



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA PODNIKATELSKÁ

FACULTY OF BUSINESS AND MANAGEMENT

ÚSTAV INFORMATIKY

INSTITUTE OF INFORMATICS

ANALÝZA VÝVOJE NEŽIVOTNÍHO POJIŠTĚNÍ POMOCÍ ČASOVÝCH ŘAD

ANALYSIS OF THE DEVELOPMENT OF NON-LIFE INSURANCE USING TIME SERIES

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Jan Fousek

VEDOUCÍ PRÁCE

SUPERVISOR

RNDr. Zuzana Chvátalová, Ph.D.

BRNO 2018

Zadání bakalářské práce

Ústav:	Ústav informatiky
Student:	Jan Fousek
Studijní program:	Kvantitativní metody v ekonomice
Studijní obor:	Matematické metody v ekonomice
Vedoucí práce:	RNDr. Zuzana Chvátalová, Ph.D.
Akademický rok:	2017/18

Ředitel ústavu Vám v souladu se zákonem č. 111/1998 Sb., o vysokých školách ve znění pozdějších předpisů a se Studijním a zkušebním řádem VUT v Brně zadává bakalářskou práci s názvem:

Analyza vývoje neživotního pojištění pomocí časových řad

Charakteristika problematiky úkolu:

Úvod
Vymezení problému a cíle práce
Teoretická východiska práce
Analýza problému a současné situace
Vlastní návrhy řešení, přínos návrhů řešení
Závěr
Seznam použité literatury
Přílohy

Cíle, kterých má být dosaženo:

Cílem této bakalářské práce je vymezit pojmy v pojišťovnictví, zejména v oblasti neživotního pojištění a jeho struktury, dále vysvětlit vybrané principy regresní analýzy a časových řad vhodné pro použití v pojišťovnictví. V praktické části bude poté těmito metodami odhadnut budoucí vývoj různých druhů neživotního pojištění. Bude využita databáze České asociace pojišťoven.

Základní literární prameny:

ČERMÁK, Libor a Rudolf HLAVIČKA. Numerické metody. 2. vyd. Brno: Akademické nakladatelství CERM, s.r.o, 2008. ISBN 978-80-214-3752-4.

DAÑHEL, Jaroslav. Pojistná teorie. 1. vyd. Praha: Professional Publishing, 2005. ISBN 80-864-19-4-3.

DUCHÁČKOVÁ, Eva. Pojištění a pojišťovnictví. 1. vyd. Praha: Ekopress, 2015. ISBN 978-80-878-5-25-5.

HINDLS, Richard. Statistika pro ekonomy. 6. vyd. Praha: Professional Publishing, 2006. ISBN 80-86-1999-1.

KROPÁČ, Jiří. Statistika B: jednorozměrné a dvourozměrné datové soubory, regresní analýza, časové řady. 2. dopl. vyd. Brno: Vysoké učení technické v Brně, 2009. ISBN 80-214-3295-0.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2017/18

V Brně dne 28.2.2018

L. S.

doc. RNDr. Bedřich Půža, CSc.
ředitel

doc. Ing. et Ing. Stanislav Škapa, Ph.D.
děkan

Abstrakt

Bakalářská práce se zaměřuje na analýzu dat v oblasti neživotního pojištění. Představuje oblasti pojišťovnictví v České republice, vybrané statistické metody, zejména prostředky teorie časových řad. Výchozí data jsou čerpána z portálu České asociace pojišťoven. Výpočty a vizualizace jsou provedeny s podporou software STATISTICA a systému MATLAB.

Abstract

The bachelor thesis focuses on the analysis of data of general insurance. It introduces the area of insurance in the Czech Republic, selected statistic methods, especially the means of time series theory. The default data is drawn from the portal of the Czech Insurance Association. Calculations and visualizations are performed with the support of STATISTICA software and MATLAB.

Klíčová slova

časové řady, regresní analýza, STATISTICA, neživotní pojištění, predikce

Key words

time series, regression analysis, STATISTICA, general insurance, prediction

Bibliografická citace

FOUSEK, J. *Analýza vývoje neživotního pojištění pomocí časových řad*. Brno: Vysoké učení technické v Brně, Fakulta podnikatelská, 2018. 81 s. Vedoucí bakalářské práce RNDr. Zuzana Chvátalová, Ph.D.

Čestné prohlášení

Prohlašuji, že předložená bakalářská práce je původní a zpracoval jsem ji samostatně. Prohlašuji, že citace použitých pramenů je úplná, že jsem ve své práci neporušil autorská práva (ve smyslu Zákona č. 121/2000 Sb., o právu autorském a o právech souvisejících s právem autorským).

V Brně dne 15. května 2018

.....

podpis autora

Poděkování

Tímto bych rád poděkoval vedoucí své práce RNDr. Zuzaně Chvátalové, Ph.D. za odborné vedení a věcné rady, kterými mi pomohla k vypracování této bakalářské práce.

OBSAH

ÚVOD	9
CÍL A METODIKA PRÁCE	11
1 TEORETICKÁ VÝCHODISKA	13
1.1. Pojišťovnictví.....	13
1.1.1 Pojistné, pojistné plnění.....	13
1.1.2 Pojistné riziko	14
1.1.3 Pojistný trh a jeho subjekty	14
2 ANALÝZA SOUČASNÉHO STAVU	17
2.1.1 Neživotní pojištění	17
2.2 Analýza časových řad	22
2.2.1 Základní pojmy	22
2.2.2 Dekompozice časových řad.....	24
2.2.3 Trend v časové řadě.....	25
2.3 Regresní analýza	38
2.3.1 Metoda nejmenších čtverců.....	40
2.3.2 Testování hypotéz.....	41
3 VLASTNÍ NÁVRHY ŘEŠENÍ	44
3.1 Pojištění budov	44
3.2 Pojištění domácností	56
3.3 Havarijní pojištění.....	66
4 PŘÍNOSY NÁVRHŮ ŘEŠENÍ	72
ZÁVĚR	74
SEZNAM POUŽITÝCH ZDROJŮ	77
SEZNAM TABULEK	79
SEZNAM GRAFŮ	80
SEZNAM PŘÍLOH	81

ÚVOD

Pojišťovnictví je oblast, která v každé vyspělé ekonomice hraje významnou úlohu jak pro život každého občana, tak pro subjekty podnikatelské sféry, ale i pro stát jako celek. V kontextu své široké působnosti souvisí pojišťovnictví s celou řadou vnějších vlivů (ekonomický růst, vývoj hrubého domácího produktu, inflace aj.).

Pojišťovnictví je důležité jednak z pohledu individuálního pro občana či subjektu, kterému poskytne jistou pojistnou ochranu, jednak z pohledu širšího makroekonomického, protože přispívá k bezporuchovosti chodu ekonomiky státu, a nakonec je třeba poznamenat, že i na finančním a investičním trhu je konkurentem či partnerem bankovního sektoru. (1, s. 6).

V období zhruba do konce 80. roků minulého století v naší republice výpočty pojistné matematiky byly považovány vesměs za nadbytečné. Po roce 1989 se již diverzifikované odvětví (pojišťovnictví) u nás transformuje a obnovuje fungování pojistného trhu, a tedy výpočty pojistné matematiky se stávají nezbytnými (1, s. 1).

Vývoj pojišťovnictví se odráží významnou měrou i v legislativě České republiky (zákon o pojišťovnictví, jeho postupné novelizace s ohledem k vývoji společnosti, ale například i s ohledem k implementaci směrnic Evropské unie).

V roce 1994 byla založena Česká asociace pojišťoven. *„Jejími posláními je koordinovat, zastupovat, hájit a prosazovat společné zájmy pojišťoven ve vztahu k orgánům státní správy a dalším osobám i ve vztahu k zahraničí.“* (2). Česká asociace pojišťoven i Český statistický úřad zveřejňují pravidelně celou řadu ukazatelů z oblasti pojišťovnictví.

Každý občan či subjekt by měl být odpovědný vůči svému životu, majetku i aktivitám. Druhy pojištění lze třídit z různých hledisek. Oblasti životního i neživotního pojištění nabízejí velké množství různých pojistných produktů. Při výběru vyvstává otázka, který z produktů vybrat a podle jakých kritérií rozhodovat tak, aby pojištění splnilo svůj účel i očekávání. Čím více informací jak statického, tak dynamického charakteru je k dispozici, tím správnější by měla být volba pojistného produktu. Zejména dynamický vývoj ukazatelů může vhodně zachytit závislosti a vliv různých faktorů. V současnosti, při

dostupnosti prostředků informačních a komunikačních technologií jak výpočty a vizualizace v pojistné matematice, tak modelování pojistných ukazatelů jsou dosažitelné.

Proto jsem se rozhodl ve své bakalářské práci zabývat analýzou dat v oblasti neživotního pojištění využitím pokročilejších prostředků statistických metod v prostředí vhodného software.

CÍL A METODIKA PRÁCE

Cílem této bakalářské práce je analyzovat data poskytnutá Českou asociací pojišťoven. K tomu je využito teorie časových řad a programů STATISTICA a MATLAB. U jednotlivých časových řad je nejdříve zkoumán jejich průběh, na základě kterého je poté zvolena vhodná metoda pro jejich dekompozici. Mezi tyto metody patří popis trendu pomocí matematických křivek, exponenciální vyrovnávání a metoda klouzavých průměrů. Podle jednotlivých metod je poté určena predikce budoucích hodnot sledovaného ukazatele.

U každého sledovaného ukazatele je nejdříve znázorněn graf hodnot zveřejněných Českou asociací pojišťoven, na základě kterého určíme, zda lze pro dekompozici použít některý z trendů. Pokud ano, porovnáme jejich index determinace. Ten, který nabývá nejvyšší hodnoty, je pro nás nejvhodnější a analyzovaná data jím proložíme. Pokud žádný z trendů použít nelze, využijeme pro dekompozici exponenciální vyrovnávání, případně metody klouzavých průměrů.

Hlavním cílem této bakalářské práce je analyzovat charakteristická data vybraných pojistných produktů z oblasti neživotního pojištění v rámci České republiky v období roků 2000 až 2016 využitím prostředků matematicko-statistických metod ve vhodném počítačovém software (STATISTICA, MATLAB). K získání zdrojových dat je využita veřejně dostupná databáze České asociace pojišťoven.

Ke splnění hlavního cíle práce je nutné dosáhnout několika podcílů:

- (1) Monitoring odborných zdrojů z oblasti (pokročilejších) statistických, resp. matematických metod, pojišťovnictví, příhodných počítačových software a dostupnost zdrojových dat. Formulace teoretických východisek.
- (2) Analýza současného stavu v oblasti pojišťovnictví v České republice, pojistné produkty, charakteristická pojistná data pro zvolené produkty.
- (3) Návrh vlastního řešení problematiky pro analyzovaná pojistná data, která budou zvolena z oblasti produktů *neživotního pojištění*, a to hojně frekventované: *pojištění budov, domácností a havarijní*. Charakteristická data pro tyto produkty budou vybrána v podstatných kategoriích: *předepsané pojistné, počet (uzavřených) smluv, počet vyřízených pojistných událostí*. U jednotlivých časových řad (danou množinu pojistných dat) bude nejdříve zkoumán jejich

průběh informativním testem, na základě kterého bude poté zvolena vhodná metoda pro jejich dekompozici. Jde o matematické modelování trendu při respektu metod statistiky, vizualizace trendu křivkou, resp. více křivkami (jde o regresní model, jednoduché či dvojitě exponenciální vyrovnávání, Holtovu metodu, metodu klouzavých průměrů aj.) a dále bude proveden test hypotézy pomocí testového kritéria Studentova t -rozdělení pro vhodnost parametrů nejlepšího modelu. Krátkodobá predikce vývoje bude zachycena bodovými a intervalovými odhady, přitom pro intervalovou předpověď bude samostatně vytvořen program v systému MATLAB. Následovat bude interpretace analýzy.

1 TEORETICKÁ VÝCHODISKA

V části „*Teoretická východiska*“ jsem se zaměřil na vysvětlení základních pojmů v oblasti pojišťovnictví.

1.1. Pojišťovnictví

Pojišťovací instituce jsou významnou součástí finančního systému národního hospodářství. Plní důležité funkce, které pomáhají zajišťovat fungování ekonomiky a její stabilitu. Mezi tyto funkce řadíme například pojištění finančních subjektů a tím zajištění jejich finanční bezpečnosti, anebo investování na kapitálovém trhu (3, s. 100).

Z právního hlediska lze na činnosti pojišťoven nahlížet jako na převzetí rizika pojišťovnou od pojistného subjektu na základě pojistné smlouvy. Z hlediska principu je pojištění založeno na vzájemném rozložení rizika mezi velké množství pojištěných (3, s. 101).

1.1.1 Pojistné, pojistné plnění

Pojistné představuje částku, kterou pojištěný subjekt platí za přenesení rizika na pojišťovnu, tj. za přenesení negativních finančních důsledků, vzniklých náhodnými jevy (4, s. 50).

Pojistné plnění lze označit jako částku, kterou pojišťovna vyplatí jako náhradu škody na majetku, či výplatu z úrazového, důchodového a životního pojištění, přičemž k vyplacení dojde, nastala-li nahodilá událost, případně jiná událost, která je v životě člověka něčím významná, například překročení věkové hranice pro odchod do důchodu (4, s. 50).

Obecně platí, že výše pojistného plnění by měla být taková, aby pojištěnému vytvořila stejnou hmotnou situaci, jako by k pojistné události nedošlo. V praxi se tato výše odvíjí od aktuální časové ceny předmětu pojištění. To navíc znamená, že pro zajištění postačitelnosti daného pojistného musí pojišťovna při každém přepracování pojištění dané věci upravovat pojistné podle současné ceny věci (4, s. 50).

1.1.2 Pojistné riziko

Pojistné riziko představuje možnost vzniku pojistné události, při které pojišťovna, na základě uzavřené pojistné smlouvy vyplatí pojištěnému pojistné plnění. Pojištění v sobě zahrnuje krytí pouze tzv. čistého rizika, tedy rizika prokazatelně náhodného charakteru. Tím může být například úraz, dopravní nehoda atd. Nahodilost pojistné události lze rozdělit na *absolutní*, (například požár, úraz) a *relativní*, (například úmrtí) (5, s. 177).

Samotný pojem pojištění lze poté chápat jako ochranu proti pojistnému riziku. Pojištěný přenesl rizika, která by pro něj mohla být neúnosná, na pojistitele. Ten, za předpokladu velkého množství souboru pojistných smluv, které tvoří tzv. *pojistný kmen*, je schopen převzít rizika v důsledku inkasovaného pojistného nejen zvládnout, ale i utržit zisk (5, s. 178).

1.1.3 Pojistný trh a jeho subjekty

„Existence solidního, důvěryhodného pojistného trhu je symptomem zdravé, úspěšné ekonomiky a dobrého fungování finanční sféry na daném teritoriu“ (4, s. 45).

K nabídce pojištění a poptávce po pojištění dochází na pojistném trhu, který je podobný jakémukoliv jinému trhu. Z důvodu jeho hladkého chodu vstupují někdy mezi klienty a pojistitele tzv. *zprostředkovatelé*, kterými mohou být agenti, poradci atd. (4, s. 45).

a) Pojišťovny

Pojišťovny jsou finanční instituce, které slouží klientovi jako ochrana před finanční ztrátou, která vznikla v důsledku pojistné události.

Nejčastější členění pojišťoven bývá podle následujících hledisek:

- z hlediska šíře zaměření jejich činnosti,
- z hlediska právní formy,
- z hlediska předmětu poskytovaného pojištění (3, s. 101).

Rozdělení pojišťoven z hlediska šíře zaměření jejich činnosti

Na základě tohoto členění se pojišťovny dále dělí na:

- **specializované**, pro které je charakteristické zaměření pouze na určitý druh pojištění, na určité riziko či na některou skupinu pojištěných,

- **univerzální**, které naopak pojišťují v podstatě všechny druhy rizik v oblasti životního i neživotního pojištění (3, s. 101–102).

Rozdělení pojišťoven z hlediska právní formy

Podle tohoto hlediska lze pojišťovny rozdělit na:

- **akciové společnosti**, které se v praxi vyskytují nejčastěji, jedná se o podnikatelské subjekty, které jsou licencovány a regulovány,
- **vzájemné pojišťovny**, fungující na zásadě, že škoda, která vznikla jednomu členovi bude nesena všemi členy, tedy i těmi, kterým škoda nevznikla. Z toho vyplývá, že vlastníci pojišťovny jsou zároveň jejími klienty a pojištění je poskytováno pouze jim,
- **státní pojišťovny**, které jsou zakládány pro pojištění v oblastech zdravotnictví a sociálního pojištění (3, s. 101).

Rozdělení pojišťoven z hlediska předmětu poskytovaného pojištění

Dle tohoto hlediska se pojišťovny dělí do dvou základních odvětví:

- **pojišťovny působící v odvětví životního pojištění – jejich** úkolem je jednak finančně chránit blízké osoby pojištěného před důsledky, které mohou vzniknout nahodilou událostí, a také samotného pojištěného pro případ, že dojde k výrazné změně jeho postavení ve společnosti, na základě které by mu poklesly příjmy,
- **pojišťovny působící v odvětví neživotního pojištění – zaměřují** se na pojištění rizik, které by mohly způsobit škodu na majetku nebo zranění osob (3, s. 102–105).

b) Zajišťovny

Jedná se o instituce, které poskytují tzv. *zajištění* neboli „pojištění pojišťoven“. Zatímco zajišťovny mohou nabízet pouze zajištění, pojišťovny, mají-li k tomu potřebnou licenci, mohou kromě pojišťovacích produktů nabízet také zajištění (3, s. 106).

c) Pojistní brokeri

Pojistní brokeri fungují jako mezičlánek mezi klientem a pojišťovnou a usnadňují styk mezi nimi. Podle stupně nezávislosti lze brokery (makléře) rozdělit do dvou skupin, a to na makléře, kteří pracují pro určité pojišťovny a sjednávají pro ně pojištění, a na *makléře*,

kteří pracují nezávisle na pojišťovnách a jejichž úkolem je sjednat pro svého klienta nejvýhodnější pojistnou ochranu (4, s. 47).

2 ANALÝZA SOUČASNÉHO STAVU

Česká asociace pojišťoven ve výročních zprávách poskytuje pouze konkrétní data týkající se jednotlivých druhů neživotního pojištění a naměřených hodnot jednotlivých ukazatelů, jako je předepsané pojistné, počet vyřízených pojistných událostí apod. Co ale nezkoumá, je dlouhodobý trend ve vývoji jednotlivých ukazatelů neživotního pojištění a predikci budoucích hodnot. Protože to jsou důležité údaje, rozhodl jsem se tomu sám věnovat.

2.1.1 Neživotní pojištění

„Neživotní pojištění zahrnuje krytí celé škály pojistných nebezpečí neživotního charakteru. V rámci neživotního pojištění jsou kryta rizika různého charakteru, a to ohrožující zdraví a životy osob, rizika vyvolávající přímě věcné škody, rizika vyvolávající finanční ztráty.“ (6, s. 178)

Jedna z možností, jak lze členit neživotní pojištění, je rozdělit ho do následujících oblastí:

- neživotní pojištění osob,
- majetkové pojištění,
- pojištění finančních ztrát,
- pojištění odpovědnosti za škody (6, s.178).

V následujících podkapitolách jsou popsány ty druhy pojištění, se kterými se setkáváme nejčastěji.

a) Neživotní pojištění osob

Úrazové pojištění

„Úrazové pojištění zahrnuje výplatu pojistného plnění v případě, že v důsledku úrazu dojde k přechodnému nebo trvalému tělesnému poškození nebo smrti pojištěného.“ (6, s. 178).

Úrazové pojištění může být sjednáno samostatně jako druh pojištění, případně se může vyskytovat jako doplňkové pojištění některého z dalších pojištění, například jako součást pojištění motorových vozidel (7, s. 120).

K naplnění podstaty tohoto pojištění je nutné, aby došlo k úrazu. V praxi platí, že pojišťovny určují ve smluvních podmínkách výluky z pojistného plnění. Mezi ty patří například následující skutečnosti:

- pokud k úrazu došlo v souvislosti s trestným činem,
- pojištěný si úraz způsobil úmyslně,
- pracovní úrazy,
- nemoci z povolání atd. (6, s. 179).

V počátcích vývoje pojištění se úrazové pojištění nabízelo pouze jako pojištění pro případ úplné invalidity. V dnešní době nabízejí pojišťovny v rámci svých produktů také pojištění, v jehož rámci je i částečná invalidita (6, s. 180).

Příklad procentuálního ohodnocení trvalých následků po úrazu prezentuje Tabulka 1, kde v procentech je uveden stupeň invalidity:

Tabulka 1: Hodnocení stupně invalidity (Zdroj: vlastní úprava dle (8))

Loketní kloub a předloktí	
Úplná ztuhlost loketního kloubu v nepříznivém postavení (úplné natažení nebo úplné ohnutí a postavení jim blížká)	25 %
Úplná ztuhlost loketního kloubu v příznivém postavení nebo v postaveních jemu blížkých (ohnutí v úhlu 90 až 95 st.)	16 %
Omezení pohyblivosti loketního kloubu	do 14 %
Úplná ztuhlost kloubů radioulnárních (s nemožností přivrácení nebo odvrácení předloktí)	15 %
Omezení převrácení a odvrácení předloktí	do 12 %
Viklavý kloub loketní	do 20 %
Ztráta jedné paže do výše nad loket	65 %
Ztráta jedné paže do výše pod loket nebo jedné ruky	60 %

Stupeň invalidity určuje velikost pojistného plnění podle následujících vzorců (1), (2), (3) dle zdroje (8):

$$PP = P\check{C} \cdot SI \quad (1)$$

$$PP = P\check{C} \cdot (30 \% + 5 \cdot (SI - 30 \%)) \quad (2)$$

$$PP = P\check{C} * (280 \% + 26 \cdot (SI - 80 \%)) \quad (3)$$

kde PP vyjadřuje pojistné plnění, $P\check{C}$ pojistnou částku a SI stupeň invalidity. Pokud $SI \in \langle 0,001; 30,00 \rangle$, použijeme pro výpočet vzorec (1), pro $SI \in \langle 30,01; 80,00 \rangle$ vzorec (2), pro $SI \in \langle 80,01; 100,00 \rangle$ vzorec (3).

b) Pojištění majetku

„Pojištění majetku zahrnuje krytí rizik, jejichž realizací dochází ke škodám na majetku (poškození, zničení, ztráta věcných hodnot). Pojištění majetku zahrnuje krytí celé řady rizik, při jejichž realizaci dochází ke vzniku přímých věcných škod.“ (6, s. 185)

Pojištění majetku obyvatelstva

Pojištění majetku obyvatelstva je jedno z nejvýznamnějších pojištění vůbec. Jeho jednotlivé druhy jsou popsány v následujících odstavcích *Pojištění domácnosti*, *Pojištění budov* a *Havarijní pojištění*.

Pojištění domácnosti

Toto pojištění zahrnuje pojištění zařízení domácnosti, které slouží členům domácnosti. Na základě ohodnocení předmětů, které se nacházejí v domácnosti, je určena pojistná částka, která se vytváří na principu pojištění na plnou hodnotu. Pokud se v domácnosti nachází předměty, jejichž cena se vymyká standardní výbavě, jako jsou například starožitnosti, umělecká díla, pak jsou v základním pojištění domácnosti kryty jen do určité výše, přičemž lze pro tyto předměty sjednat *doplňkové pojištění* (7, s.123).

Ve většině případů jsou v pojištění zahrnuta tato rizika:

- vodovodní riziko,
- riziko odcizení,
- živelná rizika (7, s. 123).

V rámci tohoto pojištění je také možnost sjednání tzv. *pojištění rekreační domácnosti*, které se od běžného pojištění liší v několika znacích. Mezi ně patří například vyšší tarif či nemožnost sjednání pojištění proti odcizení cennosti (4, s. 164).

Pojištění budov

Stejně jako v pojištění domácnosti, i zde je sdruženo krytí několika různých rizik. Rozsah těchto rizik lze v praxi rozdělit podle následujících variant:

- *základní*, která zahrnuje krytí rizika požáru, blesku, výbuchu, vichřice, odcizení, vandalství,
- *širší*, v které je kromě výše uvedeného obsaženo navíc riziko pádu předmětů, zřícení skal, vodovodní riziko,
- *nejširší*, tzv. All Risks, která obsahuje určité výluky týkající se válečných a politických rizik (6, s. 189).

Havarijní pojištění

Tento druh pojištění slouží ke krytí škod, které vznikly na motorových vozidlech, a to v případě, kdy je řidič sám způsobil, nezpůsobil, případně se na škodě částečně podílel.

V rámci havarijního pojištění jsou kryta tato rizika:

- riziko havárie,
- živelní riziko,
- riziko odcizení,
- vandalství,
- šomážní a strojní riziko (6, s. 190).

Pojištění podnikatelských a průmyslových rizik

V rámci této kategorie pojištění lze hovořit o mnoha druzích pojistných produktů. V následující části práce jsou uvedeny ty nejdůležitější z nich: *živelní pojištění* a *dopravní pojištění*.

Živelní pojištění

Jedná se o nejklasičtější pojištění v rámci pojištění podnikatelských subjektů, jelikož kryje škody, které vznikly na majetku podnikatele v důsledku realizace živelního rizika (6, s. 192).

Dopravní pojištění

Tento druh pojištění v sobě zahrnuje jak pojištění *kaska*, tzn. pojištění škod, které vznikly na dopravních prostředcích, tak pojištění *karga*, tj. pojištění škod, které mohou vzniknout na věcech, které jsou přepravovány (6, s. 194).

c) Pojištění finančních ztrát

Pojištění pro případ přerušení provozu

Jak název napovídá, tento druh pojištění má za úkol krýt škody, mezi které patří ušlý zisk vzniklý z přerušení provozu či náklady vynaložené například na hlídání objektu.

Účelem tohoto pojištění je krýt důsledky, které vznikají při nesplacení poskytnutého úvěru. Úvěrové pojištění mohou využít jak podnikatelské, tak bankovní subjekty.

Úvěrové pojištění

Účelem tohoto pojištění je krýt důsledky, které vznikají při nesplacení poskytnutého úvěru. Úvěrové pojištění mohou využít jak podnikatelské, tak bankovní subjekty.

d) Pojištění odpovědnosti za škody

Pojištění odpovědnosti za škody je důležité v situacích, kdy svou činností způsobíme škodu jinému subjektu, přičemž se může jednat o škodu na majetku, na zdraví i na životě.

Toto pojištění je charakteristické tím, že za pojistnou událost je považována povinnost pojištěného zaplatit škodu, kterou způsobil. O tom, zda má subjekt, kterému vznikla škoda, povinnost tuto škodu uhradit, rozhoduje soud. Na základě soudního rozhodnutí poté pojišťovna vyplácí pojistné plnění, nikoliv pojištěnému, avšak subjektu, kterému vznikla škoda.

Pojištění odpovědnosti za škody v sobě zahrnuje následující:

- náhradu škody,
- náklady na obhajobu pojištěného,
- náklady na soudní řízení o náhradě škody (7, s. 133).

2.2 Analýza časových řad

K analýze určitého ekonomického jevu často potřebujeme hodnoty, které sledujeme v jejich časovém vývoji, tzv. časové řady. Zjednodušeně řečeno, časová řada je posloupnost hodnot sledovaného ekonomického ukazatele, které jsou uspořádány v čase. Takovou analýzu lze do jisté míry chápat jako regresní analýzu, přičemž vysvětlující proměnnou je čas (9, s. 245).

Pomocí časových řad je možné, kromě kvantitativní analýzy ekonomického ukazatele v jeho dosavadním průběhu, odhadnout i jeho budoucí vývoj.

2.2.1 Základní pojmy

Chceme-li analyzovat ekonomické ukazatele pomocí časových řad, je nutné nejprve zavést základní pojmy.

Časová řada

„Časovou řadou rozumíme řadu hodnot určitého ukazatele, uspořádaných z hlediska přirozené časové posloupnosti. Přitom je nutné, aby věcná náplň ukazatele i jeho prostorové vymezení byly shodné v celém sledovaném časovém úseku“ (10, s. 115).

Časové řady lze dělit z několika hledisek:

- Podle rozhodného časového hlediska na časové řady *intervalové* a *okamžikové*.
- Podle periodicity, s jakou jsou údaje sledovány, na řady *krátkodobé*, ve kterých probíhá sledování čtvrtletně, měsíčně atd. a na řady *dlouhodobé*, kde probíhá sledování ročně.
- Podle způsobu vyjádření sledovaných údajů na časové řady *peněžní* a *naturální*. Pro peněžní platí, že údaje jsou vyjádřeny v penězích a naturální v naturálních jednotkách, například počty smluv (9, s.246).

Charakteristiky časových řad

Abychom získali o časových řadách potřebné informace, musíme definovat charakteristiky časových řad. Pro jejich zavedení předpokládejme, že hodnoty dané časové řady v časových okamžicích, resp. intervalech t , kde $t = 1, 2, 3, \dots, n$, jsou okamžikové řady označené y_t , a jsou kladné. Dále předpokládejme, že vzdálenosti mezi

sousedními časovými okamžiky, resp. středy časových intervalů, jsou stejné. Za předpokladu splnění daných podmínek můžeme zavést následující charakteristiky (10, s. 118).

Průměr okamžikové řady

Průměr okamžikové řady je označen \bar{y} a lze jej vypočítat pomocí vzorce (4) (10, s. 118).

$$\bar{y} = \frac{1}{n-1} \cdot \left[\frac{y_1}{2} + \sum_{t=2}^{n-1} y_t + \frac{y_n}{2} \right] \quad (4)$$

Průměr intervalové řady

Průměr intervalové řady, značen rovněž \bar{y} , vychází ze vzorce (5) (10, s. 118).

$$\bar{y} = \frac{1}{n} \cdot \sum_{t=1}^n y_t. \quad (5)$$

První diference

První diference Δ_t^1 je nejjednodušší charakteristika, kterou lze popsat vývoj časové řady. Vyjadřuje, o kolik se změnila hodnota časové řady v daném okamžiku oproti okamžiku bezprostředně předcházejícímu. Pro její výpočet platí vzorec (6) (10, s. 120), (9, s. 253).

$$\Delta_t^1 = y_t - y_{t-1}, t = 2, 3, \dots, \quad (6)$$

Druhá diference

Jestliže řada prvních diferencí roste nebo klesá, tzn. nelze ji popsat lineárně (tj. hodnoty graficky proložit přímkou), počítáme tzv. *diference vyšších řádů*, přičemž pro nás jsou důležité především *druhé diference* Δ_t^2 , vyjadřující rozdíl dvou sousedních prvních diferencí (10, s. 120), (9, s. 253). Pro její výpočet platí vzorec (7):

$$\Delta_t^2 = \Delta_t^1 - \Delta_{t-1}^1 \quad (7)$$

Koeficient růstu

Koeficient růstu vyjadřuje, kolikrát se zvýšila, případně snížila, hodnota časové řady v určitém okamžiku (období), oproti okamžiku (období) předcházejícímu. Kolísají-li hodnoty koeficientu růstu kolem určité konstanty, lze říci, že danou časovou řadu lze popsat pomocí exponenciály (10, s. 120), (9, s. 253). Výpočet lze provést dle vzorce (8):

$$k_t = \frac{y_t}{y_{t-1}}, t = 2, 3, \dots, n \quad (8)$$

Průměrný koeficient růstu

Průměrný koeficient růstu \bar{k} vyjadřuje, kolikrát se průměrně změní hodnota časové řady za jednotkový časový interval (10, s. 121), (9, s. 253). Výpočet lze provést dle vzorce (9):

$$\bar{k} = (k_2 k_3 \dots k_n)^{\frac{1}{n-1}} = \sqrt[n-1]{(k_2 k_3 \dots k_n)} \quad (9)$$

2.2.2 Dekompozice časových řad

Tradiční princip modelování časových řad vychází z jednorozměrného modelu

$$y_t = f(t, \varepsilon_t), \quad (10)$$

kde y_t vyjadřuje hodnotu modelovaného ukazatele v čase t a ε_t náhodnou složku (9, s. 254).

K výše zmíněnému modelu lze přistupovat tak, že časovou řadu rozdělíme (*dekomponujeme*) na čtyři složky:

- složku *trendovou* T_t , stručně označenou jako trend,
- složku *sezónní* S_t ,
- složku *cyklickou* C_t a
- složku *náhodnou* ε_t ,

pro které platí rovnice (9, s. 254)

$$y_t = T_t + S_t + C_t + \varepsilon_t. \quad (11)$$

Trendová složka

Trendová složka T_t vyjadřuje hlavní tendenci dlouhodobého vývoje sledovaného ukazatele v čase. V praxi se můžeme setkat s trendem *rostoucím*, *klesajícím* či *konstantním*, kdy hodnoty sledovaného ukazatele kolísají kolem neměnné úrovně. Poslední případ, ač to není z exaktního hlediska korektní, se nazývá „*časová řada bez trendu*“ (9, s. 254).

Sezónní složka

Sezónní složka S_t popisuje pravidelně se opakující změny v časové řadě způsobující výchyly od trendové složky, které lze pozorovat v průběhu jednoho roku a platí, že se každý rok opakují, přičemž jejich perioda může být rovna například měsíci, čtvrtletí apod. (9, s. 255).

Cyklická složka

Cyklickou složku C_t chápeme jako kolísání okolo trendu, které je způsobeno dlouhodobým cyklickým vývojem, jehož délka vlny je delší než jeden rok. Můžeme se setkat také s rozdílným pohledem na tuto složku, kdy někteří odborníci hovoří spíše o fluktuaci okolo trendu. Faktorů, které ovlivňují cyklickou složku může být několik. Řadíme mezi ně například změny ekonomické, které lze často obtížně popsat, nebo změny mimo ekonomickou oblast, například cyklické změny v módě, které se projeví cyklickou změnou odbytu v různých odvětvích oděvního průmyslu (9, s. 255), (10, s. 124).

Náhodná složka

Obecně lze říci, že *náhodná složka* ε_t nám zbyde v časové řadě poté, co odstraníme složku trendovou, sezónní a cyklickou. Platí, že ji nelze popsat pomocí funkce času (9, s. 255).

2.2.3 Trend v časové řadě

Následující část bakalářské práce je věnována metodám, kterými lze z časové řady eliminovat trendovou složku a předpovídat její budoucí vývoj. Hovoříme-li o trendu v časové řadě, můžeme se často setkat také s tzv. *vyrovnáváním časové řady*. Tento pojem vysvětluje Kropáč následovně:

„Při zkoumání dlouhodobé vývojové tendence ukazatele časové řady, tj. trendu v časové řadě, je nutné „očistit“ zadané údaje od ostatních vlivů, které tuto vývojovou tendenci zastírají. Postup, kterým se toho dosahuje, se nazývá vyrovnávání časových řad“ (10, s. 125).

Pro eliminaci trendu lze použít různé metody, z nichž jsem pro tuto práci zvolil *popis trendu matematickými křivkami, exponenciální vyrovnávání a metodu klouzavých průměrů*.

Popis trendu pomocí matematických křivek

Nejpoužívanější metodou pro popis vývoje trendu časové řady je regresní analýza, o které pojednává kapitola 2.3.

Lineární trend

Lineární trend (tzv. trendová přímka), je nejpoužívanější typ trendové funkce. Vychází to z faktu, že jej lze použít vždy, když stačí alespoň hrubý přehled o vývoji sledované časové řady. Navíc platí, že na omezeném časovém intervalu jím lze aproximovat ostatní trendové funkce (9, s. 257).

Z hlediska regresní analýzy se jedná o tzv. *regresní přímku* (lineární závislost), popsanou vzorcem (12)

$$T_t = \beta_0 + \beta_1 t, \quad (12)$$

kde β_0, β_1 jsou neznámé parametry a t je časová proměnná nabývající hodnot $t = 1, 2, \dots, n$. Jelikož je funkce lineární v parametrech, lze pro odhady ozn. b_0, b_1 koeficientů β_0, β_1 využít *metodu nejmenších čtverců*, která je popsána v kapitole 2.3.1 (9, s. 258).

Z té vyplývají následující vzorce (13) pro výpočet *odhadu koeficientů* (9, s. 258):

$$b_0 = \bar{y} - b_1 \bar{t}, \quad (13)$$

$$b_1 = \frac{\sum_{t=1}^n t y_t - \bar{t} \sum_{t=1}^n y_t}{\sum_{t=1}^n t^2 - n \bar{t}^2}.$$

Důležitým pojmem, který bude použit dále v praktické části, je tzv. *intervalová předpověď*. Udává, v jakém intervalu se budou pohybovat budoucí hodnoty sledované časové řady. V případě trendové přímky lze tvar intervalové předpovědi vyjádřit vzorci (14) až (16) (9, s. 259):

$$T_P - t_{1-\frac{\alpha}{2}}[n-2]sh_p < T_P < T_P + t_{1-\frac{\alpha}{2}}[n-2]sh_p, \quad (14)$$

$$s = \sqrt{\frac{\sum_{t=1}^n y_t^2 - \frac{(\sum_{t=1}^n T_t)^2}{n}}{n-2}}, \quad (15)$$

$$h_p = \sqrt{1 + \frac{1}{n} + \frac{(P - \bar{t})^2}{\sum_{t=1}^n t^2 - n\bar{t}^2}}, \quad (16)$$

kde T_p značí hodnotu v čase P , $P > n$. Výraz $t_{1-\frac{\alpha}{2}}[n-2]$ značí $(1 - \frac{\alpha}{2})100$ % kvantil Studentova t -rozdělení s $n-2$ stupni volnosti, přičemž $(1 - \frac{\alpha}{2}) \cdot 100$ % značí spolehlivost intervalové předpovědi uváděnou v procentech (9, s. 259).

Parabolický trend

Parabolický trend lze obecně zapsat pomocí výrazu (17) (9, s. 262):

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2. \quad (17)$$

Pro určení vzorců pro odhad parametrů nejprve zavedeme časovou proměnnou t' . Ta, na rozdíl od časové proměnné t , nemá počátek v bodě, ve kterém začínáme pozorování, ale je zavedena tak, aby součet hodnot t' byl roven 0 (9, s. 261).

Stejně jako lineární, i parabolický trend je lineární v parametrech, a proto lze k odhadu parametrů využít opět metodu nejmenších čtverců. Vycházíme tedy ze soustavy rovnic (18) (9, s. 262–263)

$$\begin{aligned} \sum_{t=1}^n y_t &= nb_0 + b_1 \sum_{t=1}^n t' + b_2 \sum_{t=1}^n t'^2 \\ \sum_{t=1}^n y_t t' &= b_0 \sum_{t=1}^n t' + b_1 \sum_{t=1}^n t'^2 + b_2 \sum_{t=1}^n t'^3 \\ \sum_{t=1}^n y_t t'^2 &= b_0 \sum_{t=1}^n t'^2 + b_1 \sum_{t=1}^n t'^3 + b_2 \sum_{t=1}^n t'^4. \end{aligned} \quad (18)$$

Z druhé rovnice vyjádříme odhad parametru β_1 dle (19) (9, s. 263):

$$b_1 = \frac{\sum_{t=1}^n y_t t'}{\sum_{t=1}^n t'^2}. \quad (19)$$

Odhady parametrů β_0 a β_2 poté získáme ze soustavy rovnic (20) (9, s. 263):

$$\sum_{t=1}^n y_t = nb_0 + b_2 \sum_{t=1}^n t'^2 \quad (20)$$

$$\sum_{t=1}^n y_t t'^2 = b_0 \sum_{t=1}^n t'^2 + b_2 \sum_{t=1}^n t'^4,$$

Jako (21) a (22) (9, s. 263):

$$b_0 = \frac{\sum_{t=1}^n y_t \sum_{t=1}^n t'^4 - \sum_{t=1}^n t'^2 \sum_{t=1}^n y_t t'^2}{n \sum_{t=1}^n t'^4 - (\sum_{t=1}^n t'^2)^2}, \quad (21)$$

$$b_2 = \frac{n \sum_{t=1}^n y_t t'^2 - \sum_{t=1}^n y_t \sum_{t=1}^n t'^2}{n \sum_{t=1}^n t'^4 - (\sum_{t=1}^n t'^2)^2}. \quad (22)$$

Stejně jako u lineárního trendu, i v tomto případě lze určit intervalovou předpověď (9, s. 262–265) užitím vzorců (23) až (26):

$$T_P - t_{1-\frac{\alpha}{2}}[n-3]s \cdot h_p < T_P < T_P + t_{1-\frac{\alpha}{2}}[n-3]s \cdot h_p, \quad (23)$$

$$s = \sqrt{\frac{\sum_{t=1}^n y_t^2 - \sum_{t=1}^n T_t^2}{n-3}}, \quad (24)$$

$$h_p = \sqrt{1 + (1PP^2)(\mathbf{XX}')^{-1}(1PP^2)'}, \quad (25)$$

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ \vdots & \vdots & \vdots \\ 1 & n & n^2 \end{pmatrix}. \quad (26)$$

Exponenciální trend

Hovoříme-li o *exponenciálním trendu*, jedná se o typ trendové funkce, zapsaný ve tvaru vztahu (27):

$$T_t = \beta_0 \cdot \beta_1^t, \quad (27)$$

kde β_0 a β_1 vyjadřují neznámé parametry trendu a proměnná t je časová proměnná nabývající hodnot $t = 1, 2, \dots, n$ (9, s. 266).

Protože funkce (27) není lineární v parametrech, nelze pro odhad parametrů využít metodu nejmenších čtverců. Abychom tuto metodu mohli využít, využijeme tzv. *metodu linearizující transformace*, pomocí níž provedeme *logaritmickou transformaci* funkce (27). Logaritmováním této funkce dostaneme následující vztah (28) (9, s.266):

$$\log(T_t) = \log(\beta_0) + t \cdot \log(\beta_1). \quad (28)$$

Protože tvar (28) již je lineární v parametrech, lze pro odhad parametrů využít metodu nejmenších čtverců, tzn. lze minimalizovat výraz (29) (9, s. 266), kde symbolem \sum se vždy rozumí součet přes t od 1 do n :

$$Q = \sum (\log(y_t) - \log(T_t))^2 \dots \min. \quad (29)$$

Hledané odhady parametrů b_0 a b_1 poté získáme z rovnic (30) (9, s. 267)

$$\sum \log(y_t) = n \cdot \log(b_0) + \log(b_1) \sum t, \quad (30)$$

$$\sum t \log(y_t) = \log(b_0) \sum t + \log(b_1) \sum t^2.$$

Pokud převedeme časovou proměnnou t na časovou proměnnou t' , splňující podmínku $\sum t'^k = 0$, pro $k = 1, 3, 5, \dots$, lze soustavu (30) zjednodušit. Získáme tak soustavu (31) (9, s. 267)

$$\begin{aligned} \log(b_0) &= \frac{\sum \log(y_t)}{n}, \\ \log(b_1) &= \frac{\sum t' \log(y_t)}{\sum t'^2}. \end{aligned} \quad (31)$$

Problém, který se pojí s metodou linearizující transformace je ovšem v tom, že může poskytovat statisticky zkreslené výsledky. Z toho důvodu lze využít váženou metodu nejmenších čtverců. Ta namísto výrazu (29) vychází z minimalizace výrazu (32) s vahami w_t (9, s.267)

$$Q = \sum (\log(y_t) - \log(T_t))^2 w_t \dots \min. \quad (32)$$

Zavedením váhy $w_t = y_t^2$, dostaneme následující rovnice (33) (9, s.267)

$$\begin{aligned}\sum y_t^2 \log(y_t) &= \log(b_0) \sum y_t^2 + \log(b_1) \sum ty_t^2, \\ \sum ty_t^2 \log(y_t) &= \log(b_0) \sum ty_t^2 + \log(b_1) \sum t^2 y_t^2.\end{aligned}\quad (33)$$

Řešením těchto rovnic získáme výraz (34) (9, s. 267)

$$\log(b_0) = \frac{\sum y_t^2 \log(y_t) \sum t^2 y_t^2 - \sum ty_t^2 \sum ty_t^2 \log(y_t)}{\sum y_t^2 \sum t^2 y_t^2 - (\sum ty_t^2)^2} \quad (34)$$

$$\log(b_1) = \frac{\sum y_t^2 \sum ty_t^2 \log(y_t) - \sum y_t^2 \log(y_t) \sum ty_t^2}{\sum y_t^2 \sum t^2 y_t^2 - (\sum ty_t^2)^2}$$

Volba vhodného modelu trendu

Velmi důležitou otázkou, kterou si musíme klást u analýzy časových řad je, jaký trend použít, aby co nejpřesněji popisoval závislost získaných (naměřených) dat.

Jednou z možností, na základě které se můžeme rozhodnout, který trend použít, je **analýza grafu** zobrazené časové řady. Tato metoda je velmi jednoduchá, často ale nemusí být spolehlivá, jelikož velkou roli zde hraje subjektivita, tzn., že dva lidé mohou na základě analýzy grafu stejné časové řady dojít ke dvěma různým názorům z pohledu volby trendové křivky (9, s. 287).

Jako spolehlivější metodu můžeme využít **reziduálního součtu čtverců SSR** neboli součet čtverců odchylek empirických hodnot od hodnot vyrovnaných. Používá se k posouzení, jak těsně přiléhá zadaná funkce k naměřeným datům a zjistíme jej ze vzorce (35):

$$SSR = \sum_{t=1}^n (y_t - \hat{\eta}(x_t))^2, \quad (35)$$

kde y_t jsou empirické hodnoty a $\hat{\eta}_t$ hodnoty odhadovaného trendu analyzované časové řady. Jako nejspolehlivější trendovou funkci poté zvolíme takovou, pro kterou platí, že její reziduální součet čtverců je nejmenší (11, s. 38, s. 43).

Index determinace

Někdy také nazýván jako *koeficient determinace* R^2 . Pokud chceme posoudit, jak dobře daná funkce vystihuje zkoumanou funkční závislost, využijeme indexu determinace. Ten nabývá hodnot z intervalu $\langle 0,1 \rangle$, přičemž platí, že čím blíže se jeho hodnota blíží jedné, tím lépe je daná funkční závislost zvolenou regresní funkcí vystižena. Je možné jej také vyjádřit pomocí procent, kdy 1 odpovídá 100 % a 0 odpovídá 0 %. Výpočet lze provést užitím vzorců (36) až (38) (11, s. 44-45):

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{SSR}{TSS} \quad (36)$$

TSS zde značí celkový součet čtverců, tzn.

$$TSS = \sum_{t=1}^n (y_t - \bar{y})^2 \quad (37)$$

a RSS regresní součet čtverců

$$RSS = \sum_{t=1}^n (\hat{\eta}(x_i) - \bar{y})^2. \quad (38)$$

Informativní testy charakteristik

Jako pomocnou metodu, která může napovědět, která z trendové funkce je vhodná, můžeme využít také *informativní* (orientační) *testy*. Ty jsou založeny na vlastnostech charakteristik časových řad, viz Tabulka 2.

Tabulka 2: Informativní testy charakteristik (Vlastní zpracování dle (12, s. 274))

Vhodný trend	Informativní testy
Lineární trend	1. diference přibližně konstantní
Kvadratický trend	2. diference přibližně konstantní
Exponenciální trend	koeficienty růstu přibližně konstantní

Graf

Kvalitu modelu lze ověřit také pomocí grafu $(f(t), \hat{f}(t))$, kde $f(t)$ jsou body původních hodnot a $\hat{f}(t)$ body hodnot vyrovnaných. Pokud se body řadí přibližně do přímky se směrnici 1, můžeme model označit za dostatečně kvalitní (13, s. 36).

Exponenciální vyrovnávání

Základní charakteristikou tohoto přístupu ke trendové složce je, že hodnoty vyrovnávané řady jsou váženy do minulosti exponenciálně klesajícími váhami, odtud název *exponenciální vyrovnávání*. Jinými slovy hodnotám, které jsou vzdálenější do minulosti se přiřazují nižší váhy než hodnotám novějším (12, s. 288).

Při exponenciálním vyrovnávání vycházíme z aditivního modelu časové řady, tzn. (9, s. 322):

$$y_{t-k} = T_{t-k} + \varepsilon_{t-k}. \quad (39)$$

Pro určení hodnoty trendové složky T_{t-k} lze použít funkci danou vzorcem (40) (9, s. 322):

$$T_{t-k} = \beta_0 - \beta_1 k + \beta_2 k^2 + \dots + (-1)^k \beta_k k^k, \quad (40)$$

ve které β_0, \dots, β_k vystupují jako koeficienty a k proměnná značící stáří pozorování a platí, že $k = 0, 1, 2, \dots, t - 1$. Například proměnná T_t tedy představuje pozorování v přítomném čase, T_{t-1} pozorování v předešlém časovém okamžiku atd. (9, s. 322).

K tomu, abychom získali odhady parametrů β_0, \dots, β_k funkce (40) využijeme metody nejmenších čtverců, tzn. budeme minimalizovat výraz (41) (9, s. 323):

$$\sum_{k=0}^{t-1} (y_{t-k} - T_{t-k})^2 w_k \dots \min \quad (41)$$

ve kterém w_k značí váhy, pro které platí, že se vzrůstajícím „věkem“ klesají. Pomocí exponenciální funkce lze tyto váhy zapsat vztahem (42) jako (9, s. 323):

$$w_k = \alpha^k, \alpha \in (0; 1), k = 0, 1, \dots, t - 1, \quad (42)$$

kde α se nazývá *vyrovnávací konstanta*. Dosadíme-li do vzorce (41) tuto vyrovnávací konstantu, dostaneme základní vzorec pro odhad parametrů modelu (43) (9, s. 323–324):

$$\sum_{k=0}^{t-1} (y_{t-k} - T_{t-k})^2 \alpha^k \dots \min. \quad (43)$$

V praxi se můžeme setkat s několika druhy exponenciálního vyrovnávání, z nichž jsou v této práci zmíněny *jednoduché exponenciální vyrovnávání*, ve kterém v rovnici (40)

položíme $k = 0$, *dvojité exponenciální vyrovnávání*, kde $k = 1$, a *Holtova metoda*, která je níže uvedena a popsána především proto, že pomocí této metody pracuje program STATISTICA v případě, že mu zadáme vypracovat exponenciální vyrovnávání s lineárním trendem (9, s. 324).

Abychom měli jistotu, že vybereme správnou metodu exponenciálního vyrovnávání pro získaná data, tj. zda bude použito exponenciální vyrovnávání bez lineárního trendu či s lineárním trendem, určíme sumu čtverců reziduí pro jednotlivé metody. Jako lépe vyhovující metodu poté vybereme tu, pro kterou bude suma čtverců reziduí nabývat menší hodnoty, jelikož ta bude těsněji prokládat analyzovaná data (14).

Jednoduché exponenciální vyrovnávání

Jak bylo řečeno výše, v případě jednoduchého exponenciálního vyrovnávání platí, že ve vzorci (40) položíme $k = 0$. Jedná se o řady, u kterých lze prohlásit, že trend je v krátkých úsecích řady konstantní, tzn. platí vztah (44):

$$T_{t-k} = \beta_0, \quad (44)$$

kde β_0 značí vyrovnanou hodnotu v čase $t - k$, kde $k = 0, 1, \dots, t - 1$ (12, s. 288).

Pro odhad parametru β_0 , získaný na základě pozorování y_t, y_{t-1}, \dots , která máme k dispozici v čase t , využijeme vzorec (45) (9, s. 324):

$$b_0 = \hat{y}_{t-k} = (1 - \alpha) \sum_{k=0}^{t-1} \alpha^k y_{t-k}, \quad (45)$$

který lze pro jednoduchost přepsat jako výraz (46) (9, s. 324):

$$\hat{y}_{t-k} = (1 - \alpha)y_{t-k} + \alpha\hat{y}_{t-k-1}. \quad (46)$$

Pokud chceme pomocí této metody určit hodnotu v čase $t + 1$, tzn. v období následujícím po posledním měření, potom tuto hodnotu ztotožníme s vyrovnanou hodnotou v období t , tzn. získáme vztah (47) (9, s. 324):

$$P_t(1) = y_t. \quad (47)$$

Dvojité exponenciální vyrovnávání

Tento druh exponenciálního vyrovnávání můžeme využít v případě, že lze trend považovat v krátkých úsecích za lineární, tzn. platí (48) (9, s. 326):

$$T_{t-k} = \beta_0 - \beta_1 k. \quad (48)$$

Využitím vzorce (43) získáme následující rovnice (49) pro odhad parametrů b_0 a b_1 (9, s. 326):

$$\begin{aligned} \sum_{k=0}^{t-1} \alpha^k y_{t-k} &= b_0 \sum_{k=0}^{t-1} \alpha^k - b_1 \sum_{k=0}^{t-1} k \alpha^k, \\ - \sum_{k=0}^{t-1} k \alpha^k y_{t-k} &= -b_0 \sum_{k=0}^{t-1} k \alpha^k + b_1 \sum_{k=0}^{t-1} k^2 \alpha^k. \end{aligned} \quad (49)$$

Řešením těchto rovnic získáme vzorce (50) a (51) (9, s. 326):

$$b_0 = \frac{\sum_{k=0}^{n-1} \alpha^k y_{n-k} \sum_{k=0}^{n-1} k^2 \alpha^k - \sum_{k=0}^{n-1} k \alpha^k \sum_{k=0}^{n-1} k \alpha^k y_{n-k}}{\sum_{k=0}^{n-1} \alpha^k \sum_{k=0}^{n-1} k^2 \alpha^k - (\sum_{k=0}^{n-1} k \alpha^k)^2}, \quad (50)$$

$$b_1 = \frac{-\sum_{k=0}^{n-1} \alpha^k \sum_{k=0}^{n-1} k \alpha^k y_{n-k} + \sum_{k=0}^{n-1} \alpha^k y_{n-k} \sum_{k=0}^{n-1} k \alpha^k}{\sum_{k=0}^{n-1} \alpha^k \sum_{k=0}^{n-1} k^2 \alpha^k - (\sum_{k=0}^{n-1} k \alpha^k)^2}. \quad (51)$$

Stejně jako v případě exponenciálního vyrovnávání bez trendu lze i v tomto případě provádět predikce budoucích hodnot. Predikci i – tého budoucího časového období lze získat pomocí vzorce (52) (9, s. 326–329):

$$P_t(i) = b_{0t} + b_{1t} i. \quad (52)$$

Holtova metoda

Tato metoda je v této práci uvedena mimo jiné z toho důvodu, že ji využívá software STATISTICA, ve kterém provádím většinu výpočtů a vizualizací. Jak již bylo naznačeno, jedná se tedy o speciální případ, konkrétně o zobecnění, exponenciálního vyrovnávání s lineárním trendem. Holtova metoda je rozdílná v tom, že kromě vyrovnávací konstanty α využívá také vyrovnávací konstantu γ . Jejich funkci vyjadřují následující vztahy (53) a (54):

- konstanta α slouží pro vyrovnání úrovně L_t , ta je rovna vyrovnané hodnotě časové řady v čase t (12, s. 295):

$$L_t = \alpha y_{t-k} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (53)$$

- konstanta γ se využívá pro vyrovnání směrnice trendu T_t dané řady (12, s. 295):

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}. \quad (54)$$

Vyrovnaná hodnota je poté rovna (55) a (56) (12, s. 295):

$$\hat{y}_t = L_t, \quad (55)$$

$$\hat{y}_{t+i}(t) = L_t + T_t i. \quad (56)$$

Důležitou otázkou, kterou si musíme položit, je volba konstant α a γ . Protože různé literární prameny doporučují vybírat hodnoty z různých intervalů, například Cipra ve Finanční ekonometrii (12, s. 290), se omezuje na interval $0 < \alpha \leq 3$, na druhou stranu Hindls v knize Statistika pro ekonomy doporučuje výběr z intervalu $0,7 \leq \alpha < 1$ (9, s. 329).

Protože ve své práci využívám software STATISTICA, pro volbu nejvhodnějších hodnot konstant α a γ využiji funkci *Síťové hledání*. Pomocí ní získáme tabulku, ve které je zaznamenáno 10 nejlepších hodnot jednotlivých konstant, přičemž ta nejlepší z nich je zvýrazněna červeně (15).

Klouzavé průměry

Tato metoda předpokládá, že trend časové řady mění svůj charakter, a proto nelze pro vyrovnání využít žádnou z matematických křivek, které nemění v čase svůj charakter. Počítá ovšem s tím, že v krátkých časových úsecích je použití matematických křivek možné. Obvykle se jedná o křivky, které mají v různých úsecích různé parametry (12, s. 294).

Existuje několik druhů klouzavých průměrů, z nichž se v této bakalářské práci budu věnovat klouzavým průměrům, které pro vyrovnávání úseků časových řad využívají polynomicke křivky. Princip této metody lze zjednodušeně přiblížit například takto: prvních pět hodnot časové řady proložíme polynomem třetího stupně, pomocí něhož

vyrovnáme první dvě hodnoty a hodnotu prostřední. Poté posuneme na časové ose vyhlazovací okénko o jeden časový úsek doprava a těchto nových pět hodnot proložíme novým polynomem třetího stupně, pomocí něhož vyrovnáme prostřední hodnotu vyhlazovacího okénka. Takto se posouváme po časové ose do doby, než ve vyhlazovacím okénku máme posledních pět hodnot. Tyto opět proložíme polynomem třetího stupně a pomocí něj vyrovnáme prostřední hodnotu a poslední dvě hodnoty (10, s.127).

Pokud uvažujeme polynom třetího řádu, lze využít zápisu (57) (10, s. 128):

$$\eta(\tau, \mathbf{b}_k) = b_{k1} + b_{k2}\tau + b_{k3}\tau^2 + b_{k4}\tau^3, \quad (57)$$

kde k označuje k – té okénko, τ , nabývající hodnot $\tau = -2, -1, 0, 1, 2$ značí hodnoty ve vyhlazovacím okénku, přičemž 0 náleží prostřední hodnotě a \mathbf{b}_k vektor regresních koeficientů $(b_{k1}, b_{k2}, b_{k3}, b_{k4})^T$. Pro odhad parametrů vyrovnávajícího polynomu \mathbf{b}_k použijeme metodu nejmenších čtverců. Budeme tedy minimalizovat výraz (10, s. 128):

$$\sum_{\tau=-2}^2 (y_{k+2+\tau} - \beta_{k1} - \beta_{k2}\tau - \beta_{k3}\tau^2 - \beta_{k4}\tau^3)^2. \quad (58)$$

Pokud spočteme parciální derivace této funkce podle jednotlivých koeficientů a položíme je rovny nule, získáme soustavu normálních rovnic (59) (10, s. 128):

$$\begin{aligned} 5b_{k1} + 10b_{k3} &= \sum_{\tau=-2}^2 y_{t+\tau}, & (59) \\ 10b_{k2} + 34b_{k4} &= \sum_{\tau=-2}^2 \tau y_{t+\tau}, \\ 10b_{k1} + 34b_{k3} &= \sum_{\tau=-2}^2 \tau^2 y_{t+\tau}, \\ 34b_{k2} + 130b_{k4} &= \sum_{\tau=-2}^2 \tau^3 y_{t+\tau}. \end{aligned}$$

Protože nás nejprve zajímá prostřední hodnota vyhlazovacího okénka, která odpovídá polynomu $b_{k1} + b_{k2}\tau + b_{k3}\tau^2 + b_{k4}\tau^3$ v čase $\tau = 0$, stačí nám pro odhad trendu vyjádřit odhad koeficientu b_{k1} . Z první a třetí rovnice tak dostáváme výraz (60):

$$b_{k1} = \frac{1}{35}(-3y_k + 12y_{k+1} + 17y_{k+2} + 12y_{k+3} - 3y_{k+4}). \quad (60)$$

Vyrovnaná hodnota časové řady, označená jako $\hat{\eta}_{k+2}$ je tedy v k -tém okénku rovna výrazu (61) (10, s. 129):

$$\hat{\eta}_{k+2} = \frac{1}{35}(-3y_k + 12y_{k+1} + 17y_{k+2} + 12y_{k+3} - 3y_{k+4}). \quad (61)$$

Pokud požíváme délku vyhlazovacího okénka $2m + 1 = 5$, vzorec (61), který jsme odvodili, nám neříká nic o prvních m a posledních m hodnotách. Abychom mohli vyrovnat i tyto hodnoty, potřebujeme ze soustavy rovnic (59) vyjádřit také b_{k2} , b_{k3} , b_{k4} . Přepišme si soustavu rovnic (59) do maticového tvaru, který zachycují vztahy (62) (10, s. 128–129):

$$\mathbf{A}\mathbf{b}_k = \mathbf{c}_k, \quad (62)$$

kde

$$\mathbf{A} = \begin{pmatrix} 5 & 0 & 10 & 0 \\ 0 & 10 & 0 & 34 \\ 10 & 0 & 34 & 0 \\ 0 & 34 & 0 & 130 \end{pmatrix}, \quad (63)$$

$$\mathbf{b}_k = \begin{pmatrix} b_{k1} \\ b_{k2} \\ b_{k3} \\ b_{k4} \end{pmatrix}, \quad (64)$$

$$\mathbf{c}_k = \left(\sum_{\tau=-2}^2 y_{t+\tau}, \sum_{\tau=-2}^2 \tau y_{t+\tau}, \sum_{\tau=-2}^2 \tau^2 y_{t+\tau}, \sum_{\tau=-2}^2 \tau^3 y_{t+\tau} \right)^T \quad (65)$$

Námi hledané odhady koeficientů polynomu poté určíme pomocí rovnice (66) (10, s. 129):

$$\mathbf{b}_k = \mathbf{A}^{-1}\mathbf{c}_k. \quad (66)$$

První a druhou vyrovnanou hodnotu časové řady určíme z polynomu (67) (10, 131):

$$\eta(\tau, \mathbf{b}_1) = b_{11} + b_{12}\tau + b_{13}\tau^2 + b_{14}\tau^3, \quad (67)$$

kde za τ dosadíme nejdříve $\tau = -2$ pro první hodnotu a $\tau = -1$ pro druhou hodnotu. Obdobně získáme předposlední a poslední hodnotu, kdy do polynomu (68) (10, s. 132):

$$\eta(\tau, \mathbf{b}_{n-4}) = b_{n-4,1} + b_{n-4,2}\tau + b_{n-4,3}\tau^2 + b_{n-4,4}\tau^3, \quad (68)$$

dosadíme za $\tau = 1$ a $\tau = 2$.

Pomocí polynomu $\eta(\tau, \mathbf{b}_{n-4})$, který vyhlazuje poslední pěti, lze také určit predikci budoucích hodnot. Tu získáme prostým dosazením do daného polynomu za $\tau = 3$. Zde je ovšem nutné brát predikovanou hodnotu s jistou rezervou, jelikož daný polynom nemusí vždy přesně popisovat budoucí vývoj dané časové řady (10, s. 133).

O velikosti vyhlazovacího okna se rozhodujeme na základě subjektivního posouzení, obecně platí, že čím větší zvolíme průměr vyhlazovacího okna, tím dochází k většímu vyhlazení časové řady.

2.3 Regresní analýza

Jelikož značná část metody časových řad, o které pojednává kapitola 1.2, vychází z regresní analýzy, rozhodl jsem se pojmout některé důležité části teorii regresní analýzy komplexně a věnovat jim samostatnou kapitolu.

Jako první člověk, který se zabýval regresí, a z jehož prací dnes známe samotný pojem regrese, byl antropolog Francis Galton. Ten se v druhé polovině 19. století zajímal o dědičnost mezi otci a syny, konkrétně o výšku otce a jeho prvorozeného syna. Z analýzy dat získaných pozorováním zjistil, že vysocí otcové mívají i vysoké syny, kteří jsou ovšem v průměru menší než jejich otcové. Stejně tak malí otcové mívají malé syny, kteří jsou v průměru vyšší než jejich otcové. Právě tato tendence návratu další generace k dlouhodobému průměru byla nazvána regresí (16, s.20).

Regresní analýza má za úkol zkoumat závislost, která vzniká mezi nezávisle proměnnou x a závisle proměnnou y . V praxi to znamená následující: pro různě nastavené hodnoty nezávisle proměnné x sledujeme hodnoty závisle proměnné y , přičemž provádíme alespoň

3 a více pozorování. Tímto získáme n uspořádaných dvojic (x_i, y_i) , kde $i = 1, 2, 3, \dots, n$, $n > 2$ a y_i označuje hodnotu závisle proměnné přiřazenou pozorováním k nezávisle proměnné x_i v i -tém pokusu (10, s. 79–80).

Závislost mezi nezávisle proměnnou x a závisle proměnnou y lze vyjádřit vztahem (69)

$$y = \varphi(x), \quad (69)$$

přičemž $\varphi(x)$ je pro nás neznámá funkce. V důsledku působení různých náhodných vlivů a neuvažovaných činitelů, tzv. šumů, které ovlivňují pozorování, nedostaneme při opakovaném měření, jak bychom čekali, pro určité hodnoty x_i hodnoty y_i podle předpisu $y_i = \varphi(x_i)$, ale obecně jinou hodnotu. Jinými slovy, budeme-li pro stejná x_i opakovaně sledovat hodnoty y_i , dostaneme pokaždé jiné hodnoty závisle proměnné. Závisle proměnná y se tedy chová jako náhodná veličina, označovaná Y (10, s. 79–80).

Budeme-li posuzovat velikost odchylek, můžeme brát v úvahu různá vysvětlení jejich příčiny:

- odchylky mohou být vyvolány dalšími faktory, které ovlivňují zjištěné hodnoty y ,
- odchylky jsou způsobeny přirozenou variabilitou výběrových dat kolem jejich střední hodnoty,
- odchylky vznikají z důvodu špatně zvoleného modelu (17, s. 125).

Říkáme tedy, že závislost mezi proměnnými x a y je do jisté míry ovlivněna šumem, který značíme písmenem e . Šum lze definovat jako náhodnou veličinu, která vyjadřuje vliv náhodných a neuvažovaných proměnných na výsledky měření, jejíž střední hodnota je podle předpokladu rovna nule, ozn. $E(e) = 0$. To znamená, že se při měření nevyskytují žádné systematické chyby, způsobené například nepřesností přístroje nebo chybou člověka, který provádí měření, a výchyly od skutečných hodnot se mohou vyskytovat v kladném i záporném smyslu kolem této střední hodnoty $E(e)$ (10, s. 80).

Funkce $\eta(x)$, nazývaná *regresní funkcí*, je funkcí nezávisle proměnné x , která obsahuje neznámé parametry, tzv. *regresní koeficienty*, β_1, \dots, β_p , přičemž $p \geq 1$. V terminologii regresní analýzy se veličina x nazývá *vysvětlující proměnná* a veličina Y *vysvětlovaná proměnná* (10, s. 80).

Shrneme-li výše uvedené poznatky, lze definovat princip regresní analýzy jako úlohu určit pro zjištěná data $(x_i, y_i), i = 1, 2, \dots, n$ vhodnou funkci $\eta(x; \beta_1, \beta_2, \dots, \beta_p)$ a určit koeficienty $\beta_1, \beta_2, \dots, \beta_p$ tak, aby tato funkce co nejlépe vyrovnala hodnoty y_i .

2.3.1 Metoda nejmenších čtverců

Pro zavedení této metody předpokládejme, že jsme provedli n pozorování, ze kterých jsme získali hodnoty $(x_i, y_i), i = 1, 2, \dots, n$. Naším cílem je určit $\varphi(x)$ lineární kombinací m bázevých funkcí pro $m \leq n$, to popíšeme vztahem (70) (18, s. 54):

$$\varphi(x) \approx b_1\varphi_1(x) + b_2\varphi_2(x) + \dots + b_m\varphi_m(x) := \eta_m(x). \quad (70)$$

Uurčíme si matici \mathbf{A} , kterou nazveme *návrhovou maticí*, která má předpis (71) (18, s.54)

$$\mathbf{A} = \begin{pmatrix} \varphi_1(x_1) & \cdots & \varphi_m(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(x_n) & \cdots & \varphi_m(x_n) \end{pmatrix}. \quad (71)$$

Dále si vyjádříme odchylky neboli rezidua vztahem (72) (18, s. 54):

$$e_i = y_i - \eta_m(x_i) = y_i - \sum_{j=1}^m \varphi_j(x_i)b_j = y_i - \sum_{j=1}^m a_{ij}b_j, \text{ kde } i = 1, 2, \dots, n. \quad (72)$$

Jak již bylo řečeno, metoda nejmenších čtverců spočívá v minimalizaci součtu kvadrátů reziduí, tzn. (73) (18, s. 55):

$$\sum_{i=1}^n r_i^2 \rightarrow \min. \quad (73)$$

Chceme-li vyřešit minimalizační úlohu (73), musí pro ni být splněna nutná podmínka extrému, která udává, že má-li funkce f v bodě a lokální extrém, potom její derivace je rovna 0 ($f' = 0$), tzn. platí (74) (18, s. 56):

$$\frac{\partial}{\partial x_k} \sum_{i=1}^n \left(y_i - \sum_{j=1}^m a_{ij}x_j \right)^2 = 0, k = 1, 2, \dots, m. \quad (74)$$

Po derivování dostaneme výraz (75) (18, s. 56):

$$2 \sum_{i=1}^n \left(y_i - \sum_{j=1}^m a_{ij} x_j \right) (-a_{ik}) = 0 \quad (75)$$

a po úpravě výraz (76) (18, s. 56):

$$\sum_{j=1}^m \left(\sum_{i=1}^n a_{ik} a_{ij} \right) x_j = \sum_{i=1}^n a_{ik} y_i, k = 1, 2, \dots, m. \quad (76)$$

Maticově zapsáno výrazem (77) (18, s. 56):

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y}. \quad (77)$$

Tento zápis lze poté přepsat následovně jako vztah (78) (18, s. 56):

$$\begin{pmatrix} (\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_1) & \dots & (\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_m) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_1) & \dots & (\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_m) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} (\boldsymbol{\varphi}_1, \mathbf{y}) \\ \vdots \\ (\boldsymbol{\varphi}_m, \mathbf{y}) \end{pmatrix}, \quad (78)$$

kde $(\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_j) = \sum_{i=1}^n \varphi_k(t_i) \varphi_j(t_i)$ a $(\boldsymbol{\varphi}_k, \mathbf{y}) = \sum_{i=1}^n \varphi_k(t_i) y_i$ jsou skalárními součiny vektorů $\boldsymbol{\varphi}_k$, $\boldsymbol{\varphi}_j$ a $\boldsymbol{\varphi}_j$, \mathbf{y} .

Abychom se vyvarovali případům, kdy má úloha minimalizace nekonečně mnoho řešení, volíme funkce φ_i tak, aby sloupce matice \mathbf{A} byly lineárně nezávislé (18, s. 56).

2.3.2 Testování hypotéz

Abychom otestovali, že všechny regresní parametry dané regresní funkce jsou skutečně statisticky významné a do regresního modelu patří, bude v praktické části provedeno testování hypotéz o nulové hodnotě regresních parametrů.

Testování parametrů regresní funkce

Při testování hypotéz (na určité hladině významnosti α) o regresním parametru testujeme hypotézu (79) (9, s. 232):

$$H_0: \beta_i = \beta_i^*, \quad (79)$$

proti alternativní hypotéze (80) (9, s. 232):

$$H_1: \beta_j \neq \beta_i^*. \quad (80)$$

Označme T , hodnotu testovacího kritéria, výrazem (81) (9, s. 232):

$$T = \frac{|b_i - \beta_i^*|}{s_{b_i}}. \quad (81)$$

Přitom pro kritický obor, podle něhož určujeme, zda přijímáme nulovou hypotézu, platí (82) (9, s. 233):

$$T \geq t_{1-\frac{\alpha}{2}}(n-p), \quad (82)$$

kde s_{b_i} je směrodatná chyba regresního koeficientu b_i , α hladina významnosti, jejíž hodnota se často udává jako $\alpha = 0,05$ (tato hodnota bude uvažována i v praktické části), n je počet pozorování a p počet regresních parametrů. Přitom $t(n-p)$ je kvantil Studentova t -rozdělení s $(n-p)$ stupni volnosti, jehož hodnotu zjistíme ze statistických tabulek (9, s. 233).

V praktické části budeme jako hodnotu β_i^* uvažovat $\beta_i^* = 0$, jinými slovy budeme testovat, zda hodnota regresního parametru je různá od nuly neboli zda daný regresní parametr je statisticky významný a do modelu patří či nikoliv.

Po zavedení výše zmíněného do vzorců (79) až (82) obdržíme následující hypotézu (83), kterou budeme testovat (9, s. 233):

$$H_0: \beta_i = 0, \quad (83)$$

oproti alternativní hypotéze (84) (9, s. 233):

$$H_1: \beta_j \neq 0 \quad (84)$$

a za kritický obor považujeme (85) (9, s. 233):

$$\frac{|b_i|}{s_{b_i}} \geq t_{1-\frac{\alpha}{2}}(n-p). \quad (85)$$

Jednotlivé výše zmíněné kapitoly týkající se časových řad a statistických metod obecně jsou pro hlubší studium podrobněji rozepsány například v knize „Finanční ekonometrie“ od profesora Tomáše Cipry.

3 VLASTNÍ NÁVRHY ŘEŠENÍ

Výše popsanými metodami se v praktické části se budu zabývat vývojem vybraných produktů **neživotního pojištění**, a to *pojištěním budov, pojištěním domácností a havarijním pojištěním*. Časový vývoj bude uvažován nejčastěji v období roků 2000 až 2016. Pro konstrukci těchto modelů využiji veřejně dostupná data České asociace pojišťoven (ČAP). Ke zpracování problematiky využiji výpočetní prostředí a vizualizační nástroje software **STATISTICA** a systému **MATLAB**. U každého vybraného produktu budu postupovat v několika krocích:

- a) Vizualizace zjištěných (naměřených) dat ze tří pohledů (hladin předepsaného pojistného, počtu uzavřených smluv a uskutečnění pojistného plnění). U havarijního pojištění z důvodu nemožnosti dohledání zdrojových dat nebude uváděn počet uzavřených smluv.
- b) Identifikace trendu vývoje dat (informativní test, regresní model, index determinace, ověření statistické významnosti parametrů regresního modelu testováním), a jeho vizualizace.
- c) Predikce vývoje (bodově, intervalově).

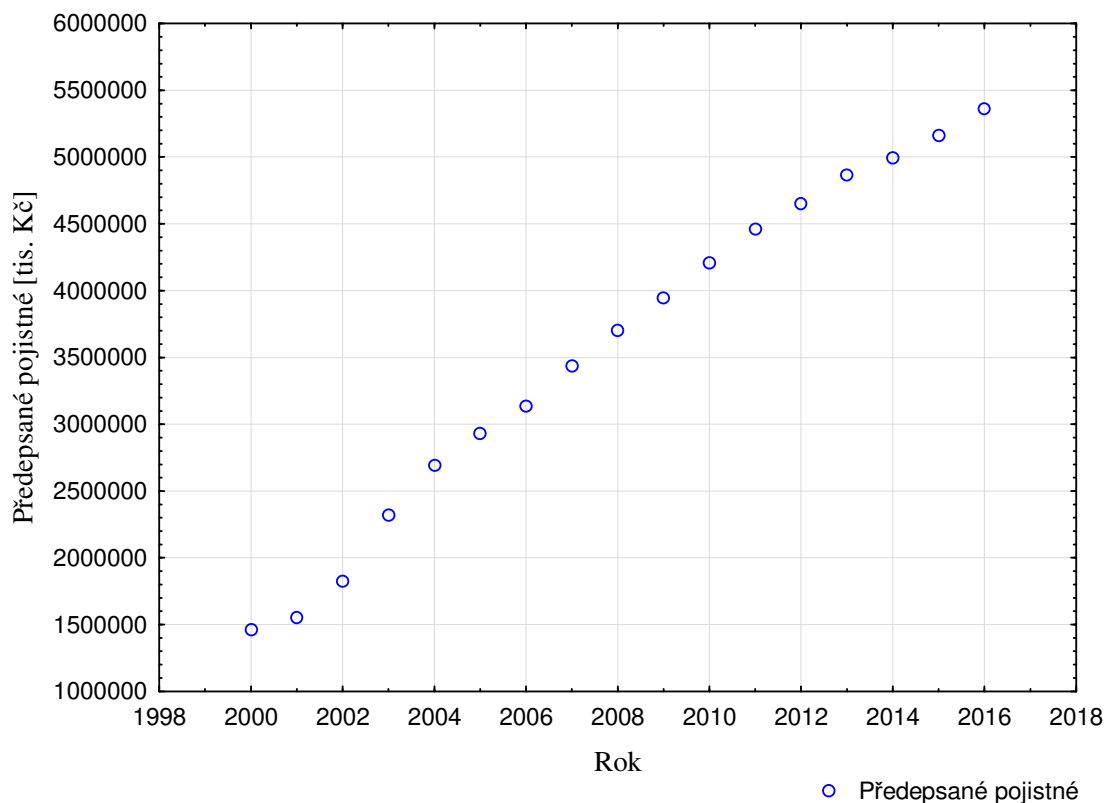
3.1 Pojištění budov

Nejprve jsou zpracována data v období roků 2000 až 2016 týkající se oblasti pojištění budov (zdrojová data viz Příloha 1).

Předepsané pojistné

Vizualizace monitorovaných dat

Nejdříve zaneseme data předepsaného pojistného do grafu, viz Graf 1.



Graf 1: Pojištění budov – Předepsané pojistné (Zdroj: Vlastní zpracování v software STATISTICA)

Informativní (orientační) test

Vizuálním vyhodnocením lze usoudit, že trend, který bude popisovat danou časovou řadu, může být *lineární*, *parabolický*, případně i *exponenciální* na daném časovém intervalu. Abychom se rozhodli, který z těchto trendů je nejvhodnější, *porovnáme indexy determinace* pro všechny tři varianty a dále ověříme *hypotézu o nulové hodnotě parametrů vybrané trendové funkce*.

Nejdříve transformujeme časovou proměnnou $t = 2000, 2001, \dots, 2016$ tak, aby její součet byl roven nule, tzn. $t' = -8, -7, \dots, 7, 8$.

Regresní trendy a index determinace (vlastní zpracování v software STATISTICA)

- *Lineární trend*

Tento trend je pak vyjádřen následujícím předpisem regresní funkce:

$$\hat{y}_t = 3571226,2 + 253130,8 \cdot t', \text{ kde } t' = -8, -7, \dots, 7, 8,$$

Příslušný index determinace vyjádřený v procentech je roven $R^2 = 98,71 \%$.

- *Parabolický trend*

Tento trend je vyjádřen předpisem

$$\hat{y}_{t'} = 3711837,4 + 253130,8t' - 5858,8t'^2, \text{ kde } t' = -8, -7, \dots, 7, 8.$$

Příslušný indexem determinace vyjádřený v procentech je roven $R^2 = 99,86 \%$.

- *Exponenciální trend*

Tento trend je vyjádřen předpisem

$$\hat{y}_{t'} = 3314177,4 * 1,0834899^{t'}, \text{ kde } t' = -8, -7, \dots, 7, 8.$$

Příslušný index determinace vyjádřený v procentech je roven $R^2 = 92,51 \%$.

Shrnutí: Porovnáním všech tří indexů determinace lze tedy prohlásit, že parabolický trend (tj. kvadratická závislost) má nejvyšší hodnotu indexu determinace, a tedy prokládá monitorovaná data nejlépe. Pro popis této časové řady je nejpřesnější *parabolický trend*.

Test hypotézy

Dále je nutné ověřit, zda všechny parametry parabolické funkce jsou statisticky významné a zda i přes nižší index determinace nedáme přednost jinému trendu. Toto ověření provedeme pomocí testu hypotéz o parametrech regresní funkce.

V souladu s kapitolou číslo 2.3.2 budeme proto testovat následující hypotézy na hladině významnosti $\alpha = 0,05$.

Nejprve určíme nulovou hypotézu

$$H_0: \beta_i = 0,$$

poté alternativní hypotézu

$$H_1: \beta_i \neq 0.$$

Pro přijetí či zamítnutí nulové hypotézy potřebujeme znát hodnoty testového kritéria T . Jejich hodnoty pro jednotlivé parametry prezentuje Tabulka 2.

Tabulka 2: Hodnoty testového kritéria T – Parabolický trend (Zdroj: Vlastní zpracování v software STATISTICA)

Parametr	T
3711837,4	138,4
253130,8	69,6
-5858,8	-7

Abychom rozhodli, zda nulovou hypotézu přijmeme či zamítneme, musíme zjistit, zda T náleží nebo nenáleží do kritického oboru. Pro určení kritického oboru uvažujeme hodnoty $\alpha = 0,05$, $n = 17$, $p = 3$, kde α značí hladinu významnosti, n počet měření (počet sledovaných roků) a p počet parametrů regresní funkce. Ze statistických tabulek zjistíme, že příslušný kvantil Studentova t -rozdělení se rovná hodnotě 2,145. Tedy kritický obor je roven (9, s. 397):

$$W_{\alpha} = \{T; |T| \geq 2,145\}.$$

Na základě hodnot z Tabulky 2 lze prohlásit, že všechna testová kritéria T (pro všechny tři parametry parabolického trendu) náleží do kritického oboru. Z toho plyne, že pro všechny parametry zamítáme nulovou hypotézu a přijímáme hypotézu alternativní. Jinými slovy *zamítáme, že regresní parametry mají nulovou hodnotu, jsou pro nás tedy statisticky významné a v trendové funkci je ponecháme.*

Predikce

Dále určíme predikci vývoje tohoto ukazatele (předepsaného pojistného – pojištění budov) pro roky 2017, 2018 a 2019. Budoucí hodnoty budeme nejdříve odhadovat bodovou předpovědí, poté, s 95 % pravděpodobností určíme intervalové předpovědi.

- Bodová předpověď

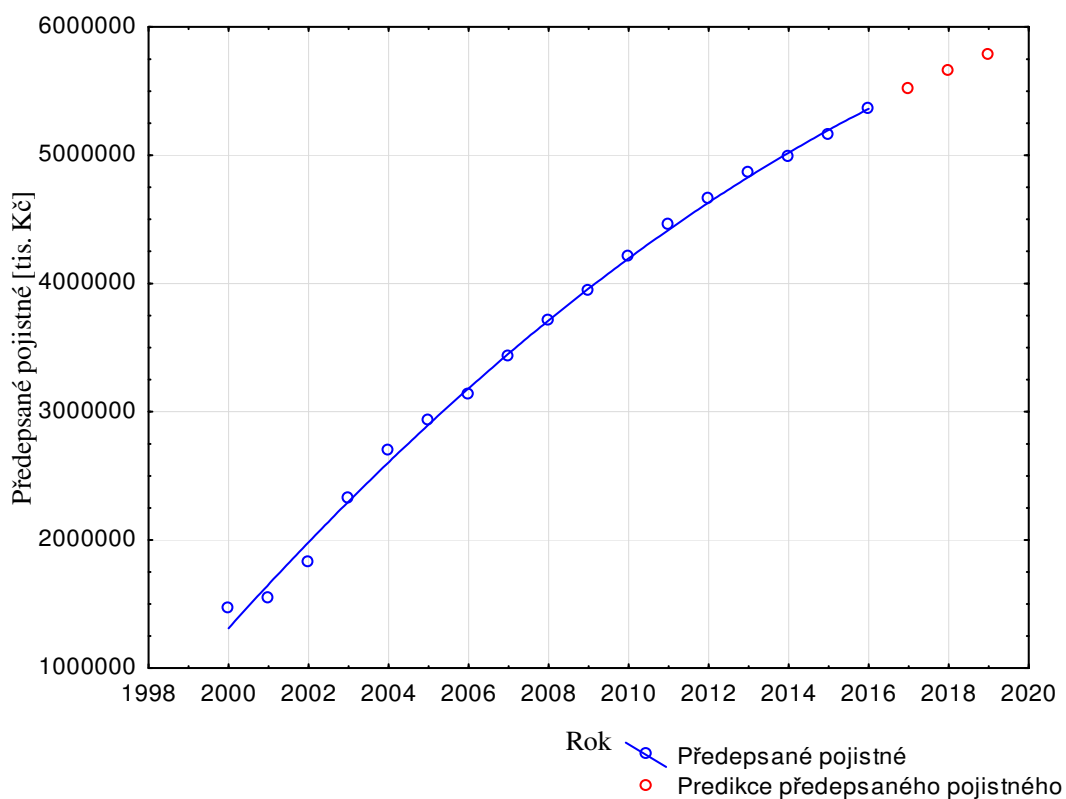
Bodovou předpověď získáme jednoduše dosazením do předpisu parabolického trendu za $t' = 9, 10, 11$ (což odpovídá rokům 2017, 2018, 2019 při zavedené transformaci proměnné t). Získáme hodnoty 5 515 451,8 tis. Kč, 5 657 265,4 tis. Kč a 5 787 361,4 tis. Kč (pořadí odpovídá pořadí predikovaných roků).

Pro jednotlivé roky jsou hodnoty předepsaného pojistného zapsány v Tabulce 3.

Tabulka 3: Pojištění budov – Bodová predikce hodnot předepsaného pojistného (Vlastní zpracování v software STATISTICA)

Rok	Predikce
2017	5 515 451 800
2018	5 657 265 400
2019	5 787 361 400

Tuto predikci spolu s analyzovanými daty prezentuje Graf 2.



Graf 2: Pojištění budov – Předepsané pojistné, proložená data s predikcí (Zdroj: vlastní zpracování v software STATISTICA)

- *Intervalová předpověď*

Kromě bodového odhadu budoucích hodnot je také vhodné určit pravděpodobnostní intervaly (intervalovou předpověď). K tomuto jsem využil kód „Intervaly“, který jsem naprogramoval v systému MATLAB. Jako vstupní funkci f zadáme do programu funkci výše vypočteného parabolického trendu a jako matici A zadáme hodnoty předepsaného pojistného. Rok, který do programu zadáváme, odpovídá predikovanému období. Tedy pro predikci na rok 2017 dosadíme do proměnné t' transformovanou hodnotu 9, na rok

2018 hodnotu 10 a na rok 2019 hodnotu 11. Vypočtené rozpětí těchto intervalů je pak pro jednotlivé roky uvedeno v Tabulce 4.

Tabulka 4: Předpovědní intervaly pro předepsané pojistné – Pojištění budov (Zdroj: Vlastní zpracování v systému MATLAB)

Rok	Interval
2017	(5 347 663 800; 5 683 239 700)
2018	(5 489 745 100; 5 824 785 700)
2019	(5 620 531 300; 5 954 191 500)

Protože jsme uvažovali hladinu významnosti $\alpha = 0,05$, můžeme říci, že hodnoty předepsaného pojistného budou v daných rocích náležet do těchto intervalů s pravděpodobností 95 %.

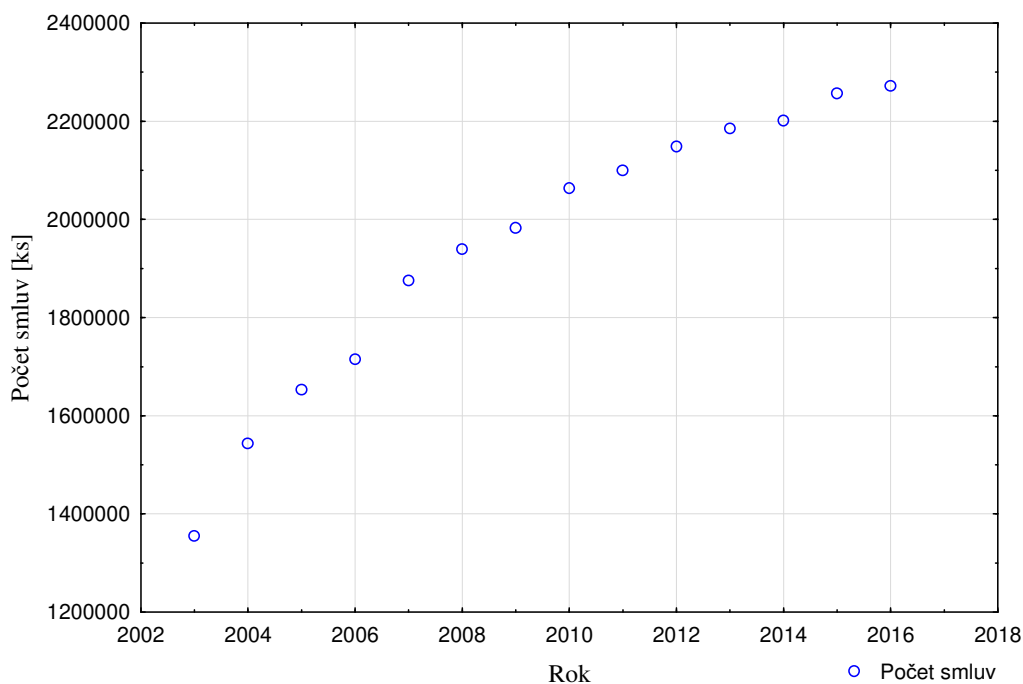
Počet (uzavřených) smluv

Poznámka: Z důvodu analogicky vedených úvah a rozsahu této práce další komentáře zstručním, případně uvedu především výsledná zjištění.

Vizualizace monitorovaných dat

Poznámka: V tomto případě ČAP uváděla hodnoty až od roku 2003.

Počet smluv u pojištění budov v období roků 2003 až 2016 je znázorněn v Grafu 3.



Graf 3: Pojištění budov – Počet smluv (Zdroj: Vlastní zpracování v software STATISTICA)

Informativní (orientační) test

Na základě vizualizace hodnot v Grafu 3 lze usoudit, že časovou řadu lze nejlépe vystihnout pomocí *parabolického trendu*. Abychom se o tom přesvědčili, zjistíme indexy determinace také pro zbylé dva trendy (lineární a exponenciální).

Regresní trendy a indexy determinace (vlastní zpracování v software STATISTIKA)

Proměnná $t' = -13, -11, \dots, 11, 13$ odpovídá transformaci proměnné $t = 2003, 2004, \dots, 2016$.

- *Lineární trend*

$$\hat{y}_{t'} = 1949467,5 + 32877,8 \cdot t', \text{ kde } t' = -13, -11, \dots, 11, 13 .$$

Index determinace, vyjádřený v procentech, je roven $R^2 = 92,75 \%$.

- *Parabolický trend*

$$\hat{y}_{t'} = 2028851,9 + 32877,8 \cdot t' - 1221,3 \cdot t'^2, \text{ pro } t' = -13, -11, \dots, 11, 13.$$

Indexem determinace vyjádřený v procentech je roven $R^2 = 99,65 \%$. (vždy dopsat že v procentech)

- *Exponenciální trend*

$$\hat{y}_{t'} = 1928258 \cdot 1,01788227^{t'}, \text{ kde } t' = -13, -11, \dots, 11, 13.$$

Index determinace vyjádřený v procentech je roven $R^2 = 88,93 \%$.

Protože index determinace pro *parabolický trend* je vyšší než pro ostatní trendy, můžeme prohlásit, že je pro data nejvhodnější.

Test hypotézy

Pomocí hypotézy o nulové hodnotě regresního parametru opět zjistíme statistickou významnost jednotlivých parametrů regresního modelu. Protože postup testování této hypotézy (nulová hypotéza $H_0: \beta_i = 0$, alternativní hypotéza $H_1: \beta_i \neq 0$, hladina významnosti $\alpha = 0,05$) byl již podrobně popsán v části předchozí (pro předepsané pojistné), uvedu dále pouze shrnující Tabulku 5 s výslednými hodnotami příslušných testových kritérií.

Tabulka 5: Hodnoty testového kritéria T – Parabolický trend (Zdroj: Vlastní zpracování v software STATISTICA)

Parametr	T
2028851,9	195,3
32877,8	38,4
-1221,3	-10,2

Pro hladinu významnosti $\alpha = 0,05$ a stupeň volnosti $n = 11$ získáme ze statistických tabulek pro kvantil Studentova t rozdělení hodnotu 2,201. Kritický obor tedy je roven (9, s. 397):

$$W_\alpha = \{T; |T| \geq 2,201\}.$$

Protože všechny hodnoty T z Tabulky 5 náleží do kritického oboru, můžeme prohlásit, že parametry jsou statisticky významné. Zamítneme tedy nulovou hypotézu a přijmeme hypotézu alternativní, tj. *zamítáme, že regresní parametry mají nulovou hodnotu, jsou pro nás tedy statisticky významné a v trendové funkci je ponecháme.*

Predikce

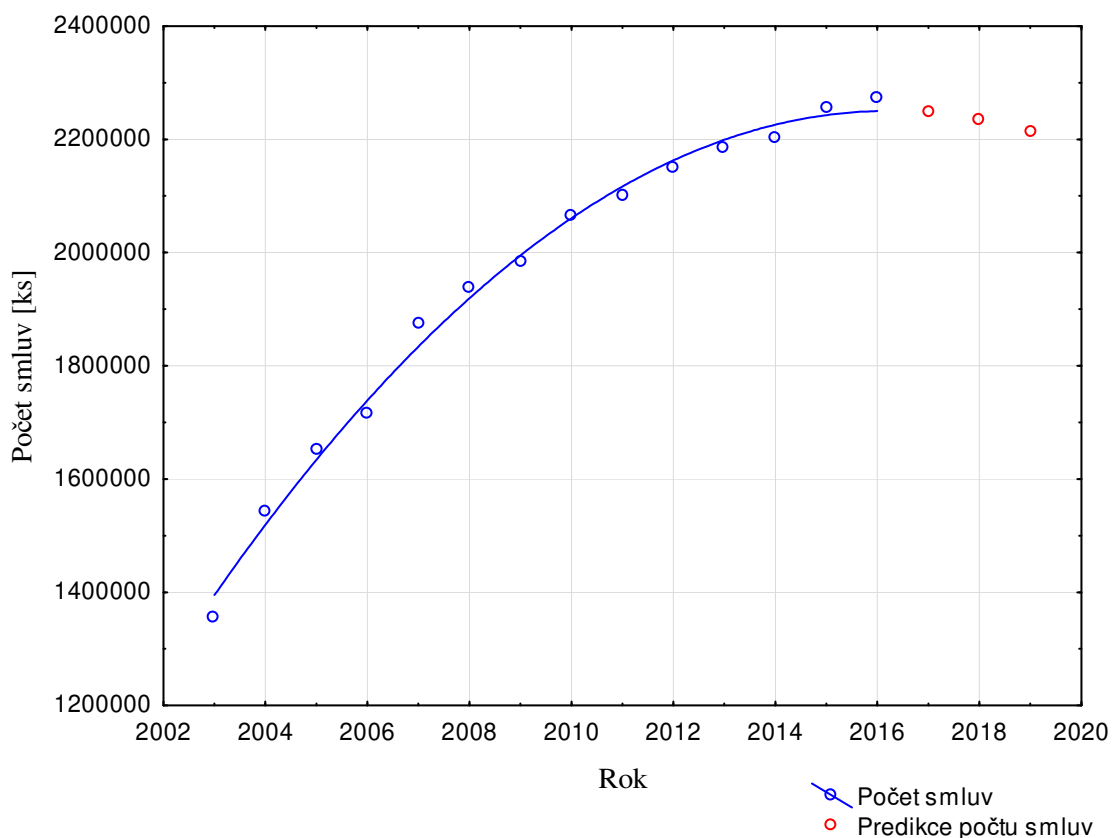
- Bodová předpověď

Chceme-li určit predikci pro roky 2017, 2018 a 2019, dostáváme po dosazení za $t' = 15, 17, 19$ hodnoty 2 247 226 smluv, 2 234 819 smluv a 2 212 641 smluv (v pořadí predikovaných roků). Pro přehlednost jsou tyto hodnoty zapsány v Tabulce 6.

Tabulka 6: Pojištěná budov – Bodová predikce počtu smluv (Zdroj: Vlastní zpracování v software STATISTIKA)

Rok	Predikce
2017	2 247 226
2018	2 234 819
2019	2 212 641

Tato predikce, společně s proložením dat, je zakreslena v Grafu 4.



Graf 4: Pojištění budov – Počet smluv, proložená data s predikcí (Zdroj: vlastní zpracování v software STATISTICA)

- *Intervalové předpovědi*

Využijeme program „Intervaly“ (MATLAB), kde jako vstupní funkci použijeme funkci parabolického trendu, prvky matice A jsou příslušné monitorované (naměřené) hodnoty a do proměnné rok dosadíme postupně hodnoty 15, 17, 19 (vypočtené výše uvedenou transformací). Výsledné intervalové předpovědi prezentuje Tabulka 7.

Tabulka 7: Předpovědní intervaly pro počet smluv – Pojištění budov (Zdroj: Vlastní zpracování v systému MATLAB)

Rok	Interval
2017	(2 169 609; 2 324 844)
2018	(2 130 901; 2 338 736)
2019	(2 070 403; 2 354 878)

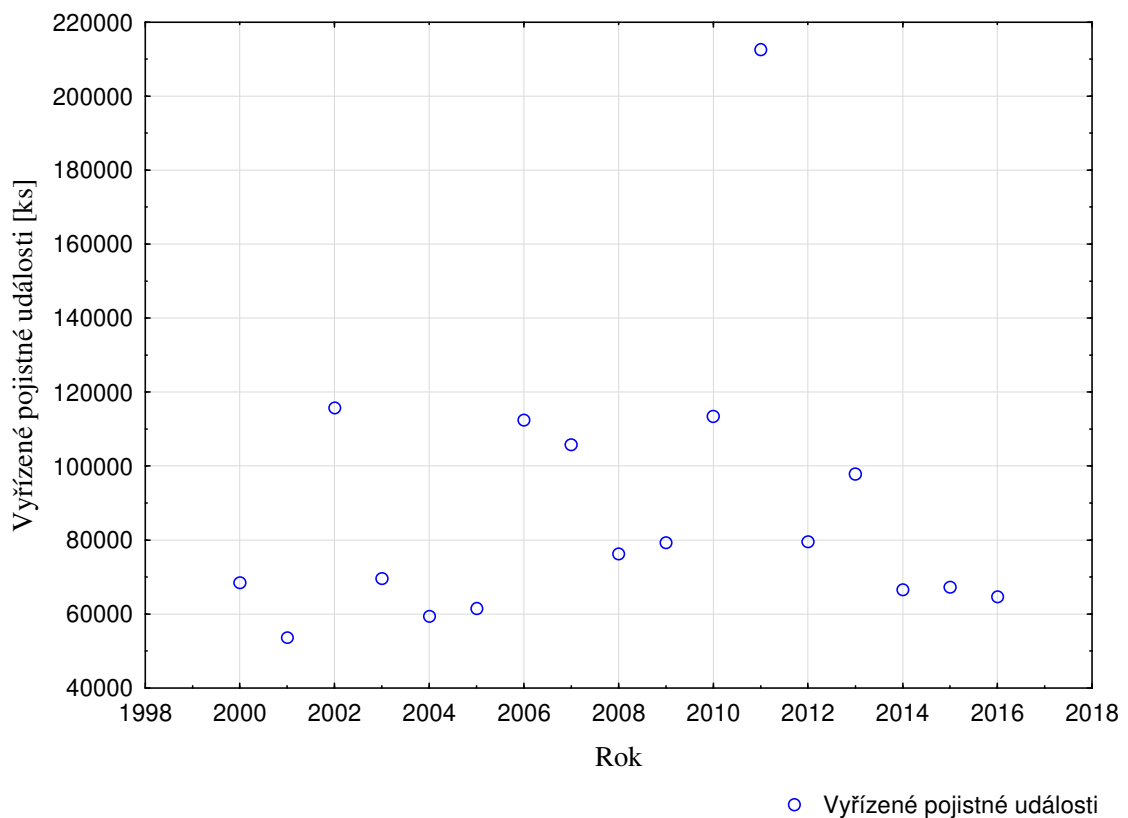
Z obdobných důvodů jako jsme popsali výše ($\alpha = 0,05$) s pravděpodobností 95 % můžeme tvrdit, že počty uzavřených smluv pro pojištění budov se v predikovaných rocích 2017, 2018 a 2019 budou vyskytovat v hodnotách těchto intervalů.

Počet vyřízených pojistných událostí

Jako poslední budeme zkoumat počet vyřízených pojistných událostí.

Informativní (orientační) test

Již vizualizací monitorovaných dat v Grafu 5 lze určit, že pro popis vývoje časové řady nelze použít jednu konkrétní matematickou křivku, jelikož ta by data neproložila s dostatečnou přesností. Pro dekompozici časové řady tedy použijeme exponenciální vyrovnávání.



Graf 5: Pojištění budov – Vyřízené pojistné události (Zdroj: Vlastní zpracování v software STATISTICA)

Nejprve pomocí funkce „Síťové hledání“ v software STATISTICA vypočteme a určíme nejlépe vyhovující hodnoty konstanty α pro jednoduché exponenciální vyrovnávání (Tabulka 8) a konstant α a γ pro Holtovu metodu (Tabulka 9).

Poznámka: zápis čísla 1,667796E+09 značí hodnot $1,66796 \cdot 10^9$.

Tabulka 8: Síťové hledání hodnoty α (Zdroj: Vlastní zpracování v software STATISTICA)

Alfa	Prům. chyba	Průměr a chyba	Suma mocniny	Průměr mocniny
0,100000	-2544,99	31950,38	2,835252E+10	1,667796E+09
0,200000	-3882,63	33308,36	2,965416E+10	1,744362E+09
0,300000	-4367,55	33864,27	3,061492E+10	1,800878E+09
0,400000	-4350,45	33942,72	3,163940E+10	1,861141E+09
0,500000	-4076,80	33719,00	3,291724E+10	1,936309E+09
0,600000	-3699,88	33272,10	3,450315E+10	2,029597E+09
0,700000	-3308,44	32632,20	3,640596E+10	2,141527E+09
0,800000	-2948,31	31924,79	3,864834E+10	2,273432E+09
0,900000	-2637,86	32277,25	4,130408E+10	2,429652E+09

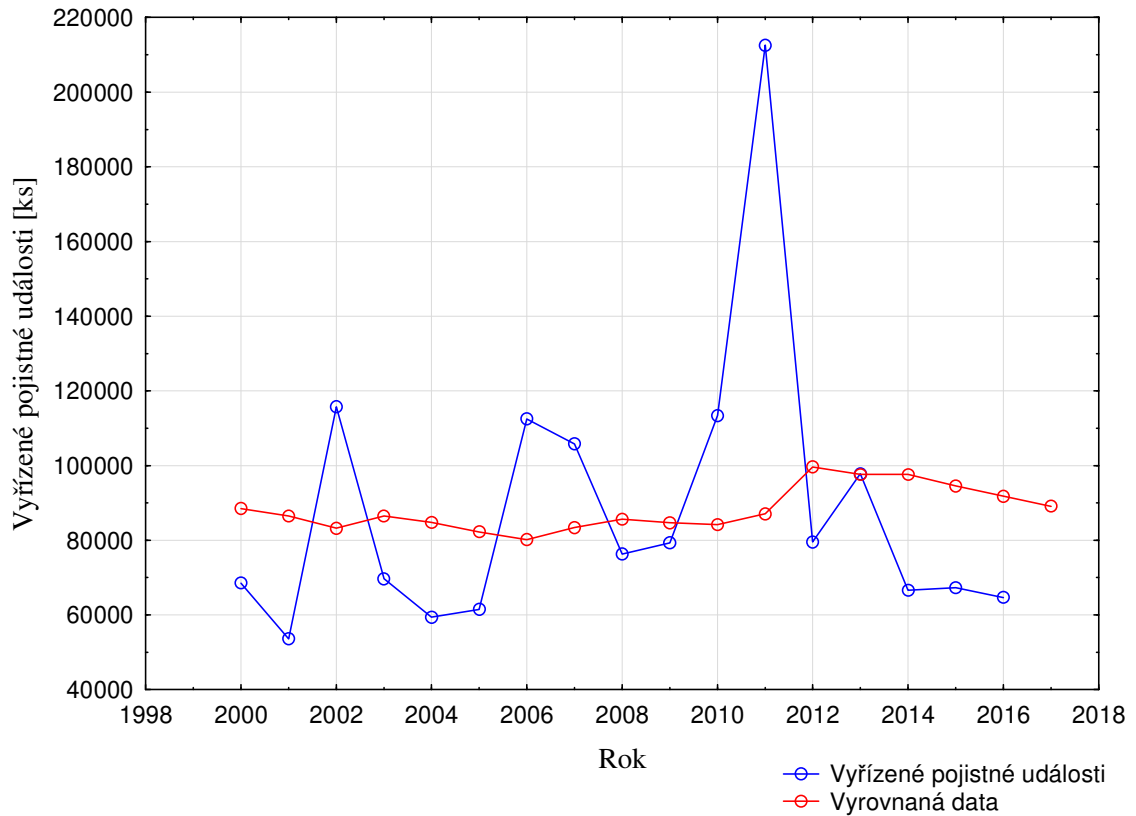
Tabulka 9: Síťové hledání hodnot α a γ (Zdroj: Vlastní zpracování v software STATISTICA)

Alfa	Gama	Prům. chyba	Průměr a chyba	Suma mocniny	Průměr mocniny
0,100000	0,100000	7856,95	26249,10	2,726639E+10	1,603905E+09
0,100000	0,200000	4829,55	28155,36	2,819862E+10	1,658742E+09
0,200000	0,100000	1760,91	30059,28	2,826344E+10	1,662555E+09
0,300000	0,100000	-842,57	31720,30	2,921550E+10	1,718559E+09
0,100000	0,300000	2278,77	29833,20	2,933404E+10	1,725532E+09
0,200000	0,200000	-1265,60	32000,06	2,988615E+10	1,758009E+09
0,400000	0,100000	-1855,17	32360,24	3,020236E+10	1,776610E+09
0,100000	0,400000	139,80	31197,63	3,048035E+10	1,792962E+09
0,300000	0,200000	-3264,55	33254,07	3,076641E+10	1,809789E+09
0,200000	0,300000	-3414,76	33314,69	3,119117E+10	1,834775E+09

V Tabulkách 8 a 9 je vypsáno deset nejlépe vyhovujících hodnot konstant α a γ , přičemž ty nejlepší z nich jsou zvýrazněny červeně. Tedy pro jednoduché exponenciální vyrovnávání použijeme hodnoty $\alpha = 0,1$ a pro Holtovu metodu $\alpha = 0,1$, $\gamma = 0,1$. Suma čtverců reziduí je v případě jednoduchého exponenciálního vyrovnávání rovna $2,5 \cdot 10^{10}$, v případě Holtovy metody $2,73 \cdot 10^{10}$. Z toho plyne, že použijeme jednoduché exponenciální vyrovnávání, což nám napovídá již samotný Graf 5.

Exponenciální vyrovnávání a predikce

V Grafu 6 jsou znázorněna vyhlazená data společně s predikcí, v Tabulce 10 jsou uvedeny jejich konkrétní hodnoty.



Graf 6: Vyřizené pojistné události – Exponenciální vyrovnávání (Zdroj: Vlastní zpracování v software STATISTICA)

Tabulka 10: Vyřizené pojistné události – Vyrovnaná data (Zdroj: Vlastní zpracování v software STATISTICA)

Rok	Data	Vyrovnaná data
2000	68545	88498
2001	53636	86503
2002	115750	83216
2003	69673	86470
2004	59404	84790
2005	61513	82251
2006	112506	80178
2007	105841	83410
2008	76332	85653
2009	79334	84721

2010	113391	84183
2011	212555	87103
2012	79551	99649
2013	97833	97639
2014	66616	97658
2015	67314	94554
2016	64678	91830
2017		89115

Z Tabulky 10 je vidět predikce pro rok 2017, kdy počet vyřízených událostí odpovídá hodnotě 89 115.

Vzhledem k tomu, že jsou zveřejněna pouze data do roku 2016 a daná metoda počítá predikci vždy z naměřené hodnoty předcházejícího období, můžeme odhadovat hodnotu pouze za rok 2017. Ve chvíli, kdy ČAP zveřejní výsledky za rok 2017, můžeme zhodnotit, zda mnou vypočítaná predikce přibližně odpovídá skutečnosti a vypočítat stejným způsobem očekávanou hodnotu za rok 2018.

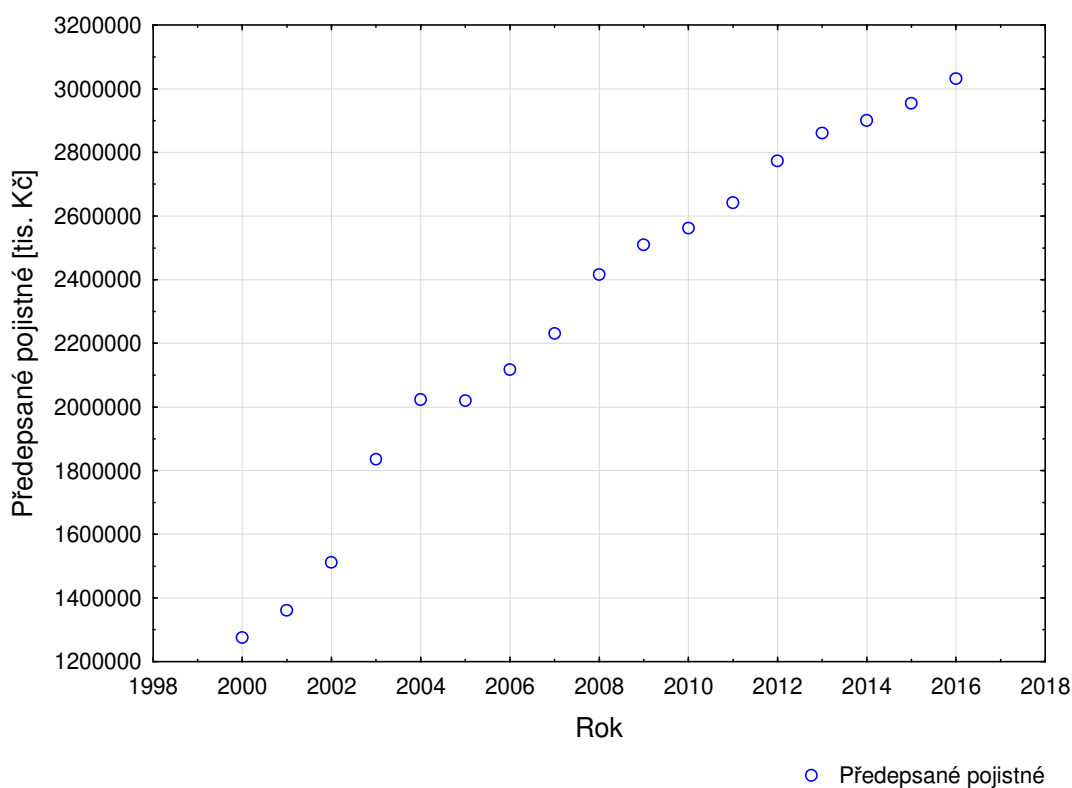
3.2 Pojištění domácností

Použitá data z ČAP jsou uvedena Příloze 2.

Předepsané pojistné

Vizualizace dat

Data popisující vývoj předepsaného pojistného v rocích 2000 až 2016 jsou znázorněna v Grafu 7:



Graf 7: Pojištění domácností – Předepsané pojistné (Zdroj: Vlastní zpracování v software STATISTICA)

Informativní (orientační) test

Podobně jako u předepsaného pojistného u pojištění budov, i zde lze na základě vizualizace dat považovat jako vhodný trend pro popis časové řady trend lineární, parabolický, případně exponenciální. Porovnáme tedy indexy determinace jednotlivých trendů, abychom zjistili, který z nich je pro zkoumaná data nejvhodnější.

Regresní model a index determinace (vlastní zpracování v software STATISTICA)

Stejně jako v případě pojištění budov, ani zde se nelze jednoznačně rozhodnout, který trend je pro danou časovou řadu nejvhodnější, rozhodneme se tedy na základě indexu determinace.

Nejprve transformujeme časovou proměnnou $t = 2000, 2001, \dots, 2016$ použijeme hodnoty $t' = -8, -7, \dots, 7, 8$.

- Lineární trend

$$\hat{y}_{t'} = 2295629,8 + 109558,7 \cdot t', \text{ pro } t' = -8, -7, \dots, 7, 8,$$

index determinace vyjádřený v procentech $R^2 = 96,62 \%$.

- Parabolický trend

$$\hat{y}_t = 2392018,62 + 109558,66 \cdot t' - 4016,2 \cdot t'^2, \text{ kde } t' = -8, -7, \dots, 7, 8,$$

index determinace vyjádřený v procentech $R^2 = 99,55 \%$.

- Exponenciální trend

$$\hat{y}_t = 2222320,1 \cdot 1,052944^{t'}, \text{ kde } t' = -8, -7, \dots, 7, 8,$$

index determinace vyjádřený v procentech $R^2 = 91,43 \%$.

Z výše uvedeného vyplývá, že nejlépe lze časovou řadu popsat pomocí *parabolického trendu*, jelikož nabývá nejvyšší hodnoty indexu determinace.

Test hypotézy

V dalším kroku zjistíme, zda regresní parametry parabolického trendu jsou statisticky významné. K tomu využijeme hypotézu o nulové hodnotě regresního parametru (nulová hypotéza $H_0: \beta_i = 0$, alternativní hypotéza $H_1: \beta_i \neq 0$, hladina významnosti $\alpha = 0,05$), jak bylo popsáno v kapitole 2.3.2.

Hodnoty T statistik pro jednotlivé parametry jsou zapsány v Tabulce 11.

Tabulka 11: Hodnoty testového kritéria T – Parabolický trend (Zdroj: Vlastní zpracování v software STATISTICA)

Parametr	T
2392018,62	114,4457
109558,66	38,6318
-4016,2	-6,1729

Kvantil Studentova t -rozdělení je pro $(17 - 3)$ stupňů volnosti a hladinu významnosti $\alpha = 0,05$ roven 2,145 (9, s. 397). Obdrželi jsme tedy kritický obor

$$W_\alpha = \{T; |T| \geq 2,145\}.$$

Protože hodnota T pro všechny regresní parametry náleží do tohoto oboru, lze prohlásit, že zamítáme nulovou hypotézu o nulové hodnotě regresních parametrů, tzn. *zamítáme, že některý z regresních parametrů je statisticky nevýznamný.*

Predikce

Dále pro pojištění domácností určíme predikci vývoje tohoto ukazatele.

- Bodová predikce

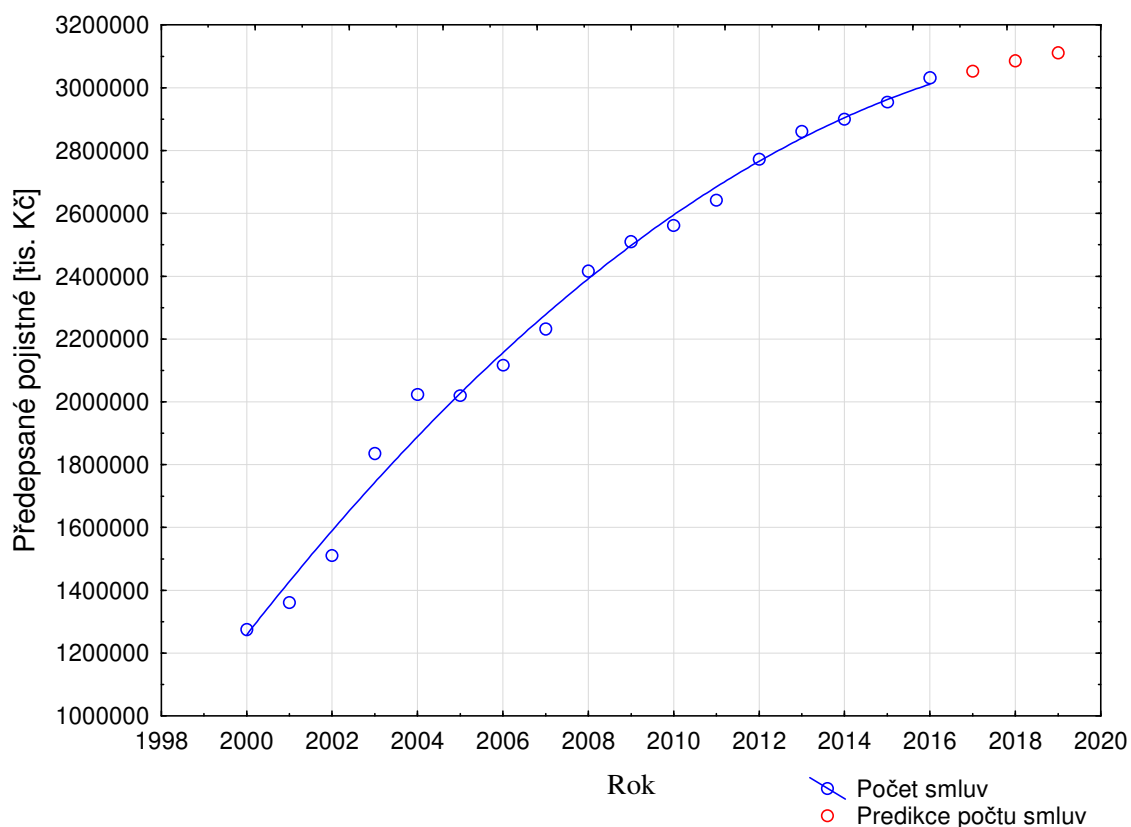
Abychom získali předpověď budoucích hodnot, stačí opět dosadit do dané rovnice za $t = 9, 10, 11$. Tímto získáme hodnoty 3 052 734,4 pro rok 2017, 3 085 985,2 pro rok 2018 a 3 111 203,7 pro rok 2019. Tyto hodnoty jsou vypsány v Tabulce 12.

Tabulka 12: Pojištění domácností – Bodová predikce hodnot předepsaného pojistného (Zdroj: Vlastní zpracování v software STATISTICA)

Rok	Predikce
2017	3 052 734 400
2018	3 085 985 200
2019	3 111 203 700

Předpovědi lze interpretovat následovně: V roce 2017 je očekávaná výše pojistného 3 052 734 400 Kč, v roce 2018 3 085 985 200 Kč a v roce 2019 3 111 203 700 Kč.

Zjištěné předpovědi, společně s původními daty a proložením danou parabolickou funkcí, jsou zobrazeny v Grafu 8.



Graf 8: Pojištění domácností – Předepsané pojistné, proložená data s predikcí (Zdroj: vlastní zpracování v software STATISTICA)

- *Intervalová předpověď*

Kromě předpovězení konkrétní hodnoty pro dané roky určíme také pravděpodobnostní intervaly těchto hodnot. Pomocí programu „Intervaly“ vypočítáme, že pravděpodobnostní intervaly nabývají hodnot z Tabulky 13, hodnoty jsou uváděny v tis. Kč. Jako vstupní funkci v tomto programu použijeme funkci parabolického trendu, matice A je matice naměřených hodnot a za proměnnou rok postupně dosadíme hodnoty 9, 10, 11.

Tabulka 13: Předpovědní intervaly pro předepsané pojistné – Pojištění domácností (Zdroj: Vlastní zpracování v systému MATLAB)

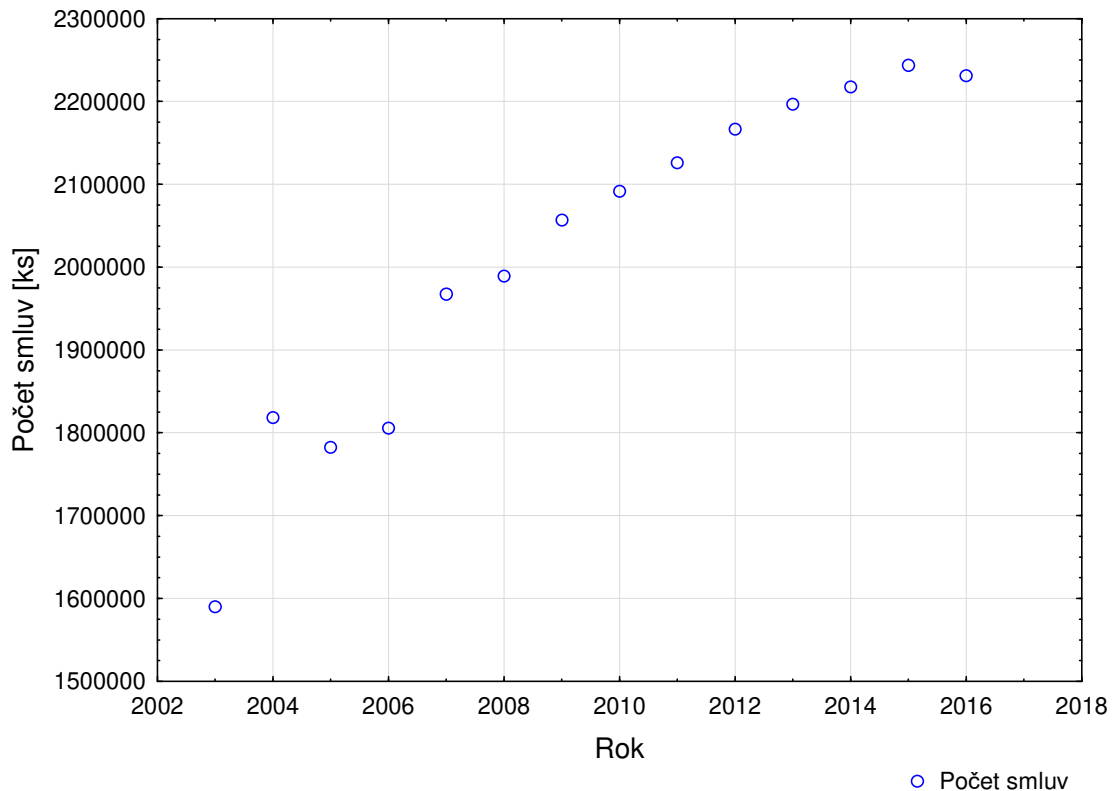
Rok	Interval
2017	(2 921 937 100; 3 183 531 600)
2018	(2 955 396 600; 3 216 573 800)
2019	(2 981 153 100; 3 241 254 300)

I zde jsme jako hladinu významnosti zvolili $\alpha = 0,05$, tedy hodnoty v daných obdobích budou ležet v těchto intervalech s pravděpodobností 95 %.

Počet smluv

Vizualizace monitorovaných dat

Počet uzavřených smluv u pojištění domácností v období roků 2003 až 2016 je znázorněn v Grafu 9.



Graf 9: Pojištění domácností – Počet smluv (Zdroj: Vlastní zpracování v software STATISTICA)

Informativní (orientační) test

Z Grafu 9 vyplývá, že jako vhodný je trend parabolický. Pro potvrzení porovnáme indexy determinace všech trendů.

Regresní trendy a indexy determinace (vlastní zpracování v software STATISTICA)

- Lineární trend

$$\hat{y}_{t'} = 2020328,6 + 23193,4 \cdot t', \text{ kde } t' = -13, -11, \dots, 11, 13,$$

index determinace vyjádřený v procentech je $R^2 = 91,39 \%$.

- *Parabolický trend*

$$y_t = 2071200,9 + 23193,4 \cdot t' - 782,7 \cdot t'^2, \text{ kde } t' = -13, -11, \dots, 11, 13,$$

index determinace vyjádřený v procentech $R^2 = 98,34 \%$.

- *Exponenciální trend*

$$\hat{y}_{t'} = 2010338 \cdot 1,011884^{t'}, \text{ pro který } t' = -13, -11, \dots, 11, 13,$$

index determinace vyjádřený v procentech $R^2 = 88,9 \%$

Protože index determinace pro parabolický trend nabývá nejvyšší hodnoty, jako nejlépe vyhovující trend vybereme trend parabolický.

Test hypotézy

Nyní opět ověříme statistickou významnost jednotlivých regresních parametrů pomocí testování hypotéz o jejich nulové hodnotě ($H_0: \beta_i = 0, H_1: \beta_i \neq 0$). Hodnoty T statistik jsou zapsány v Tabulce 14

Tabulka 14: Hodnoty testového kritéria T – Parabolický trend (Zdroj: Vlastní zpracování v software STATISTICA)

Parametr	T
2071200,9	128,6
23193,4	17,49
-782,7	-4,22

Stejně jako u předepsaného pojistného jsme pro kvantil Studentova t -rozdělení pro $(17 - 3)$ stupňů volnosti a hladinu významnosti $\alpha = 0,05$ získali hodnotu 2,145, která udává kritický obor (9, s. 397):

$$W_\alpha = \{T; |T| \geq 2,145\}.$$

Protože pro všechny regresní koeficienty platí, že jejich T hodnota náleží do kritického oboru, tzn. *zamítáme, že některý z regresních parametrů je statisticky nevýznamný.*

Predikce

Dále určíme nejdříve bodové, poté intervalové předpovědi počtu smluv pojištění domácnosti pro roky 2017, 2018 a 2019.

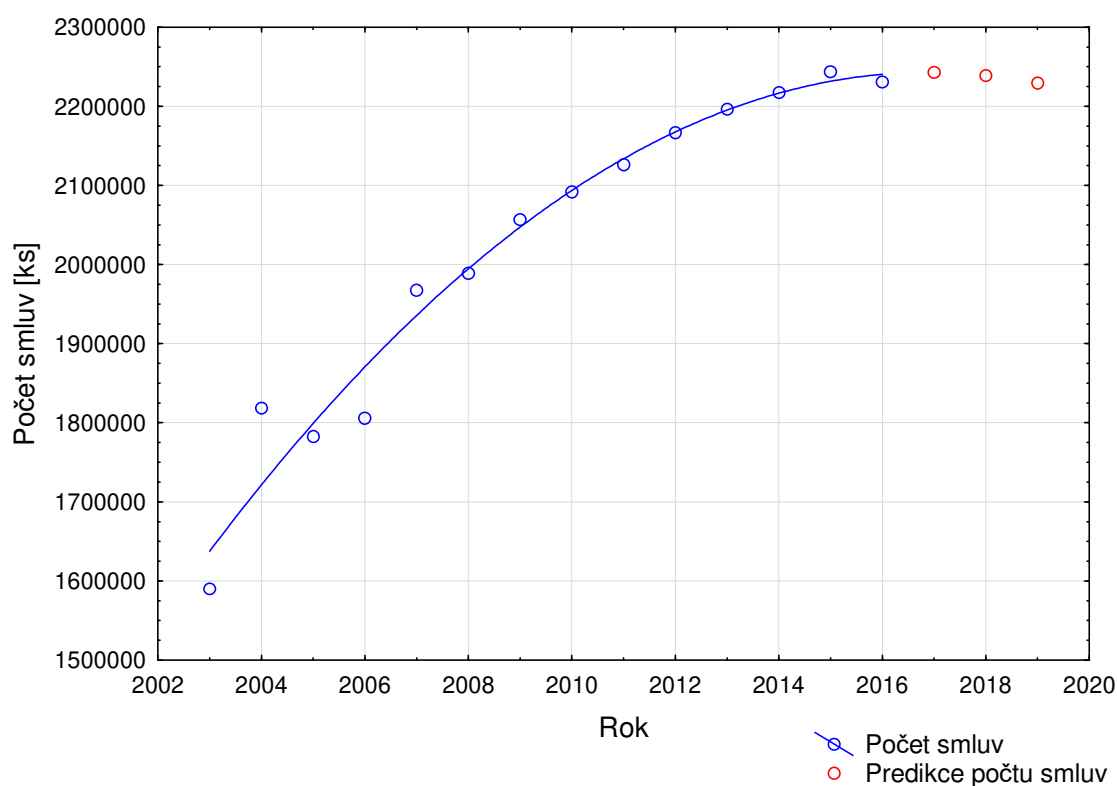
- *Bodová predikce*

Pro zjištění budoucích hodnot počtu uzavřených smluv dosadíme do rovnice parabolického trendu za $t' = 15, 17, 19$. Tyto hodnoty jsou zapsány v Tabulce 15.

Tabulka 15: Pojištění domácností – Bodová predikce hodnot počtu smluv (Zdroj: Vlastní zpracování v software STATISTICA)

Rok	Predikce
2017	2 242 994
2018	2 239 288
2019	2 229 321

Tyto hodnoty a jejich predikce jsou znázorněny v Grafu 10.



Graf 10: Pojištění domácností – Počet smluv, proložená data s predikcí (Zdroj: vlastní zpracování v software STATISTICA)

- *Intervalová predikce*

Pomocí programu „Intervaly“ určíme 95% pravděpodobnostní intervaly hodnot pro roky 2017, 2018 a 2019. Tyto intervaly jsou zapsány v Tabulce 16. Za vstupní funkci zde dosadíme funkci parabolického trendu, matice A je tvořena naměřenými hodnotami a do proměnné rok postupně dosadíme hodnoty 15, 17, 19.

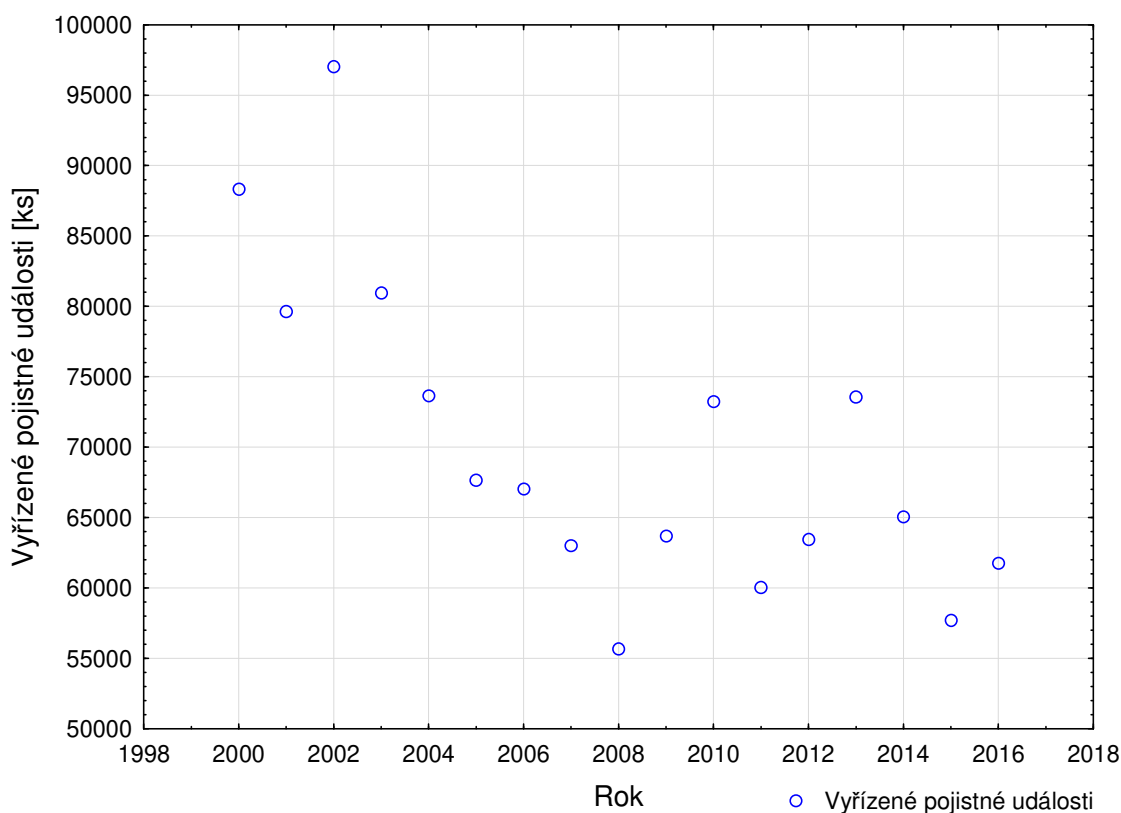
Tabulka 16: Předpovědní intervaly pro počet smluv – Pojištění budov (Zdroj: Vlastní zpracování v systému MATLAB)

Rok	Interval
2017	(2 122 658; 2 363 331)
2018	(2 078 176; 2 400 400)
2019	(2 008 798; 2 449 844)

Počet vyřízených pojistných událostí

Informativní (orientační) test

Podle Grafu 11 je patrné, že pro popis vývoje časové řady nelze použít jednu konkrétní matematickou křivku tak, aby dostatečně přesně proložila naměřená data. Pro dekompozici řady tedy využijeme exponenciální vyrovnávání.



Graf 11: Pojištění domácností – Vyřízené pojistné události (Zdroj: Vlastní zpracování)

Abychom si byli jisti, kterou metodu exponenciálního vyrovnávání použít, porovnáme sumy reziduí metody jednoduchého exponenciálního vyrovnávání a Holtovy metody. Nejlépe vyhovující hodnoty konstant α a γ pro porovnávané metody jsou červeně

zvýrazněny v Tabulkách 17 a 18, které jsme získali funkcí Síťové hledání v software STATISTICA.

Tabulka 17: Síťové hledání hodnot α (Zdroj: Vlastní zpracování v software Statistica)

Alfa	Prům. chyba	Průměr a chyba	Suma mocniny	Průměr mocniny
0,600000	-832,49	7999,518	1,513085E+09	89004975
0,700000	-734,45	8178,313	1,518777E+09	89339819
0,500000	-950,00	7954,743	1,538152E+09	90479502
0,800000	-646,82	8379,432	1,554548E+09	91443971
0,400000	-1105,19	8055,048	1,594633E+09	93801950
0,900000	-565,69	8596,292	1,620118E+09	95301033
0,300000	-1324,90	8310,616	1,684225E+09	99072068
0,200000	-1602,10	8716,515	1,813010E+09	106647626
0,100000	-1657,86	9138,918	1,979220E+09	116424690

Tabulka 18: Síťové hledání hodnot α a γ (Zdroj: Vlastní zpracování v software Statistica)

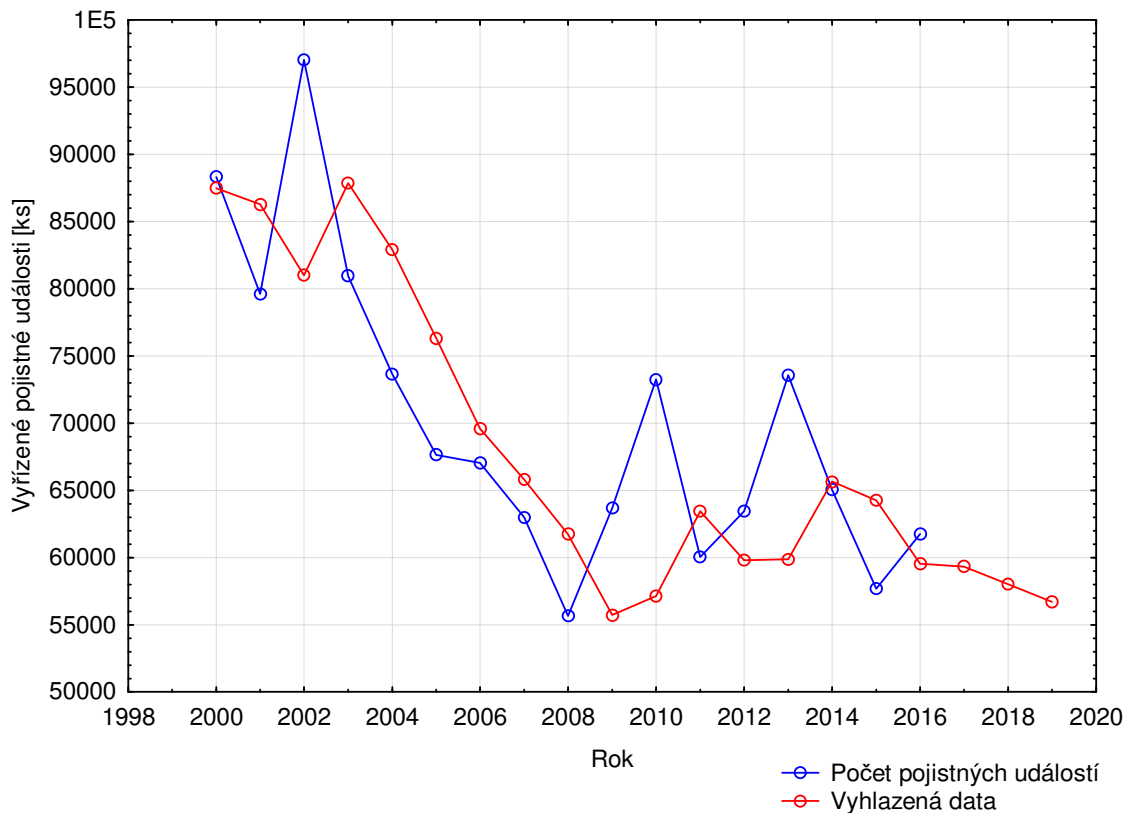
Alfa	Gama	Prům. chyba	Průměr a chyba	Suma mocniny	Průměr mocniny
0,500000	0,100000	408,070	6701,304	1,144822E+09	67342470
0,400000	0,100000	507,304	6768,600	1,148297E+09	67546883
0,600000	0,100000	325,303	6835,203	1,168666E+09	68745030
0,300000	0,100000	552,303	6939,326	1,180411E+09	69435965
0,500000	0,200000	460,715	6765,377	1,215666E+09	71509788
0,700000	0,100000	272,431	7010,165	1,216653E+09	71567818
0,400000	0,200000	677,768	6722,045	1,222210E+09	71894712
0,200000	0,100000	316,003	7034,199	1,227705E+09	72217915
0,100000	0,100000	-793,981	7192,218	1,229312E+09	72312466
0,600000	0,200000	323,917	6890,974	1,242733E+09	73101920

Nejlépe vyhovující konstanty jsou $\alpha = 0,6$ pro jednoduché exponenciální vyrovnávání a $\alpha = 0,5$, $\gamma = 0,1$ pro Holtovu metodu. Provedeme-li exponenciální vyrovnávání pro tyto metody s danými konstantami, dostaneme hodnoty sumy čtverců reziduí rovny $1,51 \cdot 10^9$ pro jednoduché vyrovnávání a $1,14 \cdot 10^9$ pro Holtovu metodu. Z toho plyne, že Holtova metoda je pro tato data přesnější, jelikož suma čtverců reziduí nabývá pro tuto metodu menší hodnoty.

Exponenciální vyrovnávání a predikce

V Grafu 12 jsou znázorněna vyhlazená naměřená data společně s jejich predikcí.

V Tabulce 19 jsou konkrétní hodnoty těchto dat.



Graf 12: Vyřízené pojistné události – Exponenciální vyrovnávání (Zdroj: Vlastní zpracování v software STATISTICA)

Tabulka 19: Vyřízené pojistné události – Vyrovnaná data (Zdroj: Vlastní zpracování v software STATISTICA)

Rok	Data	Vyrovnaná data
2000	88319	87489
2001	79630	86285
2002	97028	81006
2003	80956	87867
2004	73647	82915
2005	67652	76322
2006	67043	69594
2007	63002	65798
2008	55668	61740
2009	63682	55740

2010	73255	57144
2011	60041	63438
2012	63440	59809
2013	73563	59875
2014	65052	65654
2015	57687	64258
2016	61754	59549
2017		59338
2018		58024
2019		56711

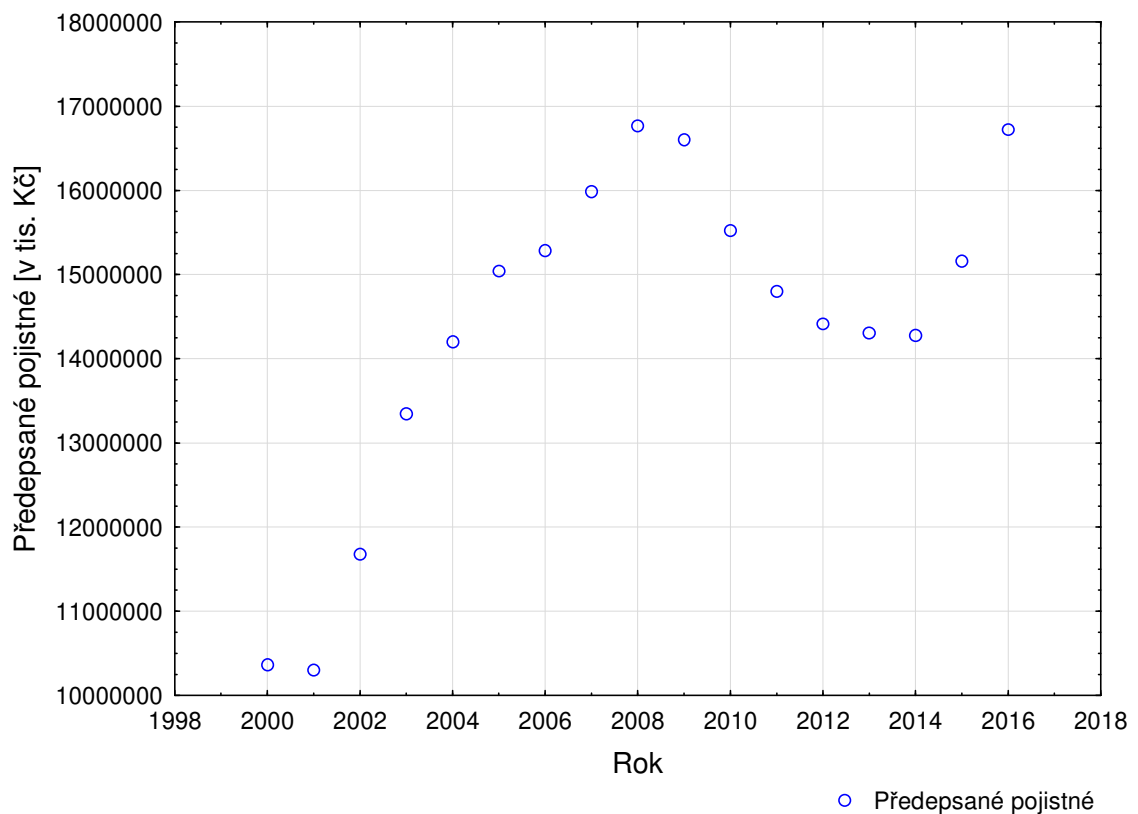
3.3 Havarijní pojištění

Jelikož ČAP neposkytuje pro toto pojištění na svých webových stránkách pro mnou monitorované období hodnoty počtu uzavřených smluv, v této části práce jsem mohl zpracovat pouze data pro ukazatel předepsané pojistné a ukazatel počet vyřízených pojistných událostí. Přehled akceptovaných a zpracovávaných dat je uveden v Příloze 3.

Předepsané pojistné

Vizualizace dat

U tohoto druhu pojištění vycházíme z dat, která jsou zakreslena do následujícího Grafu 13:



Graf 13: Havarijní pojištění – Předepsané pojistné (Zdroj: Vlastní zpracování v software STATISTICA)

Informativní (orientační) test

Na základě Grafu 13 lze prohlásit, že průběh zkoumaných dat nelze s dostatečnou přesností popsat pomocí žádného z trendů. Pro vyrovnání řady tentokrát využijeme

metodu klouzavých průměrů, konkrétně polynomu třetího stupně se šířkou vyhlazovacího okna 5.

Pro určení koeficientu polynomu, pomocí něhož vyrovnáme první dvě hodnoty, jsem naprogramoval v systému MATLAB proceduru „Klouzave_prumery“:

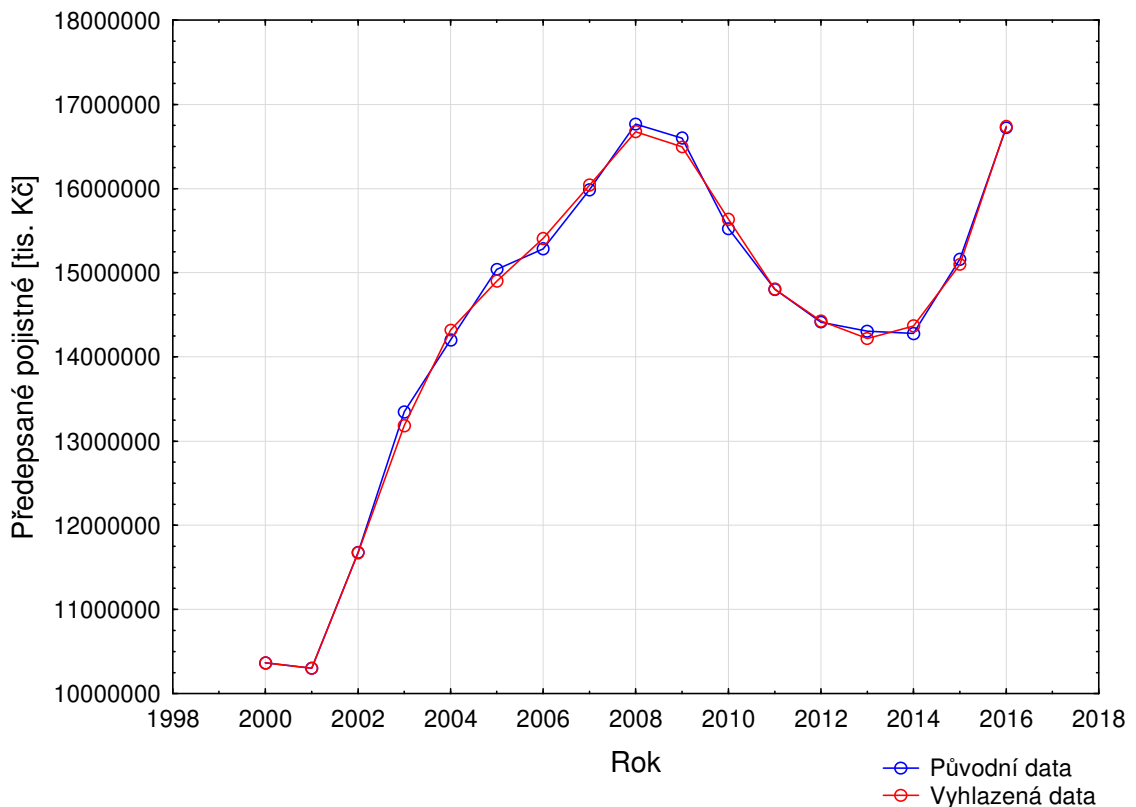
$$\eta(\tau, \mathbf{b}_1) = 11\,673\,649,25714 + 1\,712\,984,6666\tau + 152\,589,0714\tau^2 - 188\,596,1666\tau^3.$$

Po dosazení za $\tau = -2, -1$ dostaneme první a druhou vyrovnanou hodnotu.

Stejně tak dostaneme polynom pro vyrovnání posledních pěti hodnot:

$$\eta(\tau, \mathbf{b}_{13}) = 14\,368\,715,5428 + 376378,6667 \cdot \tau + 303986,9286 \cdot \tau^2 + 50224,8333 \cdot \tau^3.$$

Za τ tentokrát dosadíme hodnoty $\tau = 1$ a $\tau = 2$, tím získáme poslední dvě vyrovnané hodnoty časové řady. Vyrovnání hodnot je znázorněno na Grafu 14.



Graf 14: Předepsané pojistné – Vyrovnaná data (Zdroj: Vlastní zpracování v software STATISTICA a systému MATLAB)

Pomocí polynomu vyhlazujícího poslední pětičky lze po dosazení za $\tau = 3$ predikovat budoucí hodnotu předepsaného pojistného. Tato hodnota je rovna 19 589 804,399 tis. Kč. Vzhledem k tomu, že je tato hodnota výrazně vyšší než poslední naměřená hodnota a zároveň predikce vytvořené pomocí klouzavých průměrů mohou být značně nepřesné, je třeba tuto predikci brát s rezervou.

Konkrétní hodnoty po vyhlazení řady jsou (v tis. Kč) zapsány v Tabulce 20.

Tabulka 20: Předepsané pojistné – Vyrovnaná data (Zdroj: Vlastní zpracování v software MATLAB)

Rok	Data	Vyrovnaná data
2000	10367261	10366805,54
2001	10300028	10301849,83
2002	11676382	11673649,26
2003	13348805	13183992,63
2004	14201661	14321158,71
2005	15042125	14900848,4
2006	15283232	15407086,83
2007	15986244	16041273,51
2008	16767450	16676300,49
2009	16601128	16496055,23

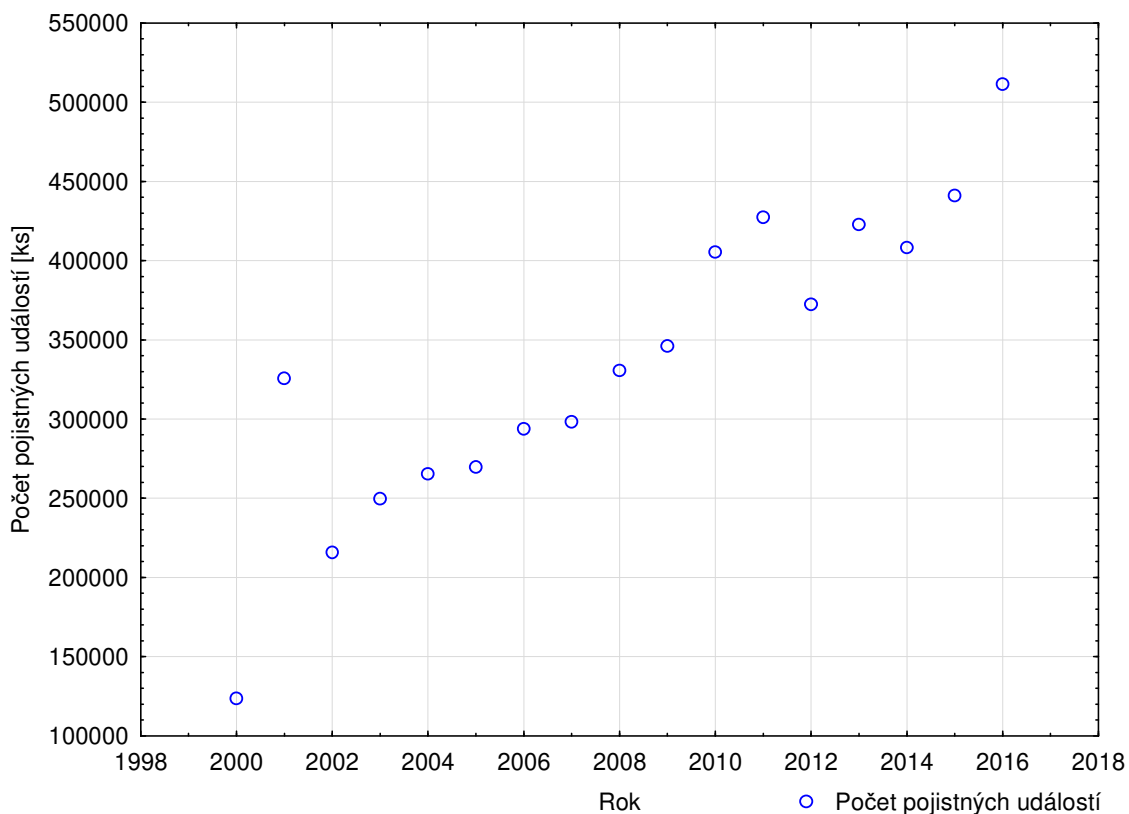
2010	15524967	15634878,17
2011	14802505	14805680,37
2012	14414983	14427288,91
2013	14306596	14218304,11
2014	14277970	14368715,54
2015	15159803	15099305,97
2016	16724095	16739219,26

Počet pojistných událostí

Jako poslední ukazatel budeme zkoumat počet vyřízených pojistných událostí.

Informativní (orientační) test

Na základě vizualizace naměřených dat zakreslených v Grafu 15 lze říci, že pro popis vývoje časové řady nelze použít žádnou z matematických křivek. Abychom dekomponovali časovou řadu, využijeme exponenciální vyrovnávání.



Graf 15: Havarijní pojištění – Počet pojistných událostí (Zdroj: Vlastní zpracování v software STATISTICA)

Pomocí funkce Síťové hledání zjistíme nejlepší hodnoty konstant α , γ , které využijeme pro vyrovnání dat jednoduchým exponenciálním vyrovnáváním a Holtovou metodou.

Pro jednoduché exponenciální vyrovnávání dosadíme jako hodnotu vyrovnávací konstanty $\alpha = 0,5$, viz Tabulka 21, pro Holtovu metodu $\alpha = 0,3$ a $\gamma = 0,1$, viz Tabulka 22. Pro tyto hodnoty dostaneme sumu čtverců reziduí pro jednoduché exponenciální vyrovnávání rovnu $8,37 \cdot 10^{10}$ a pro Holtovu metodu rovnu $4,46 \cdot 10^{10}$. Protože hodnota pro Holtovu metodu je menší, lépe vystihuje zkoumaná data a pro vyrovnání použijeme právě tuto metodu.

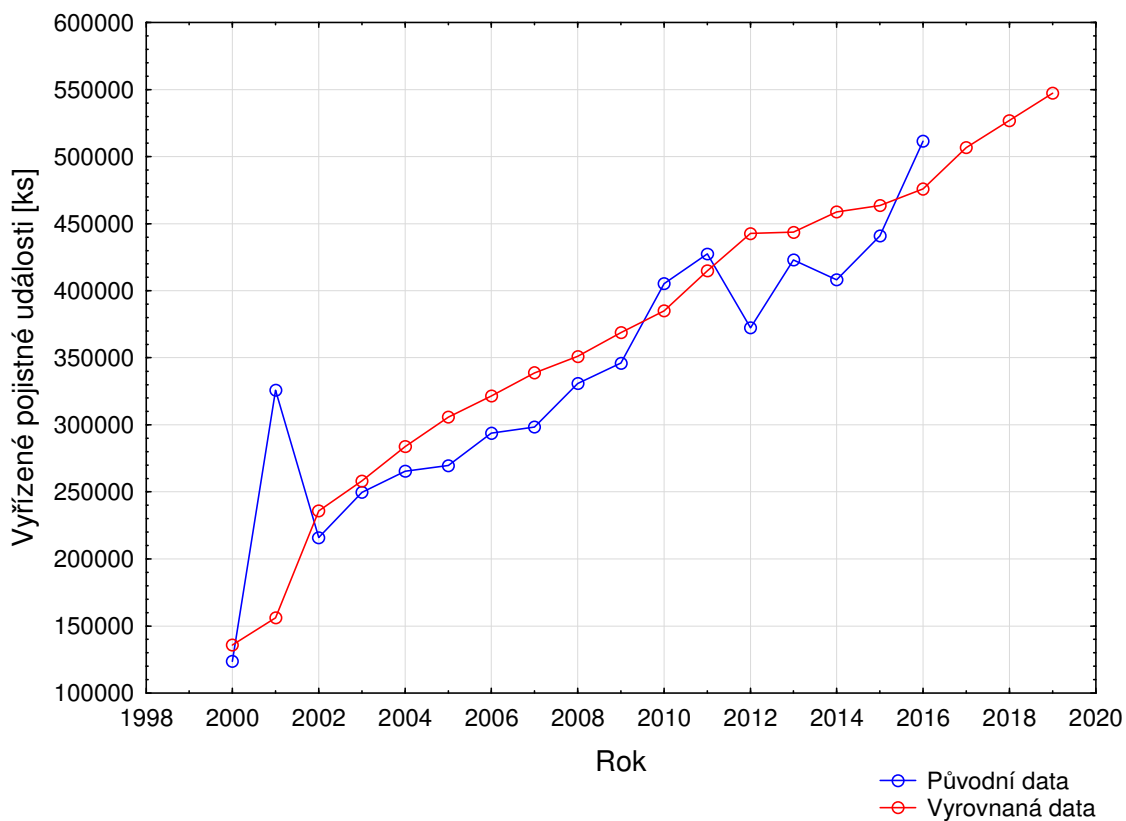
Tabulka 21: Síťové hledání hodnoty α (Zdroj: Vlastní zpracování v software STATISTICA)

Alfa	Prům. chyba	Průměr a chyba	Suma mocniny	Průměr mocniny
0,500000	15495,80	50601,96	8,366035E+10	4,921197E+09
0,600000	13947,50	50382,56	8,509702E+10	5,005707E+09
0,400000	17515,75	52011,01	8,588225E+10	5,051897E+09
0,700000	12747,31	51098,03	8,927132E+10	5,251254E+09
0,300000	20019,05	55450,82	9,372428E+10	5,513193E+09
0,800000	11789,73	52311,27	9,587519E+10	5,639717E+09
0,900000	11001,86	54067,30	1,049800E+11	6,175295E+09
0,200000	22301,22	62865,42	1,104272E+11	6,495715E+09
0,100000	20599,02	71134,75	1,367743E+11	8,045547E+09

Tabulka 22: Síťové hledání hodnot α a γ (Zdroj: Vlastní zpracování v software STATISTICA)

Alfa	Gama	Prům. chyba	Průměr a chyba	Suma mocniny	Průměr mocniny
0,300000	0,100000	-7808,9	35811,67	4,466410E+10	2,627300E+09
0,400000	0,100000	-5603,2	36002,77	4,533881E+10	2,666989E+09
0,300000	0,200000	-6660,4	37356,49	4,598228E+10	2,704840E+09
0,200000	0,100000	-10444,0	35014,62	4,612620E+10	2,713306E+09
0,200000	0,200000	-10800,9	36541,03	4,692118E+10	2,760070E+09
0,500000	0,100000	-3985,3	35577,27	4,714382E+10	2,773166E+09
0,200000	0,300000	-9311,3	37712,79	4,734685E+10	2,785109E+09
0,400000	0,200000	-4115,2	37181,32	4,742936E+10	2,789962E+09
0,300000	0,300000	-4648,8	38233,84	4,758379E+10	2,799046E+09
0,200000	0,400000	-7090,7	38396,38	4,816655E+10	2,833326E+09

Vyrovnaní naměřených dat Holtovou metodou je zakresleno v Grafu 16, konkrétní hodnoty jsou zapsány v Tabulce 23.



Graf 16: Předepsané pojistné – Vyrovnaná data (Zdroj: Vlastní zpracování v software STATISTICA)

Tabulka 23: Počet pojistných událostí – Vyrovnaná data (Zdroj: Vlastní zpracování v software STATISTICA)

Rok	Data	Vyrovnaná data
2000	123715	135831
2001	325723	156065
2002	215887	235921
2003	249687	258268
2004	265421	283794
2005	269600	305831
2006	293810	321424
2007	298322	338773
2008	330719	351058
2009	345995	368766
2010	405382	385061
2011	427413	414894
2012	372456	442762
2013	422838	443673
2014	408185	458800
2015	440964	463475
2016	511431	475905
2017		506813
2018		527063
2019		547312

4 PŘÍNOSY NÁVRHŮ ŘEŠENÍ

Jevy nelze vnímat pouze staticky, v reálném světě je rozumné sledovat jejich dynamiku. K získání kvalitních informací je proto třeba mít povědomí o minulém i budoucím vývoji. Tzn. informace kvantitativního charakteru může mít základ v kvalitativní změně, což je přínosem pro rozhodování. To byl jeden z důvodů, proč mě zajímala tato problematika.

Srovnám-li vývoj *předepsaného pojistného* u všech tří analyzovaných produktů, například porovnáním průměrné difference jednotlivých druhů pojištění, zjistím, že nejrychleji se vyvíjelo předepsané pojistné u havarijního pojištění. Tedy v segmentu pojištění vozidel je třeba pružně reagovat na vnější vlivy, jako jsou změny v legislativě, průměrné stáří automobilu apod. Pokles předepsaného pojistného u havarijního pojištění v období roků 2008 až 2012 lze částečně vysvětlit ekonomickou krizí, kterou v daném období byla ovlivněna ekonomika České republiky.

Z pohledu *predikovaných hodnot předepsaného pojistného* lze předpokládat, že předepsané pojistné pro pojištění budov a domácností kontinuálně poroste.

Porovnám-li první difference u *počtu smluv* v rámci *pojištění budov* a *pojištění domácností*, zjistím, že rychleji se vyvíjelo pojištění budov, které v průměru vzrostlo o 70 536 smluv, zatímco pojištění domácností v průměru o 49 299 smluv za sledované období, což poukazuje na větší zájem o tento pojistný produkt ze strany klientů pojišťoven.

Na rozdíl od předepsaného pojistného zde na základě *predikcí počtu smluv* neočekávám stálý výrazný růst počtu uzavřených smluv u sledovaných druhů neživotního pojištění.

Na základě prvních diferencí *počtu vyřízených pojistných událostí* zjistím, že nejrychleji se ve sledovaném období vyvíjelo havarijní pojištění, které nejpružněji reaguje na vnější vlivy. V průměru u tohoto pojištění vzrostl počet pojistných událostí během sledovaného období o 24 232 za rok. Přestože ve sledovaném období kontinuálně rostl počet uzavřených pojistných smluv v rámci pojištění budov i pojištění domácností, tento rostoucí trend se neprojevil u počtu vyřízených pojistných událostí. Naopak došlo k poklesu průměrné difference u těchto druhů pojištění, tedy během sledovaného období se průměrně za jeden rok snížil počet pojistných událostí u pojištění budov i u pojištění domácností. Toto snížení může být způsobeno několika faktory, mezi které lze zařadit

například využívání stále důmyslnějších bezpečnostních prvků, které pojišťovny často vyžadují v pojistných smlouvách.

Na základě *predikovaných hodnot počtu vyřízených pojistných událostí* očekávám růst počtu pojistných událostí u havarijního pojištění a jejich pokles u pojištění budov a pojištění domácností.

Přínosem těchto analýz jak pro individuální subjekty, tak pro managementy institucí působících v oblasti pojišťovnictví je skutečnost, že mohou zpětně vyhodnocovat kvalitu rozhodnutí v minulosti a ovlivnit svá rozhodnutí budoucí. To může podpořit „rovnováhu“ mezi poptávkou po pojistných produktech a jejich nabídkou na pojistném trhu, a tím přispět k optimalizaci ekonomických výstupů pro obě strany. Rovněž mohou být ovlivněny konkurence či partnerství bankovního sektoru, pokud bude získána dobře vyhodnocená znalost z oblasti pojišťovnictví. Dalším přínosem je i to, že navrhované matematicko-statistické postupy lze aplikovat pro celou řadu dalších pojistných produktů, doplňovat je či rozšiřovat v čase. Pojistný kalkul (výpočty potřebné pro určení závislostí, predikce, přehledné vizualizace) lze provádět přesně a s časovou úsporou v důsledku rychlého rozvoje informačních a komunikačních technologií.

ZÁVĚR

Hlavním cílem této bakalářské práce bylo analyzovat charakteristická data vybraných pojistných produktů neživotního pojištění využitím matematicko-statistických metod ve vhodném software. K získání zdrojových dat je využita veřejně dostupná databáze České asociace pojišťoven, která data poskytuje na spolehlivé úrovni.

Data byla uvažována ponejvíce (pokud byla dána v databázi) v časovém období roků 2000 až 2016. Byly zvoleny často frekventované produkty *neživotního pojištění: pojištění budov, domácností a havarijní*, a to v následujících charakteristických datech: *předepsané pojistné, počet (uzavřených) smluv, počet vyřízených pojistných událostí*.

Analýza byla realizována v několika krocích. Časové řady dat pojistných produktů byly dekomponovány a byl určen jejich *vývojový trend* (tj. jejich funkční vývoj v závislosti na čase). K tomu sloužily různé matematicko-statistické postupy pro konstrukci a ověření vhodnosti regresních modelů, včetně *testování přilohodnosti jejich parametrů*. Následná krátkodobá *bodová a intervalová predikce vývoje* pojistného ukazatele (pojistných dat) byla provedena pro roky 2017, 2018, 2019 (pokud to bylo možné v důsledku využití konkrétní metody pro zjištěný trend).

Výpočetní a vizualizační proces byl převážně veden v software STATISTICA, některé výpočty však byly provedeny v systému MATLAB s využitím samostatně zkonstruovaného potřebného programovacího kódu.

Dále jsem vysvětlil důležité pojmy týkající se analýzy časových řad. Kromě charakteristik časových řad jsem zde uvedl metody užívané pro jejich dekompozici, konkrétně se jedná o popis trendu pomocí regresní analýzy, metodu klouzavých průměrů a exponenciální vyrovnávání. Samostatnou kapitolu jsem věnoval také metodě nejmenších čtverců a testování hypotéz.

- Pro splnění *prvního podcíle* bylo třeba prostudovat řadu teoretických zdrojů, fyzických i internetových, v oblastech pojišťovnictví a kvantitativních metod (zejména šlo o časové řady a jejich charakteristiky, metody užívané pro jejich dekompozici - popis trendu pomocí regresní analýzy, metodu klouzavých průměrů a exponenciální vyrovnávání, dále o metodu nejmenších čtverců, testování hypotéz), zorientovat se i v možných metodikách, jak je či není možné přistupovat k obdobným problematikám mnou zvoleného cíle.

Zajímavé zdroje a materiály z oblasti pojistné matematiky jsou prezentovány profesorem Ciprou z Vysoké školy ekonomické v Praze (i když jsem zdaleka v práci všechny, se kterými jsem se seznámil, nevyužil). Dále se mi přínosnými jeví informace z webovského portálu České asociace pojišťoven. Seznámil jsem se s pracovními prostředními několika pokročilejších software (Maple, MATLAB, STATISTICA), abych nebyl odkázaný pouze na systém MS Excel. Nakonec jsem pro práci zvolil systémy MATLAB a STATISTICA.

-Pro splnění *druhého podcíle* jsem se podrobně zabýval aktuálním stavem systému pojišťovnictví a pojištění v rámci České republiky. Byly charakterizovány jednotlivé druhy neživotního pojištění. Dále jsem se zabýval volbou pojistných dat a informacemi, které uveřejňuje na svých webovských stránkách Česká asociace pojišťoven. Zvolil jsem období roků 2000 až 2016 pro analýzu dat (data roku 2017 dosud nebyla k dispozici) a pro predikci jsem zvolil tři následující roky 2017, 2018, 2019.

-Pro splnění *třetího podcíle* jsem provedl samotnou analýzu dat vybraných tří produktů *životního pojištění*.

U *pojištění budov* jsem se zabýval všemi třemi druhy časových řad: (a) *předepsané pojistné*: nejprve jsem řadu dat informativně vyhodnotil, poté jsem získaná data proložil různými křivkami a na základě porovnávání příslušných indexů determinace jsem vybral nejlépe vyhovující křivku (*parabolický trend*). Pro ověření správnosti volby regresní funkce byla testována hypotéza o nulové hodnotě regresních parametrů, kterou jsem zamítl, žádný ze tří parametrů v regresním modelu není nulový. Z regresního modelu jsem určil bodovou predikci budoucích hodnot předepsaného pojistného. Intervalové predikce jsem získal v software MATLAB užitím vlastního programovacího kódu. Obdobně jsem postupoval u další časové řady pojistných dat: (b) *počet smluv*, kde rovněž nejvhodnějším modelem byl *parabolický trend*. Poslední jsem u tohoto produktu (pojištění budov) analyzoval: (c) *počet vyřízených pojistných událostí*. Protože zde nebyl zřetelný trend, který by popisoval vývoj celé časové řady, využil jsem metodu exponenciálního vyrovnávání. Nejdříve jsem pomocí funkce Síťové hledání určil nejlépe vyhovující konstanty, pomocí nichž jsem proložil zkoumaná data a spočítal sumu čtverců reziduí. Tímto jsem zjistil, zda data lépe proloží jednoduché exponenciální vyrovnávání

či Holtova metoda. Nakonec jsem provedl predikci počtu vyřízených pojistných událostí. Výsledný *model vyrovnaných hodnot* je prezentován na Grafu 6.

Stejným způsobem jsem poté pokračoval u zbývajících druhů neživotního pojištění.

Pro *pojištění domácností* jsem u časové řady (a) *předpis pojistného*: obdržel *parabolický regresní model* vývoje; (b) *počet uzavřených smluv*: jsem obdržel *parabolický regresní model* vývoje; (c) *počet vyřízených pojistných událostí*: výsledný *model vyrovnaných hodnot* je prezentován na Grafu 12.

Pro poslední produkt *havarijní pojištění* jsem měl k dispozici pouze dvě datové řady – na portálu ČAP nebyla k dispozici data: počet (uzavřených) smluv. Rozdíl nastal pouze u časové řady (a) *předepsaného pojistného*, kde není zřetelný globální trend a pro vyrovnání hodnot jsem použil metodu klouzavých průměrů. Pro tuto metodu jsem využil druhý z mnou zkonstruovaných kódů v software MATLAB, pomocí něhož jsem proložil zkoumaná data a určil predikci předepsaného pojistného. Výsledný *model vyrovnaných hodnot* je prezentován na Grafu 14; (b) *počet vyřízených pojistných událostí*: výsledný *model vyrovnaných hodnot* je prezentován na Grafu 16.

Zkoumáním výše uvedených pojistných produktů a jejich charakteristik se otevřely možnosti analýzy finančního trhu, například rovnováhy mezi nabídkou a poptávkou po daných pojistných produktech.

SEZNAM POUŽITÝCH ZDROJŮ

- (1) CIPRA, Tomáš. *Pojistná matematika v praxi*. Praha: HZ, 1994. ISBN 80-901-4956-1
- (2) Základní údaje. *Česká asociace pojišťoven* [online]. ©2014 [cit. 2018-05-17]. Dostupné z: <http://cap.cz/o-nas/organizacni-struktura/zakladni-udaje>
- (3) REJNUŠ, Oldřich. *Finanční trhy: učebnice s programem na generování cvičných testů*. Praha: Grada Publishing, 2016. ISBN 978-80-247-5871-8
- (4) DAŇHEL, Jaroslav. *Pojistná teorie*. Praha: Professional Publishing, 2005. ISBN 80-86419-84-3
- (5) CIPRA, Tomáš. *Praktický průvodce finanční a pojistnou matematikou*. Vydání III., v Ekopressu II. Praha: Ekopress, 2015. ISBN 978-80-87865-18-7.
- (6) DUCHÁČKOVÁ, Eva. *Pojištění a pojišťovnictví*. Praha: Ekopress, 2015. ISBN 978-80-87865-25-5
- (7) DUCHÁČKOVÁ, Eva. *Principy pojištění a pojišťovnictví*. 2., aktualiz. vyd. Praha: Ekopress, 2005. ISBN 80-86119-92-0
- (8) Úrazové pojištění. *Generali* [online]. [cit. 2018-05-14]. Dostupné z: <https://www.generali.cz/pojisteni/zivotni-a-urazove-pojisteni/urazove-pojisteni#ke-stazeni>
- (9) HINDLS, Richard. *Statistika pro ekonomy*. 6. vyd. Praha: Professional Publishing, c2006. ISBN 80-86419-99-1
- (10) KROPÁČ, Jiří. *Statistika B: jednorozměrné a dvourozměrné datové soubory, regresní analýza, časové řady*. Brno: Vysoké učení technické v Brně, 2007. ISBN 80-214-3295-0
- (11) DOUBRAVSKÝ, K. *Ekonometrie* [přednáška]. Brno: VUT, 3. 4. 2018
- (12) CIPRA, Tomáš. *Finanční ekonometrie*. Praha: Ekopress, 2008. ISBN 978-80-86929-43-9
- (13) BUDÍKOVÁ, Marie, Tomáš LERCH a Štěpán MIKOLÁŠ. *Základní statistické metody*. Brno: Masarykova univerzita v Brně, 2005. ISBN 80-210-3886-1

- (14) Single Exponential Smoothing. *Engineering statistics handbook* [online]. [cit. 2018-05-14]. Dostupné z <https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc431.htm>
- (15) How To Identify Patterns in Time Series Data: Time Series Analysis. *Tibco* [online]. [cit. 2018-05-13]. Dostupné z: <http://www.statsoft.com/Textbook/Time-Series-Analysis>
- (16) HEBÁK, Petr, Jiří HUSTOPECKÝ, Eva JAROŠOVÁ a Ivana MALÁ. *Vícerozměrné statistické metody*. Praha: Informatorium, 2005.
ISBN 80-7333-036-9
- (17) SOUČEK, Eduard. *Statistika pro ekonomy*. Praha: Vysoká škola ekonomie a managementu, 2006. ISBN 80-86730-06-9.
- (18) ČERMÁK, Libor a Rudolf HLAVIČKA. *Numerické metody*. 2. vyd. Brno: Akademické nakladatelství CERM, s.r.o, 2008. ISBN 978-80-214-3752-4.
- (19) Výroční zprávy. *Česká asociace pojišťoven* [online]. ©2014 [cit. 2018-05-12]. Dostupné z: <http://cap.cz/o-nas/vyrocní-zpravy>

SEZNAM TABULEK

Tabulka 1: Hodnocení stupně invalidity	18
Tabulka 2: Hodnoty testového kritéria T – Parabolický trend.....	47
Tabulka 3: Pojištění budov – Bodová predikce předepsaného pojistného	48
Tabulka 4: Předpovědní intervaly pro předepsané pojistné – Pojištění budov.....	49
Tabulka 5: Hodnoty testového kritéria T – Parabolický trend.....	51
Tabulka 6: Pojištění budov – Bodová predikce počtu smluv.....	51
Tabulka 7: Předpovědní intervaly pro počet smluv – Pojištění budov	52
Tabulka 8: Síťové hledání hodnoty α	54
Tabulka 9: Síťové hledání hodnot α a γ	54
Tabulka 10: Vyřízené pojistné události – Vyrovnaná data.....	55
Tabulka 11: Hodnoty testového kritéria T – Parabolický trend.....	58
Tabulka 12: Pojištění domácností – Bodová predikce předepsaného pojistného	58
Tabulka 13: Předpovědní intervaly pro předepsané pojistné – Pojištění domácností	59
Tabulka 14: Hodnoty testového kritéria T – Parabolický trend.....	61
Tabulka 15: Pojištění domácností – Bodová predikce počtu smluv.....	62
Tabulka 16: Předpovědní intervaly pro počet smluv – Pojištění budov	63
Tabulka 17: Síťové hledání hodnot α	64
Tabulka 18: Síťové hledání hodnot α a γ	64
Tabulka 19: Vyřízené pojistné události – Vyrovnaná data.....	65
Tabulka 20: Předepsané pojistné – Vyrovnaná data.....	68
Tabulka 21: Síťové hledání hodnoty α	70
Tabulka 22: Síťové hledání hodnot α a γ	70
Tabulka 23: Počet pojistných událostí – Vyrovnaná data.....	71

SEZNAM GRAFŮ

Graf 1: Pojištění budov – Předepsané pojistné	45
Graf 2: Pojištění budov – Předepsané pojistné, proložená data s predikcí	48
Graf 3: Pojištění budov – Počet smluv	49
Graf 4: Pojištění budov – Počet smluv, proložená data s predikcí	52
Graf 5: Pojištění budov – Vyřízené pojistné události	53
Graf 6: Vyřízené pojistné události – Exponenciální vyrovnávání	55
Graf 7: Pojištění domácností – Předepsané pojistné	56
Graf 8: Pojištění domácností – Předepsané pojistné, proložená data s predikcí	59
Graf 9: Pojištění domácností – Počet smluv	60
Graf 10: Pojištění domácností – Počet smluv, proložená data s predikcí	62
Graf 11: Pojištění domácností – Vyřízené pojistné události	63
Graf 12: Vyřízené pojistné události – Exponenciální vyrovnávání	65
Graf 13: Havarijní pojištění – Předepsané pojistné	66
Graf 14: Předepsané pojistné – Vyrovnaná data	67
Graf 15: Havarijní pojištění – Počet pojistných událostí	69
Graf 16: Předepsané pojistné – Vyrovnaná data	71

SEZNAM PŘÍLOH

Příloha 1: Data – Pojištění budov

Příloha 2: Data – Pojištění domácností

Příloha 3: Data – Havarijní pojištění

Příloha 1: Data – Pojištění budov (Vlastní zpracování dle (19))

Rok	Předepsané pojistné	Počet pojistných událostí	Počet smluv
2000	1463663	68545	
2001	1551235	53636	
2002	1826522	115750	
2003	2319404	69673	1354979
2004	2693431	59404	1543369
2005	2930296	61513	1653176
2006	3134089	112506	1715374
2007	3436920	105841	1875523
2008	3703609	76332	1939381
2009	3943743	79334	1983081
2010	4208507	113391	2063544
2011	4463184	212555	2099976
2012	4654376	79551	2148778
2013	4867939	97833	2185604
2014	4993488	66616	2201116
2015	5160761	67314	2256693
2016	5359679	64678	2271952

Příloha 2: Data – Pojištění domácností (Vlastní zpracování dle (19))

Rok	Předepsané pojistné	Počet pojistných událostí	Počet smluv
2000	1275450	88319	
2001	1361094	79630	
2002	1510582	97028	
2003	1835789	80956	1590188
2004	2022998	73647	1818563
2005	2020291	67652	1782437
2006	2117087	67043	1805773
2007	2231756	63002	1967681
2008	2416834	55668	1988977
2009	2509711	63682	2057074
2010	2561935	73255	2091644
2011	2641372	60041	2126020
2012	2773725	63440	2166718
2013	2860653	73563	2196814
2014	2900277	65052	2217790
2015	2954325	57687	2243848
2016	3031828	61754	2231074

Příloha 3: Data – Havarijní pojištění (Vlastní zpracování dle (19))

Rok	Předepsané pojistné	Počet pojistných událostí
2000	10367261	123715
2001	10300028	325723
2002	11676382	215887
2003	13348805	249687
2004	14201661	265421
2005	15042125	269600
2006	15283232	293810
2007	15986244	298322
2008	16767450	330719
2009	16601128	345995
2010	15524967	405382
2011	14802505	427413
2012	14414983	372456
2013	14306596	422838
2014	14277970	408185
2015	15159803	440964
2016	16724095	511431