Nenahlásené poistné udalosti a výpočet rezerv

Diplomová práca

Bc. Elena Mrózová

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
KATEDRA APLIKOVANEJ MATEMATIKY A ŠTATISTIKY

Pravdepodobnosť a matematická štatistika

Školiteľ záverečnej práce
doc. RNDr. Rastislav Potocký, CSc.

BRATISLAVA

2009

Čestné vyhlásenie:

Čestne vyhlasujem, že som túto diplomovú prácu vypracovala samostatne, len s pomocou nadobudnutých teoretických vedomostí a použitej literatúry.

………………………………

Elena Mrózová

Abstrakt

Názov: Nenahlásené poistné udalosti a odhad rezerv.

Autor: Bc. Elena Mrózová

Škola: Univerzita Komenského

Fakulta matematiky, fyziky a informatiky

Katedra aplikovanej matematiky a štatistiky

Školiteľ: doc. RNDr. Rastislav Potocký, CSc.

Miesto: Bratislava

Rok: 2009

Táto diplomová práca najprv opisuje metódu na odhad počtu nenahlásených poistných udalostí, ktorú navrhol D. Scollnik. Neskôr sa venuje metódam na výpočet IBNR rezerv, špeciálne Mackovej metóde chain ladder. Tu je uvedený aj vzorec na výpočet strednej kvadratickej chyby odhadu IBNR rezerv. Metódy na odhad počtu nenahlásených poistných udalostí ako aj na výpočet IBNR rezerv sa v poslednej kapitole aplikujú na konkrétne dáta.

Kľúčové slová: nenahlásené poistné udalosti, IBNR, chain ladder, separačná metóda.

Firstly, this diploma work present method to estimate number of unreported claims, which has been proposed by D. Scollnik. The next part deals with methods to estimate IBNR reserves, specially Mack`s chain ladder model. We introduce formula for mean square error of estimate IBNR reserves, too. In the last chapter are methods to estimate number of unreported claims and to estimate IBNR reserves are applied on specifically data.

Key words: unreported claims, IBNR, chain ladder, separation method

Predhovor

Odhadovanie IBNR rezerv a v tejto dobe už aj odhadovanie počtu nenahlásených poistných udalostí je dnes už bežnou náplňou práce poistného matematika. Ten si môže vyberať zo širokej škály metód výpočtu IBNR rezerv. Medzi najpoužívanejšie však patrí metóda chain ladder, častokrát sa ale využíva viacero metód pre porovnanie výsledkov.

Ľudí z praxe zaujíma hlavne konečný výsledok – odhad IBNR rezerv. Prácu sa preto snažíme orientovať viac aplikačne, teda nás tiež zaujíma hlavne odhad IBNR rezerv. No pri metóde chain ladder uvádzame aj vzorec na výpočet strednej kvadratickej chyby tohto odhadu.

Najskôr sa ale venujeme novinke vo výkazníctve v poisťovníctve – odhadu počtu nenahlásených poistných udalostí. Na tento problém sa v minulosti v prípade potreby aplikovala metóda chain ladder. My v práci predstavujeme metódu navrhnutú D. Scollnikom, ktorá je zameraná špeciálne na tento problém.

Cieľom práce je aplikovať teoreticky opísané metódy na reálne dáta. Pre uľahčenie si metódy naprogramujeme v programovacom jazyku R, ktorý je voľne stiahnuteľný z internetu na http://cran.r-project.org/.

Touto cestou chcem poďakovať doc. RNDr. Rastislavovi Potockému CSc. za poskytnutie odborných článkov, za odpovede na moje otázky a za cenné rady.

Obsah

Úvod

Ešte pred pár desiatkami rokov sa plnenia na poistných udalostiach v neživotnom poistení vyplácali systémom *pay-as-you-go*, teda plnenia vyplatené v danom roku sa financovali z prijatého poistného v tom istom roku, bez ohľadu na rok vzniku poistnej udalosti. V niektorých poistných odvetviach je však častým javom, že sa poistné udalosti nevyplácajú v tom istom roku ako vznikli. Príčiny môžu byť rôzne, či už je to neskoré nahlásenie poistnej udalosti, alebo zdĺhavá likvidácia zapríčinená súdnymi konaniami. Napríklad pri povinnom zmluvnom poistení zodpovednosti za škodu spôsobenú prevádzkou motorového vozidla sa škoda na zdraví môže objaviť aj niekoľko rokov po dopravnej nehode.

Keďže poistné na nejaké obdobie, v neživotnom poistení najčastejšie jeden rok, slúži na krytie rizika, ktoré môže vzniknúť v tomto období, je logické, že poisťovne tvoria rezervy. V neživotnom poistení rozlišujeme dva základné druhy rezerv: *RBNS rezervy* a *IBNR rezervy*. *RBNS* (z angl. Reported But Not Settled) rezervy sú rezervy na ohlásené poistné udalosti, ktoré ešte neboli vyplatené, resp. neboli vyplatené úplne. *IBNR* (z angl. Incurred But Not Reported) rezervy sú rezervy na poistné udalosti, ktoré sa stali, ale neboli ohlásené poisťovni, teda sú to poistné udalosti, o ktorých poisťovňa nevie. Tieto rezervy v sebe zahŕňajú aj rezervy na nedostatočne zarezervované poistné udalosti.

V posledných rokoch sa zvýšil záujem aj o počet neohlásených poistných udalostí.

Prvá kapitola obsahuje prehľad rozdelení, ktoré sa v tejto problematike často využívajú. V druhej kapitole sa zaoberáme odhadom počtu nenahlásených poistných udalostí. V tretej kapitole sa venujeme odhadu IBNR rezerv. Zaoberáme sa najmä najbežnejšou metódou chain ladder. Štvrtá kapitola je venovaná príkladom. Na jednu množinu dát aplikujeme v práci uvedené metódy na odhad IBNR rezerv a uvedieme aj príklad na odhad počtu nenahlásených poistných udalostí.

1 Rozdelenia

1.1 Binomické rozdelenie

Opakujme *nezávisle n-krát* Bernoulliho pokus, ktorý môže mať len dva možné výsledky :

- nastane udalosť A a to s pravdepodobnosťou p,

- nenastane udalosť A a to s pravdepodobnosťou $1-p$.

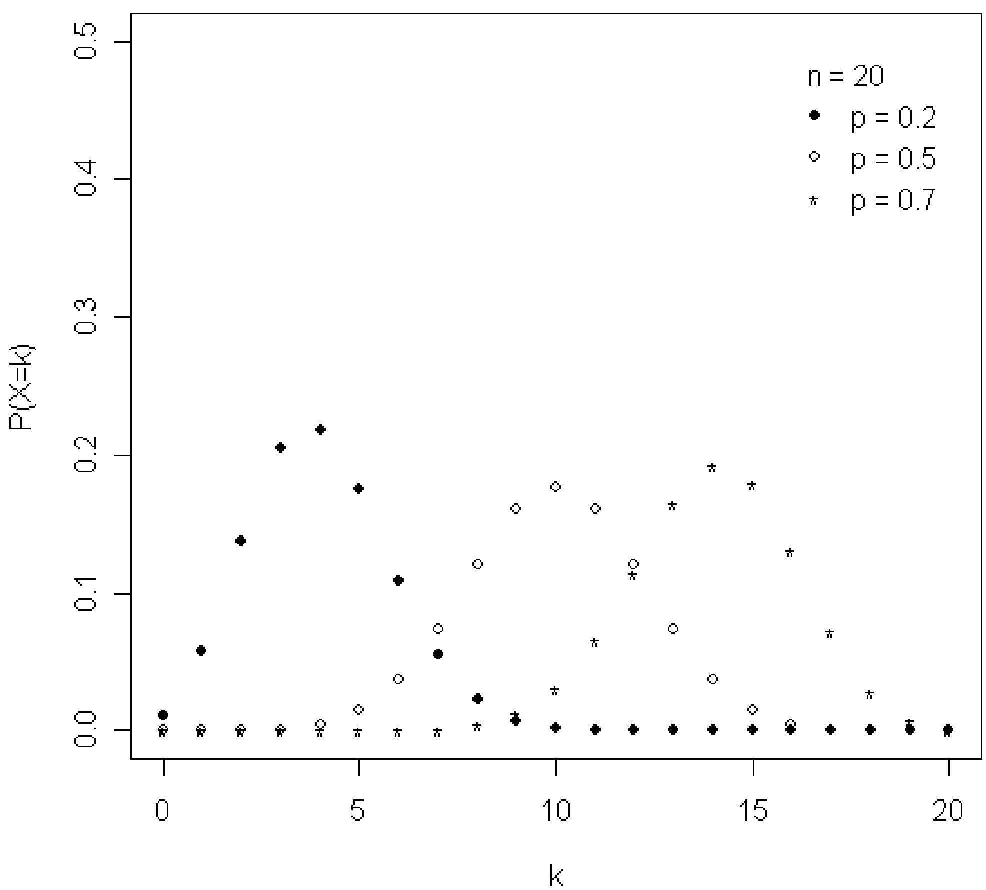
Náhodnou premennou X je počet výskytov udalosti A v n nezávislých pokusoch s rovnakou pravdepodobnosťou p jej výskytu v každom pokuse. ([5])

Náhodná premenná X, ktorá môže nadobúdať hodnoty $0,1,\ldots,n$, kde n je prirodzené číslo, má *binomické rozdelenie s parametrami n a p*; $p \in (0,1)$; ak platí:

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}; k = 0,1,\ldots,n.$$

Označenie $X \sim Bi(n,p)$.

Obr. 1: *Binomické rozdelenie*



Základné charakteristiky:

- stredná hodnota $E[X] = np$,

- rozptyl $D[X] = np(1-p)$.

Špeciálnym prípadom binomického rozdelenia je *alternatívne rozdelenie*, ktoré popisuje Bernoulliho pokus. V tomto prípade $n = 1$.

1.2 Poissonovo rozdelenie

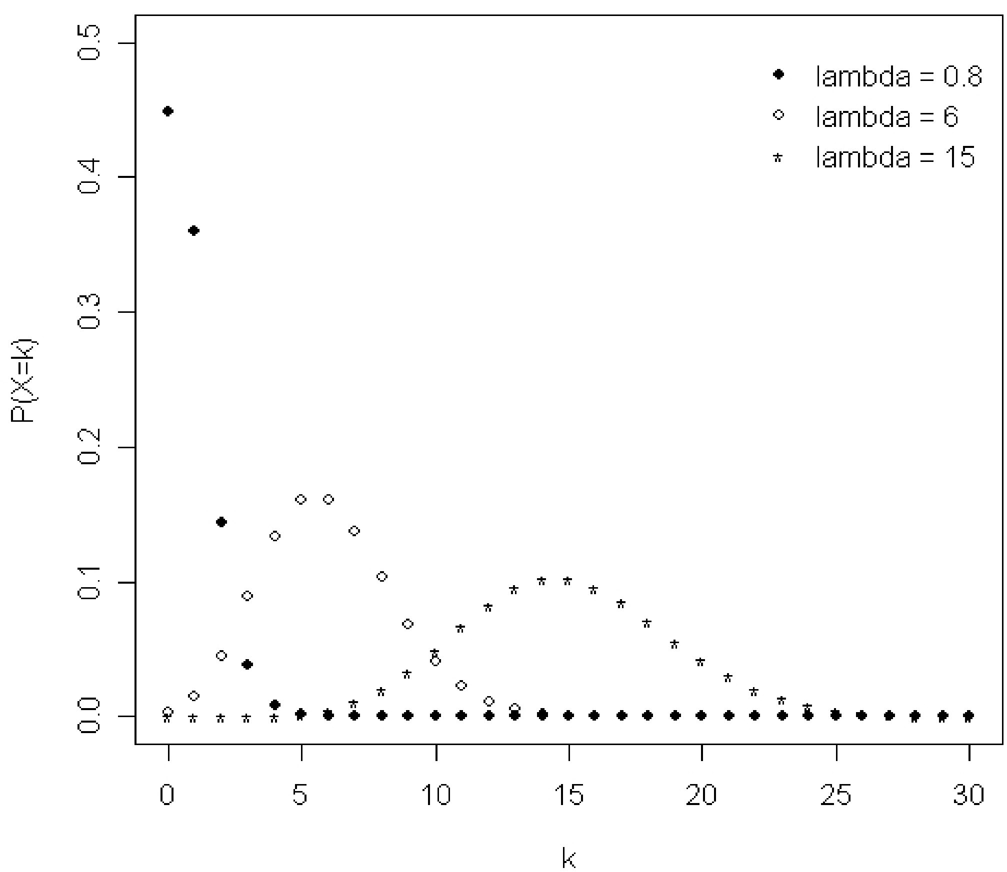
Poissonovo rozdelenie vzniká ako limitný prípad binomického rozdelenia $Bi(n, p)$, ak súčasne platí $n \to \infty$ a $p \to 0$, pričom stredná hodnota np je konštantná. ([5])

Náhodná premenná X, ktorá môže nadobúdať hodnoty $0,1,\ldots$, má *Poissonovo rozdelenie s parametrom* λ, kde $\lambda > 0$ je dané číslo, ak platí:

$$\Pr(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}; k = 0,1,\ldots.$$

Označenie $X \sim Po(\lambda)$.

Obr. 2: *Poissonovo rozdelenie*



Je to rozdelenie výskytu zriedkavých udalostí pri vysokom počte nezávislých pokusoch. ([5])

<u>*Základné charakteristiky:*</u>

- stredná hodnota $E[X] = \lambda$,
- rozptyl $D[X] = \lambda$.

1.3 Negatívne binomické rozdelenie

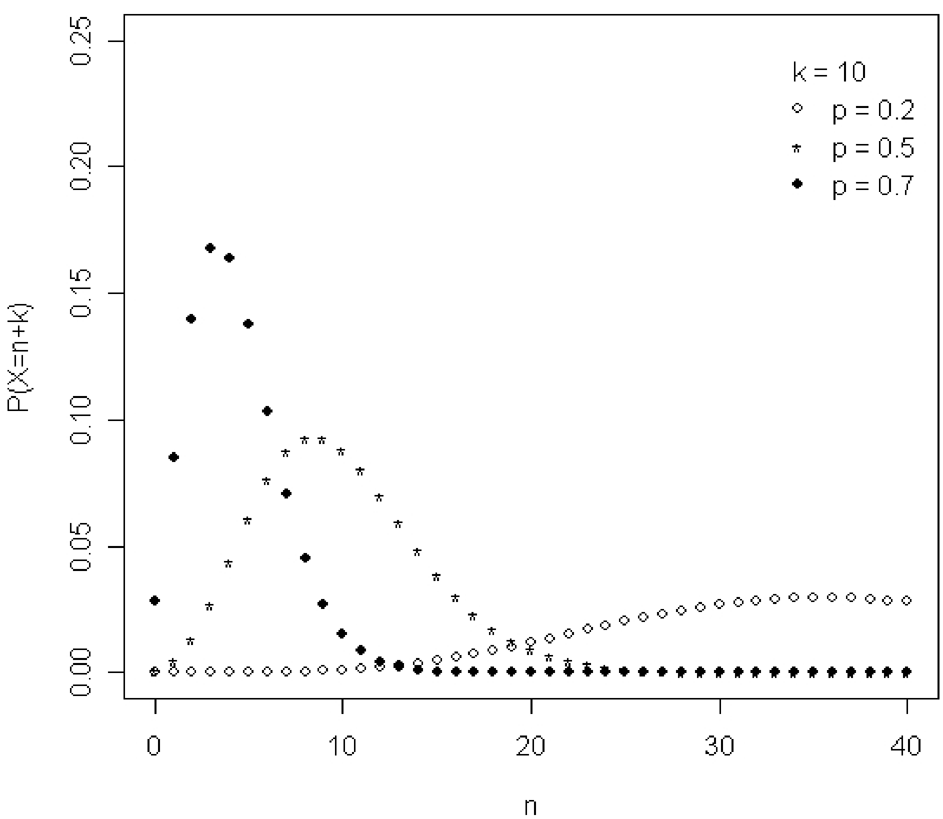
V tomto rozdelení náhodná premenná reprezentuje počet náhodných pokusov, pri ktorých nastala udalosť A práve k-krát a potom už pokusy nepokračovali.

Náhodná premenná X, ktorá nadobúda hodnoty $k, k+1, \dots$, má *negatívne binomické rozdelenie s parametrami k a p;* $p \in (0,1)$ a k je celé kladné číslo; ak platí:

$$\Pr(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}; x = k, k+1, \dots .$$

Označenie $X \sim NB(k, p)$. ([5])

Obr. 3: *Negatívne binomické rozdelenie*



Základné charakteristiky:

- stredná hodnota $E[X] = \frac{k(1-p)}{p}$, ([5])

- rozptyl $D[X] = \frac{k(1-p)}{p^2}$. ([5])

1.4 Kvázi - binomické rozdelenie

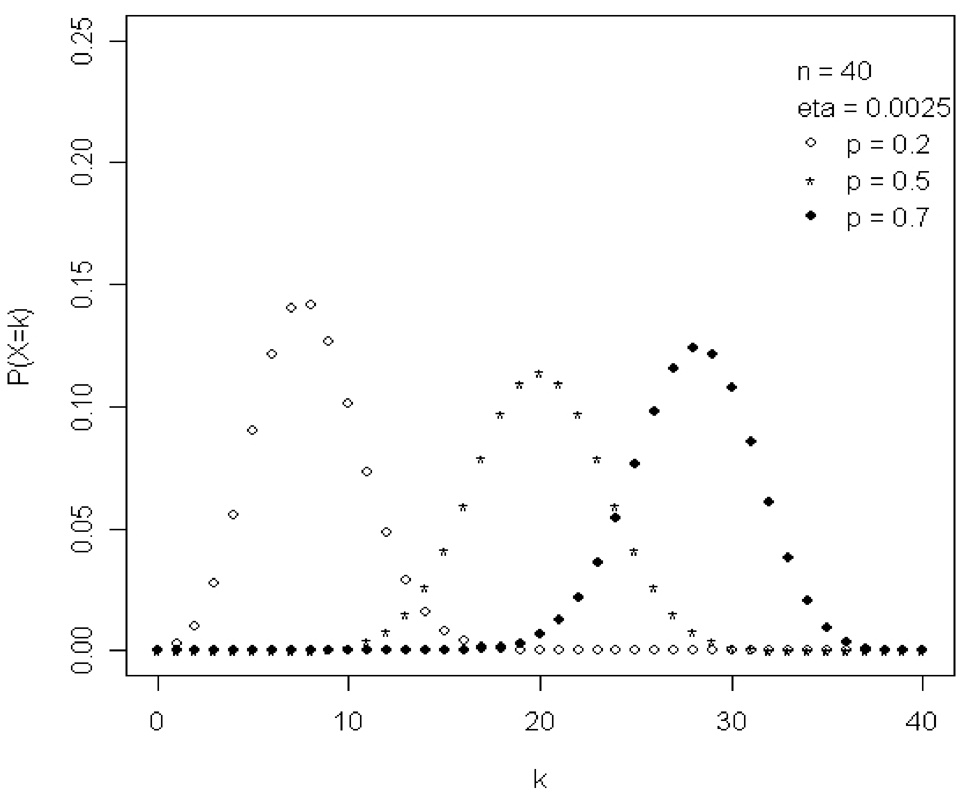
Náhodná premenná X, ktorá nadobúda hodnoty $0,1,\ldots,n$, má *kvázi – binomické rozdelenie II. typu s parametrami n, p a ξ;* $p \in (0,1)$, *n* je celé kladné číslo a $-\frac{p}{n} < \xi < \frac{1-p}{n}$; ak platí:

$$\Pr(X = k) = \binom{n}{k} \frac{p(1-p)}{1+n\xi} \left(\frac{p+k\xi}{1+n\xi}\right)^{k-1} \left(\frac{(1-p)+(n-x)\xi}{1+n\xi}\right)^{n-k-1}; k = 0,1,\ldots,n.$$

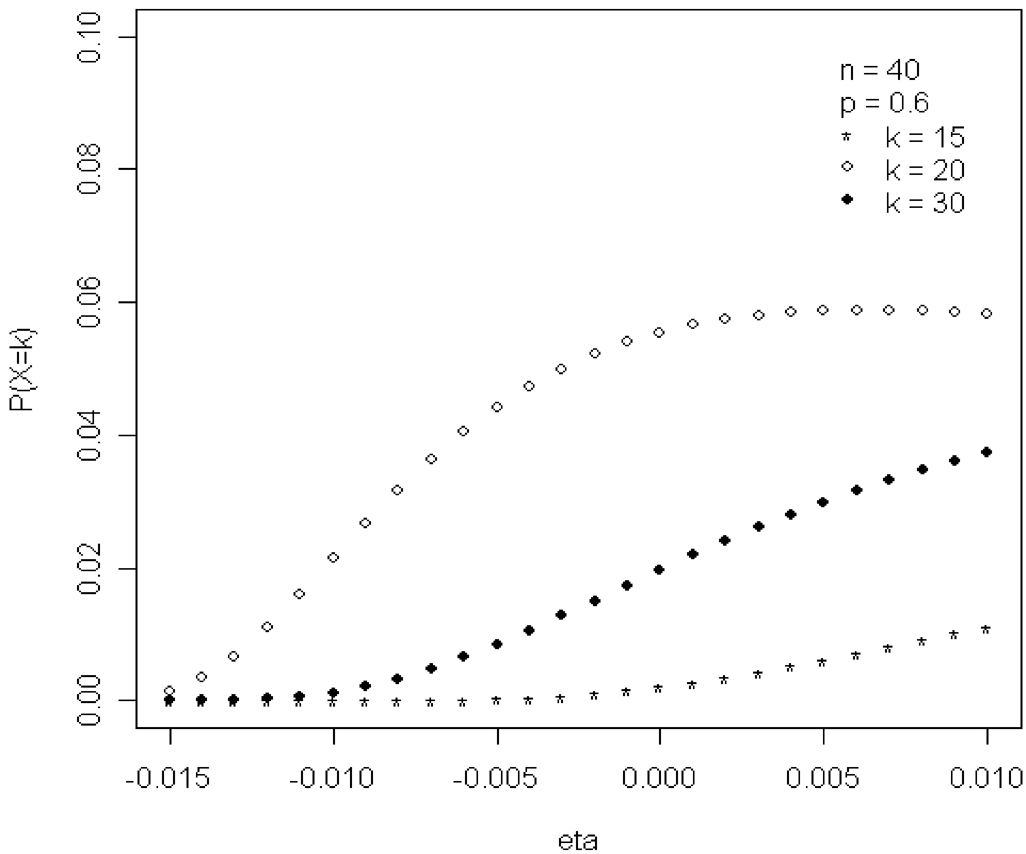
Označenie $X \sim QBD(n, p, \xi)$. ([4] , [6])

Parametre n a p označujú, ako pri binomickom rozdelení, počet pokusov a počiatočnú pravdepodobnosť úspechu v každom pokuse. Parameter ξ hovorí o tom, ako pravdepodobnosť úspechu p klesá alebo rastie s počtom úspechov ([4]).

Obr. 4: *Kvázi – binomické rozdelenie II.typu*



Obr. 5: *Kvázi – binomické rozdelenie: závislosť na parametri ξ*



Základné charakteristiky:

- stredná hodnota $E[X] = np$,
- jednoduché vyjadrenie pre rozptyl neexistuje, je ho však možné, pri konečnom rozsahu náhodného výberu, vypočítať. ([6])

1.5 Zovšeobecnené Poissonovo rozdelenie

Pre náhodnú premennú s Poissonovým rozdelením platí, že sa teoretická stredná hodnota rovná teoretickej disperzii. Preto toto rozdelenie nie je vhodné, ak výberová variancia pozorovaní výrazne presahuje výberový priemer, čo je v množinách reálnych dát bežné.

Zovšeobecnené Poissonovo rozdelenie je dvoj-parametrické diskrétne rozdelenie s pravdepodobnostnou funkciou

$$\Pr(X = k) = \theta(\theta + k\lambda)^{k-1}\frac{e^{-\theta-k\lambda}}{k!}; k = 0,1,2,\ldots,$$

s parametrami $\theta > 0$ a $0 \le \lambda < 1$. Označenie $X \sim GPD(\theta,\lambda)$. ([6])

Základné charakteristiky:

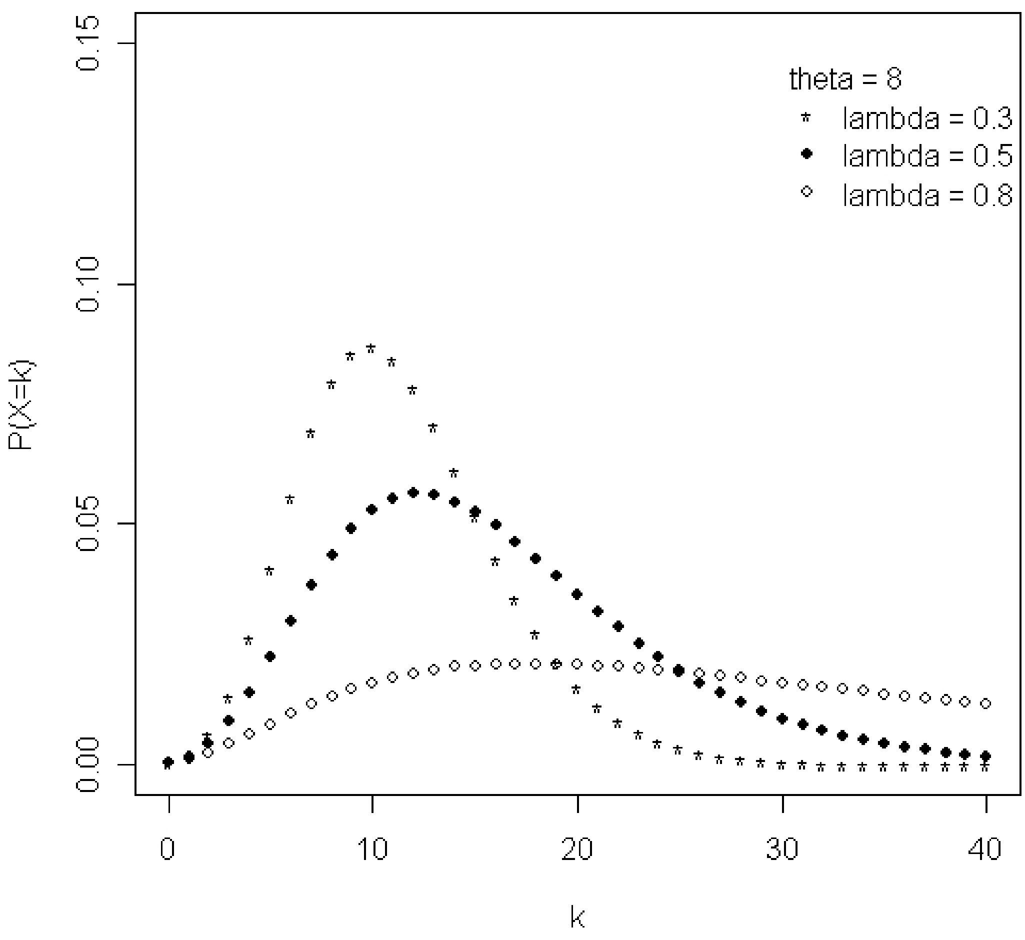
- stredná hodnota $E[X] = \frac{\theta}{1-\lambda}$, ([6])

- disperzia $D[X] = \frac{\theta}{(1-\lambda)^3}$ ([6])

a platí: $D[X] \geq E[X]$.

Ak $\lambda = 0$, potom dostávame štandardné Poissonovo rozdelenie.

Obr. 6: *Zovšeobecnené Poissonovo rozdelenie*



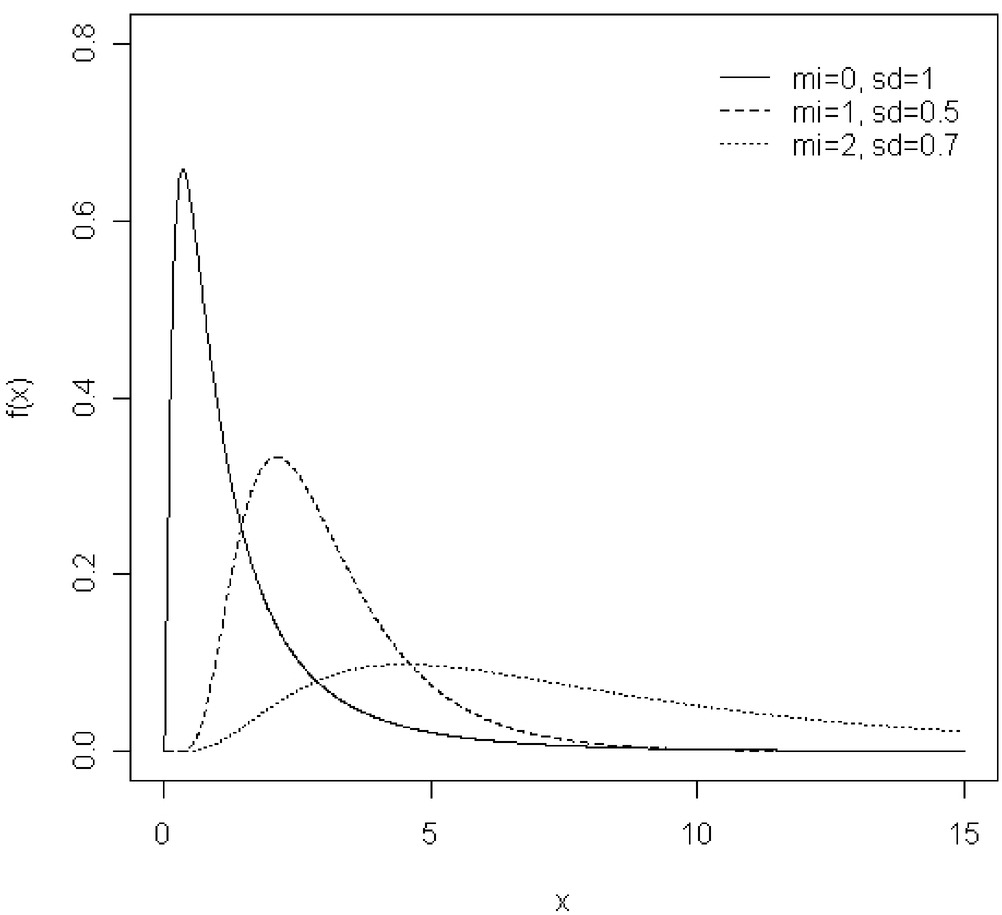
1.6 Logaritmicko – normálne rozdelenie

Náhodná premenná X má *logaritmicko normálne rozdelenie s parametrami* μ *a* σ^2; $\mu \in \mathbf{R}$ a $\sigma^2 > 0$; ak pre jej hustotu pravdepodobnosti platí:

$$f(x) = \begin{cases} \dfrac{1}{\sigma x\sqrt{2\pi}} \exp\left(-\dfrac{(\ln x - \mu)^2}{2\sigma^2}\right) & x > 0 \\ 0 & x \le 0 \end{cases}$$

Označenie $X \sim NL(\mu, \sigma^2)$. ([5])

Obr. 7: *Logaritmicko – normálne rozdelenie*



<u>*Základné charakteristiky:*</u>

- stredná hodnota $E[X] = \exp\left(\mu + \dfrac{\sigma^2}{2}\right)$, ([5])

- rozptyl $D[X] = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$. ([5])

Ak náhodná premenná $Y = \ln X$ a $X \sim NL(\mu, \sigma^2)$, potom sa dá dokázať, že $Y \sim N(\mu, \sigma^2)$.

2 Odhad počtu nenahlásených poistných udalostí

V ostatných rokoch sa zvýšil záujem poisťovní ako aj príslušných kontrolných úradov o počet nenahlásených poistných udalostí. Jednou z možností, ako tento počet odhadnúť, je použiť metódu chain ladder, ktorá ale prvotne slúži na odhad IBNR rezerv. Iný spôsob na odhad počtu nenahlásených poistných udalostí nám ponúka David P. M. Scollnik vo svojom článku *A dameged generalised Poisson model and its application to reported and unreported accident counts* [6], ktorý táto kapitola cituje.

Majme nezápornú náhodnú premennú X, nadobudajúcu celočíselné hodnoty, s funkciou pravdepodobnosti $\Pr(X = x)$. Táto náhodná premenná je rozložiteľná, ak $X = Y + Z$, kde Z je celočíselná nezáporná nepozorovateľná náhodná premenná a Y je pozorovateľná časť X. Pri danom $X = x$, podmienené rozdelenie $\Pr(Y = y \mid X = x)$ je známe ako rozdelenie prežívania. O tom, za akých okolností rozdelenie prežívania charakterizuje rozdelenie náhodnej premennej X a naopak, hovorí nasledujúce tvrdenie.

<u>*Tvrdenie1:*</u>

Ak rozdelenie prežívania $\Pr(Y = y \mid X = x)$ je binomické $Bi(x, \varphi)$, potom Y je náhodná premenná taká, že

$$\Pr(Y = y) = \Pr(Y = y \mid X = Y) = \Pr(Y = y \mid X > Y) \tag{1}$$

vtedy a len vtedy, ak X má Poissonovo rozdelenie. ([6])

Naviac platí, že ak X má Poissonovo rozdelenie s parametrom θ, potom Y má Poissonovo rozdelenie s parametrom $\varphi\theta$ a Z má Poissonovo rozdelenie s parametrom $(1-\varphi)\theta$, a tieto premenné sú navzájom nezávislé. Dokázalo sa, že ak X má Poissonovo rozdelenie a platí (1), potom rozdelenie prežívania musí byť binomické ([6]).

2.1 Poissonov model

Nech X je rozložiteľná náhodná premenná s Poissonovým rozdelením s parametrom θ a s binomickým rozdelením prežívania, $X = Y + Z$. Predpokladajme, že kedykoľvek robíme pozorovanie premennej Y na náhodnej premennej X, je možné určiť, či boli nejaké nehody neohlásené, avšak nevieme určiť ich množstvo. Zaveďme si preto indikátorovú náhodnú premennú D, ktorá nadobúda hodnotu 1, ak neohlásená nehoda bola pozorovaná, a hodnotu 0 ak neohlásená nehoda pozorovaná nebola. Udalosť $[D=0]$ je ekvivalentná s udalosťou $[Z=0]$ a udalosť $[D=1]$ je ekvivalentná s udalosťou $[Z>0]$, a teda aj Y a D sú nezávislé náhodné premenné a pre ich pravdepodobnostné funkcie platí:

$$\Pr(Y=y)=e^{-\varphi\theta}\frac{(\varphi\theta)^y}{y!} \text{ pre } y=0,1,2,\ldots,$$

$$\Pr(D=0)=\Pr(Z=0)=e^{-(1-\varphi)\theta},$$

$$\Pr(D=1)=\Pr(Z>0)=1-e^{-(1-\varphi)\theta}.$$

Z nezávislosti Y a D vyplýva

$$\Pr(Y=y,D=d)=e^{-\varphi\theta}\frac{(\varphi\theta)^y}{y!}\left(1-e^{-(1-\varphi)\theta}\right)^d\left(e^{-(1-\varphi)\theta}\right)^{1-d},$$

pre $y=0,1,2,\ldots,$ a $d=0,1$.

Majme náhodný výber $(y_i,d_i);\ i=1,2,\ldots,n$, potom prirodzený logaritmus funkcie vierohodnosti má tvar

$$l=\ln L=-n\varphi\theta+n\bar{y}\ln(\varphi\theta)+n\bar{d}\ln\left(1-e^{-(1-\varphi)\theta}\right)-n\left(1-\bar{d}\right)(1-\varphi)\theta-\sum_{i=1}^{n}\ln(y_i!),$$

kde $\bar{y}$ je výberový priemer hodnôt y_i, $\bar{d}$ je výberový priemer hodnôt d_i; $i=1,2,\ldots,n$. Riešením normálnych rovníc

$$\left.\frac{\partial l}{\partial\theta}\right|_{\varphi=\hat{\varphi},\theta=\hat{\theta}}=-n\hat{\varphi}+\frac{n\bar{y}}{\hat{\theta}}+\frac{n\bar{d}(1-\hat{\varphi})}{e^{(1-\hat{\varphi})\hat{\theta}}-1}-n\left(1-\bar{d}\right)(1-\hat{\varphi})=0$$

$$\left.\frac{\partial l}{\partial\varphi}\right|_{\varphi=\hat{\varphi},\theta=\hat{\theta}}=-n\hat{\theta}+\frac{n\bar{y}}{\hat{\varphi}}-\frac{n\bar{d}\hat{\theta}}{e^{(1-\hat{\varphi})\hat{\theta}}-1}+n\left(1-\bar{d}\right)\hat{\theta}=0$$

dostávame maximálne vierohodné odhady pre parametre θ a φ:

$$\hat{\theta}=\bar{y}-\ln\left(1-\bar{d}\right), \tag{2}$$

$$\hat{\varphi}=\frac{\bar{y}}{\hat{\theta}}=\frac{\bar{y}}{\bar{y}-\ln\left(1-\bar{d}\right)}. \tag{3}$$

Toto je prípad, keď pri každom pozorovaní máme rovnakú mieru rizika. Ďalej sa budeme zaoberať prípadom, keď nemáme rovnakú mieru rizika. Zavedieme alternatívny model, založený na zovšeobecnenom Poissonovom rozdelení náhodnej premennej X a kvázi – binomickom rozdelení prežívania $\Pr(Y=y\mid X=x)$.

2.2 Prípad nerovnej miery rizika

Nech celkový počet nehôd X má Poissonovo rozdelenie s parametrom $w\theta$, kde parameter θ predstavuje stredné množstvo nehôd na jednotku rizika a w predstavuje jednotky miery rizika spojené s X. Nech $X = Y + Z$, kde Y označuje počet nahlásených nehôd a Z počet nenahlásených nehôd. Predpokladajme, že nehody sú nezávislé s rovnakou pravdepodobnosťou nahlásenia φ. A teda rozdelenie prežívania $\Pr(Y = y \mid X = x)$ je binomické $Bi(x, \varphi)$. Z tvrdení uvedených na začiatku kapitoly vyplýva, že Y má Poissonovo rozdelenie s parametrom $w\theta\varphi$, Z má Poissonovo rozdelenie s parametrom $w\theta(1-\varphi)$ a tieto náhodné premenné sú nezávislé.

Predpokladajme, že kedykoľvek robíme pozorovanie Y na X, vieme povedať, či nejaké nehody boli neohlásené, ale ich počet nepoznáme. Nech indikátorová náhodná premenná D nadobúda hodnotu 1, ak neohlásená nehoda bola pozorovaná, a nadobúda hodnotu 0 ak takáto nehoda pozorovaná nebola. Udalosť $[D = 0]$ je ekvivalentná s udalosťou $[Z = 0]$ a udalosť $[D = 1]$ je ekvivalentná s udalosťou $[Z > 0]$. Náhodné premenné Y a D sú nezávislé. Teda

$$\Pr(Y = y, D = d) = e^{-w\varphi\theta} \frac{(w\varphi\theta)^y}{y!} \left(1 - e^{-w(1-\varphi)\theta}\right)^d \left(e^{-w(1-\varphi)\theta}\right)^{1-d}, \tag{4}$$

pre $y = 0, 1, 2, \ldots,$ a $d = 0, 1$.

Na základe náhodného výberu $(y_i, d_i);\, i = 1, 2, \ldots, n$, ktorý je spojený s vektorom jednotiek miery rizika $w = (w_1, \ldots, w_n)^T$, môžeme vyjadriť logaritmus funkcie vierohodnosti nasledovne:

$$l = -n\varphi\theta\overline{w} + n\overline{y}\ln(\varphi\theta) + \sum_{i=1}^{n} d_i \ln\left(1 - e^{-w_i(1-\varphi)\theta}\right) - (1-\varphi)\theta\sum_{i=1}^{n}(1 - d_i)w_i + \sum_{i=1}^{n} y_i \ln(w_i) - \sum_{i=1}^{n}\ln(y_i!),$$

kde $\overline{y}$ je výberový priemer hodnôt y_i, pre $i = 1, 2, \ldots, n$, a $\overline{w} = \frac{1}{n}\sum_{i=1}^{n} w_i$ je priemerná miera rizika. Normálne rovnice

$$\left.\frac{\partial l}{\partial \theta}\right|_{\varphi=\hat{\varphi},\theta=\hat{\theta}} = \frac{n\overline{y}}{\hat{\theta}} + (1-\hat{\varphi})\sum_{i=1}^{n}\frac{d_i w_i}{1 - e^{-w_i(1-\hat{\varphi})\hat{\theta}}} - n\overline{w} = 0$$

$$\left.\frac{\partial l}{\partial \varphi}\right|_{\varphi=\hat{\varphi},\theta=\hat{\theta}} = \frac{n\overline{y}}{\hat{\varphi}} - \hat{\theta}\sum_{i=1}^{n}\frac{d_i w_i}{1 - e^{-w_i(1-\hat{\varphi})\hat{\theta}}} = 0$$

možno zredukovať na rovnice

$$\hat{\varphi} = \frac{\bar{y}}{\hat{\theta}\bar{w}}$$

$$\sum_{i=1}^{n}(1-d_i)w_i = \sum_{i=1}^{n}\frac{d_i w_i}{e^{w_i(1-\hat{\varphi})\hat{\theta}}-1},$$

ktorých riešením sú maximálne vierohodné odhady pre parametre θ a φ. V špeciálnom prípade, ak $w_i = 1$ pre každé $i = 1,2,\ldots,n$, sa riešenie týchto rovníc zredukuje na riešenie (2) a (3).

2.3 Zovšeobecnený Poissonov model

Nech X je celočíselná nezáporná náhodná premenná taká, že $X = Y + Z$, kde Z je celočíselná nezáporná nepozorovateľná náhodná premenná a Y je pozorovateľná časť X. Tvrdenie 1 môžeme prepísať nasledovne:

<u>*Tvrdenie 2:*</u>

Nech rozdelenie prežívania $\Pr(Y = y \mid X = x)$ je kvázi – binomické s pravdepodobnostnou funkciou

$$\Pr(Y = y \mid X = x) = \binom{x}{y}\frac{\varphi(1-\varphi)}{1+x\xi}\left(\frac{\varphi + y\xi}{1+x\xi}\right)^{y-1}\left(\frac{(1-\varphi)+(x-y)\xi}{1+x\xi}\right)^{x-y-1}; y = 0,1,\ldots,x, \tag{5}$$

s parametrami $\varphi \in (0,1)$ a $-\frac{\varphi}{x} < \xi < \frac{1-\varphi}{x}$ pri danom $X = x$. Potom náhodná premenná Y je taká, že

$$\Pr(Y = y) = \Pr(Y = y \mid X = Y) = \Pr(Y = y \mid X > Y) \tag{1a}$$

vtedy a len vtedy, ak X má zovšeobecnené Poissonovo rozdelenie s parametrami θ a $\lambda = \xi\theta$, teda $\Pr(X = x) = \theta(\theta + x\xi\theta)^{x-1}\frac{e^{-\theta - x\xi\theta}}{x!}; x = 0,1,2,\ldots$, kde $\theta > 0$ a $0 \le \xi\theta < 1$ ([6]).

Naviac, ak X má zovšeobecnené Poissonovo rozdelenie s parametrami θ a $\xi\theta$, potom Y má zovšeobecnené Poissonovo rozdelenie s parametrami $\varphi\theta$ a $\xi\theta$, Z má zovšeobecnené Poissonovo rozdelenie s parametrami $(1-\varphi)\theta$ a $\xi\theta$, a tieto náhodné premenné sú nezávislé ([6]). Ak X má zovšeobecnené Poissonovo rozdelenie s parametrami θ a $\xi\theta$ a platí (1a), potom rozdelenie prežívania musí byť kvázi – binomické ([6]).

Binomické rozdelenie prežívania pre množstvo nahlásených nehôd predpokladá, že nehody sú nezávislé, s rovnakou pravdepodobnosťou nahlásenia. Tento predpoklad často neplatí, pretože existujú mnohé poistné schémy, ktoré zvýhodňujú klientov za bezškodový priebeh trvania ich poistnej zmluvy. Preto je vhodnejšie uvažovať kvázi – binomické rozdelenie prežívania, pre ktoré platí, že sa pravdepodobnosť úspechu úmerne mení s počtom úspechov.

Predpokladajme, že celkové množstvo nehôd v sledovanom období, ktoré je spojené s jednou jednotkou miery rizika, má zovšeobecnené Poissonovo rozdelenie sa parametrami $\theta > 0$ a $0 \le \lambda < 1$. Ak X je počet nehôd v období, ktoré je spojené s w jednotkami miery rizika, potom z konvolučnej vlastnosti zovšeobecneného Poissonovho rozdelenia, X má zovšeobecnené Poissonovo rozdelenie s parametrami $w\theta$ a λ ([6]). Nech $X = Y + Z$, kde Y označuje počet ohlásených a Z počet neohlásených nehôd. Z predpokladu o rozdelení náhodnej premennej X vyplýva, že Y má zovšeobecnené Poissonovo rozdelenie s parametrami $w\varphi\theta$ a λ, Z má zovšeobecnené Poissonovo rozdelenie s parametrami $w(1-\varphi)\theta$ a λ, a tieto náhodné premenné sú nezávislé. Rozdelenie prežívania $\Pr(Y = y \mid X = x)$ je kvázi – binomické s parametrami φ a ξ a s pravdepodobnostnou funkciou danou vzťahom (5).

Predpokladajme, že kedykoľvek robíme pozorovanie Y na X, vieme, či nejaké nehody boli neohlásené, ale nepoznáme ich počet. Nech indikátorová náhodná premenná D nadobúda hodnotu 1, ak neohlásená nehoda bola pozorovaná, a nadobúda hodnotu 0 neohlásená nehoda nebola pozorovaná. Udalosť $[D = 0]$ je ekvivalentná s udalosťou $[Z = 0]$ a udalosť $[D = 1]$ je ekvivalentná s udalosťou $[Z > 0]$. Náhodné premenné Y a D sú nezávisle. Preto

$$\Pr(Y = y, D = d) = w\varphi\theta(w\varphi\theta + y\lambda)^{y-1} \frac{e^{-w\varphi\theta - y\lambda}}{y!} \left(1 - e^{-w(1-\varphi)\theta}\right)^d \left(e^{-w(1-\varphi)\theta}\right)^{1-d},$$

pre $y = 0,1,2,\ldots$ a $d = 0,1$.

Na základe náhodného výberu $(y_i, d_i);\, i = 1,2,\ldots,n$, spojeného s vektorom jednotiek miery rizika $w = (w_1,\ldots,w_n)^T$, prirodzený logaritmus vierohodnostnej funkcie je daný vzťahom:

$$\begin{aligned} l = -n\varphi\theta\overline{w} - \lambda n\overline{y} + n\ln(\varphi\theta) + \sum_{i=1}^{n}(y_i - 1)\ln(w_i\varphi\theta + y_i\lambda) + \sum_{i=1}^{n} d_i \ln\left(1 - e^{-w_i(1-\varphi)\theta}\right) - \\ -(1-\varphi)\theta\sum_{i=1}^{n}(1 - d_i)w_i + \sum_{i=1}^{n}\ln(w_i) - \sum_{i=1}^{n}\ln(y_i!) \end{aligned},$$

kde $\bar{y}$ je výberový priemer hodnôt y_i, pre $i = 1,2,\ldots,n$, a $\bar{w} = \frac{1}{n}\sum_{i=1}^{n} w_i$ je priemerná miera rizika. Normálne rovnice

$$\left.\frac{\partial l}{\partial \theta}\right|_{\varphi=\hat{\varphi},\theta=\hat{\theta},\lambda=\hat{\lambda}} = \frac{n}{\hat{\theta}} + \sum_{i=1}^{n}(y_i - 1)\frac{w_i\hat{\varphi}}{w_i\hat{\varphi}\hat{\theta} + y_i\hat{\lambda}} + (1-\hat{\varphi})\sum_{i=1}^{n}\frac{d_i w_i}{1-e^{-w_i(1-\hat{\varphi})\hat{\theta}}} - n\bar{w} = 0$$

$$\left.\frac{\partial l}{\partial \varphi}\right|_{\varphi=\hat{\varphi},\theta=\hat{\theta},\lambda=\hat{\lambda}} = \frac{n}{\hat{\varphi}} + \sum_{i=1}^{n}(y_i - 1)\frac{w_i\hat{\theta}}{w_i\hat{\varphi}\hat{\theta} + y_i\hat{\lambda}} + \hat{\theta}\sum_{i=1}^{n}\frac{d_i w_i}{1-e^{-w_i(1-\hat{\varphi})\hat{\theta}}} = 0$$

$$\left.\frac{\partial l}{\partial \lambda}\right|_{\varphi=\hat{\varphi},\theta=\hat{\theta},\lambda=\hat{\lambda}} = -n\bar{y} + \sum_{i=1}^{n}(y_i - 1)\frac{y_i}{w_i\hat{\varphi}\hat{\theta} + y_i\hat{\lambda}} = 0$$

nie je ľahké zjednodušiť a nemožno ich explicitne vyrátať. Odhady parametrov možno získať pomocou štandardných štatistických softvérov.

3 Metódy na výpočet odhadu IBNR rezerv

Neživotné poisťovne majú zo zákona povinnosť vytvárať okrem rezerv na nahlásené poistné udalosti aj rezervy na nenahlásené poistné udalosti. Je podstatné, aby tieto rezervy boli dostatočné vysoké, teda aby poisťovňa mala v každom okamihu dostatok finančných prostriedkov na vyplatenie poistného plnenia klientovi.

3.1 Chain ladder – Mackov model

Metóda *chain ladder* je jedna z najpožívanejších metód na výpočet IBNR rezerv. Je to pre jej jednoduchosť a nezávislosť na rozdelení dát, či už uvažujeme množstvo poistných udalostí alebo poistné plnenia. Je však známe, že je citlivá na zmeny v pozorovaných dátach. V ostatných rokoch bolo publikovaných viacero metód na výpočet IBNR rezerv, ktorých výsledky sa líšili od výsledkov Mackovej chain ladder metódy. A preto uvedieme aj vzorec na výpočet štandardnej chyby pre odhad IBNR rezerv metódou chain ladder.

Označme C_{ik} celkový objem poistných plnení, na ktoré vznikol nárok v roku $i; 1 \le i \le I$ a ktoré boli vyplatené do roku $k; 1 \le k \le I$, kde I je počet sledovaných období. C_{ik} uvažujeme ako náhodnú premennú, ktorej realizácie pre $i + k \le I + 1$ poznáme, tzv. run-off triangle. Našim cieľom je odhadnúť konečný objem poistných plnení C_{iI} a rezervy na nenahlásené poistné udalosti $R_i = C_{iI} - C_{iI+1-i}$ pre rok vzniku $i = 2, 3, \ldots, I$.

Základným predpokladom metódy chain ladder je existencia vývojových faktorov $\lambda_1, \lambda_2, \ldots, \lambda_{I-1} > 0$ takých, že

$$E\left[C_{ik+1} \mid C_{i1}, \ldots, C_{ik}\right] = C_{ik}\lambda_k; 1 \le i \le I, 1 \le k \le I-1. \qquad (6)$$

Prvým krokom metódy chain ladder je odhad faktorov λ_k a to nasledovne:

$$\hat{\lambda}_k = \frac{\sum_{j=1}^{I-k} C_{jk+1}}{\sum_{j=1}^{I-k} C_{jk}}; \qquad 1 \le k \le I-1. \text{ ([3])}$$

Pomocou týchto faktorov potom odhadneme konečný objem poistných plnení C_{iI}

$$\hat{C}_{iI} = C_{iI+1-i} \cdot \hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{I-1}, \text{ ([3])}$$

alebo rezervy R_i

$$\hat{R}_i = C_{iI+1-i}\left(\hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{I-1} - 1\right). \text{ ([3])}$$

Keďže algoritmus metódy chain ladder neberie do úvahy akúkoľvek závislosť medzi rokmi vzniku nárokov, ďalším predpokladom tejto metódy je nezávislosť premenných $\{C_{i1},\cdots,C_{iI}\},\{C_{j1},\cdots,C_{jI}\}; i \neq j$. ([3]) (7)

Tvrdenie 3:

Nech $M = \{C_{ik} \mid i + k \leq I + 1\}$ je množina všetkých doposiaľ pozorovaných dát. Za predpokladov (6) a (7) platí

$$E[C_{iI} \mid M] = C_{iI+1-i}\lambda_{I+1-i} \cdot \ldots \cdot \lambda_{I-1}.$$

Dôkaz tohto tvrdenia ako aj samotné tvrdenie možno nájsť v [3].

Tvrdenie 3 ukazuje, že odhad $\hat{C}_{iI}$ má rovnaký tvar ako podmienená stredná hodnota $E[C_{iI} \mid M]$, ktorá je najlepším odhadom pre C_{iI} založeným na pozorovaniach M. Nasledujúce tvrdenie, ktoré aj s dôkazom nájdeme v [3], hovorí o vlastnostiach odhadov $\hat{\lambda}_k$.

Tvrdenie 4:

Za platnosti (6) a (7) sú odhady $\hat{\lambda}_k ; 1 \leq k \leq I - 1$, nevychýlené a nekorelované.

Dôkaz:

Z teórie pravdepodobnosti platí:

$$E[X] = E(E[X \mid Y]). \tag{\#}$$

$\rightarrow$ <u>nevychýlenosť</u>: chceme ukázať, že $E[\hat{\lambda}_k] = \lambda_k$.

Označme si $B_k = \{C_{ij}; j \leq k, i + j \leq I + 1\}$. Za platnosti predpokladov (6) a (7) platí:

$$E[C_{ik+1} \mid B_k] \overset{(7)}{=} E[C_{ik+1} \mid C_{i1},\ldots,C_{ik}] \overset{(6)}{=} C_{ik}\lambda_k. \tag{\#\#}$$

Teda

$$E[\hat{\lambda}_k \mid B_k] = E\left[\frac{\sum_{i=1}^{I-k} C_{ik+1}}{\sum_{i=1}^{I-k} C_{ik}} \mid B_k\right] = \frac{E\left[\sum_{i=1}^{I-k} C_{ik+1} \mid B_k\right]}{\sum_{i=1}^{I-k} C_{ik}} \overset{(\#\#)}{=} \lambda_k.$$

Teraz už máme všetko na to, aby sme ukázali nevychýlenosť odhadov $\hat{\lambda}_k$:

$$E[\hat{\lambda}_k] \overset{(\#)}{=} E[E(\hat{\lambda}_k \mid B_k)] = E[\lambda_k] = \lambda_k; 1 \leq k \leq I - 1.$$

→ nekorelovanosť:

Pre $j < k$ platí:

$$E\left[\hat{\lambda}_j \hat{\lambda}_k\right] \overset{(\#)}{=} E\left[E\left(\hat{\lambda}_j \hat{\lambda}_k \mid B_k\right)\right] = E\left[\hat{\lambda}_j E\left(\hat{\lambda}_k \mid B_k\right)\right] = E\left[\hat{\lambda}_j \lambda_k\right] = E\left[\hat{\lambda}_j\right]\lambda_k = E\left[\hat{\lambda}_j\right] \cdot E\left[\hat{\lambda}_k\right]$$

□

Toto tvrdenie potvrdzuje, že je rozumné súčin $\lambda_{I+1-i} \cdot \ldots \cdot \lambda_{I-1}$ odhadovať súčinom $\hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{I-1}$. Teda $\hat{C}_{iI} = C_{iI+1-i} \cdot \hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{I-1}$ je nevychýleným odhadom pre $E\left[C_{iI} \mid M\right] = C_{iI+1-i}\lambda_{I+1-i} \cdot \ldots \cdot \lambda_{I-1}$ a odhad rezerv $\hat{R}_i = \hat{C}_{iI} - C_{iI+1-i}$ je nevychýleným odhadom pre $R_i = C_{iI} - C_{iI+1-i}$. ([3])

Na Mackov článok [3] nadviazal v roku 2003 Greg Taylor článkom [7], ktorý sa zaoberal aj prípadom, že sú nezávislé prírastky $X_{ij} = C_{ij} - C_{ij-1}$ pre všetky $i, j = 1, \ldots, I$.

3.1.1 Prípad nezávislých prírastkov

Nahraďme predpoklady (6) a (7) nasledujúcimi predpokladmi: ([7])

- všetky prírastky X_{ij} sú nezávislé (7a)
- $\dfrac{E[C_{ij+1}]}{E[C_{ij}]} = \eta_j$. (6a)

Označme

$S_1(i) = \left\{(g,h) : g \leq I - k - 1, h \leq k + 1, k = I - i, \ldots, I - 1\right\}$

$S_2(i) = \left\{(g,h) : g \leq I - k - 1, h \leq k, k = I - i, \ldots, I - 1\right\}$.

- $\sum_{k=1}^{i} C_{kj} > 0$ pre každú dvojicu $(i, j) \in S_2(i)$. (8)

Tvrdenie 5:

Nech platí (6a), (7a) a (8). Nech pre aspoň jednu dvojicu $(g,h) \in S_1(i)$ platí, že $X_{gh} > 0$. Potom odhad $\hat{C}_{iI}$ je vychýleným odhadom pre $E\left[C_{iI}\right]$ v zmysle

$$E\left[\hat{C}_{iI} \mid C_{iI+1-i}\right] > C_{iI+1-i} \frac{E\left[C_{iI}\right]}{E\left[C_{iI+1-i}\right]},$$

čo môžeme na základe nezávislosti prírastkov prepísať nasledovne

$$E\left[\hat{C}_{iI}\right] > E\left[C_{iI}\right].$$

([7])

Vieme, že odhad rezerv $\hat{R}_i = C_{iI+1-i}\left(\hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{I-1} - 1\right)$. Ak platí (7a), potom R_i a C_{iI+1-i} sú nezávislé, pretože sú definované na disjunktných množinách prírastkov X_{ij}. Potom z tvrdenia 5 vyplýva, že $E\left[\hat{R}_i\right] > E\left[R_i\right]$, teda odhad rezerv je v prípade nezávislých prírastkov vychýleným odhadom.

3.1.2 Stredná kvadratická chyba a štandardná odchýlka

Vráťme sa k prípadu nezávislých rokov vzniku nárokov. Zaujíma nás, aká je podmienená stredná kvadratická chyba odhadu $\hat{C}_{iI}$ celkového objemu plnení C_{iI}. Definujeme si túto podmienenú strednú kvadratickú chybu $mse\left(\hat{C}_{iI}\right)$ nasledovne:

$$mse\left(\hat{C}_{iI}\right) = E\left[\left(\hat{C}_{iI} - C_{iI}\right)^2 \mid M\right], \tag{9}$$

kde $M = \left\{C_{ik} \mid i + k \le I + 1\right\}$ je množina všetkých doposiaľ pozorovaných dát. Platí:

$$mse\left(\hat{R}_i\right) = E\left[\left(\hat{R}_i - R_i\right)^2 \mid M\right] = E\left[\left(\hat{C}_{iI} - C_{iI}\right)^2 \mid M\right] = mse\left(\hat{C}_{iI}\right).$$

Podľa všeobecne známeho vzorca $E\left[(X-a)^2\right] = Var(X) + \left(E[X] - a\right)^2$ môžeme (9) napísať v tvare $mse\left(\hat{C}_{iI}\right) = Var\left(C_{iI} \mid M\right) + \left(E\left[C_{iI} \mid M\right] - \hat{C}_{iI}\right)^2$. Na výpočet $mse\left(\hat{C}_{iI}\right)$ potrebujeme nájsť vzorec na výpočet $Var\left(C_{iI} \mid M\right)$. Zo spôsobu výpočtu odhadu faktora $\hat{\lambda}_k$, ktorý je vlastne C_{ik}-váženým priemerom podielov $\frac{C_{ik+1}}{C_{ik}}, 1 \le i \le I - k$, môžeme tvrdiť, že $Var\left(\frac{C_{ik+1}}{C_{ik}} \mid C_{i1}, \ldots, C_{ik}\right)$ je nepriamo úmerná C_{ik}, čo je možné zapísať nasledovne:

$$Var\left(C_{ik+1} \mid C_{i1}, \ldots, C_{ik}\right) = C_{ik}\sigma_k^2; 1 \le i \le I, 1 \le k \le I - 1, \text{ ([3])} \tag{10}$$

kde $\sigma_k^2; 1 \le k \le I - 1$, je neznámy parameter, ktorý musíme odhadnúť. Dá sa ukázať, že

$$\hat{\sigma}_k^2 = \frac{1}{I - k - 1}\sum_{i=1}^{I-k} C_{ik}\left(\frac{C_{ik+1}}{C_{ik}} - \hat{\lambda}_k\right)^2; 1 \le k \le I - 2 \text{ ([3])}$$

je nevychýleným odhadom pre $\sigma_k^2; 1 \le k \le I-2$. Ak $\hat{\lambda}_{I-1} = 1$ a po $(I-1)$ rokoch už neočakávame žiadne plnenia, môžeme položiť $\hat{\sigma}_{I-1} = 0$. V inom prípade odhadneme rad $\hat{\sigma}_1, \ldots, \hat{\sigma}_{I-2}$ s jedným pridaným členom, napríklad log-lineárnou regresiou.

Teraz môžeme vysloviť tvrdenie, ktoré aj s dôkazom nájdeme v [3].

Tvrdenie 6:

Za predpokladu platnosti (6), (7) a (10) môžeme strednú kvadratickú chybu $mse(\hat{R}_i)$ odhadnúť nasledovne:

$$\widehat{mse}(\hat{R}_i) = \hat{C}_{iI} \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{\lambda}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right),$$

kde $\hat{C}_{ik} = C_{iI+1-i} \cdot \hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{k-1}; k > I+1-i$, sú odhadnuté budúce hodnoty C_{ik} a $\hat{C}_{iI+1-i} = C_{iI+1-i}$ je známa napozorovaná hodnota.

Dôkaz:

Označme si $E_i[X] = E[X \mid C_{i1}, \ldots, C_{iI+1-i}]$ a

$Var_i[X] = Var[X \mid C_{i1}, \ldots, C_{iI+1-i}]$.

Z teórie pravdepodobnosti vieme, že platí

$$E[X] = E(E[X \mid Y]), \tag{*}$$

$$Var[X] = Var(E[X \mid Y]) + E(Var[X \mid Y]). \tag{**}$$

Ako sme už skôr ukázali, $mse(\hat{R}_i) = mse(\hat{C}_{iI})$. Potom môžeme písať

$$mse(\hat{R}_i) = Var(C_{iI} \mid M) + \left(E[C_{iI} \mid M] - \hat{C}_{iI}\right)^2.$$

Aby sme získali odhad $\widehat{mse}(\hat{R}_i)$ pre $mse(\hat{R}_i)$, potrebujeme odhadnúť výrazy $Var(C_{iI} \mid M)$ a $\left(E[C_{iI} \mid M] - \hat{C}_{iI}\right)^2$.

Odhadnime najprv výraz $Var(C_{iI} \mid M)$. Za predpokladu nezávislosti rokov vzniku nárokov (7) platí: $Var(C_{iI} \mid M) = Var(C_{iI} \mid C_{i1}, \ldots, C_{iI+1-i}) = Var_i(C_{iI})$. Pomocou (*), (**) a za platnosti predpokladov (6) a (10) si $Var_i\left(C_{iI}\right)$ vyjadríme nasledovne:

$$Var_i\left(C_{iI}\right) \overset{(**)}{=} Var_i\left(E\left[C_{iI} \mid C_{i1},\dots,C_{iI-1}\right]\right) + E_i\left(Var\left[C_{iI} \mid C_{i1},\dots,C_{iI-1}\right]\right) \overset{(1),(10)}{=}$$

$$= Var_i\left(C_{iI-1}\right)\lambda_{I-1}^2 + E_i\left(C_{iI-1}\right)\sigma_{I-1}^2 \overset{(*),(**)}{=}$$

$$= Var_i\left(E\left[C_{iI-1} \mid C_{i1},\dots,C_{iI-2}\right]\right)\lambda_{I-1}^2 + E_i\left(Var\left[C_{iI-1} \mid C_{i1},\dots,C_{iI-2}\right]\right)\lambda_{I-1}^2 +$$

$$+ E_i\left(E\left[C_{iI-1} \mid C_{i1},\dots,C_{iI-2}\right]\right)\sigma_{I-1}^2 \overset{(1),(10)}{=}$$

$$= Var_i\left(C_{iI-2}\right)\lambda_{I-2}^2\lambda_{I-1}^2 + E_i\left(C_{iI-2}\right)\sigma_{I-2}^2\lambda_{I-1}^2 + E_i\left(C_{iI-2}\right)\lambda_{I-2}\sigma_{I-1}^2 =$$

$$\vdots$$

$$= \sum_{k=I+1-i}^{I-1} \lambda_{I+1-i}\cdot\ldots\cdot\lambda_{k-1}\sigma_k^2\lambda_{k+1}^2\cdot\ldots\cdot\lambda_{I-1}^2\cdot C_{iI+1-i}$$

pretože C_{iI+1-i} je známa hodnota, teda $E_i\left(C_{iI+1-i}\right) = C_{iI+1-i}$ a $Var_i\left(C_{iI+1-i}\right) = 0$.

Odhad pre $Var\left(C_{iI} \mid M\right)$ získame tak, že neznáme parametre λ_k a σ_k^2 nahradíme ich odhadmi $\hat{\lambda}_k$ a $\hat{\sigma}_k^2$.

Teda

$$\widehat{Var\left(C_{iI} \mid M\right)} = C_{iI+1-i}\left(\sum_{k=I+1-i}^{I-1} \hat{\lambda}_{I+1-i}\cdot\ldots\cdot\hat{\lambda}_{k-1}\cdot\hat{\sigma}_k^2\cdot\hat{\lambda}_{k+1}^2\cdot\ldots\cdot\hat{\lambda}_{I-1}^2\right) =$$

$$= C_{iI+1-i}\left(\sum_{k=I+1-i}^{I-1} \frac{\hat{\lambda}_{I+1-i}^2\cdot\ldots\cdot\hat{\lambda}_{k-1}^2\cdot\hat{\sigma}_k^2\cdot\hat{\lambda}_{k+1}^2\cdot\ldots\cdot\hat{\lambda}_{I-1}^2}{\hat{\lambda}_{I+1-i}\cdot\ldots\cdot\hat{\lambda}_{k-1}}\right) =$$

$$= \sum_{k=I+1-i}^{I-1} \frac{\left(C_{iI+1-i}\cdot\hat{\lambda}_{I+1-i}\cdot\ldots\cdot\hat{\lambda}_{I-1}\right)^2\cdot\hat{\sigma}_k^2}{\left(C_{iI+1-i}\cdot\hat{\lambda}_{I+1-i}\cdot\ldots\cdot\hat{\lambda}_{k-1}\right)\hat{\lambda}_k^2} =$$

$$= \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{C_{ik}\hat{\lambda}_k^2}$$

V konečnom vyjadrení už poznáme potrebné hodnoty, teda odhad pre $Var\left(C_{iI} \mid M\right)$ už poznáme. Poďme teraz odhadnúť $\left(E\left[C_{iI} \mid M\right] - \hat{C}_{iI}\right)^2$, druhú zložku vyjadrenia $mse\left(\hat{R}_i\right)$.

Z tvrdenia 3 vyplýva:

$$\left(E\left(C_{iI} \mid M\right) - \hat{C}_{iI}\right)^2 = C_{iI+1-i}^2\left(\lambda_{I+1-i}\cdot\ldots\cdot\lambda_{I-1} - \hat{\lambda}_{I+1-i}\cdot\ldots\cdot\hat{\lambda}_{I-1}\right)^2.$$

V tomto výraze ale nemôžeme λ_k nahradiť odhadom $\hat{\lambda}_k$, pretože by sa tento výraz potom rovnal nule. Použijeme preto iný prístup.

Označme si $F_k = \hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{k-1}\left(\lambda_k - \hat{\lambda}_k\right)\lambda_{k+1} \cdot \ldots \cdot \lambda_{I-1}$.

Potom

$$F = \lambda_{I+1-i} \cdot \ldots \cdot \lambda_{I-1} - \hat{\lambda}_{I+1-i} \cdot \ldots \cdot \hat{\lambda}_{I-1} = F_{I+1-i} + \ldots + F_{I-1},$$

a

$$F^2 = \sum_{k=I+1-i}^{I-1} F_k^2 + 2 \sum_{j,k:j<k} F_j F_k .$$

Tento výraz aproximujeme nad najmenšou možnou množinou pozorovaných dát, teda F_k^2 nahradíme $E\left(F_k^2 \mid B_k\right)$ a $F_j F_k$ nahradíme $E\left(F_j F_k \mid B_k\right)$, kde $B_k = \left\{C_{ij}; j \le k, i + j \le I + 1\right\}$. Z dôkazu tvrdenia 4 vyplýva: $E\left[\lambda_k - \hat{\lambda}_k \mid B_k\right] = 0$. Potom platí: $E\left[F_j F_k \mid B_k\right] = 0$ pre $j < k$.

Pretože

$$E\left[\left(\lambda_k - \hat{\lambda}_k\right)^2 \mid B_k\right] = Var\left(\hat{\lambda}_k \mid B_k\right) = Var\left(\frac{\sum_{i=1}^{I-k} C_{ik+1}}{\sum_{i=1}^{I-k} C_{ik}} \mid B_k\right) = \frac{Var\left(\sum_{i=1}^{I-k} C_{ik+1} \mid B_k\right)}{\left(\sum_{i=1}^{I-k} C_{ik}\right)^2} \overset{(7)}{=}$$

$$= \frac{\sum_{i=1}^{I-k} Var\left(C_{ik+1} \mid B_k\right)}{\left(\sum_{i=1}^{I-k} C_{ik}\right)^2} \overset{(10)}{=} \frac{\sigma_k^2}{\sum_{i=1}^{I-k} C_{ik}},$$

pre $E\left(F_k^2 \mid B_k\right)$ platí:

$$E\left(F_k^2 \mid B_k\right) = \frac{\hat{\lambda}_{I+1-i}^2 \cdot \ldots \cdot \hat{\lambda}_{k-1}^2 \cdot \sigma_k^2 \cdot \lambda_{k+1}^2 \cdot \ldots \cdot \lambda_{I-1}^2}{\sum_{i=1}^{I-k} C_{ik}} > 0 .$$

Pretože všetky takéto výrazy sú kladné, tak v nich môžeme λ_k nahradiť odhadom $\hat{\lambda}_k$ a σ_k^2 nahradíme odhadom $\hat{\sigma}_k^2$.

Dostávame odhad pre F^2.

$$\hat{F}^2 = \sum_{k=I+1-i}^{I-1} \frac{\hat{\lambda}_{I+1-i}^2 \cdot \ldots \cdot \hat{\lambda}_{k-1}^2 \cdot \hat{\sigma}_k^2 \cdot \hat{\lambda}_{k+1}^2 \cdot \ldots \cdot \hat{\lambda}_{I-1}^2}{\sum_{j=1}^{I-k} C_{jk}} = \hat{\lambda}_{I+1-i}^2 \cdot \ldots \cdot \hat{\lambda}_{I-1}^2 \cdot \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{\lambda}_k^2 \cdot \sum_{j=1}^{I-k} C_{jk}}$$

Teraz už poznáme všetky potrebné výrazy a môžeme odvodiť odhad pre $mse\left(\hat{R}_i\right)$:

$$\widehat{mse}\left(\hat{R}_i\right) = \hat{C}_{iI}^2 \cdot \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{C}_{ik}\hat{\lambda}_k^2} + C_{I+1-i}^2 \cdot \hat{\lambda}_{I+1-i}^2 \cdot \ldots \cdot \hat{\lambda}_{I-1}^2 \cdot \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{\lambda}_k^2 \sum_{j=1}^{I-k} C_{jk}} =$$

$$= \hat{C}_{iI}^2 \cdot \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{\lambda}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

□

Štandardná chyba $SE\left(\hat{R}_i\right)$ je definovaná ako druhá odmocnina z odhadu strednej kvadratickej chyby. Nás okrem $SE\left(\hat{R}_i\right)$ zaujíma aj štandardná chyba celkovej odhadnutej rezervy $\hat{R} = \hat{R}_2 + \ldots + \hat{R}_I$. O odhade strednej kvadratickej chyby pre odhad celkovej rezervy hovorí dôsledok tvrdenia 6 ([3] aj s dôkazom).

<u>*Dôsledok T6:*</u>

Za platnosti predpokladov tvrdenia 6 a pri nezmenenom označení strednú kvadratickú chybu odhadu celkovej rezervy $\hat{R} = \hat{R}_2 + \ldots + \hat{R}_I$ odhadujeme takto

$$\widehat{mse}\left(\hat{R}\right) = \sum_{i=2}^{I} \left\{ \left(SE\left(\hat{R}_i\right)\right)^2 + \hat{C}_{iI} \left(\sum_{j=i+1}^{I} \hat{C}_{jI} \right) \sum_{k=I+1-i}^{I-1} \frac{2\hat{\sigma}_k^2 / \hat{\lambda}_k^2}{\sum_{n=1}^{I-k} C_{nk}} \right\}.$$

Odhad strednej kvadratickej chyby odhadu celkovej rezervy $\hat{R} = \hat{R}_2 + \ldots + \hat{R}_I$ nemôžme rátať ako súčet odhadov stredných kvadratických chýb, pretože tieto odhady $\hat{R}_i$ sú navzájom korelované cez spoločné odhady faktorov $\hat{\lambda}_k$ a odhady $\hat{\sigma}_k$.

3.2 Metóda chain ladder podľa Roba Kaasa

Rob Kaas vo svojej knihe *Modern Actuarial Risk Theory* [2] založil metódu chain ladder na modeli $X_{ij} \approx \alpha_i \beta_j \gamma_k ; k = i + j - 1$, kde X_{ij} sú, ako v predošlej podkapitole, prírastky. Predpokladajme, že $X_{ij} \sim Po\left(\alpha_i \beta_j\right); \gamma_k \equiv 1$ a sú nezávislé.

Idea tejto chain ladder metódy je taká, že v j-tom roku vývoja sa vyplatí rovnaké percento z celkového objemu plnení v každom roku vzniku nároku. Toto percento v našom modeli reprezentuje parameter β_j. Znázorníme si túto situáciu schémou run-off triangle:

$$\begin{array}{cccccc} & 1 & \cdots & I-i+1 & \cdots & I \\ 1 & X_{11} & & X_{1I-i+1} & & X_{1I} \\ \vdots & & & & & \\ i & X_{i1} & & X_{iI-i+1} & & \\ \vdots & & & & & \\ I & X_{I1} & & & & \end{array} = \begin{array}{cccccc} & 1 & \cdots & I-i+1 & \cdots & I \\ 1 & \alpha_1\beta_1 & & \alpha_1\beta_{I-i+1} & & \alpha_1\beta_I \\ \vdots & & & & & \\ i & \alpha_i\beta_1 & & \alpha_i\beta_{I-i+1} & & \\ \vdots & & & & & \\ I & \alpha_I\beta_1 & & & & \end{array},$$

kde stĺpce reprezentujú rok vývoja a riadky rok vzniku nároku. Ďalej predpokladajme, že vývoj je po I rokoch po vzniku nároku ukončený, z čoho vyplýva, že $\beta_1 + \ldots + \beta_I = 1$ ak α_i je celkový objem plnení v i-tom roku vzniku nároku. Našou úlohou je na základe pozorovaní $X_{ij}; i, j = 1, \ldots, I; i + j \leq I$ odhadnúť parametre α_i a β_j.

<u>*Vlastnosť 1:*</u> (Property 8.3.9 v [2])

Predpokladajme, že počet poistných udalostí n_{ij} v bunke (i, j) má Poissonovo rozdelenie s parametrom $\lambda_{ij} = \alpha_i\beta_j$. Potom odhady parametrov α_i a β_j metódou maximálnej vierohodností sú rovnaké ako odhady metódou marginálnych súčtov.

Označme si V_i celkový objem známych plnení v i-tom roku vzniku nárokov a K_j nech je celkový objem známych plnení vyplatených v j-tom roku vývoja. Podľa vlastnosti 1 platí: $V_i = \sum_j \hat{\alpha}_i \hat{\beta}_j$ a $K_j = \sum_i \hat{\alpha}_i \hat{\beta}_j$. Nech $\beta_1 + \ldots + \beta_I = 1$ a teda aj $\hat{\beta}_1 + \ldots + \hat{\beta}_I = 1$. Pri odhadovaní parametrov postupujeme nasledovne: ([2])

1. Pre prvý riadok platí rovnosť $\hat{\alpha}_1\left(\hat{\beta}_1 + \ldots + \hat{\beta}_I\right) = V_1$. Odtiaľ dostávame $\hat{\alpha}_1 = V_1$. Z rovnosti $\hat{\alpha}_1\hat{\beta}_I = K_I$ vypočítame hodnotu odhadu parametra β_I.

2. Predpokladajme, že pre nejaké $t < I$ už poznáme odhady $\hat{\beta}_{t+1}, \ldots, \hat{\beta}_I$ a $\hat{\alpha}_1, \ldots, \hat{\alpha}_{I-t}$. Riešime rovnice:

$$\hat{\alpha}_{I-t+1}\left(\hat{\beta}_1 + \ldots + \hat{\beta}_t\right) = V_{I-t+1}$$

$$\left(\hat{\alpha}_1 + \ldots + \hat{\alpha}_{I-t+1}\right)\hat{\beta}_t = K_t.$$

Prvou z týchto rovníc vypočítame hodnotu odhadu parametra $\hat{\alpha}_{I-t+1}$ a potom pomocou druhej vypočítame hodnotu odhadu parametra $\hat{\beta}_t$.

3. Opakujeme krok 2 pre $t = I-1, I-2, \dots, 1$.

Ak poznáme všetky hodnoty odhadovaných parametrov, môžeme run-off triangle doplniť na štvorec a dorátať odhad IBNR rezerv a to tak, že sčítame hodnoty práve odhadnutých budúcich plnení.

Iný prístup zvolili England a Verral v [1], ktorí si model $X_{ij} \approx \alpha_i \beta_j \gamma_k ; \gamma_k \equiv 1$ linearizovali. Tu sa autori zaoberajú aj s modelom predpokladajúcim negatívne binomické rozdelenie, ktorý sa spolu aj s Poissonovým modelom od Mackovho modelu líši výpočtom strednej kvadratickej chyby.

3.3 Separačná metóda

Táto metóda, ktorú nájdeme v [2], je založená na modeli $X_{ij} \approx \alpha_i \beta_j \gamma_k ; k = i + j - 1$. Pri tejto metóde nám na prírastky X_{ij} vplývajú dva časové aspekty, efekt kalendárneho roku γ_k, kde $k = i + j - 1$, a efekt roku vývoja β_j. V tomto modeli opäť predpokladáme, že v každom roku vzniku nároku sa v *j*-tom vývojovom roku vyplatí rovnaké percento β_j objemu plnení a vývoj je po *I* rokoch ukončený. V tejto kapitole si ukážeme dve metódy – aritmetickú a geometrickú separačnú metódu.

3.3.1 Aritmetická separačná metóda

V tomto modeli predpokladáme

$$X_{ij} \sim Po(\beta_j \gamma_k), \alpha_i \equiv 1, \text{ nezávislé,}$$

kde $k = i + j - 1$. Parametre β_j a γ_k vo všeobecnosti odhadujeme metódou maximálnej vierohodnosti. Keďže my predpokladáme Poissonovo rozdelenie, môžeme tieto parametre podľa vlastnosti 1 odhadovať metódou marginálnych súčtov, ktorá dáva tie isté výsledky, ako metóda maximálnej vierohodnosti. Marginálne súčty v tomto prípade sú:

- $K_j = \sum_k \hat{\beta}_j \hat{\gamma}_k$ súčet známych hodnôt X_{ij} v *j*-stĺpci,
- $G_k = \sum_j \hat{\beta}_j \hat{\gamma}_k$ súčet hodnôt X_{ij} po *k*-tej diagonále, kde $k = i + j - 1$.

Postup je podobný ako pri Kaasovej chain ladder metóde. Opäť predpokladajme, že $\beta_1 + \ldots + \beta_I = 1$ a teda aj $\hat{\beta}_1 + \ldots + \hat{\beta}_I = 1$. Run-off triangle v tomto prípade vyzerá nasledovne:

$$\begin{array}{c|ccccc} & 1 & \cdots & I-i+1 & \cdots & I \\ 1 & X_{11} & & X_{1I-i+1} & & X_{1I} \\ \vdots & & & & & \\ i & X_{i1} & & X_{iI-i+1} & & \\ \vdots & & & & & \\ I & X_{I1} & & & & \end{array} = \begin{array}{c|ccccc} & 1 & \cdots & I-i+1 & \cdots & I \\ 1 & \beta_1\gamma_1 & & \beta_{I-i+1}\gamma_{I-i+1} & & \beta_I\gamma_I \\ \vdots & & & & & \\ i & \beta_1\gamma_i & & \beta_{I-i+1}\gamma_I & & \\ \vdots & & & & & \\ I & \beta_1\gamma_I & & & & \end{array}$$

Postup je nasledovný: ([2])

1. Pre ostatnú známu diagonálu v run-off triangle platí $\left(\hat{\beta}_1 + \ldots + \hat{\beta}_I\right)\hat{\gamma}_I = G_I$. Odtiaľ dostávame $\hat{\gamma}_I = G_I$. Z rovnosti $X_{1I} = K_I = \hat{\beta}_I \hat{\gamma}_I$ dostávame odhad $\hat{\beta}_I = \dfrac{X_{1I}}{\hat{\gamma}_I}$.

2. Predpokladajme, že pre nejaké $t < I$ už poznáme odhady $\hat{\beta}_{t+1}, \ldots, \hat{\beta}_I$ a $\hat{\gamma}_{t+1}, \ldots, \hat{\gamma}_I$. Riešime rovnice:

$$\left(\hat{\beta}_1 + \ldots + \hat{\beta}_t\right)\hat{\gamma}_t = G_t$$

$$\hat{\beta}_t\left(\hat{\gamma}_t + \ldots + \hat{\gamma}_I\right) = K_t$$

Pomocou prvej z rovníc vypočítame odhad $\hat{\gamma}_t$ parametra γ_t a potom pomocou druhej rovnice vypočítame odhad $\hat{\beta}_t$ parametra β_t.

3. Krok 2 opakujeme pre $t = I-1, I-2, \ldots, 1$.

Na doplnenie nášho run-off triangle na štvorec potrebujeme okrem už odhadnutých parametrov poznať aj parametre $\gamma_{I+1}, \ldots, \gamma_{2I}$. Tieto parametre odhadneme na základe hodnôt $\hat{\gamma}_1, \ldots, \hat{\gamma}_I$. Jednou z techník je loglineárna extrapolácia.

Pre túto aritmetickú separačnú metódu platí, že $E\left[X_{ij}\right] = \beta_j \gamma_{i+j-1}$.

3.3.2 Geometrická separačná metóda

Táto metóda je založená na modeli

$$\ln\left(X_{ij}\right) \sim N\left(\ln\left(\beta_j \gamma_k\right), \sigma^2\right); \alpha_i \equiv 1, \text{ nezávislé},$$

kde $k = i + j - 1$. σ^2 je neznámy parameter. Dostávame regresný model $E\left[\ln\left(X_{ij}\right)\right] = \ln\left(\beta_j\right) + \ln\left(\gamma_{i+j-1}\right)$, ktorého parametre môžeme odhadovať obvyklým spôsobom alebo rekurentne, ako v predošlej metóde. Pri rekurentnej metóde predpokladáme, že $\prod_{j=1}^{I}\beta_j = 1$ ([2]). Označme si $SK_j = \prod_k \hat{\beta}_j \hat{\gamma}_k$ súčin známych hodnôt X_{ij} v j-tom stĺpci a $SG_k = \prod_j \hat{\beta}_j \hat{\gamma}_k$ súčin známych hodnôt X_{ij} na k-tej diagonále, kde $k = i + j - 1$. Postupujeme podobne, ako v predošlej metóde:

1. Za predpokladu $\prod_{j=1}^{I}\beta_j = 1$ si z rovnice $SG_I = \prod_j \hat{\beta}_j \hat{\gamma}_I = \hat{\gamma}_I^I \cdot \prod_{j=1}^{I} \hat{\beta}_j = \hat{\gamma}_I^I$ pre ostatnú známu diagonálu vypočítame hodnotu odhadu $\hat{\gamma}_I = \sqrt[I]{SG_I}$, pomocou ktorej potom vypočítame hodnotu odhadu $\hat{\beta}_I = \dfrac{SK_I}{\hat{\gamma}_I} = \dfrac{X_{1I}}{\hat{\gamma}_I}$.

2. Predpokladajme, že pre nejaké $t < I$ už poznáme odhady $\hat{\beta}_{t+1},\ldots,\hat{\beta}_I$ a $\hat{\gamma}_{t+1},\ldots,\hat{\gamma}_I$. Riešime rovnice:

$$\left(\hat{\beta}_1 \cdot \ldots \cdot \hat{\beta}_t\right)\hat{\gamma}_t^t = SG_t$$

$$\beta_t^{I-t+1}\left(\hat{\gamma}_t \cdot \ldots \cdot \hat{\gamma}_I\right) = SK_t .$$

Z prvej rovnice vyrátame odhad $\hat{\gamma}_t = \sqrt[t]{SG_t \cdot \left(\hat{\beta}_{t+1} \cdot \ldots \cdot \hat{\beta}_I\right)}$, pomocou ktorého potom z druhej rovnice vyrátame odhad $\hat{\beta}_t = \sqrt[I-t+1]{\dfrac{SK_t}{\left(\hat{\gamma}_t \cdot \ldots \cdot \hat{\gamma}_I\right)}}$.

3. Krok 2 opakujeme pre $t = I-1, I-2, \ldots, 1$.

Na doplnenie nášho run-off triangle na štvorec potrebujeme okrem už odhadnutých parametrov poznať aj parametre $\gamma_{I+1},\ldots,\gamma_{2I}$. Tieto parametre odhadneme na základe hodnôt $\hat{\gamma}_1,\ldots,\hat{\gamma}_I$.

Pri tejto metóde hodnoty $\beta_j \gamma_{i+j-1}$ nie sú strednými hodnotami náhodných premenných X_{ij}. Sú to mediány, teda platí: $\Pr\left(X_{ij} \leq \beta_j \gamma_{i+j-1}\right) = \dfrac{1}{2}$. Pre strednú hodnotu hodnotu platí: $E\left[X_{ij}\right] = e^{\sigma^2/2}\beta_j \gamma_{i+j-1}$. ([2])

Iný prístup môžeme opäť nájsť v [1].

4 Aplikácia opísaných metód na reálne dáta

V tejto kapitole aplikujeme teoretické poznatky uvedené v predošlých kapitolách na reálne dáta. Snahou bude vytvoriť program, ktorý vypočíta požadované údaje. Programovať budeme v prostredí programovacieho jazyka R, ktorý je voľne dostupný na http://cran.r-project.org/. Zdrojové kódy programov obsahuje príloha tejto práce.

4.1 Odhad počtu nenahlásených poistných udalostí

Riešenie tohto problému spočíva v hľadaní argumentu maxima funkcie vierohodnosti, ktorý definuje maximálne vierohodný odhad pre vektor parametrov. Platí, že argument maxima funkcie vierohodnosti sa rovná argumentu maxima prirodzeného logaritmu funkcie vierohodnosti, pretože tieto dve funkcie majú rovnaký priebeh. My budeme pracovať s prirodzeným logaritmom funkcie vierohodnosti, pretože je to v tomto prípade jednoduchšie. Postupujeme nasledovne: naprogramujeme si funkciu, ktorá nám na základe vstupných parametrov vypočíta prirodzený logaritmus funkcie vierohodnosti. Vstupné parametre sú parametre rozdelení, na ktorých je model založený, a dáta, ktoré už poznáme, teda vektor pozorovaných počtov, vektor núl a jednotiek, ktorý je realizáciou indikátorovej náhodnej premennej D, a vektor jednotiek miery rizika. V prípade, že riziko nechceme uvažovať, za vektor jednotiek miery rizika volíme vektor so samými jednotkami.

V ďalšom kroku hľadáme maximum naprogramovanej funkcie. Využívame na to funkciu *optim*, ktorú v programe R obsahuje knižnica *stats4*. Táto funkcia je naprogramovaná tak, aby našla minimum danej funkcie a parametre funkcie, pri ktorých nadobúda toto minimum. My ale riešime maximalizačnú úlohu, preto si musíme zadať aj hodnotu škálovacieho parametra *fnscale*. Parameter *fnscale* nastavíme na hodnotu -1.

Naprogramované funkcie aplikujeme na dáta zo Štatistickej ročenky o vývoji na poistnom trhu v SR 1991 – 2001 [8], konkrétne na počet vybavených poistných udalostí v neživotnom poistení – poistenie majetku a poistenie zákonnej zodpovednosti z prevádzky motorového vozidla. Ako vektor jednotiek miery rizika uvažujeme stav kmeňa v neživotnom poistení. Tieto dáta sme zvolili z toho dôvodu, že sa nepodarilo nájsť reálne voľne prístupné údaje o počte hlásených poistných udalostí.

V práci predpokladáme, že vieme určiť, či boli pozorované nejaké nehody, ktoré neboli nahlásené poisťovni. Tento údaj vystupuje ako známy parameter. V skutočnosti však nevieme povedať, či sú nejaké poistné udalosti nenahlásené, a preto je poistný matematik nútený tento údaj na základe svojich skúseností odhadnúť. Je vhodné riadiť sa vývojom z minulosti. Pretože odvetvia neživotného poistenia sú rôznorodé, je ťažké vo všeobecnosti povedať ako postupovať. Pozrime sa napríklad na *havarijné poistenie motorového vozidla*. Tento produkt neživotného poistenia slúži na krytie rizika na konkrétnom motorovom vozidle. Predpokladáme, že ak nastane nejaká poistná udalosť a klient sa rozhodne túto udalosť poisťovni nahlásiť, nebude čakať niekoľko rokov, aby tak urobil. Iný priebeh má napríklad *povinné zmluvné poistenie zodpovednosti za škodu spôsobenú prevádzkou motorového vozidla*, ktoré okrem iného slúži aj na krytie rizika škody na zdraví. Nie je výnimočným javom, že sa zdravotné komplikácie, ktoré boli spôsobené dopravnou nehodou, objavia aj po niekoľkých rokoch.

Keďže my pri týchto dátach nepoznáme celkový vývoj, tak si realizácie indikátorovej premennej D zvolíme nasledovne: pri poistení majetku predpokladáme nenahlásené poistné udalosti len štyri roky dozadu a pri poistení zákonnej zodpovednosti z prevádzky motorového vozidla sedem rokov dozadu.

4.1.1 Poistenie majetku

Vstupné údaje použité pri odhade nenahlásených poistných udalostí v poistení majetku obsahuje tabuľka 1. V prípade, že nechceme uvažovať mieru rizika, jednotky miery rizika budú 1.

Tabuľka 1: *Vstupné dáta – poistenie majetku*

Rok	Počet PU	Stav kmeňa	Jednotky miery rizika
1991	200598	1844566	1,844566
1992	258694	2010371	2,010371
1993	260570	2034462	2,034462
1994	274343	1995357	1,995357
1995	252049	2003209	2,003209
1996	255056	2194779	2,194779
1997	245929	2607016	2,607016
1998	217694	2528047	2,528047
1999	283984	2650081	2,650081
2000	241394	2791316	2,791316
2001	217552	3038205	3,038205

V tabuľke 2 si zhrnieme výstupné údaje, ktoré sme vypočítali pomocou naprogramovaných funkcií.

Tabuľka 2: *Odhadnuté parametre a odhad počtu nenahlásených poistných udalostí – poistenie majetku*

	φ	θ	λ	**Odhad počtu nenahlásených poistných udalostí**
Poissonovo rozdelenie bez miery rizika	1,000000	246 169,700	-	0,027
Zovšeobecnené Poissonovo rozdelenie bez miery rizika	0,998998	1 623,838	0,993407	2 714,523
Poissonovo rozdelenie s mierou rizika	1,000000	105 375,100	-	0,027
Zovšeobecnené Poissonovo rozdelenie s mierou rizika	0,999206	312,293	0,997325	2 381,788

Na základe výsledkov v tabuľke 2 môžeme tvrdiť, že Poissonovo rozdelenie nie je vhodné pre charakterizáciu rozdelenia počtu poistných udalostí pre poistenie majetku. Keďže my očakávame nejaké nenahlásené poistné udalosti, odhad ich počtu založený na klasickom Poissonovom rozdelení nie je pre nás uspokojivý, pretože hovorí, že už žiadne nenahlásené poistné udalosti nie sú. Preto, na základe týchto výsledkov, tvrdíme, že zovšeobecnené Poissonovo rozdelenie je v našom prípade lepšie na charakteristiku rozdelenia počtu poistných udalostí. Potom rozdelenie prežívania je kvázi – binomické. Úlohou poistného matematika v tomto momente je rozhodnúť, pre ktorý zo zvyšných odhadov sa rozhodne. V našom prípade sú výsledky podobné, no môže nastať situácia, kedy sa tieto výsledky budú značne líšiť. Tu je z hľadiska praxe vhodné orientovať sa na základe vedomostí o vývoji z minulosti.

4.1.2 Poistenie zákonnej zodpovednosti z prevádzky motorového vozidla

V prípade poistenia akejkoľvek zodpovednosti je bežné, že sú poistné udalosti nahlásené aj niekoľko rokov po vzniku. Ako sme už spomínali, môže to byť napríklad v prípade škody na zdraví. V tomto prípade nie je stanovený presný časový interval, do kedy od okamihu vzniku je nutné takúto udalosť ohlásiť. Preto pri odhadovaní počtu nenahlásených poistných udalostí v tomto odvetví volím viac rokov do minulosti hodnotu indikátorovej premennej 1. Tabuľka 3 obsahuje vstupné údaje pre toto odvetvie.

Tabuľka 3: *Vstupné dáta – poistenie zmluvnej zodpovednosti z prevádzky motorového vozidla*

Rok	Počet PU	Stav kmeňa	Jednotky miery rizika
1991	36093	1844566	1,844566
1992	42825	2010371	2,010371
1993	54807	2034462	2,034462
1994	64348	1995357	1,995357
1995	68480	2003209	2,003209
1996	84215	2194779	2,194779
1997	106994	2607016	2,607016
1998	138091	2528047	2,528047
1999	153032	2650081	2,650081
2000	145177	2791316	2,791316
2001	157392	3038205	3,038205

Na základe týchto vstupných údajov získavame odhady zhrnuté v tabuľke 4.

Tabuľka 4: *Odhadnuté parametre a odhad počtu nenahlásených poistných udalostí – poistenie zmluvnej zodpovednosti z prevádzky motorového vozidla*

	φ	θ	λ	Odhad počtu nenahlásených poistných udalostí
Poissonovo rozdelenie bez miery rizika	1,000000	95 587,400	-	0,011
Zovšeobecnené Poissonovo rozdelenie bez miery rizika	0,998022	561,693	0,994152	2 089,657
Poissonovo rozdelenie s mierou rizika	1,000000	40 917,010	-	0,011
Zovšeobecnené Poissonovo rozdelenie s mierou rizika	0,985005	29,640	0,999489	22 327,990

Aj v tomto prípade vidíme, že obyčajné Poissonovo rozdelenie nie je vhodné na charakterizáciu rozdelenia počtu poistných rozdelení. My predpokladáme, že ešte sú nejaké poistné udalosti neohlásené, no odhad je rovný 0. Opäť je teda vhodné uvažovať zovšeobecnené Poissonovo rozdelenie s kvázi – binomickým rozdelením prežívania. V tabuľke 4 vidíme, že v tomto odvetví sa odhady počtu nenahlásených poistných udalostí značne líšia. Na základe historického vývoja sa rozhodneme pre niektorý z odhadov, alebo ak je zrejmé, že ani jeden z odhadov nezodpovedá historickému vývoju, zopakujeme proces odhadovania s pozmenenými vstupnými parametrami, napríklad zvolíme inú realizáciu indikátorovej premennej.

4.2 Výpočet IBNR rezerv

Aby sme mohli vypočítať odhad IBNR rezerv, potrebujeme mať dáta v tvare run-off triangle, podľa potreby buď hodnoty prírastkov, alebo kumulatívne hodnoty. Metódy aplikujeme na dáta nemeckej spoločnosti Münchener Rück [9], pretože sa nám nepodarilo získať potrebné údaje zo slovenského poistného trhu. Hodnoty sú v miliónoch Eur. Takto vyzerajú naše dáta:

Run-off 1: *Poistné plnenia*

	1	2	3	4	5	6	7	8	9	10	11
1998	7 161	4 557	2 628	1 696	1 512	937	1 259	569	622	633	779
1999	3 406	2 983	1 290	881	483	417	302	154	243	140	
2000	3 395	2 529	1 166	708	507	477	289	109	190		
2001	3 393	2 872	1 491	829	582	435	428	341			
2002	3 799	2 771	1 192	637	362	239	288				
2003	3 938	2 097	844	381	369	291					
2004	3 730	2 717	854	374	403						
2005	3 450	3 375	1 516	481							
2006	3 348	2 311	1 252								
2007	4 126	2 602									
2008	4 167										

Run-off 2: *Stav RBNS rezerv ku koncu daného obdobia*

	1	2	3	4	5	6	7	8	9	10	11
1998	17 184	12 420	8 901	7 315	7 133	6 294	5 491	6 422	6 191	6 208	5 738
1999	6 174	3 905	2 837	2 269	1 971	1 612	1 532	1 411	1 169	1 143	
2000	5 881	4 156	3 207	2 862	2 407	2 260	1 925	1 977	1 715		
2001	7 478	5 708	4 752	3 820	3 456	3 077	2 591	2 313			
2002	9 011	5 717	4 753	2 887	2 449	2 209	1 929				
2003	7 713	5 307	3 742	3 242	2 577	2 141					
2004	7 171	4 510	3 651	3 137	2 396						
2005	8 515	4 999	3 601	2 853							
2006	7 139	4 585	3 215								
2007	215	4 868									
2008	8 094										

Run-off 3: *Konečné výdavky na poistné udalosti*

	1	2	3	4	5	6	7	8	9	10	11
1998	24 345	24 138	23 247	23 357	24 687	24 785	25 241	26 741	27 132	27 782	28 091
1999	9 580	10 294	10 516	10 829	11 014	11 072	11 294	11 327	11 328	11 442	
2000	9 276	10 080	10 297	10 660	10 712	11 042	10 996	11 157	11 085		
2001	10 871	11 973	12 508	12 405	12 623	12 679	12 621	12 684			
2002	12 810	12 287	12 515	11 286	11 210	11 209	11 217				
2003	11 651	11 342	10 621	10 502	10 206	10 061					
2004	10 901	10 957	10 952	10 812	10 474						
2005	11 965	11 824	11 942	11 675							
2006	10 487	10 244	10 126								
2007	11 341	11 596									
2008	12 261										

Run-off 1 obsahuje poistné plnenia rozdelené podľa roku vzniku a podľa dátumu vyplatenia, Run-off 2 obsahuje stavy rezerv ku koncu jednotlivých období a Run-off 3 obsahuje celkové údaje o plneniach, resp. budúcich plneniach, o ktorých poisťovňa vie.

4.2.1 Mackova metóda chain ladder

Pri tejto metóde využijeme už naprogramovanú knižnicu „ChainLadder" pre programovací jazyk R, ktorá je voľne stiahnuteľná tu: http://cran.r-project.org/web/packages/ChainLadder/. Metóda je založená na kumulatívnych dátach, teda vstupné údaje sú tvaru Run-off 3. Pre porovnanie, aplikujme túto metódu aj na kumulatívne dáta, ktoré nemajú v sebe zahrnuté RBNS rezervy. Tieto dáta sú tvaru:

Run-off 4: *Kumulatívne dáta bez RBNS rezervy*

	1	2	3	4	5	6	7	8	9	10	11
1998	7 161	11 718	14 346	16 042	17 554	18 491	19 750	20 319	20 941	21 574	22 353
1999	3 406	6 389	7 679	8 560	9 043	9 460	9 762	9 916	10 159	10 299	
2000	3 395	5 924	7 090	7 798	8 305	8 782	9 071	9 180	9 370		
2001	3 393	6 265	7 756	8 585	9 167	9 602	10 030	10 371			
2002	3 799	6 570	7 762	8 399	8 761	9 000	9 288				
2003	3 938	6 035	6 879	7 260	7 629	7 920					
2004	3 730	6 447	7 301	7 675	8 078						
2005	3 450	6 825	8 341	8 822							
2006	3 348	5 659	6 911								
2007	4 126	6 728									
2008	4 167										

V nasledujúcej tabuľke si zhrnieme výsledky.

Tabuľka 5: *Odhady vývojových parametrov a IBNR rezerv, a ich štandardné chyby*

bez RBNS		s RBNS	
Parametre λ_j	**štandardné chyby**	**Parametre λ_j**	**štandardné chyby**
1,725	0,04072	1,012	0,01533
1,198	0,01232	0,996	0,01139
1,089	0,01004	0,990	0,01406
1,066	0,00739	1,012	0,01289
1,046	0,00430	1,005	0,00528
1,046	0,00800	1,008	0,00548
1,024	0,00484	1,029	0,01500
1,027	0,00296	1,007	0,00645
1,025	0,00771	1,020	0,00633
1,036	0,00387	1,011	0,00663
IBNR	**štandardná chyba**	**IBNR**	**štandardná chyba**
24 417	1 908	7 384	2 903

V tabuľke 5 vidíme odhady vývojových parametrov $\hat{\lambda}_j$ a ich štandardné chyby. Na základe týchto odhadov sme potom odhadli výšku IBNR rezerv. Všimnime si, že v prípade, ak uvažujeme celkové plnenia aj s RBNS rezervami, odhad IBNR rezerv je nižší. S pohľadu poisťovne je správnejšie počítať IBNR rezervy na základe plnení s RBNS rezervami, pretože RBNS rezervy sú finančné prostriedky určené na výplatu poistných plnení z poistných udalostí, o ktoré sú nahlásené, ale ešte nie sú (celkom) vyplatené.

Tabuľka 6: *Run-off triangle doplnený na štvorec – bez RBNS (kumulatív)*

	1	2	3	4	5	6	7	8	9	10	11
1998	7 161	11 718	14 346	16 042	17 554	18 491	19 750	20 319	20 941	21 574	22 353
1999	3 406	6 389	7 679	8 560	9 043	9 460	9 762	9 916	10 159	10 299	10 671
2000	3 395	5 924	7 090	7 798	8 305	8 782	9 071	9 180	9 370	9 603	9 950
2001	3 393	6 265	7 756	8 585	9 167	9 602	10 030	10 371	10 649	10 913	11 307
2002	3 799	6 570	7 762	8 399	8 761	9 000	9 288	9 512	9 767	10 009	10 371
2003	3 938	6 035	6 879	7 260	7 629	7 920	8 287	8 487	8 714	8 931	9 253
2004	3 730	6 447	7 301	7 675	8 078	8 452	8 843	9 057	9 299	9 530	9 875
2005	3 450	6 825	8 341	8 822	9 401	9 835	10 291	10 540	10 822	11 091	11 491
2006	3 348	5 659	6 911	7 527	8 021	8 392	8 781	8 993	9 233	9 463	9 805
2007	4 126	6 728	8 059	8 778	9 353	9 786	10 240	10 487	10 767	11 035	11 433
2008	4 167	7 188	8 610	9 378	9 993	10 455	10 939	11 203	11 503	11 789	12 215

Tabuľka 7: *Run-off triangle doplnený na štvorec – s RBNS (kumulatív)*

	1	2	3	4	5	6	7	8	9	10	11
1998	24 345	24 138	23 247	23 357	24 687	24 785	25 241	26 741	27 132	27 782	28 091
1999	9 580	10 294	10 516	10 829	11 014	11 072	11 294	11 327	11 328	11 442	11 569
2000	9 276	10 080	10 297	10 660	10 712	11 042	10 996	11 157	11 085	11 305	11 431
2001	10 871	11 973	12 508	12 405	12 623	12 679	12 621	12 684	12 766	13 020	13 165
2002	12 810	12 287	12 515	11 286	11 210	11 209	11 217	11 545	11 620	11 851	11 982
2003	11 651	11 342	10 621	10 502	10 206	10 061	10 144	10 440	10 508	10 717	10 836
2004	10 901	10 957	10 952	10 812	10 474	10 526	10 612	10 922	10 993	11 211	11 336
2005	11 965	11 824	11 942	11 675	11 815	11 873	11 970	12 320	12 400	12 647	12 787
2006	10 487	10 244	10 126	10 020	10 140	10 190	10 274	10 574	10 643	10 854	10 975
2007	11 341	11 596	11 553	11 433	11 570	11 626	11 722	12 064	12 143	12 384	12 522
2008	12 261	12 411	12 366	12 236	12 383	12 444	12 546	12 912	12 996	13 255	13 402

4.2.2 Kaasova metóda chain ladder

Odhad IBNR rezerv touto metódou je založený na spomínanej vlastnosti 1, ktorá predpokladá, že výška poistných plnení má Poissonovo rozdelenie. Tu ale môže nastať situácia, že ak uvažujeme plnenia s RBNR rezervami, prírastky môžu byť aj záporné, čo nezodpovedá Poissonovmu rozdeleniu, ktoré nadobúda len kladné hodnoty. Aj napriek tomuto porušeniu predpokladu vlastnosti 1, aplikujeme túto metódu aj na tieto prírastky. Vzhľadom na to, že táto metóda odhaduje veľa parametrov, uvádzame len výsledky – run-off triangle doplnený na štvorec a odhad IBNR rezervy.

Tabuľka 8: *Run-off triangle doplnený na štvorec – bez RBNS (prírastky)*

	1	2	3	4	5	6	7	8	9	10	11
1998	7 161	4 557	2 628	1 696	1 512	937	1 259	569	622	633	779
1999	3 406	2 983	1 290	881	483	417	302	154	243	140	372
2000	3 395	2 529	1 166	708	507	477	289	109	190	233	347
2001	3 393	2 872	1 491	829	582	435	428	341	278	265	394
2002	3 799	2 771	1 192	637	362	239	288	224	255	243	361
2003	3 938	2 097	844	381	369	291	367	200	227	217	323
2004	3 730	2 717	854	374	403	374	392	213	242	231	344
2005	3 450	3 375	1 516	481	579	435	456	248	282	269	401
2006	3 348	2 311	1 252	616	494	371	389	212	241	230	342
2007	4 126	2 602	1 331	719	576	433	454	247	281	268	399
2008	4 167	3 021	1 422	768	615	462	485	264	300	286	426

Celková IBNR rezerva, odhadnutá touto metódou aplikovanou na dáta v Run-off 1 je 24 417 miliónov Eur.

Tabuľka 9: *Run-off triangle doplnený na štvorec – s RBNS (prírastky)*

	1	2	3	4	5	6	7	8	9	10	11
1998	24 345	-207	-891	110	1 330	98	456	1 500	391	650	309
1999	9 580	714	222	313	185	58	222	33	1	114	127
2000	9 276	804	217	363	52	330	-46	161	-72	220	126
2001	10 871	1 102	535	-103	218	56	-58	63	83	254	145
2002	12 810	-523	228	-1 229	-76	-1	8	328	75	231	132
2003	11 651	-309	-721	-119	-296	-145	83	296	68	209	119
2004	10 901	56	-5	-140	-338	52	87	310	71	218	125
2005	11 965	-141	118	-267	140	58	98	350	80	246	141
2006	10 487	-243	-118	-106	120	50	84	300	69	211	121
2007	11 341	255	-43	-121	137	57	96	342	78	241	138
2008	12 261	150	-46	-129	146	61	102	367	84	258	147

Celková rezerva odhadnutá touto metódou, ktorú sme aplikovali na prírastky poistných plnení po započítaní RBNS rezerv je 7 384 miliónov Eur.

Všimnime si, že ak sme na tie isté dáta aplikovali Mackovu chain ladder aj Kaasovu chain ladder, získali sme rovnaký odhad IBNR rezerv. Je to v súlade s tvrdeniami článku [1], kde sa tieto metódy líšia strednou kvadratickou chybou.

4.2.3 Aritmetická separačná metóda

V tejto metóde opäť postupujeme rekurentne. Parametre, ktoré vieme odhadnúť z dát, odhadneme metódou marginálnych súčtov. Potrebujeme odhadnúť parametre $\gamma_{I+1},\ldots,\gamma_{2I}$. Na to použijeme funkciu *approxExtrap (knižnica Hmisc)*, ktorá na základe ostatných dvoch členov odhadne lineárnou extrapoláciou nasledujúce parametre. Extrapolujeme funkciou $\gamma_i = a\ln i + b$. Z dôvodu veľkého množstva odhadovaných parametrov uvádzame len výsledky, teda doplnený run-off triangle na štvorec a celkový odhad IBNR rezerv. Metóda sa aplikuje na dáta vo forme prírastkov.

Tabuľka 10: *Run-off triangle doplnený na štvorec – bez RBNS (prírastky)*

	1	2	3	4	5	6	7	8	9	10	11
1998	7 161	4 557	2 628	1 696	1 512	937	1 259	569	622	633	779
1999	3 406	2 983	1 290	881	483	417	302	154	243	140	766
2000	3 395	2 529	1 166	708	507	477	289	109	190	376	753
2001	3 393	2 872	1 491	829	582	435	428	341	343	370	742
2002	3 799	2 771	1 192	637	362	239	288	285	337	365	731
2003	3 938	2 097	844	381	369	291	491	281	332	359	721
2004	3 730	2 717	854	374	403	434	483	276	327	354	711
2005	3 450	3 375	1 516	481	547	427	475	272	323	350	703
2006	3 348	2 311	1 252	672	538	420	469	269	319	345	694
2007	4 126	2 602	1 197	661	529	414	462	265	315	341	686
2008	4 167	2 476	1 177	651	522	408	456	262	311	337	679

Odhad celkovej IBNR v tomto prípade je 28 807 miliónov Eur.

Tabuľka 11: *Run-off triangle doplnený na štvorec – s RBNS (prírastky)*

	1	2	3	4	5	6	7	8	9	10	11
1998	24 345	-207	-891	110	1 330	98	456	1 500	391	650	309
1999	9 580	714	222	313	185	58	222	33	1	114	315
2000	9 276	804	217	363	52	330	-46	161	-72	393	320
2001	10 871	1 102	535	-103	218	56	-58	63	112	399	324
2002	12 810	-523	228	-1 229	-76	-1	8	446	114	405	329
2003	11 651	-309	-721	-119	-296	-145	117	454	115	410	333
2004	10 901	56	-5	-140	-338	65	119	460	117	415	336
2005	11 965	-141	118	-267	143	66	121	466	118	420	340
2006	10 487	-243	-118	-124	145	67	123	472	120	424	343
2007	11 341	255	-44	-126	147	68	124	477	121	429	347
2008	12 261	145	-44	-128	149	69	126	483	122	433	350

Odhad celkovej IBNR v tomto prípade je 12 587 miliónov Eur.

4.2.4 Geometrická separačná metóda

Opäť postupujeme rekurentne. Parametre odhadujeme metódou marginálnych súčinov, ktorá funguje na podobnom princípe ako metóda marginálnych súčtov. O rozdelení dát predpokladáme, že je logaritmicko - normálne. Pri extrapolovaní postupujeme ako pri aritmetickej separačnej metóde. Problém nastáva, ak chceme odmocniť záporné číslo, preto túto metódu môžeme pri našich aplikovať len na prírastky poistných plnení bez RBNS rezerv. Opäť odhady parametrov neuvádzame.

Tabuľka 12: *Run-off triangle doplnený na štvorec – bez RBNS (prírastky)*

	1	2	3	4	5	6	7	8	9	10	11
1998	7 161	4 557	2 628	1 696	1 512	937	1 259	569	622	633	779
1999	3 406	2 983	1 290	881	483	417	302	154	243	140	740
2000	3 395	2 529	1 166	708	507	477	289	109	190	275	704
2001	3 393	2 872	1 491	829	582	435	428	341	277	262	670
2002	3 799	2 771	1 192	637	362	239	288	203	263	249	639
2003	3 938	2 097	844	381	369	291	339	194	251	238	610
2004	3 730	2 717	854	374	403	321	322	184	239	227	583
2005	3 450	3 375	1 516	481	383	306	307	176	228	217	557
2006	3 348	2 311	1 252	466	364	291	293	168	218	207	533
2007	4 126	2 602	873	443	347	278	279	160	209	198	510
2008	4 167	1 871	831	422	331	265	267	153	199	190	488

Celkový odhad IBNR rezervy je 20 815 miliónov Eur.

4.2.5 Porovnanie výsledkov

Tabuľka 13 obsahuje odhady IBNR rezerv, ktoré sme získali viacerými metódami. Pri aritmetickej a geometrickej metóde si môžeme zvoliť, akú extrapoláciu použijeme a tak pre porovnanie uvádzame výsledky aj ďalšie dva spôsoby extrapolácie.

Tabuľka 13: *Odhad IBNR rezerv*

Metóda	odhad IBNR	
	len plnenia	plnenia s RBNS
Mackov model chain ladder	24 417	7 384
Kaasov model chain ladder	24 417	7 384
Aritmetická separačná metóda, parametre $\gamma_{I+1},\dots,\gamma_{2I}$ *extrapolujeme funkciou:*		
a) $\gamma_i = a\ln i + b$	28 807	12 587
b) $\gamma_i = ai + b$	28 334	12 834
c) $\ln\gamma_i = ai + b$	28 460	12 911
Geometrická separačná metóda, parametre $\gamma_{I+1},\dots,\gamma_{2I}$ *extrapolujeme funkciou:*		
a) $\gamma_i = a\ln i + b$	20 815	-
b) $\gamma_i = ai + b$	19 624	-
c) $\ln\gamma_i = ai + b$	20 467	-

Vo vyššie uvedenej tabuľke vidíme, že použitím Mackovej metódy chain ladder a Kaasovovej metódy chain ladder získame totožné odhady IBNR rezerv. Kaasov model predpokladá Poissonovo rozdelenie pre dáta, teda vstupné dáta – prírastky - očakávame kladné. Presvedčili sme sa však, že aj keď prírastky nie sú kladné, prípad, keď uvažujeme aj RBNS rezervy, môžeme túto metódu použiť. To isté platí aj pre aritmetickú separačnú metódu, bez ohľadu na výber extrapolačnej funkcie. Geometrická separačná metóda predpokladá logaritmicko - normálne rozdelenie prírastkov, teda tiež predpokladá kladné hodnoty prírastkov. Táto metóda pri záporných prírastkoch zlyhala. Je však možné, že ak vhodne upravíme program, tak sa bude dať táto metóda aplikovať aj na záporné prírastky.

Ak sa pozrieme na konkrétne výsledky, môžeme tvrdiť, že ani pri aritmetickej ani pri geometrickej separačnej metóde nezáleží na výbere extrapolačnej funkcie. Výsledky sa od seba veľmi nelíšia.

Ďalej si môžeme všimnúť, že odhady IBNR rezerv sú v prípade, že uvažujeme aj RBNS rezervy, nižšie.

Záver

V práci sme sa zaoberali odhadom počtu nenahlásených poistných udalostí a výpočtom IBNR rezerv v neživotnom poistení. Na odhad počtu nenahlásených poistných udalostí sme predstavili metódu navrhnutú D. Scollnikom. Na odhad IBNR rezerv sme predstavili viacero metód. K výberu metód sme pristupovali aplikačne, teda v prípade viacerých alternatív jednej metódy sme vybrali tú, ktorú jednoduchšie aplikujeme do praxe. Pre najpoužívanejšiu metódu chain ladder, Mackov model, sme dokázali tvrdenie o výpočte strednej kvadratickej chyby odhadu IBNR rezerv.

Cieľom tejto práce bolo teoreticky opísané metódy aplikovať na reálne dáta. V prípade odhadovania IBNR metód sme potrebovali dáta v špeciálnom tvare, v tzv. run-off triangle. Nebolo jednoduché takéto dáta nájsť, pretože žiadna zo slovenských poisťovní nemá tieto dáta voľne prístupné a nechceli sme použiť dáta z odborných článkov, pretože by sme nezískali žiadne nové hodnoty.

V práci sme používali programovací jazyk R, ktorý je voľne stiahnuteľný z internetu, teda nie je nutné vlastniť licenciu na jeho používanie aj v poisťovniach.

Dovolím si tvrdiť, že sa nám úspešne podarilo naprogramovať opísané metódy. V prípade metódy chain ladder sme použili už naprogramovanú knižnicu „ChainLadder“, ktorá je tiež voľne stiahnuteľná z internetu.

Metódu na odhad počtu nenahlásených poistných udalostí sme použili na dáta z dvoch poistných odvetví – poistenie majetku a poistenie zmluvnej zodpovednosti z prevádzky motorového vozidle, dnešné povinné zmluvné poistenie motorového vozidla. V obidvoch odvetviach sme rozumné odhady dostali vtedy, keď sme predpokladali, že celkový počet poistných udalostí má zovšeobecnené Poissonovo rozdelenie s kvázi – binomickým rozdelením prežívania.

V prípade odhadovania IBNR rezerv sme spomenuté metódy aplikovali na jedny dáta. Výsledky jednotlivých metód môžeme vidieť v tabuľke 13. Vidíme, že Mackova a Kaasova metóda chain ladder dávajú totožné výsledky. Aritmetickú separačnú metódu sme použili aj na dáta, ktoré nezodpovedajú Poissonovmu rozdeleniu (záporné hodnoty). Geometrická separačná metóda pri aplikácii na dáta obsahujúce záporné hodnoty zlyhala. Stalo sa to, že sme dostali záporné číslo pod odmocninou, čo aj v prípade zrátateľnosti, programovací jazyk R nedokáže vyrátať.

Literatúra

[1] England, P. D. - Verral, R. J.: Stochastic claims reserving in general insurance. Presented to the Institute of Actuaries, 28 January 2002.

[2] Kaas, R.: Modern Actuarial Risk Theory. Hingham, USA: Kluwer Academic Publisher, 2001, 9. kapitola, pp. 203 – 220.

[3] Mack, T.: Distribution-free calculation of the standard error of chain ladder reserve estimates. In: Astin Bulletin, 1993, Vol. 23, No. 2, pp. 213 – 225.

[4] Nandi, S. B., Das, K. K.: A family of the Abel series distribution. In: Sankhya, 1994, Vol. 56 Series B, pp. 147 – 164.

[5] Pacáková, V.: Aplikovaná poistná štatistika. Bratislava: IURA EDITION 2004.

[6] Scollnik, D. P. M: A damaged generalised Poisson model and its application to reported and unreported accicent counts. In: Astin Bulletin, 2006, Vol. 36, No. 2, pp. 463 – 487.

[7] Taylor, G.: Chain ladder bias. In: Astin Bulletin, 2003, Vol. 33, No. 2, pp. 313 – 330.

[8] Slovenská asociásia poisťovní: Štatistická ročenka o vývoji na poistnom trhu v SR 1991 – 2001. Bratislava, jún 2002. www.slaspo.sk.

[9] Münchener Rück, Munich Re Group: Annual Report 2008. http://report.munichre.com/reports/munichre/annual/2008/gb/English/603530/_21_-provision-for-outstanding-claims.html

Prílohy

```
library(stats4)
likelihood<-function(par,y,d,w)
{
phi<-par[1]
theta<-par[2]
#w je vektor mier rizika
#y je vektor reported claims
#d je vektor nul a jednotiek, ci su unreported claims alebo nie
#theta a phi su parametre rozdelenia

pozorovani<-length(y)
suma1<-0
suma2<-0
suma3<-0
suma4<-0
for (i in 1:pozorovani) suma1<-suma1+d[i]*log(1-exp(-w[i]*(1-phi)*theta))
for (j in 1:pozorovani) suma2<-suma2+(1-d[j])*w[j]
for (k in 1:pozorovani) suma3<-suma3+y[k]*log(w[k])
for (l in 1:pozorovani) suma4<-suma4+lfactorial(y[l])
L<- (-phi*theta*pozorovani*mean(w)
   + pozorovani*mean(y)*log(phi*theta)
   + suma1
   -(1-phi)*theta*suma2
   + suma3
   - suma4)
return(L)
}



likelihoodGPD<-function(par, y, d, w)
{
phi<-par[1]
theta<-par[2]
lambda<-par[3]
suma1<-0
suma2<-0
suma3<-0
suma4<-0
suma5<-0
pozorovani<-length(y)
for (i in 1:pozorovani){
		suma1<-suma1+(y[i]-1)*log(w[i]*phi*theta + y[i]*lambda)
		suma2<-suma2+d[i]*log(1-exp(-w[i]*(1-phi)*theta))
		suma3<-suma3+(1-d[i])*w[i]
		suma4<-suma4+log(w[i])
		suma5<-suma5+lfactorial(y[i])
		}
L_GPD<- (-phi*theta*pozorovani*mean(w)
	-lambda*pozorovani*mean(y)
	+pozorovani*log(phi*theta)
	+suma1
	+suma2
	-(1-phi)*theta*suma3
	+suma4
	-suma5)
return(L_GPD)
}
```

```
KaasChainLadder<-function(Triangle)
{
I<-length(Triangle[1,])
K<-c()
R<-c()

alfa<-rep(0,I)
beta<-rep(0,I)

Rezervy<-rep(0,I)

FullTriangle<-Triangle
for(i in 1:I)
        {
        R[i]<-sum(Triangle[i,1:(I-i+1)])
        K[i]<-sum(Triangle[1:(I-i+1),i])
         }

alfa[1]<-R[1]
beta[I]<-K[I]/alfa[1]

for(i in 2:I)
        {
        alfa[i]<-R[i]/(1-sum(beta[(I-i+2):I]))
        beta[I-i+1]<-K[I-i+1]/sum(alfa[1:i])
        }
for(i in 2:I)
   for(j in (I+2-i):I)
        {
        FullTriangle[i,j]<-alfa[i]*beta[j]
        }

for(i in 1:I) Rezervy[i]<-sum(FullTriangle[i,])-R[i]

beta<-round(beta, digits=4)
FullTriangle<-round(FullTriangle, digits=1)
Rezervy<-round(Rezervy, digits=3)
Celkova.Rezerva<-sum(Rezervy)
output<-list()

output[["Triangle"]] <- Triangle
output[["alfa"]] <- alfa
output[["beta"]] <- beta
output[["FullTriangle"]] <- FullTriangle
output[["Rezervy.podla.rokov.vzniku"]] <- Rezervy
output[["Celkova.rezerva"]] <- Celkova.Rezerva

return(output)
}
```

```
library(Hmisc)

KaasAritmeticSeparation<-function(Triangle)
{
I<-length(Triangle[1,])
K<-c()
D<-c()
R<-c()

beta<-rep(0,I)
gama<-rep(0,2*I)

Rezervy<-rep(0,I)

FullTriangle<-Triangle
for(i in 1:I)
        {
        R[i]<-sum(Triangle[i,1:(I-i+1)])
        K[i]<-sum(Triangle[1:(I-i+1),i])
        D[i]<-0
        for(j in 1:i) D[i]<-D[i]+Triangle[j,(i-j+1)]
         }

gama[I]<-D[I]
beta[I]<-K[I]/gama[I]

for(i in 1:(I-1))
        {
        gama[I-i]<-D[I-i]/(1-sum(beta[(I-i+1):I]))
        beta[I-i]<-K[I-i]/sum(gama[(I-i):I])
        }

aprox<-approxExtrap(log(seq(1,I,by=1)),gama,log(seq(I+1,2*I,by=1)))$y
for (i in 1:I) gama[I+i]<-aprox[i]

for(i in 2:I)
  for(j in (I+2-i):I)
        {
        FullTriangle[i,j]<-gama[i+j-1]*beta[j]
        }

for(i in 1:I) Rezervy[i]<-sum(FullTriangle[i,])-R[i]

beta<-round(beta, digits=4)
FullTriangle<-round(FullTriangle, digits=1)
Rezervy<-round(Rezervy, digits=3)
Celkova.Rezerva<-sum(Rezervy)
output<-list()

output[["Triangle"]] <- Triangle
output[["beta"]] <- beta
output[["gama"]] <- gama
output[["FullTriangle"]] <- FullTriangle
output[["Rezervy.podla.rokov.vzniku"]] <- Rezervy
output[["Celkova.rezerva"]] <- Celkova.Rezerva

return(output)
}
```

```
library(Hmisc)

KaasGeometricSeparation.B<-function(Triangle)
{
I<-length(Triangle[1,])
R<-c()

SK<-rep(1,I)
SG<-rep(1,I)

beta<-rep(0,I)
gama<-rep(0,2*I)

Rezervy<-rep(0,I)

FullTriangle<-Triangle
for(i in 1:I)
        {
        R[i]<-sum(Triangle[i,1:(I-i+1)])
        SK[i]<-prod(Triangle[1:(I-i+1),i])
        for(j in 1:i) SG[i]<-SG[i]*Triangle[j,(i-j+1)]
        }

gama[I]<-(SG[I])^(1/I)
beta[I]<-(SK[I])/gama[I]

for(i in 1:(I-1))
        {
        gama[I-i]<-(SG[I-i]*prod(beta[(I-i+1):I]))^(1/(I-i))
        beta[I-i]<-(SK[I-i]/prod(gama[(I-i):I]))^(1/(i+1))
        }

aprox<-approxExtrap(log(seq(1,I,by=1)),gama,log(seq(I+1,2*I,by=1)))$y

for (i in 1:I) gama[I+i]<-aprox[i]

for(i in 2:I)
  for(j in (I+2-i):I)
        {
        FullTriangle[i,j]<-gama[i+j-1]*beta[j]
        }

for(i in 1:I) Rezervy[i]<-sum(FullTriangle[i,])-R[i]

beta<-round(beta, digits=4)
FullTriangle<-round(FullTriangle, digits=1)
Rezervy<-round(Rezervy, digits=3)
Celkova.Rezerva<-sum(Rezervy)
output<-list()

output[["Triangle"]] <- Triangle
output[["beta"]] <- beta
output[["gama"]] <- gama
output[["FullTriangle"]] <- FullTriangle
output[["Rezervy.podla.rokov.vzniku"]] <- Rezervy
output[["Celkova.rezerva"]] <- Celkova.Rezerva

return(output)
}
```