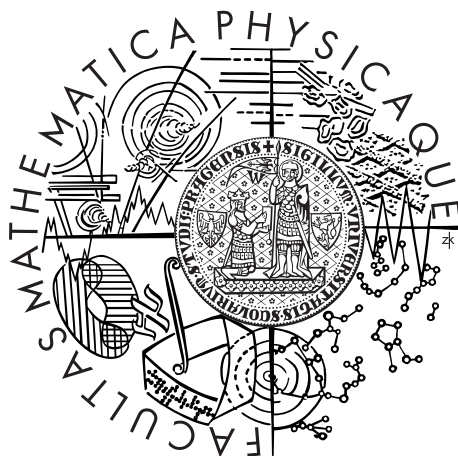


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Bc. Marika Stroukalová

Modelování systémů bonus - malus

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Lucie Mazurová, Ph.D.

Studijní program: Matematika

Studijní obor: Finanční a pojistná matematika

Praha 2013

Na tomto místě bych ráda poděkovala vedoucí diplomové práce RNDr. Lucii Mazurové, Ph.D. za užitečné rady, připomínky a nápady při psaní této práce. Také bych chtěla poděkovat celé své rodině za morální podporu, zejména pak Mgr. Davidu Stroukalovi za přínosné podněty při realizaci této práce. Dále děkuji společnosti Wüstenrot pojišťovna a.s. za poskytnutí statistických dat.

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

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

V Praze dne 31. 7. 2013

Marika Stroukalová

Název práce: Modelování systémů bonus - malus

Autor: Marika Stroukalová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Lucie Mazurová, Ph.D., KPMS MFF UK

Abstrakt: V této práci se zabýváme tarifními systémy bonus-malus užívanými k úpravě apriorně stanoveného pojistného na základě individuálního škodního průběhu v pojištění odpovědnosti za škodu způsobenou provozem vozidla. Záměrem předložené práce je popsat klasický model pro průchod klienta systémem v podobě Markovova řetězce. Pro každou bonus-malus třídu je stanovena relativní sazba pojistného. Dalším cílem této práce je najít optimální hodnoty pro relativní sazby, zohledňující apriorně stanovené pojistné. Teoretický model, založený na stacionárním rozdělení bonus-malus tříd, aplikujeme na konkrétní systém v České republice s využitím reálných českých dat. Výsledkem praktické části práce je porovnání optimálních a skutečných relativních sazeb a posouzení vhodnosti použití zvoleného teoretického modelu pro český systém bonus-malus.

Klíčová slova: bonus-malus systém, apriorní segmentace, stacionární rozdělení, relativní sazba, kvadratická ztrátová funkce

Title: Modelling Bonus - Malus Systems

Author: Marika Stroukalová

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Lucie Mazurová, Ph.D., KPMS MFF UK

Abstract: In this thesis we deal with bonus-malus tariff systems commonly used to adjust the a priori set premiums according to the individual claims during motor third party liability insurance. The main aim of this thesis is to describe the standard model based on the Markov chain. For each bonus-malus class we also determine the relative premium ("relativity"). Another objective of this thesis is to find optimal values for the relativities taking into account the a priori set premiums. We apply the theoretical model based on the stationary distribution of bonus-malus classes on real-world data and a particular real bonus-malus system used in the Czech Republic. The empirical part of this thesis compares the optimal and the real relativities and assesses the suitability of the chosen theoretical model for the particular bonus-malus system.

Keywords: bonus-malus system, a priori segmentation, stationary distribution, relativity, quadratic loss function

Obsah

Úvod	2
1 Modelování (očekávaného) počtu škod	4
1.1 Negativně binomické rozdělení	4
1.2 Regresní model počtu škod	6
1.2.1 Odhady parametrů	6
1.2.2 Roční očekávané počty škod	9
2 Použití Markovova řetězce v BMS	11
2.1 Trajektorie řidiče v BMS	11
2.2 Pravděpodobnosti přechodu mezi třídami	12
2.3 Chování BMS v dlouhém období	13
2.3.1 Konvergence ke stacionárnímu rozdělení	15
3 Optimální bonus-malus sazby	16
3.1 Bayesovské relativity s kvadratickou ztrátovou funkcí	16
3.2 Potřeba přidání fiktivních bonus-malus tříd	20
3.3 Další možnosti pro stanovení optimálních relativit	22
4 Praktická část	24
4.1 Bonus-malus systémy v České republice	24
4.2 Vstupní data	26
4.3 Modifikace a zjednodušení českého bonus-malus systému	31
4.4 Přechodová pravidla, stacionární rozdělení	32
4.5 Relativity	39
4.6 Systém <i>-1/Top</i>	41
Závěr	44
Literatura	46
A Odhady parametrů segmentačních tříd	48
B Konvergence ke stacionárnímu rozdělení	50

Úvod

Pojišťovny v České republice používají ke stanovení výše pojistného v automobilovém pojištění segmentaci na základě apriorní informace o klientech, a následně systém slev (bonusů) a přírážek (malusů) na apriorním (základním) pojistném, zohledňující škodní historii každého řidiče. Bonus-malus systém (BMS) se používá zejména v pojištění odpovědnosti za škodu způsobenou provozem vozidla, běžně nazývaném jako povinné ručení. Každý vlastník motorového vozidla, který jej chce provozovat na pozemních komunikacích v České republice, musí mít toto pojištění sjednáno ze zákona, ale může si vybrat, u jaké pojišťovny z těch, které ho nabízí, povinné ručení uzavře. Naopak pojišťovna je ze zákona povinna zohlednit předcházející bezeškodní průběh klienta.

Český trh v oblasti povinného ručení v současnosti zaznamenává poměrně silný konkurenční boj o zákazníky, apriorní klasifikace i bonus-malus systémy se liší pojišťovna od pojišťovny. Účelem použití aposteriorní rizikové klasifikace v podobě bonus-malus systému společně s apriorním segmentováním klientů je vytvoření co nejhomogennějšího pojistného kmene a stanovení pojistného, které bude adekvátně odpovídat skutečnému rizikovému profilu každého řidiče, tj. jeho aktuální očekávané roční škodní frekvenci. Kromě lepšího ohodnocení individuálního rizika má bonus-malus systém sloužit i k nabádání k bezpečnější jízdě a také k odrazení řidičů od hlášení škod malého rozsahu.

Bonus-malus systém se skládá z určitého počtu tříd (stupňů, úrovní) relativních sazeb pojistného (bonus-malus sazeb), kdy se klient každý rok zařadí do jedné z nich na základě počtu škodních nároků uplatněných v předchozím roce a stanovených pravidel. Klient pak platí pojistné ve výši součinu sjednané (základní) pojistné sazby a bonus-malus sazby příslušející aktuálnímu zařazení do systému.

Cílem této diplomové práce je popsat tradiční přístup s využitím Markovova řetězce pro modelování průchodu klienta systémem bonus-malus tříd, a dále možný způsob stanovení optimálních bonus-malus sazeb pro jednotlivé třídy systému, založený na chování systému v dlouhodobém časovém horizontu. Do výpočtu optimálních sazeb zahrneme apriorní cenovou segmentaci povinného ručení. Důraz bude kladen na praktickou část práce, ve které aplikujeme teoretický model na reálný český systém.

Diplomová práce je rozdělena do čtyř kapitol. V kapitole 1 se budeme zabývat použitím vhodného pravděpodobnostního rozdělení pro počet pojistných událostí v pojistném kmeni povinného ručení a vhodného regresního modelu pro odhad očekávaných škodních frekvencí.

V kapitole 2 se budeme věnovat klasickému modelu pro pohyb klienta v systému v podobě homogenního Markovova řetězce. Pro popis chování systému v dlouhodobém (teoreticky nekonečném) časovém horizontu použijeme stacionární roz-

dělení bonus-malus tříd a uvedeme možnou metodu jeho výpočtu.

Teoretických výsledků předchozích kapitol bude využito v kapitole 3 k určení optimálních relativních sazeb pojistného (s ohledem na apriorně stanovené pojistné), které dostaneme minimalizováním zvolené ztrátové funkce. Zmíníme se také o nutnosti v některých případech zařadit do stávajícího systému další (smyšlené) třídy.

Teoretický model pak implementujeme na vybraný systém bonus-malus v České republice, čemuž je věnována kapitola 4. Budeme tento systém zkoumat, modifikovat a z poskytnutých dat odhadneme neznámé parametry. Na stejná data použijeme také jeden poměrně přísný, hypotetický bonus-malus systém. V závěru aplikační části budeme interpretovat a porovnávat výsledky v podobě zvolených numerických charakteristik obou systémů a provedeme srovnání optimálních bonus-malus sazeb reálného systému se sazbami skutečně používanými. Cílem aplikační části této diplomové práce je odpovědět na otázku, zda-li je zvolený teoretický model vhodný pro český bonus-malus systém.

Hlavním teoretickým zdrojem k problematice modelování bonus-malus systémů bude odborná publikace z roku 2007, *Actuarial Modelling of Claim Counts* ([5]), z níž bude vycházet i většina vzorců a značení použitých v této práci.

Kapitola 1

Modelování (očekávaného) počtu škod

Cílem této kapitoly je popsat nejběžněji používané pravděpodobnostní rozdělení počtu škod v povinném ručení a konstrukci odhadů neznámých parametrů z údajů o apriorní segmentaci a počtech škod řidičů. Budeme vycházet zejména z [5, Kapitola 1 a Kapitola 2].

Mějme n pojištěnců v portfoliu povinného ručení určité pojišťovny. N_i nechť je náhodná veličina popisující počet škod a d_i je příslušná riziková expozice (smlouva nemusí být platná po celé sledované období) na i -té pojistné smlouvě za jeden rok. Pojišťovny segmentují řidiče do rizikových tříd na základě tzv. *apriorních proměnných* (např. věk, bydliště, některé technické údaje o vozidle). U každého pojistníka známe jeho (apriorní) zařazení do určité tarifní třídy, reprezentované náhodným vektorem \mathbf{x}_i . I po roztržení klientů do skupin s podobnou rizikovostí ale existují rozdíly v chování jednotlivých řidičů. Tento reziduální efekt (předem neznámé a těžko měřitelné charakteristiky jako jsou např. řidičské zkušenosti, rychlost reflexů apod.) modelujeme pomocí náhodných veličin $\Theta_1, \dots, \Theta_n$, které jsou nezávislé a stejně rozdělené s kladným nosičem rozdělení a jednotkovou střední hodnotou.

1.1 Negativně binomické rozdělení

Předpokládáme, že náhodná veličina N_i má při $\Theta_i = \theta$ Poissonovo rozdělení se střední hodnotou $\lambda_i^* \theta$, přičemž $\lambda_i^* > 0$ je očekávaná škodní frekvence a θ relativní riziková úroveň (větší rizikovost pro $\theta > 1$ v porovnání s ostatními řidiči ve stejné rizikové skupině a menší rizikovost pro $\theta \in (0, 1)$). Rozdělení počtu škod je v tomto případě definováno pravděpodobnostmi

$$P[N_i = k, | \mathbf{x}_i, \Theta_i = \theta] = e^{-\lambda_i^* \theta} \frac{(\lambda_i^* \theta)^k}{k!}, \quad k = 0, 1, \dots$$

Parametr λ_i^* jsme označili horním indexem *, abychom ho odlišili od roční očekávané frekvence škod λ_i . Počet škod na i -té smlouvě pozorujeme za období délky d_i , $d_i \in (0, 1)$, měřené v letech (riziková expozice je minimálně jeden den, maximálně jeden rok), takže střední hodnota počtu škod (odhadnutá a priori) tohoto pojistníka je

$$\lambda_i^* := d_i \lambda_i,$$

skutečná střední hodnota (tj. včetně aposteriorního rizikového parametru θ) je

$$E [N_i | \mathbf{x}_i, \Theta_i = \theta] = \text{var} [N_i | \mathbf{x}_i, \Theta_i = \theta] = \lambda_i^* \theta = d_i \lambda_i \theta.$$

Výskyty pojistných událostí na i -té pojistné smlouvě tedy tvoří Poissonův proces s intenzitou $\lambda_i \theta$, tj. počet pojistných událostí za časový interval délky d_i má Poissonovo rozdělení se střední hodnotou $(\lambda_i \theta) d_i$.

Pro počty škod v portfoliu povinného ručení je typická tzv. *overdisperte*, která vznikne v důsledku použití náhodného parametru Θ respektive Θ_i pro i -tého klienta. Overdisperte znamená, že rozdělení počtu škod bude mít větší rozptyl než střední hodnotu. Nejčastěji používaným pravděpodobnostním rozdělením náhodné veličiny N_i je negativně binomické rozdělení, které také v této práci uvažujeme. i -tá nezávislá kopie Θ_i náhodného parametru Θ má v tomto případě gama rozdělení s kladnými parametry $\alpha = a$ a $\beta = a$, tj. pro všechna i od 1 do n platí

$$\Theta_i \sim \Gamma(a, a),$$

$$E[\Theta_i] = 1, \tag{1.1}$$

$$\text{var}[\Theta_i] = \frac{1}{a}. \tag{1.2}$$

Hustota veličiny Θ_i má pak tvar

$$f_{\Theta_i}(\theta) = \frac{1}{\Gamma(a)} a^a \theta^{a-1} e^{-a\theta}, \quad \theta > 0 \tag{1.3}$$

V praxi hodnotu parametru Θ_i neznáme, takže rozdělení počtu škodních nároků i -tého klienta se známou rizikovou třídou lze na základě (1.3) vyjádřit pomocí parametrů a a λ_i^* jako

$$\begin{aligned} P[N_i = k | \mathbf{x}_i] &= \int_0^{\infty} P[N_i = k | \mathbf{x}_i, \Theta_i = \theta] f_{\Theta_i}(\theta) d\theta = \\ &= \int_0^{\infty} e^{-\lambda_i^* \theta} \frac{(\lambda_i^* \theta)^k}{k!} \frac{1}{\Gamma(a)} a^a \theta^{a-1} e^{-a\theta} d\theta = \\ &= \frac{(\lambda_i^*)^k a^a}{\Gamma(a) k!} \int_0^{\infty} e^{-(a+\lambda_i^*)\theta} \theta^{a+k-1} d\theta = \\ &= \frac{\Gamma(a+k)}{\Gamma(a) k!} \left(\frac{\lambda_i^*}{a+\lambda_i^*} \right)^k \left(\frac{a}{a+\lambda_i^*} \right)^a = \end{aligned} \tag{1.4}$$

$$= \binom{a+k-1}{k} \left(\frac{\lambda_i^*}{a+\lambda_i^*} \right)^k \left(\frac{a}{a+\lambda_i^*} \right)^a, \quad k = 0, 1, \dots \tag{1.5}$$

K odvození tvaru (1.4) jsme použili substituci

$$(a + \lambda_i^*) \theta = y,$$

$$d\theta = \frac{1}{a + \lambda_i^*} dy$$

a funkci gama vyjádřenou integrálem

$$\Gamma(p) = \int_0^{\infty} e^{-y} y^{p-1} dy,$$

kde p může být jakékoliv komplexní číslo s kladnou reálnou částí. K (1.5) jsme použili vyjádření gama funkce $\Gamma(p)$ pomocí faktoriálu $(p-1)!$.

Veličina N_i má tedy negativně binomické rozdělení s parametry λ_i^* a a . Dle (1.1) a (1.2) pro ni platí

$$\mathbb{E}[N_i | \mathbf{x}_i] = \mathbb{E} \left[\mathbb{E}[N_i | \mathbf{x}_i, \Theta_i] \right] = \mathbb{E}[\lambda_i^* \Theta_i] = \lambda_i^* \mathbb{E}[\Theta_i] = \lambda_i^*, \quad (1.6)$$

$$\begin{aligned} \text{var}[N_i | \mathbf{x}_i] &= \mathbb{E} \left[\text{var}[N_i | \mathbf{x}_i, \Theta_i] \right] + \text{var} \left[\mathbb{E}[N_i | \mathbf{x}_i, \Theta_i] \right] = \\ &= \mathbb{E}[\lambda_i^* \Theta_i] + \text{var}[\lambda_i^* \Theta_i] = \lambda_i^* \mathbb{E}[\Theta_i] + (\lambda_i^*)^2 \text{var}[\Theta_i] = \\ &= \mathbb{E}[N_i | \mathbf{x}_i] + (\lambda_i^*)^2 \frac{1}{a} > \mathbb{E}[N_i | \mathbf{x}_i], \end{aligned} \quad (1.7)$$

tedy heterogenita mezi řidiči nemá vliv na očekávanou škodní frekvenci a přidání náhodného efektu Θ_i do pravděpodobnostního modelu pro počet výskytů pojistných událostí dobře postihuje vlastnost overdisperte.

Existují i další možná pravděpodobnostní rozdělení parametru Θ , např. inverzní Gaussovo (pak máme Poissonovo-inverzní Gaussovo rozdělení jako smíšený pravděpodobnostní model počtu škod) nebo logaritmicke-normální rozdělení, která jsou také (okrajově) představena v [5, Kapitola 1].

1.2 Regresní model počtu škod

K modelování očekávaných škodních frekvencí $\lambda_1^*, \dots, \lambda_n^*$ použijeme regresi s negativně binomickým rozdělením pro odezvy N_1, \dots, N_n . Předpokládáme nezávislá pozorování N_i a známé hodnoty kategoriálních vysvětlujících proměnných $\mathbf{x}_i^\top = (1, x_{i1}, \dots, x_{im})$ vyjádřené jako 0-1 „dummy“ proměnné. Pro očekávané počty pojistných událostí pak máme

$$\lambda_i^* = d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) = \exp \left(\log(d_i) + \beta_0 + \sum_{j=1}^m x_{ij} \beta_j \right), \quad (1.8)$$

kde $\log(d_i)$ je tzv. *offset* sloužící k úpravě modelu kvůli různým délkám platnosti smluv a $\boldsymbol{\beta}^\top = (\beta_0, \beta_1, \dots, \beta_m)$ je vektor neznámých regresních koeficientů (β_0 je koeficient pro referenční rizikovou skupinu). Vektory \mathbf{x}_i a $\boldsymbol{\beta}$ tedy uvažujeme jako sloupcové.

1.2.1 Odhady parametrů

1. Homogenní portfolio

Nejdříve na chvíli předpokládejme homogenní portfolio, v kterém i -tý řidič ($i = 1, \dots, n$) nahlásil během období délky d_i celkem k_i škod. Posloupnost uplatněných škodních nároků tvoří Poissonův proces s intenzitou $\lambda > 0$, takže počet škod

i -tého řidiče během sledovaného období má rozdělení

$$N_i \sim \text{Pois}(\lambda d_i).$$

Střední hodnotu počtu škod za jeden rok můžeme odhadnout metodou maximální věrohodnosti. Mějme věrohodnostní funkci

$$\mathcal{L}(\lambda) = \prod_{i=1}^n \mathbf{P}[N_i = k_i] = \prod_{i=1}^n e^{-\lambda d_i} \frac{(\lambda d_i)^{k_i}}{k_i!}.$$

Logaritmická věrohodnostní funkce má pak vyjádření

$$\begin{aligned} \ell(\lambda) &= \log \mathcal{L}(\lambda) = \sum_{i=1}^n \log \left(e^{-\lambda d_i} \frac{(\lambda d_i)^{k_i}}{k_i!} \right) = \\ &= -\lambda \sum_{i=1}^n d_i + \sum_{i=1}^n \left(k_i (\log(\lambda) + \log d_i) - \log(k_i!) \right) = \\ &= -\lambda \sum_{i=1}^n d_i + \log(\lambda) \sum_{i=1}^n k_i + \textit{konstanta}, \end{aligned}$$

kde *konstanta* označuje členy nezávislé na λ . Položíme-li první derivaci podle λ rovnu nule, tj.

$$-\sum_{i=1}^n d_i + \frac{1}{\lambda} \sum_{i=1}^n k_i = 0,$$

dostaneme maximálně věrohodný odhad¹

$$\hat{\lambda} = \frac{\sum_{i=1}^n k_i}{\sum_{i=1}^n d_i}, \quad (1.9)$$

tedy očekávaný roční počet škod v homogenním portfoliu s Poissonovým rozdělením můžeme odhadnout metodou maximální věrohodnosti jako podíl celkového počtu škod v portfoliu ku celkové rizikové expozici.

2. Poissonovská regrese, tj. bez zahrnutí parametru Θ

V případě, že klienty zařazujeme do rizikových tříd dle apriorních proměnných, ale neuvažujeme zbytkovou heterogenitu v chování řidičů, platí

$$\mathbf{E}[N_i | \mathbf{x}_i] = \text{var}[N_i | \mathbf{x}_i] = \lambda_i^*.$$

Pro Poissonovskou regresi platí vztahy (1.6) a (1.8). Ze záznamů o klientech potřebujeme odhadnout neznámé regresní koeficienty $\beta_0, \beta_1, \dots, \beta_m$ vektoru $\boldsymbol{\beta}$, k čemuž použijeme metodu maximální věrohodnosti. Opět máme ke každému klientovi i pozorování k_i nehod během období délky d_i . Věrohodností funkce má tvar

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{i=1}^n \mathbf{P}[N_i = k_i | \mathbf{x}_i] = \prod_{i=1}^n e^{-\lambda_i^*} \frac{(\lambda_i^*)^{k_i}}{k_i!},$$

¹Druhá derivace podle λ je záporná ($-\lambda^{-2} \sum_{i=1}^n k_i < 0$), takže se skutečně jedná o maximálně věrohodný odhad.

kde vektor koeficientů $\boldsymbol{\beta}$ je schován ve vyjádření λ_i^* (viz (1.8)). Logaritmická věrohodností funkce pak bude tvaru

$$\begin{aligned}\ell(\boldsymbol{\beta}) &= \log \mathcal{L}(\boldsymbol{\beta}) = \sum_{i=1}^n \log \left(e^{-\lambda_i^*} \frac{(\lambda_i^*)^{k_i}}{k_i!} \right) = \\ &= \sum_{i=1}^n \left(-\lambda_i^* + k_i \log(\lambda_i^*) - \log(k_i!) \right) = \\ &= \sum_{i=1}^n \left(-d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) + k_i \mathbf{x}_i^\top \boldsymbol{\beta} \right) + \textit{konstanta},\end{aligned}$$

kde *konstanta* označuje členy nezávislé na $\boldsymbol{\beta}$, přičemž jsme využili vztahu

$$\log(\lambda_i^*) = \log \left(d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \right) = \log(d_i) + \mathbf{x}_i^\top \boldsymbol{\beta}.$$

Položíme-li první derivace logaritmické věrohodnostní funkce podle β_j pro $j = 0, 1, \dots, m$ rovny nule, tj.

$$\sum_{i=1}^n \left(-d_i x_{ij} \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) + k_i x_{ij} \right) = \sum_{i=1}^n x_{ij} (k_i - \lambda_i^*) = 0, \quad j = 0, 1, \dots, m, \quad (1.10)$$

získáme řešením těchto $m+1$ rovnic maximálně věrohodné odhady pro koeficienty $\beta_0, \beta_1, \dots, \beta_m$. Odhady regresních koeficientů jsou softwarovou záležitostí (rovnice (1.10) se řeší numericky).

3. Regrese s negativně binomickým rozdělením, tj. se zahrnutím parametru Θ

Nyní konečně sestavíme věrohodnostní funkci pro parametry $\boldsymbol{\beta}$ a a v uvažovaném pravděpodobnostním modelu pro počty škod s pravděpodobnostmi tvaru (1.5). Máme

$$\begin{aligned}\mathcal{L}(\boldsymbol{\beta}, a) &= \prod_{i=1}^n \mathbb{P}[N_i = k_i \mid \mathbf{x}_i] = \\ &= \prod_{i=1}^n \frac{(\lambda_i^*)^{k_i}}{k_i!} a^a (a + \lambda_i^*)^{-k_i - a} \frac{\Gamma(a + k_i)}{\Gamma(a)},\end{aligned}$$

logaritmickou věrohodnostní funkci pak lze zapsat jako

$$\begin{aligned}\ell(\boldsymbol{\beta}, a) &= \log \mathcal{L}(\boldsymbol{\beta}, a) = \\ &= \sum_{i=1}^n \log \left(\frac{(\lambda_i^*)^{k_i}}{k_i!} a^a (a + \lambda_i^*)^{-k_i - a} \frac{\Gamma(a + k_i)}{\Gamma(a)} \right) = \\ &= \sum_{i=1}^n \left(k_i \log(\lambda_i^*) - \log(k_i!) + a \log(a) - (k_i + a) \log(a + \lambda_i^*) + \right. \\ &\quad \left. + \log(\Gamma(a + k_i)) - \log(\Gamma(a)) \right) = \\ &= \sum_{i=1}^n \left(k_i \mathbf{x}_i^\top \boldsymbol{\beta} - (k_i + a) \log(a + d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta})) \right) + \\ &\quad + \textit{funkce}(a) + \textit{konstanta},\end{aligned} \quad (1.11)$$

kde *funkce*(a) označuje členy, které nejsou závislé na regresních koeficientech, ale jsou závislé na parametru a , a *konstanta* označuje členy nezávislé jak na vektorovém parametru $\boldsymbol{\beta}$, tak na parametru a . Položením prvních (parciálních) derivací (1.11) podle β_j pro $j = 0, 1, \dots, m$ rovných nule dostaneme soustavu $m + 1$ rovnic

$$\begin{aligned} 0 &= \sum_{i=1}^n \left(k_i x_{ij} - (k_i + a) \frac{1}{a + \lambda_i^*} \lambda_i^* x_{ij} \right) = \\ &= \sum_{i=1}^n x_{ij} \left(k_i - \lambda_i^* \frac{a + k_i}{a + \lambda_i^*} \right), \quad j = 0, 1, \dots, m. \end{aligned}$$

Řešením těchto rovnic jsou maximálně věrohodné odhady $\widehat{\beta}_0, \widehat{\beta}_1, \dots, \widehat{\beta}_m$, vektorově tedy $\widehat{\boldsymbol{\beta}}$. Jak je řečeno v [5], odhadnutý parametr $\widehat{\beta}_j$ představuje efekt j -té nezávislé proměnné včetně vlivu s ní korelujících nepozorovatelných proměnných (skrytých charakteristik řidičů). Maximalizací logaritmické věrohodnostní funkce vzhledem k a pak dostaneme odhad \widehat{a} parametru a . Maximálně věrohodné odhady $\widehat{\boldsymbol{\beta}}$ a \widehat{a} se získají opět numericky (nelze je explicitně vyjádřit), pomocí zvoleného iteračního algoritmu. V kapitole 4 si povíme, jakou metodu používá program R.

Pro iterační proceduru se musí vhodně zvolit počáteční hodnoty parametrů. Jako rozumná startovací hodnota parametru $\boldsymbol{\beta}$ se jeví maximálně věrohodný odhad z Poissonovské regrese, tj. numerické řešení rovnice (1.10). Označíme ho jako $\widehat{\boldsymbol{\beta}}_{poc}$. Vstupní hodnotu parametru a označíme \widehat{a}_{poc} a dostaneme ji momentovou metodou z rovnice pro podmíněný rozptyl (1.7), s použitím vstupní hodnoty $\widehat{\boldsymbol{\beta}}_{poc}$. Tedy za použití empirických odhadů rozptylu a střední hodnoty jakožto jejich momentových odhadů postupně dostaneme

$$\begin{aligned} \text{var}[N_i | \mathbf{x}_i] &= \text{E}[N_i | \mathbf{x}_i] + (\lambda_i^*)^2 \frac{1}{a}, \\ \frac{1}{n} \sum_{i=1}^n \left(k_i - d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \right)^2 - \frac{1}{n} \sum_{i=1}^n k_i - \frac{1}{n} \sum_{i=1}^n \left(d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \right)^2 \frac{1}{a} &= 0, \\ \sum_{i=1}^n \left(\left(k_i - d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \right)^2 - k_i - \left(d_i \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \right)^2 \frac{1}{a} \right) &= 0, \end{aligned}$$

tj. dostaneme startovní hodnotu parametru a vyjádřenou jako

$$\widehat{a}_{poc} = \frac{\sum_{i=1}^n \left(d_i \exp(\mathbf{x}_i^\top \widehat{\boldsymbol{\beta}}_{poc}) \right)^2}{\sum_{i=1}^n \left(\left(k_i - d_i \exp(\mathbf{x}_i^\top \widehat{\boldsymbol{\beta}}_{poc}) \right)^2 - k_i \right)}.$$

1.2.2 Roční očekávané počty škod

Jako vstupní hodnoty do vzorců optimálních bonus-malus sazeb, které si představíme v kapitole 4, potřebujeme odhadnout roční očekávané škodní frekvence a určit váhy jednotlivých tarifních tříd, tedy

$$\widehat{\lambda}_k = \exp \left(\widehat{\beta}_0 + \sum_{j=1}^m x_{kj} \widehat{\beta}_j \right) = \exp \left(\widehat{\beta}_0 + \sum_{j=1|x_{kj}=1}^m x_{kj} \widehat{\beta}_j \right), \quad k = 1, \dots, K, \quad (1.12)$$

kde K je počet v portfoliu zastoupených rizikových tříd, $\exp(\widehat{\beta}_0)$ je odhadnutá roční očekávaná škodní frekvence pro zvolenou referenční rizikovou skupinu², a $\exp(\widehat{\beta}_j)$ pro $j = 1, \dots, m$ představuje efekt j -té tarifní proměnné.

Váhu w_k k -té rizikové třídy spočteme jako součet rizikových expozičních pojistných smluv v této třídě vydělený celkovou rizikovou expozicí portfolia. Nechť k -tá tarifní třída zahrnuje t_k smluv. Pak platí

$$w_k = \frac{\sum_{i=1}^{t_k} d_i}{\sum_{i=1}^n d_i}. \quad (1.13)$$

²Např. program R za referenční třídu volí kombinaci kategorií, které jsou u jednotlivých segmentačních kritérií uvedeny na prvních místech ve výčtu používaných kategorií (konkrétní příklad bude popsán v kapitole 4.

Kapitola 2

Použití Markovova řetězce v BMS

System bonus-malus je tvořen určitým (konečným) počtem tříd s relativní výší pojistného (bonus-malus sazbou)¹. Každý pojistný rok je motorista zařazen do jedné z bonus-malus tříd na základě počtu škod hlášených během předešlého pojistného roku, přechodových pravidel a znalosti třídy, do které byl přiřazen v předchozím pojistném roce. Průchod pojistníka systémem budeme modelovat pomocí homogenního Markovova řetězce $\{L_n, n \in \mathbb{N}_0\}$, jehož stavy představují bonus-malus třídy. Kromě konečnosti a homogenity budeme také předpokládat neperiodičnost stavů a nerozložitelnost Markovova řetězce².

Budeme vycházet z všeobecného porozumění těmto pojmům, jejichž definice a vlastnosti lze nalézt např. v univerzitních skriptech [8]. V této kapitole vyložíme Markovovy řetězce s ohledem na záměry této práce. Opět se budeme držet struktury a značení hlavního zdroje [5, Kapitola 4].

2.1 Trajektorie řidiče v BMS

Mějme množinu stavů $S = \{1, \dots, s\}$, kde stav 1 určuje nejlepší třídu (klient zařazený do této třídy dostane nejvyšší bonus neboli slevu na pojistném) a s nejhorší (v této třídě bude klientovi vyměřen nejvyšší malus neboli maximální přírážka k pojistnému). Trajektorii pojištěnce v systému modelujeme pomocí posloupnosti náhodných veličin L_0, L_1, L_2, \dots , kde L_n pro časy $n = 0, 1, \dots$ může nabývat hodnot z S a představuje bonus-malus třídu, ve které je klient v časovém intervalu $[n, n + 1)$. L_0 představuje tzv. vstupní třídu pro nového řidiče (do ní je řidič zařazen v čase 0, tj. když přijde do systému). Vstupní třída je určena spolu s přechodovými pravidly systému³. Pro počet uplatněných škodních nároků pojištěnce během n -tého pojistného roku mějme náhodnou veličinu N_n . Hodnota L_n závisí na hodnotách L_{n-1} a N_n a na daných přechodových pravidlech systému. Např. dle [5] může L_n splňovat následující stochastickou rekurzivní rovnici

$$L_n = \max \left\{ \min \left\{ L_{n-1} + (N_n \times pen - 1), s \right\}, 1 \right\}, \quad (2.1)$$

¹Bliže o relativních sazbách pojednává teoretická kapitola 3 a praktická kapitola 4.

²Nerozložitelnost řetězce znamená, že se lze z každého stavu dostat do ostatních stavů v konečném čase.

³V České republice je to trochu jinak, viz kapitola 4.

kde symbol *pen* znamená penalizaci v podobě sestupu o daný počet tříd za každou způsobenou škodu, např. $pen = 2$ (viz kapitola 4). V tomto případě za každý bezeškodný rok klient postoupí do sousední lepší třídy, (do stavu $L_{n-1} - 1$), za jednu škodu sestoupí do sousední horší třídy (do stavu $L_{n-1} + 1$), při dvou nastalých pojistných událostech se klient v systému posune o tři bonus-malus třídy níže směrem k nejhůře třídě s (přejde do stavu $L_{n-1} + 3$) atd., s ohledem na to, že pojistník se může dostat nejlépe do stavu 1 a nejhůře do stavu s . Vidíme zde vlastnost „bez paměti“ (markovská vlastnost), charakteristickou pro Markovův řetězec, kdy se budoucí trajektorie odvíjí jen od přítomného stavu.

2.2 Pravděpodobnosti přechodu mezi třídami

Zařazování řidičů do jednotlivých tříd probíhá na základě stanovených přechodových pravidel v (konkrétním) systému bonus-malus. Pravidla můžeme zapsat pomocí čtvercových matic $\mathbf{T}(k) = \{t_{ij}(k), i, j \in S\}$, kde

$$t_{ij}(k) = \begin{cases} 1 & \text{pro přechod z } i\text{-té do } j\text{-té třídy pro následující rok,} \\ 0 & \text{jinak} \end{cases}$$

při uplatnění k škod během současného roku.

Nechť počty pojistných událostí N_1, N_2, \dots pojistníka v jednotlivých letech jsou nezávislé a poissonovsky rozdělené náhodné veličiny s roční očekávanou škodní frekvencí λ , tj.

$$N_n \sim Poiss(\lambda), n \in \mathbb{N}.$$

Rozdělení veličin L_1, L_2, \dots závisí na parametru λ představujícím individuální riziko klienta. Nyní můžeme vyjádřit pravděpodobnost přechodu mezi bonus-malus třídami i a j tohoto klienta za jeden pojistný rok jako

$$\begin{aligned} p_{ij} &= \mathbb{P}[L_{n+1} = j \mid L_n = i, L_{n-1} = i_{n-1}, \dots, L_0 = i_0] = \mathbb{P}[L_{n+1} = j \mid L_n = i] = \\ &= \sum_{k=0}^{\infty} \mathbb{P}[L_{n+1} = j \mid L_n = i, N_{n+1} = k] \mathbb{P}[N_{n+1} = k \mid L_n = i] = \\ &= \sum_{k=0}^{\infty} t_{ij}(k) e^{-\lambda} \frac{\lambda^k}{k!}, \quad i, j \in S \end{aligned} \tag{2.2}$$

pro všechny časy $n \in \mathbb{N}_0$ (uvažujeme homogenní Markovův řetězec) a všechny stavy i, \dots, i_0 takové, že pravděpodobnost první podmínky je kladná. Dále jsme v (2.2) využili markovské vlastnosti a nezávislosti N_{n+1} na L_n , neboť (např.) dle (2.1) závisí veličina L_n na hlášených škodách N_1, \dots, N_n a škodní nároky v jednotlivých letech jsou dle předpokladu nezávislé.

Všechny pravděpodobnosti z (2.2) uspořádáme do (čtvercové) matice

$$\mathbf{P} = \{p_{ij}, i, j \in S\}.$$

Matice pravděpodobností přechodu \mathbf{P} je stochastická matice, neboť p_{ij} jsou nezáporné pro všechny stavy $i, j \in S$ a platí $\sum_{j \in S} p_{ij} = 1$ pro všechny $i \in S$.

Pravděpodobnosti p_{ij} tedy můžeme přepsat maticově jako

$$\mathbf{P} = \sum_{k=0}^{\infty} \mathbf{T}(k) e^{-\lambda} \frac{\lambda^k}{k!}.$$

Ovšem matice \mathbf{P} obsahuje pravděpodobnosti přechodu mezi jednotlivými stavy (respektive bonus-malus třídami) Markovova řetězce během jednoho pojistného roku (obecně jednoho období). K dalšímu výkladu budeme potřebovat přechodové pravděpodobnosti za více let.

2.3 Chování BMS v dlouhém období

Potřebujeme mít k dispozici pravděpodobnosti přechodu mezi třídami za více pojistných let, tj. pro každé $n, m \in \mathbb{N}_0$ a $i, j \in S$ pravděpodobnost přechodu ze třídy i do třídy j za n let

$$p_{ij}^{(n)} = \mathbf{P} [L_{m+n} = j \mid L_m = i]$$

či v maticovém vyjádření

$$\mathbf{P}^{(n)} = \left\{ p_{ij}^{(n)}, i, j \in S \right\}.$$

Matice pravděpodobností přechodu $\mathbf{P}^{(n)}$ po n krocích, příslušná matici \mathbf{P} , je také stochastická (tj. všechny prvky z intervalu $(0,1)$ a jednotkové řádkové součty).

Dále zavedeme označení pro počáteční rozdělení pravděpodobností (tj. v čase 0) zařazení pojistníka do jednotlivých bonus-malus tříd

$$\mathbf{p}(0) = \{p_i(0) := \mathbf{P} [L_0 = i], i \in S\}$$

a pro rozdělení tříd v čase n

$$\mathbf{p}(n) = \{p_j(n) := \mathbf{P} [L_n = j], j \in S\}.$$

Uvažujeme $\mathbf{p}(0)$ a $\mathbf{p}(n)$ jako sloupcové vektory. Nepodmíněné pravděpodobnosti $p_j(n)$, nazývané *absolutní pravděpodobnosti* v čase n ([8]), můžeme vyjádřit jako

$$p_j(n) = \sum_{i \in S} p_i(0) p_{ij}^{(n)},$$

maticově pak

$$\mathbf{p}^\top(n) = \mathbf{p}^\top(0) \mathbf{P}^{(n)}.$$

Podobně také vyjádříme přechodové pravděpodobnosti po $n + 1$ krocích jako

$$p_{ij}^{(n+1)} = \sum_{l \in S} p_{il}^{(n)} p_{lj}, \quad (2.3)$$

maticově pak

$$\mathbf{P}^{(n+1)} = \mathbf{P}^{(n)} \mathbf{P},$$

z čehož plyne

$$\mathbf{P}^{(n)} = \mathbf{P}^n.$$

Od bonus-malus systému očekáváme, že se stabilizuje v dlouhodobém časovém horizontu. O tom, jak bude systém vypadat po ustálení, se dozvíme prostřednictvím stacionárního rozdělení tříd ([8]):

Definice 1 (Stacionární rozdělení). *Nechť $\{L_n, n \in \mathbb{N}_0\}$ je homogenní Markovův řetězec s množinou stavů S a maticí pravděpodobností přechodu P . Dále nechť $\boldsymbol{\pi} = \{\pi_j, j \in S\}$ je pravděpodobnostní rozdělení na S , tj. $\pi_j \geq 0$ pro $j \in S$ a $\sum_{j \in S} \pi_j = 1$. Potom $\boldsymbol{\pi}$ je stacionární rozdělení, pokud platí $\pi_j = \sum_{i \in S} \pi_i p_{ij}$ pro každé $j \in S$, nebo v maticovém vyjádření $\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top \mathbf{P}$, když uvažujeme $\boldsymbol{\pi}$ jako sloupcový vektor.*

Díky předpokladům, které jsme stanovili na začátku pro Markovův řetězec, víme, že stacionární rozdělení existuje právě jedno a platí ([8])

$$\pi_l = \lim_{n \rightarrow \infty} p_{il}^{(n)} = \lim_{n \rightarrow \infty} p_l(n) > 0 \quad \text{pro všechna } i, l \in S. \quad (2.4)$$

π_l je tedy pravděpodobnost, že motorista s rizikovým parametrem λ skončí v nekonečném horizontu (po ustálení systému) ve třídě l . Stacionární rozdělení tříd nezávisí na tom, jaké je počáteční rozdělení, tj. pomocí matic můžeme (2.4) zapsat jako

$$\begin{aligned} \Pi &= \begin{pmatrix} \boldsymbol{\pi}^\top \\ \vdots \\ \boldsymbol{\pi}^\top \end{pmatrix} = \lim_{n \rightarrow \infty} \mathbf{P}^{(n)}, \\ \boldsymbol{\pi}^\top &= \lim_{n \rightarrow \infty} \mathbf{p}^{(n)}. \end{aligned} \quad (2.5)$$

Z definice 1 vidíme, že $\pi_j = \sum_{i \in S} \pi_i p_{ij}$ je vlastně (2.3) po aplikování limity pro $n \rightarrow \infty$ na obě strany rovnosti.

Přímá metoda pro výpočet $\boldsymbol{\pi}$

Stacionární rozdělení $\boldsymbol{\pi}$ můžeme spočítat např. přímou metodou uvedenou v [10]. Mějme matici \mathbf{E} složenou z $s \times s$ jedniček, jednotkovou matici \mathbf{I} řádu s a vektor jedniček \mathbf{e} délky s .

Definice 2 (Regulární stochastická matice). *Stochastickou matici nazveme regulární, pokud existuje konečné přirozené číslo n_0 takové, že všechny prvky matice \mathbf{P}^{n_0} jsou kladné.*

Dle našich vstupních předpokladů jsou všechny stavy Markovova řetězce dosažitelné ze všech stavů po konečném počtu kroků (tj. řetězec je nerozložitelný), tedy matice \mathbf{P} splňuje regularitu z definice 2.

Věta 1. *Mějme matici pravděpodobností přechodu \mathbf{P} , o níž předpokládáme, že je regulární. Potom je matice $\mathbf{I} - \mathbf{P} + \mathbf{E}$ invertibilní a platí*

$$\boldsymbol{\pi}^\top = \mathbf{e}^\top (\mathbf{I} - \mathbf{P} + \mathbf{E})^{-1}. \quad (2.6)$$

Důkaz nalezneme v [10].

2.3.1 Konvergence ke stacionárnímu rozdělení

Bonus-malus systém se dostane do rovnovážného stavu bez ohledu na výchozí rozdělení tříd, každopádně zvolené počáteční rozdělení má vliv na to, jak rychle se bude systém stabilizovat. Rozdělení tříd se bude blížit stacionárnímu rozdělení tím rychleji (systém se ustálí za kratší dobu), čím blíže bude počáteční rozdělení stacionárnímu rozdělení. Dle [5] či [11] můžeme systém považovat za ustálený za n_0 let, jestliže

$$\begin{aligned} K_{n_0} &= \sum_{l \in S} \left| p_l(n_0) - \pi_l \right| = \\ &= \sum_{l \in S} \left| \sum_{i \in S} p_i(0) p_{il}^{(n_0)} - \pi_l \right| < \epsilon = 0.05, \end{aligned} \quad (2.7)$$

pokud ovšem známe očekávaný roční počet škod λ . Pro náhodně vybraného řidiče z pojistného kmene, tj. s neznámou očekávanou roční škodní frekvencí λ_k příslušející rizikové třídě k , do které je řidič zařazen na základě apriorních informací, a s neznámou relativní rizikovou úrovní θ , rychlost konvergence zavedeme v následující kapitole.

Kapitola 3

Optimální bonus-malus sazby

Na úvod této kapitoly objasníme pojem **relativita**¹. Relativitu neboli relativní sazbu pojistného v l -té třídě systému bonus-malus značíme r_l a interpretujeme ji následovně: motorista, nacházející se v l -té bonus-malus třídě, platí pojistné rovné součinu základního (apriorního) pojistného příslušejícího segmentační třídě, do které je zařazen při příchodu do systému na základě apriorních proměnných, a této relativní výše pojistného, čili platí $r_l \times 100\%$ z apriorního pojistného.

Cílem této kapitoly je představit jeden z možných způsobů, jak lze optimálně nastavit relativní sazby r_l se zahrnutím apriorní segmentace do modelu. K problematice optimálních relativních (bonus-malus) sazeb budeme přistupovat hlavně na základě [5, Kapitola 4]. Chceme, aby klienti z dlouhodobého hlediska platili pojistné co nejvíce odpovídající jejich aktuální (ale neznámé) škodní frekvenci, tedy jejich skutečnému rizikovému profilu.

3.1 Bayesovské relativity s kvadratickou ztrátovou funkcí

Mějme náhodně vybraného řidiče z pojistného kmene, u kterého neznáme očekávaný počet škod na základě apriorní segmentace ani relativní rizikovou úroveň. Zavedeme pro něj tedy dvě náhodné veličiny dle [5]:

- Λ - představuje (apriorně) očekávanou roční škodní frekvenci
- Θ - představuje relativní výši pojistného.

Řidič pak bude mít aktuální (neznámou) roční očekávanou škodní frekvenci $\Lambda\Theta$. O veličinách Λ a Θ předpokládáme, že jsou vzájemně nezávislé. V první kapitole jsme odhadli roční očekávané počty událostí $\hat{\lambda}_k$ dle (1.12) a spočetli váhy w_k dle (1.13) pro tarifní třídy $k = 1, \dots, K$. Veličina Λ tedy může nabývat hodnot $\hat{\lambda}_1, \dots, \hat{\lambda}_K$ s pravděpodobnostmi

$$\mathbb{P}[\Lambda = \hat{\lambda}_1] = w_1, \dots, \mathbb{P}[\Lambda = \hat{\lambda}_K] = w_K. \quad (3.1)$$

Dále zavedeme náhodnou veličinu

- L_∞ - představuje bonus-malus třídu, v které se ustálí náhodně vybraný motorista v nekonečném časovém horizontu.

¹Relativity - překlad z anglického výrazu *relativities*.

Vzorce pro optimální relativity, které zde za chvíli představíme, jsou založeny na stacionárním rozdělení. Pro účely dalšího výkladu místo π budeme psát $\pi(\widehat{\lambda}_k\theta)$ pro zdůraznění, že se jedná o stacionární rozdělení bonus-malus tříd klienta se známou hodnotou aktuálního ročního počtu škod $\widehat{\lambda}_k\theta$ (to samé pro pravděpodobnosti přechodu či absolutní pravděpodobnosti). Pravděpodobnost, že klient s tímto rizikovým profilem bude po (nekonečně) dlouhém období zařazen do l -té bonus-malus třídy, je

$$P[L_\infty = l \mid \Lambda = \widehat{\lambda}_k, \Theta = \theta] = \pi_l(\widehat{\lambda}_k\theta), \quad l \in S,$$

nepodmíněné pravděpodobnosti výskytu jednotlivých hodnot veličiny L_∞ určíme dle vzorce

$$\begin{aligned} P[L_\infty = l] &= \sum_{k=1}^K P[L_\infty = l \mid \Lambda = \widehat{\lambda}_k] P[\Lambda = \widehat{\lambda}_k] = \\ &= \sum_{k=1}^K w_k \int_0^\infty P[L_\infty = l \mid \Lambda = \widehat{\lambda}_k, \Theta = \theta] f_\Theta(\theta) d\theta = \\ &= \sum_{k=1}^K w_k \int_0^\infty \pi_l(\widehat{\lambda}_k\theta) f_\Theta(\theta) d\theta, \quad l \in S, \end{aligned} \quad (3.2)$$

kde $f_\Theta(\theta)$ je hustota neznámé relativní sazby Θ , o níž v této práci předpokládáme, že pochází z gama rozdělení s jednotkovou střední hodnotou, tedy s hustotou ve tvaru (1.3). Pravděpodobnosti $P[L_\infty = 1]$ až $P[L_\infty = s]$ lze také interpretovat jako podíly pojistníků v jednotlivých bonus-malus třídách po uvedení systému do rovnovážného stavu.

Optimální relativity dostaneme minimalizováním ztrátové funkce. Zvolili jsme kvadratickou ztrátovou funkci, tedy chceme takové hodnoty, pro které je průměrná čtvercová odchylka skutečné²(neznámé) relativní sazby pojistného Θ od odhadnuté relativní sazby r_{L_∞} minimální. Kvadratická ztrátová funkce má tedy tvar

$$\begin{aligned} \text{MSE}(r_{L_\infty}) &= \mathbf{E} [(\Theta - r_{L_\infty})^2] = \\ &= \sum_{l=1}^s \mathbf{E} [(\Theta - r_{L_\infty})^2 \mid L_\infty = l] P[L_\infty = l] = \\ &= \sum_{l=1}^s \int_0^\infty (\theta - r_l)^2 P[L_\infty = l \mid \Theta = \theta] f_\Theta(\theta) d\theta = \\ &= \sum_{l=1}^s \int_0^\infty (\theta - r_l)^2 \sum_{k=1}^K P[L_\infty = l \mid \Theta = \theta, \Lambda = \widehat{\lambda}_k] P[\Lambda = \widehat{\lambda}_k] f_\Theta(\theta) d\theta = \\ &= \sum_{k=1}^K w_k \int_0^\infty \sum_{l=1}^s (\theta - r_l)^2 \pi_l(\theta \widehat{\lambda}_k) f_\Theta(\theta) d\theta. \end{aligned} \quad (3.3)$$

Ztrátovou funkci jsme označili MSE podle anglického výrazu *mean squared error*.

²V publikaci [5] se píše o veličině Θ jako o „true“ *relative premium*, můžeme ji tedy chápat jako spravedlivou či správnou relativní pojistnou sazbu.

Pomocné výsledky:

Bayesův vzorec: aposteriorní rozdělení Θ (za podmínky, že veličiny L_∞ a Λ nabývají konkrétních hodnot) pomocí apriorního rozdělení Θ společně s na ní nezávislou veličinou Λ (tj. pomocí apriorní informace o Θ vyjádřené hustotou (1.3) a marginálního rozdělení Λ v podobě (3.1)) vyjádříme prostřednictvím aposteriorní hustoty $g(\theta | L_\infty = l, \Lambda = \hat{\lambda}_k)$ veličiny Θ jako

$$\begin{aligned} g(\theta | L_\infty = l, \Lambda = \hat{\lambda}_k) &= \frac{\mathbb{P}[L_\infty = l, \Lambda = \hat{\lambda}_k | \Theta = \theta] f_\Theta(\theta)}{\mathbb{P}[L_\infty = l, \Lambda = \hat{\lambda}_k]} = \\ &= \frac{\mathbb{P}[L_\infty = l | \Theta = \theta, \Lambda = \hat{\lambda}_k] \mathbb{P}[\Lambda = \hat{\lambda}_k] f_\Theta(\theta)}{\mathbb{P}[L_\infty = l, \Lambda = \hat{\lambda}_k]} = \\ &= \frac{\pi_l(\theta \hat{\lambda}_k) w_k f_\Theta(\theta)}{\mathbb{P}[L_\infty = l, \Lambda = \hat{\lambda}_k]}. \end{aligned}$$

Podmíněná střední hodnota parametru Θ pak bude

$$\begin{aligned} \mathbb{E}[\Theta | L_\infty = l, \Lambda = \hat{\lambda}_k] &= \int_0^\infty \theta g(\theta | L_\infty = l, \Lambda = \hat{\lambda}_k) d\theta = \\ &= w_k \int_0^\infty \theta \frac{\pi_l(\theta \hat{\lambda}_k)}{\mathbb{P}[L_\infty = l, \Lambda = \hat{\lambda}_k]} f_\Theta(\theta) d\theta. \end{aligned} \quad (3.4)$$

Optimální relativity

Minimalizováním kvadratické ztrátové funkce (3.3) dostaneme pro každou bonus-malus třídu l optimální relativní sazbu r_l rovnou střední hodnotě relativního rizikového parametru Θ za podmínky, že náhodný řidič bude v nekonečném horizontu zařazen do této třídy, tj. bayesovský odhad neznámého parametru Θ . S pomocí (3.2) a (3.4) tak získáme optimální relativity

$$\begin{aligned} r_l &= \mathbb{E}[\Theta | L_\infty = l] = \\ &= \mathbb{E} \left[\mathbb{E}[\Theta | L_\infty = l, \Lambda] | L_\infty = l \right] = \\ &= \sum_{k=1}^K \mathbb{E}[\Theta | L_\infty = l, \Lambda = \hat{\lambda}_k] \mathbb{P}[\Lambda = \hat{\lambda}_k | L_\infty = l] = \\ &= \sum_{k=1}^K w_k \int_0^\infty \theta \frac{\pi_l(\theta \hat{\lambda}_k)}{\mathbb{P}[L_\infty = l, \Lambda = \hat{\lambda}_k]} f_\Theta(\theta) d\theta \frac{\mathbb{P}[L_\infty = l, \Lambda = \hat{\lambda}_k]}{\mathbb{P}[L_\infty = l]} = \\ &= \frac{\sum_{k=1}^K w_k \int_0^\infty \theta \pi_l(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}{\sum_{k=1}^K w_k \int_0^\infty \pi_l(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}, \quad l \in S. \end{aligned} \quad (3.5)$$

Finanční rovnováha

Pro střední hodnotu veličiny r_{L_∞} odvodíme

$$\begin{aligned}
 E[r_{L_\infty}] &= \sum_{l=1}^s r_l P[L_\infty = l] = \\
 &= \sum_{l=1}^s \left(\frac{\sum_{k=1}^K w_k \int_0^\infty \theta \pi_l(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}{\sum_{k=1}^K w_k \int_0^\infty \pi_l(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta} \sum_{k=1}^K w_k \int_0^\infty \pi_l(\hat{\lambda}_k \theta) f_\Theta(\theta) d\theta \right) = \\
 &= \sum_{k=1}^K w_k \int_0^\infty \theta \left(\sum_{l=1}^s \pi_l(\theta \hat{\lambda}_k) \right) f_\Theta(\theta) d\theta = \\
 &= \sum_{k=1}^K w_k \int_0^\infty \theta \cdot 1 \cdot f_\Theta(\theta) d\theta = \sum_{k=1}^K w_k E[\Theta] = \sum_{k=1}^K w_k \cdot 1 = 1.
 \end{aligned}$$

Střední hodnota relativity r_{L_∞} je 1 neboli 100%³, což znamená, že „bourající“ řidiči zaplatí na pojistném více v porovnání s těmi, kteří jezdí bez nehod v rámci téže segmentační skupiny, ale v součtu pojišťovna od klientů vybere stejnou částku jako kdyby optimální bonus-malus systém nepoužívala.

Konvergence ke stacionárnímu rozdělení náhodně vybraného řidiče

V sekci 2.3.1 jsme se zmiňovali o rychlosti konvergence ke stacionárnímu rozdělení pojištěnce s konkrétní rizikovostí. Pro náhodně vybraného jedince z portfolia budeme systém bonus-malus považovat za stabilizovaný za n_0 let, pokud

$$\begin{aligned}
 K_{n_0} &= \sum_{l=1}^s \left| P[L_{n_0} = l] - P[L_\infty = l] \right| = \\
 &= \sum_{l=1}^s \left| \sum_{k=1}^K w_k \int_0^\infty \left(p_l(n_0)(\theta \hat{\lambda}_k) - \pi_l(\theta \hat{\lambda}_k) \right) f_\Theta(\theta) d\theta \right| = \\
 &= \sum_{l=1}^s \left| \sum_{k=1}^K w_k \int_0^\infty \left(\sum_{i \in S} p_i(0)(\theta \hat{\lambda}_k) p_{il}^{(n_0)}(\theta \hat{\lambda}_k) - \pi_l(\theta \hat{\lambda}_k) \right) f_\Theta(\theta) d\theta \right| < \epsilon, \quad (3.6)
 \end{aligned}$$

kde opět uvažujeme $\epsilon = 0.05$.

³Uvedený výsledek plyne rovněž z $E[r_{L_\infty}] = E[E[\Theta | L_\infty]] = E[\Theta] = 1$.

Apriorní a aposteriorní sazbování

Dle [5] můžeme vzájemné ovlivňování apriorního a aposteriorního sazbování popsat pomocí střední hodnoty aposteriorního rozdělení Λ , když víme, v které bonus-malus třídě v dlouhodobém horizontu náhodně vybraný řidič skončí. Se znalostí apriorního rozdělení Λ (3.1) a s pomocí (3.2) dostaneme

$$\begin{aligned} E[\Lambda \mid L_\infty = l] &= \sum_{k=1}^K \hat{\lambda}_k P[\Lambda = \hat{\lambda}_k \mid L_\infty = l] = \\ &= \sum_{k=1}^K \hat{\lambda}_k \frac{P[L_\infty = l \mid \Lambda = \hat{\lambda}_k] P[\Lambda = \hat{\lambda}_k]}{P[L_\infty = l]} = \\ &= \frac{\sum_{k=1}^K \hat{\lambda}_k w_k \int_0^\infty \pi_l(\hat{\lambda}_k \theta) f_\Theta(\theta) d\theta}{\sum_{k=1}^K w_k \int_0^\infty \pi_l(\hat{\lambda}_k \theta) f_\Theta(\theta) d\theta}, \quad l \in S. \end{aligned} \quad (3.7)$$

Intuitivně očekáváme, že tato podmíněná střední hodnota, podobně jako optimální bonus-malus sazba, bude rostoucí funkcí l , což ale nemusí vzhledem k různým speciálním přechodovým pravidlům obecně platit.

Předpokládejme nyní rostoucí průběh střední hodnoty aposteriorního rozdělení Λ a střední hodnoty aposteriorního rozdělení Θ v závislosti na l . Mějme klienta s pozorovatelnými charakteristikami beznehodového řidiče. Při uzavření pojistné smlouvy je tento klient zařazen do tarifní skupiny s nízkým základním pojistným, tj. pojišťovna ještě před započítáním škodního průběhu očekává, že tento klient nebude způsobovat mnoho škod a adekvátně ho tímto (apriorně) ocení. Pak mějme klienta s apriorním profilem nehodového řidiče. Ten bude na počátku zařazen do skupiny s vysokým základním pojistným, neboť pojistitel očekává vysokou škodní frekvenci.

Bude-li se klient po ustálení systému vyskytovat v nejlepší bonusové třídě, bude střední hodnota Λ nízká a klient bude aposteriorně oceněn, protože této třídě odpovídá nízká optimální relativní pojistná sazba. Bude-li se klient v dlouhodobém horizontu pohybovat v malusových třídách, bude střední hodnota Λ vysoká a klient bude pojišťovnou aposteriorně penalizován v podobě vysokých relativních sazeb pojistného, příslušejících těmto třídám.

V této práci se zabýváme stanovením optimálních relativních pojistných sazeb, které do výpočtu zahrnují apriorní segmentaci řidičů. Oproti relativitám, které při výpočtu neberou v úvahu tuto segmentaci, zaznamenávají vždy nižší hodnoty v případě malusových tříd (tj. nižší přírážky k základnímu pojistnému) a vyšší hodnoty v třídách bonusových (tj. nižší slevy na základním pojistném), neboť klienti jsou pojistitelem již částečně oceněni či penalizováni apriorně, tj. než u něj započnou svůj škodní průběh. Tuto skutečnost dokládají i různé příklady uvedené v [5].

3.2 Potřeba přidání fiktivních bonus-malus tříd

Některé systémy bonus-malus jsou zkonstruovány tak, že ve stávající podobě nelze průchod klienta takto definovaným systémem modelovat pomocí Markovova

řetězce (nesplňují markovskou vlastnost). Je potřeba přidat do systému fiktivní třídy tak, aby byly splněny předpoklady, které jsme zmínili v úvodu kapitoly 2. V publikaci [5, Podkapitola 4.7], je tato problematika popsána v souvislosti se speciálním bonusovým pravidlem v dřívějším jednotném belgickém bonus-malus systému. V kapitole 4 se budeme zabývat konkrétním systémem jedné české pojišťovny, kde bude také potřeba systém bonus-malus tříd mírně upravit, takže na tomto místě popíšeme problematiku pouze teoreticky.

Máme nějaký systém bonus-malus s množinou tříd $S = \{1, \dots, s\}$. Třídy, které je potřeba upravit, rozdělíme do podtříd, takže ke každé třídě l z S máme určitý počet podtříd označený písmenem f_l . Pokud jsme k některé třídě i nepřidávali žádné fiktivní třídy, bude $f_i = 1$. Místo upravované třídy l pak do výpočtu stacionárního rozdělení budou vstupovat fiktivní třídy, které značíme l_{-j} (j od 1 do f_l), u ostatních tříd ponecháváme původní značení l . Matice pravděpodobností přechodu pak bude mít rozměry $(\sum_{l=1}^s f_l) \times (\sum_{l=1}^s f_l)$ namísto $s \times s$, vektor stacionárních pravděpodobností pak délku $\sum_{l=1}^s f_l$ oproti původní délce s .

Optimální relativity

Optimální relativní sazby pojistného r_l pro $l \in S = \{1, \dots, s\}$ dostaneme opět jako argument minima kvadratické ztrátové funkce, definované v (3.3), ovšem s omezující podmínkou

$$r_l = r_{l-1} = \dots = r_{l-f_l}, \quad l = 1, \dots, s, \quad (3.8)$$

protože relativity stále počítáme pro původní počet bonus-malus tříd a tedy všechny podtřídy příslušné jedné, takto modifikované, třídě mají stejnou relativní sazbu. Dle [3] pak v upraveném systému určíme optimální relativní sazby pojistného podle vzorce

$$r_l = \frac{\sum_{k=1}^K w_k \int_0^\infty \theta \sum_{j=1}^{f_l} \pi_{l-j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}{\sum_{k=1}^K w_k \int_0^\infty \sum_{j=1}^{f_l} \pi_{l-j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}, \quad l \in S. \quad (3.9)$$

Pro střední hodnotu r_{L_∞} opět odvodíme

$$\begin{aligned} \mathbb{E}[r_{L_\infty}] &= \sum_{l=1}^s r_l \mathbb{P}[L_\infty = l] = \\ &= \sum_{l=1}^s \left(\frac{\sum_{k=1}^K w_k \int_0^\infty \theta \sum_{j=1}^{f_l} \pi_{l-j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}{\sum_{k=1}^K w_k \int_0^\infty \sum_{j=1}^{f_l} \pi_{l-j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta} \sum_{k=1}^K w_k \int_0^\infty \sum_{j=1}^{f_l} \pi_{l-j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta \right) = \\ &= \sum_{k=1}^K w_k \int_0^\infty \theta \left(\sum_{l=1}^s \sum_{j=1}^{f_l} \pi_{l-j}(\theta \hat{\lambda}_k) \right) f_\Theta(\theta) d\theta = \\ &= \sum_{k=1}^K w_k \int_0^\infty \theta f_\Theta(\theta) d\theta = \sum_{k=1}^K w_k \mathbb{E}[\Theta] = \sum_{k=1}^K w_k = 1, \end{aligned}$$

k čemuž jsme využili nového vyjádření rozdělení pravděpodobností původních s hodnot veličiny L_∞ ve tvaru

$$P[L_\infty = l] = \sum_{k=1}^K w_k \int_0^\infty \sum_{j=1}^{f_l} \pi_{l,j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta, \quad l \in S. \quad (3.10)$$

Tedy stačí zaměnit pravděpodobnosti $\pi_l(\theta \hat{\lambda}_k)$ za součet složek vektoru stacionárních pravděpodobností přes všechny podtřídy příslušející modifikované třídě l , tj. za $\sum_{j=1}^{f_l} \pi_{l,j}(\theta \hat{\lambda}_k)$. To samé platí pro podmíněnou střední hodnotu veličiny Λ , kterou v situaci s přidáním podtřídami vyjádříme jako

$$E[\Lambda | L_\infty = l] = \frac{\sum_{k=1}^K \hat{\lambda}_k w_k \int_0^\infty \sum_{j=1}^{f_l} \pi_{l,j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}{\sum_{k=1}^K w_k \int_0^\infty \sum_{j=1}^{f_l} \pi_{l,j}(\theta \hat{\lambda}_k) f_\Theta(\theta) d\theta}, \quad l \in S. \quad (3.11)$$

3.3 Další možnosti pro stanovení optimálních relativit

Existují také jiné způsoby, jak stanovit optimální relativní sazby pojistného, jiné ztrátové funkce nebo jiné předpoklady, např.

- *Lineární relativita* jako rostoucí funkce proměnné l , tj. ve tvaru

$$r_l^{lin} = \alpha + \beta l, \quad \beta > 0, l \in S,$$

kteří dostaneme minimalizováním ztrátové funkce. V případě kvadratické ztrátové funkce minimalizujeme

$$E \left[(\Theta - r_{L_\infty}^{lin})^2 \right] = E \left[(\Theta - \alpha - \beta L_\infty)^2 \right].$$

Odvození tvaru optimálních lineárních relativit najdeme např. v [7]. Výhodou linearizace relativit je pravidelnost růstu r_l^{lin} vzhledem ke třídě l .

- *Geometrické relativita* $r_l^{geo} = \alpha \beta^l$ s kladnými parametry α a β , představené např. v [1] jako alternativa k lineárním relativitám, hledaná proto, že lineární relativní sazby pro nejlepší bonusové třídy mohou (v případě značné heterogenity mezi klienty i po apriorním roztržení) teoreticky vyjít záporné.
- Relativita s *exponenciální ztrátovou funkcí*, která je tvaru

$$E \left[\exp(-c(\Theta - r_{L_\infty})) \right],$$

kde parametr $c > 0$ určuje přísnost bonus-malus systému. Optimální relativita s exponenciální ztrátovou funkcí se blíží relativitám s kvadratickou ztrátovou funkcí pro parametr c blížící se zprava k nule. Čím vyšší je hodnota c , tím je systém méně přísný, tj. snižují se přírážky na pojistném. Při vysoké hodnotě parametru c jsou ale i bonusy nižší. Volbou tohoto parametru může pojišťovna zpřísnit nebo naopak zjemnit aposteriorní úpravy pojistného. Tato ztrátová funkce dává větší důležitost chybám z nadhodnocení odhadnuté relativy ($r_{L_\infty} > \Theta$) než na chyby vzniklé z podhodnocení ([5]). Popis systémů s exponenciální ztrátovou funkcí lze nalézt např. v [4].

- V této práci předpokládáme, že náhodná veličina Λ , reprezentující neznámou apriorně očekávanou roční škodní frekvenci, se nemění v čase a tedy můžeme použít tradiční přístup s použitím homogenního Markovova řetězce a stationárního rozdělení. Systémům se stochastickou migrací mezi segmenty je věnován stejnojmenný článek v [2].

Kapitola 4

Praktická část

Teoretické poznatky předchozích kapitol nyní aplikujeme na reálný bonus-malus systém vybrané pojišťovny v České republice a reálná data, která nám poskytla tato pojišťovna. Hlavním cílem této kapitoly je na základě údajů o počtech pojistných událostí a použité segmentaci řidičů stanovit optimální bonus-malus sazby pro analyzovaný systém a porovnat je se sazbami, které pojišťovna aktuálně používá.

Modelování a výpočty budeme realizovat pomocí statistického softwaru R (verze 3.0.1). Program R a všechny knihovny (balíčky) jsou dostupné z centrálního repozitáře CRAN. Budeme čerpat z dokumentace k jednotlivých knihovnám, zejména [9] a [6], a také pro nás bude užitečný materiál [14] o implementaci regresních modelů v R.

4.1 Bonus-malus systémy v České republice

Teoretická část této práce je, jak již bylo výše uvedeno, založena na publikaci [5], kde autoři ilustrovali teoretické poznatky na belgických datech (z roku 1997), týkající se smluv povinného ručení jednotného systému bonus-malus. V České republice není předepsán jednotný systém přiznávání bonusů a uplatňování malusů v povinném ručení, každá pojišťovna může na český trh povinného ručení přijít s vlastním systémem. Pojišťovny mezi sebou soutěží ve vytváření modelů pro apriorní sazbování a na rozdíl od [5] si také konkurují v nabídce bonusů a malusů. V České republice je používán zcela jiný princip, založený na tzv. **rozhodné době**.

Pro popis takového systému jsme vybrali systém bonus-malus povinného ručení, který používá pojišťovna Wüstenrot. Na tomto místě vycházíme přímo z pojistných podmínek pojišťovny [12], veřejně dostupných na internetu.

- **Rozhodná doba** je doba nepřerušeno trvání pojištění odpovědnosti pojistníka, která se počítá v celých ukončených měsících, snižena za každou pojistnou událost o 24 měsíců.
- **Bonus** je sleva na pojistném za kladnou rozhodnou dobu.
- **Malus** je přírážka k pojistnému za zápornou rozhodnou dobu.
- Při uzavírání pojistné smlouvy se započítává rozhodná doba a přiznává bonus resp. uplatňuje malus podle údajů v databázi škod České kanceláře

pojistitelů. Přiznání bonusu, resp. uplatnění malusu v průběhu pojištění odpovědnosti se uskutečňuje na základě vyhodnocení rozhodné doby a zohlední se v předpisu pojistného nejpozději na následující pojistné období.

- Pojišťovna Wüstenrot používá systém bonus-malus, znázorněný v tabulce 4.1.

Stupeň bonusu	Rozhodná doba v měsících	Bonus (sleva)
B10	120 a více	60%
B9	108 až 119	55%
B8	96 až 107	50%
B7	84 až 95	45%
B6	72 až 83	40%
B5	60 až 71	30%
B4	48 až 59	20%
B3	36 až 47	15%
B2	24 až 35	10%
B1	12 až 23	5%
Základní stupeň	Rozhodná doba v měsících	Bonus (sleva)
Z	0 až 11	0%
Stupeň malusu	Rozhodná doba v měsících	Malus (přirážka)
M1	-12 až -1	30%
M2	-36 až -13	90%
M3	-37 a více	150%

Tabulka 4.1: Systém bonus-malus používaný pojišťovnou Wüstenrot pro povinné ručení.

Systém z tabulky 4.1 má celkem 14 stupňů (tříd); z toho 10 bonusových (klient dostane slevu na pojistném), jeden základní a tři malusové (klientovi je vyměřena přirážka k základnímu pojistnému). Při stanovování pojistného pro další pojistný rok se hodnotí individuální škodní průběh řidiče v předchozích dvanácti měsících. Za každý beznehodový rok řidič dostane bonus v procentuální výši odpovídající jeho aktuálnímu zařazení v systému. Při dlouholetém bezeškodním průběhu může klient dostat slevu až 60%. Pojišťovna Wüstenrot nehodnotí jednu pojistnou událost příliš přísně, když se klientovi jen sníží jeho rozhodná doba a na začátku dalšího období sestoupí v systému maximálně o jeden stupeň. Při častém hlášení škod pojišťovně ale může platit až o 150% více než je jeho základní pojistné, stanovené při uzavření pojistné smlouvy.

Když klient změní pojišťovnu, dají se přesunout jeho bonusy získané u dřívějšího pojistitele. Bohužel, pokud je jeho aktuální rozhodná doba záporná, změnou pojišťovny se nevyhne přesunu jeho malusu. V České republice totiž od 1.10.2007 funguje centrální Databáze škod povinného ručení, spravovaná Českou kanceláří pojistitelů, jak jsme již zmínili výše. Databáze obsahuje historická data počínaje

dnem 1.1.2000, kdy skončil monopol České pojišťovny na poskytování tohoto pojištění. Nový klient se tedy do systému zařadí podle jeho aktuální rozhodné doby. Pokud osoba uzavírá svou první smlouvu o povinném ručení, je její rozhodná doba 0 měsíců a tedy se zařadí do základní bonus-malus třídy (Základní stupeň).

Vzniká tu jeden zásadní problém: rozhodná doba může být v nekonečném horizontu nekonečná a to kladná i záporná, tedy český systém bude třeba modifikovat, aby se dal průchod klienta systémem modelovat pomocí homogenního Markovova řetězce. Celkově bude nutné systém, definovaný pomocí rozhodné doby, upravit tak, aby se dal převést do řeči stavů Markovových řetězců a aby byly splněny všechny předpoklady na něj kladené.

4.2 Vstupní data

Data jsme získali s laskavým svolením pojišťovny Wüstenrot. Máme k dispozici údaje k celkem $n = 112224$ pojistným smlouvám povinného ručení za rok 2012, tj. ke smlouvám, které byly během kalendářního roku 2012 platné alespoň jeden den. Ke každé smlouvě i , $i = 1, \dots, n$ známe

- její expozici v riziku, tj. po jakou část roku 2012 byla v platnosti. . . v teoretické části práce jsme značili d_i
- počet pojistných událostí v roce 2012. . . značili jsme k_i jako realizaci veličiny N_i
- používanou třídu bonusu tj. do jaké bonus-malus třídy je klient zařazen pro příští období, po zohlednění počtu škod v roce 2012.

Ke smlouvám povinného ručení používáme (apriorní) cenovou segmentaci podle frekvence placení pojistného, rozlišování mezi fyzickou a právnickou osobou, věku a místa bydliště majitele auta. Měli jsme původně k dispozici více apriorních proměnných, ale některé jsme vyřadili z důvodu velké výpočetní náročnosti. Vyřadili jsme takové proměnné, které se zdály být nejméně statisticky významné pro počty škod (např. analýzou rozptylu v R). Výsledky optimálních sazeb byly hodně podobné, ať jsme tyto proměnné použili či nikoliv. V této práci není naším úkolem určit optimální apriorní segmentaci, nýbrž optimální aposteriorní bonus-malus sazby, takže budeme brát upravenou a zjednodušenou segmentaci podle čtyř kritérií jako daný předpoklad. V programu R si pro lepší názornost označíme apriorní proměnné (segmentační kritéria) a příslušné kategorie jako

- frekvence placení: proměnná **FREKV** s kategoriemi obecně zapsanými jako **fr1** až **fr3**
- fyzická/právnická osoba: proměnná **F0_PO** s dvěma kategoriemi **o1** a **o2**
- věk řidiče: proměnná **VEK** se čtyřmi kategoriemi **vek1** až **vek4**
- bydliště řidiče: proměnná **REGION** s pěti kategoriemi **reg1** až **reg5**.

Dále pak v R označíme

- rizikovou expozici, nabývající hodnot mezi 0 a 1, proměnnou **EXPOZICE**

- počet pojistných událostí proměnnou POCET_PU.

V R pracujeme s výchozím formátem dat v této podobě:

```
'data.frame': 112224 obs. of 6 variables:
 $ FREKV   : Factor w/ 3 levels "fr1","fr2","fr3": 2 3 2 2 2 3 ...
 $ FO_PO   : Factor w/ 2 levels "o1","o2": 1 1 1 1 1 1 1 1 1 1 ...
 $ EZPOZICE: num 1 0.749 1 1 1 ...
 $ VEK     : Factor w/ 4 levels "vek1","vek2",...: 3 3 3 3 3 3 3 ...
 $ REGION  : Factor w/ 5 levels "reg1","reg2",...: 3 3 5 5 5 5 4 ...
 $ POCET_PU: int 0 0 0 0 0 0 0 0 0 0 ...
```

Za rok 2012 máme celkovou rizikovou expozici 80015.31, tj. součet dob platnosti na 112224 smlouvách. Pojistných událostí máme celkem 2667. Kdybychom uvažovali homogenní pojistný kmen bez apriorní segmentace, dostali bychom průměrnou roční škodní frekvenci jako

$$\hat{\lambda} = \frac{\sum_{i=1}^n k_i}{\sum_{i=1}^n d_i} = \frac{2667}{80015.31} = 0.0333,$$

viz podkapitola 1.2.1, bod 1., odhad (1.9). Hodnota 3.33% se zdá být celkem malá, v případě 100 totožných smluv jsoucích v platnosti po celý rok očekáváme vznik 3.33 pojistných událostí za rok. Když se vrátíme k našim pozorováním, tak v 97.362% případech, tj. ve valné většině, nedošlo během sledovaného roku k žádné nahlášené dopravní nehodě, na 2.530% pojistných smlouvách došlo k jedné pojistné události, 0.104% řidičů hlásilo dvě nehody a zanedbatelných 0.004% smluv zaznamenalo tři pojistné události. K více než třem škodám nedošlo u žádného klienta.

Odhad parametrů v R

Pro počet pojistných událostí, jak jsme psali v kapitole 1, používáme negativně binomickou regresi. Není to klasický zobecněný lineární model (GLM, *generalized linear model*) s negativně binomickým rozdělením závisle proměnné, ale spíš rozšíření GLM s odhadováním parametrů modelu metodou maximální věrohodnosti se zahrnutím přidaného parametru a . Hodnotu overdispersionního parametru a , znamenajícího převrácenou hodnotu rozptylu náhodného parametru Θ , neznáme, a je tedy třeba ji odhadnout z dat. Taktéž neznámé regresní koeficienty β_0, \dots, β_m z (1.8) musíme odhadnout z dat.

Dle [14] v případě, kdy neznáme parametr a , není tedy negativně binomická regrese speciálním případem GLM, nicméně maximálně věrohodné odhady mohou být snadno spočítány opětovným použitím GLM metodologie - iteračním odhadováním β , když známe a , a naopak. To vede k maximálně věrohodným odhadům pro oba parametry β a a , které mohou být spočteny použitím funkce `glm.nb()` z balíčku (knihovny) MASS ([9]). Parametry modelu odhadneme pomocí příkazu

```
glm_pov <- glm.nb(POCET_PU ~ FREKV+REGION+
FO_PO+VEK+offset(log(EXPOZICE)), data = pov)
```

POCET_PU v zadané formuli představuje závisle proměnnou (odezvu), tedy vektor $(N_1, \dots, N_n)^\top$ o délce n pozorování počtu pojistných událostí u jednotlivých smluv během jednoho roku, FREKV+REGION+FO_PO+VEK+offset(log(EXPOZICE)) je pak řada členů specifikujících lineární prediktor $\sum_{j=0}^9 x_{ij}\beta_j + \log(d_i)$ z (1.8) odezvy N_i (prediktorů je 9, viz dále). Do lineárního prediktoru je zahrnut offset (jak jsme psali v podkapitole 1.2) jakožto člen se známým koeficientem 1. V proměnné pov máme uložená data (datová tabulka s 112224 pozorováními o šesti argumentech).

Stručně si nyní popíšeme výstup s odhadnutými parametry regresního modelu. Relevantní část výstupu po aplikování funkce `summary()` vypadá následovně:

Call:

```
glm.nb(formula = POCET_PU ~ FREKV + REGION + FO_PO + VEK
       + offset(log(EXPOZICE)),
       data = pov, init.theta = 0.5089103115, link = log)
```

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.01927	0.14877	-13.574	< 2e-16	***
FREKVfr2	0.13371	0.06302	2.122	0.033872	*
FREKVfr3	-0.32559	0.05619	-5.795	6.84e-09	***
REGIONreg2	-0.11385	0.08511	-1.338	0.180971	
REGIONreg3	-0.20713	0.05463	-3.792	0.000150	***
REGIONreg4	-0.34816	0.07801	-4.463	8.08e-06	***
REGIONreg5	0.13419	0.06977	1.923	0.054441	.
FO_POo2	-0.66949	0.18831	-3.555	0.000378	***
VEKvek2	-0.41055	0.19301	-2.127	0.033412	*
VEKvek3	-1.10533	0.13681	-8.080	6.50e-16	***
VEKvek4	NA	NA	NA	NA	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(0.5089) family taken to be 1)

```
Theta: 0.5089
Std. Err.: 0.0761
```

Byla použita logaritmická linková funkce (defaultní nastavení funkce `glm.nb()`). Argument `init.theta` představuje optimální vstupní hodnotu pro overdispersion parametr a (v R ho nazývají `Theta`), kterou jsme nezadali, takže byl použit momentový odhad na základě vstupních odhadů regresních koeficientů v modelu Poissonovské regrese (GLM model s Poissonovým rozdělením), tak jak jsme zmiňovali v podkapitole 1.2.1, bod 3. Ponechali jsme výchozí nastavení pro odhadovou metodu (`method = "glm.fit"`). Jedná se o iterační algoritmus, anglicky nazývaný jako *iteratively reweighted least squares* (IRLS), česky pak metoda iterativních vážených nejmenších čtverců. Algoritmus IRLS se používá k nalezení maximálně věrohodných odhadů GLM modelu a používá k tomu Fisherovu skórovací metodu (*Fisher scoring*). Ve výstupu z R vidíme odhadnuté (ma-

ximálně věrohodné) regresní koeficienty spolu se směrodatnými odchylkami odhadů a výsledky testu signifikantnosti parametrů v podobě hodnot testových statistik a p -hodnot. Vidíme, že většina parametrů je významná, tj. nulová hypotéza (s oboustrannou alternativou), že skutečná hodnota parametru je rovna nule, je zamítnuta na hladině významnosti testu $\alpha = 0.05$, neboť příslušná p -hodnota je pod touto hladinou. Koeficient ke kategorii **vek4** není definován a to z důvodu shodnosti s jinou kategorií (nebudeme blíže specifikovat). Koeficient značený v R jako (**Intercept**) patří k referenční tarifní třídě, kterou R zvolil jako kombinaci kategorií **fr1**, **reg1**, **o1** a **vek1**. Dále ve výstupu vidíme maximálně věrohodný odhad parametru a se směrodatnou odchylkou odhadu.

j	Název parametru v R	Matem. označení	Odhad $\widehat{\beta}_j$	$\exp(\widehat{\beta}_j)$
0	Intercept	$\widehat{\beta}_0$	-2.01927	0.1327520
1	fr2	$\widehat{\beta}_1$	0.13371	1.1430611
2	fr3	$\widehat{\beta}_2$	-0.32559	0.7220976
3	reg2	$\widehat{\beta}_3$	-0.11385	0.8923884
4	reg3	$\widehat{\beta}_4$	-0.20713	0.8129106
5	reg4	$\widehat{\beta}_5$	-0.34816	0.7059840
6	reg5	$\widehat{\beta}_6$	0.13419	1.1436121
7	o2	$\widehat{\beta}_7$	-0.66949	0.5119678
8	vek2	$\widehat{\beta}_8$	-0.41055	0.6632832
9	vek3	$\widehat{\beta}_9$	-1.10533	0.3311018
	Theta	\widehat{a}	0.5089	

Tabulka 4.2: Odhady regresních koeficientů jako argumentů přirozené exponenciální funkce a odhad overdispersionního parametru z negativně binomické regrese.

Hodnoty, které potřebujeme k výpočtu ročních očekávaných škodních frekvencí a k dosažení do vzorců pro optimální relativní sazby pojistného, jsme názorně zapsali do tabulky 4.2. Jak tyto hodnoty interpretovat a jak z nich dostat výsledné roční očekávané škodní frekvence jednotlivých tarifních tříd, si ukážeme na příkladu klienta se známým apriorním zařazením.

Mějme řidiče i , který je v rizikové třídě s kategoriemi **fr1** (spadá do referenční třídy), **reg4** (této kategorii odpovídá prediktor x_{i5} a odhadnutý koeficient $\widehat{\beta}_5$), **o1** (spadá do referenční třídy) a **vek2** (prediktor x_{i8} a odhadnutý koeficient $\widehat{\beta}_8$). Tuto tarifní třídu označme indexem k . Klientovo zařazení do této třídy můžeme symbolicky zapsat pomocí vektoru prediktorů

$$\begin{aligned} \mathbf{x}_i &= (1, x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}, x_{i7}, x_{i8}, x_{i9}) = \\ &= (1, 0, 0, 0, 0, 1, 0, 0, 1, 0), \end{aligned}$$

kde 1 znamená, že klient spadá do této kategorie, 0 naopak. Pokud by klient spadl do referenční třídy (se všemi kategoriemi), měl by 1 pouze na první místě. Roční očekávanou škodní frekvenci tohoto řidiče pak dostaneme na základě (1.12) jako

$$\begin{aligned} \widehat{\lambda}_k &= \exp(\widehat{\beta}_0) \times \exp(\widehat{\beta}_5) \times \exp(\widehat{\beta}_8) = \\ &= 0.1327520 \times 0.7059840 \times 0.6632832 = 0.062163. \end{aligned}$$

Takto získáme očekávané roční počty škod $\widehat{\lambda}_k$ ke všem zastoupeným rizikovým třídám (je jich 60), pomocí rizikových expozič dostaneme zastoupení (1.13) jednotlivých tříd v pojistném kmeni, které značíme w_k . V tabulce 4.3 je část těchto výsledků (kompletní tabulka je v příloze A), po vynásobení číslem 100 bychom dostali výsledky v procentech. Průměrná očekávaná roční frekvence škod na základě apriorního rozdělení je

$$E[\Lambda] = \sum_{k=1}^{60} \widehat{\lambda}_k P[\Lambda = \widehat{\lambda}_k] = \sum_{k=1}^{60} \widehat{\lambda}_k w_k = 0.0335,$$

tedy skoro stejné číslo jako v případě homogenního pojistného kmene. Průměrná frekvence škod je poměrně nízká, ze sta totožných smluv s jednotkovou rizikovou expozič dojde k pojistné události na 3.35 z nich během jednoho roku.

k	FREKV	FO_PO	VEK	REGION	w_k	$\widehat{\lambda}_k$
1	fr1	o1	vek1	reg1	0.000169	0.132752
2	fr2	o1	vek1	reg1	0.000502	0.151744
3	fr3	o1	vek1	reg1	0.000204	0.095860
4	fr1	o1	vek2	reg1	0.000354	0.088052
5	fr2	o1	vek2	reg1	0.000776	0.100649
6	fr3	o1	vek2	reg1	0.000487	0.063582
7	fr1	o1	vek3	reg1	0.019819	0.043954
8	fr2	o1	vek3	reg1	0.032106	0.050243
9	fr3	o1	vek3	reg1	0.102423	0.031739
10	fr1	o2	vek4	reg1	0.000331	0.067965
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
50	fr2	o1	vek1	reg5	0.000186	0.173536
51	fr3	o1	vek1	reg5	0.000116	0.109627
52	fr1	o1	vek2	reg5	0.000088	0.100697
53	fr2	o1	vek2	reg5	0.000425	0.115103
54	fr3	o1	vek2	reg5	0.000200	0.072713
55	fr1	o1	vek3	reg5	0.013801	0.050267
56	fr2	o1	vek3	reg5	0.024203	0.057458
57	fr3	o1	vek3	reg5	0.069803	0.036298
58	fr1	o2	vek4	reg5	0.000389	0.077725
59	fr2	o2	vek4	reg5	0.000845	0.088845
60	fr3	o2	vek4	reg5	0.003298	0.056125

Tabulka 4.3: Váhy a odhadnuté roční očekávané počty škod pro některé segmentační třídy.

Nyní použijeme jeden test na overdisperzi v datech. Nabízí ho R prostřednictvím funkce `odTest` knihovny `pscl` [6]. Jedná se o test poměru věrohodnosti (*likelihood ratio test*) pro overdisperzi v datech, které znamenají počty, zde tedy počty škod. Jedná se o statistický test, který porovnává logaritmické věrohodnostní funkce negativně binomického regresního modelu a modelu Poissonovské regrese. Nulovou hypotézou je předpoklad Poissonovského regresního modelu (speciální případ negativně binomické regrese s nekonečnou hodnotou parametru a , tj. nulovým rozptylem Θ , kdy platí rovnost střední hodnoty a rozptylu počtu škod).

Rozdělení testové statistiky je nestandardní, jak vidíme ve výstupu z R, jelikož pro overdisperní parametr a (v R nazvaný `Theta`) se předpokládá, že je striktně pozitivní (parametry gama rozdělení jsou kladné). Ve výstupu vidíme, že testová statistika překročila kritickou hodnotu, p -hodnota je pod hladinou významnosti testu $\alpha = 0.05$, takže nulovou hypotézu zamítáme ve prospěch alternativy, tj. negativně binomický regresní model považujeme za adekvátní právě kvůli over-disperzi v datech. Následuje výstup testu:

```
> odTest(glm_pov, alpha=.05)
```

```
Likelihood ratio test of H0: Poisson, as restricted NB model:  
n.b., the distribution of the test-statistic under H0 is  
non-standard  
e.g., see help(odTest) for details/references
```

```
Critical value of test statistic at the alpha= 0.05 level: 2.7055  
Chi-Square Test Statistic = 89.1146 p-value = < 2.2e-16
```

4.3 Modifikace a zjednodušení českého bonus-malus systému

Nyní musíme systém bonusů a malusů pojišťovny Wüstenrot z tabulky 4.1 převést do jazyka Markovových řetězců a upravit ho tak, aby byl Markovův řetězec modelující průchod motoristy systémem homogenní. Bude nutné systém zjednodušit, neboť nejlepší bonusová třída pod označením B10 koresponduje s dobou nepřerušného trvání pojištění 120 měsíců (10 let) a více, tedy může v nekonečném časovém horizontu nabývat nekonečného počtu měsíců. My ale potřebujeme konečný počet tříd, respektive stavů Markovova řetězce. Totéž v záporných hodnotách platí pro nejhorší malusovou třídu pod označením M3.

Nejdříve jsme systém zkusili modifikovat tak, že jsme přidali poměrně dost fiktivních tříd a předpokládali jsme, že rozhodná doba může nabývat minimálně -180 měsíců (-15 let) a maximálně 179 měsíců. Třídu B10 s nejvyšším bonusem jsme rozdělili do pěti podtříd, třídu M3 do dvanácti a třídu M2 do dvou. Místo původních čtrnácti jsme dostali 30 stavů řetězce. Přidáním značného počtu podtříd k okrajovým třídám jsme chtěli z dlouhodobého hlediska pečlivěji rozlišovat mezi klienty s rozhodnou dobou nad 120, respektive pod -37 měsíců. Vzhledem k (3.8) pro $l = 1$ a $f_l = 5$ má pět podtříd vztahených k modifikované bonusové třídě B10 stejnou optimální relativní sazbu pojistného. Po zmíněných úpravách jsme obdrželi výsledky takové, že se téměř 100% klientů v dlouhodobém horizontu zařadilo do nejlepší třídy, s optimální relativitou $r_1 = 98\%$ čili jen s dvouprocentní slevou na pojistném. Optimální relativní sazba r_2 pro druhou nejlepší třídu B9 oproti sazbě r_1 zaznamenala velký skok, neboť v této třídě dle našich výsledků skončí minimum pojistníků z portfolia. Relativity ostatních tříd jen pozvolna rostly a podíly pojistníků v těchto třídách ve stabilizovaném systému byly takřka nulové. Tyto výsledky se nám nezdály příliš uspokojivé. Stávající bonus-malus systém upravíme jiným, více omezujícím způsobem, a ten nyní podrobněji popíšeme.

Jednotlivé stupně bonusů a malusů převedeme na stavy Markovova řetězce tímto způsobem:

- Okrajové třídy B10 a M3 již nerozdělíme do dalších podtříd, ale pro rozhodnou dobu vymezení hranice: dolní hranice -48 měsíců (-4 roky), horní hranice 131 měsíců (necelých 11 let). Rozhodnou dobu „120 a více“ měsíců omezíme na „120 až 131“ a dobu „-37 a více“ na „-48 až -37“.
- Každé stávající bonus-malus třídě s rozhodnou dobou pokrývající x až $x+11$ celých měsíců ($x = -48 + 12j$, $j = 0, 1, \dots, 14$), přiřadíme jeden stav Markovova řetězce. Každý stav řetězce se tedy bude týkat určitého polouzavřeného intervalu $[x, x + 12)$ s ohledem na nastavená přechodová pravidla, která detailněji rozebereme v následující podkapitole.
- Jelikož malusová třída M2 pokrývá dobu dvou let, rozdělíme ji do dvou podtříd pro jednoleté intervaly.

K systému o 14 třídách jsme přidali jen jednu třídu navíc, týkající se malusové třídy M2. Máme tedy 15 stavů homogenního Markovova řetězce. Pokud se klient vyskytuje v bonusové třídě B10 a během roku způsobí jednu škodu, posune se v příštím období do (horší) třídy B9 bez ohledu na to, jestli je jeho skutečná rozhodná doba mezi 120 a 131 měsíci, mezi 132 a 143 či ještě delší. K tomuto zjednodušení nás inspirovala bakalářská práce [13]. Stavy řetězce očíslovujeme od 1 (třída B10) přes 11 (základní stupeň bez bonusu i malusu) do 14 (třída M3). Dvě smyšlené třídy jako podtřídy malusové třídy M2 označíme 13.1 a 13.2 konzistentně s podkapitolou 3.2. Počty podtříd ke všem čtrnácti bonus-malus třídám jsou znázorněny v tabulce 4.4, přehled jednotlivých tříd a odpovídajících stavů Markovova řetězce spolu s nastavenými relativními sazbami pojistného pojišťovny Wüstenrot (úprava z tabulky 4.1) je v tabulce 4.5.

l	1	2	3	4	5	6	7	8	9	10	11	12	13	14
f_l	1	1	1	1	1	1	1	1	1	1	1	1	2	1

Tabulka 4.4: Počty podtříd jednotlivých bonus-malus tříd číslovaných od 1 (pro B10) do 14 (pro M3).

4.4 Přechodová pravidla, stacionární rozdělení

Máme celkem 15 stavů Markovova řetězce, tj. množinu

$$S_{15} = \{1, \dots, 13, 14, 15\}, \quad (4.1)$$

jejíž prvky ovšem dle předchozích poznámek číslujeme poněkud odlišně a v tomto případě mluvíme o množině stavů

$$S = \{1, \dots, 13.1, 13.2, 14\}. \quad (4.2)$$

V případech, kdy se stavy pracujeme jako s čísly (např. v rekurzivní stochastické rovnici vyjadřující trajektorii pojistníka), budeme používat spíše množinu

Bonus-malus třída (značení dle Wüstenrot)	Trvání pojištění v měsících (rozmezí)	Stav Markov. řetězce	Relativní sazba pojistného
B10	[120, 132)	1	0.40
B9	[108, 120)	2	0.45
B8	[96, 108)	3	0.50
B7	[84, 96)	4	0.55
B6	[72, 84)	5	0.60
B5	[60, 72)	6	0.70
B4	[48, 60)	7	0.80
B3	[36, 48)	8	0.85
B2	[24, 36)	9	0.90
B1	[12, 24)	10	0.95
Z	[0, 12)	11	1.00
M1	[-12, 0)	12	1.30
M2	[-24, -12)	13.1	1.90
	[-36, -24)	13.2	
M3	[-48, -36)	14	2.50

Tabulka 4.5: Bonus-malus třídy pojišťovny Wüstenrot s bonusy a malusy vyjádřenými v relativních sazbách (násobcích) základního pojistného, a odpovídající stavy Markovova řetězce (jedno z možných značení).

S_{15} . Každý stav představuje rozhodnou dobu v rozsahu jednoho roku, přesněji jedenácti celých měsíců. Pojišťovna klienty odměňuje, respektive penalizuje, způsobem popsáním v podkapitole 4.1, což znamená následující možnosti při průchodu klienta bonus-malus systémem.

- Po roce bez nehod se pojistník posune v systému o jeden stupeň výše směrem k nejlepší třídě B10 (pokud již v této třídě není), protože se mu rozhodná doba za uplynulý rok navýší o 12 měsíců. V řeci Markovových řetězců přejde ze stavu i do stavu $i - 1$, jestliže není ve stavu 1.
- Za jeden uplatněný pojistný nárok během uplynulého roku pojistník sestoupí o jednu třídu níže směrem k nejhorší třídě M3 (pokud již v této okrajové malusové třídě není), neboť bylo klientovi po nahlášení pojistné události okamžitě strženo 24 měsíců z jeho rozhodné doby, současně ale od předchozího do současného okamžiku stanovování pojistného pro následující pojistné období uplynulo dalších 12 měsíců. K rozhodné době tedy po roce s jednou pojistnou událostí přičítáme $12 - 24 = -12$ měsíců. V řeci stavů zvyšujeme předchozí stav o jedničku, maximální hodnota je 15.
- Po dvou nahlášených pojistných událostech v jednom roce klient sestoupí o tři třídy (pokud je to možné), neboť se k rozhodné době pojistníka po uplynutí jednoho roku přičte $12 - 2 \times 24 = -36$ měsíců. Jinak řečeno, řetězec přejde z původního do stavu i do stavu $i + 3$, jde-li to.

Obecně po k škodách nahlášených během sledovaného roku ($k = 0, 1, \dots$) přičteme k předchozí rozhodné době klienta $12 - k \times 24$ měsíců, v řeci stavů přejde řetězec ze stavu i do stavu $i + k \times 2 - 1$, pokud to jde.

Pojištěncovu trajektorii napříč systémem můžeme vyjádřit pomocí veličiny L_n (zařazení do třídy v čase n pro období $[n, n + 1)$), kdy na základě (2.1) máme pro tento systém rekurzivní rovnici

$$L_n = \max \left\{ \min \{ L_{n-1} + (N_n \times 2 - 1), 15 \}, 1 \right\}$$

Vidíme, že první škoda je penalizována sestupem jen o jednu třídu, zatímco každá další škoda (během jednoho roku) znamená pokles o (další) dvě třídy níže. Přehled zmíněných přechodových pravidel mezi jednotlivými třídami systému je znázorněn v tabulce 4.6, kde ovšem používáme značení stavů/tříd dle množiny S . Na základě této tabulky můžeme jednoduše vyjádřit nula-jedničkové matice $T(k)$ pro žádnou, jednu až „více než osm“ pojistných událostí za jeden rok. Pro k nehod bude v matici $\mathbf{T}(k)$ v i -tém řádku (startující třída i) jednička pouze ve sloupci j , který odpovídá následující třídě po nahlášení k škod dle tabulky 4.6, v ostatních sloupcích pak nuly. Tyto matice zde nebudeme uvádět.

Startující třída	Následující třída po nahlášení								
	0	1	2	3	4	5	6	7	≥ 8
	nehod během 1 roku								
1	1	2	4	6	8	10	12	13_2	14
2	1	3	5	7	9	11	13_1	14	14
3	2	4	6	8	10	12	13_2	14	14
4	3	5	7	9	11	13_1	14	14	14
5	4	6	8	10	12	13_2	14	14	14
6	5	7	9	11	13_1	14	14	14	14
7	6	8	10	12	13_2	14	14	14	14
8	7	9	11	13_1	14	14	14	14	14
9	8	10	12	13_2	14	14	14	14	14
10	9	11	13_1	14	14	14	14	14	14
11	10	12	13_2	14	14	14	14	14	14
12	11	13_1	14	14	14	14	14	14	14
13_1	12	13_2	14	14	14	14	14	14	14
13_2	13_1	14	14	14	14	14	14	14	14
14	13_2	14	14	14	14	14	14	14	14

Tabulka 4.6: Přechodová pravidla pro modifikovaný bonus-malus systém s patnácti stavy.

Na základě těchto pravidel a předpokladu Poissonova rozdělení počtu škod klienta se střední hodnotou λ můžeme pravděpodobnosti přechodu (2.2) takto rizikového klienta uspořádat do matice pravděpodobností přechodu \mathbf{P} o rozměrech 15×15 ve tvaru¹

$$\mathbf{P} = \begin{pmatrix} k_0 & k_1 & 0 & k_2 & \cdots & \cdots & 0 & k_7 & 1 - \sum_{i=0}^7 k_i \\ k_0 & 0 & k_1 & 0 & k_2 & \cdots & \cdots & k_6 & 0 & 1 - \sum_{i=0}^6 k_i \\ 0 & k_0 & 0 & k_1 & 0 & \cdots & \cdots & 0 & k_6 & 1 - \sum_{i=0}^6 k_i \\ 0 & 0 & k_0 & 0 & k_1 & \cdots & \cdots & k_5 & 0 & 1 - \sum_{i=0}^5 k_i \\ 0 & 0 & 0 & k_0 & 0 & \cdots & \cdots & 0 & k_5 & 1 - \sum_{i=0}^5 k_i \\ 0 & 0 & 0 & 0 & k_0 & \cdots & \cdots & k_4 & 0 & 1 - \sum_{i=0}^4 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & k_4 & 1 - \sum_{i=0}^4 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & k_3 & 0 & 1 - \sum_{i=0}^3 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & k_3 & 1 - \sum_{i=0}^3 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & k_2 & 0 & 1 - \sum_{i=0}^2 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & k_2 & 1 - \sum_{i=0}^2 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & k_1 & 0 & 1 - \sum_{i=0}^1 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & k_0 & 0 & 1 - \sum_{i=0}^1 k_i \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & k_0 & 0 & 1 - k_0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & k_0 & 0 & 1 - k_0 \end{pmatrix}, \quad (4.3)$$

kde

$$k_i = \frac{\lambda^i}{i!} e^{-\lambda}, \quad i = 0, 1, \dots, 7.$$

Stacionární rozdělení

Stacionární rozdělení bonus-malus tříd pojistníka se známou roční očekávanou škodní frekvencí λ zkonstruujeme přímou metodou podle věty 1. Např. pro $\lambda = 0.0333$ dostaneme vektor stacionárních pravděpodobností

$$\begin{aligned} \boldsymbol{\pi}^\top = & (0.96501, & 0.03268, & 0.001647, & 0.00062, \\ & 4.60609 \times 10^{-5}, & 9.01970 \times 10^{-6}, & 9.38893 \times 10^{-7}, & 1.28172 \times 10^{-7}, \\ & 1.61281 \times 10^{-8}, & 1.9452 \times 10^{-9}, & 2.5629 \times 10^{-10}, & 3.0949 \times 10^{-11}, \\ & 3.99 \times 10^{-12}, & 4.96 \times 10^{-13}, & 6.2 \times 10^{-14}). \end{aligned}$$

Vidíme, že řidič s nízkou očekávanou frekvencí škod 3.33% se s dosti velkou pravděpodobností (96.5%) ustálí ve třídě s nejvyšším bonusem.

Pro náhodně vybraného řidiče z portfolia povinného ručení spočítáme stacionární pravděpodobnosti kvůli přidání smyšlených tříd na základě (3.10). Pravděpodobnost, že klient s neznámou skutečnou střední hodnotou počtu škod $\Lambda\Theta$ (náhodné veličiny Λ a Θ jsme představili v podkapitole 3.1) skončí v l -té bonus-malus třídě, jsme v teoretické části práce značili $P[L_\infty = l]$. S využitím odhadnutých parametrů \hat{a} z posledního řádku tabulky 4.2 a $\{\hat{\lambda}_k, k = 1, \dots, 60\}$

¹Celá matice se sem bohužel nevešla, ale chybějící sloupce není obtížné doplnit. Matice má řádkově odstupňovaný tvar a jistou pravidelnost v opakujících se hodnotách 0 a symbolů k_i .

s příslušnými vahami $\{w_k, k = 1, \dots, 60\}$ z přílohy A tedy dostaneme stacionární rozdělení

$$\begin{aligned} P[L_\infty = l] = & (0.960632, \quad 0.029900, \quad 0.004372, \quad 0.002364, \quad 0.000877, \\ & 0.000475, \quad 0.000284, \quad 0.000194, \quad 0.000147, \quad 0.000123, \\ & 0.000111, \quad 0.000109, \quad 0.000247, \quad 0.000173). \end{aligned} \quad (4.4)$$

V tomto případě má vektor stacionárních pravděpodobností již délku 14, tj. rovnou počtu původních bonus-malus tříd. I náhodně vybraný jedinec se v nekonečném časovém horizontu s nejvyšší pravděpodobností zařadí do nejlepší bonusové třídy (B10 resp. 1), s pravděpodobností zhruba 96%. Můžeme také konstatovat, že 96% pojistníků skončí v nekonečném horizontu ve třídě s nejvyšší slevou na základním pojistném. Zhruba 3% pojistníků bude zařazeno do druhé nejlepší třídy (B9 resp. 2). Zbylé procento pojistníků skončí v ostatních třídách, pravděpodobnosti jsou odstupňované, nejméně pojistníků skončí dle očekávání v nejhorší malusové třídě (M3 resp. 14).

Podívejme se nyní na počáteční rozdělení, čili rozmístění klientů do jednotlivých tříd dle údajů o zařazení do tříd pro rok 2013. Do systému jsme přidali fiktivní třídy, tehdejší rozhodné doby neznáme, takže počáteční rozdělení

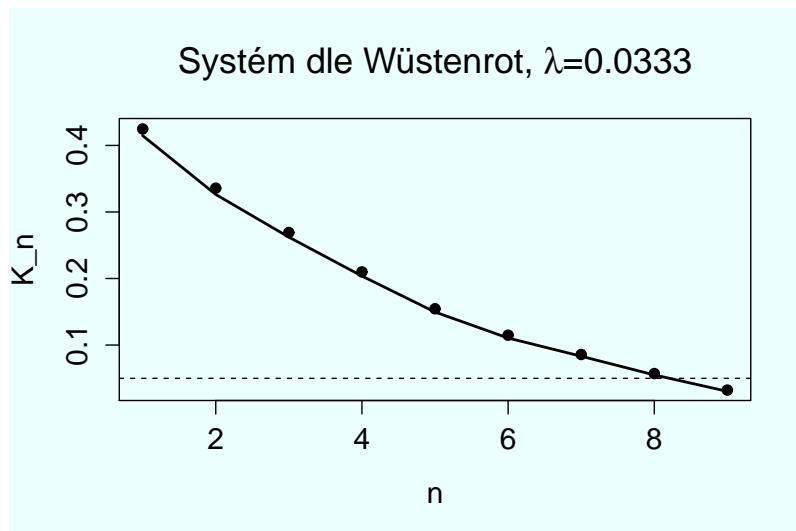
$$\begin{aligned} \mathbf{p}(0) = & (0.710804, \quad 0.067318, \quad 0.050207, \quad 0.037538, \quad 0.031999, \\ & 0.030194, \quad 0.020891, \quad 0.013802, \quad 0.014662, \quad 0.013417, \\ & 0.007517, \quad 0.000915, \quad 0.000483, \quad 0.000161, \quad 0.000092) \end{aligned} \quad (4.5)$$

je spíše přibližné.

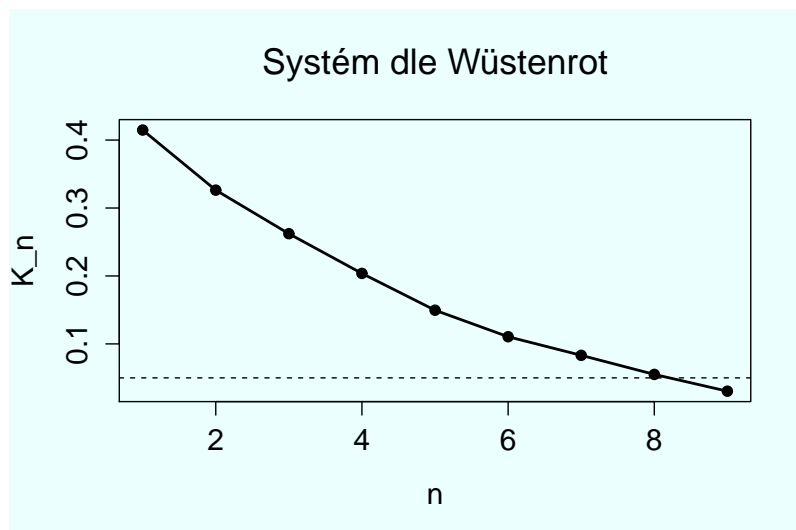
Konvergence ke stacionárnímu rozdělení

Pro zadaný parametr λ , např. 0.0333, můžeme systém považovat za stabilizovaný v okamžiku, kdy je součet absolutních odchylek absolutních pravděpodobností jednotlivých (patnácti) stavů v tomto čase od stacionárních pravděpodobností stavů (tento součet jsme v teoretické části práce značili K_n) pod nějakou zadanou hodnotou ϵ , např. 0.05. Takže podle podmínky stabilizace systému (2.7) pro řidiče s očekávaným ročním počtem škod $\lambda = 0.0333$ dostaneme hodnoty K_n , uvedené v tabulce 4.7 (2. sloupec) a graficky znázorněné na obr. 4.1.

V případě, že současnou roční očekávanou škodní frekvenci neznáme, pracujeme s podmínkou (3.6). Dopustili jsme se zde jisté úpravy, neboť máme 15 stavů Markovova řetězce, ale pravděpodobnost, že náhodná veličina L_∞ nabude hodnoty l , jsme definovali pro 14 původních bonus-malus tříd. Výchozí rozdělení sice můžeme upravit na 14 pravděpodobností (přímo z dat, neboť počáteční rozdělení s patnácti pravděpodobnostmi je právě z něj odvozené, přibližně stanovené), ale stochastickou matici \mathbf{P} máme řádu 15. Takže na místo stacionárního rozdělení definovaného v případě přidání fiktivních tříd v (3.10) zde pracujeme se stacionárními pravděpodobnostmi (3.2), ovšem pro 15 stavů l , tj. z množiny S (4.2), respektive (pro lepší zacházení) z množiny S_{15} (4.1). Podmínku stabilizace (3.6) pro součty K_n tedy uvažujeme pro tento počet stavů. Pro náhodného řidiče dostaneme hodnoty K_n , uvedené v tabulce 4.7 (3. sloupec) a graficky znázorněné na obr. 4.2.



Obrázek 4.1: Konvergence ke stacionárnímu rozdělení řidiče se známou hodnotou aktuální roční očekávané frekvence škod $\Lambda\Theta = \lambda = 0.0333$. Přerušovaná čára značí hodnotu $\epsilon = 0.05$.



Obrázek 4.2: Konvergence ke stacionárnímu rozdělení řidiče s neznámou hodnotou aktuální roční očekávané frekvence škod $\Lambda\Theta$. Přerušovaná čára značí hodnotu $\epsilon = 0.05$.

n	K_n	
	$\lambda = 0.0333$	neznámý řidič
1	0.424736	0.414717
2	0.335802	0.326310
3	0.269123	0.262206
4	0.209913	0.203748
5	0.154522	0.149591
6	0.114841	0.110575
7	0.085816	0.083194
8	0.057150	0.055105
9	0.032363	0.030435

Tabulka 4.7: Hodnoty K_n po n letech v systému, pro konkrétní a neznámou skutečnou roční očekávanou frekvenci škod.

Vidíme, že pro oba případy vyšlo přibližování systému do stabilizovaného stavu značně podobně. S ohledem na zadané počáteční rozdělení (4.5) můžeme říci, že se bonus-malus systém ustálí do devíti let. To platí pro jakéhokoliv řidiče z portfolia. Pro klienta s takovým parametrem λ , pro nějž bude dané počáteční rozdělení hodně blízké stacionárnímu, se systém dle zadané podmínky stabilizace může ustálit dříve, v případě velké odlišnosti počátečních pravděpodobností od těch stacionárních zase později. Víme, že stacionární rozdělení nezávisí na počátečním zařazení pojistníků do tříd. Na zvoleném výchozím rozdělení ovšem závisí rychlost konvergence ke stacionárnímu rozdělení, čili za kolik let budeme moci (dle zadané podmínky stabilizace) systém považovat za ustálený.

n	C_n
1	25.999723
5	18.504978
10	9.159118
15	1.052855
16	0.405197
17	0.243411
18	0.082955
19	0.049132
33	0.000001
34	0.000000

Tabulka 4.8: Hodnoty C_n po n letech v systému, pro skutečnou roční očekávanou frekvenci škod $\lambda = 0.0333$ a bez znalosti počátečního rozdělení.

Pokud neznáme počáteční rozdělení bonus-malus tříd, lze se na konvergenci ke stacionárnímu rozdělení dívat na základě (2.5). Pro námi zvolený parametr $\lambda = 0.0333$ budeme konstruovat n -té mocniny matice \mathbf{P} , tj. budeme postupně dostávat pravděpodobnosti přechodů mezi patnácti stavy Markovova řetězce (třídami bonus-malus včetně fiktivních tříd) po n krocích (letech). Matice \mathbf{P}^n se budou blížit matici $\mathbf{\Pi}$ složené z řádkových vektorů stacionárních pravděpodobností. V příloze B jsou (ve formátu z R) zobrazeny matice \mathbf{P}^n postupně pro n rovné číslu 1, 25, 33 a 34. Pro porovnání prvků matice \mathbf{P} a matice $\mathbf{\Pi}$ zavedeme kontrolní

součet

$$C_n = \sum_{i=1}^{15} \sum_{j=1}^{15} \left| p_{ij}^{(n)} - \pi_j \right|, \quad n \in \mathbb{N}.$$

V tabulce 4.8 vidíme tyto kontrolní součty pro určité hodnoty n .

S ohledem na zaokrouhlování na šest desetinných míst můžeme říci, že k dosažení stacionárního rozdělení dojde do 34 let. Ať je tedy počáteční rozdělení bonus-malus tříd jakékoliv, určili jsme pro zadaný parametr λ horní hranici pro počet let potřebných pro dosažení stacionárního rozdělení. Pro porovnání, kontrolní součet K_n v situaci se zadaným počátečním rozdělením a stejnou zadanou hodnotou parametru λ , po zaokrouhlení K_n na 6 desetinných míst, vyjde roven nule při $n = 23$.

4.5 Relativity

Nyní máme všechny poklady pro to, abychom určili optimální relativity pro 14 bonus-malus tříd, minimalizující kvadratickou ztrátovou funkci (3.3). Jelikož jsme do systému přidali dvě fiktivní podtřídy na místo jedné malusové třídy M2, spočítáme relativity na základě (3.9). Dále dle (3.11) určíme podmíněnou střední hodnotu očekávané roční škodní frekvence Λ za podmínky, že náhodně vybraný klient skončí v dlouhodobém časovém horizontu v l -té bonus-malus třídě. Rozmístění klientů do jednotlivých tříd v teoreticky nekonečném čase jsme spočetli v (4.4). Všechny tyto výsledky jsou zapsané v tabulce 4.9.

Třída	l	$\mathbb{P}[L_\infty = l]$	$r_l = \mathbb{E}[\Theta \mid L_\infty = l]$	$\mathbb{E}[\Lambda \mid L_\infty = l]$
B10	1	0.960632	0.908140	0.033248
B9	2	0.029900	2.642688	0.036375
B8	3	0.004372	4.170799	0.040210
B7	4	0.002364	4.662700	0.041904
B6	5	0.000877	5.768414	0.046083
B5	6	0.000475	6.428912	0.049402
B4	7	0.000284	7.039359	0.053122
B3	8	0.000194	7.505049	0.056737
B2	9	0.000147	7.852306	0.060244
B1	10	0.000123	8.128771	0.063816
Z	11	0.000111	8.333367	0.067450
M1	12	0.000109	8.506273	0.071368
M2	13	0.000247	8.863649	0.079120
M3	14	0.000173	9.085688	0.087919

Tabulka 4.9: Numerické charakteristiky systému dle pojišťovny Wüstenrot s přidáním fiktivními třídami.

Hodnota kvadratické ztrátové funkce $\text{MSE}(r_{L_\infty}) = \mathbb{E}[(\Theta - r_{L_\infty})^2]$ je

$$\text{MSE}(r_{L_\infty}) = 1.695601.$$

Můžeme se snadno přesvědčit, že očekávaná hodnota $E[r_{L_\infty}]$ relativní sazby r_{L_∞} je (po zaokrouhlení) rovna jedné neboli sto procentům, neboť platí

$$E[r_{L_\infty}] = \sum_{l=1}^{14} r_l P[L_\infty = l] = 1.$$

Je tedy splněna finanční rovnováha systému po dosažení stacionárního rozdělení tříd. Od „nebourajícího“ řidiče pojišťovna vybere na pojistném méně než od „bou-
rajícího“, ale od všech řidičů dohromady obdrží stejnou částku jako v situaci bez použití systému bonus-malus, tj. dostane součet základního ročního pojistného přes celé portfolio smluv povinného ručení.

V tabulce vidíme rostoucí průběhy $E[\Theta | L_\infty = l]$ i $E[\Lambda | L_\infty = l]$ jako funkcí proměnné l , což je určitě vítaná vlastnost vzhledem k tomu, že v každé další třídě směrem k okrajové třídě M3 by měl klient zaplatit větší násobek základního pojistného než v předešlé (lepší) třídě, a vyšším třídám l by měla odpovídat vyšší očekávaná hodnota aposteriorního rozdělení veličiny Λ .

Numerické výsledky nyní budeme interpretovat v procentech. Naprostá většina pojistníků (96%), bude v dlouhodobém horizontu zařazena do nejlepší třídy B10 a bude odměněna slevou na pojistném ve výši (jen) 9.2%. Jinými slovy bude platit 90.8% ze základního pojistného. Střední hodnota očekávané roční škodní frekvence Λ za předpokladu ustálení v této třídě je nízká, zhruba 3.32%. V druhé nejlepší třídě B9 skončí jen necelých 3% klientů a ti již budou platit přírážku ve výši 164.3% k základnímu pojistnému, tj. budou nakonec platit 264.3% apriorně nastaveného pojistného. Zbylé necelé 1% klientů skončí v teoreticky nekonečném časovém horizontu v ostatních dvanácti třídách, pravděpodobnosti jsou seřazeny sestupně od třídy B10 až k třídě M1. Pravděpodobnost, že klient skončí ve třídě M2, je vyšší než pravděpodobnost zařazení do lepší třídy M1, neboť třída M2 zahrnuje oproti ostatním třídám období dvou let. Poslední stacionární pravděpodobnost, týkající se třídy M3, tj. třídy s největší zápornou rozhodnou dobou, je nižší než pravděpodobnost zařazení do třídy M2, ale vyšší než pravděpodobnosti týkající se pro klienta finančně přijatelnějších tříd M1 až B2. To můžeme chápat jako důsledek principu zařazování do tříd dle rozhodné doby, která pro okrajovou malusovou třídu M3 může být v záporném smyslu teoreticky až nekonečná, a tedy se netýká období jen jednoho roku dle našeho omezení.

Co se týče hlavního výsledku, optimálních relativit, dostali jsme sazby poněkud odlišné od těch, které používá pojišťovna Wüstenrot. Optimální relativní sazby se zvyšují s rostoucí hodnotou l , což je žádoucí jev. Ovšem jedinou slevu na pojistném zaznamenává nejlepší bonusová třída B10, zatímco všechny ostatní (včetně těch, které pojišťovna nastavila jako bonusové) zaznamenávají přírážku k pojistnému, a to poměrně značnou (až 808.6%, respektive 908.6% v relativním vyjádření). Jelikož má ale $E[r_{L_\infty}]$ být rovna jedné a podíly pojistníků v třetí až čtrnácté třídě jsou minimální, musí pak být o to větší malusy v těchto třídách.

4.6 Systém $-1/Top$

Zkusme nyní na data od pojišťovny Wüstenrot použít daleko přísnější systém, který je v [5] nazýván jako systém $-1/Top^2$. Jedná se o systém se šesti třídami, číslovanými 1 (třída s nejvyšším bonusem) až 6 (třída s nejvyšším malusem). Nový klient je zařazen do vstupní třídy 6, za každý rok bez nehody se posune z třídy i do třídy $i - 1$, je-li to možné, za každou pojistnou událost sestoupí do nejhorší třídy 6, pokud v ní zařazen již není. Odtud pramení pojmenování tohoto teoretického systému. I ojedinělá nehoda klienta je tedy tvrdě penalizována a řidiči pak bude trvat dalších pět let, bude-li jezdit bez nehod, aby se dostal do nejlepší třídy 1. Další škoda ve stejném roce pak již penalizována není, protože klient již v nejhorší třídě je, a tak v ní i nadále zůstane. Přejímová pravidla jsou jednoduchá a jsou znázorněna v tabulce 4.10. Matice $T(0)$ a $T(k)$ pro $k \geq 1$ z ní dostaneme jednoduše a matici pravděpodobností přechodu \mathbf{P} též.

Startující třída	Následující třída po nahlášení	
	0	≥ 1
	nehod během 1 roku	
1	1	6
2	1	6
3	2	6
4	3	6
5	4	6
6	5	6

Tabulka 4.10: Přejímová pravidla systému $-1/Top$.

Zvolme počáteční rozdělení tříd např.

$$\mathbf{p}(0) = (0.95, 0.01, 0.01, 0.01, 0.01, 0.01).$$

Jak je to s přibližováním absolutních pravděpodobností rozmístění pojištěnců do tříd (po n letech od výchozího rozdělení) k rozdělení stacionárnímu (po teoreticky nekonečném počtu let)? Hodnoty K_n vypočtené dle (3.6) pro $n = 1, \dots, 5$ se zvoleným počátečním rozdělením jsou zapsány v tabulce 4.11. Konvergence systému $-1/Top$ do ustáleného stavu je též graficky znázorněna na obr. 4.3.

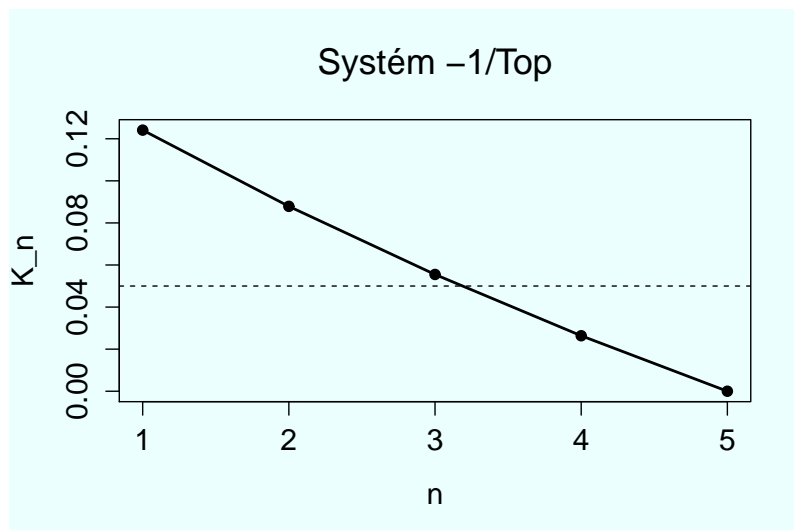
n	K_n
1	0.124103
2	0.087876
3	0.055503
4	0.026370
5	$4.651 \times 10^{-17} \approx 0$

Tabulka 4.11: Hodnoty K_n po n letech v systému $-1/Top$, pro neznámou skutečnou roční očekávanou frekvenci škod.

Vidíme, že součet absolutních odchylek stacionárních pravděpodobností od absolutních pravděpodobností v konečném čase je pod hodnotou 0.05 již po čtyřech

²V [5] jsou třídy číslované od 0 do 5, my se však držíme číslování od č. 1 v konzistenci se značením v celé této práci.

letech ($n_0 = 4$). Obecně se bez ohledu na volbu počátečního rozdělení systém ustálí do pěti let, což lze nahlédnout např. v [5], kde pro zadaný parametr λ je pátá mocnina matice pravděpodobností přechodu složena z řádkových vektorů stacionárních pravděpodobností, tj. $\mathbf{P}^5 = \mathbf{\Pi}$. Můžeme tento výsledek interpretovat tak, že řidiči jezdícímu bez nehod, který je ale na počátku automaticky zařazen do nejhorší třídy 6, trvá následujících pět let, než projde systémem do nejlepší třídy 1.



Obrázek 4.3: Konvergence ke stacionárnímu rozdělení řidiče s neznámou hodnotou aktuální roční očekávané frekvence škod $\Lambda\Theta$ v systému $-1/Top$. Přerušovaná čára značí hodnotu $\epsilon = 0.05$.

Numerické charakteristiky tohoto systému v podobě stacionárních pravděpodobností náhodného řidiče vypočtených dle (3.2), optimálních relativit (minimalizujících kvadratickou ztrátovou funkci) vypočítaných podle vzorce (3.5) a podmíněných středních hodnot veličiny Λ (za předpokladu ustálení systému v jednotlivých třídách) vypočtených dle (3.7) jsou zobrazeny v tabulce 4.12.

l	$P[L_\infty = l]$	$r_l = E[\Theta L_\infty = l]$	$E[\Lambda L_\infty = l]$
1	0.86750440	0.7595206	0.03300218
2	0.02207979	2.2728750	0.03558792
3	0.02394178	2.3921403	0.03588972
4	0.02610555	2.5270669	0.03625935
5	0.02865479	2.6820286	0.03672863
6	0.03171402	2.8641087	0.03735700

Tabulka 4.12: Numerické charakteristiky systému $-1/Top$.

V porovnání s předchozím systémem vyšly maximální přírážky k základnímu pojistnému výrazně nižší (maximální malus je 186.4% oproti 808.6%, respektive maximální relativní sazba je 286.4% oproti 908.6%), neboť systém daleko více penalizuje řidiče za způsobené nehody. Z toho důvodu se do nejlepší třídy 1 zařadí méně pojistníků než v předchozím systému (86.8% oproti 96.1%) s vyšší slevou na pojistném (24% oproti 9.2%, respektive minimální relativní sazba 76% oproti 90.8%). Tedy, když už se pojistník v dlouhodobém časovém horizontu

bude přes přísné penalizační pravidlo pohybovat v nejlepší třídě, je třeba ho také ocenit výraznějším bonusem, než jak tomu bylo v předchozím, ne tak přísném systému. Vzhledem k nastaveným přechodovým pravidlům skončí v nejhorší třídě 6 dokonce více klientů než ve třídě 2, 3, 4 nebo 5. Dále opět platí jednotková očekávaná hodnota optimální relativity r_{L_∞} , tj.

$$\mathbf{E}[r_{L_\infty}] = \sum_{l=1}^6 r_l \mathbf{P}[L_\infty = l] = 1$$

a hodnota kvadratické ztrátové funkce vyjde

$$\text{MSE}(r_{L_\infty}) = 1.580489,$$

což je o jednu desetinu méně než v předchozím případě.

Závěr

V této práci jsme se zabývali systémy bonus-malus užívanými pojišťovnami v oblasti povinného ručení. Pomocí Markovova řetězce jsme modelovali pohyb klienta v systému v závislosti na počtu škodních nároků uplatněných v předchozím roce. Objasnili jsme, proč je adekvátní použít negativně binomické rozdělení pro počty pojistných událostí v portfoliu povinného ručení pojišťovny. Uvažovali jsme, že pojišťovna apriorně segmentuje řidiče. Jednoduše řečeno to znamená, že při uzavření pojistné smlouvy je klient s pozorovatelnými charakteristikami „dobrého“ řidiče pojišťovnou oceněn základní pojistnou sazbou v nízké cenové hladině, zatímco klientovi, který je apriorně ohodnocen jako „špatný“ řidič, pojišťovna vyměří vyšší pojistnou sazbu. Očekávání samozřejmě nemusí odpovídat skutečnosti.

Bonus-malus systém se skládá z určitého počtu tříd a pravidel přechodu mezi nimi. Každé třídě odpovídá určitá relativní sazba pojistného, kterou jsme nazývali také relativitou či bonus-malus sazbou. Řekli jsme si, že relativní sazba v desetinném vyjádření udává hodnotu, kterou musíme přenásobit základní pojistnou sazbu, abychom dostali spravedlivější výslednou sazbu, kterou má z dlouhodobého hlediska klient hradit s ohledem na jeho aktuální rizikovost (nehodovost). Naším úkolem bylo uvést, jakým způsobem lze tyto relativní sazby nastavit optimálně.

V teoretické části práce jsme tedy vyložili možný způsob určení optimálních bonus-malus sazeb na základě minimalizace střední kvadratické funkce. Uvedené vzorce jsou založeny na stacionárním rozdělení, které ukazuje, jak bude stávající pojistný kmen v dlouhodobém časovém horizontu rozdělen do bonus-malus tříd. Do výpočtu optimálních sazeb jsme zahrnuli též informaci o apriorním roztržení klientů.

V praktické části jsme se zaměřili na bonus-malus systémy používané pojišťovnami v České republice, kde je pro přiznání bonusu a uplatnění malusu rozhodující tzv. rozhodná doba klienta. Pro aplikaci teoretického aparátu, popsaného v prvních třech kapitolách, jsme vybrali bonus-malus systém, který aktuálně používá pojišťovna Wüstenrot v povinném ručení. Reálný systém jsme museli nejprve upravit vymezením hranic pro rozhodnou dobu a přidáním fiktivních tříd kvůli splnění Markovské vlastnosti. Pojišťovna Wüstenrot nám poskytla data o segmentaci řidičů a údajích o počtech pojistných událostí za rok 2012, na kterých jsme odhadli potřebné parametry a následně spočítali optimální relativní sazby a další numerické charakteristiky.

Analyzovali jsme systém, uplatňovaný touto pojišťovnou v České republice, a můžeme konstatovat, že není pro řidiče příliš přísný. Jeden uplatněný škodní nárok za rok je penalizován sestupem pouze maximálně o jednu třídu, další škoda pak maximálně o dvě třídy. Došli jsme k závěru, že v dlouhodobém časovém horizontu několik málo nahlášených nehod nebude hrát výraznou roli, protože

rozhodná doba se s délkou trvání smlouvy stále navyšuje. Čím delší je rozhodná doba, tím vyššího stupně bonusu lze dosáhnout, přičemž rozhodná doba nad 10 let již pro klienta představuje maximální možnou slevu na základním pojistném.

Řekli jsme si, že systém s daným počátečním rozdělením, které jsme přibližně určili z dostupných dat na základě aktuálního zařazení klientů do bonus-malus tříd, můžeme považovat za finančně rovnovážný zhruba po devíti letech. Absolutní pravděpodobnosti ale nejsou ani po dalších letech totožné stacionárním, jsou si jen víc a víc blízké. Optimální relativity jsou počítány s (na počátečním rozdělení nezávislým) rozdělením tříd v nekonečnu, což pro systém založený na rozhodné době klienta s nízkou průměrnou škodní frekvencí (kolem 3.3%) a s poměrně mírnou penalizací znamená, že v tomto teoreticky nekonečném čase bude většina smluv zařazena do nejlepší třídy a jen malý zlomek klientů se bude pohybovat ve třídách s rozhodnou dobou o délce několika málo let či rozhodnou dobou zápornou. Např. v nejhorší třídě skončí jen přibližně 17 smluv z 100000. Budou to řidiči způsobující značný počet nehod, pak je na místě těmto třídám přiřadit adekvátně vysoké relativní sazby, což také potvrzují naše numerické výstupy.

Přísnost systému, můžeme-li to takto nazvat, má opravdu značný vliv na výsledné optimální relativity. Vypozorovali jsme to z porovnání zmíněných výsledků pro zkoumaný systém s výsledky pro hypotetický a poměrně přísný systém - $1/Top$, s využitím stejných dat, kde výsledkem bylo daleko nižší malusové navýšení pojistného a vyšší relativní sazba pro nejlepší třídu.

Pokud bychom chtěli zodpovědět v úvodu položenou otázku, tak musíme bohužel konstatovat, že se aplikovaný model pro stanovení optimálních bonus-malus sazeb příliš nehodí pro bonus-malus systém navržený pojišťovnou Wüstenrot. Dovolíme si toto tvrzení rozšířit na systémy užívané pojišťovnami na českém trhu povinného ručení, které používají shodný princip založený na rozhodné době.

Hlavní problém tkví ve stacionárním rozdělení, na němž jsou založeny všechny výpočty. Pro české bonus-malus systémy by se hodil model, který není založený na chování systému v nekonečném čase. Závěrem můžeme ještě dodat, že vinou současně nastaveného konkurenčního prostředí v oblasti povinného ručení v České republice, kdy jsou spotřebitelům nabízeny pojistné produkty v nízké cenové relaci, nelze z hlediska rentability aplikovat optimální výsledky pojistné matematiky, protože by pojištění byla příliš drahá.

Literatura

- [1] Andrade e Silva, J. M. a Centeno, M. d. L. (2005). A Note on Bonus Scales. *Journal of Risk and Insurance*, vol. **72**, no. 4, s. 601–607. ISSN: 0022-4367.
- [2] Brouhns, N., Guillén, M., Denuit, M. a Pinguet, J. (2003). Bonus-Malus Scales in Segmented Tariffs With Stochastic Migration Between Segments. *Journal of Risk and Insurance*, vol. **70**, no. 4, s. 577–599. ISSN: 0022-4367.
- [3] Centeno, M. d. L. a Andrade e Silva J. M. (2002). Optimal Bonus Scales Under Path-Dependent Bonus Rules. *Scandinavian Actuarial Journal*. vol. **2002**, no. 2, s. 129–136. ISSN: 0346-1238.
- [4] Denuit, M. a Dhaene, J. (2001). Bonus-Malus scales using exponential loss functions. *Blätter der DGVM*, vol. **25**, no. 1, s. 13–27. ISSN: 1864-0281.
- [5] Denuit, M., Maréchal, X., Pitrebois, S. a Walhin, J.-F. (2007). *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*. John Wiley & Sons, Chichester. ISBN: 978-0-470-02677-9.
- [6] Jackman, S. (2012). Package "pscl": verze 1.04.4. CRAN. [online]. 2012-06-13 [cit. 2013-07-13]. Dostupné z: <http://www.cran.r-project.org/web/packages/pscl/pscl.pdf>.
- [7] Pitrebois, S., Denuit, M. a Walhin, J.F. (2004). Bonus-malus scales in segmented tariffs: Gilde & Sundt's work revisited. *Australian Actuarial Journal*, vol. **10**, s. 107–125. ISSN: 1442-3065.
- [8] Prášková, Z. a Lachout, P. (2001). *Základy náhodných procesů*. Karolinum, Praha. ISBN 80-7184-688-0.
- [9] Ripley, B. (2002). Package "MASS": verze 7.3-27. CRAN. [online]. 2013-07-01 [cit. 2013-07-13]. Dostupné z: <http://www.cran.r-project.org/web/packages/MASS/MASS.pdf>.
- [10] Rolski, T., Schmidli, H., Schmidt, V. a Teugels, J (1999). *Stochastic Processes for Insurance and Finance*. John Wiley & Sons, Chichester. ISBN: 978-0-471-95925-0.
- [11] Slavíček, O. (2011). Bonus-malus Systems: Theory and Practice. *Studia ubezpieczeniowe: Zarządzanie ryzykiem i finansami*, vol. Neuveden, no. 182, s. 308–317. ISSN: 1689-7374.

- [12] *Všeobecné pojistné podmínky pro pojištění odpovědnosti za škodu způsobenou provozem vozidla VPP - POV 2011*. Wüstenrot pojišťovna a.s., Praha, 2011. [online]. [cit. 2013-07-20]. Dostupné z: <http://www.wuestenrot.cz/o-nas/pojistovna/Pojistne-podminky/>.
- [13] Wagnerova, E. *Porovnání modelu kredibilitního a systému bonus-malus*. Plzeň, 2007. Bakalářská práce. Západočeská univerzita v Plzni, Fakulta aplikovaných věd, Katedra matematiky.
- [14] Zeileis, A., Kleiber, C. a Jackman, S. (2008). Regression Models for Count Data in R. *Journal of Statistical Software*, vol. **27**, no. 8, s. 1–25. ISSN: 1548-7660. [online]. [cit. 2013-07-18]. Dostupné z: <http://www.jstatsoft.org/v27/i08/>.

Příloha A

Odhady parametrů segmentačních tříd

Váhy w_k a roční očekávané škodní frekvence $\hat{\lambda}_k$ pro 60 segmentačních tříd.

k	FREKV	FO_PO	VEK	REGION	w_k	$\hat{\lambda}_k$
1	fr1	o1	vek1	reg1	0.000169	0.132752
2	fr2	o1	vek1	reg1	0.000502	0.151744
3	fr3	o1	vek1	reg1	0.000204	0.095860
4	fr1	o1	vek2	reg1	0.000354	0.088052
5	fr2	o1	vek2	reg1	0.000776	0.100649
6	fr3	o1	vek2	reg1	0.000487	0.063582
7	fr1	o1	vek3	reg1	0.019819	0.043954
8	fr2	o1	vek3	reg1	0.032106	0.050243
9	fr3	o1	vek3	reg1	0.102423	0.031739
10	fr1	o2	vek4	reg1	0.000331	0.067965
11	fr2	o2	vek4	reg1	0.000527	0.077688
12	fr3	o2	vek4	reg1	0.002468	0.049077
13	fr1	o1	vek1	reg2	0.000067	0.118466
14	fr2	o1	vek1	reg2	0.000181	0.135414
15	fr3	o1	vek1	reg2	0.000109	0.085544
16	fr1	o1	vek2	reg2	0.000119	0.078577
17	fr2	o1	vek2	reg2	0.000435	0.089818
18	fr3	o1	vek2	reg2	0.000189	0.056740
19	fr1	o1	vek3	reg2	0.010319	0.039224
20	fr2	o1	vek3	reg2	0.017331	0.044836
21	fr3	o1	vek3	reg2	0.045246	0.028324
22	fr1	o2	vek4	reg2	0.000112	0.060651
23	fr2	o2	vek4	reg2	0.000188	0.069328
24	fr3	o2	vek4	reg2	0.000777	0.043796
25	fr1	o1	vek1	reg3	0.000847	0.107915
26	fr2	o1	vek1	reg3	0.002004	0.123354
27	fr3	o1	vek1	reg3	0.001627	0.077926
28	fr1	o1	vek2	reg3	0.001297	0.071579
29	fr2	o1	vek2	reg3	0.001958	0.081819

Pokračování na další straně

Pokračování z předchozí strany						
k	FREKV	FO_PO	VEK	REGION	w_k	$\hat{\lambda}_k$
30	fr3	o1	vek2	reg3	0.002040	0.051687
31	fr1	o1	vek3	reg3	0.071616	0.035731
32	fr2	o1	vek3	reg3	0.089750	0.040843
33	fr3	o1	vek3	reg3	0.347456	0.025801
34	fr1	o2	vek4	reg3	0.000599	0.055249
35	fr2	o2	vek4	reg3	0.000691	0.063153
36	fr3	o2	vek4	reg3	0.004284	0.039895
37	fr1	o1	vek1	reg4	0.000354	0.093721
38	fr2	o1	vek1	reg4	0.000410	0.107129
39	fr3	o1	vek1	reg4	0.000536	0.067676
40	fr1	o1	vek2	reg4	0.000347	0.062163
41	fr2	o1	vek2	reg4	0.000636	0.071057
42	fr3	o1	vek2	reg4	0.000570	0.044888
43	fr1	o1	vek3	reg4	0.016420	0.031031
44	fr2	o1	vek3	reg4	0.016691	0.035470
45	fr3	o1	vek3	reg4	0.090201	0.022407
46	fr1	o2	vek4	reg4	0.000084	0.047982
47	fr2	o2	vek4	reg4	0.000100	0.054846
48	fr3	o2	vek4	reg4	0.000832	0.034648
49	fr1	o1	vek1	reg5	0.000057	0.151817
50	fr2	o1	vek1	reg5	0.000186	0.173536
51	fr3	o1	vek1	reg5	0.000116	0.109627
52	fr1	o1	vek2	reg5	0.000088	0.100697
53	fr2	o1	vek2	reg5	0.000425	0.115103
54	fr3	o1	vek2	reg5	0.000200	0.072713
55	fr1	o1	vek3	reg5	0.013801	0.050267
56	fr2	o1	vek3	reg5	0.024203	0.057458
57	fr3	o1	vek3	reg5	0.069803	0.036298
58	fr1	o2	vek4	reg5	0.000389	0.077725
59	fr2	o2	vek4	reg5	0.000845	0.088845
60	fr3	o2	vek4	reg5	0.003298	0.056125

Příloha B

Konvergence ke stacionárnímu rozdělení

Mocnění matice pravděpodobností přechodu \mathbf{P} a stacionární rozdělení $\boldsymbol{\pi}$ pro $\lambda = 0.0333$, zaokrouhlování na 6 desetinných míst. Pro důraz na zvolený parametr λ píšeme $\mathbf{P}(0.0333)$ a $\boldsymbol{\pi}(0.0333)$.

Následují výstupy z R. Matice jsou ve formátu jakýchsi tabulek o patnácti řádcích a patnácti sloupcích číslovaných dle množiny stavů S_{15} . Kvůli větším rozměrům jsou matice rozdělené na více částí, vždy pro všech patnáct řádků.

- $\mathbf{P}(0.0333)$

	1	2	3	4	5	6
1	0.967248	0.032209	0.000000	0.000536	0.000000	0.000006
2	0.967248	0.000000	0.032209	0.000000	0.000536	0.000000
3	0.000000	0.967248	0.000000	0.032209	0.000000	0.000536
4	0.000000	0.000000	0.967248	0.000000	0.032209	0.000000
5	0.000000	0.000000	0.000000	0.967248	0.000000	0.032209
6	0.000000	0.000000	0.000000	0.000000	0.967248	0.000000
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.967248
8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
14	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

	7	8	9	10	11	12
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.000006	0.000000	0.000000	0.000000	0.000000	0.000000
3	0.000000	0.000006	0.000000	0.000000	0.000000	0.000000
4	0.000536	0.000000	0.000006	0.000000	0.000000	0.000000
5	0.000000	0.000536	0.000000	0.000006	0.000000	0.000000
6	0.032209	0.000000	0.000536	0.000000	0.000006	0.000000
7	0.000000	0.032209	0.000000	0.000536	0.000000	0.000006

8	0.967248	0.000000	0.032209	0.000000	0.000536	0.000000
9	0.000000	0.967248	0.000000	0.032209	0.000000	0.000536
10	0.000000	0.000000	0.967248	0.000000	0.032209	0.000000
11	0.000000	0.000000	0.000000	0.967248	0.000000	0.032209
12	0.000000	0.000000	0.000000	0.000000	0.967248	0.000000
13	0.000000	0.000000	0.000000	0.000000	0.000000	0.967248
14	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

		13	14	15
1	0.000000	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.000000
7	0.000000	0.000000	0.000000	0.000000
8	0.000006	0.000000	0.000000	0.000000
9	0.000000	0.000006	0.000000	0.000000
10	0.000536	0.000000	0.000006	0.000000
11	0.000000	0.000536	0.000006	0.000000
12	0.032209	0.000000	0.000542	0.000000
13	0.000000	0.032209	0.000542	0.000000
14	0.967248	0.000000	0.032752	0.000000
15	0.000000	0.967248	0.032752	0.000000

• $P^{25}(0.0333)$

	1	2	3	4	5	6	7	8	9
1	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
2	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
3	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
4	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
5	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
6	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
7	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
8	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
9	0.965005	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
10	0.965005	0.032676	0.001648	0.000615	4.6e-05	9e-06	1e-06	0	0
11	0.965004	0.032677	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
12	0.965004	0.032675	0.001649	0.000615	4.6e-05	9e-06	1e-06	0	0
13	0.964991	0.032688	0.001647	0.000617	4.6e-05	9e-06	1e-06	0	0
14	0.964991	0.032672	0.001663	0.000615	4.8e-05	9e-06	1e-06	0	0
15	0.964903	0.032760	0.001650	0.000628	4.7e-05	1e-05	1e-06	0	0

	10	11	12	13	14	15
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0

4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0

• $P^{33}(0.0333)$

	1	2	3	4	5	6	7	8	9
1	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
2	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
3	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
4	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
5	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
6	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
7	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
8	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
9	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
10	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
11	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
12	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
13	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
14	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
15	0.965005	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0

	10	11	12	13	14	15
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0

• $P^{34}(0.0333)$

	1	2	3	4	5	6	7	8	9
1	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
2	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
3	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
4	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
5	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
6	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
7	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
8	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
9	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
10	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
11	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
12	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
13	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
14	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0
15	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0	0

	10	11	12	13	14	15
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0

• $\pi(0.0333)$

	1	2	3	4	5	6	7	8
	0.965006	0.032676	0.001647	0.000615	4.6e-05	9e-06	1e-06	0

	9	10	11	12	13	14	15
	0	0	0	0	0	0	0