

Maria Jadamus-Hacura

Akademia Ekonomiczna w Katowicach

METODY BAYESOWSKIE SZACOWANIA WARTOŚCI REZERWY SZKODOWEJ

1. Wstęp

W niniejszym opracowaniu przedstawiono bayesowski model szacowania rezerwy szkodowej. Punktem wyjścia budowy tego modelu jest powszechnie stosowana przez aktuariuszy prosta metoda kalkulacji poziomu rezerwy szkodowej, tzw. metoda *chain ladder*. Algorytm tej metody przedstawiono w drugiej części pracy. W części następnej opisano metodę bayesowską szacowania wartości rezerwy szkodowej. Najistotniejszym punktem tej metody jest założenie, że współczynniki rozwoju szkodowości nie są stałymi, ale zmiennymi losowymi o rozkładzie normalnym. Dla parametrów tego rozkładu przyjęto z kolei hierarchiczny rozkład *a priori*. W części ostatniej opracowania przedstawiono przykład zaczerpnięty z pracy [3] i porównano wyniki wartości rezerwy oszacowane klasyczną metodą *chain ladder* i metodą bayesowską. Obliczenia zostały przeprowadzone za pomocą programu WINBUGS – specjalistycznego programu pozwalającego na wykorzystanie metod symulacji MCMC (*Markov chain Monte Carlo*) do oszacowania parametrów modelu bayesowskiego.

2. Metoda *chain ladder* kalkulacji rezerwy szkodowej

Oszacowania całkowitej wartości wypłaconych odszkodowań w przyjętym przedziale czasowym dokonuje się w praktyce na podstawie danych zapisanych w postaci macierzy, zwanej trójkątem szkód. W trójkącie tym wyróżnia się następujące wielkości:

- i* – rok wystąpienia szkody (okres wypadkowy),
- j* – opóźnienie w wypłacie odszkodowania,

$Y[i, j]$ – wartość odszkodowań dla szkód, które zaszły w okresie wypadkowym i , a zostały wypłacone z opóźnieniem j ,

$C[i, j]$ – skumulowana wartość odszkodowań dla szkód, które zaszły w okresie wypadkowym i , a zostały wypłacone z opóźnieniem j ,

$$C[i, j] = \sum_{k=1}^j Y[i, k] \quad (1)$$

dla $i = 1, \dots, r$ oraz $j = 1, \dots, r - i + 1$.

W metodzie *chain ladder* podstawą obliczenia rezerw szkodowych są następujące współczynniki:

- indywidualny współczynnik rozwoju szkody (*single or individual development factor*)

$$f[i, j] = \frac{C[i, j+1]}{C[i, j]}, \quad (2)$$

- współczynnik rozwoju szkody (*volume weighted development factors*)

$$\hat{f}[i, j] = \frac{\sum_{i=1}^{r-j} C[i, j+1]}{\sum_{i=1}^{r-j} C[i, j]}. \quad (3)$$

Metodę *chain ladder* stosuje się, jeżeli współczynniki $\hat{f}[j]$ są w przybliżeniu takie same dla wszystkich analizowanych okresów wypadkowych. Często przyjmuje się oszacowanie współczynników rozwoju szkody $\hat{f}[j]$ na poziomie średniej arytmetycznej lub geometrycznej z wartości $f[i, j]$.

Oszacowania całkowitej wartości wypłaconych odszkodowań dla okresu wypadkowego i dokonuje się na podstawie wzoru rekurencyjnego:

$$\hat{C}[i, j] = \hat{C}[i, j-1] \cdot \hat{f}[j], \quad (4)$$

gdzie $\hat{C}[i, n+i-1] = C[i, n+i-1]$.

Ostatecznie całkowita wartość wypłaconych odszkodowań dla okresu wypadkowego i -tego wynosi:

$$\hat{C}[i] = C[i, n-i-1] \cdot \hat{f}[0] \dots \hat{f}[n-i-1], \quad (5)$$

a wartość rezerwy szkodowej dla tego okresu jest równa różnicy:

$$\hat{C}[i] - C[i, n-i-1].$$

Przykładowy trójkąt szkód (zawierający wartości skumulowane $C[i, j]$), zaczerpnięty z cytowanej już pracy [3] przedstawiono w tab. 1.

Tabela 1. Przykład trójkąta szkód zawierającego wartości skumulowane odszkodowań

Okres wypadkowy	Opóźnienie w wypłacie odszkodowania					
	1	2	3	4	5	6
1991	115 239	171 294	185 985	196 240	201 770	204 872
1992	121 528	179 439	194 660	203 999	209 335	
1993	115 427	169 384	181 882	190 425		
1994	113 008	159 674	170 724			
1995	109 881	163 289				
1996	128 982					

Źródło: [3, s. 34].

Wartości współczynników $f[i, j]$ i oszacowań wartości $\hat{f}[j]$ zgodnie ze wzorem (3) podano z kolei w tab. 2.

Tabela 2. Indywidualne współczynniki rozwoju $f[i, j]$ i oszacowane wartości $\hat{f}[j]$

Rok	2	3	4	5	6
1991	1,4864	1,0858	1,0551	1,0282	1,0154
1992	1,4765	1,0848	1,0480	1,0262	
1993	1,4675	1,0738	1,0470		
1994	1,4129	1,0692			
1995	1,4861				
$\hat{f}[j]$	1,4660	1,0786	1,0500	1,0271	1,0154

Źródło: obliczenia własne.

Oszacowania całkowitej wartości wypłaconych odszkodowań dla kolejnych okresów wypadkowych dokonano na podstawie wzoru rekurencyjnego (4) i zamieszczono je w tab. 3.

Tabela 3. Oszacowania całkowitej wartości wypłaconych odszkodowań dla okresów wypadkowych 1991-1996

Okres wypadkowy	Opóźnienie w wypłacie odszkodowania					
	1	2	3	4	5	6
1991	115 239	171 294	185 985	196 240	201 770	204 872
1992	121 528	179 439	194 660	203 999	209 335	212 553,3
1993	115 427	169 384	181 882	190 425	195 594,8	198 601,9
1994	113 008	159 674	170 724	179 263,4	184 130,2	186 961
1995	109 881	163 289	176 130,3	184 940,2	189 961,1	192 881,5
1996	128 982	189 089,5	203 959,8	214 161,7	219 975,9	223 357,8

Źródło: obliczenia własne.

Tabela 4. Wartości rezerwy szkodowej w latach 1992-1996

1992	1993	1994	1995	1996
3 218,30	8 176,87	16 237,02	29 078,37	94 375,80

Źródło: obliczenia własne.

Na podstawie wyników podanych w tab. 3 obliczono wartości rezerw szkodowych dla poszczególnych lat w okresie 1991-1996. Wartości tych rezerw, których oszacowanie jest celem przedstawionej metody *chain ladder*, zawiera tab. 4.

3. Model bayesowski szacowania wartości rezerwy szkodowej

Analiza indywidualnych współczynników rozwoju szkody zamieszczonych w tab. 2, pozwala na stwierdzenie, że współczynniki te dla danego opóźnienia j mają zbliżone wartości. W literaturze najczęściej przyjmuje się założenie, że są one zmiennymi losowymi o rozkładzie normalnym:

$$f[i, j] \sim N(\theta_j, \tau_{i,j})$$

dla $i = 1, \dots, r$ oraz $j = 1, \dots, r-1$.

Parametr $\tau_{i,j}$ oznacza precyzję (odwrotność wariancji) rozkładu normalnego i może być modelowany na wiele sposobów. Najprostsze jest założenie, że parametr ten jest taki sam dla wszystkich wartości i, j , czyli $\tau_{ij} = \tau$. W literaturze przedmiotu można również spotkać propozycję przyjęcia założenia o proporcjonalności precyzji τ_{ij} do wielkości składki ubezpieczeniowej w danym okresie lub liczby szkód.

W bayesowskich modelach statystyki ubezpieczeniowej przyjmuje się najczęściej hierarchiczne rozkłady *a priori*. Jest to pomysł wynikający z potrzeby określenia hiperparametrów. Na parametry rozkładu *a priori* nakłada się więc kolejny rozkład *a priori*, a powstała w ten sposób hierarchia liczy zwykle 2 lub 3 szczeble, przy czym na ostatnim z nich pojawia się zwykle rozkład nieinformacyjny lub rozproszony (*diffused prior*); może nim być np. rozkład gamma $G(0,001; 0,001)$.

W pracy [3] przyjęto następujący hierarchiczny rozkład *a priori*:

$$\theta_j \sim N(\mu_0, \tau_0),$$

gdzie: $\tau \sim$ rozkład gamma $G(0,001; 0,001)$,

$\mu_0 \sim$ rozkład normalny $N(1; 0,001)$,

$\tau_0 \sim$ rozkład gamma $G(0,001; 0,001)$.

Rozkłady parametrów τ , τ_0 i μ_0 są rozkładami rozproszonymi ze średnią 1 i wariancją równą 1000. Powyższe założenia zapisane w języku programu WINBUGS przyjmują następującą postać (por. [3, s. 11-12]):

```

# Define the development factors for the observed data.
for( i in 1 : ( r - 1 ) ) {
  for( j in 1 : ( r - i ) ) {
    devfac[ i, j ] <- z[ i, j + 1 ] / z[ i, j ]
  }
}
# Define the stochastic model for future cumulative claims.
for( i in 2 : r ) {
  for( j in ( r - i + 1 ) : ( r - 1 ) ) {
    z[ i, j + 1 ] <- z[ i, j ] * devfac[ i, j ]
    y[ i, j + 1 ] <- z[ i, j + 1 ] - z[ i, j ]
  }
}
# Define a distribution on the development factors devfac.
for( i in 1 : r ) {
  for( j in 1 : ( r - 1 ) ) {
    # Uncomment / modify the I( 0.95, ) terms as desired.
    devfac[ i, j ] ~ dnorm( mu.df[ i, j ], tau.df[ i, j ] ) # I( 0.95, )
    mu.df[ i, j ] <- theta[ j ]
    tau.df[ i, j ] <- tau * weight[ i, j ]
    weight[ i, j ] <- 1 # Or whatever is appropriate.
  }
}
# Code for unrestricted theta[ j ] nodes.
for( j in 1 : ( r - 1 ) ) {
  theta[ j ] ~ dnorm( mu.theta, tau.theta )
}
# Illustrative code for order restricted (decreasing) theta[ j ] nodes.
# theta[ 1 ] ~ dnorm( mu.theta, tau.theta ) I( theta[ 2 ], )
# for( j in 2 : ( r - 2 ) ) {
#   theta[ j ] ~
#   dnorm( mu.theta, tau.theta ) I( theta[ j + 1 ], theta[ j - 1 ] )
# }
# theta[ r - 1 ] ~ dnorm( mu.theta, tau.theta ) I( 0, theta[ r - 2 ] )
tau ~ dgamma( 0.001, 0.001 )
mu.theta ~ dnorm( 1, 0.001 )
tau.theta ~ dgamma( 0.001, 0.001 )

```

Program instalacyjny programu WINBUGS można znaleźć na stronie internetowej: <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>. Program ten pozwala na wykorzystanie metod symulacji komputerowej MCMC, niezbędnych do otrzymania rozkładów *a posteriori* parametrów opisanego modelu bayesowskiego.

W tabeli 5 przedstawiono otrzymane wyniki, tzn. podstawowe statystyki rozkładu predyktywnego rezerw szkodowych, dla poszczególnych lat: wartości średnie, odchylenia standardowe i 95% przedziały ufności dla wyznaczonych średnich. W ostatniej kolumnie tej tabeli pokazano również – w celu porównania – wartości rezerwy szkodowej obliczonej klasyczną metodą *chain ladder*.

Tabela 5. Wyniki estymacji parametrów rozkładu predyktywnego rezerw szkodowych

Okres wypadkowy	Wartość średnia (odchylenie standardowe) 95% przedział ufności dla średniej	Wartości rezerwy szkodowej obliczone metodą <i>chain ladder</i>
1992	3,682 (8,050) (-12,260, 19,880)	3 218,30
1993	8,780 (9,964) (-10,610, 29,100)	8 176,87
1994	16,950 (10,930) (-4,100, 39,440)	16 237,02
1995	30,300 (12,560) (6,239, 56,470)	29 078,37
1996	94,970 (15,240) (66,050, 126,700)	94 375,80

Źródło: [3, s. 38].

Jak widać, wartości średnie predyktywnych rozkładów rezerw szkodowych są wyższe od tych otrzymanych klasyczną metodą *chain ladder*. Ponadto odchylenia standardowe tych rozkładów są stosunkowo duże. Jest to wynik przyjętych założeń *a priori*, w szczególności założenia o rozproszonych rozkładach parametrów τ , τ_0 i μ_0 . W praktyce aktuarialnej w roku następnym otrzymane rozkłady *a posteriori* służą jako rozkłady *a priori* i często ma miejsce sytuacja, w której po każdym uzupełnieniu danych i uaktualnieniu rozkładu *a priori* wariancja interesujących nas rozkładów parametrów spada, co sugeruje wzrost precyzji oszacowania. Taki sposób postępowania jest właściwy dla empirycznych metod bayesowskich. Przydatność tych metod w prowadzeniu działalności ubezpieczeniowej powinna być przedmiotem dalszych badań.

Literatura

- [1] Alba de E., *Bayesian Estimation of Outstanding Claim Reserves*, „North American Actuarial Journal” 2002, 6(4), 1-20.
- [2] Klugman S.A., *Bayesian Statistics in Actuarial Science*, Kluwer, Boston 1992.
- [3] Scollnik D.P.M., *Actuarial Modeling with MCMC and BUGS*, „North American Actuarial Journal” 2001, 5(2), 96-124.

- [4] Verrall R.J., *An Investigation into Stochastic Claims Reserving Models and the Chain-Ladder Technique*, „Insurance: Mathematics and Economics” 2000, vol. 26, 91-99.
- [5] Verrall R.J., *A Bayesian Generalised Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving*, „North American Actuarial Journal” 2004, 8.
- [6] Verrall R.J., *Bayes and Empirical Bayes Estimation for the Chain Ladder Model*, „ASTIN Bulletin” 1990, 20(2), 217-243.

BAYESIAN ESTIMATION OF OUTSTANDING CLAIMS RESERVES

Summary

This paper examines Bayesian model for loss reserving inspired by a consideration of some of the methods appearing in the traditional chain ladder literature.

This includes possible order restricted hierarchical Bayesian model for the year over year development factors. Described Bayesian model can be implemented using Markov chain Monte Carlo (MCMC) methods and can serve as a starting point for more advanced models. This paper describes its implementation in WINBUGS (a specialized software programme for MCMC simulations).