

UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA

DIPLOMOVÁ PRÁCE

Využití aparátu teorie fuzzy množin v
pojišťovnictví



Katedra matematické analýzy a aplikací matematiky
Vedoucí diplomové práce: **RNDr. Ondřej Pavlačka, Ph.D.**
Vypracovala: **Bc. Ivana Pavlů**
Studijní program: N1103 Aplikovaná matematika
Studijní obor Aplikace matematiky v ekonomii
Forma studia: prezenční
Rok odevzdání: 2020

BIBLIOGRAFICKÁ IDENTIFIKACE

Autor: Bc. Ivana Pavlů

Název práce: Využití aparátu teorie fuzzy množin v pojišťovnictví

Typ práce: Diplomová práce

Pracoviště: Katedra matematické analýzy a aplikací matematiky

Vedoucí práce: RNDr. Ondřej Pavlačka, Ph.D.

Rok obhajoby práce: 2020

Abstrakt: Tématem této diplomové práce je využití teorie fuzzy množin v pojišťovnictví. Úvodní kapitola stručně shrnuje terminologii teorie fuzzy množin, druhá kapitola pak obsahuje základní pojmy z oblasti pojišťovnictví. Ve třetí kapitole jsou na základě průzkumu literatury stručně představeny některé dosavadní aplikace aparátu teorie fuzzy množin v pojišťovnictví. Závěrečná čtvrtá kapitola se pak blíže věnuje výpočtu rezervy na pojistná plnění v neživotním pojištění metodou Chain-Ladder a různým možnostem fuzzifikace této metody. Dále jsou zde příslušné metody aplikovány na reálná data a porovnány skrze výsledné odhady rezervy.

Klíčová slova: fuzzy množina, fuzzy číslo, fuzzy aritmetika, pojišťovnictví, technické rezervy v pojišťovnictví, metoda Chain-Ladder, metoda fuzzy Chain-Ladder

Počet stran: 72

Počet příloh: 1

Jazyk: český

BIBLIOGRAPHICAL IDENTIFICATION

Author: Bc. Ivana Pavlů

Title: Application of fuzzy set theory in insurance

Type of thesis: Master's

Department: Department of Mathematical Analysis and Application of Mathematics

Supervisor: RNDr. Ondřej Pavlačka, Ph.D.

The year of presentation: 2020

Abstract: The topic of this master thesis is fuzzy sets theory and its possible use in insurance industry. The opening chapter presents a brief introduction of fuzzy sets theory, the second chapter provides an overview of insurance industry and its terminology. In the third chapter, some already existing use of fuzzy sets theory in insurance industry is described. The final chapter presents calculating of the claims reserve in non-life insurance by the Chain-Ladder method and its possible ways of fuzzification, respectively. Different methods are used on adopted data and compared through their reserve estimates.

Key words: fuzzy set, fuzzy number, fuzzy arithmetic, insurance industry, technical reserves in insurance, Chain-Ladder method, fuzzy Chain-Ladder method

Number of pages: 72

Number of appendices: 1

Language: Czech

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracovala samostatně pod vedením pana RNDr. Ondřeje Pavlačky, Ph.D. a všechny použité zdroje jsem uvedla v seznamu literatury.

V Olomouci dne

.....

podpis

Obsah

Úvod	8
1 Základy teorie fuzzy množin	9
1.1 Motivace k zavedení fuzzy množin	9
1.2 Základní pojmy z teorie fuzzy množin	11
1.3 Fuzzy čísla	12
1.3.1 Operace s fuzzy čísly	17
1.3.2 Defuzifikace fuzzy čísel	20
2 Pojišťovnictví	23
2.1 Pojištění	23
2.2 Technické rezervy	25
3 Využití aparátu teorie fuzzy množin v pojišťovnictví	27
3.1 Aplikace TFM v neživotním pojištění	27
3.1.1 Detekce pojistných podvodů v havarijním pojištění	27
3.1.2 Fuzzifikace jazykových termů v pojištění proti zemětřesení	29
3.1.3 Fuzzy regresní analýza	30
3.2 Aplikace TFM v životním pojištění	31
3.2.1 Fuzzy cashflow	31
3.2.2 Volba životního pojištění	32
4 Fuzzifikace metod využívajících trojúhelníková schémata	35
4.1 Obecný úvod k trojúhelníkovým schématům	35
4.2 Metoda Chain-Ladder	38
4.3 Fuzzifikace metody Chain-Ladder	43
4.3.1 Fuzzy Chain-Ladder	43
4.3.2 Fuzzy Chain-Ladder s využitím minimálních a maximálních podílů $C_{i,j}$	50
4.3.3 Fuzzy Chain-Ladder s pevně nastaveným rozpětím fuzzy koeficientů	54
4.3.4 Fuzzy Chain-Ladder využívající průměr podílů kumulativních pojistných plnění	55

4.3.5	Fuzzy Chain-Ladder s využitím regresní analýzy	56
4.4	Aplikace variant metody fuzzy Chain-Ladder	59
4.4.1	Fuzzy Chain-Ladder	60
4.4.2	Fuzzy Chain-Ladder - minimální a maximální podíly	61
4.4.3	Fuzzy Chain-Ladder - průměrné podíly	63
4.4.4	Fuzzy Chain-Ladder - regresní analýza	64
4.4.5	Shrnutí výsledků	65
	Závěr	68
	Literatura	70

Poděkování

Ráda bych na tomto místě poděkovala vedoucímu své práce, RNDr. Ondřeji Pavlačkovi, Ph.D., za spolupráci, obětovaný čas a kladení těch správných otázek na správných místech. Dále děkuji své rodině, přátelům, spolužákům a všem, kteří mi byli oporou i v hektických chvílích studia.

Úvod

Podnětem pro tuto práci byla snaha o propojení dvou problematik, kterým se studenti magisterského studijního oboru Aplikace matematiky v ekonomii během studia věnují, a to teorie fuzzy množin a pojistné matematiky. Teorie fuzzy množin patří mezi relativně mladé matematické disciplíny - její počátky se datují do šedesátých let minulého století, přesto její aparát již našel využití v mnoha aspektech lidské činnosti. Pojišťovnictví oproti tomu patří mezi klíčové složky finančních sektorů všech ekonomicky vyspělých zemí.

Vzhledem k zaměření své bakalářské práce (téma *Metody stanovení rezervy na pojistná plnění v neživotním pojištění*) bylo přirozeným nápadem zkusit aplikovat poznatky z teorie fuzzy množin právě v oblasti technických rezerv v neživotním pojištění - této oblasti zatím nebylo věnováno mnoho pozornosti (za zmínku stojí články [11, 14]). Po dohodě s vedoucím práce jsme jako hlavní náplň diplomové práce zvolili možnosti fuzzifikace nejpoužívanější metody pro stanovení rezervy na pojistná plnění v neživotním pojištění, která je založená na tzv. trojúhelníkových schématech, a sice metody Chain-Ladder.

Pro uvedení do problematiky bude práce zahrnovat kapitoly, které jak teorii fuzzy množin, tak pojišťovnictví blíže představí, a to včetně používané terminologie. Zároveň bude proveden průzkum literatury zabývající se dosavadním využitím aparátu teorie fuzzy množin v pojišťovnictví, jehož výstup zde bude obsažen. Relativně samostatnou částí práce pak bude kapitola věnující se metodě Chain-Ladder. Zde se podrobněji rozebere její fuzzifikovaná varianta, kterou jedinou jsem v literatuře objevila. Cílem práce by zde mělo být navržení vlastních modifikací této metody, které by eliminovaly některé její negativní stránky.

Kapitola 1

Základy teorie fuzzy množin

S ideou fuzzy množin jako první přišel Lotfi Zadeh, původem ázerbajdžánský matematik, který během své profesní kariéry působil na několika předních amerických univerzitách. Základní kameny celé teorie fuzzy množin položil ve svém článku *Fuzzy sets* [30] z roku 1965. Této problematice se nadále věnoval ve svých dalších rozšiřujících publikacích, stejně tak jako další matematikové. Odvětví tak bylo obohaceno o fuzzy logiku, fuzzy rozhodovací systémy, fuzzy algoritmy, fuzzy regulátory a kontrolní systémy aj.

Prostudování této kapitoly by mělo čtenáře vybavit dostatečnými znalostmi z teorie fuzzy množin pro potřeby této práce. V úvodu kapitoly budou představeny některé základní aspekty problematiky a provede se srovnání fuzzy množin s jejich „klasickým“ protějškem. Dále bude přiblížena základní terminologie, při které bude kladen důraz na speciální třídu fuzzy množin, fuzzy čísla, včetně některých jejich důležitých struktur. Zároveň se čtenář dozví, jak lze snadno zobecnit základní algebraické operace z klasické aritmetiky právě pro fuzzy čísla. Někdy je vhodné fuzzy číslo interpretovat pomocí čísla reálného - budou zde proto uvedeny i některé ze způsobů defuzzifikace fuzzy čísel.

1.1. Motivace k zavedení fuzzy množin

V základních kurzech matematiky či matematické analýzy se seznamujeme s pojmem *množina*. Množinou M rozumíme soubor prvků - lze ji zadefinovat

výčtem, příp. nějakou charakteristickou vlastností, přičemž o každém objektu lze rozhodnout, zda do dané množiny M patří nebo ne. Podle počtu prvků může být množina konečná, konečná spočetná i nekonečná. V rámci této práce budeme tuto množinu označovat jako klasickou, příp. ostrou. Klasické množiny se dají popsat i pomocí charakteristické funkce uvedené v následující definici.

Definice 1.1 (Charakteristická funkce). Nechť U je neprázdná množina, tzv. universum. *Charakteristická funkce množiny M* je zobrazení $\chi_M : U \rightarrow \{0, 1\}$ definované vztahem

$$\chi_M(x) = \begin{cases} 1, & \text{jestliže } x \in M, \\ 0, & \text{jestliže } x \notin M. \end{cases}$$

Laicky lze definici 1.1 popsat následovně: těm prvkům universa U , které patří do množiny M , přiřadíme logickou hodnotu 1, zatímco prvkům, které do M nepatří, hodnotu 0. V kontextu další části lze tuto skutečnost chápat tak, že prvek do množiny M patří buď zcela (1), nebo vůbec (0). Příkladem klasických množin jsou např. množina přirozených čísel \mathbb{N} nebo množina všech studentů ve třídě.

Uvažujme nyní některé termíny používané v běžné mluvě - mladý člověk, vysoká budova, dlouhá kniha. Cítíme, že v tomto případě není tak jednoduché určit, zda např. 30letého člověka jednoznačně přiřadit do ostré množiny mladých lidí.¹ V tuto chvíli lze vzít na pomoc právě *fuzzy množiny*, které nabízejí způsob, jak popsat pouze částečnou příslušnost prvku do množiny. Analogií charakteristické funkce v případě klasické množiny je pro fuzzy množinu A tzv. *funkce příslušnosti* μ_A detailněji přiblížená v samotné definici fuzzy množiny.

Definice 1.2 (Fuzzy množina). Nechť U je neprázdná množina, tzv. universum. Pak *fuzzy množina A* na U je určena zobrazením

$$\mu_A : U \rightarrow \langle 0, 1 \rangle.$$

Funkci μ_A nazýváme *funkcí příslušnosti fuzzy množiny A* . Pro každé $x \in U$ nazýváme hodnotu $\mu_A(x)$ *stupněm příslušnosti prvku x k A* .

¹Právě vágní jazykové pojmy jsou jedním z typických příkladů, kdy jsou fuzzy množiny užitečným a praktickým nástrojem.

Výrazem $F(U)$ se bude nadále rozumět (ostrá) množina všech fuzzy množin definovaných na universu U . Dále výraz pro stupeň příslušnosti $\mu_A(x)$ se často pro zjednodušení nahrazuje zkráceným zápisem $A(x)$ - jeho význam zůstává nezměněn. Stupně příslušnosti tedy udávají, do jaké míry jednotlivé prvky do fuzzy množiny patří - díky tomu, že se pohybují spojitě na intervalu $\langle 0, 1 \rangle$, lze s jejich pomocí kvantifikovat i nepřesné či neurčité pojmy. Tato vlastnost je důvodem, proč fuzzy množiny dostaly své jméno (z angl. *fuzzy* - nejasný, neostrý, mlhavý). Z definice fuzzy množiny je také zřejmé, že ostrá množina je zde jen speciálním případem obecněji definované fuzzy množiny.

Je dobré na tomto místě zmínit, že zejména v případě jazykových termínů se jejich význam může značně lišit v závislosti na osobě, která tyto termíny použije. Dva jedinci si tak pod stejným termínem mohou představit dvě naprosto rozdílné fuzzy množiny.²

1.2. Základní pojmy z teorie fuzzy množin

V této části bude definováno několik důležitých pojmů, které jsou důležité jak pro bližší pochopení konceptu fuzzy množin, tak pro lepší orientaci v praktické části práce.

Definice 1.3 (Základní pojmy). Nechť U je neprázdná množina, tzv. universum, a A je fuzzy množina na U , tj. $A \in F(U)$.

- Pro $\alpha \in \langle 0, 1 \rangle$ nazveme α -řezem fuzzy množiny A ostrou množinu A_α , pro kterou platí: $A_\alpha = \{x \in U \mid A(x) \geq \alpha\}$.
- Nosičem fuzzy množiny A nazveme ostrou množinu $\text{Supp } A$, pro kterou platí: $\text{Supp } A = \{x \in U \mid A(x) > 0\}$.
- Jádrem fuzzy množiny A nazveme ostrou množinu $\text{Ker } A$, pro kterou platí: $\text{Ker } A = \{x \in U \mid A(x) = 1\}$.

²Vraťme se k příkladu s termínem *mladý člověk* - lze očekávat, že představa 17letého a 70letého jedince o tom, kdo je „mladý člověk“, se bude značně lišit.

- *Výškou fuzzy množiny* A nazveme číslo $\text{hgt } A = \sup \{A(x) \mid x \in U\}$.

Poznámka 1.1. Pokud existuje alespoň jeden prvek $x \in U$ takový, že $A(x) = 1$, tj. fuzzy množina A má neprázdné jádro, pak takovou fuzzy množinu označujeme jako *normální*.

Právě vlastnost uvedená v poznámce 1.1 je klíčová pro zavedení fuzzy čísel v následujícím textu.

1.3. Fuzzy čísla

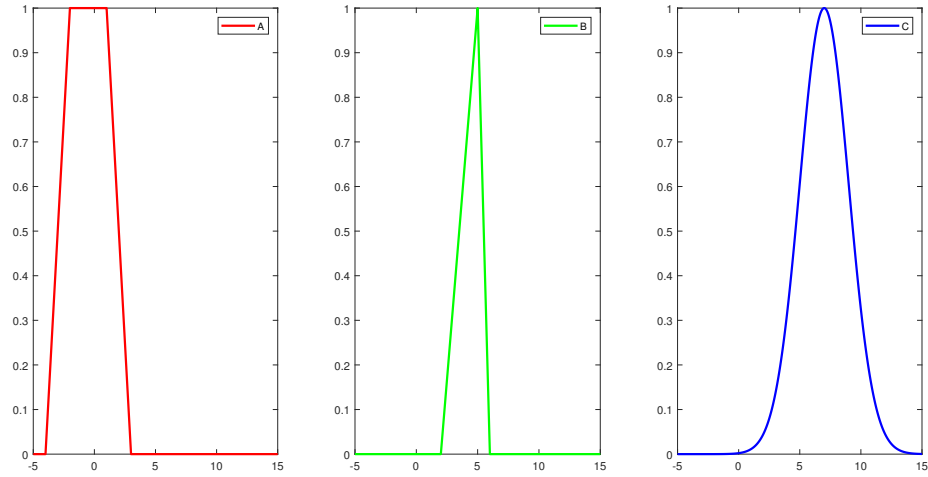
Speciálním případem fuzzy množin jsou tzv. *fuzzy čísla*. Universem, na kterém je fuzzy číslo definováno, je v tomto případě množina všech reálných čísel \mathbb{R} , resp. nějaký uzavřený interval $\langle a, b \rangle \subset \mathbb{R}$.

Definice 1.4 (Fuzzy číslo). Fuzzy množina $C \in F(\mathbb{R})$ se nazývá *fuzzy číslem*, jestliže splňuje vlastnosti:

- $\exists x_0 \in \mathbb{R} : C(x_0) = 1$, tj. C je normální fuzzy množina,
- $\forall \alpha \in (0, 1) : C_\alpha$ je uzavřený interval,
- nosič fuzzy množiny C ($\text{Supp } C$) je omezený.

Poznámka 1.2. V literatuře lze najít více způsobů, jak fuzzy číslo definovat. Výše uvedená definice je výhodná z toho důvodu, že s její pomocí lze reálná čísla, resp. uzavřené intervaly na reálné ose, zahrnout mezi fuzzy čísla jakožto jejich speciální případy. Pro obě varianty lze všechny vlastnosti fuzzy čísla z definice 1.4 snadno ověřit - pro číslo $A = a \in \mathbb{R}$ a uzavřený interval $B = \langle b, c \rangle \subset \mathbb{R}$ platí:

- $A(a) = 1, \quad A(x) = 0 \quad \forall x \in \mathbb{R} \setminus \{a\},$
 $B(x) = 1 \quad \forall x \in \langle b, c \rangle, \quad B(x) = 0 \quad \forall x \in \mathbb{R} \setminus \langle b, c \rangle;$
- $A_\alpha = \{a\} \quad \forall \alpha \in (0, 1), \quad B_\alpha = \langle b, c \rangle \quad \forall \alpha \in (0, 1);$
- $\text{Supp } A = \{a\}, \text{Supp } B = \langle b, c \rangle.$



Obrázek 1.1: Různé typy fuzzy čísel.

Pro zkrácení zápisu budeme nadále používat značení $C \in F_N(\mathbb{R})$ vyjadřující fuzzy číslo C z množiny fuzzy čísel definovaných na množině reálných čísel. Obdobné značení lze zavést pro fuzzy čísla definovaná na uzavřeném intervalu $\langle a, b \rangle \subset \mathbb{R}$ (tj. pokud $\text{Supp } C \subseteq \langle a, b \rangle$) - množinu takových fuzzy čísel budeme značit $F_N(\langle a, b \rangle)$. Několik různých fuzzy čísel je znázorněno na obrázku 1.1.

O vlastnosti fuzzy čísel, díky které je lze vyčlenit z obecných fuzzy množin, pojednává věta 1.1.

Věta 1.1. Fuzzy množina $C \in F(\mathbb{R})$ je fuzzy číslem právě tehdy, když existují reálná čísla $x_1 \leq x_2 \leq x_3 \leq x_4$ taková, že:

$$C(x) = \begin{cases} L(x), & \text{jestliže } x \in (-\infty, x_2), \\ 1, & \text{jestliže } x \in \langle x_2, x_3 \rangle, \\ P(x), & \text{jestliže } x \in (x_3, \infty), \end{cases}$$

kde $L(x)$ a $P(x)$ jsou funkce s následujícími vlastnostmi:

- $L(x)$ je funkce neklesající, zprava spojitá pro $x \in (-\infty, x_2)$ a nulová pro $x \in (-\infty, x_1)$,
- $P(x)$ je funkce nerostoucí, zleva spojitá pro $x \in (x_3, \infty)$ a nulová pro $x \in (x_4, \infty)$.

Důkaz. Viz [15]. □

Hodnoty x_1, x_2, x_3, x_4 ve větě 1.1 nejsou určeny jednoznačně, podstatné je, že nějaké takové hodnoty existují. Speciální čtveřicí těchto hodnot jsou tzv. *význačné hodnoty* fuzzy čísla C .

Definice 1.5 (Význačné hodnoty fuzzy čísla). Nechť $C \in F_N(\mathbb{R})$. *Význačnými hodnotami fuzzy čísla C* rozumíme reálná čísla $c_1 \leq c_2 \leq c_3 \leq c_4$, pro která platí:

- $\overline{\text{Supp } C} = \langle c_1, c_4 \rangle$,
- $C_1 = \langle c_2, c_3 \rangle$,

kde $\overline{\text{Supp } C}$ označuje uzávěr nosiče fuzzy čísla C .

Význačné hodnoty fuzzy čísla tedy poskytují základní představu o fuzzy číslu. Zatímco hodnoty c_1 a c_4 ohraničují nosič fuzzy čísla C , interval $\langle c_2, c_3 \rangle$ je jeho jádrem - kromě spojitosti a intervalů monotonie ale nevíme nic bližšího o tvaru funkce příslušnosti.

Pro zápis fuzzy čísel se často používá dvojice funkcí $\underline{c}(\alpha)$ a $\bar{c}(\alpha)$ definovaných na intervalu $\langle 0, 1 \rangle$. Jedná se o alternativu zápisu pomocí systému α -řezů - funkce $\underline{c}(\alpha)$ popisuje nejmenší hodnotu spadající do α -řezu pro dané α , zatímco $\bar{c}(\alpha)$ hodnotu největší. Pomocí těchto funkcí lze tedy fuzzy číslo C zapsat vztahem $C = \{ \langle \underline{c}(\alpha), \bar{c}(\alpha) \rangle, \alpha \in \langle 0, 1 \rangle \}$, kde:

$$\underline{c}, \bar{c} : \langle 0, 1 \rangle \rightarrow \mathbb{R} :$$

$$C_\alpha = \langle \underline{c}(\alpha), \bar{c}(\alpha) \rangle, \alpha \in (0, 1),$$

$$\overline{\text{Supp } C} = \langle \underline{c}(0), \bar{c}(0) \rangle.$$

Pokud tedy máme k dispozici fuzzy číslo C ve tvaru $C = \{ \langle \underline{c}(\alpha), \bar{c}(\alpha) \rangle, \alpha \in \langle 0, 1 \rangle \}$, lze potom pro každé $x \in \mathbb{R}$ určit stupeň příslušnosti $\mu_C(x)$ vztahem

$$\mu_C(x) = \begin{cases} \max \{ \alpha \mid x \in \langle \underline{c}(\alpha), \bar{c}(\alpha) \rangle \}, & \text{jestliže } \{ \alpha \mid x \in \langle \underline{c}(\alpha), \bar{c}(\alpha) \rangle \} \neq \emptyset, \\ 0, & \text{jinak.} \end{cases}$$

Je důležité uvědomit si rozdíl mezi funkcí příslušnosti μ_C a funkcemi $\underline{c}(\alpha)$ a $\bar{c}(\alpha)$. Zatímco μ_C zobrazuje z množiny reálných čísel \mathbb{R} na interval $\langle 0, 1 \rangle$, u funkcí $\underline{c}(\alpha)$ a $\bar{c}(\alpha)$ je tomu přesně naopak.

Definice 1.4 umožňuje zavádět fuzzy čísla velmi obecně, v praxi se ovšem běžně vyskytují fuzzy čísla spadající do některé ze speciálních struktur fuzzy čísel. Tyto struktury se vyznačují specifickými vlastnostmi, které práci s danými fuzzy čísly z velké části usnadňují. Mezi nejběžnější (a také nejjednodušší) struktury fuzzy čísel patří lineární fuzzy čísla, resp. po částech lineární fuzzy čísla třídy n .

Definice 1.6 (Lineární fuzzy číslo). $C \in F_N(\mathbb{R})$ je lineární fuzzy číslo, jestliže existují reálná čísla $c_1 \leq c_2 \leq c_3 \leq c_4$ taková, že platí:

$$C(x) = \begin{cases} \frac{x-c_1}{c_2-c_1}, & \text{jestliže } x \in \langle c_1, c_2 \rangle, \\ 1, & \text{jestliže } x \in \langle c_2, c_3 \rangle, \\ \frac{c_4-x}{c_4-c_3}, & \text{jestliže } x \in \langle c_3, c_4 \rangle, \\ 0, & \text{jestliže } x \in (-\infty, c_1) \cup (c_4, \infty). \end{cases}$$

Z definice 1.6 je patrné, že funkce příslušnosti μ_C je na jednotlivých intervalech tvořena lineárními funkcemi - odtud vyplývá název této struktury. Výhodou lineárních fuzzy čísel je jejich jednoduchost - pro znalost celého fuzzy čísla nám stačí znát pouze jeho význačné hodnoty, protože právě ty nám jednoznačně udávají tvar jeho funkce příslušnosti.³ Podle vztahu těchto význačných hodnot lze lineární čísla rozdělit ještě na další specifické skupiny:

- reálné číslo k : ($k =$) $c_1 = c_2 = c_3 = c_4$,
- reálný interval $\langle l, m \rangle$: ($l =$) $c_1 = c_2 < c_3 = c_4$ ($= m$),
- trojúhelníkové fuzzy číslo: $c_1 < c_2 = c_3 < c_4$,
- lichoběžníkové fuzzy číslo: $c_1 < c_2 < c_3 < c_4$,

³Právě pro tuto vlastnost a snadnou interpretaci nacházejí lineární fuzzy čísla využití např. při expertním hodnocení, kdy expertovi stačí u nějakého kvantitativního kritéria zadat hodnotu předpokládanou, minimální a maximální možnou (v případě trojúhelníkového fuzzy čísla).

- fuzzy číslo typu S: $c_1 < c_2 \leq c_3 = c_4$,
- fuzzy číslo typu Z: $c_1 = c_2 \leq c_3 < c_4$.

Poznámka 1.3. V dalším textu budou hojně využívána právě trojúhelníková fuzzy čísla. Ta se budou nadále zapisovat pomocí význačných hodnot, a to ve tvaru $A = (a_A, c_A, b_A)$ (vynechá se tedy zdvojený zápis jádra). Dolní index zde odpovídá fuzzy číslu, které význačné hodnoty popisují.

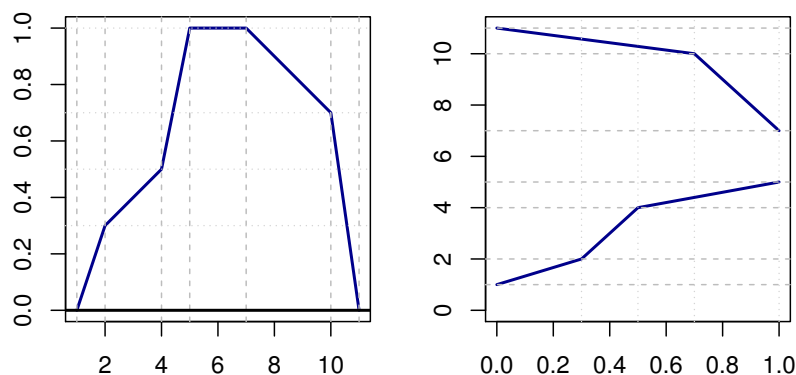
Jak už název napovídá, po částech lineární fuzzy čísla třídy n jsou do jisté míry podobná lineárním fuzzy číslům. Pro jejich zápis ale potřebujeme znát $n+1$ α -řezů tohoto fuzzy čísla a také uzávěr jeho nosiče $\overline{\text{Supp } C}$. Díky tomu máme posloupnost čísel $0 < \alpha_1 < \alpha_2 < \dots < \alpha_n < 1$ a příslušné hodnoty $\underline{c}(0), \underline{c}(\alpha_1), \dots, \underline{c}(\alpha_n), \underline{c}(1)$ a $\bar{c}(0), \bar{c}(\alpha_1), \dots, \bar{c}(\alpha_n), \bar{c}(1)$, tedy jakousi diskrétní aproximaci fuzzy čísla. Lineárním proložením příslušných hodnot pak získáme příslušné po částech lineární fuzzy číslo. Tento způsob zavedení fuzzy čísla se často využívá v případě, kdy jej nelze zavést analyticky, případně by analytické vyjádření bylo zbytečně složité (dosažená přesnost by za cenu složitosti nebyla výhodná). Je přirozené, že s rostoucím n , tj. s větším počtem zavedených α -řezů, bude aproximace fuzzy čísla přesnější. Příklad po částech lineárního fuzzy čísla lze vidět na obrázku 1.2.

Vraťme se na okamžik k trojúhelníkovým fuzzy číslům, která budou hojně využívána v části práce věnované aplikaci teorie fuzzy množin. Jak už bylo zmíněno výše, lze je popsat pomocí trojice význačných hodnot - zde pro fuzzy číslo A v podobě $A = (a_A, c_A, b_A)$. Někdy se ovšem používá i alternativní zápis, kterým získáme tzv. *LR fuzzy čísla* (více např. v [31]) v podobě $A = (c_A, l_A, r_A)_{LR}$, kdy hodnota c_A zůstává nezměněna (odpovídá jádru fuzzy čísla A) a hodnoty l_A , resp. r_A , udávají jeho neurčitost na levou, resp. pravou stranu. Oba zápisy poskytují stejnou informaci a je mezi nimi jednoznačný přechod; platí zde vztah

$$(a_A, c_A, b_A) = (c_A, c_A - a_A, b_A - c_A)_{LR},$$

resp.

$$(c_A, l_A, r_A)_{LR} = (c_A - l_A, c_A, c_A + r_A).$$



Obrázek 1.2: Po částech lineární fuzzy číslo a jeho příslušné funkce $\underline{c}(\alpha)$ a $\bar{c}(\alpha)$.

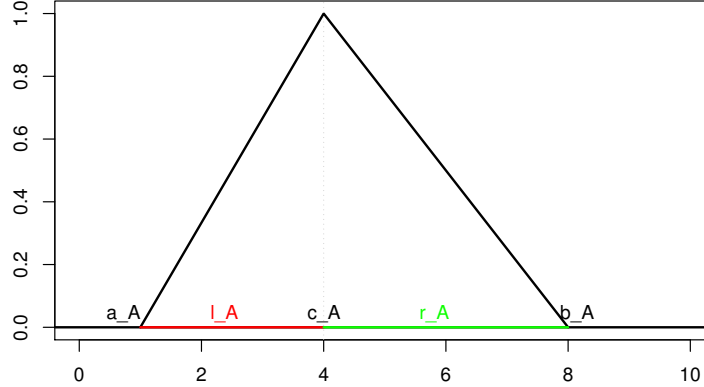
Vztah mezi oběma zápisy je přiblížen na obrázku 1.3. Tento LR zápis se může hodit při některých matematických operacích s fuzzy čísly, protože výrazně zjednoduší některé výpočetní vztahy.

1.3.1. Operace s fuzzy čísly

Aby bylo možné zadefinovat základní aritmetické operace (sčítání, odčítání, násobení, dělení) pro fuzzy čísla, je nejprve potřeba zobecnit zobrazení pro případ fuzzy množin, tj. ošetřit, jakým způsobem bude naloženo se stupni příslušnosti jednotlivých zobrazovaných prvků. K tomu slouží tzv. *princip rozšíření*.

Definice 1.7 (Princip rozšíření). Nechť $U, V \neq \emptyset$ a $f : U \rightarrow V$. *Fuzzy rozšířením* (fuzzifikací) *zobrazení* f rozumíme zobrazení $f_F : F(U) \rightarrow F(V)$, které každé fuzzy množině $A \in F(U)$ přiřazuje fuzzy množinu $f_F(A) \in F(V)$, která je určena následující funkcí příslušnosti:

$$\forall y \in V : f_F(A)(y) = \begin{cases} \sup \{A(x) \mid x \in U, f(x) = y\}, & \text{jestliže } f^{-1}(y) \neq \emptyset, \\ 0, & \text{jestliže } f^{-1}(y) = \emptyset, \end{cases}$$



Obrázek 1.3: Zápisy trojúhelníkového fuzzy čísla - graficky.

kde

$$f^{-1}(y) = \{x \in U \mid f(x) = y\}.$$

Základní myšlenka principu rozšíření je následující: prvky z množiny U se do množiny V zobrazují včetně svých stupňů příslušnosti. Pokud se více prvků z U zobrazí na tentýž prvek y z V , potom se prvku y přiřadí nejvyšší stupeň příslušnosti z těchto prvků (resp. přesněji supremum z těchto stupňů příslušnosti).

V následující definici vyžadujeme, aby fuzzy čísla A a B byla neinteraktivní. Jednoduše řečeno to znamená, že požadujeme, aby byly možné všechny kombinace hodnot s nenulovými stupni příslušnosti obou fuzzy čísel (tj. aby neexistovala nějaká vnitřní závislost, resp. podmíněnost, kvůli které by některé kombinace nemohly nastat - blíže je tato vlastnost rozebrána např. v [7]).

Definice 1.8 (Aritmetika fuzzy čísel). Nechtě $A, B \in F_N(\mathbb{R})$ jsou neinteraktivní fuzzy čísla, kdy $A = \{\langle \underline{a}(\alpha), \bar{a}(\alpha) \rangle, \alpha \in \langle 0, 1 \rangle\}$, $B = \{\langle \underline{b}(\alpha), \bar{b}(\alpha) \rangle, \alpha \in \langle 0, 1 \rangle\}$ a $k \in \mathbb{R}$. Pak platí:

- součin reálné konstanty k a fuzzy čísla A je dán vztahem

$$k \cdot A = \{\langle \min \{k \cdot \underline{a}(\alpha), k \cdot \bar{a}(\alpha)\}, \max \{k \cdot \underline{a}(\alpha), k \cdot \bar{a}(\alpha)\} \rangle, \alpha \in \langle 0, 1 \rangle\},$$

- *součet fuzzy čísel* A a B je dán vztahem

$$A + B = \{ \langle \underline{a}(\alpha) + \bar{a}(\alpha), \underline{b}(\alpha) + \bar{b}(\alpha) \rangle, \alpha \in \langle 0, 1 \rangle \},$$

- *rozdíl fuzzy čísel* A a B je dán vztahem

$$A - B = \{ \langle \underline{a}(\alpha) - \bar{b}(\alpha), \bar{a}(\alpha) - \underline{b}(\alpha) \rangle, \alpha \in \langle 0, 1 \rangle \},$$

- *součin fuzzy čísel* A a B je dán vztahem

$$A \cdot B = \{ \langle \min \{ \underline{a}(\alpha) \cdot \underline{b}(\alpha), \underline{a}(\alpha) \cdot \bar{b}(\alpha), \bar{a}(\alpha) \cdot \underline{b}(\alpha), \bar{a}(\alpha) \cdot \bar{b}(\alpha) \}, \max \{ \underline{a}(\alpha) \cdot \underline{b}(\alpha), \underline{a}(\alpha) \cdot \bar{b}(\alpha), \bar{a}(\alpha) \cdot \underline{b}(\alpha), \bar{a}(\alpha) \cdot \bar{b}(\alpha) \} \rangle, \alpha \in \langle 0, 1 \rangle \},$$

- necht' $0 \notin \overline{\text{Supp } B}$, pak *podíl fuzzy čísel* A a B je dán vztahem

$$\frac{A}{B} = \left\{ \left\langle \min \left\{ \frac{\underline{a}(\alpha)}{\underline{b}(\alpha)}, \frac{\underline{a}(\alpha)}{\bar{b}(\alpha)}, \frac{\bar{a}(\alpha)}{\underline{b}(\alpha)}, \frac{\bar{a}(\alpha)}{\bar{b}(\alpha)} \right\}, \max \left\{ \frac{\underline{a}(\alpha)}{\underline{b}(\alpha)}, \frac{\underline{a}(\alpha)}{\bar{b}(\alpha)}, \frac{\bar{a}(\alpha)}{\underline{b}(\alpha)}, \frac{\bar{a}(\alpha)}{\bar{b}(\alpha)} \right\} \right\rangle, \alpha \in \langle 0, 1 \rangle \right\}.$$

Poznámka 1.4. Výše uvedené vztahy je možné pro trojúhelníková fuzzy čísla zjednodušit. Je ale potřeba upozornit na to, že zatímco součet a rozdíl dvou trojúhelníkových fuzzy čísel je opět trojúhelníkovým fuzzy číslem, u součinu tato skutečnost neplatí - funkce příslušnosti je zde obecně nelineární. V praxi se však součin často aproximuje níže uvedeným trojúhelníkovým fuzzy číslem - proto je v příslušných vztazích uvedeno znaménko \approx .

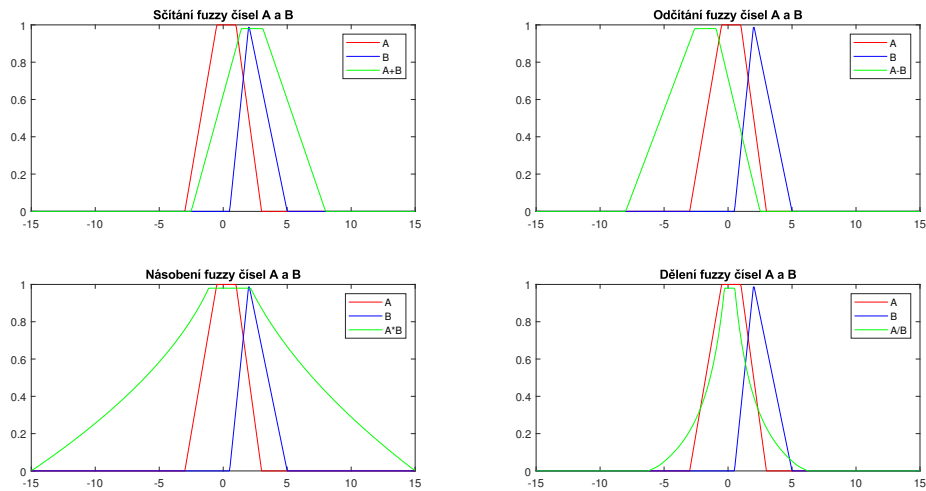
Pro $k \in \mathbb{R}^+$ a trojúhelníková fuzzy čísla $A = (a_A, c_A, b_A)$ a $B = (a_B, c_B, b_B)$ tedy platí:

- $k \cdot A = (k \cdot a_A, k \cdot c_A, k \cdot b_A),$
- $A + B = (a_A + a_B, c_A + c_B, b_A + b_B),$
- $A - B = (a_A - b_B, c_A - c_B, b_A - a_B),$

- jestliže $a_A, c_A, b_A, a_B, c_B, b_B \in \mathbb{R}^+$, pak i $A \cdot B \approx (a_A \cdot a_B, c_A \cdot c_B, b_A \cdot b_B)$,

pro trojúhelníková fuzzy čísla $C = (c_C, l_C, r_C)_{LR}$ a $D = (c_D, l_D, r_D)_{LR}$, kde $c_C, l_C, r_C, c_D, l_D, r_D \in \mathbb{R}^+$, potom obdobně:

- $k \cdot C = (k \cdot c_C, k \cdot l_C, k \cdot r_C)_{LR}$
- $C + D = (c_C + c_D, l_C + l_D, r_C + r_D)_{LR}$,
- $C - D = (c_C - c_D, l_C + r_D, r_C + l_D)_{LR}$,
- $C \cdot D \approx (c_C c_D, c_C l_D + c_D l_C - l_C l_D, c_C r_D + c_D r_C + r_C r_D)_{LR}$.



Obrázek 1.4: Fuzzy aritmetika.

1.3.2. Defuzzifikace fuzzy čísel

Při práci a výpočtech s fuzzy čísly je někdy nutné výsledek v podobě fuzzy čísla „sesumírovat“ do podoby reálného čísla, tj. defuzzifikovat. Defuzzifikace je potřebná např. ve chvíli, kdy chceme fuzzy čísla uspořádat nebo je použít v rámci nějakého rozhodovacího problému, kde jsou nutné ostré vstupy. V této kapitole bude uvedeno několik způsobů defuzzifikace fuzzy čísla.

- *Těžištěm fuzzy čísla* A nazveme reálné číslo t_A , pro které platí:

$$t_A = \begin{cases} \frac{\int_{\mathbb{R}} x \cdot A(x) dx}{\int_{\mathbb{R}} A(x) dx}, & \text{jestliže } A \notin \mathbb{R}, \\ A, & \text{jestliže } A \in \mathbb{R}. \end{cases}$$

- *Zobecněným těžištěm fuzzy čísla* A nazveme reálné číslo t_A^k , pro které platí:

$$t_A^k = \begin{cases} \frac{\int_{\mathbb{R}} x \cdot A(x)^{k+1} dx}{\int_{\mathbb{R}} A(x)^{k+1} dx}, & \text{jestliže } A \notin \mathbb{R}, \\ A, & \text{jestliže } A \in \mathbb{R}. \end{cases}$$

- *Středem maxim fuzzy čísla* A nazveme reálné číslo $\text{MOM}(A)$, pro které platí:

$$\text{MOM}(A) = \frac{a(1) + \bar{a}(1)}{2}.$$

- *Středem průměrného intervalu fuzzy čísla* A nazveme reálné číslo $\bar{E}(A)$, pro které platí:

$$\bar{E}(A) = \int_0^1 \frac{a(\alpha) + \bar{a}(\alpha)}{2} d\alpha.$$

- *Mediánem fuzzy čísla* A nazveme reálné číslo m_A , pro které platí:

$$\int_{-\infty}^{m_A} A(x) dx = \int_{m_A}^{\infty} A(x) dx.$$

Poznámka 1.5. Jestliže $k \rightarrow \infty$, potom výraz pro zobecněné těžiště $\lim_{k \rightarrow \infty} t_A^k$ bude splývat se středem maxim $\text{MOM}(A)$.

Na tomto místě opět zmíníme určité zjednodušení vzorců pro trojúhelníková fuzzy čísla. Uvažujme nyní trojúhelníkové fuzzy číslo $A = (a_A, c_A, b_A)$. Nejjednodušší je vztah pro střed maxim fuzzy čísel - na první pohled je zřejmé, že u trojúhelníkového fuzzy čísla A bude pro střed maxim platit

$$\text{MOM}(A) = \frac{c_A + c_A}{2} = c_A,$$

protože trojúhelníkové fuzzy číslo má jednoprvkové jádro. Zjednodušit lze také vztah pro střed průměrného intervalu lineárního fuzzy čísla - tato charakteristika se spočítá následovně [18]:

$$\overline{E}(A) = \frac{a_A + 2c_A + b_A}{4}.$$

Co se týče výpočtu těžiště fuzzy čísla, lze dokázat (např. [18]), že pro trojúhelníkové fuzzy číslo platí

$$t_A = \frac{1}{3} \cdot (a_A + c_A + b_A).$$

Na závěr této podkapitoly je dobré zmínit, že použití různých způsobů defuzzifikace pro stejná fuzzy čísla může zapříčinit rozdílné výsledky např. v rámci zmíněného uspořádání fuzzy čísel.

Kapitola 2

Pojišťovnictví

Vedle odvětví bankovníctví je pojišťovnictví jedním ze základních pilířů finančního sektoru každé vyspělé ekonomiky. Jeho fungování pomáhá udržet ekonomickou stabilitu země tím, že komerční pojišťovny z podstaty své činnosti dokáží krýt škody, které by pro samostatné subjekty (osoby, domácnosti, podniky) mohly mít likvidační následky.

Tato kapitola čtenáře uvede do problematiky pojišťovnictví: budou vysvětleny základní pojmy a stručně představeno fungování pojistného trhu. Měla by poskytnout dostatečný přehled pro lepší orientaci v kapitolách věnujících se aplikacím teorie fuzzy množin v pojišťovnictví.

2.1. Pojištění

Pod pojmem *pojištění* obecně rozumíme jednu z možností, jak se mohou jednotlivé ekonomické subjekty chránit před některými riziky prostřednictvím (nejčastěji) komerčních pojišťoven⁴. Je to efektivní způsob tvorby a užití finančních rezerv pro úhradu peněžních potřeb ekonomických subjektů, které jsou v jednotlivých případech výskytu náhodné, vcelku však odhadnutelné. V praxi to tedy znamená, že pokud je více subjektů ohroženo stejným pojistným rizikem, je možné odhadnout pravděpodobnost, s jakou se dané riziko realizuje. Škoda, která realizací rizika vznikla, se pak uhradí z prostředků, které dohromady složily všechny

⁴V tomto textu budeme uvažovat pouze komerční pojišťovny, proto slovo *komerční* bude často vynecháno.

ohrožené subjekty. Nosnou myšlenkou komerčního pojištění je skutečnost, že subjekt raději dobrovolně zaplatí relativně malou částku - v případě, že se riziko realizuje právě u něj, mu ale může být uhrazena škoda mnohonásobně vyšší.

Peněžní potřeby vznikají jako následek realizace tzv. *pojistné události*, tj. nějaké rizikové skutečnosti specifikované v pojistné smlouvě, jejímž důsledkem může subjekt utrpět ztrátu. Takové riziko, kdy může nastat pouze negativní vychýlení od současného stavu, se nazývá *čisté riziko*. Aby pojišťovna vůbec mohla riziko pojistit, musí riziko splňovat několik kritérií pojistitelnosti: musí být přesně identifikovatelné, nahodilé, vyčíslitelné, ekonomicky přijatelné pro pojišťovnu a pojišťovna jeho přijetím nesmí porušovat morální zásady [10].

Pojištění se dělí na dvě základní odvětví: *životní* a *neživotní*, kterému se v rámci této práce budeme věnovat blíže. U životního pojištění jsou možnými pojistnými událostmi smrt nebo dožití se určitého věku - všechna ostatní pojistná nebezpečí jsou kryta v rámci pojištění neživotního. Mezi nejběžnější druhy neživotního pojištění patří pojištění majetku (např. havarijní pojištění), pojištění osob nebo pojištění odpovědnosti.

Uzavření pojištění je ve většině případů podmíněno uzavřením *pojistné smlouvy*, a to mezi dvěma stranami - pojistitelem a pojistníkem. *Pojistitel* je právnická osoba, která má od orgánu dozoru (u nás je jím Česká národní banka) uděleno povolení k provozování pojistné činnosti. Nejčastěji jsou pojistiteli právě komerční pojišťovny. Pojistitel se zavazuje, že v případě realizace pojistné události specifikované v pojistné smlouvě vyplatí pojištěnému smluvně stanovenou náhradu škody, tzv. *pojistné plnění*. *Pojistníkem* je fyzická nebo právnická osoba, které uzavřením pojistné smlouvy vzniká povinnost vyplácet pojistiteli úplatu za pojistnou ochranu, tzv. *pojistné*, ve stanovené výši a termínech. V neživotním pojištění se dále v pojistné smlouvě vyskytuje *pojištěný* (tj. osoba, na jejíž život, zdraví, majetek, odpovědnost nebo jiné hodnoty pojistného zájmu se pojištění vztahuje), případně *oprávněná osoba* (tj. osoba, které v případě realizace pojistné události vzniká právo na pojistné plnění).

2.2. Technické rezervy

Činnost pojišťoven je z velké části ovlivňována náhodou - dopředu se přesně neví, jestli, kdy, příp. kolikrát pojistná událost u daného jedince nastane. Pojišťovna sice má k dispozici údaje z minulosti, díky kterým si může určit očekávaný průběh škod, skutečný stav však může být od tohoto odhadu vychýlen jak v kladném, tak v záporném směru. Finance pro vyplacení pojistných plnění jsou navíc tvořeny kumulací prostředků přijatých na pojistném (tzv. *nettopojistným*). Rozsah pojistného kmene (tj. počet jedinců uzavírajících podobné pojištění) také není dopředu známý, stejně tak doba, po kterou se pojistné plnění bude muset vyplácet. Z těchto důvodů musí mít pojišťovna k dispozici prostředky, ze kterých může hradit pojistná plnění i v době, kdy jí zdroje z placeného pojistného nestačují. Tyto prostředky se nazývají *technické rezervy* - v této práci bude pozornost později věnována konkrétně technickým rezervám v neživotním pojištění.

Až do nedávné minulosti u nás dle předchozí legislativy (tzv. *Solvency I*) bylo zavedeno striktní dělení jednotlivých složek technických rezerv v neživotním pojištění. Jednalo se o rezervu na pojistné plnění, rezervu na nezasloužené pojistné, rezervu na prémie a slevy, rezervu pojistného neživotních pojištění, vyrovnávací rezervu, případné další rezervy pak musel pojišťovně schválit orgán státního dozoru. Zároveň byla stanovena minimální výše rezerv, kterou musela pojišťovna disponovat. Od prosince roku 2016 platí v České republice nová úprava této legislativy, tzv. *Solvency II*, která toto dělení již nevyžaduje - pojišťovna musí nově vykazovat jen celkovou výši technických rezerv (vliv změny legislativy je blíže popsán např. v [6]).

I přesto, že pojišťovny nyní už nemají povinnost vykazovat jednotlivé složky rezerv zvlášť, pro výpočet celkových technických rezerv se tyto složky i nadále počítají. V této práci se budeme věnovat nejvýznamnější složce technických rezerv, a to rezervě na pojistná plnění v neživotním pojištění. Tato rezerva je tvořena dvěma částmi - rezervou IBNR a RBNS. Rezerva IBNR (*incurred but not reported*) je tvořena pro škody z pojistných událostí, které nastaly v aktuálním účetním období, ale ještě nebyly pojišťovně nahlášený. Naproti tomu rezerva RBNS (*re-*

ported but not settled) je pro případy, kdy pojistná události již byla v daném účetním období nahlášena, ale nebyla zcela zlikvidována, příp. vyplacena.

Ačkoli existují způsoby, jak složky IBNR a RBNS odhadovat zvlášť, zde se budeme v kapitole 4 věnovat způsobu, kde tyto složky odhadujeme dohromady pomocí tzv. trojúhelníkových schémat. Ještě předtím ale budou stručně představeny některé z dosavadních aplikací teorie fuzzy množin v pojišťovnictví.

Kapitola 3

Využití aparátu teorie fuzzy množin v pojišťovnictví

I přes poměrně krátkou historii problematiky fuzzy množin byla zaznamenána řada způsobů, jak teorii fuzzy množin (TFM) implementovat v pojišťovnictví. Tato kapitola čerpá především ze staršího přehledového článku [23] a jeho aktualizované verze [22], dále pak z článků novějších [2, 3, 24, 25, 26]. Cílem této kapitoly je především poskytnout stručné přiblížení vybraných aplikací a motivaci pro jejich podrobnější studium.

Poznámka 3.1. Aby se v následujícím textu předešlo problému s identifikací ostrých a fuzzy hodnot, budou nadále fuzzifikované hodnoty označeny vlnovkou, tj. např. \tilde{A} .

3.1. Aplikace TFM v neživotním pojištění

3.1.1. Detekce pojistných podvodů v havarijním pojištění

Havarijní pojištění je jednou z oblastí neživotního pojištění, kde se pojišťovny relativně běžně setkávají s pokusy o pojistný podvod. Může se jednat o „nahranou“ nehodu, případně se může žádat o pojistné plnění za událost, která nastala ještě před uzavřením pojištění. Při množství pojistných smluv, které je v rámci havarijního pojištění uzavřeno, nemusí být snadné kontrolovat podrobně všechny souvislosti pojistné události, na druhou stranu při velkém počtu po-

jistných podvodů se může zkreslit představa pojišťovny o četnosti a závažnosti pojistných událostí - jejich umělé navýšení by zapříčinilo zdražení pojištění pro „pocitivé“ klienty, pro které by se potom ovšem mohlo stát nevýhodným. Snahou tedy je vypořádat podezřelé pojistné smlouvy a na ně se soustředit.

V [24] byl nastíněn způsob, jak pro vyhledávání těchto smluv využít fuzzi-fikovanou metodu nehierarchického shlukování - metodu k průměrů. Principem „klasické“ metody k průměrů je vytvořit shluky tak, aby se minimalizoval součet čtvercových vzdáleností prvku od centra shluku v rámci skupin - tj. při zařazení libovolného prvku do jiného shluku by se celková hodnota kritéria zvýšila. Výsledkem je tedy ostré zařazení prvků do příslušných shluků. Oproti tomu metoda fuzzy k průměrů umožňuje překrývání shluků - prvek tedy může patřit do více shluků najednou v nějakém stupni (součet těchto stupňů musí být pro každý prvek roven 1).

Uvažujme nyní n pozorování představující jednotlivé pojistné smlouvy. Souřadnicemi každého pozorování d jsou údaje z této smlouvy (obvykle znormované, aby jednotky jednotlivých údajů nezkrasovaly výsledek). Dále označme dopředu určený počet shluků (jejich optimální počet lze určit z dat různými způsoby) jako c a v_i jako střed i -tého shluku. Cílem je vytvořit shluky tak, aby minimalizovaly funkci

$$J_m = \sum_{j=1}^n \sum_{i=1}^c (u_{ij}^m) B_{ij}(v_i, d_j), \quad (3.1)$$

kde u_{ij} značí stupeň příslušnosti j -tého prvku k i -tému shluku a B_{ij} je (zde euklidovská) vzdálenost mezi středem i -tého shluku a prvkem d_j . Koeficient m potom popisuje míru překrývání shluků.

Na úvod metody se prvky rozdělí do shluků (např. pomocí cross-validace) - dále se aplikuje algoritmus metody fuzzy k průměrů z [24] minimalizující hodnotu funkce (3.1). Iteračním procesem zde tedy vznikne c shluků se svými středy. Vzniklé shluky potom odpovídají bezproblémovým pojistným smlouvám. Pro každý prvek se spočítá jeho vzdálenost od středů shluků. Tyto vzdálenosti se porovnají s dopředu pevně určenými hodnotami β_L a β_U ($\beta_L < \beta_U$), které sta-

novují jakousi dolní a horní hranici vzdáleností. Jestliže je vzdálenost prvku od shluku menší než β_L , je prvek, potažmo smlouva, pokládána za bezproblémovou. Jestliže vzdálenost překročí hodnotu β_U (tj. odpovídá odlehlému pozorování), je smlouva označena jako podvodná. Pro hodnoty mezi β_L a β_U platí, že jsou příslušné smlouvy podrobeny další kontrole a zkoumání. Kladné stránky metody jsou podle autorů zejména ve snížení počtu smluv, které jsou falešně označeny jako podvodné (ve srovnání s dalšími přístupy). Pokud se totiž zohlední, že prvek může částečně patřit do několika shluků, zmenší se obvykle jeho vzdálenost od středů - prvek proto bude spíše označen za bezproblémový, potažmo „jen“ za podezřelý.

3.1.2. Fuzzifikace jazykových termů v pojištění proti zemětřesení

Při uzavírání pojištění, případně při likvidaci pojistné události, musí zainteresované osoby pracovat jak s údaji, které jsou bez problému kvantifikovatelné (pojistná částka/hodnota, věk osoby, stáří auta atp.), tak s údaji kvalitativního charakteru. I tyto údaje je ovšem třeba zahrnout do používaných modelů (např. pro výpočet pojistného, odhad škodního průběhu apod.). V některých případech se ukázalo užitečným využít fuzzy množin právě pro jejich schopnost kvantifikovat jazykové pojmy. Ve chvíli, kdy se kvalitativní charakteristice přiřadí význam v podobě fuzzy množiny, je možné s těmito hodnotami (skrže princip rozšíření, fuzzy aritmetiku, ...) kalkulovat v matematických modelech.

Jedním z prvních případů, kde se fuzzy množin využilo při likvidaci pojistné události, byla studie [2] týkající se pojištění proti zemětřesení. Mezi vlastnostmi, které by se klasicky obtížně kvantifikovaly, byly například kvalita konstrukce nebo přítomnost určitých architektonických prvků, které mohly zabránit škodám nebo je alespoň zmírnit. V jiném případě se jednalo o údaje, které sice bylo možné vyjádřit číselně (např. míra poškození budovy), ale nebylo snadné je přesně určit (např. v procentech). V tuto chvíli se vzala na pomoc jazyková škála termů, kterým byla přiřazena konkrétní fuzzy množina. Oslovený likvidátor, příp. jiný

expert, pak mohl vybrat z několika málo termů takový, který nejlépe popisoval daný stav (např. poškození mírné/závažné/totální). Došlo tak k zjednodušení a urychlení některých fází procesu likvidace pojistné události.

3.1.3. Fuzzy regresní analýza

V [26] byla poprvé nastíněna možnost, jak rozšířit lineární regresní analýzu do fuzzy roviny. Předpokládá se zde, že odchylky mezi skutečně naměřenými hodnotami a hodnotami odhadnutými z modelu jsou způsobeny neurčitostí struktury systému, nikoli chybami měření. Parametry regresní křivky jsou zde nově stanoveny pomocí fuzzy čísel.

Uvažujme ostré množiny X a Y a zobrazení $f(x, \mathbf{a}) : X \rightarrow Y$, kde \mathbf{a} je vektor parametrů dané funkce. Jestliže tyto parametry budou dány pomocí vektoru fuzzy čísel (značeno $\tilde{\mathbf{A}}$), bude pro každou vstupní hodnotu x_i výstupem fuzzy množina \tilde{Y}_i , jejíž funkce příslušnosti bude dána předpisem

$$\mu_{\tilde{Y}_i}(y) = \begin{cases} \max \{ \mu_{\tilde{\mathbf{A}}}(\mathbf{a}) \mid y = f(x_i, \mathbf{a}) \}, & \text{jestliže } \exists \mathbf{a} \in \tilde{\mathbf{A}} : y = f(x_i, \mathbf{a}), \\ 0, & \text{jinak.} \end{cases}$$

V případě dat s ostrými hodnotami vysvětlující i vysvětlované proměnné je potom snahou fuzzy lineární regresní analýzy najít množinu parametrů $\tilde{\mathbf{A}}^*$ tak, aby fuzzy množina $\tilde{Y}_i^* = \tilde{\mathbf{A}}^* x_i$ obsahovala hodnotu y_i ve stupni příslušnosti rovném alespoň (nějaké) hodnotě h - a to pro všechna pozorování (tj. pro všechny hodnoty indexu i). Problém hledání $\tilde{\mathbf{A}}^*$ je v článku podrobně popsán - uplatňuje se zde metodika lineárního programování. Fuzzy lineární regrese je následně využita pro popis chování cen u montovaných domů v Japonsku. Autoři zdůrazňují možnost volby stupně h podle typu problému - fuzzy regrese podle nich může lépe reagovat na požadavky uživatele.

Aplikaci této myšlenky v neživotním pojištění je věnován článek [14], kdy se s využitím minulých dat použil fuzzy lineární regresní model pro určení rezervy na pojistná plnění pro pojistné události, které již nastaly, ale zatím nebyly pojišťovně ohlášené (tzv. IBNR rezerva). Náhled na problematiku pomocí fuzzy množin podle článku pomáhá pojistným matematikům zohlednit v jinak striktně

přesném modelu možnou nejistotu. Vzhledem k možnosti volby h tak pojišťovna má k dispozici matematický aparát, v němž lze s nejistotou pracovat (tj. na základě dostupných dat postupně upřesňovat fuzzy množiny odpovídající regresním parametrům).

3.2. Aplikace TFM v životním pojištění

3.2.1. Fuzzy cashflow

V celém finančním sektoru, tedy i v pojišťovnictví, hraje úlohu hodnota peněz, která se v čase mění. Aby se mohly porovnávat peněžní částky v různých časových okamžicích, je vždy nutné je úročit, resp. diskontovat ke stejnému časovému okamžiku pomocí úrokové míry i . Podle [3] mohou být údaje o uvažovaném množství peněz, úrokové míře a případně i o úrokovém období nepřesné („přibližně 4% úrok“, „15000 - 18000 Kč“ atp.) - proto je lze vnímat jako fuzzy množiny. K popisu těchto hodnot autor využívá fuzzy čísel - dle principu rozšíření pak lze fuzziifikovat standardní předpisy pro úročení a diskontování. Při fuzziifikované současné hodnotě peněz \widetilde{PV} , budoucí hodnotě peněz \widetilde{FV} a úrokové míře \tilde{i} jde o vztahy:

$$\widetilde{FV} = \widetilde{PV} \cdot (1 + \tilde{i})^n,$$

$$\widetilde{PV} = \widetilde{FV} \cdot (1 + \tilde{i})^{-n}.$$

Obdobně lze postupovat i při použití vzorce pro spojitě úročení nebo při využití tzv. efektivní úrokové míry. Podrobněji se článek zabývá fuzziifikací úrokového období \tilde{n} a jeho vlivu na výslednou hodnotu peněz, a to i s příslušnými důkazy.

Aplikace uvedených vztahů je provedena na porovnání několika investičních variant, kdy se při jejich zhodnocení využijí fuzziifikované peněžní toky (cash flow). Lze tedy určit fuzzy ekvivalent čisté současné hodnoty jakožto součtu všech peněžních toků za uvažované období pro každou variantu a na jejich základy projekty porovnat. Hlavní pozitivum tohoto postupu vidí autor v možném zahrnutí nejistoty v každém kroku postupu - protože jsou vstupy popsány pomocí fuzzy

čísel, popisuje výsledek skutečnost lépe než v případě, že by vstupy musely být v ostré podobě, a tím paradoxně nepřesnější.

Na tuto práci je odkazováno v [16] - autor článku se zabývá výpočtem pojistného u pojištění dožití se určitého věku. Kromě fuzzy úrokové míry \tilde{i} pracuje i s fuzzy pravděpodobnostmi dožití, (např. ${}_{10}\tilde{p}_{55}$ - pravděpodobnost, že se 55letý jedinec dožije věku 65 let). Pro pojistnou částku S a roční úrokovací období se potom příslušné fuzzy jednorázové pojistné \tilde{P} určí ze vztahu

$$\tilde{P} = S \cdot {}_{10}\tilde{p}_{55} \cdot \left(\frac{1}{1 + \tilde{i}} \right)^{10}.$$

Zůstává ovšem nezodpovězena otázka, jak potom tuto hodnotu fuzzy pojistného defuzzifikovat.

3.2.2. Volba životního pojištění

Aplikaci, jejíž přínos může pocítit pojistník přímo, je věnován článek [1]. Jedná se o výběr konkrétního produktu životního pojištění z nabídky různých pojišťoven, a to prostřednictvím fuzzifikovaného vícekriteriálního rozhodovacího modelu.

Rozhodovatelem je v tomto případě pojistník - v případě, že se sám v pojistných produktech neorientuje, lze za rozhodovatele v modelu brát i jednoho, popřípadě několik expertů, které si pojistník vezme na pomoc. Principem celého rozhodovacího modelu je výběr takové varianty (zde pojistného produktu), která nejlépe naplňuje rozhodovatelův cíl. Přínos oproti klasickému vícekriteriálnímu rozhodování vidí autoři v možnosti zohlednění „váhání“ rozhodovatele (potažmo expertů) prostřednictvím zahrnutí více pojmů z jazykové škály - expert tak může lépe vyjádřit svou znalost.

V prvním kroce si rozhodovatel vytvoří soubor kritérií, která jsou pro něj při uzavření pojištění důležitá (např. cena pojištění, spolehlivost pojišťovny atp.) a popisují tak jeho cíl. Tato kritéria mohou být ještě dále rozdělena na dílčí kritéria. Rozhodovatel, příp. povolání experti potom individuálně tato dílčí kritéria párově porovnají z hlediska jejich důležitosti, tedy určí, jak důležitá je preference jednoho

kritéria před druhým. Porovnání zde má podobu tabulek - důležitost je v nich popisována pojmy z jazykové škály, a to buď prostřednictvím jednoho termu (např. „málo důležité“) nebo využitím rozšířené jazykové škály (např. „středně až absolutně důležité“). Těmto jazykovým termům zároveň odpovídá bodové ohodnocení - příklad z článku [1] lze vidět v tabulce 3.1.

jazykový term	bodové ohodnocení
zcela nedůležité	0
velmi málo důležité	1
málo důležité	2
středně důležité	3
vysoce důležité	4
velmi vysoce důležité	5
absolutně důležité	6

Tabulka 3.1: Jazyková škála s příslušícím bodovým ohodnocením.

Ze zmíněných individuálních preferenčních tabulek se posléze určí finální preferenční tabulka: pro každou dvojici kritérií (tj. pro každé pole tabulky) se ze všech individuálních preferenčních tabulek vyberou jazykové termy, které odpovídají nejnižšímu, resp. nejvyššímu bodovému ohodnocení. Ta potom tvoří krajní hodnoty jakéhosi „jazykového intervalu“, který kromě těchto krajních termů zahrnuje i všechny jazykové termy mezi nimi. Jestliže tedy například 1. expert přiřadí preferenci prvního kritéria před druhým jazykový term „málo důležité“ a 2. expert „vysoce důležité až velmi vysoce důležité“, bude takto vzniklý jazykový interval pro tuto dvojici kritérií obsahovat pojmy „málo důležité“, „středně důležité“, „vysoce důležité“, „velmi vysoce důležité“. Takto vzniklý jazykový interval se potom prostřednictvím bodového ohodnocení odpovídajících jazykových termů převede do reálných čísel. Dále se zvolí charakteristika, která ho bude reprezentovat (např. zde aritmetický průměr). Po znormování těchto průměrů na hodnotu 1 je pro každé dílčí kritérium stanoven poměr, jakým se dané kritérium podílí na celkovém cíli.

Podle důsledků jednotlivých variant vzhledem ke kritériu se potom určí hodnocení variant dle jednotlivých kritérií - hodnocení všech variant podle jednoho

kritéria musí být rovno podílu tohoto kritéria na celkovém cíli. Jestliže se toto hodnocení provede pro všechna kritéria, součet těchto dílčích hodnocení pak pro každou variantu udává, do jaké míry tato varianta naplňuje celkový cíl. Logicky varianta s nejvyšším hodnocením je pro rozhodovatele variantou optimální.

Mezi další dílčí problémy, k jejichž řešení bylo také využito aparátu teorie fuzzy množin, patří např. modelování délky dožití v životním pojištění jakožto fuzzy náhodné veličiny [21], stanovení pojistného u pojištění rozsáhlých projektů (např. výstavba ropných plošin, elektráren atp.) [17], určení pro pojišťovnu optimální výše škody, při jejímž překročení zbytek škody hradí zajišťovna [16], volba optimálního skupinového zdravotního pojištění [29] aj. Těmito aplikacemi se již blíže zabývat nebudeme, pro zájemce jsou uvedeny odkazy na literaturu.

Kapitola 4

Fuzzifikace metod využívajících trojúhelníková schémata

Zatímco v předchozí kapitole byly příklady jednotlivých aplikací teorie fuzzy množin uvedeny pouze stručně, následující téma bude rozvedeno podrobněji. Jedná se o možnost fuzzifikace metody Chain-Ladder, která se používá pro výpočet rezervy na pojistná plnění v neživotním pojištění. Tato metoda patří mezi metody založené na tzv. trojúhelníkových schématech, která budou čtenáři přiblížena v úvodu kapitoly. Následovat bude popis klasické metody Chain-Ladder s aplikací na jednoduchá data pro lepší představu (pro tuto část byly informace čerpány především z [5] a [28]). Poté už bude prostor pro fuzzifikaci této metody, a to prostřednictvím různých způsobů výpočtu vývojových koeficientů.

4.1. Obecný úvod k trojúhelníkovým schématům

Trojúhelníková schémata představují jeden z nástrojů používaných pro výpočet rezervy na pojistná plnění v neživotním pojištění. Tato schémata slouží k odhadu obou složek rezervy na pojistná plnění (IBNR a RBNS) dohromady, jako celku. Mezi přednosti trojúhelníkových schémat patří jejich přehlednost a názornost. Pomocí jednoduchého grafického nástroje totiž zobrazují jak minulá, tak odhadovaná budoucí pojistná plnění.

Vstupními daty pro trojúhelníková schémata jsou údaje o pojistných plněních v neživotním pojištění vyplacených v minulosti. Tyto hodnoty jsou charakteri-

zovány dvěma roky, a to tzv. *rokem původu i* a *rokem vývoje j* .⁵ Rok původu indikuje, ve kterém roce došlo k dané pojistné události, zatímco rok vývoje vyjadřuje, kolik let od vzniku pojistné události uběhlo, resp. kolik let již probíhá likvidace pojistné události (ta v praxi může spolu s vyplácením pojistného plnění trvat několik let, v některých případech až desítky).

Na základě těchto dvou faktorů lze tedy minulá pojistná plnění sloučit - vždy se sečtou hodnoty se stejným rokem původu i a rokem vývoje j . Tyto součty se dosadí do tzv. *nekumulativního vývojového trojúhelníku* (jeho prvky se budou značit $X_{i,j}$). Hodnota $X_{i,j}$ tedy vyjadřuje množství peněz vyplacené v j -tém roce vývoje na pojistná plnění, která se vztahují k pojistným událostem z roku původu i . Příklad jednoduchého nekumulativního vývojového trojúhelníku, převzatého z [5] je možné vidět v tabulce 4.1 (hodnoty jsou uváděny v milionech Kč).

$i \backslash j$	0	1	2	3	4	5
1	9	12	10	11	8	0
2	13	15	8	10	14	
3	14	15	15	16		
4	16	8	18			
5	12	14				
6	11					

Tabulka 4.1: Nekumulativní vývojový trojúhelník.

Pro jiné výpočty může být důležité, jaká částka byla vyplacena v rámci pojistného plnění za pojistné události z jednoho konkrétního roku původu od doby jejich vzniku, tedy za všechny dosavadní roky vývoje. Tuto informaci nesou řádkové součty nekumulativního trojúhelníku. Bývá obsažena v tzv. *kumulativním vývojovém trojúhelníku* (s prvky $C_{i,j}$) - hodnota $C_{i,j}$ udává součet dosavadních pojistných plnění na události z roku původu i až do j -tého roku vývoje. Odpovídající kumulativní vývojový trojúhelník pro předchozí data je uveden v tabulce 4.2.

⁵V praxi jsou často letopočty nahrazeny indexy, které se počítají např. od 0 či od 1, aby se zjednodušil zápis a zpřehlednila práce se vzorci.

$i \backslash j$	0	1	2	3	4	5
1	9	21	31	42	50	50
2	13	28	36	46	60	
3	14	29	44	60		
4	16	24	42			
5	12	26				
6	11					

Tabulka 4.2: Kumulativní vývojový trojúhelník.

Z předchozího popisu jasně vyplývá, že nekumulativní vývojový trojúhelník jednoznačně charakterizuje trojúhelník kumulativní a naopak. Prvky nekumulativního vývojového trojúhelníku získáme z kumulativního vztahem

$$X_{i,j} = C_{i,j} - C_{i,j-1}, \quad i = 1, \dots, n-1, \quad j = 1, \dots, n-i, \quad (4.1)$$

$$X_{n,0} = C_{n,0},$$

pro opačný směr platí vztah

$$C_{i,j} = \sum_{k=0}^j X_{i,k}, \quad i = 1, \dots, n \quad j = 0, \dots, n-i. \quad (4.2)$$

Závěrem této části nebude od věci shrnout značení pro následující metody. V dalším textu budou roky původu a roky vývoje nahrazovány indexy. Pro roky původu se bude i počítat od jedničky, pro roky vývoje půjde j od nuly. Důvodem bude interpretovatelnost hodnot v trojúhelnících. Předpokládáme totiž, že určitá část pojistného plnění může být vyplacena již v samotném roce původu (jedná se o hodnoty v prvním sloupci nekumulativního vývojového trojúhelníku). V tom případě ale od vzniku pojistné události neuběhl ani rok (resp. jsme stále ve stejném kalendářním roce), a proto $j = 0$.

Poznámka 4.1. Od indexů roku původu je možné snadno přejít ke skutečným letopočtům. Pokud roku vývoje $i = 1$ odpovídá skutečný kalendářní rok K_0 ,

pro každý další rok původu lze příslušný kalendářní rok získat ze vztahu $K_i = K_0 + (i - 1)$. K_0 tak lze pokládat za referenční rok pro dané trojúhelníkové schéma.

Při bližším pohledu na nekumulativní trojúhelník si lze povšimnout, že prvky na jednotlivých diagonálách mají stejný součet indexů i a j , tzn. $i + j = k$. Pro tyto prvky platí, že byly vyplaceny ve stejném kalendářním roce. Součet těchto diagonálních prvků proto udává kumulativní pojistné plnění vyplacené do kalendářního roku $K_k = K_0 + (k - 1)$ včetně.

Dále se předpokládá, že poslední rok původu, ze kterého jsou data o pojistných plněních k dispozici, bude označen indexem n . Z tohoto důvodu se bude nadále pracovat s tabulkou/maticí o rozměrech $n \times n$, kde se rok původu i bude pohybovat od 1 do n a rok vývoje j od 0 do $n - 1$. „Levý horní roh“ této tabulky bude tvořen historickými daty (v podobě vývojového trojúhelníku), zatímco zbytek tabulky budou data odhadovaná.

4.2. Metoda Chain-Ladder

Tato kapitola se bude zabývat v praxi nejpoužívanější metodou pro stanovení rezerv na pojistná plnění v neživotním pojištění využívající trojúhelníková schémata, a to *metodou Chain-Ladder*. Teoretické základy této metody jsou blíže rozepsány např. v [5, 28]. Model metody Chain-Ladder stojí na třech základních předpokladech:

1. kumulativní pojistná plnění $C_{i,j}$ jsou pro různé roky původu i nezávislá;
2. poměr kumulativních pojistných plnění mezi sousedními roky vývoje zůstává pro jednotlivé roky původu přibližně stejný, tzn.

$$\frac{C_{i_1,j+1}}{C_{i_1,j}} \approx \frac{C_{i_2,j+1}}{C_{i_2,j}}, \quad j = 0, \dots, n - 2; \quad (4.3)$$

3. existují taková nezáporná čísla $\lambda_0, \lambda_1, \dots, \lambda_{n-2}$, pro která platí:

$$E[C_{i,j+1} | C_{i,0}, \dots, C_{i,j}] = E[C_{i,j+1} | C_{i,j}] = \lambda_j \cdot C_{i,j}, \quad (4.4)$$

$$i = 1, \dots, n, \quad j = 0, \dots, n - 2.$$

Praktický postup výpočtů provedených metodou Chain-Ladder lze potom shrnout do následujících kroků:

1. určit odhady tzv. *vývojových koeficientů* $\hat{\lambda}_0, \hat{\lambda}_1, \dots, \hat{\lambda}_{n-2}$, pro které bude platit vztah (4.4);
2. odhadnout budoucí kumulativní pojistná plnění $\hat{C}_{i,j}$ pro $i + j > n$;
3. spočítat celkový odhad rezervy, tj. \hat{R}_{ChL} .

Pro odhad vývojových koeficientů využijeme předpokladu (4.3) o podobnosti poměru kumulativních pojistných plnění mezi sousedními roky vývoje. Z tohoto důvodu obecně stačí odhadnout jeden vývojový koeficient pro každý index j . Klasicky se určuje jako podíl součtu pojistných plnění ve dvou po sobě jdoucích letech vývoje, tedy

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j-1} C_{i,j+1}}{\sum_{i=1}^{n-j-1} C_{i,j}}, \quad j = 0, \dots, n-2. \quad (4.5)$$

Ke koeficientům $\hat{\lambda}_j$ lze se stejným výsledkem dojít i jinou úvahou. Z dostupných historických dat jsme schopni určit analogické koeficienty k vývojovým koeficientům, tzv. *individuální vývojové koeficienty* $\lambda_{i,j}$. Ty ovšem neurčují střední hodnotu z rovnosti (4.4), ale skutečnou realizaci pojistného plnění v daném, již uplynulém, roce. Jednoduše lze dojít ke vztahu, který pro tyto individuální vývojové koeficienty platí, a to:

$$\lambda_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}, \quad i = 1, \dots, n-1, \quad j = 0, \dots, n-i-1.$$

Tyto koeficienty sice přesně popisují stavy minulé, pro predikce budoucích hodnot ale samy o sobě vhodné nejsou. Platí totiž pouze pro danou kombinaci roku původu i a roku vývoje j . Aby se tato vlastnost oslabila a jejich použití mohlo být obecnější, je vhodné koeficienty zprůměrovat. Místo klasického aritmetického průměru se použije průměr vážený, a to s vahami ve tvaru $\frac{C_{i,j}}{\sum_{k=1}^{n-j-1} C_{k,j}}$. Ty udávají,

jaká část prostředků vyplacených do roku j byla použita na hrazení pojistných událostí z roku i . Lze snadno ověřit, že součet těchto vah přes všechny indexy $i = 1, \dots, n-j-1$ je roven jedné. Po dosazení hodnot a příslušných aritmetických úpravách dostaneme stejný vztah jako v rovnici (4.5):

$$\begin{aligned} \sum_{i=1}^{n-j-1} \frac{C_{i,j}}{\sum_{k=1}^{n-j-1} C_{k,j}} \cdot \lambda_{i,j} &= \sum_{i=1}^{n-j-1} \frac{C_{i,j}}{\sum_{k=1}^{n-j-1} C_{k,j}} \cdot \frac{C_{i,j+1}}{C_{i,j}} = \\ &= \frac{\sum_{i=1}^{n-j-1} C_{i,j+1}}{\sum_{i=1}^{n-j-1} C_{i,j}} = \lambda_j, \quad j = 0, \dots, n-2. \end{aligned}$$

Poznámka 4.2. Alternativně lze za odhad vývojových koeficientů brát aritmetický průměr jednotlivých podílů kumulativních pojistných plnění, tj.

$$\hat{\lambda}_j = \frac{1}{n-j-1} \sum_{i=1}^{n-j-1} \frac{C_{i,j+1}}{C_{i,j}}, \quad j = 0, \dots, n-2.$$

Další postup metody je v obou případech totožný.

Je potřeba si uvědomit, že pro rok vývoje $n-1$ už nejsme s naší znalostí dat schopni určit vývojový koeficient $\hat{\lambda}_{n-1}$. To ovšem není na škodu, protože by byl využit pro odhad hodnot $\hat{C}_{i,n}$, které ale v tomto modelu neuvažujeme (nejvyšší uvažovaný index j je roven $n-1$). Dále je vidět, že s vyšším rokem vývoje j do vzorce vstupuje menší počet hodnot - pokud by tyto hodnoty byly nějakým způsobem extrémní nebo vychýlené, může dojít k zavádějícím výsledkům.

Vraťme se k ilustrativnímu kumulativnímu trojúhelníku 4.2. Pro příklad uvedeme podrobně výpočet vývojového koeficientu $\hat{\lambda}_1$. Ten získáme dosazením do vztahu (4.5) (ostatní koeficienty dostaneme analogicky):

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^4 C_{i,2}}{\sum_{i=1}^4 C_{i,1}} = \frac{31 + 36 + 44 + 42}{21 + 28 + 29 + 24} = \frac{153}{102} = 1,5.$$

Všechny dopočítané vývojové koeficienty jsou uvedeny v tabulce 4.3. S pomocí dopočítaných vývojových koeficientů je již možné doplnit kumulativní vývojový

j	$\hat{\lambda}_j$
0	2,00
1	1,50
2	1,33
3	1,25
4	1

Tabulka 4.3: Vývojové koeficienty $\hat{\lambda}_j$.

trojúhelník odhady budoucích pojistných plnění. Pro prázdná pole tabulky totiž platí

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

resp.

$$\hat{C}_{i,j} = \hat{\lambda}_{j-1} \cdot \hat{C}_{i,j-1} \quad i + j > n + 1.$$

Vyčíslení kumulativních pojistných plnění v celé tabulce je pro představu jistě užitečné, pro výpočet rezervy ale není nezbytné - do výpočtu rezervy vstupuje pouze poslední sloupec tabulky, tj. hodnoty $\hat{C}_{i,n-1}$. Pokud tedy nepotřebujeme doplnit celou tabulku, ale zajímají nás jen celková očekávaná pojistná plnění, lze si určit tzv. *kumulativní vývojové koeficienty* $\hat{\lambda}_j^c$ vztahem

$$\hat{\lambda}_j^c = \prod_{k=j}^{n-2} \hat{\lambda}_k, \quad j = 0, \dots, n-2,$$

a potom jen dopočítat odhady $\hat{C}_{i,n-1}$ následujícím způsobem:

$$\hat{C}_{i,n-1} = C_{i,n-i} \cdot \hat{\lambda}_{n-i}^c, \quad i = 2, \dots, n.^6$$

Kompletní data doplněná o dopočítané hodnoty lze vidět v tabulce 4.4.

Nyní lze určit hledanou výši rezervy na pojistná plnění. Ta je dána jako rozdíl očekávaných nákladů na pojistná plnění a již vyplacených pojistných plnění. Od součtu hodnot ve sloupci $j = n - 1$, tedy od očekávaného celkového pojistného plnění ze všech roků původu k poslednímu uvažovanému roku vývoje, se odečtou

⁶Hodnotu $\hat{C}_{1,n-1}$ není potřeba odhadovat, protože je k dispozici skutečné kumulativní pojistné plnění $C_{1,n-1}$.

prvky $C_{i,j}$, kde $i + j = n$, což jsou dosud již vyplacená pojistná plnění za události z jednotlivých let původu. Rezerva určená metodou Chain-Ladder se tedy určí vztahem

$$\hat{R}_{ChL} = \sum_{i=1}^n \hat{C}_{i,n-1} - \sum_{i=1}^n C_{i,n-i}, \quad (4.6)$$

pro použitá ilustrační data to bude

$$\hat{R}_{ChL} = (50+60+75+70+65+55) - (50+60+60+42+26+11) = 375 - 249 = 126,$$

tedy 126 mil. Kč.

$i \backslash j$	0	1	2	3	4	5
1	9	21	31	42	50	50
2	13	28	36	46	60	60
3	14	29	44	60	75	75
4	16	24	42	56	70	70
5	12	26	39	52	65	65
6	11	22	33	44	55	55

Tabulka 4.4: Kumulativní trojúhelník z tabulky 4.2 doplněný metodou Chain-Ladder

Jak je z výpočtu vidět, výsledkem metody Chain-Ladder je „pouze“ číslo, tedy bodový odhad rezervy. Nabízí se otázka, zda je takováto predikce dostatečná. Jedním ze způsobů, jak tento problém řešit, je použití tzv. Mackovy varianty Chain-Ladder metody. V rámci této metody se pro každý odhad vývojového koeficientu $\hat{\lambda}_j$ počítá i jeho rozptyl - ten lze potom zahrnout i do výpočtu rezervy. Výsledkem této metody je tedy kromě bodového odhadu také odhad rozptylu celkové rezervy. Tato varianta je blíže přiblížena např. v [28].

Doposud ovšem v rámci metody nebyl brán v úvahu jeden důležitý ukazatel, a tím je lidský faktor. Pojistní matematici, kteří se této problematice dlouhodobě věnují, mohou mít ze svých zkušeností vlastní způsoby, jak upravit vývojové faktory v průběhu výpočtu. V praxi navíc mohou vstupovat do hry další faktory, např. rezervotvorná politika jednotlivých pojišťoven, kvalita dat, které mají pojistní matematici k dispozici apod. V následující kapitole tak bude představen

jeden z postupů, jak zahrnout do výpočtu nejistotu související s kvalitou dostupných dat (a tím ovlivněnou „důvěru“ v tato data), a to převedení problému do fuzzy roviny.

4.3. Fuzzifikace metody Chain-Ladder

Tato sekce bude věnována možnostem fuzzifikace metody Chain-Ladder. Tomuto přístupu zatím zjevně nebylo věnováno mnoho pozornosti - podařilo se dohledat pouze článek [11]. V následující části bude představena zde uvedená metodologie a diskutovány její silné a slabé stránky. Dále budou představeny původní alternativní metody, jak metodu fuzzifikovat, spolu se svými výhodami (či nevýhodami). Tyto metody budou představeny na stejných ilustrativních datech, jako klasická metoda Chain-Ladder.

Poznámka 4.3. Pokud nebude uvedeno jinak, v této části se bude používat LR zápis trojúhelníkového fuzzy čísla ve tvaru $A = (c_A, l_A, r_A)_{LR}$.

Dále bude v této kapitole platit, že i přesto, že fuzzy vývojové koeficienty jsou stále „jen“ odhady, bude se u nich vynechávat stříška $\hat{}$ - tou budou označeny jen význačné hodnoty těchto fuzzy koeficientů.

4.3.1. Fuzzy Chain-Ladder

Již bylo zmíněno v úvodu kapitoly, že tato metoda je blíže popsána v [11]. Jak už název napovídá, metoda *fuzzy Chain-Ladder* vychází z klasického (ostrého) modelu metody Chain-Ladder. Pro obě verze metody jsou nutná stejná vstupní data, tedy kumulativní, resp. nekumulativní vývojový trojúhelník zobrazující vyplacená pojistná plnění za uplynulá účetní období. Princip obou metod je také velmi podobný. I pro výpočet rezerv metodou fuzzy Chain-Ladder je třeba stanovit vývojové koeficienty, které budou udávat, jak se objem vyplaceného pojistného plnění mění s plynoucím časem. Tyto koeficienty se následně použijí pro odhady budoucích pojistných plnění a z nich již lze snadno odhadnout očekávanou výši rezervy.

Předpoklad nezávislosti pojistných plnění v různých letech původu zůstává i ve fuzzifikované metodě, přidávají se ale další předpoklady:

1. nekumulativní pojistná plnění $X_{i,j} = C_{i,j+1} - C_{i,j}$, jsou nezáporná;⁷
2. existují taková trojúhelníková fuzzy čísla $\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{n-2}$, pro která bude platit

$$\tilde{C}_{i,j+1} = \tilde{f}_j \cdot \tilde{C}_{i,j}, \quad i = 1, \dots, n, \quad j = 0, \dots, n-2, \quad (4.7)$$

kde

$$\tilde{C}_{i,j} = (C_{i,j}, l_{\tilde{C}_{i,j}}, r_{\tilde{C}_{i,j}})_{LR}, \quad \tilde{f}_j = (f_j, l_{\tilde{f}_j}, r_{\tilde{f}_j})_{LR}.$$

Poznámka 4.4. Ve vztahu (4.7) se objevuje kumulativní pojistné plnění v podobě fuzzy čísla, což může být matoucí. Ačkoli budoucí hodnoty skutečně budou odhadovány jako trojúhelníkové fuzzy číslo, původní kumulativní a nekumulativní trojúhelník zůstane ve své počáteční „ostré“ podobě. V zápise se pro sjednocení ale pracuje se zápisem fuzzy čísel, protože každé reálné číslo x lze zapsat jako trojúhelníkové fuzzy číslo ve tvaru $\tilde{x} = (x, 0, 0)_{LR}$. Vztah (4.7) proto obecně platí pro minulá i budoucí pozorování.

Právě vývojové koeficienty jsou tedy v metodě faktorem, který zajišťuje jakési rozostření výsledku (tj. odhadované rezervy): místo ostrých vývojových koeficientů λ_j pro každý rok vývoje se tak bude pracovat s koeficienty v podobě fuzzy čísel. V rámci modelu se předpokládá, že tyto odhady těchto koeficientů budou trojúhelníková fuzzy čísla $\tilde{f}_j = (\hat{f}_j, \hat{l}_{\tilde{f}_j}, \hat{r}_{\tilde{f}_j})_{LR}$, která zohlední vágnost, resp. neurčitost, se kterou pojistný matematik pracuje. Trojúhelníková fuzzy čísla jsou zároveň snadno interpretovatelná a dobře se s nimi pracuje v rámci aritmetických operací (což bylo přiblíženo v kapitole 1).

Autoři citovaného článku přišli s postupem, jak odhadnout hodnoty $\hat{f}_j, \hat{l}_{\tilde{f}_j}, \hat{r}_{\tilde{f}_j}$ těchto fuzzy vývojových koeficientů. Střed, resp. jádro trojúhelníkového fuzzy čísla se stanoví analogicky jako v klasické metodě Chain-Ladder, tedy:

$$\hat{f}_j (= \hat{\lambda}_j) = \frac{\sum_{i=1}^{n-j-1} C_{i,j+1}}{\sum_{i=1}^{n-j-1} C_{i,j}}, \quad j = 0, \dots, n-2, \quad (4.8)$$

⁷Tento předpoklad je zcela přirozený: logicky očekáváme, že celkové pojistné plnění bude v průběhu času růst (příp. zůstane stejné), ale nebude se snižovat.

aby zůstal splněn vztah (4.5). Otázkou zůstává, jak určit levé a pravé rozpětí fuzzy čísel. Zmíněný model určuje obě hranice stejně, a to vztahem

$$\hat{l}_{\tilde{f}_j} = \hat{r}_{\tilde{f}_j} = \frac{\sum_{i=1}^{n-j-1} X_{i,j+1}}{\sum_{i=1}^{n-j-1} C_{i,j}}, \quad j = 0, \dots, n-2. \quad (4.9)$$

Úvaha pro tuto volbu je následující: vzhledem k tomu, že pro hodnoty z nekumulativního trojúhelníku platí

$$X_{i,j+1} = C_{i,j+1} - C_{i,j}, \quad i = 1, \dots, n, \quad j = 0, \dots, n-2,$$

dostaneme ze vztahu (4.9) další vyjádření rozpětí pro $\hat{l}_{\tilde{f}_j}$ (stejně i pro $\hat{r}_{\tilde{f}_j}$), a to:

$$\begin{aligned} \hat{l}_{\tilde{f}_j} &= \frac{\sum_{i=1}^{n-j-1} X_{i,j+1}}{\sum_{i=1}^{n-j-1} C_{i,j}} = \frac{\sum_{i=1}^{n-j-1} (C_{i,j+1} - C_{i,j})}{\sum_{i=1}^{n-j-1} C_{i,j}} = \frac{\sum_{i=1}^{n-j-1} C_{i,j+1} - \sum_{i=1}^{n-j-1} C_{i,j}}{\sum_{i=1}^{n-j-1} C_{i,j}} \\ &= \frac{\sum_{i=1}^{n-j-1} C_{i,j+1}}{\sum_{i=1}^{n-j-1} C_{i,j}} - 1 = \hat{f}_j - 1, \end{aligned}$$

z čehož potom logicky vyplývá

$$\hat{f}_j - \hat{l}_{\tilde{f}_j} = 1. \quad (4.10)$$

Z předchozího vztahu lze vidět, že jako nejmenší prvek nosiče fuzzy čísla \tilde{f}_j dostáváme číslo 1, a to z důvodu, že v předpokladech modelu požadujeme nezápornost hodnot $X_{i,j}$ (není proto možné, aby kumulativní pojistné plnění v jednom roce vývoje bylo menší než hodnota v roce předchozím). Volba $\hat{l}_{\tilde{f}_j}$ ze vztahu (4.9) proto umožňuje uvažovat maximální možnou, ale zároveň reálnou, neurčitost fuzzy vývojového koeficientu \tilde{f}_j směrem vlevo. Pro jednoduchost je pak stejný vztah aplikován i pro rozpětí směrem vpravo, tedy pro hodnotu $\hat{r}_{\tilde{f}_j}$.

Pro představu opět uvedeme výpočet na příkladě, a to na stejných datech jako u metody Chain Ladder. Pro fuzzy vývojový koeficient $\tilde{f}_1 = (\hat{f}_1, \hat{l}_{\tilde{f}_1}, \hat{r}_{\tilde{f}_1})_{LR}$ nyní bude platit:

$$\hat{f}_1 = \frac{\sum_{i=1}^4 C_{i,2}}{\sum_{i=1}^4 C_{i,1}} = \frac{31 + 36 + 44 + 42}{21 + 28 + 29 + 24} = \frac{153}{102} = 1,5,$$

$$\hat{l}_{\tilde{f}_1} = \hat{r}_{\tilde{f}_1} = \frac{\sum_{i=1}^4 X_{i,2}}{\sum_{i=1}^4 C_{i,1}} = \frac{10 + 8 + 15 + 18}{21 + 28 + 29 + 24} = \frac{51}{102} = 0,5,$$

tj. $\tilde{f}_1 = (1,5, 0,5, 0,5)_{LR}$. Všechny fuzzy vývojové koeficienty je pak možné vidět v tabulce 4.5.

j	\tilde{f}_j
0	$(2, 00, 1, 00, 1, 00)_{LR}$
1	$(1, 50, 0, 50, 0, 50)_{LR}$
2	$(1, 33, 0, 33, 0, 33)_{LR}$
3	$(1, 25, 0, 25, 0, 25)_{LR}$
4	$(1, 00, 0, 00, 0, 00)_{LR}$

Tabulka 4.5: Zaokrouhlené fuzzy vývojové koeficienty pro schéma 4.2.

V souladu s klasickou metodou se pak budoucí pojistná plnění určit obdobným způsobem. I zde tedy lze dopočítat celou tabulku vztahem

$$\tilde{C}_{i,j+1} = \tilde{f}_j \cdot \tilde{C}_{i,j} \quad i + j > n,$$

nebo využít fuzzy analogii kumulovaných vývojových koeficientů a určit si jen poslední sloupec tabulky (kromě hodnoty $C_{1,n-1}$, která je známá), tedy

$$\tilde{C}_{i,n-1} = C_{i,n-i} \cdot \prod_{k=n-i}^{n-2} \tilde{f}_k, \quad i = 2, \dots, n. \quad (4.11)$$

Pro usnadnění výpočtu fuzzy kumulovaných vývojových koeficientů ze vztahu (4.11) lze využít následující věty (dle [11]):

Věta 4.1. Pro součin $\prod_{k=n-i}^{n-2} \tilde{f}_k$ ze vztahu (4.11), kde jsou význačné hodnoty jednotlivých fuzzy vývojových koeficientů spočítány dle vztahů (4.8), resp. (4.9), platí:

$$\prod_{k=n-i}^{n-2} \tilde{f}_k \approx \left(\hat{F}_i, \hat{l}_{\hat{F}_i}, \hat{r}_{\hat{F}_i} \right)_{LR} \quad \forall i \in \{2, \dots, n\},$$

kde

$$\hat{F}_i = \prod_{k=n-i}^{n-2} \hat{f}_k,$$

$$\hat{l}_{\hat{F}_i} = \hat{F}_i - 1,$$

$$\hat{r}_{\hat{F}_i} = \prod_{k=n-i}^{n-2} (2\hat{f}_k - 1) - \hat{F}_i.$$

Důkaz. V kapitole 1 bylo uvedeno, že trojúhelníkové fuzzy číslo v LR zápise lze převést na „klasický“ zápis. Pro fuzzy vývojové koeficienty \tilde{f}_j proto bude platit:

$$\tilde{f}_j = \left(\hat{f}_j, \hat{l}_{\tilde{f}_j}, \hat{r}_{\tilde{f}_j} \right)_{LR} = \left(\hat{f}_j - \hat{l}_{\tilde{f}_j}, \hat{f}_j, \hat{f}_j + \hat{r}_{\tilde{f}_j} \right).$$

Díky vztahu (4.10) a rovnosti $l_{\tilde{f}_j}$ a $r_{\tilde{f}_j}$ lze tento zápis ještě upravit:

$$\left(\hat{f}_j - \hat{l}_{\tilde{f}_j}, \hat{f}_j, \hat{f}_j + \hat{r}_{\tilde{f}_j} \right) = \left(1, \hat{f}_j, 2\hat{f}_j - 1 \right).$$

Odtud už ze vzorce pro násobení trojúhelníkových fuzzy čísel uvedeného v poznámce 1.4 vyplývá, že

$$\prod_{k=n-i}^{n-2} \left(1, \hat{f}_k, 2\hat{f}_k - 1 \right) \approx \left(1, \prod_{k=n-i}^{n-2} \hat{f}_k, \prod_{k=n-i}^{n-2} (2\hat{f}_k - 1) \right).$$

Jestliže tento vztah převedeme zpět do LR zápisu, dostaneme

$$\left(1, \prod_{k=n-i}^{n-2} \hat{f}_k, \prod_{k=n-i}^{n-2} (2\hat{f}_k - 1) \right) = \left(\prod_{j=n-i}^{n-2} \hat{f}_k, \prod_{k=n-i}^{n-2} \hat{f}_k - 1, \prod_{k=n-i}^{n-2} (2\hat{f}_k - 1) - \prod_{k=n-i}^{n-2} \hat{f}_k \right)_{LR},$$

což po přeznačení $\prod_{k=n-i}^{n-2} \hat{f}_k = \hat{F}_i$ dává

$$\left(\hat{F}_i, \hat{F}_i - 1, \prod_{k=n-i}^{n-2} (2\hat{f}_k - 1) - \hat{F}_i \right)_{LR}.$$

□

Výše uvedenými postupy lze dopočítat fuzzy odhady budoucích pojistných plnění tak, jak je vidět v tabulce 4.6. Celková výše odhadnuté rezervy je pak opět

i	$\tilde{C}_{i,0}$	$\tilde{C}_{i,1}$	$\tilde{C}_{i,2}$
1	$(9, 0, 0)_{LR}$	$(21, 0, 0)_{LR}$	$(31, 0, 0)_{LR}$
2	$(13, 0, 0)_{LR}$	$(28, 0, 0)_{LR}$	$(36, 0, 0)_{LR}$
3	$(14, 0, 0)_{LR}$	$(29, 0, 0)_{LR}$	$(44, 0, 0)_{LR}$
4	$(16, 0, 0)_{LR}$	$(24, 0, 0)_{LR}$	$(42, 0, 0)_{LR}$
5	$(12, 0, 0)_{LR}$	$(26, 0, 0)_{LR}$	$(\mathbf{39}, \mathbf{13}, \mathbf{13})_{LR}$
6	$(11, 0, 0)_{LR}$	$(\mathbf{22}, \mathbf{11}, \mathbf{11})_{LR}$	$(\mathbf{33}, \mathbf{22}, \mathbf{33})_{LR}$

i	$\tilde{C}_{i,3}$	$\tilde{C}_{i,4}$	$\tilde{C}_{i,5}$
1	$(42, 0, 0)_{LR}$	$(50, 0, 0)_{LR}$	$(50, 0, 0)_{LR}$
2	$(46, 0, 0)_{LR}$	$(60, 0, 0)_{LR}$	$(\mathbf{60}, \mathbf{0}, \mathbf{0})_{LR}$
3	$(60, 0, 0)_{LR}$	$(\mathbf{75}, \mathbf{15}, \mathbf{15})_{LR}$	$(\mathbf{75}, \mathbf{15}, \mathbf{15})_{LR}$
4	$(\mathbf{56}, \mathbf{14}, \mathbf{14})_{LR}$	$(\mathbf{70}, \mathbf{28}, \mathbf{35})_{LR}$	$(\mathbf{70}, \mathbf{28}, \mathbf{35})_{LR}$
5	$(\mathbf{52}, \mathbf{26}, \mathbf{35})_{LR}$	$(\mathbf{65}, \mathbf{39}, \mathbf{65})_{LR}$	$(\mathbf{65}, \mathbf{39}, \mathbf{65})_{LR}$
6	$(\mathbf{44}, \mathbf{33}, \mathbf{66})_{LR}$	$(\mathbf{55}, \mathbf{44}, \mathbf{110})_{LR}$	$(\mathbf{55}, \mathbf{44}, \mathbf{110})_{LR}$

Tabulka 4.6: Schéma 4.2 doplněné metodou fuzzy Chain-Ladder.

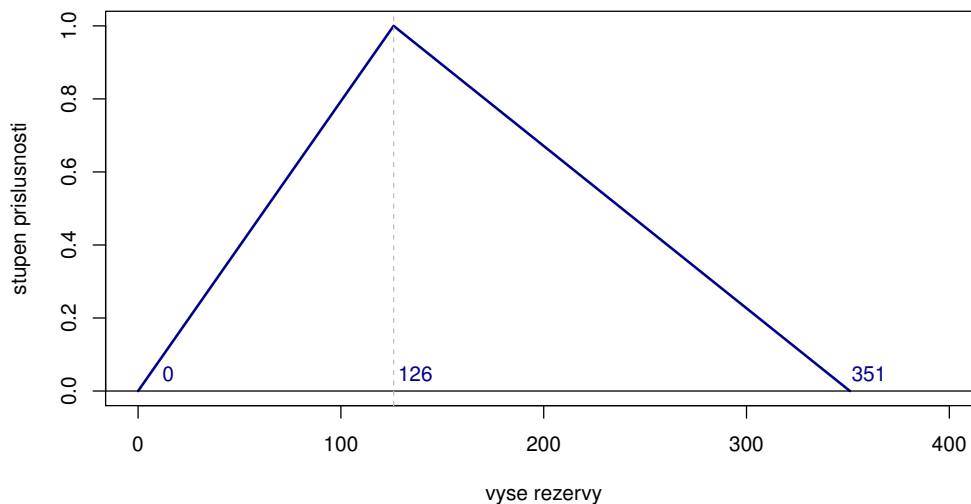
určena odhadem celkových pojistných plnění a dosud vyplacených pojistných plnění, nyní už ovšem s fuzzifikovanými hodnotami, tj.

$$\tilde{R}_{ChL} = \sum_{i=1}^n \tilde{C}_{i,n-1} - \sum_{i=1}^n C_{i,n-i} = \sum_{i=1}^n (\tilde{C}_{i,n-1} - C_{i,n-i}).$$

Pro ilustrační data bude potom hodnota rezervy následující (grafické znázornění rezervy lze vidět na obrázku 4.1):

$$\tilde{R}_{ChL} = \sum_{i=1}^6 \tilde{C}_{i,5} - \sum_{i=1}^6 C_{i,6-i} = (375, 126, 225)_{LR} - (249, 0, 0)_{LR} = (126, 126, 225)_{LR}.$$

Vidíme, že jádro rezervy spočítané metodou fuzzy Chain-Ladder je totožné s hodnotou získanou pomocí klasické metody (což je při stejném výpočtu koeficientů \hat{f}_j a $\hat{\lambda}_j$ a stávajícím postupu výpočtu rezervy logické). Důsledkem fuzzifikace vývojových koeficientů je výsledek poměrně dosti neurčitý, což samo o sobě nemusí být na škodu - neurčitost zachycuje jak postoj autora ke kvalitě vstupních dat, tak nejistotu v souvislosti s budoucím vývojem pojistných plnění. Ačkoli se tedy pojišťovna v praxi musí rozhodnout pro nějakou pevnou/ostrou hodnotu



Obrázek 4.1: Odhad fuzzy rezervy pro schéma 4.2.

rezerv, kterou bude tvořit, vzhled fuzzy čísla představujícího rezervu může sloužit jako jednoduchý grafický nástroj, díky kterému lze do stanovení výše tvořené rezervy zahrnout i informaci o naší nejistotě.

Pozitivně lze hodnotit i skutečnost, že při výpočtu nemůžeme dostat zápornou rezervu (tj. do nosiče výsledného fuzzy čísla nebude patřit žádná záporná hodnota) - opět to vychází z předpokladu nezáporných hodnot $X_{i,j}$. Naopak ve směru vpravo od jádra má rezerva mnohem větší neurčitost, protože explicitně nepředpokládáme žádné omezení. I z praktického hlediska má neurčitost směrem vpravo význam - při vhodně zvolené defuzzifikační charakteristice pak dostaneme odhad rezervy vyšší, než hodnota z jádra. Pojišťovna by proto měla být lépe připravená na případné negativní odchylky od odhadovaného stavu.

Co se týče požadavků na teoretické znalosti při aplikaci této metody, vzhledem k využití trojúhelníkových fuzzy čísel, pro která lze potřebnou fuzzy aritmetiku shrnout do několika vzorců, není tato metodologie náročná na použití. Postup metody jinak zůstává i ve fuzzifikované podobě téměř totožný. Doba výpočtu může mírně vzrůst vzhledem k nutnosti dopočítávat 2 hodnoty navíc pro každý

fuzzy koeficient ($\hat{l}_{\tilde{f}_j}$ a $\hat{r}_{\tilde{f}_j}$) - na druhou stranu trojúhelníková schémata, se kterými se pracuje, nebývají natolik rozsáhlá, aby tato skutečnost působila přílišné nároky na čas či paměť.

Za slabší stránky této metody lze pokládat problematiku interpretaci pro rozpětí $\hat{l}_{\tilde{f}_j}$, resp. $\hat{r}_{\tilde{f}_j}$. Zatímco levá strana $\hat{l}_{\tilde{f}_j}$ má své opodstatnění alespoň ve výsledném zajištění nezáporné rezervy, pro pravou stranu žádné podobné vysvětlení neexistuje (sami autoři volbu $\hat{r}_{\tilde{f}_j}$ shodně s $\hat{l}_{\tilde{f}_j}$ motivují pouze zjednodušením modelu).

Při bližším pohledu na vztah původního trojúhelníkového schématu a odhadu fuzzy rezervy vyvstává další otázka: jak moc je reálné, aby nastala situace, že skutečné potřebné rezervy pro budoucí pojistná plnění budou téměř nulové? Z tvaru fuzzy čísla je vidět, že kladná hodnota libovolně blízká nule patří do nosiče odhadu fuzzy rezervy, tedy její stupeň příslušnosti je nenulový. To je sice teoreticky správné, z trojúhelníkového schématu lze ale očekávat, že zejména u pojistných událostí z vyšším rokem původu i , tj. s i blízkým n , k nějakému dodatečnému pojistnému plnění v následujících letech (resp. do roku $i + (n - 1)$, který nejvyšší ve schématu pro dané i uvažujeme), téměř jistě dojde. Možnost, že by pojišťovna vyplácela na pojistném plnění ony velmi nízké sumy, které ale spadají do nosiče fuzzy rezervy, je velice nízká až mizivá. Je proto správné, že jsou takové možnosti ve výsledném odhadu rezervy vůbec uvažované?

Následující text by měl přinést několik alternativních možností, jak metodu Chain-Ladder fuzzifikovat, a to se zmíněnými negativy předchozí metody na paměti. Nejprve bude pozornost zaměřena na jiné možnosti odhadu rozpětí $\hat{l}_{\tilde{f}_j}$ a $\hat{r}_{\tilde{f}_j}$, později bude alternativně počítáno i jádro fuzzy vývojového koeficientu \hat{f}_j .

4.3.2. Fuzzy Chain-Ladder s využitím minimálních a maximálních podílů $C_{i,j}$

Jak už bylo popsáno v předchozí části, jedním ze zásadních nedostatků metody fuzzy Chain-Ladder dle [11] je špatná interpretovatelnost hodnot $\hat{l}_{\tilde{f}_j}$ a $\hat{r}_{\tilde{f}_j}$. Zde se chceme věnovat právě tomuto nedostatku.

Hodnota \hat{f}_j (ze vztahu (4.8)) je dána jako podíl součtů kumulativních pojistných plnění ve dvou po sobě jdoucích letech vývoje, tj. v letech j a $j + 1$. Tento vztah je účinný pro potlačení nějakým způsobem vychýlených hodnot (příliš vysokých, nízkých). Nesmíme ale zapomínat na to, že všechny tyto hodnoty odpovídají skutečnosti - pro konkrétní rok původu i víme, že hodnoty $C_{i,j}$ a $C_{i,j+1}$ odpovídají scénářům, které v minulosti reálně nastaly a pojišťovna odpovídající pojistná plnění $X_{i,j}$ a $X_{i,j+1}$ opravdu vyplatila. Pokud se podíváme na interpretaci podílu $\frac{C_{i,j+1}}{C_{i,j}}$, jedná se o hodnotu vyjadřující, kolikrát se zvýšila hodnota kumulativního pojistného plnění oproti předchozímu roku vývoje.⁸ To znamená, že jsme pro konkrétní dvojici po sobě jdoucích sloupců j a $j + 1$ schopni určit takový (v minulosti proběhlý) scénář, který odpovídá situaci pro pojišťovnu nejhorší (hodnota $C_{i,j+1}$ se oproti hodnotě $C_{i,j}$ navýšila nejvíce), ale také nejlepší (hodnota $C_{i,j+1}$ se oproti hodnotě $C_{i,j}$ navýšila nejméně).

Nejlepší možný scénář pro pojišťovnu odpovídá situaci, kdy výše pojistného plnění, které by měla poskytnout, bude minimální, tzn. hodnota $C_{i,j+1}$ se oproti $C_{i,j}$ zvýší co nejméně. Tuto „imaginární“ hodnotu skutečného, ale stále možného minima neznáme, za její odhad ale lze vzít právě minimální hodnotu $\frac{C_{i,j+1}}{C_{i,j}}$ pro danou dvojici let vývoje - tato hodnota je pro nás minimem ze situací, o kterých víme, že skutečně nastaly. Obdobnou úvahu lze použít pro nejhorší možný scénář pro pojišťovnu, která vyplyne v maximální hodnotu $\frac{C_{i,j+1}}{C_{i,j}}$.

Pokud tuto myšlenku rozvedeme pro všechny podíly $\frac{C_{i,j+1}}{C_{i,j}}$ ze všech let vývoje, pro které máme informace k dispozici, můžeme získat „celkový nejlepší“ a „celkový nejhorší“ jistě dosažitelný scénář. Nejlepší scénář bude zahrnovat celkové minimální navyšování pojistného plnění v průběhu času, nejhorší naopak navyšování největší. A právě tyto dva scénáře nám poskytnou odhady pro rozpětí celkového odhadu fuzzy rezervy.

Vraťme se nyní k výpočtu jednotlivých fuzzy koeficientů. Vzhledem k charakteru těchto odhadů se přesuneme od LR zápisu fuzzy čísel ke „klasickému“

⁸I nadále zde pracujeme s předpokladem, že nekumulativní pojistná plnění $X_{i,j}$ jsou nezáporná, a tedy $\frac{C_{i,j+1}}{C_{i,j}} \geq 1$.

tvary fuzzy čísla, tj. fuzzy vývojový koeficient bude tvaru $\tilde{f}_j = (\hat{a}_{\tilde{f}_j}, \hat{c}_{\tilde{f}_j}, \hat{b}_{\tilde{f}_j})$. Pro konkrétní rok vývoje j zůstane jádro trojúhelníkového fuzzy koeficientu stejné, jako u klasické metody Chain-Ladder:

$$\hat{c}_{\tilde{f}_j} = \frac{\sum_{i=1}^{n-j-1} C_{i,j+1}}{\sum_{i=1}^{n-j-1} C_{i,j}}, \quad j = 0, \dots, n-2.$$

Příslušný levý okraj nosiče $\hat{a}_{\tilde{f}_j}$ bude dán nejmenším podílem kumulativních pojistných plnění odpovídající skutečnému historickému scénáři, tj.

$$\hat{a}_{\tilde{f}_j} = \min \left\{ \frac{C_{i,j+1}}{C_{i,j}}, i = 1, \dots, n-j-1 \right\}, \quad j = 0, \dots, n-2, \quad (4.12)$$

zatímco pravý okraj nosiče $b_{\tilde{f}_j}$ se stanoví podílem nejvyšším:

$$\hat{b}_{\tilde{f}_j} = \max \left\{ \frac{C_{i,j+1}}{C_{i,j}}, i = 1, \dots, n-j-1 \right\}, \quad j = 0, \dots, n-2. \quad (4.13)$$

Z ilustračního schématu 4.2 a vztahů (4.8), (4.12) a (4.13) dostáváme opět pro $\tilde{f}_1 = (\hat{a}_{\tilde{f}_1}, \hat{c}_{\tilde{f}_1}, \hat{b}_{\tilde{f}_1})$:

$$\begin{aligned} \hat{c}_{\tilde{f}_1} &= \frac{31 + 36 + 44 + 42}{21 + 28 + 29 + 24} = \frac{153}{102} = 1,5, \\ \hat{a}_{\tilde{f}_1} &= \min \left\{ \frac{31}{21}, \frac{36}{28}, \frac{44}{29}, \frac{42}{24} \right\} = \frac{36}{28} = 1,29, \\ \hat{b}_{\tilde{f}_1} &= \max \left\{ \frac{31}{21}, \frac{36}{28}, \frac{44}{29}, \frac{42}{24} \right\} = \frac{42}{24} = 1,75, \end{aligned}$$

všechny fuzzy vývojové koeficienty jsou potom uvedeny v tabulce 4.7. Aby se daly metody mezi sebou snáze porovnávat, byly fuzzy koeficienty převedeny zároveň zpět do LR zápisu.

Zbývajícím postup směřující k dopočítání odhadu fuzzy rezervy je analogický s klasickou, příp. dříve popsanou fuzzifikovanou metodou Chain-Ladder.

Celkový odhad fuzzy rezervy s využitím minimálních a maximálních podílů $C_{i,j}$ pak bude mít podobu

$$\tilde{R}_{ChL} = \sum_{i=1}^6 \tilde{C}_{i,6} - \sum_{i=1}^6 C_{i,6-i+1} = (375, 47, 49)_{LR} - (249, 0, 0)_{LR} = (126, 47, 49)_{LR}.$$

j	\tilde{f}_j	j	\tilde{f}_j
0	(1,50, 2,00, 2,33)	0	(2, 00, 0, 50, 0, 33) _{LR}
1	(1,29, 1,50, 1,75)	1	(1, 50, 0, 21, 0, 25) _{LR}
2	(1,28, 1,33, 1,36)	2	(1, 33, 0, 06, 0, 03) _{LR}
3	(1,19, 1,25, 1,30)	3	(1, 25, 0, 06, 0, 05) _{LR}
4	(1,00, 1,00, 1,00)	4	(1, 00, 0, 00, 0, 00) _{LR}

Tabulka 4.7: Zaokrouhlené fuzzy vývojové koeficienty pro schéma 4.2 s využitím minimálních a maximálních podílů.

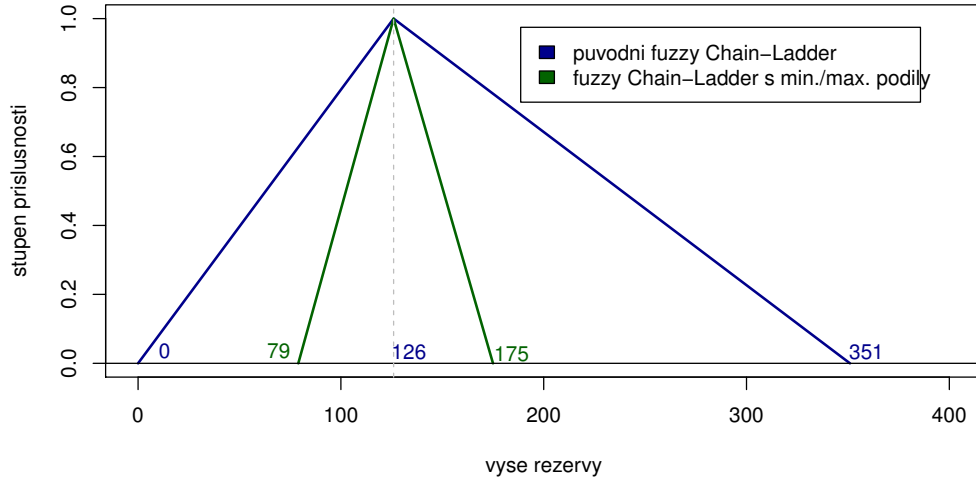
i	$\tilde{C}_{i,0}$	$\tilde{C}_{i,1}$	$\tilde{C}_{i,2}$
1	(9, 0, 0) _{LR}	(21, 0, 0) _{LR}	(31, 0, 0) _{LR}
2	(13, 0, 0) _{LR}	(28, 0, 0) _{LR}	(36, 0, 0) _{LR}
3	(14, 0, 0) _{LR}	(29, 0, 0) _{LR}	(44, 0, 0) _{LR}
4	(16, 0, 0) _{LR}	(24, 0, 0) _{LR}	(42, 0, 0) _{LR}
5	(12, 0, 0) _{LR}	(26, 0, 0) _{LR}	(39, 5,6, 6,5) _{LR}
6	(11, 0, 0) _{LR}	(22, 5,5, 3,7) _{LR}	(33, 12, 12) _{LR}

i	$\tilde{C}_{i,3}$	$\tilde{C}_{i,4}$	$\tilde{C}_{i,5}$
1	(42, 0, 0) _{LR}	(50, 0, 0) _{LR}	(50, 0, 0) _{LR}
2	(46, 0, 0) _{LR}	(60, 0, 0) _{LR}	(60, 0, 0) _{LR}
3	(60, 0, 0) _{LR}	(75, 3,6, 3,3) _{LR}	(75, 3,6, 3,3) _{LR}
4	(56, 2,3, 1,3) _{LR}	(70, 6,1, 4,7) _{LR}	(70, 6,1, 4,7) _{LR}
5	(52, 9,3, 10) _{LR}	(65, 14, 16) _{LR}	(65, 14, 16) _{LR}
6	(44, 17, 17) _{LR}	(55, 23, 25) _{LR}	(55, 23, 25) _{LR}

Tabulka 4.8: Schéma 4.2 doplněné metodou fuzzy Chain-Ladder při použití minimálních a maximálních podílů.

Porovnání výsledných odhadů dosažených oběma fuzzifikovanými metodami je pak znázorněno na obrázku 4.2.

Vztah (4.12) zajišťuje, že $\hat{a}_{\tilde{f}_j}$ bude vždy větší nebo rovno 1 (tj. $\hat{l}_{\tilde{f}_j} \geq 0$), tzn. opět máme zaručeno, že odhad fuzzy rezervy nebude ve svém nosiči zahrnovat žádný záporný prvek. Nevýhodou tohoto postupu je, že obecně nemáme k dispozici příliš mnoho minulých scénářů - nemůžeme tedy předvídat, zdali nemůže dojít i k výrazně lepší (či horší) situaci. Tento problém by bylo možné řešit navýšením hodnot $\hat{l}_{\tilde{f}_j}$, resp. $\hat{r}_{\tilde{f}_j}$, pomocí nějakých kladných koeficientů k_{min} , resp. k_{max} , která by dosavadní uskutečněné scénáře uměle zlepšila či zhoršila a tím pádem zvětšila



Obrázek 4.2: Porovnání odhadů fuzzy rezervy pro schéma 4.2.

celkovou neurčitost výsledku. Upravená rozpětí by se potom řídila vztahem:

$$\hat{l}_{\tilde{f}_j}^{adj} = k_{min} \cdot \min \left\{ \frac{C_{i,j+1}}{C_{i,j}}, \quad i = 1, \dots, n - j - 1 \right\}, \quad j = 0, \dots, n - 2,$$

$$\hat{r}_{\tilde{f}_j}^{adj} = k_{max} \cdot \max \left\{ \frac{C_{i,j+1}}{C_{i,j}}, \quad i = 1, \dots, n - j - 1 \right\}, \quad j = 0, \dots, n - 2.$$

I zde se ale nabízí otázka, jakým způsobem tyto koeficienty volit tak, aby jejich použití vedlo k odhadům lepším, a ne zbytečně nadhodnoceným či podhodnoceným.

4.3.3. Fuzzy Chain-Ladder s pevně nastaveným rozpětím fuzzy koeficientů

Zde se jedná o jednodušší přístup, než tomu bylo v předchozím případě. Vrátime se ke klasickému zápisu fuzzy čísla, kdy budeme opět pracovat s jádry fuzzy koeficientů $\hat{c}_{\tilde{f}_j}$, které jsme získali ze vztahu (4.8). Význačné hodnoty $\hat{a}_{\tilde{f}_j}$ a $\hat{b}_{\tilde{f}_j}$ se zde nyní odvodí přímo z hodnoty $\hat{c}_{\tilde{f}_j}$. Myšlenka je následující: expertně se zavedou kladná reálná čísla A , resp. B představující koeficienty, které okraje nosiče určí

poměrným snížením, resp. zvýšením hodnoty $\hat{c}_{\tilde{f}_j}$. Pro $\hat{a}_{\tilde{f}_j}$ a $\hat{b}_{\tilde{f}_j}$ tedy bude platit:

$$\hat{a}_{\tilde{f}_j} = A \cdot \hat{c}_{\tilde{f}_j}, \quad A \in \left\langle \frac{1}{\hat{c}_j^{min}}, 1 \right\rangle, \quad j = 0, \dots, n-2,$$

$$\hat{b}_{\tilde{f}_j} = B \cdot \hat{c}_{\tilde{f}_j}, \quad B \geq 1, \quad j = 0, \dots, n-2,$$

kde $\hat{c}_j^{min} = \min \{ \hat{c}_{\tilde{f}_j} \mid j = 0, \dots, n-2 \}$.

Zastavme se na chvíli u podmínek pro koeficienty A a B . Přirozeně vidíme, že aby $\hat{a}_{\tilde{f}_j}$ nebylo větší než $\hat{c}_{\tilde{f}_j}$, nemůže být koeficient A nikdy větší než 1. Zároveň chceme zachovat jistotu, že při předpokladu nezáporných hodnot $X_{i,j}$ během výpočtů nedostaneme ve finálním odhadu rezervy v nosiči záporné hodnoty. V kapitole 4.3.1 se dospělo k závěru že pro LR zápis fuzzy čísel tuto skutečnost zajišťuje podmínka (4.10), tedy když nejmenší možný přípustný rozdíl čísel \hat{f}_j a $\hat{l}_{\tilde{f}_j}$ je 1. Převedením do klasického zápisu fuzzy čísel snadno dostaneme, že musí platit $A \geq \frac{1}{\hat{c}_{\tilde{f}_j}}$ pro každé j - do podmínky proto stačí zahrnout nejmenší hodnotu \hat{c}_j^{min} . Vzhledem k tomu, že pro koeficient B nemáme žádná doplňující omezení, stačí, když nesníží hodnotu $\hat{c}_{\tilde{f}_j}$, proto $B \geq 1$.

Poznámka 4.5. Ačkoli pro oba koeficienty A, B při jejich zavedení připouštíme, že mohou být rovny 1, jedná se o extrémní případ, kdy by výsledný trojúhelníkový fuzzy vývojový koeficient byl deformovaný, případně tvořený „jen“ reálným číslem $\hat{c}_{\tilde{f}_j}$ při $A = B = 1$.

Tento způsob určení fuzzy koeficientů je velmi intuitivní a výpočetně nenáročný, na druhou stranu je obtížné vhodně a objektivně zvolit koeficienty A , potažmo B , proto se tímto přístupem blíže zabývat nebudeme.

4.3.4. Fuzzy Chain-Ladder využívající průměr podílů kumulativních pojistných plnění

V tuto chvíli přestáváme využívat vztah 4.8 pro výpočet jader fuzzy koeficientů \tilde{f}_j , i nadále se ovšem budou využívat známé hodnoty kumulativních pojistných plnění $C_{i,j}$. Jak již název této kapitoly napovídá, namísto předchozího

postupu, kdy jsme jádra fuzzy koeficientů určovali podílem součtů kumulativních pojistných plnění v po sobě jdoucích letech vývoje, tentokrát bude jádro určeno prostým průměrem těchto podílů (tj. individuálních vývojových koeficientů), tedy:

$$\hat{c}_j = \frac{1}{n-j-1} \sum_{i=1}^{n-j-1} \frac{C_{i,j+1}}{C_{i,j}}, \quad j = 0, \dots, n-2. \quad (4.14)$$

Tímto způsobem dáváme stejnou váhu všem podílům/scénářům, které v minulosti nastaly. V praxi je ovšem běžné, že je vyšší vliv kladen na „novější“ pozorování, tj. ta, která mají časově blíže k současnosti a dá se tedy předpokládat, že budou lépe zohledňovat aktuální situaci. Tuto úvahu lze ve výpočtu jádra fuzzy koeficientu zohlednit využitím vah $w_{i,j}$ následovně:

$$\hat{c}_j^w = \frac{1}{n-j-1} \sum_{i=1}^{n-j-1} w_{i,j} \cdot \frac{C_{i,j+1}}{C_{i,j}}, \quad j = 0, \dots, n-2, \quad (4.15)$$

kde

$$\sum_{i=1}^{n-j-1} w_{i,j} = 1, \quad w_{i,j} \leq w_{i+1,j} \quad \forall j \in \{0, \dots, n-2\}.$$

Co se týče výpočtu význačných hodnot $\hat{a}_{\hat{f}_j}$ a $\hat{b}_{\hat{f}_j}$, lze v podstatě použít již dříve popsané metody (s případným převodem do LR zápisu). Jedinou podmínkou, která musí být nutně splněna, je, aby $\hat{a}_{\hat{f}_j} \geq 1$ z důvodu nezápornosti konečného fuzzy odhadu rezervy.

4.3.5. Fuzzy Chain-Ladder s využitím regresní analýzy

Tato část nabídne trochu jiný pohled na problém volby hodnot $\hat{a}_{\hat{f}_j}$ a $\hat{b}_{\hat{f}_j}$ pro fuzzy vývojové koeficienty. Jednotlivé (ostré) vývojové koeficienty spočítané vztahem (4.5) budou v tuto chvíli představovat podkladová data pro konstrukci regresní křivky, tzn. bude se zkoumat, jaký vliv má příslušný rok vývoje na vývojový koeficient. Po určení typu regresní funkce a jejích parametrů se za nové odhady jádra pro jednotlivé fuzzy koeficienty vezmou hodnoty ležící na zkonstruované

regresní křivce. Dále je také možné pro každou pozorovanou hodnotu sestavit konfidenční interval na hladině spolehlivosti $\alpha \in \langle 0, 1 \rangle$, tj. $(1 - \alpha)\%$ interval spolehlivosti. Krajní body tohoto intervalu pro každý vývojový koeficient se potom vezmou za příslušné odhady $\hat{a}_{\tilde{f}_j}$, resp. $\hat{b}_{\tilde{f}_j}$.

Prvním krokem je volba regresní funkce pro zmíněná data. Vzhledem k charakteru ostrých vývojových koeficientů (nerostoucí posloupnost, jejíž hodnoty nikdy nebudou menší než 1), lze podle [8] sledovat lineární trend v logaritmech těchto vývojových koeficientů, které byly sníženy o 1. Takto modifikované koeficienty (značené λ_j^{mod}) tedy budou tvaru

$$\lambda_j^{mod} = \log(\hat{\lambda}_j - 1), \quad j = 0, \dots, n - 2.$$

Poznámka 4.6. Problém mohou tvořit situace, kdy odhad vývojového koeficientu bude roven přesně jedné (obvykle to může nastat u nejvyššího vývojového koeficientu, tj. $\hat{\lambda}_{n-2}$), kdy by logaritmus modifikovaného koeficientu nebyl definován - což je případ i zde používaných ilustračních dat. Jedním z možným řešení je prosté vynechání tohoto koeficientu pro určení parametrů regresní přímky.

Původní i modifikované vývojové koeficienty pro ilustrační data z tabulky 4.2 jsou vyznačeny na obrázku 4.3. Je vidět, že hodnoty modifikovaných vývojových koeficientů „přibližně“ leží na pomyslné přímce.⁹ Proto lze tyto hodnoty λ_j^{mod} proložit přímkou, jejíž parametry se budou hledat prostřednictvím metody nejmenších čtverců.

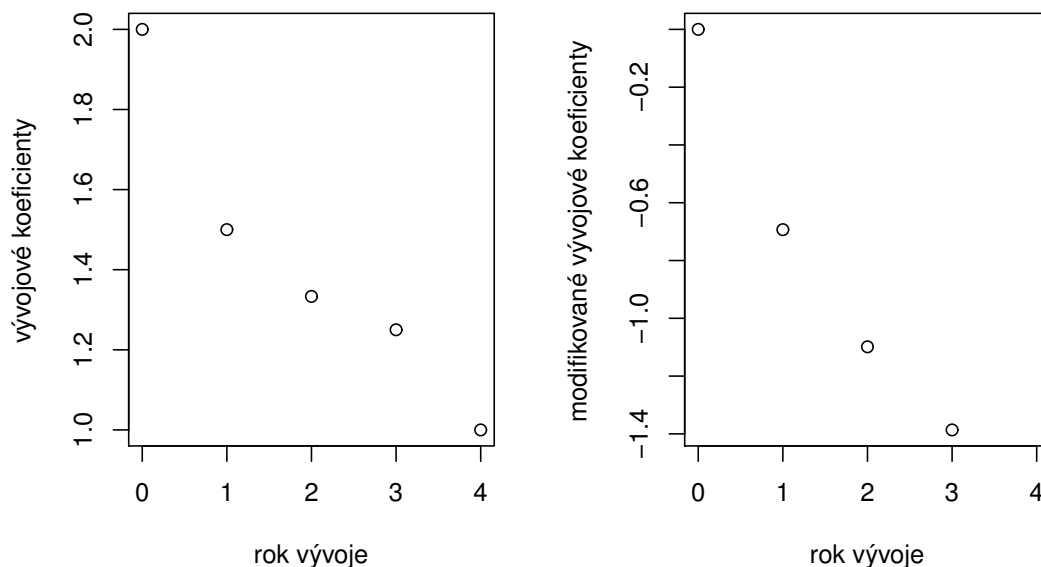
Řekněme, že pro pozorované hodnoty modifikovaných vývojových koeficientů platí

$$\lambda_j^{mod} = \beta_0 + \beta_1 j + \epsilon_j, \quad j = 0, \dots, n - 2, \quad (4.16)$$

kde β_0 a β_1 jsou parametry regresní přímky a ϵ_j je náhodná chyba pro j -té pozorování. Předpokládejme dále, že odchylky ϵ_j mají nulovou střední hodnotu, konstantní rozptyl a jsou vzájemně nekorelované. Modelovanou regresní přímku potom popíšeme vztahem

$$y = \hat{\beta}_0 + \hat{\beta}_1 j,$$

⁹Zde je potřeba brát v úvahu ilustrativní povahu dat a také malý počet pozorování. Proto se zde také myšlenka jen nastíní, ale finální odhad počítán nebude.



Obrázek 4.3: Původní a modifikované vývojové koeficienty.

kdy odhady jejích parametrů $\hat{\beta}_0$ a $\hat{\beta}_1$ určíme tak, aby minimalizovaly výraz

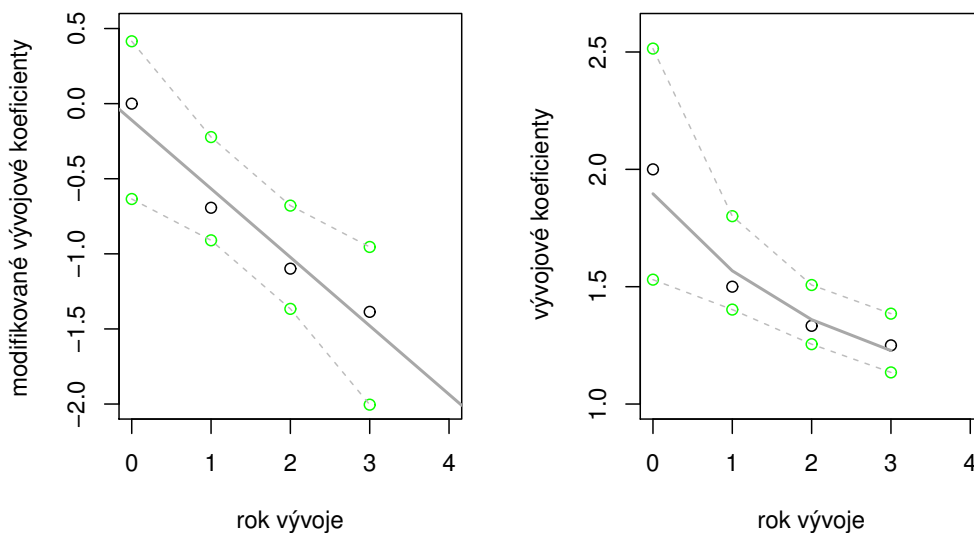
$$\sum_{j=0}^{n-2} [\lambda_j^{mod} - (\beta_0 + \beta_1 j)]^2,$$

což je podstata metody nejmenších čtverců. Nově odhadnutá jádra se tedy určí ze vztahu

$$\hat{c}_j = \hat{\beta}_0 + \hat{\beta}_1 j, \quad j = 0, \dots, n - 2.$$

Nyní zbývá určit zmíněné konfidenční intervaly pro jednotlivá pozorování. Jejich konstrukce je blíže popsána např. v [12], pro účely této práce byla využita syntaxe funkce *lm* v softwaru R. Na obrázku 4.4 jsou tyto odhady znázorněny jak pro modifikované, tak po převedení i pro původní vývojové koeficienty, a to pro hladinu citlivosti $\alpha = 0,05$ (což odpovídá 95% intervalu spolehlivosti).

Stejnou myšlenku lze aplikovat i při využití prostých průměrů individuálních vývojových koeficientů (analogie s kapitolou 4.3.4). Nyní se tedy za hodnoty, které budeme prokládat regresní křivkou, budou brát právě průměrné hodnoty jednot-



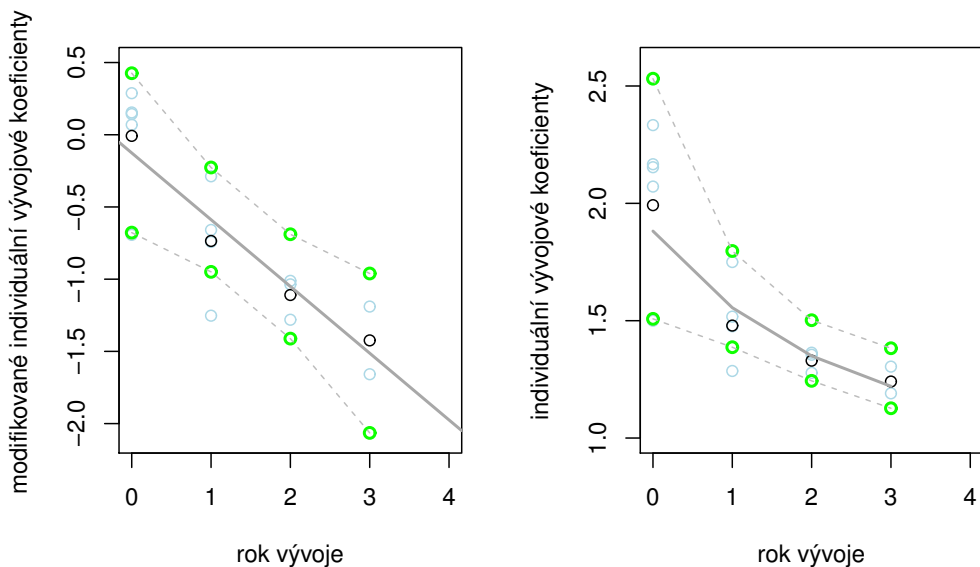
Obrázek 4.4: Grafické znázornění konfidenčních intervalů.

livých podílů kumulativních pojistných plnění. Obrázek 4.5 zobrazuje kromě samotných průměrů a jejich intervalových odhadů i původní individuální vývojové koeficienty, ze kterých se ve výpočtu vycházelo.

Zatímco dosud bylo pro vysvětlování jednotlivých metod využito ilustrační (a proto nepříliš obsáhlé) trojúhelníkové schéma z tabulky 4.2, následující část se bude věnovat aplikaci těchto metod na datech obsáhlejších, která lépe odpovídají skutečnosti.

4.4. Aplikace variant metody fuzzy Chain-Ladder

V této sekci již budou představené metody aplikovány na stejná data, se kterými pracovali autoři článku [11]. Jedná se o schéma o rozměrech 10×10 (tzn. $n = 10, i = 1, \dots, 10, j = 0, \dots, 9$), které je zobrazeno v tabulce 4.9. Jednotky, ve kterých jsou jednotlivé peněžní částky, bohužel nejsou ve zdroji uvedeny, ovšem pro samotnou metodologii to není podstatné - proto i nadále budou uváděné hodnoty bez jednotek.



Obrázek 4.5: Grafické znázornění postupu pro individuální vývojové koeficienty.

Veškeré provedené výpočty v rámci uvedených metod včetně výchozího trojúhelníkového schématu jsou uloženy v souboru *Pavlu_DP.xlsx* - dále v této kapitole bude odkazováno na jednotlivé listy tohoto souboru.¹⁰ Vzhledem k rozsahu trojúhelníkového schématu, které bude po fuzzifikaci odhadovaných pojistných plnění ještě obsáhlejší, nebudou tato schémata dále uváděna přímo zde v textu.

Na úvod jsem na zmíněná data aplikovala metodu Chain-Ladder bez jakékoli fuzzifikace - výpočet je obsažen na listu *ChL_puvodni*. Z něj je patrné, že odhad rezervy pro tato data je roven 18 680 856. Následně jsem již přikročila k jednotlivým způsobům fuzzifikace.

4.4.1. Fuzzy Chain-Ladder

Výpočet uvedený na listu *FChL_puvodni* byl v tomto případě zcela analogický postupu popsanému v [11], proto i výsledky jsou totožné. Výsledný odhad rezervy je znázorněn na obrázku 4.6 (kvůli přehlednosti jsou jednotlivé význačné hodnoty odhadu fuzzy rezervy zaokrouhleny).

¹⁰I pro zmíněný soubor platí, že v něm nejsou odhadované hodnoty značeny stříškou.

$i \backslash j$	0	1	2	3	4
1	357 848	1 124 788	1 735 330	2 218 270	2 745 596
2	352 118	1 236 239	2 170 033	3 353 322	3 799 067
3	290 507	1 292 306	2 218 525	3 235 179	3 985 995
4	310 608	1 418 858	2 195 047	3 757 447	4 029 929
5	443 160	1 136 350	2 128 333	2 897 821	3 402 672
6	396 132	1 333 217	2 180 715	2 985 752	3 691 712
7	440 832	1 288 463	2 419 861	3 483 130	
8	359 480	1 421 128	2 864 498		
9	376 686	1 363 294			
10	344 014				

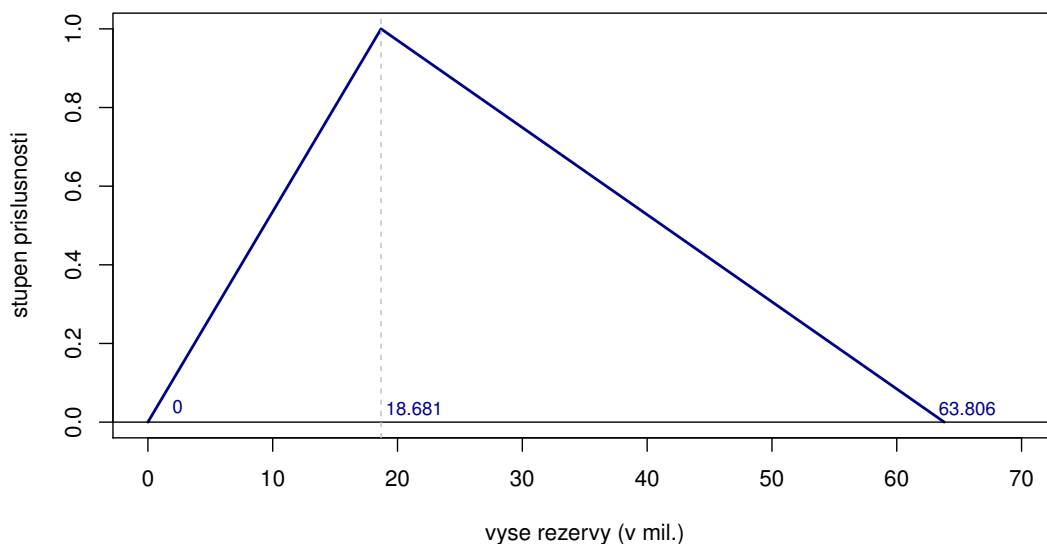
$i \backslash j$	5	6	7	8	9
1	3 319 994	3 466 336	3 606 286	3 833 515	3 901 463
2	4 120 063	4 647 867	4 914 039	5 339 085	
3	4 132 918	4 628 910	4 909 315		
4	4 381 982	4 588 268			
5	3 873 311				
6					
7					
8					
9					
10					

Tabulka 4.9: Trojúhelníkové schéma pro aplikaci jednotlivých metod.

Z obrázku je patrné, že nosič odhadu výsledné fuzzy rezervy je zleva ohraničen nulou, což bylo zdůvodněno v kapitole 4.3.1. Jádrem odhadu je hodnota „ostrého“ odhadu, tj. 18 680 856, směrem vpravo je potom fuzzy číslo výrazně neurčitější než vlevo.

4.4.2. Fuzzy Chain-Ladder - minimální a maximální podíly

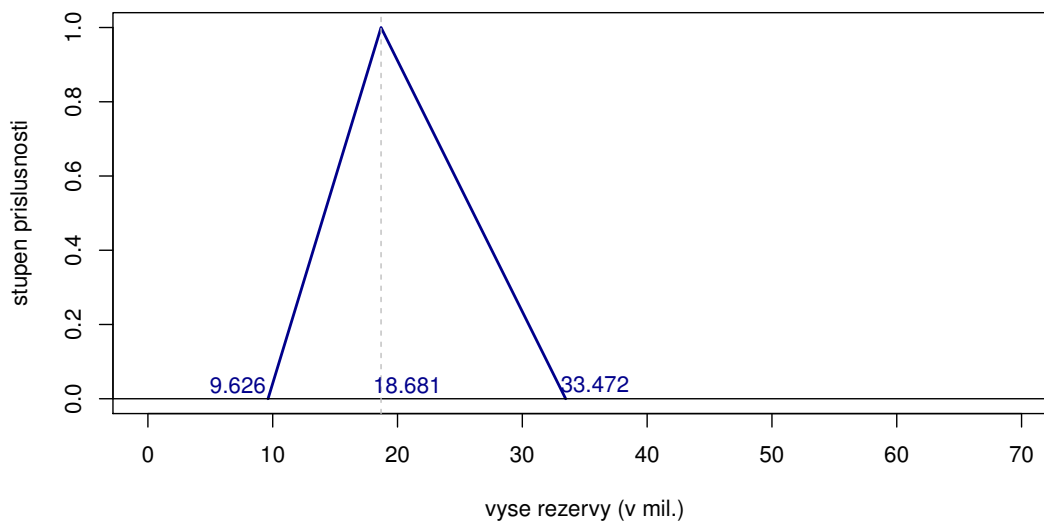
Tato metoda je použita na listu *FChL_min_max* - pro lepší přehlednost jsou uvedeny význačné hodnoty fuzzy koeficientů jak v klasickém zápise (tj. $(\hat{a}_{\tilde{f}_j}, \hat{c}_{\tilde{f}_j}, \hat{b}_{\tilde{f}_j})$), tak v LR zápise $(\hat{c}_{\tilde{f}_j}, \hat{l}_{\tilde{f}_j}, \hat{r}_{\tilde{f}_j})_{LR}$. Z výpočtu je patrné, že při vy-



Obrázek 4.6: Odhad rezervy metodou fuzzy Chain-Ladder.

užití této metody bude poslední fuzzy koeficient, tj. \tilde{f}_{n-2} , roven reálnému číslu odpovídajícímu ostrému vývojovému koeficientu. Důvodem je, že při výpočtu tohoto koeficientu je k dispozici jediný podíl hodnot, který je tím pádem maximálním, minimálním i „průměrným“ podílem. Při uvedeném výpočtu je proto jasné, že rozpětí tohoto fuzzy koeficientu bude na obě strany nulové - pro výpočet rezervy to znamená předpoklad, že v roce vývoje $n - 1$ se už žádné pojistné plnění vzhledm k rokům původu i nevyplatí.

Z grafu 4.7 je patrné, že v tomto případě je rozpětí odhadu směrem vlevo výrazně menší, tzn. předpokládá se, že určité pojistné plnění bude jistě vyplaceno (což se vzhledem k uvažovanému časovému horizontu jeví jako pravděpodobné). Na druhou stranu se poměrně výrazně zmenšilo i rozpětí směrem vpravo - nejvyšším hodnota v nosiči fuzzy rezervy v tomto případě nedosahuje ani dvojnásobku očekávané hodnoty.

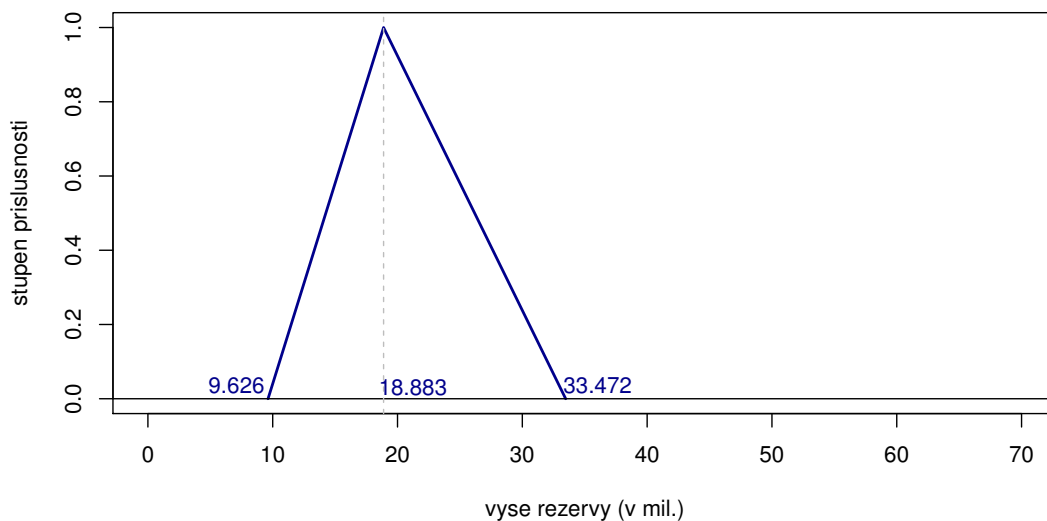


Obrázek 4.7: Odhad rezervy metodou fuzzy Chain-Ladder s využitím minimálních a maximálních podílů.

4.4.3. Fuzzy Chain-Ladder - průměrné podíly

List *FChL_prumerne_podily* obsahuje konstrukci jader fuzzy vývojových koeficientů jako průměrných podílů jednotlivých kumulativních pojistných plnění. Co se týče volby rozpětí fuzzy čísla, zvolila jsem způsob využívající minimální a maximální podíly pro co největší využití právě jednoduchých dílčích podílů - při jejich použití se samozřejmě opakuje výše uvedený „problém“ s koeficientem \hat{f}_{n-2} , který je zde opět reálným číslem.

Z grafu je patrné, že odhad rezervy se oproti metodě z kapitoly 4.4.2 příliš nezměnil - jádro rezervy se mírně zvýšilo na 18 883 073. Nelze ovšem jednoznačně říci, že se jedná o všeobecně platnou skutečnost - pro různá data mohou být průměrné podíly jednotlivých pojistných plnění vyšší, ale i nižší, než je tomu v případě podílu jejich součtů. Faktem ale je, že při použití stejného způsobu pro odhad význačných hodnot $\hat{l}_{\tilde{f}_j}$ a $\hat{r}_{\tilde{f}_j}$ jednotlivých fuzzy koeficientů bude zachována stejná neurčitost výsledné rezervy (což vychází ze vztahu pro násobení trojúhelníkových fuzzy čísel).

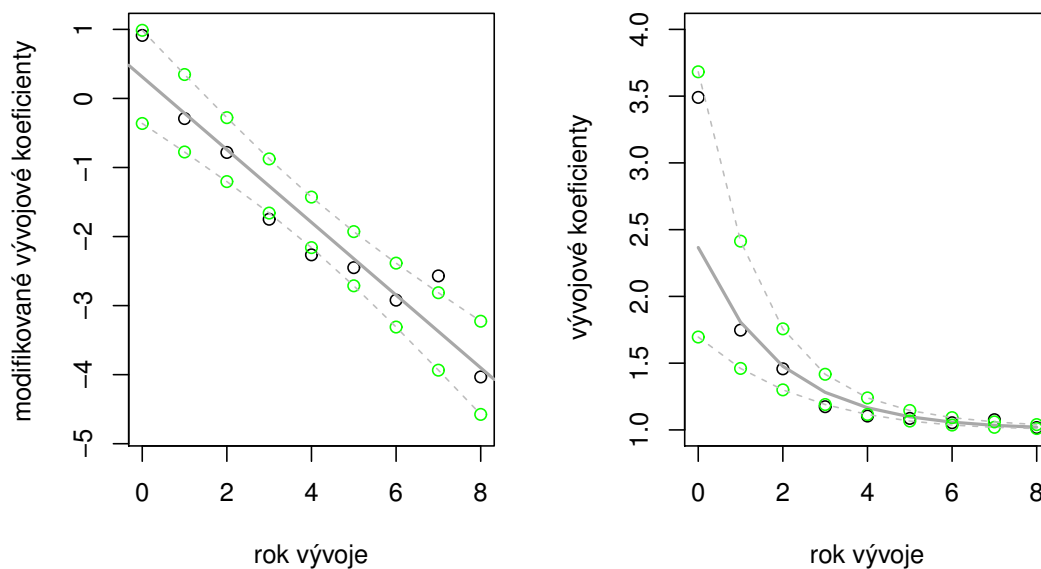


Obrázek 4.8: Odhad rezervy metodou fuzzy Chain-Ladder s využitím průměrných podílů pojistných plnění.

4.4.4. Fuzzy Chain-Ladder - regresní analýza

Nejprve jsem aplikovala regresní analýzu na vývojové koeficienty jakožto podíly součtů kumulativních pojistných plnění. Zkonstruovaná regresní křivka včetně konfidenčních intervalů je znázorněna na obrázku 4.9. Postup výpočtu, kdy jsou za jádra fuzzy koeficientů $\hat{c}_{\tilde{f}_j}$ brány fitované hodnoty z regresní přímky a za hodnoty $\hat{a}_{\tilde{f}_j}$, resp. $\hat{b}_{\tilde{f}_j}$ krajní hodnoty příslušných konfidenčních intervalů, je uveden na listě *FChL_regrese1*.

V tomto případě došlo u celkového odhadu rezervy ke znatelnějšímu navýšení hodnoty jádra fuzzy rezervy (v řádu jednotek procent oproti předchozím metodám), zároveň se oproti jiným modifikacím původní metody Chain-Ladder zvýšila i neurčitost ve směru vysokých hodnot. Pro porovnání jsem vyzkoušela i regresi průměrných hodnot jednoduchých podílů - výpočty jsou provedeny na listě *FChL_regrese2*. V obou případech jsem pro stanovení konfidenčních intervalů zvolila α rovno 0,05, jednalo se tedy o 95% interval spolehlivosti. Oproti výše uvedenému se v tomto případě všechny význačné hodnoty fuzzy čísla popisujícího



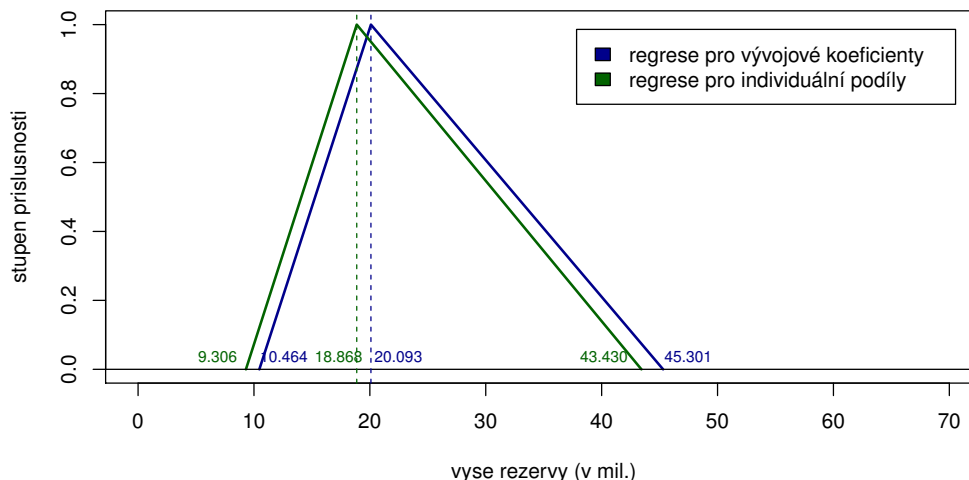
Obrázek 4.9: Grafické zobrazení regresní analýzy pro vývojové koeficienty.

odhad rezervy snížily (viz obrázek 4.10).

4.4.5. Shrnutí výsledků

Následující tabulky shrnují výsledky dosažené jednotlivými modifikacemi metody fuzzy Chain-Ladder. V tabulce 4.10 lze vidět význačné hodnoty jednotlivých fuzzy odhadů rezervy, v tabulce 4.11 jsou pak uvedeny defuzzifikované hodnoty reprezentující jednotlivé odhady, a to konkrétně těžiště a střed průměrného intervalu.

Celkově lze vidět, že výsledná rezerva původní metody fuzzy Chain-Ladder má obě defuzzifikované charakteristiky nejvyšší - způsobeno je to především velkou neurčitostí odhadu v kladném směru. V ostatních, modifikovaných metodách se potom eliminovala záporná stránka původní verze metody fuzzy Chain-Ladder, a to připouštění až nesmyslně nízkých hodnot v nosiči celkové fuzzy rezervy. Ve všech těchto případech se nejmenší hodnota náležející do nosiče pohybovala těsně pod 0,5násobkem hodnoty jádra. Na druhou stranu, všechny modifikované



Obrázek 4.10: Porovnání odhadů dosažených prostřednictvím regresní analýzy.

metody se oproti původní metodě Chain-Ladder „drží při zemi“ co se týče uvažovaných velkých hodnot rezervy.

Vezmeme-li do úvahy defuzzifikační charakteristiky modifikovaných metod, v porovnání s výchozí ostrou metodou Chain-Ladder nejpodobněji vychází obě modifikace, které za odhad význačných hodnot \hat{l}_{f_j} , resp. \hat{r}_{f_j} fuzzy vývojových koeficientů berou minimální, resp. maximální individuální podíly kumulativních pojistných plnění. Patrné je to při pohledu na podobu odhadu fuzzy rezervy - kromě toho, že mají jádro srovnatelné (příp. stejné) s ostrým odhadem rezervy, je i jejich neurčitost směrem vlevo a vpravo relativně nejrovnoměrněji rozdělena oproti ostatním odhadům.

Oproti tomu odhady u metod založených na regresní analýze mají výrazně vyšší neurčitost směrem v pravo, tj. výsledný odhad zahrnuje ve svém nosiči daleko vyšší hodnoty. Důvodem je koncept určení fuzzy vývojových koeficientů - zatímco u modifikovaných vývojových koeficientů ze vztahu (4.16) je neurčitost na obě strany stejná (z lineární regresní funkce), po logaritmizaci těchto hodnot se neurčitost směrem vpravo zvýší, což se projeví i na celkovém odhadu rezervy. Samozřejmě by zde bylo možné pohybovat s hladinou citlivosti α a tím i se

šířkou konfidenčních intervalů - při širších konfidenčních intervalech by neurčitost směrem vpravo, a tím i defuzzifikační charakteristiky, ještě více vzrostly.

použitá metoda	$\hat{a}_{\tilde{R}_{ChL}}$	$\hat{c}_{\tilde{R}_{ChL}}$	$\hat{b}_{\tilde{R}_{ChL}}$
Chain-Ladder	—	18 680 856	—
fuzzy Chain-Ladder	0	18 680 856	63 805 733
FCh-L: minimální + maximální podíly	9 626 333	18 680 856	33 471 909
FCh-L: průměrné podíly	9 626 333	18 883 073	33 471 909
FCh-L: regrese vývojových koeficientů	10 463 674	20 092 656	45 300 830
FCh-L: regrese průměrných podílů	9 305 666	18 868 470	43 430 006

Tabulka 4.10: Odhady rezervy - význačné hodnoty

použitá metoda	$t_{\tilde{R}_{ChL}}$	$\bar{E}(\tilde{R}_{ChL})$
Chain-Ladder	—	—
fuzzy Chain-Ladder	27 495 529	25 291 861
fuzzy Ch-L: minimální + maximální podíly	20 593 032	20 114 988
fuzzy Ch-L: průměrné podíly	20 660 438	20 216 097
fuzzy Ch-L: regrese vývojových koeficientů	25 285 720	23 987 454
fuzzy Ch-L: regrese průměrných podílů	23 868 047	22 618 153

Tabulka 4.11: Odhady rezerv - defuzzifikace

Závěr

Aparát teorie fuzzy množin poskytuje nástroje, pomocí kterých lze do (nejen) matematických problémů zahrnout neurčitou informaci, resp. nejistotu. Ve své práci jsem se věnovala tomu, jak lze koncept fuzzy množin využít v pojišťovnictví. Na základě průzkumu literatury jsem stručně přiblížila některé doposud popsané aplikace teorie fuzzy množin v pojišťovnictví - pro lepší porozumění jsem práci doplnila i o úvodní kapitoly věnující se terminologii z oblasti fuzzy množin a pojišťovnictví. Podrobněji rozebraným tématem byla možnost fuzzifikace metody Chain-Ladder používané ke stanovení rezervy na pojistná plnění v neživotním pojištění, kdy jsem nejprve popsala „ostrou“ verzi této metody a potom její publikovanou fuzzifikovanou variantu. Vzhledem k některým negativním stránkám této metody jsem se potom pokusila vymyslet vlastní alternativní postupy, které by tyto negativa potlačily.

Na základě výsledků shrnutých v sekci 4.4.5 lze vidět, že i za pomoci poměrně jednoduchých matematických úvah bylo možné původní metodu fuzzy Chain-Ladder modifikovat tak, aby (alespoň některé) nově vzniklé odhady rezervy byly víceméně srovnatelné s odhadem získaným původní ostrou metodou. V případě modifikovaných fuzzy metod bylo potom výrazným pozitivem dosažení jasné a logické interpretace význačných hodnot $\hat{l}_{\tilde{f}_j}$, resp. $\hat{r}_{\tilde{f}_j}$ u jednotlivých vývojových koeficientů. Nabízí se navíc řada způsobů, jak s těmito modifikovanými metodami dále pracovat - některé z nich byly v práci i naznačeny (práce s váženým průměrem individuálních podílů kumulativních pojistných plnění oproti průměru aritmetickému, pohyb hladiny citlivosti při využití regresní analýzy apod.). Rozhodně se tedy dá říct, že celá problematika není vyčerpána, prostor pro další

hledání dalších metod tu zcela jistě je.

Otázkou zůstává, zda má v dnešní době využití teorie fuzzy množin v pojišťovnictví své místo. Již během průzkumu literatury mě překvapilo, že aplikace, které mě zaujaly, pocházely z velké části ze starších zdrojů (cca 15 let a starší). Nedá se říct, že by v dnešní době nevycházely články či studie zabývající se aplikacemi teorie fuzzy množin v pojišťovnictví, spíše se ale jedná o poměrně okrajové problematiky - alespoň z mého pohledu. Pro jejich popis by byl nutný mnohem podrobnější úvod do problematiky pojišťovnictví, který by podle mě celkovému charakteru práce neprospěl.

Pojišťovny jsou dosti specifickými institucemi a disponují ohromnými objemy dat. Dá se proto očekávat, že i v nadcházejících letech či desítkách let se s těmito daty bude různými způsoby nakládat - z novějších postupů v úvahu připadají např. metody strojového učení či používání složitějších mnohorozměrných statistických modelů. Jestli mezi těmito způsoby zůstanou i metody založené na využití teorie fuzzy množin, o tom lze jen spekulovat.

Literatura

- [1] Adem, A., Dağdeviren, M.: *A life insurance policy selection via hesitant fuzzy linguistic decision making model*. Procedia Computer Science, 2016.
- [2] Boissonnade, A.C.: *Earthquake Damage and Insurance Risk*. Ph.D. Dissertation, Stanford University, 1984.
- [3] Buckley, J.J.: *The Fuzzy Mathematics of Finance*. Fuzzy Sets and Systems, 1987.
- [4] Cipra, T.: *Finanční a pojistné vzorce*. Praha: Grada Publishing, 2006.
- [5] Cipra, T.: *Pojistná matematika: teorie a praxe (2. vydání)*. Praha: Ekopress, 2006.
- [6] Cipra, T.: *Riziko ve financích a pojišťovnictví: Basel III a Solvency II*. Praha: Ekopress, 2015.
- [7] Fullér, R., Majlender, P.: *On interactive fuzzy numbers*. Fuzzy Sets and Systems, 2004.
- [8] Carrato, A. et al: *Claims reserving with R: ChainLadder-0.2.11 Package Vignette*.
- [9] DeWit, G. W.: *Underwriting and Uncertainty*. Insurance: Mathematics and Economics, 1982.
- [10] Ducháčková, E.: *Principy pojištění a pojišťovnictví (2. vydání)*. Praha: Ekopress, 2005.
- [11] Heberle, J., Thomas, A.: *Combining chain-ladder claims reserving with fuzzy numbers*. Insurance: Mathematics and Economics, 2014.
- [12] Hron, K., Kunderová, P.: *Základy počtu pravděpodobnosti a metod matematické statistiky (2. vydání)*. Vydavatelství Univerzity Palackého v Olomouci, 2015.
- [13] Chen, J.J.G., He, Z.: *Using Analytical Hierarchy Process and Fuzzy Sets Theory to Rate and Rank the Disability*. Praha: Ekopress, 2015.

- [14] Kerkez, M.: *Estimating Reserve for Incurred but not Reported (IBNR) Claims with Fuzzy Regression Model*. Advances in Accounting, Auditing and Risk Management, 2018.
- [15] Klir, G. J., Yuan, B.: *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, New Jersey, 1996
- [16] Lemaire, J.: *Fuzzy insurance*. ASTIN Bulletin: The Journal of the IAA, 1990.
- [17] Luuka, P., Collan, M.: *New fuzzy insurance pricing method for giga-investment project insurance*. Insurance: Mathematics and Economics, 2015.
- [18] Novosádová, L.: *Metody defuzzifikace fuzzy čísel*. Diplomová práce . Univerzita Palackého v Olomouci - Přírodovědecká fakulta, 2012.
- [19] Ostaszewski, K.: *Fuzzy Sets Methods in Actuarial Science*. Schaumburg, IL: Society of Actuaries, 1993.
- [20] Pavlů, I.: *Metody stanovení rezervy na pojistná plnění v neživotním pojištění*. Bakalářská práce . Univerzita Palackého v Olomouci - Přírodovědecká fakulta, 2018.
- [21] Shapiro, A.: *Modeling future lifetime as a fuzzy random variable*. Insurance: Mathematics and Economics, 2013.
- [22] Shapiro, A.: *An Overview of Insurance Uses of Fuzzy Logic*. Computational Intelligence in Economics and Finance, 2007.
- [23] Shapiro, A.: *Insurance applications of fuzzy logic*. Insurance: Mathematics and Economics, 2004.
- [24] Subudhi, S., Panigrahi, S.: *Use of optimized Fuzzy C-Means clustering and supervised classifiers for automobile insurance fraud detection*. Journal of King Saud University – Computer and Information Sciences, 2017.
- [25] Tanaka, H., Hayashi, I., Watada, J.: *Possibilistic linear regression analysis for fuzzy data*. European Journal of Operational Research, 1989.
- [26] Tanaka, H., Uejima, S., Asai, K.: *Linear Regression Analysis with Fuzzy Model*. IEEE Transaction on Systems, Man and Cybernetics, 1982.
- [27] Taylor, G.C., Ashe, F.R.: *Second moments of estimates of outstanding claims*. Journal of Econometrics, 1983.
- [28] Wüthrich, M.; Merz, M.: *Stochastic claims reserving methods in non-life insurance*. John Wiley & Sons, 2008.

- [29] Young, V.R.: *The Application of Fuzzy Sets to Group Health Underwriting*. Transactions of the Society of Actuaries, 1993.
- [30] Zadeh, L.: *Fuzzy sets*. Information and Control 8, 1965.
- [31] Zhou, J., Yang, F., Wang, K.: *Fuzzy arithmetic on LR fuzzy numbers with applications to fuzzy programming*. Journal of Intelligent & Fuzzy Systems, 2016.
- [32] Zákon č.277/2009 Sb. o pojišťovnictví, v platném znění