4 Markovovy řetězce

Definice 7 *Posloupnost celočíselných náhodných veličin* $\{X_n\}_{n=0}^{\infty}$ *se nazývá Markovův řetězec (markovský řetězec), jestliže*

$$\mathbb{P}(X_{n+1} = j | X_n = i_n, \cdots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i_n) \tag{17}$$

pro každé $n \geq 0$, *a* $i_0, \cdots, i_n \in \mathbb{Z}$ *takové, že* $\mathbb{P}(X_n = i_n, \cdots, X_0 = i_0) > 0$.

Hodnoty n interpretujeme jako časové okamžiky, hodnoty, kterých řetězec X_n může nabývat, interpretujeme jako stavy. Množinu stavů řetězce budeme značit S.

Vztah (17) vyjadřuje markovskou vlastnost, tedy skutečnost, že výsledek v čase $n+1$ závisí pouze na stavu řetězce v čase n (stav v přítomném čase), nikoli na starší minulosti (posloupnosti realizovaných stavů před časem n).

Příklady – hody kostkou, náhodná procházka

Definice 8 *Hodnotu*

$$p_{ij}(n, n+1) = \mathbb{P}(X_{n+1} = j | X_n = i)$$

budeme nazývat pravděpodobnost přechodu ze stavu i *v čase* n *do stavu* j *v čase* $n+1$. *Markovův řetězec se nazývá homogenní, pokud hodnoty* $p_{ij} = p_{ij}(n, n+1)$ *nezávisí na hodnotě* n. *Označme* $p_i = \mathbb{P}(X_0 = i)$, *pak* $(p_i, i \in S)$ *označuje počáteční rozdělení řetězce.*

Dále budeme předpokládat, že daný Markovův řetězec je homogenní.

Věta 3 *Nechť* $\{X_n\}_{n=0}^{\infty}$ *je homogenní Markovův řetězec s množinou stavů* $S \subseteq \mathbb{Z}$, *počátečním rozdělením* $(p_i, i \in S)$ *a pravděpodobnostmi přechodu* $(p_{ij}, i, j \in S)$. *Pak*

$$\mathbb{P}(X_0 = i_0, \cdots, X_n = i_n) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}.$$

Důkaz: Indukcí podle n. Pro $n = 0, 1$ tvrzení zřejmě platí (z definice, resp. elementárním podmíněním). Pro $n = 2$ máme podle věty o podmíněné pravděpodobnosti

$$\begin{aligned}\mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2) &= \mathbb{P}(X_0 = i_0)\mathbb{P}(X_1 = i_1 | X_0 = i_0)\mathbb{P}(X_2 = i_2 | X_1 = i_1, X_0 = i_0) \\ &= p_{i_0} p_{i_0 i_1} p_{i_1 i_2},\end{aligned}$$

protože z markovské vlastnosti platí

$$\mathbb{P}(X_2 = i_2 | X_1 = i_1, X_0 = i_0) = \mathbb{P}(X_2 = i_2 | X_1 = i_1) = p_{i_1 i_2}.$$

Podobně pro $n > 2$:

$$\begin{aligned}\mathbb{P}(X_0 = i_0, \cdots, X_n = i_n) &= \mathbb{P}(X_n = i_n | X_0 = i_0, \cdots, X_{n-1} = i_{n-1})\mathbb{P}(X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) \\ &= p_{i_{n-1} i_n}\mathbb{P}(X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n},\end{aligned}$$

kde jsme nejdříve využili markovské vlastnosti a poté indukčního předpokladu. $\square$

Definujme si také pravděpodobnosti přechodu vyšších řádů (po více krocích):

$$p_{ij}^{(0)} = I\{i = j\}, \quad p_{ij}^{(1)} = p_{ij}, \quad p_{ij}^{(n+1)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}.$$

Tyto pravděpodobnosti se hodí zapisovat v maticovém tvaru, pak tedy pracujeme s maticí pravděpodobností přechodu (přechodovou maticí) $\mathbf{P} = (p_{ij})_{i,j\in S}$ a maticemi pravděpodobností přechodu po více krocích $\mathbf{P}^{(n)} = (p_{ij}^{(n)})_{i,j\in S}$.

Smysluplnost této definice pravděpodobností přechodu po více krocích ukazuje následující tvrzení.

Věta 4 *Nechť $m \geq 0$ je celé číslo, $i, j \in S$, pak $\mathbb{P}(X_{m+n} = j | X_n = i) = p_{ij}^{(m)}$.*

Důkaz: Indukcí podle m. □

Rovněž platí $\mathbf{P}^{(n)} = \mathbf{P}^n$ pro $n \geq 0$ (důkaz opět indukcí). Pravděpodobnosti přechodu vyšších řádů tedy můžeme získat jednoduše násobením matic.

Poznámka: Uvědomme si, že jak $\mathbf{P}$ tak i $\mathbf{P}^{(n)}$ jsou tzv. stochastické matice, tj. obsahují pouze čísla mezi 0 a 1 a součet každého řádku se rovná 1.

Skládání pravděpodobností přechodu ukazuje následující věta.

Věta 5 *(Chapmanova-Kolmogorova rovnost) pro všechna celá čísla $m, n \geq 0$ platí*

$$p_{ij}^{(m+n)} = \sum_{k\in S} p_{ik}^{(m)} p_{kj}^{(n)}.$$

Důkaz: Použijeme rozklad uvažovaného jevu podle všech možných stavů v čase m a postupné podmiňování:

$$\begin{aligned} p_{ij}^{(m+n)} &= \mathbb{P}(X_{m+n} = j | X_0 = i) = \sum_{k\in S} \mathbb{P}(X_{m+n} = j, X_m = k | X_0 = i) \\ &= \sum_{k\in S} \mathbb{P}(X_{m+n} = j | X_m = k, X_0 = i)\mathbb{P}(X_m = k | X_0 = i) = \sum_{k\in S} p_{ik}^{(m)} p_{kj}^{(n)}. \end{aligned}$$

V poslední rovnosti jsme použili markovskou vlastnost a homogenitu řetězce. □

4.1 Klasifikace stavů Markovova řetězce

Definice 9 *Řekneme, že stav $j \in S$ je dosažitelný ze stavu $i \in S$, pokud existuje $n > 0$ takové, že $p_{ij}^{(n)} > 0$. Píšeme $i \to j$. Markovův řetězec nazveme nerozložitelným, pokud je každý stav dosažitelný z každého stavu.*

Definice 10 *Řekneme, že $\emptyset \neq C \subset S$ je uzavřená množina, pokud žádný stav vně množiny C není dosažitelný z žádného stavu uvnitř množiny C.*

Věta 6 *$C \subset S$ je uzavřená množina právě tehdy, když platí $p_{jk} = 0$ pro každé $j \in C$ a $k \notin C$.*

Důkaz: Implikace zleva doprava je zřejmá. Opačnou implikaci dokážeme indukcí. Musíme ukázat, že $p_{jk}^{(l)} = 0$ pro každé $l \in \mathbf{N}$, $j \in C$ a $k \notin C$. Pro $l = 1$ to platí z předpokladu, pro $l > 1$ pišme

$$p_{jk}^{(l)} = \sum_{i\in S} p_{ji}^{(l-1)} p_{ik} = \sum_{i\in C} p_{ji}^{(l-1)} p_{ik} + \sum_{i\notin C} p_{ji}^{(l-1)} p_{ik} = \sum_{i\in C} p_{ji}^{(l-1)} \cdot 0 + \sum_{i\notin C} 0 \cdot p_{ik} = 0,$$

kde jsme v předposlední rovnosti využili indukční předpoklad. □

Věta 7 *Řetězec je rozložitelný právě tehdy, když existuje uzavřená množina $C \neq S$.*

Důkaz: Zřejmý.

Věta 8 *Markovův řetězec je rozložitelný právě tehdy, když po vhodném přečíslování stavů je matice přechodu tvaru* $\begin{pmatrix} P_1 & 0 \\ Q & R \end{pmatrix}$, *kde P_1, R jsou čtvercové matice.*

Důkaz: Přečíslujeme stavy tak, aby nejdříve byly očíslované stavy z C. Pak snadno nahlédneme, že nová matice pravděpodobností přechodu bude mít přesně žádaný tvar. $\square$

Definice 11 *Buď $i \in S$, potom*

$$\nu_i = \inf\{n > 0, X_n = i\}$$

se nazývá doba prvního vstupu (někdy se více hodí doba prvního návratu) řetězce do stavu i. Stav $i \in S$ se nazývá trvalý, pokud

$$\mathbb{P}_i(\nu_i < \infty) = \mathbb{P}(\nu_i < \infty | X_0 = i) = 1,$$

tedy řetězec, který vychází ze stavu i, se s pravděpodobností 1 vrátí do stavu i po konečně mnoha krocích. V opačném případě (kladná pravděpodobnost, že se řetězec do stavu i nikdy nevrátí) řekneme, že stav i je přechodný. Trvalý stav $i \in S$ se nazývá nulový, pokud $\mathbb{E}_i \nu_i = \infty$, a nenulový, pokud $\mathbb{E}_i \nu_i < \infty$.

Poznámka: Do trvalého stavu se Markovův řetězec s pravděpodobností 1 vrátí nekonečně-krát. Do přechodného stavu se Markovův řeězec vrátí nekonečně-krát s pravděpodobností 0.

Označme $f_i^{(n)} = P_i(\nu_i = n)$ a $f_i^{(0)} = 0$. Zřejmě stav i je trvalý právě tehdy, když $\sum_{n=1}^{\infty} f_i^{(n)} = 1$. Dále platí $\mathbb{E}_i \nu_i = \sum_{n=1}^{\infty} n f_i^{(n)}$.

Poznámka: $p_{ii}^{(n)} = P_i(X_n = i) = \sum_{k=1}^{n} f_i^{(k)} p_{ii}^{(n-k)}$ platí pro $n \geq 1$.

Definice 12 *Stav $i \in S$ homogenního Markovova řetězce se nazývá periodický s periodou d, pokud*

$$d = NSD\{n > 0, p_{ii}^{(n)} > 0\} > 1.$$

Pokud vychází $d = 1$, říkáme, že stav $i \in S$ je aperiodický (neperiodický).
Řekneme, že dva stavy jsou stejného typu, pokud jsou oba dva trvalé/přechodné, nulové/nenulové, neperiodické/periodické se stejnou periodou.

Věta 9 *Stav $i \in S$ je trvalý právě tehdy, když $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.*

Trvalý stav $i \in S$ je nulový právě tehdy, když $p_{ii}^{(n)} \to 0$ pro $n \to \infty$.

Důkaz: Pro $x \in (0, 1)$ definujme vytvořující funkce $P(x)$, $F(x)$ posloupností $p_{ii}^{(n)}, f_i^{(n)}$

$$P(x) = \sum_{n=0}^{\infty} p_{ii}^{(n)} x^n, \quad \text{a} \quad F(x) = \sum_{n=0}^{\infty} f_i^{(n)} x^n.$$

Z poznámky plyne, že pro $n \geq 1$ platí

$$p_{ii}^{(n)} x^n = \sum_{k=1}^{n} f_i^{(k)} p_{ii}^{(n-k)} x^n = \sum_{k=1}^{n} f_i^{(k)} x^k p_{ii}^{(n-k)} x^{n-k}$$

Sečtením této rovnosti přes $n \geq 1$ dostaneme

$$P(x) - 1 = \sum_{n=1}^{\infty} p_{ii}^{(n)} x^n = \sum_{n=1}^{\infty} \sum_{k=1}^{n} f_i^{(k)} x^k p_{ii}^{(n-k)} x^{n-k} = \left(\sum_{k=1}^{\infty} f_i^{(k)} x^k \right) \left(\sum_{l=0}^{\infty} p_{ii}^{(l)} x^l \right) = F(x)P(x).$$

Platí tedy $P(x) = 1/(1 - F(x))$. Dále

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \lim_{x \to 1^-} P(x) = \lim_{x \to 1^-} \frac{1}{1 - F(x)} = \infty$$

nastává právě tehdy, když

$$1 = \lim_{x \to 1^-} F(x) = \sum_{n=0}^{\infty} f_i^{(n)} = P_i(\nu_i < \infty).$$

Odtud je vidět, že stav $i \in S$ je trvalý právě tehdy, když $\sum_{i=0}^{\infty} p_{ii}^{(n)} = \infty$.

Náznak důkazu druhé části* : Zbylou část důkazu pouze naznačíme. Zřejmě

$$\begin{aligned} \frac{1 - F(x)}{1 - x} &= \frac{1}{1 - x} \left(\sum_{n=0}^{\infty} f_i^{(n)} - \sum_{n=0}^{\infty} f_i^{(n)} x^n \right) = \sum_{n=0}^{\infty} f_i^{(n)} \frac{1 - x^n}{1 - x} \\ &= \sum_{n=0}^{\infty} f_i^{(n)} (1 + \cdots + x^{n-1}) \to \sum_{n=0}^{\infty} f_i^{(n)} n = \mathbb{E}_i \nu_i, \end{aligned} \tag{18}$$

pro $x \to 1^-$. Limitní přechod v (18) dostaneme z Fatouova lemmatu

$$\liminf_{x \to 1^-} \sum_{n=0}^{\infty} f_i^{(n)} (1 + \cdots x^{n-1}) \geq \sum_{n=0}^{\infty} f_i^{(n)} n,$$

a z nerovnosti $f_i^{(n)} (1 + \cdots + x^{n-1}) \leq f_i^{(n)} n$ pro $x \in (0, 1)$. Víme tedy, že trvalý stav $i \in S$ je nulový právě tehdy, když

$$\lim_{x \to 1^-} (1 - x) P(x) = \lim_{x \to 1^-} \frac{1 - x}{1 - F(x)} = 0.$$

Zbývá tedy ukázat, že $(\lim_{x \to 1^-} (1 - x) P(x) = 0)$ je ekvivalentní $(p_{ii}^{(n)} \to 0)$ pro $n \to \infty$. Jednu implikaci lze obdržet z Tauberovy věty, druhá je o něco (hodně) složitější. Podrobnosti a odkazy na další literaturu lze najít ve skriptech Z. Práškové a P. Lachouta: Základy náhodných procesů. □

Věta 10 *Nechť $i \to j \to i$. Pak stavy i a j jsou stejného typu. Pokud $i \to j \not\to i$, pak stav i je přechodný. V nerozložitelném řetězci jsou všechny stavy stejného typu.*

Věta 11 *V řetězci s konečně mnoha stavy nemohou být všechny stavy přechodné. V řetězci s konečně mnoha stavy neexistují stavy trvalé nulové.*

Definice 13 *Stav, ze kterého není dosažitelný žádný jiný stav, nazveme stavem absorpčním. Stav, který je trvalý, nenulový a neperiodický, nazveme stavem ergodickým.*

4.2 Limitní a stacionární rozdělení Markovova řetězce

V této podkapitole se podíváme na dlouhodobé chování markovského řetězce, tedy jaké je chování náhodné veličiny X_n pokud $n \to \infty$.

Definice 14 *Rozdělení $\pi = (\pi_i, i \in S)$ nazveme stacionárním, pokud platí $\pi_j = \sum_{i \in S} \pi_i p_{ij}$ pro všechna $j \in S$. Maticově lze tuto soustavu rovností zapsat ve tvaru $\pi^T = \pi^T \mathbf{P}$.*
Rozdělení $\mathbf{a} = (a_i, i \in S)$ nazveme limitním, pokud pro každé $i \in S$ a každé počáteční rozdělení platí $a_i = \lim_{n\to\infty} p_i(n)$, kde $p_i(n) = \mathbb{P}(X_n = i)$.

Poznámka: Všimněte si, že na rozdíl od stacionárního rozdělení, které se vztahuje k pravděpodobnostem přechodu, se limitní rozdělení vztahuje k rozdělení celého markovského řetězce, jež je ovlivněno i počátečním rozdělením. Nicméně snadno nahlédneme, že je-li počáteční rozdělení markovského řetězce rovno stacionárnímu, pak i rozdělení řetězce ve všech následujících časech $(p_i(n), i \in S), n \in \mathbb{N}$, je rovno stacionárnímu rozdělení.

Věta 12 *Nechť existuje limitní rozdělení $\mathbf{a}$ řetězce $\{X_n\}_{n=0}^{\infty}$, pak toto rozdělení je jeho stacionárním rozdělením.*

Důkaz: Zřejmě

$$p_j(n+1) = \sum_{i \in S} p_i(n) p_{ij}.$$

Podle Fatouova lemmatu

$$a_j = \lim_{n\to\infty} p_j(n+1) = \liminf_{n\to\infty} \sum_{i \in S} p_i(n) p_{ij} \geq \sum_{i \in S} \liminf_{n\to\infty} p_i(n) p_{ij} = \sum_{i \in S} a_i p_{ij}. \tag{19}$$

Posčítáme-li předchozí nerovnosti přes $j \in S$, dojdeme k nerovnosti

$$1 = \sum_{j \in S} a_j \geq \sum_{i,j \in S} a_i p_{ij} = \sum_{i \in S} a_i \sum_{j \in S} p_{ij} = \sum_{i \in S} a_i = 1.$$

V neostré nerovnosti výše se tedy nabývá rovnosti, což dále znamená, že ve všech nerovnostech (19) musí nastávat rovnost. □

Věta 13 *V nerozložitelném řetězci existuje stacionární rozdělení právě tehdy, když všechny stavy jsou trvalé nenulové. V tom případě existuje právě jedno stacionární rozdělení. Jsou-li navíc stavy řetězce neperiodické, je stacionární rozdělení i rozdělením limitním, dokonce*

$$\lim_{n\to\infty} p_j(n) = \lim_{n\to\infty} p_{ij}^{(n)} = \pi_j = \frac{1}{\mathbb{E}_j \nu_j},$$

kde $\pi = (\pi_i, i \in S)$ je stacionární rozdělení a $\mathbb{E}_j \nu_j = \sum_{n=1}^{\infty} n f_j(n)$ je střední hodnota doby prvního návratu do stavu j. Jsou-li všechny stavy periodické, platí

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_j(k) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} = \pi_j.$$

Poznámka: V konečném nerozložitelném řetězci stacionární rozdělení vždy existuje.

Stacionární rozdělení nerozložitelného řetězce necharakterizuje jen limity absolutních pravděpodobností jednotlivých stavů. Charakterizuje také četnosti návratu do jednotlivých stavů, resp. čas řetězcem v jednotlivých stavech strávený.

Věta 14 *Označme*

$$N_j(n) = \sum_{k=1}^{n} I(X_k = j), \quad j \in S,$$

náhodnou veličinu udávající, kolikrát v prvních n krocích projde řetězec stavem j. V nerozložitelném řetězci s trvalými nenulovými neperiodickými stavy platí

$$\lim_{n\to\infty} \frac{N_j(n)}{n} = \pi_j \quad \text{s pravděpodobností 1}$$

pro každé $j \in S$.

4.3 Systém bonus-malus v pojišťovnictví

Tento systém se používá v havarijním pojištění. Systém má určitý počet tříd výše pojistného (bonusové třídy), do kterých je pojištěnec přiřazován na základě počtu škod uplatněných za pojistné období. Cílem tohoto systému je snížení heterogenity pojistného kmene, odrazení pojištěnců od uplatnění nízkých nároků a umožnění stanovení pojistného, jehož výše lépe odpovídá individuálním rizikům.

Aktuální bonusová třída pojištěnce a počet škod v daném období určí bonusovou třídu a tím i výši pojistného v dalším období. Vhodným modelem je proto homogenní markovský řetězec.

Označme $S = \{1, \cdots, k\}$ bonusové třídy a $a = (a_1, \cdots, a_k)^T$ výše pojistného jako podílnou část základu. Hodnota X_n bude označovat bonusovou třídu pojištěnce v období n, jinými slovy, řetězec $\{X_n\}$ bude v čase n ve stavu i, pokud bude pojištěnec zařazen do i-té bonusové třídy. Symbolem k_0 budeme označovat třídu, do které je pojištěnec zařazen na začátku, tedy v době příchodu do systému.

Španělský systém

$k = 5, a = \frac{1}{100}(60, 70, 80, 90, 100)^T, k_0 = 5.$

Pokud bude pojištěnec v předcházejícím období uplatňovat nárok na výplatu, přesune se v následujícím kroku do 5-té třídy, pokud nebude, posune se o třídu níže, je-li to možné.

Nechť pravděpodobnost neuplatnění žádného nároku v daném období je p. Matice přechodu mezi třídami je potom tvaru

$$P = \begin{pmatrix} p & 0 & 0 & 0 & 1-p \\ p & 0 & 0 & 0 & 1-p \\ 0 & p & 0 & 0 & 1-p \\ 0 & 0 & p & 0 & 1-p \\ 0 & 0 & 0 & p & 1-p \end{pmatrix}$$

Obvykle předpokládáme, že individuální riziko pojištěnce je charakterizováno rizikovým parametrem λ tak, že počet pojistných událostí j má Poissonovo rozdělení: $p_j = \frac{\lambda^j}{j!} e^{-\lambda}$,

$j = 0, 1, 2, \ldots$ O výši škod Y_l budeme předpokládat, že jsou to nezávislé, stejně rozdělené náhodné veličiny, nezávislé na hodnotě parametru λ.

Britský systém

$k = 7, a = \frac{1}{100}(33, 40, 45, 55, 65, 75, 100)^T, k_0 = 6.$

Pokud v minulém období pojištěnec neuplatňoval nárok na výplatu, posune se o jednu třídu níže, je-li to možné.

Nechť naopak v minulém období uplatňoval nárok na výplatu jednou. Pokud byl v první třídě, posune se o 3 třídy výše. Pokud byl ve druhé nebo třetí, posune se o dvě třídy. V ostatních případech se posune o jednu třídu výše, je-li to možné.

Za každý další nárok na výplatu (druhý, třetí, atd.) se pojištěnec posune ještě o dvě třídy výše. Pokud není možné, aby se pojištěnec posunul o předepsaný počet tříd, skončí v poslední třídě. Matice přechodu tak bude tvaru

$$P = \begin{pmatrix} e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 0 & \frac{\lambda^2}{2}e^{-\lambda} & 1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}) \\ e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 0 & \frac{\lambda^2}{2}e^{-\lambda} & 1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}) \\ 0 & e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 0 & 1 - e^{-\lambda}(1 + \lambda) \\ 0 & 0 & e^{-\lambda} & 0 & \lambda e^{-\lambda} & 0 & 1 - e^{-\lambda}(1 + \lambda) \\ 0 & 0 & 0 & e^{-\lambda} & 0 & \lambda e^{-\lambda} & 1 - e^{-\lambda}(1 + \lambda) \\ 0 & 0 & 0 & 0 & e^{-\lambda} & 0 & 1 - e^{-\lambda} \\ 0 & 0 & 0 & 0 & 0 & e^{-\lambda} & 1 - e^{-\lambda} \end{pmatrix}$$

Poznámka: Nechť matice pravděpodobností přechodu (homogenního markovského řetězce s konečnou množinou stavů) má kladný sloupec, potom existuje stacionární rozdělení $\pi = (\pi_i, i \in S)$, přičemž $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$.

Střední hodnotu pojistného $\mathbb{E}C$ určíme pomocí stacionárního rozdělení. $\mathbb{E}C$ je rovna (vzhledem k základnímu pojistnému odpovídajícímu číslu 1):

$$EC = \pi_1 a_1 + \cdots + \pi_k a_k$$

Příklad: Španělský model, pravděpodobnost, že nedošlo ke škodě je $p = e^{-\lambda}$, a tedy pravděpodobnost uplatnění nároku je $1 - p = 1 - e^{-\lambda}$. Stacionární rozdělení je potom rovno

$$\pi = (p^4, p^3(1-p), p^2(1-p), p(1-p), 1-p)^T.$$

V tomto modelu můžeme počítat i další charakteristiky jako například pravděpodobnost uplatnění pojistné události.

Pravděpodobnost uplatnění pojistné události

Předpokládejme, že k uplatnění nastalé pojistné události dochází, pokud je škoda Y vyšší, než je odpovídající navýšení pojistného c v dalším roce po uplatnění. Tedy

$$\mathbb{P}(\text{uplatnění}) = \mathbb{P}(\text{uplatnění}|\text{nastává})\mathbb{P}(\text{nastává}) = \mathbb{P}(Y > c)\left(1 - e^{-\lambda}\right),$$

kde $1 - e^{-\lambda}$ je pravděpodobnost výskytu pojistné události.

Předpokládejme, že výše škody má logaritmicko-normální rozdělení s parametry μ a $\sigma^2 > 0$, tedy $\log Y \sim N(\mu, \sigma^2)$. Pak

$$\mathbb{P}(Y > c) = \mathbb{P}\left(\frac{\ln Y - \mu}{\sigma} > \frac{\ln c - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\ln c - \mu}{\sigma}\right).$$

Výsledná pravděpodobnost uplatnění je pak rovna

$$\left[1 - \Phi\left(\frac{\ln c - \mu}{\sigma}\right)\right]\left(1 - e^{-\lambda}\right).$$

Úplný model

Model můžeme rozšířit tak, aby zahrnoval různou individuální rizikovost chování jednotlivých pojištěnců, tedy každému pojištěnci bude odpovídat obecně různý parametr $\lambda \in \Lambda$ a individuální matice pravděpodobností přechodu $P(\lambda)$ na něm závisí. Obvykle se předpokládá, že rizikové parametry λ mají Γ-rozdělení popsané hustotou

$$f(\lambda) = \frac{\tau^h \lambda^{h-1}}{\Gamma(h)} \exp^{-\tau h}, \quad \lambda > 0,$$

kde h a τ jsou parametry. Ty se pro potřeby modelu odhadnou z dat o škodách pomocí momentové metody, protože střední hodnota Γ-rozdělení je rovna $\frac{h}{\tau}$ a rozptyl $\frac{h}{\tau^2}$.

Pojištěnce ještě rozvrstvíme podle doby, po kterou jsou pojištění. Označme příslušné relativní četnosti $w_1(\lambda), w_2(\lambda), \ldots, w_n(\lambda), \ldots$ Zde připouštíme, že doba pojištění a rizikovost pojištěnce se ovlivňují, například v tom smyslu, že vysoce rizikoví klienti budou typicky brzy odcházet k jiným pojišťovnám, které jim nabídnou vstřícnější podmínky. Dále označme symbolem A základní ryzí pojistné. Průměrné ryzí pojistné pojištěnce s rizikovým parametrem λ je potom

$$C(\lambda) = A \sum_{n=1}^{\infty} w_n(\lambda) \sum_{j \in S} p_j(n, \lambda) a_j,$$

kde $a = (a_1, \ldots a_k)$ je stupnice bonusu.

Ve vyváženém modelu by mělo platit

$$\mathbb{E}C(\lambda) = \int_0^{\infty} C(\lambda) f(\lambda) d\lambda = \mathbb{E}N\, \mathbb{E}Y,$$

kde $\mathbb{E}N$ je průměrný počet škod a $\mathbb{E}Y$ je průměrná výše škody. Tato rovnice slouží pro určení základního ryzího pojistného A. Řekněme, že rozdělení doby, po kterou jsou pojištěnci pojištěni je geometrické nezávislé na λ:

$$w_n(\lambda) = q^{n-1}(1-q), \quad n = 1, 2, \ldots.$$

Průměrná doba pojištění je pak $\frac{1}{1-q}$ a platí

$$C(\lambda) = A(1-q)p(1)^T \sum_{n=1}^{\infty} q^{n-1} P^{n-1}(\lambda) a = A(1-q)p(1)^T (I - qP(\lambda))^{-1} a,$$

kde I je jednotková matice, neboli

$$C(\lambda) = A(1-q) r_{k_0}(\lambda),$$

kde $r(\lambda)$ je řešením rovnice $(I - qP(\lambda))r(\lambda) = a$ a $r_{k_0}(\lambda)$ je řádek $r(\lambda)$ odpovídající počáteční bonusové třídě.

4.4 * Markovovy řetězce se stavy z množiny reálných čísel

Uvažujme následující lineární model

$$X_{n+1} = aX_n + \varepsilon_{n+1},$$

kde $\{\varepsilon_n\}_{n\in\mathbb{N}}$ je posloupnost nezávislých stejně rozdělených náhodných veličin s rozdělením $N(0,\sigma^2)$, tedy s hustotou

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}.$$

Přechodová hustota (ze stavu x do stavu y) je hustota normálního rozdělení s parametry ax a σ^2:

$$f(x,y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(y-ax)^2}{2\sigma^2}\right\}.$$

Přechodová hustota po k krocích pak odpovídá normálnímu rozdělení s parametry $a^k x$ a $\sum_{l=0}^{k-1} a^{2l}\sigma^2 = \sigma^2\frac{1-a^{2k}}{1-a^2}$, tedy:

$$f^{(k)}(x,y) = \frac{1}{\sigma\sqrt{2\pi\frac{1-a^{2k}}{1-a^2}}} \exp\left\{-\frac{(y-a^k x)^2}{2\sigma^2\frac{1-a^{2k}}{1-a^2}}\right\}.$$

Je-li $|a| < 1$, existuje limitní hustota f^∞ ve tvaru

$$f^\infty(y) = \lim_{k\to\infty} f^{(k)}(x,y) = \frac{\sqrt{1-a^2}}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(1-a^2)y^2}{2\sigma^2}\right\},$$

přičemž hodnota této funkce nezávisí na x. Tato hustota je hustotou normálního rozdělení $N(0,\sigma^2/(1-a^2))$.

Toto rozdělení je stacionární v následujícím smyslu: pokud hodnota řetězce X_0 v čase 0 má rozdělení $N(0,\sigma^2/(1-a^2))$, pak také hodnota řetězce v čase 1 má toto rozdělení, $X_1 \sim N(0,\sigma^2/(1-a^2))$. V tom případě pro každé $n \in \mathbb{N}$ platí $X_n \sim N(0,\sigma^2/(1-a^2))$.