(R_{BF} - R)]^2 = \\ &= c^2 mse(R_{CL}) + 2c(1-c)\underbrace{E[(R_{CL} - R)(R_{BF} - R)]}_{A} + \\ &+ (1-c)^2 mse(R_{BF}). \end{aligned}$$

$$\begin{aligned} A &= E[(R_{CL} - R)(R_{BF} - R)] \overset{1}{=} Cov(R_{CL} - R, R_{BF} - R) \overset{2}{=} \\ &= -Cov(R_{CL}, R) + Var(R) = Var(R) - \frac{q_k}{p_k} Cov(X_k, R) = \\ &= \underbrace{q_k E(\alpha^2(U)) + q_k^2 \left(Var(U) - E(\alpha^2(U))\right)}_{Var(R)} - \frac{q_k}{p_k} \underbrace{p_k q_k \left(Var(U) - E(\alpha^2(U))\right)}_{Cov(X_k, R)} \\ &= q_k E(\alpha^2(U)). \end{aligned}$$

$$\Downarrow$$

$$mse(R_c) = c^2 mse(R_{CL}) + 2c(1-c)q_k E(\alpha^2(U)) + (1-c)^2 mse(R_{BF}). \quad (2.18)$$

Ezzel kifejeztük a c súllyal kevert R_c módszer átlagos négyzetes hibáját a klasszikus módszerek - CL és BF - négyzetes hibáival.

2.6 Megjegyzés: A keverék négyzetes hibájának a kevert módszerek négyzetes hibáiával való levezetésében csak azt használtuk, mégpedig az 1-essel jelölt egyenlőségnél, hogy a két kiindulási módszerünk várható értékben pontos (nevezzük ezeket X és Y módszernek), azaz $E(R_X) = E(R) = E(R_Y)$, illetve a 2-essel jelölt egyenlőségnél azt, hogy a végső kár előzetes becslése független R-től és X_k-tól. Ez esetben tehát általában igaz, hogy ha a kiinduló 2.8 és 2.9 feltételek teljesülnek, akkor az

$$R_d = dR_X + (1-d)R_Y \text{ módszerre fennáll } (0 \leq d \leq 1):$$

$$mse(R_d) = d^2 mse(R_X) + 2d(1-d)q_k E(\alpha^2(U)) + (1-d)^2 mse(R_Y).$$

A klasszikus módszerek négyzetes hibái:

$$
\begin{aligned}
mse(R_{CL}) &= E(R_{CL}-R)^2 = Var(R_{CL}-R) = Var(R_{CL}) - \\
&- 2Cov(R_{CL},R) + Var(R) = \left(\frac{q_k}{p_k}\right)^2 Var(X_k) - 2\frac{q_k}{p_k}Cov(X_k,R) + \\
&+ Var(R) = \left(\frac{q_k}{p_k}\right)^2 [p_k E(\alpha^2(U)) + p_k^2\left(Var(U) - E(\alpha^2(U))\right)] - \\
&- 2\frac{q_k}{p_k}p_k q_k\left(Var(U) - E(\alpha^2(U))\right) + q_k E(\alpha^2(U)) + \\
&+ q_k^2\left(Var(U) - E(\alpha^2(U))\right) = \left(\frac{q_k^2}{p_k} + q_k\right)E(\alpha^2(U)) = \frac{q_k}{p_k}E(\alpha^2(U)).
\end{aligned}
\tag{2.19}
$$

$$
\begin{aligned}
mse(R_{BF}) &= E(R_{BF}-R)^2 = Var(R_{BF}-R) \overset{3}{=} Var(R_{BF}) + Var(R) = \\
&= q_k^2 Var(U_0) + q_k E(\alpha^2(U)) + q_k^2\left(Var(U) - E(\alpha^2(U))\right) \\
&= q_k^2 \underbrace{\left[Var(U_0) + \left(Var(U) - E(\alpha^2(U))\right)\right]}_{\frac{E(\alpha^2(U))}{t}} + q_k E(\alpha^2(U)) \\
&= E(\alpha^2(U))\left(q_k + \frac{q_k^2}{t}\right).
\end{aligned}
\tag{2.20}
$$

3-nál ismét használtuk, hogy az előzetes becslés, amiből R_{BF} származik független R-től. Írjuk be a $mse(R_c)$-t megadó 2.18-ba a klasszikus módszerek négyzetes hibáit megadó kifejezéseinket:

$$
\begin{aligned}
mse(R_c) &= c^2\frac{q_k}{p_k}E(\alpha^2(U)) + 2c(1-c)q_k E(\alpha^2(U)) + \\
&+ (1-c)^2 E(\alpha^2(U))\left(q_k + \frac{q_k^2}{t}\right) = \\
&= E(\alpha^2(U))\left[c^2\frac{q_k}{p_k} + 2c(1-c)q_k + (1-c)^2\left(q_k + \frac{q_k^2}{t}\right)\right] = \\
&= E(\alpha^2(U))\left[c^2\left(\frac{p_k}{q_k} - q_k\right) + q_k + q_k^2\frac{(1-c)^2}{t}\right] \\
&= E(\alpha^2(U))q_k^2\left[\frac{c^2}{p_k} + \frac{1}{q_k} + \frac{(1-c)^2}{t}\right].
\end{aligned}
\tag{2.21}
$$

$E(\alpha^2(U))$ és t nem függ a p_k elértségi szinttől. Rögzített c súlyra a $q_k^2 \cdot \dfrac{c^2}{p_k}$ és a $q_k^2 \cdot \dfrac{1}{q_k} = q_k$ is csökken, ha p_k nő. Vagyis: minél magasabb elértségi szintnél tartunk, minél több információt hordoz a már meglévő kártapasztalat, $\mathrm{mse}(R_c)$ annál kisebb, annál kisebb négyzetes hibával tudjuk becsülni a valódi tartalékszükségletet. (Természetesen magasabb elértségi szintnél a tartalékszükséglet is alacsonyabb, eleve egy kisebb értéket kell becsülni.)

3. Gamma eloszlású kárhányad

Az előző fejezetben levezettük, milyen optimális c^* súllyal érdemes keverni a CL és BF módszereket, hogy a keveréssel kapott R_{c^*} módszer minimalizálja mse(R_c)-t és ki is fejeztük a kevert R_c módszer négyzetes hibáját. A levezetéshez különböző feltevésekkel éltünk, melyek közül a legfontosabbak az $\frac{X_k}{U}$ hányados feltételes eloszlására vonatkoztak. Megmutatjuk, hogy a felhasznált 2.8 és 2.9 feltételek már viszonylag természetes esetben is fennállnak.

3.1. Kiinduló feltételek teljesülése

Az egyszerűbb írásmód érdekében az adatok most nem kumuláltak, vagyis Z_i jelenti azon károknak az összegét, melyeket pontosan $(i-1)$ év késéssel fizettek ki. Ez esetben feltesszük, hogy a kifutási háromszög egy oszlopára jellemző kárhányadok Γ eloszlásúak, vagyis a pontosan $(i-1)$ év késéssel kifizett károkra fennáll: $Z_i \sim \Gamma(\alpha_i, \lambda) \cdot P$ ahol P az alapfelbontáshoz tartozó előírt díj.

Általában igaz, hogy független Γ eloszlások esetén:

$$\sum_{i=1}^{k} \Gamma(n_i, \lambda) = \Gamma\left(\sum_{i=1}^{k} n_i, \lambda\right).$$

Az oszlopoknak megfelelő eloszlások függetlenségét feltesszük (lásd 1.2-ben a növekedési faktorokról leírtakat), így az $X_k = \sum_{i=1}^{k} Z_i$ kumulált kárkifizetésekre:

$$X_k \sim \Gamma\left(\sum_{i=1}^{k} \alpha_i, \lambda\right) \cdot P.$$

A végső kárra pedig:

$$U \sim \Gamma\left(\sum_{i=1}^{t} \alpha_i, \lambda\right) \cdot P = \Gamma(\alpha, \lambda) \cdot P, \qquad \alpha = \sum_{i=1}^{t} \alpha_i \text{-vel.}$$

A következőkben megmutatjuk, hogy Γ eloszlású kárhányadot feltéve teljesülnek a Mack-modell kiinduló feltételei, vagyis a X_k/U hányados feltételes várható értékére és feltételes varianciájára fennállnak:

$$E\left(\frac{X_k}{U} \,|\, U\right) = p_k, \quad \text{ahol} \quad p_k = \frac{\sum_{i=1}^{k} \alpha_i}{\sum_{i=1}^{t} \alpha_i} = \frac{\sum_{i=1}^{k} \alpha_i}{\alpha} \quad \text{és} \tag{3.1}$$

$$Var\left(\frac{X_k}{U}|U\right) = p_k q_k \beta^2(U), \quad \text{ahol} \quad q_k = 1 - p_k = \frac{\sum_{i=k+1}^{t} \alpha_i}{\sum_{i=1}^{t} \alpha_i} = \frac{\sum_{i=k+1}^{t} \alpha_i}{\alpha}. \tag{3.2}$$

Vezessük be a $D_k = \sum_{i=k+1}^{t} Z_i$ jelölést.

$$X_k + D_k = \sum_{i=1}^{k} Z_i + \sum_{i=k+1}^{t} Z_i = \sum_{i=1}^{t} Z_i = U. \tag{3.3}$$

Feltevésünk szerint $X_k \sim \Gamma(\alpha p_k, \lambda) \cdot P$ valamint $D_k \sim \Gamma(\alpha q_k, \lambda) \cdot P$. Az oszlopoknak megfelelő valószínűségi változókról feltételeztük a függetlenséget, tehát X_k és D_k is függetlenek. Vezessük be az $X = X_k$, $Y = X_k + D_k$ jelöléseket. (3.3 miatt Y=U.)

Az $X|Y$ feltételes eloszlására szeretnénk igazolni 3.1-et és 3.2-őt, ehhez meg kell határoznunk az $X|Y$ feltételes eloszlást.

$$f_{X\,|\,Y}(x\,|\,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

Vagyis szükségünk lesz X és Y együttes sűrűségfüggvényére. X_k és D_k függetlenek és sűrűségfüggvényüket ismerjük, ezt fogjuk kihasználni a következő általános tétel segítségével:

3.1 Tétel: Ha $F : \Omega \to \mathbb{R}^k$ abszolút folytonos eloszlású valószínűségi változó, ahol valamely $V \subseteq \mathbb{R}^k$ nyílt halmazra $P(F \in V) = 1$ továbbá: $g : V \to W$ $W \subseteq \mathbb{R}^k$ W nyílt, g diffeomorfizmus, akkor $G = g \circ F$ is abszolút folytonos eloszlású és sűrűségfüggvénye:

$$f_G = (f_F \circ g^{-1}) \det(Jacobi\,(g^{-1})).$$

A mi esetünkben

$$F := (X_k, D_k) \quad F : \Omega \to \mathbf{R}^2,$$
$$g(X_k, D_k) := (X_k, X_k + D_k) = (X, Y).$$

$X_k = X$ és $D_k = Y - X$ ezért $g^{-1}(X, Y) = (X, Y - X)$. g könnyen láthatóan diffeomorfizmus, g^{-1} Jacobi-mátrixa:

$$J_{g^{-1}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix},$$
$$\det J_{g^{-1}} = 1 - 0 = 1.$$

Alkalmazzuk a 3.1 tételt:

$$f_{X,Y}(x,y) = \frac{\lambda^{\alpha p_k}}{\Gamma(\alpha p_k)}\, x^{\alpha p_k - 1}\, e^{-\lambda x} \cdot \frac{\lambda^{\alpha q_k}}{\Gamma(\alpha q_k)} (y-x)^{\alpha q_k - 1}\, e^{-\lambda(y-x)} \det J_{g^{-1}}\,,$$

$$f_{X,Y}(x,y) = \frac{\lambda^{\alpha}}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, x^{\alpha p_k - 1}\, (y-x)^{\alpha q_k - 1}\, e^{-\lambda y}. \tag{3.4}$$

$$f_{X\,|\,Y}(x\,|\,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\dfrac{\lambda^{\alpha}}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, x^{\alpha p_k - 1}\, (y-x)^{\alpha q_k - 1}\, e^{-\lambda y}}{\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}\, y^{\alpha - 1}\, e^{-\lambda y}} =$$

$$= \frac{\Gamma(\alpha)}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, x^{\alpha p_k - 1}\, (y-x)^{\alpha q_k - 1}\, y^{1-\alpha}. \tag{3.5}$$

Ebből

$$\begin{aligned}
E(X\,|\,Y=y) &= \int x f(x\,|\,y) dx = \\
&= \int \frac{\Gamma(\alpha)}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, x\, x^{\alpha p_k - 1}\, (y-x)^{\alpha q_k - 1}\, y^{1-\alpha}\, dx = \\
&= \quad \ldots \text{x=ys helyettesítéssel} \ldots \quad = \\
&= \int \frac{\Gamma(\alpha)}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, (ys)^{\alpha p_k}\, (y-ys)^{\alpha q_k - 1}\, y^{1-\alpha}\, y\, ds = \\
&= \int \frac{\Gamma(\alpha)}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, y^{\alpha p_k}\, y^{\alpha q_k - 1}\, y^{1-\alpha}\, s^{\alpha p_k}\, (1-s)^{\alpha q_k - 1}\, y\, ds = \\
&= y^{\alpha-1} y^{1-\alpha} y \int \underbrace{\frac{\Gamma(\alpha)}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, s^{\alpha p_k}\, (1-s)^{\alpha q_k - 1}}_{B}\, ds = \\
&= y \left(1 \cdot \frac{\Gamma(\alpha)}{\Gamma(\alpha+1)} \frac{\Gamma(\alpha p_k + 1)}{\Gamma(\alpha p_k)} \right).
\end{aligned} \tag{3.6}$$

A B-vel jelölt függvény Beta$(\alpha p_k + 1, \alpha q_k)$ eloszlás sűrűségfüggvénye konstanstól eltekintve, a megfelelő konstanssal beszoroztunk és leosztottunk. Felhasználva, hogy $\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$ minden α esetén fennáll, kapjuk, hogy:

$$E(X\,|\,Y=y) = y \cdot \frac{1}{\alpha} \cdot \alpha p_k = p_k y.$$

Vagyis $E(X\,|\,Y) = p_k Y \;\Rightarrow\; E\left(\dfrac{X}{Y}\,|\,Y\right) = p_k$, amivel (3.1)-et igazoltuk. Továbbá:

$$Var(X\,|\,Y) = E((X - E(X\,|\,Y))^2|Y) = E(X^2\,|\,Y) - E(X\,|\,Y)^2. \qquad (3.7)$$

$$\begin{aligned}
E(X^2\,|\,Y = y) &= \int x^2 f(x\,|\,y)dx \\
&= \int \frac{\Gamma(\alpha)}{\Gamma(\alpha p_k)\Gamma(\alpha q_k)}\, x^2\, x^{\alpha p_k - 1}\,(y-x)^{\alpha q_k - 1}\, y^{1-\alpha}\, dx = \\
&= \quad \ldots \text{az előzőhöz hasonló átalakításokat alkalmazva} \ldots \quad = \\
&= y^2 \cdot \frac{\alpha p_k(\alpha p_k + 1)}{\alpha(\alpha+1)}, \qquad (3.8)
\end{aligned}$$

$$\Rightarrow E(X^2\,|\,Y) = Y^2 \cdot \frac{\alpha p_k(\alpha p_k + 1)}{\alpha(\alpha+1)}.$$

Írjuk be ezt és a már igazolt 3.1-et 3.7-be:

$$\begin{aligned}
Var(X\,|\,Y) &= Y^2\left(\frac{\alpha p_k(\alpha p_k+1)}{\alpha(\alpha+1)}\right) - (p_k Y)^2 = Y^2\left(\frac{\alpha p_k(\alpha p_k+1)}{\alpha(\alpha+1)} - p_k^2\right) = \\
&= Y^2\left(\frac{\alpha^2 p_k^2 + \alpha p_k}{\alpha^2+\alpha} - p_k^2\right) = Y^2\left(\frac{\alpha p_k^2 + p_k}{\alpha+1} - p_k^2\right) = \\
&= Y^2\left(\frac{\alpha p_k^2 + p_k - \alpha p_k^2 - p_k^2}{\alpha+1}\right) = Y^2\left(\frac{p_k - p_k^2}{\alpha+1}\right) = \\
&= Y^2\left(\frac{p_k(1-p_k)}{\alpha+1}\right) = Y^2\left(\frac{p_k q_k}{\alpha+1}\right),
\end{aligned}$$

$\Rightarrow Var\left(\frac{X}{Y}\,|\,Y\right) = \frac{p_k q_k}{\alpha+1}$, amivel 3.2-őt igazoltuk, vagyis teljesülnek a Mack-modell kiinduló feltételei $\beta^2(Y) = \beta^2(U) = \frac{1}{\alpha+1}$ U-tól nem függő konstanssal.

3.2. Az elméleti optimális súly

Emlékeztetünk, hogy a Mack-modellből nyert mse(R_c)-t minimalizáló optimális tartalék

$$R_{c^*} = c^* R_{CL} + (1 - c^*) R_{BF},$$

az optimumot adó c^* súly értéke pedig:

$$c^* = \frac{p_k}{p_k + t}, \quad \text{ahol} \quad t = \frac{E(\alpha^2(U))}{Var(U_0) + Var(U) - E(\alpha^2(U))}.$$

Ahhoz, hogy az optimális c^* súlyt megkapjuk, szükségünk lesz Var(U), Var(U_0) és $E(\alpha^2(U)) = E(U^2)\beta^2(U) = E(U^2)\beta^2$ meghatározására. [5]

Feltételezésünk szerint most $U \sim \Gamma(\alpha, \lambda) \cdot P$. Ekkor fennállnak:

$$E(U) = \frac{\alpha}{\lambda} \cdot P,$$

$$Var(U) = \frac{\alpha}{\lambda^2} \cdot P^2, \tag{3.9}$$

$$E(U^2) = Var(U) + E(U)^2 = \frac{\alpha}{\lambda^2} \cdot P^2 + \left(\frac{\alpha}{\lambda} \cdot P\right)^2 = \frac{\alpha(\alpha+1)}{\lambda^2} \cdot P^2. \tag{3.10}$$

U_0 előzetes becslésnek a következőt választjuk: nézzük a korábbi, már teljes kifutással rendelkező éveket és a végső kárhányadok átlaga lesz a végső kárhányad becslése. Ezt megszorozva az aktuális év előírt díjával becslést kapunk az aktuális évhez tartozó végső kárra. Általánosan felírva a becslést, tegyük fel, hogy a vizsgált évet megelőzően l évre ismert a teljes kárkifutás, vagyis ismertek a

$$\begin{array}{llll} P_{-l}, & P_{-(l-1)}, & \ldots, & P_{-1} \quad \text{díjelőírások és a} \\ U_{-l}, & U_{-(l-1)}, & \ldots, & U_{-1} \quad \text{végső kárkifizetések.} \end{array}$$

A korábbi évek kárhányad eloszlásairól az aktuális évével megegyező kárhányad eloszlást feltételezünk, speciálisan a végső kárkifizetésekre: $U_{-i} \sim \Gamma(\alpha, \lambda) \cdot P_{-i} \quad \forall\ i = 1 \ldots l$ esetén. Az átlag torzítatlan becslés a várható értékre, ezért

$$\hat{m} = \frac{\sum_{i=1}^{l} \frac{U_{-i}}{P_i}}{l} \quad \text{torzítatlan becslés} \quad m = \frac{\alpha}{\lambda} \text{ -ra.} \tag{3.11}$$

Az előzetes becslésünk a végső kárra: $U_0 = \hat{m} \cdot P$, ahol P a vizsgált év díjelőírása. Ez a becslés független a vizsgált alapfelbontási évhez tartozó, már megismert C_k kártól. [6]

[5] az utolsó egyenlőség azért teljesül, mert az imént beláttuk, hogy Γ eloszlású kárhányadot feltételezve $\beta^2(U)$ nem függ U-tól.

[6] ezt a függetlenséget kihasználtunk a Mack-modell levezetésében, tehát szükséges, hogy így legyen

$$
\begin{aligned}
Var(U_0) &= P^2 \cdot Var(\hat{m}) = \\
&= P^2 \cdot E\left(\frac{\sum_{i=1}^{l} \frac{U_{-i}}{P_i}}{l} - m\right)^2 = P^2 \cdot E\left(\frac{\sum_{i=1}^{l} \frac{U_{-i}}{P_i} - ml}{l}\right)^2 = \\
&= \left(\frac{P}{l}\right)^2 E\left(\sum_{i=1}^{l} \frac{U_{-i}}{P_i} - ml\right)^2 = \\
&\overset{1}{=} \left(\frac{P}{l}\right)^2 \cdot l \cdot E\left(\frac{U_{-i}}{P_{-i}} - m\right)^2 \overset{2}{=} \frac{P^2}{l} \cdot \frac{\alpha}{\lambda^2}.
\end{aligned} \tag{3.12}
$$

Az 1-essel jelölt egyenlőség azért igaz, mert $\frac{U_{-i}}{P_{-i}}$ -k m várható értékű valószínűségi változók, így ott éppen az összeg szórásnégyzete szerepel, továbbá független, azonos eloszlásúak, ezért az összeg szórásnégyzete a szórásnégyzetek összege.

A 2-essel jelölt egyenlőségnél pedig azt használtuk, hogy éppen a $\Gamma(\alpha, \lambda^2)$ eloszlású $\frac{U_{-i}}{P_{-i}}$ szórásnégyzete szerepel, ami $\frac{\alpha}{\lambda^2}$.

3.9-et, 3.10-et és 3.12-őt beírva a t-t meghatározó kifejezésbe, kapjuk:

$$
\begin{aligned}
t &= \frac{E(\alpha^2(U))}{Var(U_0) + Var(U) - E(\alpha^2(U))} = \frac{E(U^2)\beta^2}{Var(U_0) + Var(U) - E(U^2)\beta^2} = \\
&= \frac{\frac{\alpha(\alpha+1)}{\lambda^2} P^2 \frac{1}{\alpha+1}}{\frac{P^2}{l}\frac{\alpha}{\lambda^2} + \frac{\alpha}{\lambda^2} P^2 - \frac{\alpha(\alpha+1)}{\lambda^2} P^2 \frac{1}{\alpha+1}} = \frac{\frac{\alpha}{\lambda^2}}{\frac{\alpha}{l\lambda^2}} = l,
\end{aligned}
$$

$$
\text{amiből} \quad c^* = \frac{p_k}{p_k + t} = \frac{p_k}{p_k + l}. \tag{3.13}
$$

A következő megfigyeléseket tehetjük:

- Ha l nő, akkor c^* csökken, vagyis kisebb súlyt kap az aktuális évből számolt tartalék, összhangban azzal, hogy nagyobb l esetén az előzetes becslés több információn alapul.

- Mivel $0 \leq p_k \leq 1$, ezért c^* értéke még magas elértségi szint és kevés teljes kifutással rendelkező korábbi év esetén sem túl magas, vagyis kicsi súlyt kap R_{CL} és nagy súlyt kap R_{BF}. Ez persze nem meglepő: **feltettük, hogy a korábbi évek kárhányadainak eloszlása megegyezik a vizsgált évével**, így az előzetes becslés nemcsak torzítatlan, hanem l-lel fordítottan arányos szórásnégyzetű. Emiatt ez ő szórásnégyete jóval kisebb tud lenni, mint $\frac{X_k}{P}$ szórásnégyzete, a BF módszer pedig ebből a kisebb szórású előzetes becslésből, a CL módszer pedig éppen X_k-ból indul ki. Ilyen feltételek mellett tehát R_{BF} várhatóan nagyon erős lesz.

Egy konkrét numerikus példán szemléltetjük a jelenséget:
Ha $\frac{U_i}{P_i} \sim \Gamma(12, 20)$ - a végső kárhányad várható értéke tehát 60 százalék - továbbá $p_k = 0,5$ és $l = 5 \Rightarrow \frac{X_k}{P} \sim \Gamma(0,5 \cdot 12, 20)$.

$$Var\left(\frac{U_0}{P}\right) = \frac{12}{20^2} \cdot \frac{1}{5} = \frac{12}{2000}, \quad Var\left(\frac{X_k}{P}\right) = \frac{0,5 \cdot 12}{20^2} = \frac{30}{2000},$$

vagyis R_{CL} nagyobb szórású adatból indul ki, mint R_{BF}, természetes tehát - ezen feltevések mellett -, hogy R_{BF} kapja a nagyobb súlyt a keverés során, valamint várhatóan pontosabb becslést tud adni a tartalékszükségletre. Az elméleti optimális súlyokat tartalmazza a 2. ábra táblázata.

	l - teljes kifutású évek száma								
p_k	2	3	4	5	6	7	8	9	10
0,2	0,090909	0,0625	0,047619	0,038462	0,032258	0,027778	0,02439	0,021739	0,019608
0,3	0,130435	0,090909	0,069767	0,056604	0,047619	0,041096	0,036145	0,032258	0,029126
0,4	0,166667	0,117647	0,090909	0,074074	0,0625	0,054054	0,047619	0,042553	0,038462
0,5	0,2	0,142857	0,111111	0,090909	0,076923	0,066667	0,058824	0,052632	0,047619
0,6	0,230769	0,166667	0,130435	0,107143	0,090909	0,078947	0,069767	0,0625	0,056604
0,7	0,259259	0,189189	0,148936	0,122807	0,104478	0,090909	0,08046	0,072165	0,065421
0,8	0,285714	0,210526	0,166667	0,137931	0,117647	0,102564	0,090909	0,081633	0,074074
0,9	0,310345	0,230769	0,183673	0,152542	0,130435	0,113924	0,101124	0,090909	0,082569

2. ábra. elméleti optimális súlyok

3.3. Mack-modell a gyakorlatban

A 2. fejezetben az optimális súly és a különböző módszerek négyzetes hibái levezetésében p_k elértéségi szinteket ismertnek vettük, konstansként kezeltük. Majd megnéztük, hogy Γ eloszlású kárhányadból kiindulva - miután beláttuk, hogy ekkor teljesülnek a modell kiinduló feltételei - mit ad a levezetett optimális súly, ezt előrelátóan elméleti optimális súlynak hívtuk.

A valóságban ezeket a $p_k = E(X_k/U)$ értékeket nem ismerjük. Viszont a már megismert káradatokból tudjuk őket becsülni, jelöljük a becsléseket $\hat{p_k}$-pal, p_k minden becslése c^* egy becslését szolgáltatja:

$$\hat{c^*} = \frac{\hat{p_k}}{\hat{p_k} + \hat{t}}.$$

Γ eloszlású kárhányad esetén (mi továbbra is ezzel foglalkozunk)

$$\hat{c^*} = \frac{\hat{p_k}}{\hat{p_k} + l}.$$

Az egyszerűség kedvért a továbbiakban $\forall\, i$-re P_i-t és P-t 1-nek választjuk, ezzel elérjük, hogy az $X_{i,j}/P_i$ kárhányadok helyett $X_{i,j}$ károk eloszlása lesz Γ eloszlású. Az optimális súlyban P nem szerepel, az előzetes becslést U_{-j}/P_{-j} átlagai adják, de itt P_{-j} egyenlőnek való választása (speciálisan: azonosan egynek) csak annyit jelent, hogy stabil, nagyjából ugyanakkora portfóliónk van minden évben, vagyis a biztosító által viselt kockázat mértéke nem változik, hiszen a díj a kockázati kitettség egy mértéke. [7]

Célunk annak megvizsgálása, hogy az ezzel a $\hat{c^*}$-val megadott kevert $R_{\hat{c^*}}$ tartalék átlagos négyzetes hibája hogyan viszonyul a klasszikus módszerek átlagos négyzetes hibáihoz. Innentől mindig 6 lebonyolítási periódust fogunk vizsgálni, vagyis feltesszük, hogy 5 évnél nagyobb késéssel nem történik kárkifizetés, a korábbi jelöléseinkkel: $p_6 = 1$ vagy ezzel ekvivalensen $X_{i,6} = X_{i,6_+}$.

3.3.1. A becslés lépései

Egy konkrét példán végigvisszük hogyan épülnek fel a becsléseink, hogyan kapjuk meg a CL, BF és az - immáron becsült - optimális súllyal kevert tartalék négyzetes hibáit. Az első teljes kifutású évet tekintjük az első alapfelbontási periódusnak.

[7] P=1-re gondolhatunk úgy is, hogy egy kellően nagy egységben számolunk, mondjuk milliárd, vagy tízmilliárd forintban

1, Kifutási háromszög generálása

Γ eloszlású károkat fogunk generálni a kifutási háromszögbe, néhány korábbi, teljes kifutási évvel együtt, vagyis $\forall\ i$ alapfelbontási és j lebonyolítási periódusra legeneráljuk a $Z_{i,j}$ nem kumulált károknak megfelelő értékeket egy $\Gamma(\alpha_j, \lambda)$ eloszlású valószínűségi változóból. Mivel minden i-re közösek a paraméterek (a károk eloszlása nem függ az alapfelbontási periódustól), ezért ha megmondjuk az egyes lebonyolítási periódusokra jellemző α_j paramétereket és λ-t, akkor mindent megmondtunk.

Legyen $\lambda = 20$, a példánkban 4 teljes kifutású évet veszünk, a táblázatban a megfelelő α_i értékeket is feltüntetjük:

alapfelbontási	α_j értéke					
periódus	5,2	2	1,6	1,6	0,8	0,8
1	0,30138	0,08979	0,18122	0,09102	0,36639	0,02641
2	0,34699	0,03877	0,10061	0,02406	0,00087	0,08475
3	0,18364	0,14163	0,07638	0,11700	0,08978	0,14194
4	0,20697	0,09955	0,05920	0,07127	0,09553	0,10522
5	0,31836	0,14907	0,11966	0,07333	0,11750	
6	0,14893	0,18986	0,07843	0,28943		
7	0,51170	0,06961	0,11361			
8	0,16776	0,11418				
9	0,27094					

2. Kumulált háromszög

Az előző nem kumulált háromszögből legyártjuk az $X_{i,k} = \sum_{j=1}^{k} Z_{i,j}$ kumulált értékeket tartalmazó háromszöget. Ezt úgy is nézhetjük, mintha $X_{i,k}$-ket $\Gamma\left(\sum_{j=1}^{k} \alpha_j, \lambda\right)$ eloszlású valószínűségi változóból generáltuk volna.

alapfelbontási	$\sum_{j=1}^{k} \alpha_j$ értéke					
periódus	5,2	7,2	8,8	10,4	11,2	12
1	0,30138	0,39117	0,57239	0,66342	1,02981	1,05623
2	0,34699	0,38576	0,48637	0,51044	0,51131	0,59606
3	0,18364	0,32527	0,40165	0,51866	0,60844	0,75039
4	0,20697	0,30652	0,36572	0,43700	0,53253	0,63775
5	0,31836	0,46743	0,58709	0,66043	0,77793	
6	0,14893	0,33879	0,41723	0,70667		
7	0,51170	0,58132	0,69493			
8	0,16776	0,28194				
9	0,27094					

3. p_k elértségi szintek becslése

A CL és a BF módszerek esetén a p_k elértségi szintek becslése az 1.2 definícióban definiált módon történik, először a kumulált kifutási háromszögből a lánclétra-módszerrel megbecsüljük a növekedési faktorokat, majd ezek reciprokainak hátulról vett szorzata adja p_k lánclétra becslését.

p_i becsült értékei		elméleti értékek	
$\hat{p_1}$	0,32429	p_1	$5,2/12 = 0,4\dot{3}$
$\hat{p_2}$	0,45670	p_2	$7,2/12 = 0,6$
$\hat{p_3}$	0,57579	p_3	$8,8/12 = 0,7\dot{3}$
$\hat{p_4}$	0,71130	p_4	$10,4/12 = 0,8\dot{6}$
$\hat{p_5}$	0,88214	p_5	$11,2/12 = 0,9\dot{3}$

4. U_0 előzetes becslés

Előzetes becslésnek a korábbi, már teljes kifutással rendelkező évek végső kárainak átlagát választottuk. Itt most:

$$U_0 = (1,05623 + 0,59606 + 0,75039 + 0,63775)/4 = 0,76011$$

Bevezetünk egy jelölést: az M módszerrel meghatározott tartalék esetében R_{M_i}-vel fogjuk jelölni az i. lebonyolítási évhez tartozó tartalékot. Tehát R_{M_1} annak az alapfelbontási évnek a tartaléka, amikor még csak egy lebonyolítási periódus adata ismert (nekünk most a 9. alapfelbontási év ilyen), R_{M_2} esetében 2 lebonyolítási periódus adata ismert (most a 8. év ilyen) stb.

5. Az optimális súly becslése

$$\hat{c^*} = \frac{\hat{p_k}}{\hat{p_k} + \hat{t}}.$$

Γ eloszlású adatok esetén t becslésének természetesen adódik a korábbi, már teljes kifutású évek száma. De használhatjuk a t eredeti alakját használó becslést is, ez ugyanis visszaadja $l = 4$-et:

$$\hat{t} = \frac{\widehat{E(\alpha^2(U))}}{\widehat{Var(U_0)} + \widehat{Var(U)} - \widehat{E(\alpha^2(U))}}.$$

- Az Var(U) becslése a már teljes kifutású évek végső káradatainak korrigált tapasztalati szórásnégyzete lesz, jelölje S, ez torzítatlan becslés $Var(U) = \dfrac{\alpha}{\lambda^2}$-re.

- $E(U^2)$ második momentum becslése pedig

$$\widehat{E(U^2)} = \widehat{Var(U)} + (\widehat{E(U)})^2 = S + U_0^2.$$

- $\alpha = \dfrac{(E(U))^2}{Var(U)} \;\Rightarrow\; \hat{\alpha} = \dfrac{(\widehat{E(U)})^2}{\widehat{Var(U)}}.$

- k teljes kifutású korábbi év esetén

$$Var(U_0) = \frac{Var(U)}{k} \;\Rightarrow\; \widehat{Var(U_0)} = \frac{\widehat{Var(U)}}{k},$$

$$\begin{aligned}
\hat{t} &= \frac{\widehat{E(U^2)}\dfrac{1}{\hat{\alpha}+1}}{\hat{Var(U_0)} + \widehat{Var(U)} - \widehat{E(U^2)}\dfrac{1}{\hat{\alpha}+1}} = \\
&= \frac{\widehat{E(U^2)}\dfrac{1}{\dfrac{(\widehat{E(U)})^2}{\widehat{Var(U)}}+1}}{\widehat{Var(U_0)} + \widehat{Var(U)} - \widehat{E(U^2)}\dfrac{1}{\dfrac{(\widehat{E(U)})^2}{\widehat{Var(U)}}+1}} \\
&= \frac{\widehat{E(U^2)}\dfrac{\widehat{Var(U)}}{(\widehat{E(U^2)}}}{\dfrac{\widehat{Var(U)}}{k} + \widehat{Var(U)} - \widehat{E(U^2)}\dfrac{\widehat{Var(U)}}{\widehat{E(U^2)}}} = k.
\end{aligned}$$

Ennek a felírásnak az az előnye, hogy szükség van a végső kár varianciájának és második momentumának a becslésére is (és természetesen várható értékének becslésére is, de ezt már U_0-nál megtettük), ezek önmagukban értékes információkat hordozhatnak.

becsült értékek		elméleti értékek	
$\widehat{E(U)}$	0,76011	E(U)	12/20=0,6
$\widehat{Var(U)}$	0,04322	$Var(U)$	$12/20^2$=0,03
$\widehat{E(U^2)}$	0,62099	$E(U^2)$	0,39
$\hat{c}_1^*$	0,07499	c_1^*	0,0977443
$\hat{c}_2^*$	0,10247	c_2^*	0,1304347
$\hat{c}_3^*$	0,12583	c_3^*	0,1549295
$\hat{c}_4^*$	0,15097	c_4^*	0,1780821
$\hat{c}_5^*$	0,18068	c_5^*	0,1891891

6. R_{CL} és R_{BF} és R_{c^*} meghatározása

$$R_{BF_i} = (1 - \hat{p}_i)U_0 = \hat{q}_i U_0,$$
$$R_{CL_i} = X_{10-i,i}\frac{\hat{q}_i}{\hat{p}_i},$$
$$R_{c_i^*} = \hat{c_i^*}R_{CL_i} + (1 - \hat{c_i^*})R_{BF_i}.$$

$$\Downarrow$$

R_{BF_1}	0,51361	R_{CL_1}	0,56455	$R_{c_1^*}$	0,51743
R_{BF_2}	0,41296	R_{CL_2}	0,33540	$R_{c_2^*}$	0,40501
R_{BF_3}	0,32244	R_{CL_3}	0,51198	$R_{c_3^*}$	0,34629
R_{BF_4}	0,21944	R_{CL_4}	0,28681	$R_{c_4^*}$	0,22961
R_{BF_5}	0,08958	R_{CL_5}	0,10393	$R_{c_5^*}$	0,09217

7. A négyzetes hibák meghatározása

Az egyes pontban a kifutási háromszög generálásakor a "jövőbeli" károkat - vagyis az alsó háromszög értékeit - is legeneráljuk: a meglévő adatainkból ezeket az értékeket szeretnék minél pontosabban becsülni.

alapfelbontási	α_j értéke					
periódus	5	2	1,6	1,6	0,8	0,8
5					...	0,01129
6				...	0,03593	0,05155
7			...	0,15020	0,09689	0,03741
8		...	0,06023	0,07061	0,01929	0,04315
9	...	0,12527	0,18358	0,02784	0,07643	0,20362

A megfelelő sorösszegek lesznek az alapfelbontási periódusok tényleges tartalékszükségletei.[8]

[8]Emlékeztetünk, hogy R_i-vel jelöljük annak az alapfelbontási évnek a tartalékszükségletét, amelyikből i év káradatai ismertek.

R_1	0,61677
R_2	0,19329
R_3	0,28451
R_4	0,08748
R_5	0,01129

Az M módszerrel meghatározott tartalék négyzetes hibája abban az alapfelbontási évben, amelyben i év káradatai ismertek: $(R_{M_i} - R_i)^2$.

	$(R_{M_i} - R_i)^2$				
M	1	2	3	4	5
BF	0,01064	0,04825	**0,00143**	**0,01741**	**0,00612**
CL	**0,00272**	**0,020192**	0,05174	0,03973	0,00858
kevert	0,00986	0,04482	0,00381	0,02019	0,00654

Vastagon a legpontosabb tartalékot szedtük. [9]

Ezt az eljárást 10000-szer ismételjük meg, és vizsgáljuk az egyes módszerek átlagos négyzetes hibáit a 10000 generálás során. A kifutási háromszögek generálására és a fenti becslések elvégzésére Visual Basicben írtam programot.

3.2 Definíció: Standard hiba
Egy módszer n számú generálás melletti standard hibáján az n generáláshoz tartozó átlagos négyzetes hiba négyzetgyökét értjük.

A fogalom jól láthatóan analóg a szórással, ami a várható értéktől való várható négyzetes eltérés négyzetgyöke. [10] Későbbi táblázataink a módszerek 10000 generálás melletti standrad hibáit fogják tartalmazni, az előírt díj (feltevésünk szerint most az 1) arányában kifejezve.

A következő előzetes észrevételeket tesszük:

- minél több teljes kifutással rendelkező korábbi évünk van, a BF módszer annál erősebb lesz, hiszen az előzetes U_0 becslés annál pontosabb.

- a több teljes kifutású év hatására a c_i növekedési faktorokat, így a p_i elértségi szinteket is pontosabban tudjuk becsélni, tehát ennek a tényezőnek a hatására mindkét klasszikus módszer pontosabb lesz.

- különböző szórású adatokat fogunk vizsgálni, azonos "lebonyolódási minta" mellett. Ez azt jelenti, hogy a tényleges p_k elértségi szinteket nem

[9] A továbbiakban sokszor a legpontosabb tartalékra, mint "nyertes" tartalék hivatkozunk és táblázatainkban vastagon szedjük a nyertest.

[10] A standard hiba tehát NEM az abszolút eltérések átlaga, hanem az abszolút eltérések négyzetei átlagának négyzetgyöke. Előbbi a pozitív abszolút eltérések számtani, utóbbi négyzetes közepe.

α_1	$\sum_{j=1}^{2} \alpha_j$	$\sum_{j=1}^{3} \alpha_j$	$\sum_{j=1}^{4} \alpha_j$	$\sum_{j=1}^{5} \alpha_j$	$\sum_{j=1}^{6} \alpha_j$	λ	U szórása
26	36	44	52	56	60	100	7,746%
13	18	22	26	28	30	50	10,954%
5,2	7,2	8,8	10,4	11,2	12	20	17,32%
3,9	5,4	6,6	7,8	8,4	9	15	20%

2. táblázat. A Γ eloszlások paraméterei

változtatjuk, de az eloszlások szórását, amiből az adatok származnak, igen. A tényleges lebonyolódási mintánk a következő lesz: a végső kárhányad 60%, az elértségi szintek pedig:

$$p_1 = 0,4\dot{3}\,, \quad p_2 = 0,6\,, \quad p_3 = 0,7\dot{3}\,, \quad p_4 = 0,8\dot{6}\,, \quad p_5 = 0,9\dot{3}.$$

Ezek a 2. táblázatban feltüntetett paraméterekkel megadott Γ eloszlású valószínűségi változókból fognak származni. Az első két esetben viszonylag kicsi szórású, a második két esetben viszonylag nagy szórású portfólióról beszélhetünk. Az "1,96 σ szabályból" adódó konfidenciaintervallumok nagysága a végső kárra rendre 30,4%, 42,9%, 67,9%, 78,4% a díj százalékában (a várt végső kár a díj 60%-a).

3.3.2. Az elértségi szintekről

Milyen pontosan tudjuk becsülni a tényleges elértségi szinteket?

A 4. ábra mutatja a becslések standard hibáit. p_i CL becslése $6 - i$ növekedési faktor reciprokának szorzata, tehát nagyobb i-re kevesebb becslés szorzata adja p_i becslését. Emiatt nagyobb i-re annak ellenére pontosabbak a becsléseink, hogy a kifutási háromszögben "rövidebb" oszlopok adják a növekedési faktorok CL becsléseit. Ez alól jól láthatóan kivétel az első oszlop, pedig ő a legtöbb becslés szorzataként áll elő. Ennek oka, hogy $p_1 = 0,4\dot{3}$ elértségi szint alacsony, szóval csak az abszolút standard hiba kisebb, a becsült érték arányához mért relatív hiba természetesen már nagyobb.

Jegyezzük meg, hogy az elértségi szintek becslései mindkét klasszikus módszerben szerepelnek, de a CL becslés pontosságát jobban befolyásolják. A 3. ábra azt mutatja, hogy a különböző elértségi szinteken az elértségi szint becslésének adott százalékos változásának hatására a tartalékbecslés megváltozása a kiindulásként becsült tartalék hány százaléka. A 3. sorok ezeknek a relatív változásoknak az arányait mutatják.

	elértségi szintek											
	0,4	0,45	0,5	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	
1%	0,67%	0,82%	1,00%	1,22%	1,50%	1,86%	2,33%	3,00%	4,00%	5,67%	9,00%	BF
	1,65%	1,80%	1,98%	2,20%	2,48%	2,83%	3,30%	3,96%	4,95%	6,60%	9,90%	CL
	0,40	0,45	0,51	0,56	0,61	0,66	0,71	0,76	0,81	0,86	0,91	
2%	1,33%	1,64%	2,00%	2,44%	3,00%	3,71%	4,67%	6,00%	8,00%	11,33%	18,00%	BF
	3,27%	3,57%	3,92%	4,36%	4,90%	5,60%	6,54%	7,84%	9,80%	13,07%	19,61%	CL
	0,41	0,46	0,51	0,56	0,61	0,66	0,71	0,77	0,82	0,87	0,92	
3%	2,00%	2,45%	3,00%	3,67%	4,50%	5,57%	7,00%	9,00%	12,00%	17,00%	27,00%	BF
	4,85%	5,30%	5,83%	6,47%	7,28%	8,32%	9,71%	11,65%	14,56%	19,42%	29,13%	CL
	0,41	0,46	0,52	0,57	0,62	0,67	0,72	0,77	0,82	0,88	0,93	
4%	2,67%	3,27%	4,00%	4,89%	6,00%	7,43%	9,33%	12,00%	16,00%	22,67%	36,00%	BF
	6,41%	6,99%	7,69%	8,55%	9,62%	10,99%	12,82%	15,38%	19,23%	25,64%	38,46%	CL
	0,42	0,47	0,52	0,57	0,62	0,68	0,73	0,78	0,83	0,88	0,94	
5%	3,33%	4,09%	5,00%	6,11%	7,50%	9,29%	11,67%	15,00%	20,00%	28,33%	45,00%	BF
	7,94%	8,66%	9,52%	10,58%	11,90%	13,61%	15,87%	19,05%	23,81%	31,75%	47,62%	CL
	0,42	0,47	0,53	0,58	0,63	0,68	0,74	0,79	0,84	0,89	0,95	
6%	4,00%	4,91%	6,00%	7,33%	9,00%	11,14%	14,00%	18,00%	24,00%	34,00%	54,00%	BF
	9,43%	10,29%	11,32%	12,58%	14,15%	16,17%	18,87%	22,64%	28,30%	37,74%	56,60%	CL
	0,42	0,48	0,53	0,58	0,64	0,69	0,74	0,80	0,85	0,90	0,95	
7%	4,67%	5,73%	7,00%	8,56%	10,50%	13,00%	16,33%	21,00%	28,00%	39,67%	63,00%	BF
	10,90%	11,89%	13,08%	14,54%	16,36%	18,69%	21,81%	26,17%	32,71%	43,61%	65,42%	CL
	0,43	0,48	0,54	0,59	0,64	0,70	0,75	0,80	0,86	0,91	0,96	
8%	5,33%	6,55%	8,00%	9,78%	12,00%	14,86%	18,67%	24,00%	32,00%	45,33%	72,00%	BF
	12,35%	13,47%	14,81%	16,46%	18,52%	21,16%	24,69%	29,63%	37,04%	49,38%	74,07%	CL
	0,43	0,49	0,54	0,59	0,65	0,70	0,76	0,81	0,86	0,92	0,97	
9%	6,00%	7,36%	9,00%	11,00%	13,50%	16,71%	21,00%	27,00%	36,00%	51,00%	81,00%	BF
	13,76%	15,01%	16,51%	18,35%	20,64%	23,59%	27,52%	33,03%	41,28%	55,05%	82,57%	CL
	0,44	0,49	0,55	0,60	0,65	0,71	0,76	0,82	0,87	0,93	0,98	
10%	6,67%	8,18%	10,00%	12,22%	15,00%	18,57%	23,33%	30,00%	40,00%	56,67%	90,00%	BF
	15,15%	16,53%	18,18%	20,20%	22,73%	25,97%	30,30%	36,36%	45,45%	60,61%	90,91%	CL
	0,44	0,50	0,55	0,61	0,66	0,72	0,77	0,83	0,88	0,94	0,99	

3. ábra. BF és CL tartalékok változása és a változások aránya a különböző elértségi szinteknél

		p_i becsléseinek standard hibái				
		i=1	i=2	i=3	i=4	i=5
l=1	λ=100	3,4753%	3,9238%	4,0263%	3,7364%	3,2327%
	λ=50	4,9267%	5,5268%	5,6551%	5,2037%	4,4134%
	λ=20	7,6865%	8,6019%	8,8629%	8,1540%	6,9384%
	λ=15	8,7864%	9,8312%	10,0213%	9,2076%	7,8335%
l=2	λ=100	2,9537%	3,2572%	3,2318%	2,8043%	2,2690%
	λ=50	4,1901%	4,6010%	4,5348%	3,9708%	3,2043%
	λ=20	6,5966%	7,1860%	7,1461%	6,2171%	5,0398%
	λ=15	7,5640%	8,2368%	8,1569%	7,1357%	5,7627%
l=4	λ=100	2,4252%	2,5825%	2,4678%	2,0539%	1,6085%
	λ=50	3,4638%	3,6653%	3,5031%	2,8911%	2,2521%
	λ=20	5,4366%	5,7863%	5,5474%	4,5636%	3,5193%
	λ=15	6,2930%	6,6665%	6,4064%	5,2857%	4,1375%
l=10	λ=100	1,7859%	1,8215%	1,6888%	1,3533%	1,0233%
	λ=50	2,4892%	2,5700%	2,4059%	1,9168%	1,4437%
	λ=20	3,9389%	3,9918%	3,7183%	2,9913%	2,2840%
	λ=15	4,5376%	4,6748%	4,3851%	3,4704%	2,6159%

4. ábra. az elméleti elértségi szintek becsléseinek standard hibái

3.3.3. Eredmények

- A teljes kifutású évek száma sok : $l = 4$ és $l = 10$

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	**6,1298%**	**5,0774%**	**4,1657%**	**2,9533%**	**2,0881%**
	CL	9,3117%	6,5186%	4,8489%	3,1878%	2,1588%
	kevert	6,1370%	5,0780%	4,1666%	2,9557%	2,0884%
λ=50	BF	**8,6114%**	**7,1922%**	**6,0033%**	**4,1829%**	**2,9849%**
	CL	13,1119%	9,3640%	6,9997%	4,4872%	3,1023%
	kevert	8,6206%	7,1997%	6,0092%	4,1830%	2,9863%
λ=20	BF	**13,4372%**	**11,4338%**	**9,3328%**	**6,5749%**	4,6496%
	CL	20,7198%	14,9437%	10,9284%	7,0926%	4,8279%
	kevert	13,4416%	11,4473%	9,3342%	6,5759%	**4,6494%**
λ=15	BF	**15,7168%**	**13,0554%**	**11,0502%**	**7,6116%**	**5,4754%**
	CL	24,3573%	17,2923%	12,9740%	8,2314%	5,7245%
	kevert	15,7335%	13,0711%	11,0613%	7,6131%	5,4789%

5. ábra. tartalék becslések standard hibái, 10 teljes kifutású év

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	**6,3953%**	**5,4122%**	**4,4198%**	**3,0853%**	**2,2295%**
	CL	9,6668%	6,9288%	5,1724%	3,3427%	2,3128%
	kevert	6,4151%	5,4292%	4,4354%	3,0951%	2,2322%
λ=50	BF	**8,9819%**	**7,6851%**	**6,2107%**	**4,4501%**	**3,1264%**
	CL	13,7652%	9,9012%	7,3251%	4,7984%	3,2629%
	kevert	9,0122%	7,7149%	6,2358%	4,4591%	3,1322%
λ=20	BF	**14,3526%**	**12,0986%**	**9,8975%**	**6,9120%**	**4,9983%**
	CL	22,2310%	15,6346%	11,5289%	7,5465%	5,2572%
	kevert	14,3723%	12,1072%	9,9077%	6,9303%	5,0087%
λ=15	BF	**16,2186%**	**13,7122%**	**11,3264%**	**7,9856%**	**5,7599%**
	CL	25,0407%	18,0350%	13,5119%	8,7773%	6,1016%
	kevert	16,2198%	13,7365%	11,3568%	8,0068%	5,7758%

6. ábra. tartalék becslések standard hibái, 4 teljes kifutású év

Mint már előre jeleztük, a BF ilyen körülmények között nagyon erős, hiszen az előzetes becslés, amit használ, sokkal kisebb szórású adatból indul ki. Ezt egy jóval nagyobb szórású becslés hozzákeverésével nem tudjuk ellensúlyozni, hiába a szép elméleti háttér. A kisebb meglepetést inkább az jelentheti, hogy egyszer mégis nyerni tudott a kevert módszer a(z ilyen körülmények között általánosan) jóval gyengébb CL módszer hozzákeverésével. Az okok:

i, $l = 10$ miatt a CL módszert épphogy hozzákevertük a BF-hez, hiszen az elméleti optimális súlyok rendre

$c_1^* = 0,04153$	$c_2^* = 0,05660$	$c_3^* = 0,06832$	$c_4^* = 0,07975$	$c_5^* = 0,08536$

10 teljes kifutású év esetén (és

$c_1^* = 0,09774$	$c_2^* = 0,13043$	$c_3^* = 0,15492$	$c_4^* = 0,17808$	$c_5^* = 0,18918$

4 teljes kifutású év esetén.)

ii, ráadásul $l = 10$ miatt sok adatból becsüljük p_i-ket (az oszlopösszegeket legalább 10 kumulált érték adja, ezzel a becslés bizonytalansága lecsökken) és $i = 5$-re az elértségi szint becsléséhez csupán egy növekedési faktor becslésére van szükségünk. A 4.ábrából láthatjuk, hogy $i = 5$-re, különösen $l = 10$ esetén kicsi a standard hiba, ami mivel ebben az esetben nem egy generálásonként változó értéktől, hanem egy fix, elméleti értéktől vett távolságot mér, mutatja, hogy a becslések kis intervallumban szóródnak.

- A teljes kifutású évek száma 2.

Itt már izgalmasabb a kép. Kétszer akkor tudott nyerni a kevert tartalék, amikor a lebonyolításból még csak 1 év telt el, a tartalékszükséglet ekkor a legnagyobb, a pontos tartalékbecslés itt különösen sokat ér. Ráadásul most már az elméleti optimális súlyok is nagyobbak:

$c_1^* = 0,17808$	$c_2^* = 0,23076$	$c_3^* = 0,26829$	$c_4^* = 0,30232$	$c_5^* = 0,31818$

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	**6,9148%**	**5,6260%**	**4,6699%**	**3,4064%**	**2,4325%**
	CL	9,9644%	7,0820%	5,3941%	3,6513%	2,5558%
	kevert	6,9190%	5,6410%	4,6908%	3,4208%	2,4492%
λ=50	BF	9,6032%	**8,1533%**	**6,6808%**	**4,7877%**	**3,4727%**
	CL	14,0143%	10,3369%	7,7272%	5,1876%	3,6538%
	kevert	**9,5796%**	8,1903%	6,7151%	4,8132%	3,4949%
λ=20	BF	**15,2032%**	**12,6240%**	**10,3444%**	**7,3903%**	**5,5209%**
	CL	23,2813%	16,7168%	12,3311%	8,3071%	5,9406%
	kevert	15,2305%	12,6996%	10,4237%	7,4680%	5,5677%
λ=15	BF	17,7948%	**14,8765%**	**12,2183%**	**8,8349%**	**6,3407%**
	CL	27,1034%	19,3651%	14,6387%	9,9129%	6,9995%
	kevert	**17,7865%**	14,9364%	12,2802%	8,9102%	6,4207%

7. ábra. tartalék becslések standard hibái, 2 teljes kifutású év

- A teljes kifutású évek száma 1.

Ekkor előzetes becslésnek az egyetlen teljes kifutású év végső kárát választjuk, ennek a szórásnégyzete már nem kisebb, mint a CL módszer által használt aktuális kártapasztalatoké. **A Mack-modell ereje itt** már **megmutatkozik**, a 4 első felbontási évhez tartozó tartalékolás közül mindet, a 4 második felbontási évhez tartozó tartalékolás közül 3-at megnyert a kevert tartalék, valamint egyszer a harmadik felbontási periódusban is nyerni tudott. Az elméleti optimális súlyok:

$c_1^* = 0,33106$	$c_2^* = 0,38701$	$c_3^* = 0,42216$	$c_4^* = 0,47342$	$c_5^* = 0,48952$

Konklúzió: ha az előzetes becslés hasonló szórású előzetes becslésből indul ki, mint a CL módszer által használt saját kártapasztalat, akkor a Mack-modell által nyújtott keverés hatékony módszer az átlagos négyzetes hiba (vagy ezzel ekvivalensen a standard hiba) alacsonyan tartására.

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	7,5287%	**6,3085%**	5,0577%	**3,7798%**	**2,8523%**
	CL	10,1031%	7,5986%	5,6918%	4,0702%	3,0161%
	kevert	**7,3922%**	6,3087%	**5,0549%**	3,8241%	2,8953%
λ=50	BF	10,5855%	8,8336%	**7,2237%**	**5,3258%**	**4,0263%**
	CL	14,6181%	10,8048%	8,3569%	5,8160%	4,3123%
	kevert	**10,4666%**	**8,8282%**	7,3007%	5,4078%	4,1016%
λ=20	BF	16,8820%	13,9080%	**11,3172%**	**8,5039%**	**6,2500%**
	CL	24,7848%	17,7304%	13,8251%	9,8715%	7,2196%
	kevert	**16,7749%**	**13,8659%**	11,4840%	8,7530%	6,4925%
λ=15	BF	19,3812%	16,1241%	**13,2218%**	**9,6096%**	**7,1575%**
	CL	28,6403%	21,0440%	16,2361%	11,5832%	8,4082%
	kevert	**19,1817%**	**16,1018%**	13,4133%	9,9279%	7,4277%

8. ábra. tartalék becslések standard hibái, 1 teljes kifutású év

Eredményeinkből azt is láthatjuk, hogy azokban az esetekben, amikor a kevert tartalék nem tudott nyerni, a standard hibája sokszor 1 vagy akár 2 tizedesjegy pontossággal is megegyezik a nyertes BF módszerével. Sok teljes kifutású év esetén az egyik ok természetesen az, amit már megjegyeztünk korábban, viszonylag kicsi súllyal keverjük a BF módszerhez a CL módszert. Erről 2 vagy 1 teljes kifutású év esetén már nincs szó, a másik ok a keverés tényleges hatása, ha az egyik becslés egy esetleges extrém érték miatt távol van a tényleges tartalékszükséglettől, a másik módszer hozzákeverése ezen tud javítani, átlagos hibát tekintve egy jobb becslés tud megszületni.

Ezt erősíti meg a 9. ábra, melyen láthatjuk azt, hogy a kevert módszer önmagában rendre nem nyer meg túl sokat a 10000 generálásból. Mégis összeségében, standard hibát tekintve nyerni tudott több esetben, máskor pedig standard hibája tizedesjegy pontossággal megegyezett a nyertes BF módszer standard hibájával. Vagyis a keverés két különböző információból kiinduló becslésből olyan becslést hoz létre, ami az extrém nagy hibákat kiküszöböli, tehát az átlagos hibát tartja alacsonyan - persze pontosan ezt vártuk.

A 9. ábra másik tanulsága, hogy a CL módszer a lebonyolítás előrehaladtával egyre többször tudott nyerni, ilyenkor az elértségi szintek becsléséhez kevesebb növekedési faktort kell megbecsülnünk. Ha úgy gondolunk a kifutási háromszög legenerált értékeire, mint a már ismert, meglévő kártapasztalatok, akkor a lebonyolítás előrehaladtával az adott évre jellemző kártapasztalat egyre több információt hordoz, ez a mi modellünkben is visszaköszön: a CL módszer egyre többször tud nyerni.

Összefoglalva:
Abból a feltevésből indultunk ki, hogy minden év kárlebonyolítása azonos eloszlás szerint zajlik és láttuk azt is, hogy ez a BF módszernek nagyon kedvez. Standard hibában mégis rendre meg tudtuk közelíteni a kevert módszerünkkel. Még 10 teljes kifutású korábbi év esetén is tudott egyszer nyerni a kevert módszer. Két teljes kifutású korábbi év esetén 2-szer akkor tudott nyerni, amikor a tartalékszükséglet a legnagyobb. 1 teljes kifutású év esetén pedig, amikor a BF módszer elveszti a feltevésből származó előnyét, a kevert tartalék már a legfontosabb tartalékolási időszakokban (a lebonyolítás elején) rendre nyerni tudott. Azt is látjuk, hogy az, hogy az adatok mekkora szórású eloszlásból származnak (tehát a portfólió várt kárhányada kicsi, közepes vagy nagy szórású) ilyen feltételek mellett nem ad lényeges különbséget. [11]

		ki hányszor nyer, l=10					ki hányszor nyer, l=4				
		i=1	i=2	i=3	i=4	i=5	i=1	i=2	i=3	i=4	i=5
λ=100	BF	5016	5015	4989	4985	4959	5059	5125	5048	5162	5022
	CL	3365	3752	4113	4421	4625	3344	3756	4072	4302	4705
	kevert	1619	1233	898	594	416	1597	1119	880	536	273
λ=50	BF	5051	5054	5075	4937	5011	5112	5110	5186	5073	5059
	CL	3339	3750	3988	4429	4608	3310	3713	3851	4331	4544
	kevert	1610	1196	937	634	381	1578	1177	963	596	397
λ=20	BF	4920	4994	4943	4943	4860	4977	4978	4986	5038	5008
	CL	3382	3821	4207	4455	4760	3395	3811	4101	4378	4623
	kevert	1698	1185	850	602	380	1628	1211	913	584	369
λ=15	BF	5026	5060	5016	4886	4923	4950	5018	5042	5029	5076
	CL	3415	3798	4111	4546	4736	3451	3753	4076	4396	4601
	kevert	1559	1142	873	568	341	1599	1229	882	575	323
		ki hányszor nyer, l=2					ki hányszor nyer, l=1				
		i=1	i=2	i=3	i=4	i=5	i=1	i=2	i=3	i=4	i=5
λ=100	BF	4909	5018	5073	4991	5202	4730	5045	4956	5179	5258
	CL	3425	3707	3985	4417	4367	3491	3700	4033	4205	4391
	kevert	1666	1275	942	592	431	1779	1255	1011	616	351
λ=50	BF	4917	5031	5134	5121	5105	4868	5006	5047	5211	5154
	CL	3390	3800	3962	4242	4495	3447	3735	3906	4156	4406
	kevert	1693	1169	904	637	400	1685	1259	1047	633	440
λ=20	BF	4950	5075	5100	5072	5076	4926	4908	5037	5186	5246
	CL	3360	3697	4004	4344	4541	3411	3780	3985	4238	4395
	kevert	1690	1228	896	584	383	1663	1312	978	576	359
λ=15	BF	4927	5083	5004	5104	4997	4950	4920	5095	5106	5255
	CL	3507	3771	4059	4316	4638	3398	3778	3946	4257	4404
	kevert	1566	1146	937	580	365	1652	1302	959	637	341

9. ábra. ki hányszor tudott nyerni a 10000 generálás során

[11] nagyobb szórású adatokkal dolgozva a standard hibák persze külön-külön nőnek

3.3.4. Randomizált eredmények

Megvizsgáljuk azt, hogy mi történik, ha vannak "jó" és "rossz" éveink, 1/2 - 1/2 valószínűséggel. A jó évben a 2. táblázatnak megfelelően az előírt díj százalékában kifejezve rendre 26, 36, 44, 52, 56 és 60% az i. periódus után a kumulált kár várható értéke, a rossz évben ennek pont a duplája, vagyis 52, 72, 88, 104, 112 és 120%. A dupla várható értéket úgy érjük el, hogy a megfelelő α paramétert választjuk kétszeresre, a λ paramétereket nem változtatjuk.

Figyelembe véve, hogy

$$\frac{1}{2}\,\Gamma((k\cdot\lambda),\,\lambda)+\frac{1}{2}\,\Gamma((2k\cdot\lambda),\,\lambda)=\Gamma((k\cdot\lambda),\,2\lambda)+\Gamma((2k\cdot\lambda),\,2\lambda)=$$
$$=\Gamma((3k\cdot\lambda),\,2\lambda), \tag{3.14}$$

ezért minden i-re a megfelelő károk eloszlása itt is minden alapfelbontásbeli évre azonos Γ eloszlás, így az a tulajdonság is megmarad, hogy az i. felbontási periódusra jellemző várt elértségi szintek nem függnek az alapfelbontási periódustól.

Több teljes kifutású korábbi év esetén ezért túl sok változást az eredményeinkben nem várhatunk, ezt a következő táblázatok is igazolják.

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	**11,3386%**	**9,1340%**	**7,3461%**	**4,6165%**	**3,3283%**
	CL	22,7057%	14,3666%	9,7491%	5,4164%	3,6035%
	kevert	11,3458%	9,1537%	7,3574%	4,6208%	3,3308%
λ=50	BF	**13,5484%**	**11,0335%**	**8,9577%**	**5,8658%**	4,1498%
	CL	25,4817%	16,4957%	11,4741%	6,6430%	4,3930%
	kevert	13,5611%	11,0479%	8,9759%	5,8697%	**4,1489%**
λ=20	BF	**18,9837%**	**15,5413%**	**12,5311%**	**8,5669%**	**6,0219%**
	CL	32,6307%	21,6135%	15,7146%	9,5554%	6,4042%
	kevert	18,9850%	15,5457%	12,5604%	8,5746%	6,0279%
λ=15	BF	**21,0739%**	**17,5649%**	**14,3531%**	**9,7033%**	6,8277%
	CL	35,8911%	24,3542%	17,5653%	10,6842%	7,1502%
	kevert	21,0915%	17,5811%	14,3756%	9,7081%	**6,8267%**

10. ábra. standard hibák, 10 teljes kifutású év, randomizált eset

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	**11,7897%**	**9,6251%**	**7,6010%**	**4,9540%**	**3,5262%**
	CL	24,2931%	15,0981%	10,1119%	5,8354%	3,8438%
	kevert	11,8861%	9,6923%	7,6452%	4,9823%	3,5390%
λ=50	BF	**12,6940%**	**10,6533%**	**8,8276%**	**7,0007%**	**4,9670%**
	CL	24,7420%	16,5753%	11,9057%	8,4411%	5,5009%
	kevert	12,7678%	10,7368%	8,8932%	7,0471%	4,9912%
λ=20	BF	**19,5789%**	**16,3502%**	**13,2239%**	**9,2017%**	**6,4981%**
	CL	34,8341%	23,1834%	16,3411%	10,3308%	6,9634%
	kevert	19,6515%	16,4178%	13,2787%	9,2385%	6,5192%
λ=15	BF	**22,4021%**	**18,4163%**	**14,7627%**	**10,3718%**	**7,5111%**
	CL	38,4934%	25,8709%	18,3909%	11,5502%	8,0606%
	kevert	22,4653%	18,4973%	14,8368%	10,3981%	7,5391%

11. ábra. standard hibák, 4 teljes kifutású év, randomizált eset

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	**12,9118%**	**10,2802%**	**8,1286%**	**5,3139%**	**3,8397%**
	CL	24,8737%	15,7755%	10,7122%	6,2238%	4,2609%
	kevert	13,0696%	10,4488%	8,2211%	5,3711%	3,8923%
λ=50	BF	**15,3302%**	**12,4893%**	**10,0187%**	**6,7192%**	**4,9263%**
	CL	27,5123%	18,1294%	12,7582%	7,8590%	5,3726%
	kevert	15,3966%	12,5982%	10,0975%	6,8154%	4,9759%
λ=20	BF	**21,4153%**	**17,4314%**	**13,8885%**	**9,9883%**	**7,1429%**
	CL	36,3605%	24,4165%	17,5347%	11,5254%	7,9143%
	kevert	21,5267%	17,5533%	14,0167%	10,1279%	7,2444%
λ=15	BF	**24,0746%**	**19,6551%**	**15,7037%**	**11,3775%**	**8,1129%**
	CL	40,5228%	27,4658%	19,8431%	13,2002%	9,0786%
	kevert	24,1804%	19,7897%	15,8458%	11,5435%	8,2372%

12. ábra. standard hibák, 2 teljes kifutású év, randomizált eset

Ha csak 1 teljes kifutású korábbi évünk van, akkor jelentősen változik a helyzet. Ilyenkor az előzetes becslés az egyetlen már ismert végső kár, ami ha rossz év volt, a vizsgált év pedig jó, akkor nagyon pontotlan tartalékbecslést fog adni, hiszen várható értékben pont dupla annyi tartalékszükségletet jelez előre, mint amire várhatóan szükségünk lesz (és ugyanígy megfordítva). Emiatt itt a CL becslés várhatóan sokkal erősebb, ezt mutatják eredményeink is:

- A teljes kifutású évek száma 1.

		i=1	i=2	i=3	i=4	i=5
λ=100	BF	80,3977%	82,9871%	67,8418%	40,6420%	**4,4916%**
	CL	**61,4749%**	**55,5965%**	**46,8910%**	**32,4306%**	5,0403%
	kevert	69,0536%	66,0378%	55,0285%	35,8620%	4,6234%
λ=50	BF	80,8234%	83,4363%	68,4079%	41,0512%	**5,6991%**
	CL	**61,8346%**	**55,9387%**	**47,4378%**	**33,0262%**	6,5201%
	kevert	69,3072%	66,2887%	55,4864%	36,3420%	5,9176%
λ=20	BF	83,5998%	86,2414%	70,3797%	42,8836%	**8,1397%**
	CL	**63,5957%**	**58,1304%**	**49,2662%**	**35,3439%**	9,6664%
	kevert	70,9840%	68,2741%	57,0661%	38,3533%	8,5481%
λ=15	BF	85,1670%	87,0497%	71,3284%	43,5132%	**9,6900%**
	CL	**65,0757%**	**59,0995%**	**50,4482%**	**36,2648%**	11,9637%
	kevert	72,2433%	68,9375%	57,9837%	39,0930%	10,2581%

13. ábra. standard hibák, 1 teljes kifutású év, randomizált eset

Nem randomizált esetben $l = 1$ mellett a kevert módszer meglehetősen sikeresnek bizonyult, mi az oka, hogy itt nem tud nyerni? A BF módszer a fentiek alapján kifejezett rossz előzetes becslést használ, és az elméleti optimális súlyok - amelyek a CL módszer súlyát mondják meg - 1/2-nél kisebbek:

$c_1^* = 0,33106$	$c_2^* = 0,38701$	$c_3^* = 0,42216$	$c_4^* = 0,47342$	$c_5^* = 0,48952$

tehát az ilyen körülmények között sokkal gyengébb BF módszer kapja a nagyobb súlyt, így a kevert tartalék esélyeit is rontjuk.

Mi a helyzet a 5. lebonyolítási periódushoz tartozó tartalékolással? Arra az eddig elmondottak nem igazak, nem teljesített rosszabbul a BF módszer, sőt, minden esetben nyerni tudott. Az ok a következő: 1 teljes kifutású év esetén p_5 elértségi szint becslése nem más, mint a kumulált kifutási háromszögbeli $X_{1,5}$ és $X_{1,6}$ értékek hányadosa, ennek ismerjük az eloszlását:

$$\frac{\Gamma(1,68\cdot\lambda,2\lambda)}{\Gamma(1,8\cdot\lambda,2\lambda)} \Rightarrow q_k \text{ becslésének eloszlása } \left(1-\frac{\Gamma(1,68\cdot\lambda,2\lambda)}{\Gamma(1,8\cdot\lambda,2\lambda)}\right).$$

A BF tartalékbecslés nem más, mint q_k becslésének és $X_{1,6}$-nak a szorzata, az ő eloszlása :

$$\left(1-\frac{\Gamma(1,68\cdot\lambda,2\lambda)}{\Gamma(1,8\cdot\lambda,2\lambda)}\right)\cdot\Gamma(1,8\cdot\lambda,2\lambda).$$

A becslés felírásakor a $\Gamma(1,8\cdot\lambda,2\lambda)$ eloszlású valószínűségi változóból származó érték mindkétszer ugyanaz az érték, az egyetlen ismert végső kár, ezért lehet vele egyszerűsíteni! [12] $\Rightarrow$

$$\Gamma(1,8\cdot\lambda,2\lambda)-\Gamma(1,68\cdot\lambda,2\lambda)=\Gamma(0,12\cdot\lambda,2\lambda)$$

eloszlású a BF módszer tartalékbecslése.

A CL módszer esetében a $\frac{q_k}{p_k}$ becslésének eloszlása

$$\frac{\left(1-\frac{\Gamma(1,68\cdot\lambda,2\lambda)}{\Gamma(1,8\cdot\lambda,2\lambda)}\right)}{\frac{\Gamma(1,68\cdot\lambda,2\lambda)}{\Gamma(1,8\cdot\lambda,2\lambda)}}=\left(1-\frac{\Gamma(1,68\cdot\lambda,2\lambda)}{\Gamma(1,8\cdot\lambda,2\lambda)}\right)\cdot\frac{\Gamma(1,8\cdot\lambda,2\lambda)}{\Gamma(1,68\cdot\lambda,2\lambda)}=$$

$$\stackrel{*}{=}\left(\frac{\Gamma(1,8\cdot\lambda,2\lambda)}{\Gamma(1,68\cdot\lambda,2\lambda)}-1\right). \tag{3.15}$$

(*-nál azért egyszerűsíthettünk, mert a becslésben ismét ugyanazokról az értékekről - $X_{1,5}$ és $X_{1,6}$ - van szó) Egy ilyen eloszlású valószínűségi változóból származó értékre kell rászoroznunk a kifutási háromszög $X_{2,5}$-ös elemére, ami $\Gamma(1,68\cdot\lambda,2\lambda)$ eloszlású valószínűségi változóból származik.

$$\Rightarrow \text{a CL becslés } \left(\frac{\Gamma(1,8\cdot\lambda,2\lambda)}{\Gamma(1,68\cdot\lambda,2\lambda)}-1\right)\cdot\Gamma(1,68\cdot\lambda,2\lambda) \text{ eloszlású.}$$

Fontos, hogy itt már nem ugyanaz $X_{2,5}$ értéke, mint ami 3.15-ben a $\Gamma(1,68\cdot\lambda,2\lambda)$ eloszlású valószínűségi változóból származik [13], csak az eloszlásuk egyezik meg, ezért természetesen egyszerűsíteni sem tudunk. [14]

[12] tehát magában a becslésben nem két azonos eloszlású valószínűségi változóból származó érték szerepel, hanem egy adott valószínűségi változóból származó érték kétszer, ezért jogos az egyszerűsítés

[13] az egymástól független $X_{1,5}$-öt és $X_{2,5}$-öt adják meg

[14] ha tudnánk egyszerűsíteni, itt is visszakapnánk a becslésre a $\Gamma(0,12\cdot\lambda,2\lambda)$ eloszlást.

Amit becsülni szeretnék, az egy az eddigiektől független $\Gamma(0, 12 \cdot \lambda, \lambda)$ eloszlású valószínűségi változóból származó $X_{2,6}$ értéke. Ezt a BF módszer $X_{2,6}$ eloszlásával megegyező eloszlású valószínűségi változóból származó értékkel becsüli, eredményeink alapján átlagos négyzetes eltérésben jobban, mint a CL módszer, aminek az eloszlására egy összetettebb eloszlást kaptunk.

3.3 Megjegyzés: Az az előny, ami abból származott, hogy tudtunk "egyszerűsíteni" $i = 4$-re már nem lesz meg. Pontosabban, a BF módszer eloszlásának meghatározásánál tudunk ugyan $X_{1,6}$ értékével egyszerűsíteni, illetve a CL módszernél is a $\frac{q_k}{p_k}$ eloszlást felírva, de már mindkét módszer által szolgáltatott becslésre egy összetett eloszlást kapunk. A két becslés eloszlása:

$$\mathrm{BF} \sim \left(1 - \underbrace{\frac{\Gamma(2 \cdot 1, 56\lambda, 2\lambda)}{\Gamma(2 \cdot 1, 68\lambda, 2\lambda)} \frac{\Gamma(1, 68\lambda, 2\lambda)}{\Gamma(1, 8\lambda, 2\lambda)}}_{\hat{p_k} \text{ eloszlása}}\right) \cdot \Gamma(1, 8 \cdot \lambda, 2\lambda)$$

$$\mathrm{CL} \sim \underbrace{\left(\underbrace{\frac{\Gamma(2 \cdot 1, 68\lambda, 2\lambda)}{\Gamma(2 \cdot 1, 56\lambda, 2\lambda)}}_{\frac{1}{\hat{c}_4} \text{ eloszlása}} \underbrace{\frac{\Gamma(1, 8\lambda, 2\lambda)}{\Gamma(1, 68\lambda, 2\lambda)}}_{\frac{1}{\hat{c}_5} \text{ eloszlása}} - 1\right)}_{\frac{\hat{q}_k}{\hat{p}_k} \text{ eloszlása}} \cdot \underbrace{\Gamma(1, 56 \cdot \lambda, 2\lambda)}_{X_{3,4} \text{ eloszlása}}$$

Megnézve a 15. ábrát, a következőket láthatjuk:

i, $l = 2, 4, 10$ esetén változás volt abban, hogy a 10000 generálásból melyik módszer hányat tudott megnyerni: a kevert módszert összeségében több esetet nyert meg, mint a sima esetben és a többlet a CL módszertől hiányzik.

ii, Különbség továbbá: az adatok szórásának komolyabb hatása van arra, hogy ki hányszor tudott nyerni, szemben a sima esettel, haladva a kisebb szórású adatok felé a kevert módszer erősödik a CL módszerrel szemben. A 14. ábra részben erre magyarázatot ad, rögzített i-re a kisebb szórású adatok esetén a BF módszer relatíve pontosabb.

iii, $l = 1$ esetén pedig pontosan visszaköszön, amiről eddig beszéltünk, $i = 1, 2, 3, 4$-re a CL módszer tarol, majd $i = 5$-nél a BF módszer életre kel.

	standard hibák aránya (BF/CL)					standard hibák aránya, randomizált eset				
l=2	i=1	i=2	i=3	i=4	i=5	i=1	i=2	i=3	i=4	i=5
λ=100	69,395%	79,441%	86,574%	93,294%	95,178%	51,910%	65,166%	75,881%	85,380%	90,115%
λ=50	68,524%	78,876%	86,459%	92,292%	95,042%	55,721%	68,890%	78,528%	85,498%	91,693%
λ=20	65,302%	75,517%	83,889%	88,963%	92,934%	58,897%	71,392%	79,206%	86,664%	90,254%
λ=15	65,655%	76,821%	83,466%	89,125%	90,587%	59,410%	71,562%	79,140%	86,192%	89,362%

14. ábra. A klasszikus becslések standard hibáinak egymáshoz viszonyított aránya, l=2

		ki hányszor nyer, l=10, randomizált					ki hányszor nyer, l=4, randomizált				
		i=1	i=2	i=3	i=4	i=5	i=1	i=2	i=3	i=4	i=5
λ=100	BF	5068	5046	5111	5113	5090	5139	5116	5208	5222	5080
	CL	2421	3099	3632	3953	4341	2376	3085	3523	3875	4316
	kevert	2511	1855	1257	934	569	2485	1799	1269	903	604
λ=50	BF	4995	5066	5048	5039	4956	5196	5148	5240	5169	5142
	CL	2689	3290	3761	4138	4509	2520	3118	3445	3867	4171
	kevert	2316	1644	1191	823	535	2284	1734	1315	964	687
λ=20	BF	4912	4973	5060	5008	5057	4982	5115	5131	5102	5056
	CL	3110	3577	3848	4291	4520	3058	3455	3851	4173	4475
	kevert	1978	1450	1092	701	423	1960	1430	1018	725	469
λ=15	BF	5054	4964	4977	4975	4893	5006	5155	5108	5009	5089
	CL	3111	3601	3968	4332	4688	3058	3469	3863	4320	4491
	kevert	1835	1435	1055	693	419	1936	1376	1029	671	420
		ki hányszor nyer, l=2, randomizált					ki hányszor nyer, l=1, randomizált				
		i=1	i=2	i=3	i=4	i=5	i=1	i=2	i=3	i=4	i=5
λ=100	BF	5250	5208	5096	5154	5261	522	11	24	465	5232
	CL	2372	3083	3610	3871	4105	9478	9989	9976	9535	4181
	kevert	2378	1709	1294	975	634	0	0	0	0	587
λ=50	BF	5177	5084	5155	5171	5122	718	48	95	718	5247
	CL	2634	3298	3677	4027	4336	9282	9952	9905	9281	4198
	kevert	2189	1618	1168	802	542	0	0	0	1	555
λ=20	BF	5104	5075	5097	5172	5200	1231	228	399	1353	5233
	CL	2981	3447	3834	4131	4334	8760	9772	9600	8633	4347
	kevert	1915	1478	1069	697	466	9	0	1	14	420
λ=15	BF	5054	5067	5088	5159	5191	1593	466	665	1636	5202
	CL	3091	3525	3886	4183	4401	8389	9523	9321	8323	4404
	kevert	1855	1408	1026	658	408	18	11	14	41	394

15. ábra. ki hányszor tudott nyerni a 10000 generálás során, randomizált eset

3.3.5. Összefoglalás

Miután sorra vettük a kifutási háromszögön alapuló klasszikus módszereket, abból a feltevésből indultunk ki, hogy a különböző információn alapuló klasszikus módszerek keverésével egy sikeresebb tartalékolási módszert hozhatunk létre. A károk lebonyolódására tett elméleti feltevésekből Mack levezetett egy optimális súlyt. A dolgozatban megmutattuk, hogy Gamma eloszlású kárhányadra a Mack feltételek teljesülnek. Gamma eloszlású valószínűségi változókból generáltunk káradatokat, felépítettük becsléseinket és megvizsgáltuk, hogyan teljesítenek a különböző módszerek. A modellünkben minden évnek azonos eloszlást adtunk és láttuk, hogy ilyen körülmények között a Bornhuetter-Ferguson módszer nagyon erős, hiszen az aktuális kártapasztalatnál jóval kisebb szórású adatra épít.

Eredményeinkből azt láttuk, hogy a kevert tartalék reális alternatívának bizonyul: a legtöbb esetben pontossága nagyon közel volt a "nyertes" - legpontosabb - BF tartalékbecsléshez, 2 teljes kifutási év esetén pedig előfordult, hogy a kárlebonyolódás elején is nyerni tudott, amikor a nagy tartalékszükséglet miatt legfontosabb a jó becslés.

1 teljes kifutású év esetén láttuk, hogy a lebonyolítás elején szinte mindig a kevert tartalék nyert, ez azt mutatta, hogy amikor a kiinduló módszerek hasonló volitalitású adatból építkeznek, akkor a két módszer keverésével csökkenthető az átlagos négyzetes hiba.

Randomizált esetben - amikoris feltételeztünk a kárarány szempontjából jó és rossz éveket is - a helyzet 1 teljes kifutású év esetén tért el jelentősen a nem randomizált esettől, ekkor főleg a BF és a CL módszer egymáshoz való viszonyáról tettünk megállapításokat. Láttuk, hogy ilyen körülmények között a CL módszer a legpontosabb, kivéve az utolsó lebonyolítási periódusra vonatkozó tartalékolást, ahol a BF módszer.

Hivatkozások

[1] T.Mack : Credible claims reserves: The Benktander Method, ASTIN Bulletin Volume 30, No. 2, 2000

[2] G.Benktander: An Approach to Credibility in Calculating IBNR for Casualty Excess Reinsurance. The Actuarial Review, 1976 április

[3] Arató Miklós: Nem-életbiztosítási matematika, Eötvös Kiadó, Budapest, 2001

[4] Hanák Gábor: Biztosítási tartalék és szolvencia kurzus anyaga az ELTE-n 2009/2010-es év tavaszi félévében

[5] G. Taylor: Loss Reserving: An Actuarial Perspective, Kluwer, 2000

[6] P. D. England, R. J. Verrall: Stochastic claims reserving in general insurance, British Actuarial Journal (443-518), 2002

[7] K. D. Schmidt: Bornhuetter-Ferguson as a General Principle of Loss Reserving, 2008

[8] R. L. Bornhuetter, R. E. Ferguson: The Actuary and IBNR. Proc. CAS 59, 181-195., 1972