

**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Adam Němec

Aditivita metody Chain-Ladder pro projekci technických rezerv v neživotním pojištění

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Tomáš Cipra, DrSc.

Studijní program: Matematika

Studijní obor: Finanční matematika

Praha 2023

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Rád bych tímto poděkoval prof. RNDr. Tomáši Ciprovi, DrSc. za cenné rady a věcné připomínky při vedení mé bakalářské práce a za nesmírnou ochotu a čas, který mi věnoval během konzultací i mimo ně.

Název práce: Aditivita metody Chain-Ladder pro projekci technických rezerv v neživotním pojištění

Autor: Adam Němec

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Tomáš Cipra, DrSc., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Tato bakalářská práce se věnuje problematice aditivity projekcí získaných pomocí metody Chain Ladder v příslušných kumulativních vývojových trojúhelnících. Čtenáře nejprve seznamuje se samotnou metodou Chain Ladder a následně předkládá základní teoretické poznatky týkající se právě aditivity projekcí a s ní související projekční nerovnosti. Zabývá se také praktickou interpretací této problematiky, kterou demonstruje na reálných datech zajišťovny. Při numerické analýze využívá mj. směrodatných chyb, k jejichž výpočtu využívá nástrojů stručně zavedených v samostatné kapitole za využití základní teorie pravděpodobnosti.

Klíčová slova: aditivita, Chain-Ladder, IBNR, neživotní pojištění.

Title: Additivity of Chain-Ladder method for projection of technical provisions in non-life insurance

Author: Adam Němec

Department: Department of Probability and Mathematical Statistics

Supervisor: prof. RNDr. Tomáš Cipra, DrSc., Department of Probability and Mathematical Statistics

Abstract: This bachelor thesis deals with the subject of additivity of projections obtained by the Chain Ladder method in the corresponding cumulative development triangles. The reader first gets acquainted with the Chain Ladder method itself and then one presents basic theoretical insights concerning the additivity of projections and the related projection inequality. It also discusses practical interpretation, which it demonstrates using real reinsurance data. Moreover, standard errors of projections in triangles are briefly described in a separate chapter using basic theory of probability and applied in the given numerical study.

Keywords: additivity, Chain-Ladder, IBNR, non-life insurance.

Obsah

Úvod	2
1 Metoda Chain Ladder	4
1.1 Deterministický popis metody CL	4
1.2 Stochastické modely založené na metodě CL	5
1.2.1 Mackův model	6
2 Aditivita projekcí získaných pomocí metody Chain Ladder	8
2.1 Podmínky pro aditivitu projekcí	9
2.2 Projekční nerovnost pro dva trojúhelníky	12
3 Směrodatná chyba projekcí	18
3.1 Směrodatná chyba odhadu IBNR rezervy pro jeden trojúhelník . .	18
3.2 Směrodatná chyba součtu odhadů IBNR rezerv dvou korelovaných trojúhelníků	21
4 Numerická analýza	25
Závěr	34
Seznam použité literatury	35
Seznam obrázků	36
Seznam tabulek	37
A Přílohy	38
A.1 První příloha	38

Úvod

Metoda Chain Ladder je jednou z nejčastěji používaných metod pro odhadování technických rezerv v pojišтовnictví. Využívá se zejména v neživotním pojištění pro odhad tzv. IBNR rezerv na škody již vzniklé, ovšem dosud nenahlášené pojišтовně (z anglického „Incurred But Not Reported“). Aby získala tyto odhady, vytváří pojišтовna tzv. vývojové trojúhelníky.

Označme N rok, na jehož konci chceme odhadnout potřebnou rezervu. Dále označme i rok vzniku pojistné události a j rok vývoje pojistných plnění, tedy počet let, o než se zpozdilo vyplacení pojistných plnění na tuto událost. Nechtě $0 \leq i \leq N$ a $0 \leq j \leq N - i$. Vývojovým trojúhelníkem pak rozumíme takový levý horní trojúhelník matice, jehož prvek na pozici $\{i, j\}$ odpovídá celkovému pojistnému plnění na pojistné události vzniklé v roce i vyplacenému se zpožděním v roce $i + j$. Tento prvek označme $X_{i,j}$.

Poznámka. Tato matice nemusí být nutně čtvercová. V rámci této práce ovšem jiné případy uvažovat nebudeme.

Pro potřeby metody Chain Ladder se vývojový trojúhelník dále upravuje na tzv. kumulativní vývojový trojúhelník.

Definice. Kumulativním vývojovým trojúhelníkem rozumíme takový levý horní trojúhelník matice, jehož prvek na pozici $\{i, j\}$, $0 \leq i \leq N$, $0 \leq j \leq N - i$, odpovídá součtu celkových pojistných plnění na pojistné události vzniklé v roce i vyplacených se zpožděním v letech $i, \dots, i + j$. Tento prvek označme

$$C_{i,j} := \sum_{k=0}^j X_{i,k}.$$

Rok vzniku pojistné události	Rok vývoje pojistných plnění				
	0	...	j	...	N
0	$C_{0,0}$...	$C_{0,j}$...	$C_{0,N}$
\vdots	\vdots		\vdots	\ddots	
i	$C_{i,0}$...	$C_{i,j}$		
\vdots	\vdots	\ddots			
N	$C_{N,0}$				

Tabulka 1: Kumulativní vývojový trojúhelník

Aplikací metody Chain Ladder na takový kumulativní vývojový trojúhelník získáme prvky napravo od vedlejší diagonály na pozici $\{i, j\}$ pro $1 \leq i \leq N$ a $N - i + 1 \leq j \leq N$. Ty představují odhady celkových budoucích kumulovaných pojistných plnění, jež bude pojišтовna v budoucnu muset vyplácet a na něž je ze zákona povinna vytvářet rezervy. Označme je $\hat{C}_{i,j}$.

Rok vzniku pojistné události	Rok vývoje pojistných plnění				
	0	...	j	...	N
0	$C_{0,0}$...	$C_{0,j}$...	$C_{0,N}$
\vdots	\vdots		\vdots		\vdots
i	$C_{i,0}$...	$C_{i,j}$...	$\hat{C}_{i,N}$
\vdots	\vdots		\vdots		\vdots
N	$C_{N,0}$...	$\hat{C}_{N,j}$...	$\hat{C}_{N,N}$

Tabulka 2: Doplněný kumulativní vývojový trojúhelník

Odhady získané pomocí metody Chain Ladder nazveme *projekce*.

Hlavním tématem této práce bude *aditivita* těchto projekcí:

Mějme dva kumulativní vývojové trojúhelníky Δ_1 a Δ_2 vytvořené pojišťovnou pro roky vzniku i , $0 \leq i \leq N$, a roky vývoje j , $0 \leq j \leq N$.

Poznámka. Pod zkráceným pojmem *trojúhelník* budeme nadále rozumět výhradně kumulativní vývojový trojúhelník dle definice na straně 2.

Projekce trojúhelníku Δ_1 získané metodou Chain Ladder označme souhrnně $Proj(\Delta_1)$ a projekce trojúhelníku Δ_2 označme analogicky $Proj(\Delta_2)$. Ajne (1994) pojednává o tom, kdy jsou projekce dvou trojúhelníků Δ_1 a Δ_2 *aditivní*, tedy kdy platí rovnost

$$Proj_{i,N}(\Delta_1) + Proj_{i,N}(\Delta_2) = Proj_{i,N}(\Delta_1 + \Delta_2) \quad \forall i \in 1, \dots, N,$$

kde pojmem $Proj_{i,N}(\Delta)$ rozumíme odhad $\hat{C}_{i,N}$ v trojúhelníku Δ . Dále uvádí, jaké jsou nutné a postačující podmínky k dosažení této rovnosti, kdy naopak dochází k nerovnosti, a jakým směrem je tato případná nerovnost orientována.

V praxi se spíše než s pojmem *trojúhelník* setkáme s pojmem *portfolio*. Portfoliem v daném kontextu rozumíme soubor pojistných kmenů, které vykazují podobné vlastnosti. Vývoj pojistného plnění na škody typu IBNR v portfoliu je poté reprezentován příslušným trojúhelníkem.

Cílem této práce bude seznámit čtenáře s problematikou aditivity projekcí v daných trojúhelnících a diskutovat o tom, zde může dělení trojúhelníku (portfolia) na více trojúhelníků (subportfolií) vést ke zpřesnění odhadu výsledné rezervy.

Vlastní text práce je členěn do čtyř kapitol. V první kapitole se nejprve seznámíme s metodou Chain Ladder, a to jak s jejím čistě deterministickým popisem, tak se stochastickými modely na ní založenými. Poté budeme ve druhé kapitole pojednávat o tom, kdy nastává aditivita projekcí dvou trojúhelníků, a kdy naopak nastává projekční nerovnost. Ve třetí kapitole si stručně představíme nástroje k výpočtu směrodatné chyby v Chain Ladder projekcích. A nakonec ve čtvrté kapitole aplikujeme nabyté poznatky při analýze reálných dat.

1. Metoda Chain Ladder

V Úvodu jsme nastínili důležitost získávání projekcí z kumulativních vývojových trojúhelníků pomocí metody Chain Ladder. V této kapitole tuto metodu představíme. Uvedeme nejprve její klasický (deterministický) popis, přičemž budeme vycházet z knih Cipra (1999) a Cipra (2015). Poté se budeme zabývat na této metodě založenými stochastickými modely, jak je prezentoval Mack (1994). Získané poznatky následně využijeme v dalších kapitolách.

1.1 Deterministický popis metody CL

Uvažujme kumulativní vývojový trojúhelník dle definice na straně 2. Metoda Chain Ladder doplňuje tento trojúhelník na čtverec pomocí odhadů $\widehat{C}_{i,j}$, jak bylo ukázáno v tabulce 2. Pro získání těchto odhadů je nejprve potřeba odhadnout tzv. vývojové faktory.

Vývojovým faktorem rozumíme poměr kumulovaných pojistných plnění mezi sousedními roky vývoje. Metoda předpokládá, že v budoucnu bude tento poměr přibližně stejný. Na tomto principu následně odhaduje budoucí hodnoty. Odhad vývojového faktoru spočteme jako

$$\widehat{f}_j = \sum_{i=0}^{N-j} C_{i,j} / \sum_{i=0}^{N-j} C_{i,j-1}, \quad 1 \leq j \leq N. \quad (1.1)$$

Poznámka. Podstatnou nevýhodou metody Chain Ladder je její špatná schopnost reagovat na potenciální budoucí změny, které by svou velikostí výrazně narušovaly trend minulých let, na němž je metoda přímo založená. Více se této problematice budeme věnovat na konci této kapitoly.

Odhad $\widehat{C}_{i,N-i+1}$ získáme z prvku $C_{i,N-i}$ přenásobením odhadem příslušného vývojového faktoru jako

$$\widehat{C}_{i,N-i+1} := C_{i,N-i} \widehat{f}_{N-i+1} \quad \forall i = 1, \dots, N. \quad (1.2)$$

Obecně odhad $\widehat{C}_{i,j}$, $N - i + 1 \leq j \leq N$, vypočteme z $C_{i,N-i}$ jako

$$\widehat{C}_{i,j} := C_{i,N-i} \widehat{f}_{N-i+1} \cdots \widehat{f}_j \quad \forall i = 1, \dots, N. \quad (1.3)$$

Poznámka. Pro $0 \leq j \leq N - i$ se odhad $\widehat{C}_{i,j}$ přímo rovná $C_{i,j}$.

Výši rezervy odhadované na konci roku N určíme z doplněného kumulativního vývojového trojúhelníku snadno jako rozdíl součtu hodnot v posledním (N -tém) sloupci a součtu hodnot na vedlejší diagonále.

$$\widehat{IBNR}_N = \sum_{i=0}^N \widehat{C}_{i,N} - \sum_{i=0}^N C_{i,N-i}. \quad (1.4)$$

Poznámka. Pro $i = 0$ se $\widehat{C}_{0,N}$ rovná $C_{0,N}$.

1.2 Stochastické modely založené na metodě CL

Nechť nyní všechna kumulovaná pojistná plnění $C_{i,j}$ jsou *náhodné veličiny*. Pak namísto deterministické metody Chain Ladder (1.2) dostáváme zobecněný stochastický model

$$E(C_{i,j}) = E(C_{i,j-1})f_j, \quad 1 \leq j \leq N, \quad (1.5)$$

kde f_j jsou *neznámé parametry*.

Následující věta je i s důkazem převzatá z článku Mack (1994).

Věta 1.1. *Uvažujme náhodné veličiny*

$$S_{i,j} = C_{i,j} - C_{i,j-1}, \quad 1 \leq j \leq N,$$

a dodefinujme $S_{i,0} := C_{i,0}$. Pak lze (1.5) převést na ekvivalentní model

$$E(S_{i,j}) = x_i y_j, \quad (1.6)$$

kde x_i , $0 \leq i \leq N$, a y_j , $0 \leq j \leq N$, jsou *neznámé parametry* a $\sum_{j=0}^N y_j = 1$.

Důkaz. Z (1.3) a (1.5) platí rovnost

$$E(C_{i,N}) = E(C_{i,j})f_{j+1} \cdots f_N.$$

Aplikací střední hodnoty na $S_{i,j}$ a dosazením této rovnosti dostáváme

$$\begin{aligned} E(S_{i,j}) &= E(C_{i,j} - C_{i,j-1}) = E(C_{i,j}) - E(C_{i,j-1}) \\ &= E(C_{i,N})(f_{j+1} \cdots f_N)^{-1} - E(C_{i,N})(f_j \cdots f_N)^{-1} \\ &= E(C_{i,N}) \left((f_{j+1} \cdots f_N)^{-1} - (f_j \cdots f_N)^{-1} \right). \end{aligned}$$

Neznámé parametry x_i a y_j definujeme jako

$$\begin{aligned} x_i &:= E(C_{i,N}), \quad y_0 := (f_1 \cdots f_N)^{-1}, \quad y_N := 1 - (f_N)^{-1}, \\ y_j &:= \left((f_{j+1} \cdots f_N)^{-1} - (f_j \cdots f_N)^{-1} \right), \quad 1 \leq j \leq N-1. \end{aligned}$$

Pak je podmínka $\sum_{j=0}^N y_j = 1$ splněna a platí (1.5) \implies (1.6). Obrácenou implikaci dostaneme snadno, neboť

$$E(C_{i,j}) = E(S_{i,0}) + \dots + E(S_{i,j}),$$

z čehož odvodíme f_j jako

$$\frac{E(C_{i,j})}{E(C_{i,j-1})} = \frac{x_i(y_0 + \dots + y_{j-1} + y_j)}{x_i(y_0 + \dots + y_{j-1})} =: f_j, \quad 1 \leq j \leq N, \quad (1.7)$$

a tedy platí (1.6) \implies (1.5). □

Z (1.7) a ekvivalence mezi oběma modely plyne, že modely (1.5) a (1.6) disponují oba přesně $2n + 1$ parametry.

Jsou-li náhodné veličiny $S_{i,j}$ navíc *nezávislé* pro $0 \leq i, j \leq N$, dokážeme pomocí vhodné metody najít odhady parametrů x_i, y_j pro libovolné rozdělení $S_{i,j}$ (viz Mack, 1994). Příkladem takového rozdělení může být například normální rozdělení

$$S_{i,j} \sim N(x_i y_j, \sigma^2).$$

Jak ale Mack (1994) dále uvádí, výsledná rezerva vypočtená pomocí takového modelu se může výrazně lišit od rezervy získané pomocí deterministické metody Chain Ladder. To je způsobeno tím, že v klasické metodě Chain Ladder (1.2) jsou k výpočtu projekcí využívána skutečná minulá data pojišťovny $C_{i,j}$, zatímco ve stochastickém modelu (1.5) jsou použity jejich střední hodnoty $E(C_{i,j})$. Mack proto zavádí nový *stochastický* model, který stejně jako deterministická metoda Chain Ladder připisuje větší relevanci skutečným napozorovaným hodnotám. V praxi je tento model často nazýván *Mackův model*.

1.2.1 Mackův model

Uvažujme stochastický model

$$E(C_{i,j} | C_{i,0}, \dots, C_{i,j-1}) = C_{i,j-1} f_j, \quad 0 \leq i \leq N, 1 \leq j \leq N, \quad (1.8)$$

kde f_j jsou neznámé parametry. Vytvořme množinu kumulovaných napozorovaných pojistných plnění příslušných roku vzniku i

$$D_{i,N-i} := \{C_{i,j}; 0 \leq j \leq N - i\}.$$

Dosazením této množiny napozorovaných dat za podmínku v modelu (1.8) dostáváme podmíněnou střední hodnotu $E(C_{i,j} | D_{i,N-i})$, čímž jsme získali obdobu modelu (1.5), jež je ale podstatně bližší deterministické metodě Chain Ladder, neboť na rozdíl od nepodmíněné střední hodnoty $E(C_{i,j})$ v modelu (1.5) bere $E(C_{i,j} | D_{i,N-i})$ v úvahu skutečná minulá data.

Pokud jsou navíc dva různé roky vzniku $i \neq k$ navzájem nezávislé:

$$\{C_{i,0}, \dots, C_{i,N}\} \perp\!\!\!\perp \{C_{k,0}, \dots, C_{k,N}\}, \quad i \neq k, \quad (1.9)$$

pak platí následující věta.

Věta 1.2. *Za platnosti modelu (1.8) a podmínky (1.9) platí rovnost*

$$E(C_{i,j} | D_{i,N-i}) = C_{i,n-i} f_{n-i+1} \cdots f_j, \quad N - i + 1 \leq j \leq N.$$

Důkaz. viz Mack (1994, str. 234)

Odhadneme-li neznámé parametry f_j pomocí (1.1), rezerva získaná pomocí modelu (1.8) je identická s rezervou získanou pomocí deterministické metody Chain Ladder (1.2). Model (1.8) navíc disponuje pouze n parametry, což je významně méně, než kolika disponovaly modely (1.5) a (1.6).

Následující definice je převzatá z práce Wüthrich a Merz (2008).

Definice 1.1 (Mackův model). *Nechť jsou dva různé roky vzniku $i \neq k$ navzájem nezávislé, tzn. necht' je splněna podmínka (1.9). Necht' $(C_{i,j})_{0 \leq j \leq N}$ tvoří pro každý rok vzniku i markovský řetězec, a necht' existují vývojové faktory $f_1, \dots, f_N > 0$ a rozptylové parametry $\sigma_1^2, \dots, \sigma_N^2 > 0$ takové, že platí*

$$\begin{aligned} E(C_{i,j} | C_{i,j-1}) &= C_{i,j-1} f_j, & 0 \leq i \leq N, 1 \leq j \leq N, \\ \text{Var}(C_{i,j} | C_{i,j-1}) &= C_{i,j-1} \sigma_j^2, & 0 \leq i \leq N, 1 \leq j \leq N. \end{aligned}$$

Pak tento model nazveme Mackův model.

Poznámka. Takto definovaný model nám oproti modelu (1.8) díky zavedení druhých momentů dává možnost zkoumat odchylky. Této problematice se budeme stručně věnovat v kapitole 3.

Na závěr této kapitoly uvedeme Mackovu diskusi o tom, kdy se vyplatí aplikovat Mackův model, a kdy je naopak lepší sáhnout po modelu (1.5), popř. jiné metodě.

Model (1.5) odhaduje rezervu $R_i = C_{i,N} - C_{i,N-i} = S_{i,N-i+1} + \dots + S_{i,N}$ příslušnou roku vzniku i tak, že odhadne její střední hodnotu

$$E(R_i) = E(S_{i,N-i+1} + \dots + S_{i,N}).$$

Jak Mack (1994) uvádí, pokud bychom znali skutečné hodnoty parametrů x_i, y_j modelu (1.5) a skutečné hodnoty parametrů f_j Mackova modelu, byli bychom schopni ukázat rozdíl mezi oběma modely následujícím způsobem:

Rezerva získaná pomocí modelu (1.5) by byla nezávisle na napozorovaných datech

$$E(R_i) = x_i(y_{N-i+1} + \dots + y_N).$$

Naopak Mackův model by v závislosti na napozorovaných datech spočetl rezervu

$$E(R_i | D_{i,N-i}) = C_{i,N-i} (f_{N-i+1} \cdots f_N - 1).$$

To v zásadě znamená, že Mackův model (a metodu Chain Ladder obecně) bychom měli použít za předpokladu, že vychýlení

$$C_{i,N-i} - E(C_{i,N-i})$$

nám dává dobrou představu o budoucím vývoji. V opačném případě může být vhodnější zvolit jinou metodu. Konkrétně použití modelu (1.5) je nutné dobře zvážit, neboť model může být nestabilní z důvodu velkého počtu parametrů.

2. Aditivita projekcí získaných pomocí metody Chain Ladder

Pojišťovna vytváří (kumulativní) vývojové trojúhelníky pro odhad technický rezerv v daném portfoliu. V této kapitole se budeme zabývat aditivitou projekcí v uvažovaných trojúhelnících. Budeme zkoumat, za jakých podmínek jsou projekce daných trojúhelníků aditivní a kdy naopak nastává projekční nerovnost. Pokusíme se odpovědět na otázku, zda není vhodné pro určitou část portfolia projektovat rezervy zvlášť, pokud se tato část chová v jistém smyslu odlišně od jeho zbytku.

Nejprve obecně definujeme pojmy *projekce* a *aditivita projekcí*.

Definice 2.1 (Projekce). *Nechť Δ je kumulativní vývojový trojúhelník a necht $\hat{C}_{i,j}$ je Chain Ladder odhad celkových budoucích kumulovaných pojistných plnění příslušných roku vzniku i , $0 \leq i \leq N$, a roku vývoje j , $N - i + 1 \leq j \leq N$, v tomto trojúhelníku. Pak tyto odhady nazveme projekce metody Chain Ladder v trojúhelníku Δ a budeme je souhrnně značit $Proj(\Delta)$.*

Definice 2.2 (Aditivita projekcí). *Nechť $\Delta_1, \dots, \Delta_n$ jsou kumulativní vývojové trojúhelníky příslušné rokům vzniku i , $0 \leq i \leq N$, a rokům vývoje j , $0 \leq j \leq N$, a necht $Proj(\Delta_1), \dots, Proj(\Delta_n)$ jsou jejich odpovídající projekce. Pak řekneme, že projekce trojúhelníků $\Delta_1, \dots, \Delta_n$ jsou aditivní, pokud platí rovnost*

$$\sum_{k=1}^n Proj_{i,N}(\Delta_k) = Proj_{i,N}\left(\sum_{k=1}^n \Delta_k\right) \quad \forall i \in 0, \dots, N, \quad (2.1)$$

kde pojmem $Proj_{i,N}(\Delta)$ rozumíme odhad $\hat{C}_{i,N}$ v příslušném trojúhelníku Δ a součtem $\sum_{k=1}^n \Delta_k$ rozumíme maticový součet po složkách.

Povšimněme si, že v definici aditivity projekcí se vyskytují pouze projekce pro poslední (N -tý) rok vývoje. Důvodem je fakt, že pojistná plnění v trojúhelnících jsou kumulovaná, a tedy při výpočtu výsledné IBNR rezervy (1.4) stačí znát pouze poslední sloupec (a vedlejší diagonálu) dané matice.

Ajne (1994) uvádí, že odhad rozdělení kumulovaných pojistných plnění je určen odhadnutými vývojovými faktory \hat{f}_j jako

$$\hat{\Psi}(0), \dots, \hat{\Psi}(N), \quad (2.2)$$

kde

$$\hat{\Psi}(j) = \frac{1}{\hat{f}_{j+1} \cdots \hat{f}_N}, \quad 0 \leq j \leq N-1, \quad \text{a} \quad \hat{\Psi}(N) = 1. \quad (2.3)$$

Je důležité si uvědomit, že pojmem *rozdělení* zde nerozumíme pravděpodobnostní rozdělení ve smyslu definice $F(x) = P[X \leq x]$, nýbrž rozdělení ve smyslu proporcionálního rozložení celkového kumulovaného pojistného plnění napříč roky vývoje j . Metoda Chain Ladder navíc předpokládá, že toto rozložení je pro každý rok vzniku i identické.

Z (1.3) a (2.3) pak platí, že

$$C_{i,N-i} = \widehat{C}_{i,N} \widehat{\Psi}(N-i). \quad (2.4)$$

Rovnost (2.4) nám dává důležitý nástroj pro další zkoumání aditivity projekcí.

Poznámka. Prvky trojúhelníku Δ_k budeme značit $C_{i,j}^{(k)}$, odhady jeho vývojových faktorů $\widehat{f}_j^{(k)}$. Odhad rozdělení (2.2) příslušný Δ_k budeme značit $\widehat{\Psi}^{(k)}$.

2.1 Podmínky pro aditivitu projekcí

Začněme větou o nutných a postačujících podmínkách k dosažení aditivity projekcí dvou trojúhelníků. Tato věta je včetně důkazu a důsledků převzatá z článku (Ajne, 1994).

Věta 2.1. *Uvažujme dva kumulativní vývojové trojúhelníky Δ_1 a Δ_2 příslušné rokům vzniku i , $0 \leq i \leq N$, a rokům vývoje j , $0 \leq j \leq N$. Jejich projekce jsou aditivní právě tehdy, když pro každý rok vzniku i , $1 \leq i \leq N$, platí*

$$\widehat{\Psi}^{(1)}(N-i) = \widehat{\Psi}^{(2)}(N-i) \quad (2.5)$$

nebo alternativně

$$\frac{\widehat{C}_{i,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}} = \frac{\widehat{C}_{i,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}. \quad (2.6)$$

Důkaz. Definujme trojúhelník Δ_Σ pomocí maticového součtu po složkách jako

$$\Delta_\Sigma := \Delta_1 + \Delta_2. \quad (2.7)$$

Projekci $Proj_{i,N}(\Delta_\Sigma)$ budeme dále značit jako $\widehat{C}_{i,N}^{(\Sigma)}$. Z (2.1) pak přeznačením a volbou $n = 2$ dostáváme rovnost

$$\widehat{C}_{i,N}^{(\Sigma)} = \widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)} \quad \forall i \in 0, \dots, N. \quad (2.8)$$

Pro volbu $i = 0$ je tato rovnost triviální, neboť se ve všech třech případech jedná o pojišťovnou napozorované hodnoty.

Nechť $i > 0$. Pak za použití (2.4) a (2.7) dostáváme, že

$$\begin{aligned} \frac{\widehat{C}_{i,N}^{(\Sigma)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(\Sigma)}} &= \frac{\widehat{C}_{i,N}^{(\Sigma)}}{\sum_{k=0}^{i-1} C_{k,N-i}^{(\Sigma)} \left(\widehat{\Psi}^{(\Sigma)}(N-i) \right)^{-1}} = \frac{\widehat{C}_{i,N}^{(\Sigma)} \widehat{\Psi}^{(\Sigma)}(N-i)}{\sum_{k=0}^{i-1} C_{k,N-i}^{(\Sigma)}} \\ &= \frac{C_{i,N-i}^{(1)} + C_{i,N-i}^{(2)}}{\sum_{k=0}^{i-1} \left(\widehat{C}_{k,N}^{(1)} \widehat{\Psi}^{(1)}(N-i) + \widehat{C}_{k,N}^{(2)} \widehat{\Psi}^{(2)}(N-i) \right)}, \end{aligned}$$

což dále upravujeme na

$$\frac{\widehat{C}_{i,N}^{(\Sigma)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(\Sigma)}} = \frac{\widehat{C}_{i,N}^{(1)} \widehat{\Psi}^{(1)}(N-i) + \widehat{C}_{i,N}^{(2)} \widehat{\Psi}^{(2)}(N-i)}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} \widehat{\Psi}^{(1)}(N-i) + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)} \widehat{\Psi}^{(2)}(N-i)}.$$

Ekvivalentními úpravami nakonec dojdeme do tvaru

$$\frac{\widehat{C}_{i,N}^{(\Sigma)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}} = P_i * \frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(\Sigma)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}, \quad (2.9)$$

kde P_i je podíl mezi

$$\frac{\widehat{C}_{i,N}^{(1)} \widehat{\Psi}^{(1)}(N-i) + \widehat{C}_{i,N}^{(2)} \widehat{\Psi}^{(2)}(N-i)}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}} \quad (2.10)$$

a

$$\frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} \widehat{\Psi}^{(1)}(N-i) + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)} \widehat{\Psi}^{(2)}(N-i)}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}. \quad (2.11)$$

Nechť jsou projekce trojúhelníků Δ_1 a Δ_2 aditivní, tedy necht' je splněna rovnost (2.8). Z (2.9) je pak $P_i = 1$ pro každý rok vzniku i , $1 \leq i \leq N$, a tedy (2.10) = (2.11) pro každý rok vzniku i , $1 \leq i \leq N$. Tuto rovnost upravíme do tvaru

$$\begin{aligned} & \left[\frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}} - \frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} \right] \widehat{\Psi}^{(1)}(N-i) \\ & + \left[\frac{\widehat{C}_{i,N}^{(2)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}} - \frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} \right] \widehat{\Psi}^{(2)}(N-i) = 0. \end{aligned} \quad (2.12)$$

Platí-li, že $\widehat{\Psi}^{(1)}(N-i) = \widehat{\Psi}^{(2)}(N-i)$, pak je rovnost (2.12) jistě splněna.

Položme tedy $\widehat{\Psi}^{(1)}(N-i) \neq \widehat{\Psi}^{(2)}(N-i)$. Pak rovnost (2.12) nastává právě tehdy, když

$$\frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}} - \frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} = 0 \quad (2.13)$$

a

$$\frac{\widehat{C}_{i,N}^{(2)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}} - \frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} = 0. \quad (2.14)$$

Rovnosti (2.13) a (2.14) lze ekvivalentně zapsat jako

$$\frac{\frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} = 1 \quad \text{a} \quad \frac{\frac{\widehat{C}_{i,N}^{(2)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} = 1,$$

z čehož vyplývá, že

$$\frac{\frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} = \frac{\frac{\widehat{C}_{i,N}^{(2)}}{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}. \quad (2.15)$$

Pomocí ekvivalentních úprav nakonec z (2.15) dostáváme, že

$$\frac{\widehat{C}_{i,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}} = \frac{\widehat{C}_{i,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}. \quad (2.16)$$

Nechť nyní naopak pro každý rok vzniku i , $1 \leq i \leq N$, platí, že

$$\widehat{\Psi}^{(1)}(N-i) = \widehat{\Psi}^{(2)}(N-i) \quad \text{nebo} \quad \frac{\widehat{C}_{i,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}} = \frac{\widehat{C}_{i,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}.$$

Pak $P_i = 1$ pro každý rok vzniku i , $1 \leq i \leq N$, a rovnost (2.8) dostáváme z (2.9) indukcí za předpokladu, že pro $i = 0$ je rovnost (2.8) triviálně splněna. \square

Poznámka. Poměry v (2.16) nazveme *míry růstu* objemu pojistných plnění v daných portfoliích pro rok vzniku i .

Věta 2.1 pak dává následující důsledky.

Důsledek 1. Splňují-li dané dva trojúhelníky Δ_1 a Δ_2 podmínku (2.5) pro každý rok vzniku i , $0 \leq i \leq N$, pak jsou příslušné projekce v subportfoliích reprezentovaných těmito trojúhelníky aditivní a řekneme, že tato dvě subportfolia mají *stejně dlouhé chvosty*.

Důsledek 2. Splňují-li dané dva trojúhelníky Δ_1 a Δ_2 podmínku (2.6) pro každý rok vzniku i , $0 \leq i \leq N$, pak jsou příslušné projekce v subportfoliích reprezentovaných těmito trojúhelníky aditivní a řekneme, že tato dvě subportfolia mají *stejnou míru růstu* pro každý rok vzniku i , $0 \leq i \leq N$.

Na základě věty 2.1 dále zformulujeme následující větu, kterou Ajne (1994) explicitně neuvádí.

Věta 2.2. *Nechť jsou projekce trojúhelníku Δ_1 a Δ_2 aditivní. Splňují-li tyto dva trojúhelníky pro daný konkrétní rok vzniku i podmínku (2.5), pak v tomto roce vzniku i platí pro trojúhelník Δ_Σ vzniklý z Δ_1 a Δ_2 maticovým součtem po složkách, že*

$$\widehat{\Psi}^{(\Sigma)}(N-i) = \widehat{\Psi}^{(1)}(N-i) = \widehat{\Psi}^{(2)}(N-i).$$

Alternativně, splňují-li trojúhelníky Δ_1 a Δ_2 pro daný konkrétní rok vzniku i podmínku (2.6), pak v tomto roce vzniku i pro trojúhelník Δ_Σ platí, že

$$\frac{\widehat{C}_{i,N}^{(\Sigma)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(\Sigma)}} = \frac{\widehat{C}_{i,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}} = \frac{\widehat{C}_{i,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}.$$

Důkaz. Použijeme stejné značení jako v důkazu věty 2.1. Za předpokladu aditivity projekcí je pro každý rok vzniku i splněna rovnost (2.8). Aplikací (2.4) na (2.8) dostáváme, že

$$\frac{1}{\widehat{\Psi}^{(\Sigma)}(N-i)} C_{i,N-i}^{(\Sigma)} = \frac{1}{\widehat{\Psi}^{(1)}(N-i)} C_{i,N-i}^{(1)} + \frac{1}{\widehat{\Psi}^{(2)}(N-i)} C_{i,N-i}^{(2)}.$$

Nechť je splněna podmínka (2.5). Pak

$$\frac{1}{\widehat{\Psi}^{(\Sigma)}(N-i)} C_{i,N-i}^{(\Sigma)} = \frac{1}{\widehat{\Psi}^{(1)}(N-i)} \left(C_{i,N-i}^{(1)} + C_{i,N-i}^{(2)} \right),$$

a protože $\Delta_\Sigma = \Delta_1 + \Delta_2$, a tedy $C_{i,N-i}^{(\Sigma)} = C_{i,N-i}^{(1)} + C_{i,N-i}^{(2)}$, musí pro daný rok vzniku i nutně platit, že

$$\widehat{\Psi}^{(\Sigma)}(N-i) = \widehat{\Psi}^{(1)}(N-i) \stackrel{(2.5)}{=} \widehat{\Psi}^{(2)}(N-i).$$

Nechť je nyní splněna podmínka (2.6). Rovnost (2.6) rozšíříme pomocí ekvivalentních úprav do tvaru

$$\frac{\widehat{C}_{i,N}^{(2)}}{\widehat{C}_{i,N}^{(1)}} + \frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(1)}} = \frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}} + \frac{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}}.$$

Po několika dalších úpravách dostáváme, že

$$\frac{\widehat{C}_{i,N}^{(\Sigma)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(\Sigma)}} \stackrel{(2.8)}{=} \frac{\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)} + \sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} = \frac{\widehat{C}_{i,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}} \stackrel{(2.6)}{=} \frac{\widehat{C}_{i,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}},$$

a tedy se pro daný rok vzniku i *míra růstu* v trojúhelníku Δ_Σ rovná *mírák růstu* v trojúhelnících Δ_1 a Δ_2 , čímž je důkaz hotov. □

Výsledek věty 2.2 nazveme *dědičnost vlastnosti*.

Nyní se nabízí aditivitu projekcí zobecnit pro více trojúhelníků. Analogickým postupem jako v důkazu věty 2.1 bychom došli k tomu, že projekce n trojúhelníků jsou aditivní právě tehdy, když

$$\frac{\sum_{m=1}^n \widehat{C}_{i,N}^{(m)} \widehat{\Psi}^{(m)}(N-i)}{\sum_{m=1}^n \widehat{C}_{i,N}^{(m)}} = \frac{\sum_{m=1}^n \sum_{k=1}^{i-1} \widehat{C}_{k,N}^{(m)} \widehat{\Psi}^{(m)}(N-i)}{\sum_{m=1}^n \sum_{k=1}^{i-1} \widehat{C}_{k,N}^{(m)}} \quad \forall i \in 1, \dots, N.$$

Vyjadřovat pro tuto rovnost explicitní podmínky by ovšem bylo velmi obtížné už pro volbu $n = 3$.

2.2 Projekční nerovnost pro dva trojúhelníky

V praxi se s případy, kdy jsou projekce daných trojúhelníků aditivní, prakticky nesetkáme. Mnohem podstatnější je tedy zkoumat, kdy v (2.1) nastává nerovnost a jakým směrem je orientována. V rámci této práce budeme uvažovat výhradně volbu $n = 2$. Pokud bychom chtěli projekční nerovnost zobecnit na více trojúhelníků, případ by se výrazně zkomplikoval. Ajne (1994) zavádí následující pojmy.

Pokud platí, že

$$\widehat{\Psi}^{(1)}(N-i) \leq \widehat{\Psi}^{(2)}(N-i) \quad \forall i \in 1, \dots, N, \quad (2.17)$$

pak subportfolio Δ_1 má *alespoň tak dlouhé chvosty* jako subportfolio Δ_2 .

Pokud platí, že

$$\frac{\widehat{C}_{i,N}^{(1)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}} \geq \frac{\widehat{C}_{i,N}^{(2)}}{\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}} \quad \forall i \in 1, \dots, N, \quad (2.18)$$

pak subportfolio Δ_1 *roste alespoň tak rychle* jako subportfolio Δ_2 .

Tím jsme získali všechny prostředky potřebné ke zformulování následující věty o postačujících podmínkách pro projekční nerovnost dvou trojúhelníků.

Věta 2.3. *Nechť platí (2.17) a zároveň necht platí (2.18). Pak*

$$\sum_{k=1}^2 Proj_{i,N}(\Delta_k) \geq Proj_{i,N}\left(\sum_{k=1}^2 \Delta_k\right) \quad \forall i \in 1, \dots, N. \quad (2.19)$$

Nechť platí (2.17) a zároveň necht platí (2.18) s obrácenou nerovností. Pak

$$\sum_{k=1}^2 Proj_{i,N}(\Delta_k) \leq Proj_{i,N}\left(\sum_{k=1}^2 \Delta_k\right) \quad \forall i \in 1, \dots, N. \quad (2.20)$$

Důkaz. Následující důkaz je opět převzatý z článku Ajne (1994).

Poměr P_i z důkazu věty 2.1 lze interpretovat jako poměr vážených průměrů (2.10) a (2.11), kde v prvním případě jsou jako váhy použity $\widehat{C}_{i,N}^{(1)}$ a $\widehat{C}_{i,N}^{(2)}$ a v druhém případě $\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(1)}$ a $\sum_{k=0}^{i-1} \widehat{C}_{k,N}^{(2)}$.

Z (2.18) plyne, že v (2.10) má $\widehat{\Psi}^{(1)}(N - i)$ větší (nebo stejnou) váhu než v (2.11) pro každý rok vzniku i . Necht zároveň platí (2.17). Pak

$$(2.10) \leq (2.11) \quad \forall i \in 1, \dots, N \implies P_i \leq 1 \quad \forall i \in 1, \dots, N \xrightarrow{(2.9)} (2.19).$$

Analogicky, obrátíme-li v (2.18) znaménko nerovnosti, dostáváme, že v (2.10) má $\widehat{\Psi}^{(1)}(N - i)$ menší (nebo stejnou) váhu než v (2.11) pro každý rok vzniku i . Necht zároveň platí (2.17). Pak

$$(2.10) \geq (2.11) \quad \forall i \in 1, \dots, N \implies P_i \geq 1 \quad \forall i \in 1, \dots, N \xrightarrow{(2.9)} (2.20). \quad \square$$

Až do tohoto momentu jsme nejprve uvažovali více trojúhelníků (portfolií) a následně jsme zkoumali, jaká bude výše výsledné rezervy v závislosti na tom, zda trojúhelníky nejprve po složkách sčítáme, a až poté aplikujeme metodu Chain Ladder, a nebo naopak nejprve aplikujeme metodu Chain Ladder na každý z trojúhelníků zvlášť a následně sčítáme projekce v posledních sloupcích.

V praxi se zpravidla setkáváme s opačnou situací. Pojišťovna vytváří portfolia, jež se běžně skládají z více subportfolií reprezentujících různé typy škodních událostí. U těchto typů škodních událostí se předpokládá podobný trend vývoje pojistných plnění v příslušných vývojových trojúhelnících. Takový předpoklad zajistí relativní přesnost odhadu výsledné rezervy v portfoliu z nich složeném.

Ajne v závěru svého článku (viz Ajne (1994)) popisuje situaci, kdy by mohlo dojít k porušení tohoto předpokladu. Uvádí, že by v jistých případech mohlo být vhodné jedno z takových subportfolií od zbytku portfolia oddělit.

Z věty 2.3 totiž plyne, že chová-li se jedno subportfolio odlišně od zbytku portfolia, a to jak ve smyslu *délky chvostů*, tak ve smyslu *míry růstu*, jedná se o důležitý signál. Takové chování by totiž mohlo vést k výrazné odchylce ve výsledné rezervě.

Má-li toto subportfolio výrazně delší chvosty než zbytek portfolia a zároveň roste výrazně rychleji, pak by dle (2.19) mohlo dojít k podhodnocení celkové rezervy, pokud bychom toto subportfolio neoddělili. To znamená, že by výsledná celková rezerva nemusela dostatečně pokrýt vzniklé budoucí náklady.

Naopak má-li toto subportfolio výrazně delší chvosty než zbytek portfolia, ale roste výrazně pomaleji, mohlo by dle (2.20) při jeho ponechání v celkovém portfoliu dojít k nadhodnocení rezervy. To by v praxi znamenalo, že pojišťovna sice dokáže pokrýt vzniklé budoucí náklady na pojistná plnění bez problémů, ovšem na úkor alokace značných finančních prostředků na rezervotvorný účet. Tyto finanční prostředky vzhledem k nadhodnocení rezervy nejsou ve skutečnosti potřebné, a pojišťovna by je mohla v rámci své činnosti využít lépe jinde.

V obou případech je samozřejmě nutná jistá dávka opatrnosti a je vhodné provést další analýzu.

Získali jsme tak odpověď na otázku z úvodu této kapitoly.

Věta 2.3 popisuje pouze situaci, kdy je projekční nerovnost orientována stejným směrem v každém roce vzniku i . Je zřejmé, že obecně tomu tak být nemusí. Na závěr této podkapitoly budeme proto diskutovat o tom, lze-li určit orientaci znaménka nerovnosti pro každý konkrétní rok vzniku i zvlášť.

Věta 2.4. *Nechť Δ_1 a Δ_2 jsou kumulativní vývojové trojúhelníky příslušné rokům vzniku i , $0 \leq i \leq N$, a rokům vývoje j , $0 \leq j \leq N$. Jestliže pro nějaký konkrétní rok vzniku i je*

$$\widehat{\Psi}^{(1)}(N-i) \leq \widehat{\Psi}^{(2)}(N-i),$$

pak pro trojúhelník Δ_Σ vzniklý z Δ_1 a Δ_2 maticovým součtem po složkách platí

$$\widehat{\Psi}^{(1)}(N-i) \leq \widehat{\Psi}^{(\Sigma)}(N-i) \leq \widehat{\Psi}^{(2)}(N-i), \quad (2.21)$$

a analogicky pro obrácené nerovnosti.

Důkaz. Dle zavedeného značení

$$\widehat{\Psi}^{(1)}(N-i) = \frac{1}{\widehat{f}_{N-i+1}^{(1)} \cdots \widehat{f}_N^{(1)}} \quad \text{a} \quad \widehat{\Psi}^{(2)}(N-i) = \frac{1}{\widehat{f}_{N-i+1}^{(2)} \cdots \widehat{f}_N^{(2)}}.$$

Pro trojúhelník $\Delta_\Sigma = \Delta_1 + \Delta_2$ platí, že

$$\widehat{\Psi}^{(\Sigma)}(N-i) = \frac{1}{\widehat{f}_{N-i+1}^{(\Sigma)} \cdots \widehat{f}_N^{(\Sigma)}},$$

kde

$$\widehat{f}_j^{(\Sigma)} = \sum_{i=0}^{N-j} (C_{i,j}^{(1)} + C_{i,j}^{(2)}) / \sum_{i=0}^{N-j} (C_{i,j-1}^{(1)} + C_{i,j-1}^{(2)}), \quad 1 \leq j \leq N.$$

Nerovnost (2.21) pak přepíšeme do tvaru

$$\widehat{f}_{N-i+1}^{(1)} \cdots \widehat{f}_N^{(1)} \geq \widehat{f}_{N-i+1}^{(\Sigma)} \cdots \widehat{f}_N^{(\Sigma)} \geq \widehat{f}_{N-i+1}^{(2)} \cdots \widehat{f}_N^{(2)}. \quad (2.22)$$

Z (1.1) lze $\widehat{f}_j^{(1)}$ a $\widehat{f}_j^{(2)}$ zapsat jako

$$\begin{aligned} \widehat{f}_j^{(1)} &:= \frac{P_j}{Q_j}, \text{ kde } P_j = \sum_{i=0}^{N-j} C_{i,j}^{(1)}, Q_j = \sum_{i=0}^{N-j} C_{i,j-1}^{(1)}, \\ \widehat{f}_j^{(2)} &:= \frac{R_j}{S_j}, \text{ kde } R_j = \sum_{i=0}^{N-j} C_{i,j}^{(2)}, S_j = \sum_{i=0}^{N-j} C_{i,j-1}^{(2)}. \end{aligned}$$

Potom pro $\widehat{f}_j^{(\Sigma)}$ platí, že

$$\widehat{f}_j^{(\Sigma)} := \frac{P_j + R_j}{Q_j + S_j}.$$

Dosažením do (2.22) dostáváme

$$\frac{P_{N-i+1}}{Q_{N-i+1}} \cdots \frac{P_N}{Q_N} \geq \frac{P_{N-i+1} + R_{N-i+1}}{Q_{N-i+1} + S_{N-i+1}} \cdots \frac{P_N + R_N}{Q_N + S_N} \geq \frac{R_{N-i+1}}{S_{N-i+1}} \cdots \frac{R_N}{S_N}.$$

Z předpokladu tvrzení platí, že

$$\frac{P_{N-i+1}}{Q_{N-i+1}} \cdots \frac{P_N}{Q_N} \geq \frac{R_{N-i+1}}{S_{N-i+1}} \cdots \frac{R_N}{S_N}, \quad (2.23)$$

což přenásobením $\frac{S_{N-i+1} \cdots S_N}{P_{N-i+1} \cdots P_N}$ upravíme do tvaru

$$\frac{S_{N-i+1}}{Q_{N-i+1}} \cdots \frac{S_N}{Q_N} \geq \frac{R_{N-i+1}}{P_{N-i+1}} \cdots \frac{R_N}{P_N}. \quad (2.24)$$

Poznámka. Dodejme, že P_j, Q_j, R_j, S_j jsou z podstaty kumulativních vývojových trojúhelníků kladná čísla.

Nyní ukážeme, že za platnosti předpokladu (2.23) platí, že

$$\frac{P_{N-i+1}}{Q_{N-i+1}} \cdots \frac{P_N}{Q_N} \geq \frac{P_{N-i+1} + R_{N-i+1}}{Q_{N-i+1} + S_{N-i+1}} \cdots \frac{P_N + R_N}{Q_N + S_N}. \quad (2.25)$$

Tuto nerovnost upravíme opět pomocí ekvivalentní úpravy do tvaru

$$\frac{Q_{N-i+1} + S_{N-i+1}}{Q_{N-i+1}} \cdots \frac{Q_N + S_N}{Q_N} \geq \frac{P_{N-i+1} + R_{N-i+1}}{P_{N-i+1}} \cdots \frac{P_N + R_N}{P_N},$$

který lze dále upravit na

$$\left(1 + \frac{S_{N-i+1}}{Q_{N-i+1}}\right) \cdots \left(1 + \frac{S_N}{Q_N}\right) \geq \left(1 + \frac{R_{N-i+1}}{P_{N-i+1}}\right) \cdots \left(1 + \frac{R_N}{P_N}\right). \quad (2.26)$$

Nerovnost (2.24) lze interpretovat jako porovnání objemů dvou i -dimenzionálních kvádrů popořadě o stranách délek $\frac{S_j}{Q_j}$ a $\frac{R_j}{P_j}$. Prodloužíme-li každou ze stran v obou tělesech o jedna, objem tělesa s větším původním objemem se zvětší více než objem tělesa s menším původním objemem. Orientace nerovnosti je tedy zachována. Dle této úvahy pak (2.24) \implies (2.26), a tedy (2.23) \implies (2.25).

Analogicky ukážeme, že platí

$$\frac{P_{N-i+1} + R_{N-i+1}}{Q_{N-i+1} + S_{N-i+1}} \cdots \frac{P_N + R_N}{Q_N + S_N} \geq \frac{R_{N-i+1}}{S_{N-i+1}} \cdots \frac{R_N}{S_N}. \quad (2.27)$$

Postupnými úpravami dostáváme

$$\left(\frac{P_{N-i+1}}{R_{N-i+1}} + 1 \right) \cdots \left(\frac{P_N}{R_N} + 1 \right) \geq \left(\frac{Q_{N-i+1}}{S_{N-i+1}} + 1 \right) \cdots \left(\frac{Q_N}{S_N} + 1 \right).$$

Opětovným upravením (2.23) a analogickou úvahou o objemech dostáváme, že (2.23) \implies (2.27). Analogicky bychom dokázali větu pro obrácené nerovnosti. \square

Poznámka 1. Dle věty 2.4 lze odhad rozdělení $\widehat{\Psi}^{(\Sigma)}$ shora a zdola omezit pomocí odhadů rozdělení $\widehat{\Psi}^{(1)}$ a $\widehat{\Psi}^{(2)}$.

Poznámka 2. Jsou-li projekce trojúhelníků Δ_1 a Δ_2 aditivní, pak dává věta 2.4 stejný výsledek jako věta 2.2.

Věta 2.5. *Nechť Δ_1 a Δ_2 jsou kumulativní vývojové trojúhelníky příslušné rokům vzniku i , $0 \leq i \leq N$, a rokům vývoje j , $0 \leq j \leq N$, a necht Δ_Σ je trojúhelník vzniklý z Δ_1 a Δ_2 maticovým součtem po složkách. Pak*

$$\widehat{C}_{i,N}^{(\Sigma)} \leq \widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)} \quad (2.28)$$

v nějakém konkrétním roce vzniku i právě tehdy, když platí

$$\widehat{\Psi}^{(1)}(N-i) \leq \widehat{\Psi}^{(2)}(N-i) \quad \& \quad \frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(2)}} \geq \frac{\widehat{\Psi}^{(\Sigma)}(N-i) - \widehat{\Psi}^{(2)}(N-i)}{\widehat{\Psi}^{(1)}(N-i) - \widehat{\Psi}^{(\Sigma)}(N-i)}$$

nebo alternativně

$$\widehat{\Psi}^{(1)}(N-i) \geq \widehat{\Psi}^{(2)}(N-i) \quad \& \quad \frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(2)}} \leq \frac{\widehat{\Psi}^{(\Sigma)}(N-i) - \widehat{\Psi}^{(2)}(N-i)}{\widehat{\Psi}^{(1)}(N-i) - \widehat{\Psi}^{(\Sigma)}(N-i)}.$$

Důkaz. Na (2.28) aplikujeme (2.4) a (2.7). Dostáváme, že

$$\frac{C_{i,N-i}^{(1)}}{\widehat{\Psi}^{(\Sigma)}(N-i)} + \frac{C_{i,N-i}^{(2)}}{\widehat{\Psi}^{(\Sigma)}(N-i)} \leq \frac{C_{i,N-i}^{(1)}}{\widehat{\Psi}^{(1)}(N-i)} + \frac{C_{i,N-i}^{(2)}}{\widehat{\Psi}^{(2)}(N-i)},$$

což lze dále upravovat do tvaru

$$C_{i,N-i}^{(1)} \left(\frac{1}{\widehat{\Psi}^{(\Sigma)}(N-i)} - \frac{1}{\widehat{\Psi}^{(1)}(N-i)} \right) \leq C_{i,N-i}^{(2)} \left(\frac{1}{\widehat{\Psi}^{(2)}(N-i)} - \frac{1}{\widehat{\Psi}^{(\Sigma)}(N-i)} \right).$$

Nechť $\widehat{\Psi}^{(1)}(N-i) \leq \widehat{\Psi}^{(2)}(N-i)$. Pak dle věty 2.4 je $\widehat{\Psi}^{(1)}(N-i) \leq \widehat{\Psi}^{(\Sigma)}(N-i)$, takže závorka u $C_{i,N-i}^{(1)}$ je záporná, a pak

$$\frac{C_{i,N-i}^{(1)}}{C_{i,N-i}^{(2)}} \geq \frac{\frac{1}{\widehat{\Psi}^{(2)}(N-i)} - \frac{1}{\widehat{\Psi}^{(\Sigma)}(N-i)}}{\frac{1}{\widehat{\Psi}^{(\Sigma)}(N-i)} - \frac{1}{\widehat{\Psi}^{(1)}(N-i)}}.$$

Úpravou pravé strany, vynásobením obou stran faktorem $\frac{\widehat{\Psi}^{(2)}(N-i)}{\widehat{\Psi}^{(1)}(N-i)}$ a opětovnou aplikací (2.4) nakonec dostáváme, že

$$\frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(2)}} \geq \frac{\widehat{\Psi}^{(\Sigma)}(N-i) - \widehat{\Psi}^{(2)}(N-i)}{\widehat{\Psi}^{(1)}(N-i) - \widehat{\Psi}^{(\Sigma)}(N-i)},$$

čímž je dokázána platnost první dvojice podmínek.

Pro druhou dvojici podmínek volíme $\widehat{\Psi}^{(1)}(N-i) \geq \widehat{\Psi}^{(\Sigma)}(N-i)$ a postupujeme analogicky. □

Dále zformulujeme větu pro případ, kdy nastává v (2.28) obrácená nerovnost.

Věta 2.6. *Nechť Δ_1 a Δ_2 jsou kumulativní vývojové trojúhelníky příslušné rokům vzniku i , $0 \leq i \leq N$, a rokům vývoje j , $0 \leq j \leq N$, a necht Δ_Σ je trojúhelník vzniklý z Δ_1 a Δ_2 maticovým součtem po složkách. Pak*

$$\widehat{C}_{i,N}^{(\Sigma)} \geq \widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}$$

v nějakém konkrétním roce vzniku i právě tehdy, když platí

$$\widehat{\Psi}^{(1)}(N-i) \leq \widehat{\Psi}^{(2)}(N-i) \quad \& \quad \frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(2)}} \leq \frac{\widehat{\Psi}^{(\Sigma)}(N-i) - \widehat{\Psi}^{(2)}(N-i)}{\widehat{\Psi}^{(1)}(N-i) - \widehat{\Psi}^{(\Sigma)}(N-i)}$$

nebo alternativně

$$\widehat{\Psi}^{(1)}(N-i) \geq \widehat{\Psi}^{(2)}(N-i) \quad \& \quad \frac{\widehat{C}_{i,N}^{(1)}}{\widehat{C}_{i,N}^{(2)}} \geq \frac{\widehat{\Psi}^{(\Sigma)}(N-i) - \widehat{\Psi}^{(2)}(N-i)}{\widehat{\Psi}^{(1)}(N-i) - \widehat{\Psi}^{(\Sigma)}(N-i)}.$$

Důkaz. Jedná o obdobu důkazu věty 2.5.

Povšimněme si, že zlomek $\frac{\widehat{\Psi}^{(\Sigma)}(N-i) - \widehat{\Psi}^{(2)}(N-i)}{\widehat{\Psi}^{(1)}(N-i) - \widehat{\Psi}^{(\Sigma)}(N-i)}$ nabývá svého minima, je-li $\widehat{\Psi}^{(\Sigma)}(N-i) = \widehat{\Psi}^{(2)}(N-i)$, a svého maxima, je-li $\widehat{\Psi}^{(\Sigma)}(N-i) = \widehat{\Psi}^{(1)}(N-i)$.

Věty 2.5 a 2.6 tedy ukazují, že blíží-li se odhad rozdělení $\widehat{\Psi}^{(\Sigma)}(N-i)$ v daném konkrétním roce vzniku i odhadu rozdělení, jež ho *omezuje zdola*, zvyšuje se pravděpodobnost, že v daném roce vzniku i bude

$$Proj_{i,N}(\sum_{k=1}^2 \Delta_k) \geq \sum_{k=1}^2 Proj_{i,N}(\Delta_k).$$

Naopak blíží-li se odhad rozdělení $\widehat{\Psi}^{(\Sigma)}(N-i)$ v daném konkrétním roce vzniku i odhadu rozdělení, jež ho *omezuje shora*, zvyšuje se pravděpodobnost, že v daném roce i bude

$$Proj_{i,N}(\sum_{k=1}^2 \Delta_k) \leq \sum_{k=1}^2 Proj_{i,N}(\Delta_k).$$

3. Směrodatná chyba projekcí

V první kapitole jsme se vedle deterministické metody Chain Ladder věnovali také jejímu stochastickému popisu. Definovali jsme Mackův model (viz definice 1.1) a provedli jsme diskusi o některých jeho výhodách (relativně nízký počet parametrů) a nevýhodách (vysoká citlivost na vychýlení od střední hodnoty). V této kapitole budeme stavět na jeho další velmi podstatné výhodě. Mackův model totiž pamatuje také na druhé momenty. To nám umožňuje zkoumat střední čtvercové chyby (MSE) a směrodatné chyby (SE).

Přítom cílem této kapitoly není poskytnout čtenáři hluboké teoretické poznatky k odhadování směrodatných chyb Chain Ladder odhadů IBNR rezerv, nýbrž zavedení nástrojů potřebných k jejich výpočtu. Tyto nástroje využijeme při numerické analýze ve čtvrté kapitole.

Všechny poznatky prezentované v této kapitole jsou převzaté z článků Mack (1993) a Braun (2004). Veškerá podrobnější odvození a důkazy jsou k nalezení v těchto publikacích. My je zde uvádět nebudeme, neboť nejsou pro potřeby této kapitoly podstatné a naopak by narušovaly její plynulost.

3.1 Směrodatná chyba odhadu IBNR rezervy pro jeden trojúhelník

Uvažujme Mackův model dle definice 1.1. Definujme množinu

$$D_{i,k} := \{C_{i,j}; 0 \leq j \leq k\}.$$

Dále definujme množinu D_k jako množinu obsahující napozorované realizace náhodné veličiny $C_{i,j}$ až do roku vývoje k včetně napříč všemi roky vzniku i za platnosti $j \leq N - i$, tedy

$$D_k := \{C_{i,j}; 0 \leq i \leq N, 0 \leq j \leq k, j \leq N - i\}.$$

V modelu pak při daném značení platí, že

$$E(C_{i,k} | D_{i,k-1}) = C_{i,k-1} f_k, \quad 0 \leq i \leq N, 1 \leq k \leq N, \quad (3.1)$$

$$\text{Var}(C_{i,k} | D_{i,k-1}) = C_{i,k-1} \sigma_k^2, \quad 0 \leq i \leq N, 1 \leq k \leq N. \quad (3.2)$$

Připomeňme také, že pro každé dva různé roky vzniku i a m ($i \neq m$) platí, že

$$\{C_{i,0}, \dots, C_{i,N}\} \perp\!\!\!\perp \{C_{m,0}, \dots, C_{m,N}\}. \quad (3.3)$$

Náhodné vývojové faktory f_k odhadneme pomocí (1.1).

Věta 3.1. *Za platnosti podmínek (3.1) a (3.3) jsou odhady \hat{f}_k parametrů f_k získané pomocí (1.1), $1 \leq k \leq N$, nestranné a nekorelované.*

Důkaz. viz Mack (1993, Theorem 2, str. 215)

Nestranný odhad náhodného parametru σ_k^2 navrhuje Mack (1993) jako

$$\begin{aligned}\hat{\sigma}_k^2 &:= \frac{1}{N-k} \sum_{i=0}^{N-k} C_{i,k-1} \left(\frac{C_{i,k}}{C_{i,k-1}} - \hat{f}_k \right)^2, \quad 1 \leq k \leq N-1, \\ \hat{\sigma}_N^2 &= \min \left(\hat{\sigma}_{N-1}^4 / \hat{\sigma}_{N-2}^2, \min(\hat{\sigma}_{N-2}^2, \hat{\sigma}_{N-1}^2) \right).\end{aligned}\tag{3.4}$$

Střední čtvercová chyba odhadu kumulovaného pojistného plnění $\hat{C}_{i,N}$ příslušného roku vzniku i je pak definována jako

$$\text{MSE}(\hat{C}_{i,N}) := \text{E} \left((C_{i,N} - \hat{C}_{i,N})^2 \mid D_N \right).\tag{3.5}$$

V tomto kontextu se $\text{MSE}(\hat{C}_{i,N})$ nazývá *predikční chyba*.

Ačkoliv chceme pro daný rok vzniku i určit predikční chybu výsledného odhadu IBNR rezervy, můžeme namísto toho odhadovat právě $\text{MSE}(\hat{C}_{i,N})$, jelikož

$$\begin{aligned}\text{MSE}(\widehat{IBNR}_i) &= \text{E} \left((IBNR_{i,N} - \widehat{IBNR}_{i,N})^2 \mid D_N \right) \\ &\stackrel{(1.4)}{=} \text{E} \left((C_{i,N} - C_{i,N-i} - \hat{C}_{i,N} + C_{i,N-i})^2 \mid D_N \right) \\ &= \text{E} \left((C_{i,N} - \hat{C}_{i,N})^2 \mid D_N \right).\end{aligned}$$

Z (3.5) dále za použití $\text{E}(X - a)^2 = \text{Var}(X) + (\text{E}(X) - a)^2$ dostáváme, že

$$\text{MSE}(\hat{C}_{i,N}) = \text{Var}(C_{i,N} \mid D_{N-i}) + \left(\text{E}(\hat{C}_{i,N} \mid D_{N-i}) - \hat{C}_{i,N} \right)^2.\tag{3.6}$$

Věta 3.2. *Za platnosti podmínek (3.1), (3.2) a (3.3) lze predikční chybu (3.6) odhadnout jako*

$$\widehat{\text{MSE}}(\hat{C}_{i,N}) = \hat{C}_{i,N}^2 \sum_{k=N-i+1}^N \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{i,k-1}} + \frac{1}{\sum_{j=0}^{N-k} C_{j,k-1}} \right).\tag{3.7}$$

Důkaz. viz Mack (1993, Theorem 3, str. 217)

Jak je dále uvedeno v článku Braun (2004), (3.6) lze také aproximovat jako

$$\text{MSE}(\hat{C}_{i,N}) \approx \text{Var}(C_{i,N} \mid D_{N-i}) + \text{Var}(\hat{C}_{i,N} \mid D_{N-i}),$$

kde $\text{Var}(C_{i,N} \mid D_{N-i})$ je *náhodná chyba modelu* a $\text{Var}(\hat{C}_{i,N} \mid D_{N-i})$ je *chyba odhadu*.

Náhodnou chybu odhadneme rekurzivně jako

$$\widehat{\text{Var}}(C_{i,k} \mid D_{N-i}) = \widehat{\text{Var}}(C_{i,k-1} \mid D_{N-i}) \hat{f}_k^2 + \hat{C}_{i,k-1} \hat{\sigma}_k^2,\tag{3.8}$$

s počáteční hodnotou $\widehat{\text{Var}}(C_{i,N-i} \mid D_{N-i}) = 0$. Chybu odhadu odhadneme taktéž rekurzivně jako

$$\widehat{\text{Var}}(\hat{C}_{i,k} \mid D_{N-i}) = \widehat{\text{Var}}(\hat{C}_{i,k-1} \mid D_{N-i}) \hat{f}_k^2 + \hat{C}_{i,k-1}^2 \frac{\hat{\sigma}_k^2}{\sum_{i=0}^{N-k} C_{i,k-1}}\tag{3.9}$$

s počáteční hodnotou $\widehat{\text{Var}}(\hat{C}_{i,N-i} \mid D_{N-i}) = 0$. Získáme tak odhad predikční chyby $\text{MSE}(\hat{C}_{i,N})$ v rekurzivním tvaru

$$\widehat{\text{MSE}}(\hat{C}_{i,k}) = \widehat{\text{MSE}}(\hat{C}_{i,k-1}) \hat{f}_k^2 + \hat{C}_{i,k-1}^2 \left(\frac{\hat{\sigma}_k^2}{\hat{C}_{i,k-1}} + \frac{\hat{\sigma}_k^2}{\sum_{i=0}^{N-k} C_{i,k-1}} \right).\tag{3.10}$$

Poznámka. Metodou postupného dosazování lze ukázat, že Mackovo odvození (3.7) dává stejnou predikční chybu jako Braunův rekurzivní postup (3.10).

Ukázali jsme tedy, jak počítat odhady predikční chyby odhadu IBNR rezervy pro jednotlivé roky vzniku i v daném trojúhelníku. Nyní odhad predikční chyby spočteme pro odhad celkové rezervy.

Začneme postupem uvedeným v článku Braun (2004). Namísto $\text{MSE}(\widehat{IBNR}_N)$ můžeme nyní analogicky s případem pro jeden rok vzniku i odhadovat

$$\begin{aligned} \text{MSE}\left(\sum_{i=1}^N \widehat{C}_{i,N}\right) &:= \text{E}\left(\left(\sum_{i=1}^N (C_{i,N} - \widehat{C}_{i,N})\right)^2 \mid D_N\right) \\ &\approx \text{Var}\left(\sum_{i=1}^N C_{i,N} \mid D_N\right) + \sum_{i=1}^N \text{Var}(\widehat{C}_{i,N} \mid D_{N-i}) \\ &\quad + 2 \sum_{1 \leq i < m \leq N} \text{Cov}(\widehat{C}_{i,N}, \widehat{C}_{m,N} \mid D_{N-i}). \end{aligned} \quad (3.11)$$

Za použití $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$ dostáváme indukci, že

$$\text{Var}\left(\sum_{i=1}^N \widehat{C}_{i,N} \mid D_{N-i}\right) = \sum_{i=1}^N \text{Var}(\widehat{C}_{i,N} \mid D_{N-i}) + \sum_{1 \leq i < m \leq N} 2\text{Cov}(\widehat{C}_{i,N}, \widehat{C}_{m,N} \mid D_{N-i}),$$

a tedy

$$\text{MSE}\left(\sum_{i=1}^N \widehat{C}_{i,N}\right) \approx \text{Var}\left(\sum_{i=1}^N C_{i,N} \mid D_N\right) + \text{Var}\left(\sum_{i=1}^N \widehat{C}_{i,N} \mid D_{N-i}\right).$$

Opět jsme tedy predikční chybu rozložili na náhodnou chybu a chybu odhadu.

Náhodnou chybu odhadneme opět rekurzivně jako

$$\begin{aligned} \widehat{\text{Var}}\left(\sum_{i=N-k+1}^N C_{i,k} \mid D_N\right) &= \widehat{\text{Var}}\left(\sum_{i=N-k+2}^N C_{i,k-1} \mid D_N\right) \widehat{f}_k^2 \\ &\quad + \sum_{i=N-k+1}^N \widehat{C}_{i,k-1} \widehat{\sigma}_k^2, \end{aligned} \quad (3.12)$$

přičemž tato rekurze je smysluplná pro roky vývoje $k \geq 1$.

Abychom našli odhad chyby odhadu, musíme nejprve odhadnout kovarianční člen $\text{Cov}(\widehat{C}_{i,N}, \widehat{C}_{m,N} \mid D_{N-i})$. Braun (2004) uvádí následující rekurzivní tvar

$$\begin{aligned} \widehat{\text{Cov}}(\widehat{C}_{i,k}, \widehat{C}_{m,k} \mid D_{N-i}) &= \widehat{\text{Cov}}(\widehat{C}_{i,k-1}, \widehat{C}_{m,k-1} \mid D_{N-i}) \widehat{f}_k^2 \\ &\quad + \widehat{C}_{i,k-1} \widehat{C}_{m,k-1} \frac{\widehat{\sigma}_k^2}{\sum_{i=0}^{N-k} C_{i,k-1}} \end{aligned}$$

s počáteční hodnotou $\widehat{\text{Cov}}(\widehat{C}_{i,N-i}, \widehat{C}_{m,N-i} \mid D_{N-i}) = 0$. Chybu odhadu pak rekurzivně odhadneme jako

$$\begin{aligned} \widehat{\text{Var}}\left(\sum_{i=N-k+1}^N \widehat{C}_{i,k} \mid D_{N-i}\right) &= \widehat{\text{Var}}\left(\sum_{i=N-k+2}^N \widehat{C}_{i,k-1} \mid D_{N-i}\right) \widehat{f}_k^2 \\ &\quad + \left(\sum_{i=N-k+1}^N \widehat{C}_{i,k-1}\right)^2 \frac{\widehat{\sigma}_k^2}{\sum_{i=0}^{N-k} C_{i,k-1}}, \quad k \geq 1. \end{aligned} \quad (3.13)$$

Predikční chybu nakonec odhadneme rekurzivní formulí

$$\begin{aligned} \widehat{\text{MSE}} \left(\sum_{i=N-k+1}^N \widehat{C}_{i,k} \right) &= \widehat{\text{MSE}} \left(\sum_{i=N-k+2}^N \widehat{C}_{i,k-1} \right) \widehat{f}_k^2 \\ &+ \left(\sum_{i=N-k+1}^N \widehat{C}_{i,k-1} \right)^2 \left(\frac{\widehat{\sigma}_k^2}{\sum_{i=N-k+1}^N \widehat{C}_{i,k-1}} + \frac{\widehat{\sigma}_k^2}{\sum_{i=0}^{N-k} C_{i,k-1}} \right), \end{aligned} \quad (3.14)$$

přičemž tato rekurze je opět smysluplná pouze pro roky vývoje $k \geq 1$.

Mack (1993) uvádí pro odhad predikční chyby odhadu celkové IBNR rezervy následující větu.

Věta 3.3. *Za platnosti podmínek (3.1), (3.2) a (3.3) lze predikční chybu (3.11) odhadnout jako*

$$\widehat{\text{MSE}} \left(\sum_{i=1}^N \widehat{C}_{i,N} \right) = \sum_{i=1}^N \left[\widehat{\text{MSE}}(\widehat{C}_{i,N}) + \widehat{C}_{i,N} \left(\sum_{m=i+1}^N \widehat{C}_{m,N} \right) \sum_{k=N+1-i}^N \frac{2\widehat{\sigma}_k^2 \widehat{f}_k^2}{\sum_{l=0}^{N-k} C_{l,k}} \right].$$

Důkaz. viz Mack (1993, str. 220)

Poznámka. Mack na tuto větu navázal v jednom ze svých dalších článků odvozením rekurzivní formule (viz Mack, 1999, str. 366). Braun (2004) uvádí, že odhad predikční chyby odhadu celkové rezervy získaný pomocí formule z článku Mack (1999) dává stejný výsledek jako rekurze (3.14).

Směrodatnou predikční chybu definujeme jako odmocninu z predikční chyby. Analogicky pro náhodnou chybu a chybu odhadu.

3.2 Směrodatná chyba součtu odhadů IBNR rezerv dvou korelovaných trojúhelníků

V první podkapitole jsme zavedli nástroje pro odhadování chyb odhadnutých rezerv (Chain Ladder projekcí) v rámci jednoho kumulativního vývojového trojúhelníku. Nyní zavedeme tyto nástroje pro portfolio složené ze dvou subportfolií reprezentovaných příslušnými trojúhelníky, pro něž odhadujeme rezervy samostatně a u nichž budeme brát v úvahu jejich možnou korelovanost.

Stejně jako v podkapitole 3.1 nejprve zavedeme příslušný model, jak ho zavádí Braun (2004).

Uvažujme dva vývojové kumulativní trojúhelníky Δ_1 a Δ_2 . Definujme nyní množinu $D_{i,k}$ jako

$$\{C_{i,j}^{(1)}; 0 \leq j \leq k\} \cup \{C_{i,j}^{(2)}; 0 \leq j \leq k\}.$$

a analogicky D_k jako

$$\{C_{i,j}^{(1)}; 0 \leq i \leq N, 0 \leq j \leq k \leq N - i\} \cup \{C_{i,j}^{(2)}; 0 \leq i \leq N, 0 \leq j \leq k \leq N - i\}.$$

Alternativně množiny $D_{i,k}$ a D_k zavedené v podkapitole 3.1 přeznačíme jako $D_{i,k}^{(1)}$ a $D_k^{(1)}$, resp. $D_{i,k}^{(2)}$ a $D_k^{(2)}$.

Nechť dále platí, že

$$\begin{aligned} \mathbb{E}(C_{i,k}^{(1)} \mid D_{i,k-1}^{(1)}) &= C_{i,k-1}^{(1)} f_k, \quad 0 \leq i \leq N, \quad 1 \leq k \leq N, \\ \text{Var}(C_{i,k}^{(1)} \mid D_{i,k-1}^{(1)}) &= C_{i,k-1}^{(1)} \sigma_k^2, \quad 0 \leq i \leq N, \quad 1 \leq k \leq N \end{aligned} \quad (3.15)$$

a

$$\begin{aligned} \mathbb{E}(C_{i,k}^{(2)} \mid D_{i,k-1}^{(2)}) &= C_{i,k-1}^{(2)} g_k, \quad 0 \leq i \leq N, \quad 1 \leq k \leq N, \\ \text{Var}(C_{i,k}^{(2)} \mid D_{i,k-1}^{(2)}) &= C_{i,k-1}^{(2)} \tau_k^2, \quad 0 \leq i \leq N, \quad 1 \leq k \leq N, \end{aligned} \quad (3.16)$$

kde nyní neznámé parametry nebudeme značit $f_k^{(1)}, f_k^{(2)}, (\sigma_k^{(1)})^2, (\sigma_k^{(2)})^2$ podle značení zavedeného v kapitole 2, ale použijeme pro přehlednost značení f_k, g_k a σ_k^2, τ_k^2 , kde vždy první z dvojice připadá trojúhelníku Δ_1 a druhý trojúhelníku Δ_2 .

Neznámé parametry f_k a g_k odhadneme opět pomocí (1.1) jako

$$\hat{f}_k = \frac{\sum_{i=0}^{N-k} C_{i,k}^{(1)}}{\sum_{i=0}^{N-k} C_{i,k-1}^{(1)}}, \quad 1 \leq k \leq N, \quad \hat{g}_k = \frac{\sum_{i=0}^{N-k} C_{i,k}^{(2)}}{\sum_{i=0}^{N-k} C_{i,k-1}^{(2)}}, \quad 1 \leq k \leq N,$$

a neznámé parametry σ_k^2 a τ_k^2 pomocí (3.4) jako

$$\hat{\sigma}_k^2 := \frac{1}{N-k} \sum_{i=0}^{N-k} C_{i,k-1}^{(1)} \left(\frac{C_{i,k}^{(1)}}{C_{i,k-1}^{(1)}} - \hat{f}_k \right)^2, \quad 1 \leq k \leq N-1, \quad (3.17)$$

$$\hat{\sigma}_N^2 = \min \left(\hat{\sigma}_{N-1}^4 / \hat{\sigma}_{N-2}^2, \min(\hat{\sigma}_{N-2}^2, \hat{\sigma}_{N-1}^2) \right),$$

$$\hat{\tau}_k^2 := \frac{1}{N-k} \sum_{i=0}^{N-k} C_{i,k-1}^{(2)} \left(\frac{C_{i,k}^{(2)}}{C_{i,k-1}^{(2)}} - \hat{g}_k \right)^2, \quad 1 \leq k \leq N-1, \quad (3.18)$$

$$\hat{\tau}_N^2 = \min \left(\hat{\tau}_{N-1}^4 / \hat{\tau}_{N-2}^2, \min(\hat{\tau}_{N-2}^2, \hat{\tau}_{N-1}^2) \right).$$

Oproti modelu v podkapitole 3.1 nyní přidáme ještě následující předpoklad. Analogicky k (3.15) a (3.16) předpokládáme, že

$$\text{Cov}(C_{i,k}^{(1)}, C_{i,k}^{(2)} \mid D_{i,k-1}) = \sqrt{C_{i,k-1}^{(1)} C_{i,k-1}^{(2)}} \rho_k, \quad 1 \leq k \leq N, \quad (3.19)$$

kde $\rho_k \in \mathbb{R}$ je neznámý parametr. Pak

$$\begin{aligned} \text{Corr}(C_{i,k}^{(1)}, C_{i,k}^{(2)} \mid D_{i,k-1}) &= \frac{\text{Cov}(C_{i,k}^{(1)}, C_{i,k}^{(2)} \mid D_{i,k-1})}{\sqrt{\text{Var}(C_{i,k}^{(1)} \mid D_{i,k-1}) \cdot \text{Var}(C_{i,k}^{(2)} \mid D_{i,k-1})}} \\ &\stackrel{(3.15), (3.16)}{=} \frac{\sqrt{C_{i,k-1}^{(1)} C_{i,k-1}^{(2)}} \rho_k}{\sqrt{C_{i,k-1}^{(1)} \sigma_k^2 C_{i,k-1}^{(2)} \tau_k^2}} = \frac{\rho_k}{\sigma_k \tau_k}. \end{aligned} \quad (3.20)$$

To tedy znamená, jak uvádí Braun (2004), že předpokládáme konstantní korelaci mezi $C_{i,k}^{(1)}$ a $C_{i,k}^{(2)}$ pro fixní k . Z (3.3) dále plyne, že

$$\text{Cov}(C_{i,k}^{(1)}, C_{m,k}^{(2)} \mid D_{k-1}) = 0 \text{ pro } i \neq m.$$

Poznámka. V článku Braun (2004) je tato podmínka formulována pro tzv. *individuální vývojové faktory* $F_{i,k} = \frac{C_{i,k}^{(1)}}{C_{i,k-1}^{(1)}}$ a $G_{i,k} = \frac{C_{i,k}^{(2)}}{C_{i,k-1}^{(2)}}$. Lze snadno ukázat, že za daných podmínek

$$\text{Corr}(F_{i,k}, G_{i,k} \mid D_{i,k-1}) = \text{Corr}(C_{i,k}^{(1)}, C_{i,k}^{(2)} \mid D_{i,k-1}) = \frac{\rho_k}{\sigma_k \tau_k}.$$

Braun (2004) uvádí pro neznámý parametr ρ_k jeho nestranný odhad

$$\hat{\rho}_k = \frac{1}{N - k - 1 + \omega_k^2} \sum_{i=0}^{N-k} \sqrt{C_{i,k-1}^{(1)} C_{i,k-1}^{(2)}} \left(\frac{C_{i,k}^{(1)}}{C_{i,k-1}^{(1)}} - \hat{f}_k \right) \left(\frac{C_{i,k}^{(2)}}{C_{i,k-1}^{(2)}} - \hat{g}_k \right), \quad (3.21)$$

kde

$$\omega_k^2 := \frac{\left(\sum_{i=0}^{N-k} \sqrt{C_{i,k-1}^{(1)} C_{i,k-1}^{(2)}} \right)^2}{\sum_{i=0}^{N-k} C_{i,k-1}^{(1)} \sum_{i=0}^{N-k} C_{i,k-1}^{(2)}} \leq 1. \quad (3.22)$$

Nyní již můžeme odhadovat predikční chybu odhadu kumulovaného pojistného plnění $\hat{C}_{i,N} + \hat{D}_{i,N}$ pro daný konkrétní rok vzniku i . Definujeme ji jako

$$\text{MSE}(\hat{C}_{i,N}^{(1)} + \hat{C}_{i,N}^{(2)}) := \text{E} \left(\left(C_{i,N}^{(1)} + C_{i,N}^{(2)} - (\hat{C}_{i,N}^{(1)} + \hat{C}_{i,N}^{(2)}) \right)^2 \mid D_N \right)$$

a aproximujeme (viz Braun, 2004) opět jako součet náhodné chyby a chyby odhadu jako

$$\begin{aligned} \text{MSE}(\hat{C}_{i,N}^{(1)} + \hat{C}_{i,N}^{(2)}) &\approx \text{Var}(C_{i,N}^{(1)} + C_{i,N}^{(2)} \mid D_{N-i}) \\ &\quad + \text{Var}(\hat{C}_{i,N}^{(1)} + \hat{C}_{i,N}^{(2)} \mid D_{N-i}). \end{aligned} \quad (3.23)$$

Za použití $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$ platí pro náhodnou chybu, že

$$\begin{aligned} \text{Var}(C_{i,N}^{(1)} + C_{i,N}^{(2)} \mid D_{N-i}) &= \text{Var}(C_{i,N}^{(1)} \mid D_{N-i}) + \text{Var}(C_{i,N}^{(2)} \mid D_{N-i}) \\ &\quad + 2\text{Cov}(C_{i,N}^{(1)}, C_{i,N}^{(2)} \mid D_{N-i}), \end{aligned} \quad (3.24)$$

a analogicky pro chybu odhadu, že

$$\begin{aligned} \text{Var}(\hat{C}_{i,N}^{(1)} + \hat{C}_{i,N}^{(2)} \mid D_{N-i}) &= \text{Var}(\hat{C}_{i,N}^{(1)} \mid D_{N-i}) + \text{Var}(\hat{C}_{i,N}^{(2)} \mid D_{N-i}) \\ &\quad + 2\text{Cov}(\hat{C}_{i,N}^{(1)}, \hat{C}_{i,N}^{(2)} \mid D_{N-i}). \end{aligned} \quad (3.25)$$

Jelikož jsme již dříve získali odhady (3.8) a (3.9), stačí nám nyní odhadnout pouze kovarianční členy.

Pro odhad $\widehat{\text{Cov}}(C_{i,N}^{(1)}, C_{i,N}^{(2)} \mid D_{N-i})$ uvádí Braun (2004) rekurzi

$$\begin{aligned} \widehat{\text{Cov}}(C_{i,k}^{(1)}, C_{i,k}^{(2)} \mid D_{i,k-1}) &= \widehat{\text{Cov}}(C_{i,k-1}^{(1)}, C_{i,k-1}^{(2)} \mid D_{i,k-2}) \hat{f}_k \hat{g}_k \\ &\quad + \sqrt{\widehat{C}_{i,k-1}^{(1)} \widehat{C}_{i,k-1}^{(2)}} \hat{\rho}_k \end{aligned} \quad (3.26)$$

s počátečním stavem $\widehat{\text{Cov}}(C_{i,N-i}^{(1)}, C_{i,N-i}^{(2)} \mid D_{i,N-i}) = 0$.

Odhad kovariančního členu $\text{Cov}(\widehat{C}_{i,N}^{(1)}, \widehat{C}_{i,N}^{(2)} \mid D_{N-i})$ pak Braun (2004) uvádí v rekurzivním tvaru

$$\begin{aligned} \widehat{\text{Cov}}(\widehat{C}_{i,k}^{(1)}, \widehat{C}_{i,k}^{(2)} \mid D_{i,k-1}) &= \widehat{\text{Cov}}(\widehat{C}_{i,k-1}^{(1)}, \widehat{C}_{i,k-1}^{(2)} \mid D_{i,k-2}) \widehat{f}_k \widehat{g}_k \\ &+ \frac{\widehat{C}_{i,k-1}^{(1)} \widehat{C}_{i,k-1}^{(2)}}{\sum_{i=0}^{N-k} C_{i,k-1}^{(1)} \cdot \sum_{i=0}^{N-k} C_{i,k-1}^{(2)}} \widehat{\rho}_k \sum_{i=0}^{N-k} \sqrt{\widehat{C}_{i,k-1}^{(1)} \widehat{C}_{i,k-1}^{(2)}} \end{aligned} \quad (3.27)$$

s počátečním stavem $\widehat{\text{Cov}}(\widehat{C}_{i,N-i}^{(1)}, \widehat{C}_{i,N-i}^{(2)} \mid D_{i,N-i}) = 0$.

Na závěr kapitoly zbývá uvést, jak odhadnout predikční chybu celkového kumulovaného pojistného plnění $\sum_{i=1}^N (\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)})$. Tato predikční chyba je definována jako

$$\widehat{\text{MSE}} \left(\sum_{i=1}^N (\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}) \right) := \text{E} \left(\left(\sum_{i=1}^N (C_{i,N}^{(1)} + C_{i,N}^{(2)} - (\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)})) \right)^2 \mid D_N \right)$$

a lze jí aproximovat (viz Braun, 2004) jako

$$\begin{aligned} \widehat{\text{MSE}} \left(\sum_{i=1}^N (\widehat{C}_{i,N}^{(1)} + \widehat{C}_{i,N}^{(2)}) \right) &\approx \text{Var} \left(\sum_{i=1}^N (C_{i,N}^{(1)} + C_{i,N}^{(2)}) \mid D_N \right) \\ &+ \text{Var} \left(\sum_{i=1}^N \widehat{C}_{i,N}^{(1)} \mid D_{N-i} \right) + \text{Var} \left(\sum_{i=1}^N \widehat{C}_{i,N}^{(2)} \mid D_{N-i} \right) \\ &+ 2 \sum_{1 \leq i, m \leq N} \text{Cov}(\widehat{C}_{i,N}, \widehat{C}_{m,N} \mid D_{N-\min(i,m)}), \end{aligned} \quad (3.28)$$

kde první člen je náhodná chyba a ostatní členy součtu tvoří dohromady chybu odhadu. Na náhodnou chybu bychom opět aplikovali $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$ a jelikož jsme již dříve získali odhady (3.12) a (3.13), stačí už nyní najít pouze odhady kovariančních členů

$$\text{Cov} \left(\sum_{i=1}^N C_{i,N}^{(1)}, \sum_{i=1}^N C_{i,N}^{(2)} \mid D_{N-i} \right) \quad \text{a} \quad \sum_{1 \leq i, m \leq N} \text{Cov}(\widehat{C}_{i,N}^{(1)}, \widehat{C}_{i,N}^{(2)} \mid D_{N-\min(i,m)}).$$

Braun (2004) uvádí popořadě následující rekurze s počátkem v roce vývoje $k = 1$:

$$\begin{aligned} \widehat{\text{Cov}} \left(\sum_{i=N-k+1}^N C_{i,k}^{(1)}, \sum_{i=N-k+1}^N C_{i,k}^{(2)} \mid D_{k-1} \right) \\ = \widehat{\text{Cov}} \left(\sum_{i=N-k+2}^N \widehat{C}_{i,k-1}^{(1)}, \sum_{i=N-k+2}^N \widehat{C}_{i,k-1}^{(2)} \mid D_{k-2} \right) \widehat{f}_k \widehat{g}_k + \widehat{\rho}_k \sum_{i=N-k+1}^N \sqrt{\widehat{C}_{i,k-1}^{(1)} \widehat{C}_{i,k-1}^{(2)}} \end{aligned} \quad (3.29)$$

a

$$\begin{aligned} \sum_{i, m=N-k+1}^N \widehat{\text{Cov}}(\widehat{C}_{i,k}^{(1)}, \widehat{C}_{m,k}^{(2)} \mid D) &= \sum_{i, m=N-k+2}^N \widehat{\text{Cov}}(\widehat{C}_{i,k-1}^{(1)}, \widehat{C}_{m,k-1}^{(2)} \mid D) \widehat{f}_k \widehat{g}_k \\ &+ \frac{\sum_{i=N-k+1}^N \widehat{C}_{i,k-1}^{(1)} \sum_{m=N-k+1}^N \widehat{C}_{m,k-1}^{(2)}}{\sum_{i=0}^{N-k} C_{i,k-1}^{(1)} \cdot \sum_{m=0}^{N-k} C_{m,k-1}^{(2)}} \widehat{\rho}_k \sum_{i=0}^{N-k} \sqrt{\widehat{C}_{i,k-1}^{(1)} \widehat{C}_{i,k-1}^{(2)}}, \end{aligned} \quad (3.30)$$

kde $D = D_{N-\min(i,m)}$.

4. Numerická analýza

Na závěr této práce provedeme ukázkou aplikace teoretických poznatků nabytých v předchozích kapitolách na reálná data. K těmto datům přistupujeme prostřednictvím balíčku „ChainLadder“ (viz Gesmann a kol., 2023). Jedná se o stejná data, jež byla použita v článku Braun (2004), tedy o data prezentovaná RAA (Reinsurance Association of America) v jejich studii z roku 2001, avšak s tím rozdílem, že my použijeme pouze prvních 7×7 prvků z každého z trojúhelníků.

Poznamenejme, že k získání výsledků prezentovaných v této kapitole jsme využili mj. funkcí, které nabízí balíček „ChainLadder“ (viz Gesmann a kol., 2023). Dále při analyzování směrodatné predikční chyby se inspirováme analýzou uvedenou v článku Braun (2004).

Uvažujme tedy hypotetickou situaci, kdy nspecifikovaná zajišťovna vytváří portfolio obsahující pojistný kmen týkající se tzv. *General Liability* zajištění a pojistný kmen týkající se tzv. *Auto Liability* zajištění. Cílem této zajišťovny je co nejpřesněji odhadnout rezervu potřebnou na pokrytí nákladů na pojistná plnění v tomto portfolio. Portfolio budeme značit $GL + AL$.

Portfolio $GL + AL$ je reprezentováno následujícím kumulativním vývojovým trojúhelníkem:

i	0	1	2	j 3	4	5	6
0	174389	411113	567494	693864	804744	846880	894651
1	201981	458519	649549	801648	899099	968981	
2	196239	477292	732813	889070	1023264		
3	230841	580819	794412	981024			
4	269254	619830	882632				
5	261351	606710					
6	271372						

Tabulka 4.1: Portfolio $GL + AL$

Poznámka. Pomocí i značíme rok vzniku, pomocí j značíme rok vývoje.

Pokusíme se určit, zda-li a jak se změní výsledný odhad celkové rezervy, rozhodne-li se zajišťovna portfolio $GL + AL$ rozdělit na dvě subportfolia a odhadovat rezervu pro každé ze subportfolií zvlášť. První subportfolio (označme GL) bude obsahovat právě pojistný kmen příslušný *General Liability* zajištění. Druhé subportfolio (označme AL) bude pak obsahovat pojistný kmen příslušný *Auto Liability* zajištění.

Subportfolio GL představuje kumulativní vývojový trojúhelník:

i	0	1	2	j 3	4	5	6
0	59966	163152	254512	349524	433265	475778	513660
1	49685	153344	272936	383349	458791	503358	
2	51914	170048	319204	425029	503999		
3	84937	273183	407318	547288			
4	98921	278329	448530				
5	71708	245587					
6	92350						

Tabulka 4.2: Subportfolio GL

Subportfolio AL představuje kumulativní vývojový trojúhelník:

i	0	1	2	j 3	4	5	6
0	114423	247961	312982	344340	371479	371102	380991
1	152296	305175	376613	418299	440308	465623	
2	144325	307244	413609	464041	519265		
3	145904	307636	387094	433736			
4	170333	341501	434102				
5	189643	361123					
6	179022						

Tabulka 4.3: Subportfolio AL

Poznámka. Trojúhelník 4.1 vznikl z 4.2 a 4.3 maticovým součtem po složkách.

j	0	1	2	3	4	5	6
$\hat{f}_j^{(GL)}$		3.08	1.64	1.36	1.21	1.10	1.08
$\hat{\Psi}^{(GL)}(j)$	0.11	0.31	0.51	0.70	0.84	0.93	
$\hat{f}_j^{(AL)}$		2.04	1.27	1.11	1.09	1.03	1.03
$\hat{\Psi}^{(AL)}(j)$	0.30	0.61	0.78	0.87	0.95	0.97	
$\hat{f}_j^{(GL+AL)}$		2.36	1.42	1.23	1.14	1.07	1.06
$\hat{\Psi}^{(GL+AL)}(j)$	0.19	0.44	0.63	0.77	0.89	0.95	

Tabulka 4.4: Odhady vývojových charakteristik

Tabulka 4.4 uvádí odhady vývojových faktorů získané pomocí (1.1) a odhady rozdělení kumulativních pojistných plnění napříč roky vývoje získané pomocí (2.3).

Nyní aplikujeme na trojúhelníky Δ_{GL} , Δ_{AL} a Δ_{GL+AL} metodu Chain Ladder a získáme tak příslušné projekce. V tabulce 4.5 níže uvádíme hodnoty projekcí v posledním sloupci každého z trojúhelníků. Pomocí těchto projekcí následně napočítáme míry růstu v daných trojúhelnících (viz tabulka 4.6).

i	$\widehat{C}_{i,6}^{(GL)}$	$\widehat{C}_{i,6}^{(AL)}$	$\widehat{C}_{i,6}^{(GL+AL)}$
0	513660	380991	894651
1	543436	478031	1023640
2	597244	549479	1152053
3	781932	498025	1263148
4	871425	555343	1393770
5	782547	588955	1363960
6	905553	595649	1442489
Celkem	4995797	3646473	8533711

Tabulka 4.5: Chain Ladder projekce v posledním sloupci

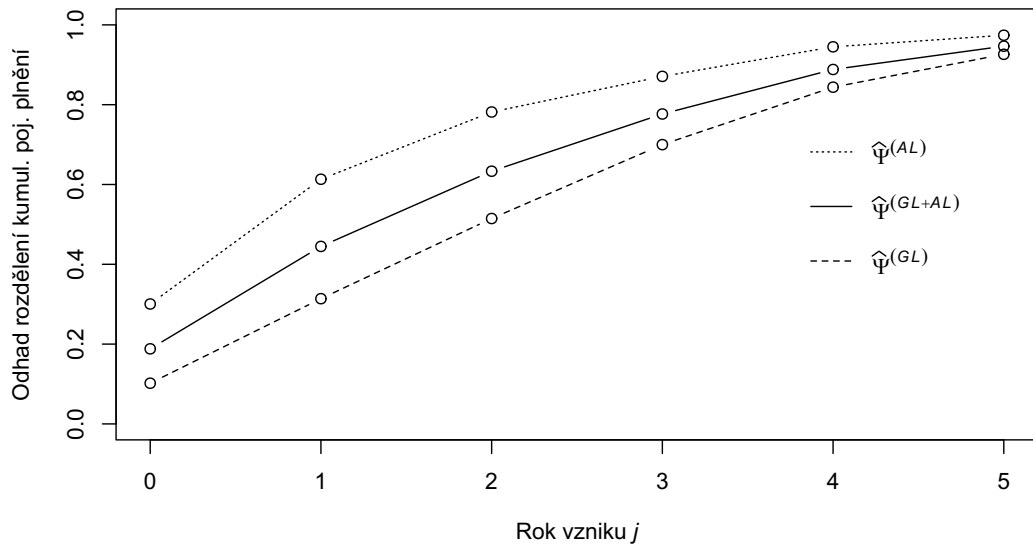
i	1	2	3	4	5	6
$\widehat{C}_{i,6}^{(GL)} / \sum_{k=0}^{i-1} \widehat{C}_{k,6}^{(GL)}$	1.06	0.56	0.47	0.36	0.24	0.22
$\widehat{C}_{i,6}^{(AL)} / \sum_{k=0}^{i-1} \widehat{C}_{k,6}^{(AL)}$	1.25	0.64	0.35	0.29	0.24	0.20
$\widehat{C}_{i,6}^{(GL+AL)} / \sum_{k=0}^{i-1} \widehat{C}_{k,6}^{(GL+AL)}$	1.14	0.60	0.41	0.32	0.24	0.20

Tabulka 4.6: Míry růstu

Z výsledků uvedených v tabulkách 4.4 a 4.6 je patrné, že projekce trojúhelníků Δ_{GL} a Δ_{AL} jistě nebudou aditivní, neboť pro většinu roků vývoje i není splněna ani jedna z podmínek uvedených ve větě 2.1. Naším cílem bude tedy určit, jakým směrem jsou orientovány projekční nerovnosti pro dané roky vzniku i , a rozhodnout, zda bude nakonec součet odhadů rezerv získaných z trojúhelníků Δ_{GL} a Δ_{AL} větší než odhad rezervy získané z trojúhelníku Δ_{GL+AL} či naopak. Pokusíme se zjistit orientace těchto nerovností pomocí aplikace poznatků uvedených v kapitolách 2 a 3. Na závěr ověříme správnost výsledků pomocí hodnot z tabulky 4.5.

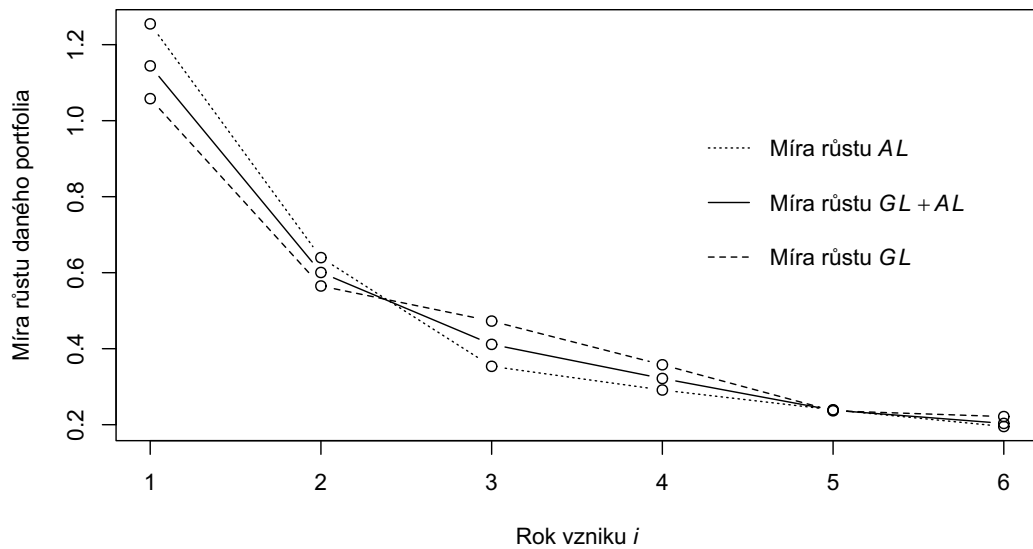
Zaměříme se nejprve na to, zda jsou splněny podmínky věty 2.3.

Hodnoty z tabulek 4.4 a 4.6 zaneseme pro přehlednost do grafů. Nejdříve vykreslíme příslušné odhady rozdělení kumulovaných pojistných plnění napříč roky vývoje j pro každý z trojúhelníků (viz tabulka 4.4).



Obrázek 4.1: Odhady rozdělení kumulovaných pojistných plnění

Poznámka. Obrázek 4.1 názorně ukazuje platnost věty 2.4. Nyní vykreslíme příslušné míry růstu (viz tabulka 4.6).



Obrázek 4.2: Míry růstu

Poznámka. Z obrázku 4.2 je patrné, že by bylo možné zformulovat analogii věty 2.4 založenou na mírách růstu.

Z obrázku 4.1 vidíme, že podmínka (2.17) je skutečně splněna pro každý rok vzniku i , $0 \leq i \leq 6$. V podmínce (2.18) se ale dle obrázku 4.2 orientace nerovnosti pro jednotlivé roky vzniku i mění. Na základě věty 2.3 tedy nelze o orientaci projekční nerovnosti rozhodnout. Povšimněme si ale, že míra růstu pro rok vzniku i stále závisí na $\widehat{C}_{k,N}$ pro $\forall k \leq i - 1$, a proto bychom mohli chtít připisovat větší váhu chvostu tohoto grafu, kde je míra růstu subportfolia GL větší než míra růstu subportfolia AL .

Nyní provedeme rozbor pro jednotlivé roky vzniku podle vět 2.5 a 2.6.

j	0	1	2	3	4	5
$\frac{\widehat{\Psi}^{(AL)}(j) - \widehat{\Psi}^{(GL+AL)}(j)}{\widehat{\Psi}^{(AL)}(j) - \widehat{\Psi}^{(GL)}(j)}$	0.57	0.56	0.56	0.55	0.56	0.56

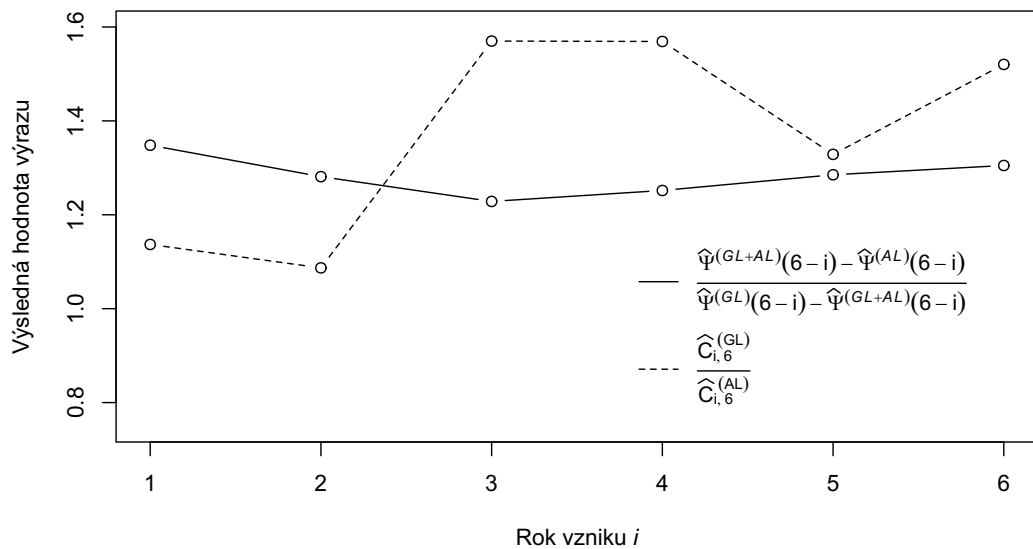
Tabulka 4.7: Potenciální poloha $\widehat{\Psi}^{(GL+AL)}$

Dle tabulky 4.7 je odhad rozdělení $\widehat{\Psi}^{(GL+AL)}(N - i)$ v každém roce vzniku i mírně blíže odhadu rozdělení, který jej omezuje zdola. V tuto chvíli ovšem nelze vyvozovat žádné zásadní závěry, neboť toto přiblížení je velmi nevýrazné.

Jelikož $\widehat{\Psi}^{(GL)}(N - i) \leq \widehat{\Psi}^{(AL)}(N - i)$ pro každý rok vzniku i (viz obrázek 4.1), stačí nám dle vět 2.5 a 2.6 porovnávat

$$\frac{\widehat{C}_{i,6}^{(GL)}}{\widehat{C}_{i,6}^{(AL)}} \quad \text{a} \quad \frac{\widehat{\Psi}^{(GL+AL)}(6 - i) - \widehat{\Psi}^{(AL)}(6 - i)}{\widehat{\Psi}^{(GL)}(6 - i) - \widehat{\Psi}^{(GL+AL)}(6 - i)}.$$

Tyto dva výrazy porovnáme opět graficky.



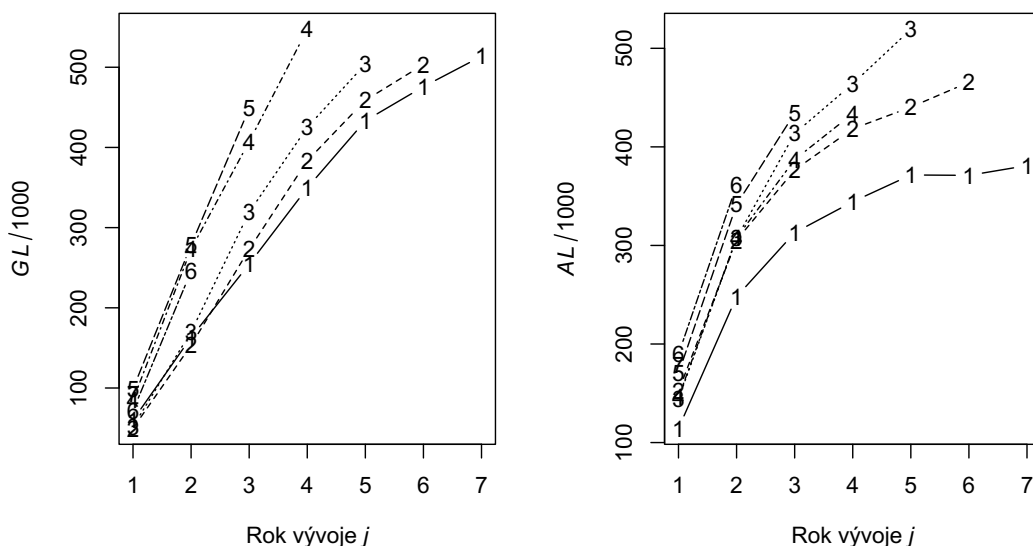
Obrázek 4.3: Porovnání výrazů vystupujících ve větách 2.5 a 2.6

Na základě obrázku 4.3 dle vět 2.5 a 2.6 dostáváme, že v roce vzniku $i = 1$ a v roce vzniku $i = 2$ platí, že

$$Proj_{i,6}(\Delta_{GL+AL}) \geq Proj_{i,6}(\Delta_{GL}) + Proj_{i,6}(\Delta_{AL}),$$

zatímco pro roky vzniku $i = 3, 4, 5$ a 6 nastává obrácená nerovnost.

Jako další provedeme následující rozbor.



Obrázek 4.4: Vývoj kumulovaných pojistných plnění v trojúhelnících Δ_{GL} a Δ_{AL} samostatně pro každý rok vzniku i

Poznámka. Hodnoty na ose y představují výši kumulovaného pojistného plnění v daném trojúhelníku (portfoliu) pro přehlednost přeškálovanou vydělením číslem tisíc. Každá z křivek přísluší podle příslušného očíslování jednomu konkrétnímu roku vzniku i , $1 \leq i \leq 6$.

Z obrázku 4.4 vidíme, že jak v trojúhelníku Δ_{GL} , tak v trojúhelníku Δ_{AL} vycházejí v každém roce vývoje j pro pozdější roky vzniku i vyšší hodnoty kumulovaného pojistného plnění. Z definice trojúhelníku Δ_{GL+AL} je patrné, že pro tento trojúhelník očekáváme obdobný vývoj. Toto zjištění by mohlo společně s nižším počtem napozorovaných dat poukazovat pro pozdější roky vzniku na vyšší variabilitu ve smyslu *predikčních chyb* (viz kapitola 3) a dalo by nám tak další argument pro to, abychom těmto rokům vzniku přikládali větší váhu při zkoumání orientace výsledné nerovnosti.

Abychom ověřili, zda tomu tak skutečně je, budeme nyní zkoumat směrodatné chyby v daných trojúhelnících. Nejprve pomocí (3.4), (3.21) a (3.22) získáme příslušné odhadnuté parametry $\hat{\sigma}_k^2$ a $\hat{\rho}_k$ a parametr ω_k^2 (viz tabulka 4.8).

j	1	2	3	4	5	6
$\hat{\sigma}_k^2 (GL)$	5369.21	4974.73	318.09	300.03	0.215	0.00015
$\hat{\sigma}_k^2 (AL)$	1515.09	549.84	35.37	494.25	689.75	494.25
$\hat{\sigma}_k^2 (GL+AL)$	1909.68	2140.83	77.19	309.63	273.21	241.08
ω_k^2	0.986	0.988	0.995	0.999	0.999	1.000
$\hat{\rho}_k$	-1102.92	898.88	-74.66	-128.92	-12.18	0.00
$\frac{\hat{\rho}_k}{\hat{\sigma}_k^{(GL)}\hat{\sigma}_k^{(AL)}}$	-0.39	0.54	-0.70	-0.33	-1.00	0.00

Tabulka 4.8: Odhady příslušných parametrů

Nyní spočteme směrodatné náhodné chyby a směrodatné chyby odhadu modelu, a z nich následně směrodatné predikční chyby. Spočtené hodnoty uspořádáme do tabulek, jejichž styl jsme převzali z článku Braun (2004).

i	$GL + AL$	GL	AL	Portfolio Λ	Implikovaná korelace v Λ	Volba kor. = -1
0	0	0	0	0		0
1	15284	9	15170	15170	0.000	15161
2	23977	356	25338	25067	-0.767	24983
3	31865	15190	38670	29667	-0.198	13480
4	35119	23418	30607	34420	-0.210	7190
5	66677	71443	36311	86371	0.200	35132
6	85532	104599	45331	112446	-0.037	59268
Celková	121753	129708	77487	151748	0.010	52221

Tabulka 4.9: Směrodatné náhodné chyby

i	$GL + AL$	GL	AL	Portfolio Λ	Implikovaná korelace v Λ	Volba kor. = -1
0	0	0	0	0		0
1	16349	9	16993	16993	0.000	16984
2	22933	267	24960	24794	-0.622	24692
3	28118	10445	24427	25130	-0.145	13981
4	31606	15482	27346	29333	-0.150	11864
5	41569	35837	30312	48957	0.089	5526
6	49654	53239	32873	61947	-0.022	20366
Celková	167795	95308	148039	173857	-0.027	52731

Tabulka 4.10: Směrodatné chyby odhadu modelu

i	$GL + AL$	GL	AL	Portfolio Λ	Implikovaná korelace v Λ	Volba kor. = -1
0	0	0	0	0		0
1	22380	13	22779	22779	0.000	22766
2	33179	445	35567	35258	-0.699	35122
3	42497	18435	37665	38880	-0.178	19229
4	47247	28073	41044	45224	-0.186	12971
5	78573	79927	47300	99281	0.163	32628
6	98926	117368	55996	128381	-0.033	61372
Celková	207314	160959	167093	230768	-0.011	6133

Tabulka 4.11: Směrodatné predikční chyby

Hodnoty v prvním sloupci tabulky vždy přísluší portfoliu $GL + AL$, hodnoty ve druhém a třetím sloupci pak subportfoliím GL a AL . Nástroje pro výpočet těchto směrodatných chyb jsme uvedli v podkapitole 3.1 (výpočty byly provedeny pomocí vzorců (3.8), (3.9), (3.10), (3.12), (3.13) a (3.14)).

Čtvrtý sloupec se vždy týká portfolia obsahujícího obě subportfolia GL a AL , pro něž jsou ale projekce projektovány samostatně a až poté sčítány. Toto portfolio označíme v rámci této kapitoly jako Λ . Nástroje pro výpočet směrodatných chyb v takovém portfoliu jsme uvedli v podkapitole 3.2 (výpočty byly provedeny pomocí vzorců (3.23), (3.24), (3.25), (3.26), (3.27), (3.28), (3.29) a (3.30)).

Předposlední sloupec uvádí vždy tzv. *implikovanou korelaci* v portfoliu Λ . Jak uvádí Braun (2004), jedná se o řešení $\rho(X,Y)$ rovnice

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\rho(X,Y)\sqrt{\text{Var}(X)\text{Var}(Y)}, \quad (4.1)$$

kde X a Y jsou příslušné rezervy v subportfoliích GL a AL a $\text{Var}(X)$, $\text{Var}(Y)$ a $\text{Var}(X + Y)$ jsou druhé mocniny příslušných směrodatných chyb z prvních tří sloupců. V posledním sloupci naopak vždy uvádíme směrodatné chyby příslušné portfoliu Λ získané z (4.1) při volbě $\rho(X,Y) = -1$.

Směrodatné chyby jsou pro pozdější roky vzniku skutečně výrazně vyšší, což potvrzuje naši úvahu o vyšší variabilitě. Zároveň z tabulek 4.9, 4.10 a 4.11 vidíme, že směrodatné chyby příslušné rokům vzniku $i = 1$ a $i = 2$ v portfoliu $GL + AL$ jsou dokonce nižší (až na jednu výjimku v tabulce 4.9) než směrodatné chyby příslušné stejným rokům vzniku v těchto tabulkách v portfoliu Λ při volbě korelačního koeficientu $\rho(X,Y) = -1$. Podle analogické úvahy jako v článku Braun (2004, str. 416) tato skutečnost naznačuje, že portfolio $GL + AL$ je opravdu vhodné rozdělit na dvě subportfolia GL a AL a odhadovat pro tato subportfolia rezervu samostatně.

Na základě výsledků numerické analýzy tedy usuzujeme, že je vhodné pozdějším rokům vzniku při zkoumání orientace nerovnosti pro celkové odhady IBNR

rezervy připisovat vyšší váhu. Dle vět 2.3, 2.5 a 2.6 je pak, za využití výše zavedeného značení, výsledný odhad IBNR rezervy v portfoliu Λ větší než výsledný odhad IBNR rezervy v portfoliu $GL + AL$. Tento závěr ověříme pomocí projekcí z tabulky 4.5 porovnáním součtu celkových součtů prvního a druhého sloupce ($4995797 + 3646473 = 8642270$) s celkovým součtem v posledním sloupci (8533711). Vidíme, že orientaci této nerovnosti jsme tedy určili správně.

Výsledný rozdíl těchto dvou součtů je navíc markantních 108559, což ve spojení s dalšími argumenty uvedenými v této kapitole naznačuje, že pokud bychom portfolio $GL + AL$ nerozdělili na příslušná dvě subportfolia, mohlo by dojít k významnému podhodnocení výsledné odhadnuté IBNR rezervy.

Závěr

V práci jsme se zabývali problematikou aditivity projekcí získaných pomocí metody Chain Ladder v příslušných kumulativních vývojových trojúhelnících. Uvedli jsme, jaký je význam odhadování IBNR rezerv v neživotním pojištění a jak přispívá metoda Chain Ladder k získávání těchto odhadů. Představili jsme deterministický popis této metody i stochastické modely na ní založené. Obecně jsme definovali pojmy *projekce* a *aditivita projekcí*. Věnovali jsme se nejprve nutným a postačujícím podmínkám pro aditivitu projekcí dvou trojúhelníků a dalším specifikům z této aditivity vyplývajícím. Následně jsme uvedli postačující podmínky pro projekční nerovnost a tuto problematiku jsme následně dále rozšířili. Hlavním cílem práce bylo diskutovat, zda by nebylo vhodné projektovat rezervy pro určitou část portfolia samostatně, chová-li se tato část portfolia (subportfolio) v jistém smyslu odlišně. Pro tyto účely jsme představili nástroje k výpočtu směrodatných chyb pro projekce rezerv v portfoliích reprezentovaných danými trojúhelníky.

Nabyté poznatky jsme nakonec aplikovali na reálná data při numerické analýze hypotetické situace týkající se nespécifikované zajišťovny. Ukázalo se, že pro tato data bylo vhodné portfolio rozdělit na dvě subportfolia a odhadovat pro každé z nich rezervy samostatně. V opačném případě by totiž, jak ukázala analýza pomocí směrodatných chyb, mohlo dojít k významnému podhodnocení výsledné odhadnuté IBNR rezervy.

Příslušný kód, jenž jsme v programu R vytvořili za účelem získání hodnot prezentovaných během numerické analýzy, je uveden v příloze práce.

Seznam použité literatury

- AJNE, B. (1994). Additivity of chain-ladder projections. *ASTIN Bulletin*, **24**, 311 – 318.
- BRAUN, C. (2004). The prediction error of the chain ladder method applied to correlated run-off triangles. *ASTIN Bulletin*, **34**, 399 – 423.
- CIPRA, T. (1999). *Pojistná matematika : teorie a praxe*. Ekopress. ISBN 80-86119-17-3.
- CIPRA, T. (2015). *Praktický průvodce finanční a pojistnou matematikou*. Ekopress. ISBN 978-80-87865-18-7.
- GESMANN, M., MURPHY, D., ZHANG, Y. W., CARRATO, A., WUTHRICH, M., CONCINA, F. a DAL MORO, E. (2023). *ChainLadder: Statistical Methods and Models for Claims Reserving in General Insurance*. URL <https://CRAN.R-project.org/package=ChainLadder>. R package version 0.2.17.
- MACK, T. C. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, **23**, 213 – 225.
- MACK, T. C. (1994). Which stochastic model is underlying the chain ladder method. *Insurance Mathematics & Economics*, **15**, 133–138.
- MACK, T. C. (1999). The standard error of chain ladder reserve estimates: Recursive calculation and inclusion of a tail factor. *ASTIN Bulletin*, **29**, 361 – 366.
- WÜTHRICH, M. V. a MERZ, M. (2008). *Stochastic Claims Reserving Methods in Insurance*. John Wiley & Sons Ltd. ISBN 978-0-470-72346-3.

Seznam obrázků

4.1	Odhady rozdělení kumulovaných pojistných plnění	28
4.2	Míry růstu	28
4.3	Porovnání výrazů vystupujících ve větách 2.5 a 2.6	29
4.4	Vývoj kumulovaných pojistných plnění v trojúhelnících Δ_{GL} a Δ_{AL} samostatně pro každý rok vzniku i	30

Seznam tabulek

1	Kumulativní vývojový trojúhelník	2
2	Doplněný kumulativní vývojový trojúhelník	3
4.1	Portfolio $GL + AL$	25
4.2	Subportfolio GL	26
4.3	Subportfolio AL	26
4.4	Odhady vývojových charakteristik	26
4.5	Chain Ladder projekce v posledním sloupci	27
4.6	Míry růstu	27
4.7	Potenciální poloha $\hat{\Psi}^{(GL+AL)}$	29
4.8	Odhady příslušných parametrů	31
4.9	Směrodatné náhodné chyby	31
4.10	Směrodatné chyby odhadu modelu	31
4.11	Směrodatné predikční chyby	32

A. Přílohy

A.1 První příloha

Elektronická příloha obsahuje RMD soubor s kódem, jenž byl v programu R vytvořen za účelem získání hodnot prezentovaných během numerické analýzy v kapitole 4.