PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU
Nr 60 — 2009
Inwestycje finansowe i ubezpieczenia – tendencje światowe a polski rynek

Alicja Wolny-Dominiak
Akademia Ekonomiczna w Katowicach

METODY TARYFIKACJI NA PRZYKŁADZIE UBEZPIECZEŃ KOMUNIKACYJNYCH

1. Wstęp

Zakłady ubezpieczeń działające w grupie ubezpieczeń majątkowych tworzą własne systemy taryfikacyjne w celu ustalenia sprawiedliwego poziomu składki dla każdego ryzyka w różnego rodzaju portfelach ubezpieczeniowych. Tworzenie systemu taryfikacyjnego bazuje głównie na analizie danych dotyczących szkodowości oraz przebiegu ubezpieczenia jednostek bądź grup (klas) ubezpieczeniowych w danym portfelu ubezpieczeń. Ustalenie składki dla konkretnego ryzyka na podstawie danego systemu taryfikacyjnego jest dwuetapowe; składają się na nie: taryfikacja *a priori* oraz taryfikacja *a posteriori*. Taryfikacja *a priori* polega na kalkulacji składki bazowej z uwzględnieniem czynników opisującch: osobę ubezpieczającą się oraz jej wcześniejszy przebieg ubezpieczenia, charakterystyczne cechy przedmiotu ubezpieczenia (pojazd, nieruchomość itd.). Natomiast taryfikacja *a posteriori* to określenie składki dodatkowej (liczonej najczęściej jako procent składki bazowej) uwzględniające liczbę szkód spowodowanych w przeszłości przez osobę ubezpieczającą się. W tym celu wykorzystywany jest najczęściej system *bonus malus* − system zniżek−zwyżek.

W pracy rozpatrywane są statystyczne metody taryfikacji *a priori* z grupy procedur zwanych ogólnie procedurą minimalnego obciążenia (*minimum bias procedure − MBP*). W celu optymalnego wyboru poziomu wszystkich składowych w procesie taryfikacji *a priori* w pracy zaproponowana została procedura bazująca na symulacji Monte Carlo analizująca zmiany przebiegu szkodowości w przypadku wystąpienia różnych rozkładów szkód.

2. Wybrane procedury MBP w taryfikacji *a priori*

W literaturze znaleźć można liczne procedury z grupy MBP. Ogólną ideą tego typu podejścia do taryfikacji *a priori* jest pomiar wpływu zmiennych taryfikacyjnych na przebieg szkodowości za pomocą tzw. wskaźników zależności (*relativities*), a następnie wyznaczenie tablicy taryfikacyjnej (w przypadku dwóch zmiennych taryfikacyjnych). Wskaźniki zależności pokazują, o ile należy zmienić poziom ustalanej składki bazowej w każdym subportfelu. Najczęściej procedury MBP są modyfikacjami czterech podstawowych podejść do badania wpływu zmiennych taryfikacyjnych na poziom szkodowości [Bailey, LeRoy 1960]. W celu zobrazowania każdego podejścia przyjęto następujące oznaczenia:

1. X, Y − zmienne taryfikacyjne.
2. $x_1,...,x_n$ − wskaźniki wpływu dla zmiennej X, $y_1,...,y_m$ − wskaźniki wpływu dla zmiennej Y.
3. r_{ij} − średnia wartość pojedynczej szkody dla i-tej realizacji zmiennej X oraz dla j-tej realizacji zmiennej Y.
4. n_{ij} − liczba szkód dla i-tej realizacji zmiennej X oraz dla j-tej realizacji zmiennej Y.
5. t_{ij} − elementy tablicy taryfikacyjnej.
6. B – bazowa średnia wartość szkody.

Podejście pierwsze zakłada, że dla każdej realizacji zmiennej taryfikacyjnej występuje tzw. zasada bilansu, którą można zapisać następująco [Fu i in. 2007]:

a) dla zmiennej X − $\sum_j n_{ij} r_{ij} = \sum_j n_{ij} B x_i y_j$,

b) dla zmiennej Y − $\sum_i n_{ij} r_{ij} = \sum_i n_{ij} B x_i y_j$.

Po przekształceniu równania bilansującego estymatory wskaźników zależności wyrażają się wzorem: $\hat{x}_i = \frac{\sum_j n_{ij} r_{ij}}{\sum_j n_{ij} B y_j}$ oraz $\hat{y}_j = \frac{\sum_i n_{ij} r_{ij}}{\sum_i n_{ij} B x_i}$.

Pozostałe podejścia bazują na minimalizacji funkcji obciążenia, która jest w różny sposób definiowana. W podejściu drugim do szacowania poziomu wskaźników zależności stosowane jest kryterium minimalizacji kwadratów odchyleń średniej wartości szkody od jej szacowanej wielkości. Zatem w tym modelu wyznaczyć należy $\min\limits_{x,y} \sum_{i,j} n_{ij} (r_{ij} - B x_i y_j)^2$. Stosując twierdzenia o istnieniu ekstremum lokalnego dla funkcji wielu zmiennych, uzyskujemy następujące równania:

a) dla zmiennej X − $\sum_j n_{ij} 2(r_{ij} - B x_i y_j)(-y_j) = 0$,

b) dla zmiennej Y − $\sum_i n_{ij} 2(r_{ij} - B x_i y_j)(-x_i) = 0$.

Przekształcając powyższe równania, uzyskujemy następujące estymatory wskaźników zależności $\hat{x}_i = \frac{\sum_j n_{ij} r_{ij} B y_j}{\sum_j n_{ij} (B y_j)^2}$ oraz $\hat{y}_j = \frac{\sum_i n_{ij} r_{ij} B x_i}{\sum_i n_{ij} (B x_i)^2}$.

Podejście trzecie za kryterium wyznaczenia estymatorów wskaźników przyjmuje minimalizację błędu χ^2, co można zapisać jako: $\min_{x,y} \sum_{i,j} n_{ij} \frac{(r_{ij} - Bx_i y_j)^2}{Bx_i y_j}$.
Ponownie stosując twierdzenie odnośnie do istnienia ekstremum, uzyskujemy następujące estymatory $\hat{x}_i = \sqrt{\frac{\sum_j n_{ij} r_{ij}^2 (By_j)^{-1}}{\sum_j n_{ij} By_j}}$ oraz $\hat{y}_j = \sqrt{\frac{\sum_i n_{ij} r_{ij}^2 (Bx_i)^{-1}}{\sum_i n_{ij} x_i}}$.

W podejściu czwartym stosowana jest metoda największej wiarygodności. W tym celu niezbędne jest przyjęcie założenia o postaci rozkładu opisującego szkodowość. W literaturze przedmiotu przeprowadzono wiele analiz, które pokazały, iż rozkłady dotyczące szkodowości charakteryzują się następującymi własnościami: przyjmują wartości dodatnie, występuje proporcjonalność wariancji do wartości oczekiwanej oraz asymetria dodatnia [McCullagh, Nelder 1999]. Dlatego też szkodowość ryzyka najczęściej modelowana jest rozkładem gama lub odwróconym normalnym. Ponadto w analizie danych ubezpieczeniowych zastosowanie znajduje grupa rozkładów zwana rozkładami Tweedie [Anderson i in. 2007]. Wszystkie wymienione rozkłady należą do tzw. dyspersyjnej rodziny rozkładów wykładniczych. Przy założeniu rozkładu gamma dla średniej wartości szkody estymatory największej wiarygodności dla zmiennych X, Y przyjmują postać $\hat{x}_i = \frac{\sum_j n_{ij} r_{ij} (By_j)^{-1}}{\sum_j n_{ij}}$ oraz $\hat{y}_j = \frac{\sum_i n_{ij} r_{ij} (Bx_i)^{-1}}{\sum_i n_{ij}}$. Natomiast jeśli zakłada się rozkład odwrócony normalny dla średniej wartości szkody, otrzymuje się $\hat{x}_i = \frac{\sum_j n_{ij} r_{ij} /(By_j)^2}{\sum_j n_{ij} /(By_j)}$ oraz $\hat{y}_j = \frac{\sum_i n_{ij} r_{ij} / (Bx_i)^2}{\sum_i n_{ij} / (Bx_i)}$. W celu wyznaczenia wartości wskaźników zależności niezbędne jest wykonanie algorytmu iteracyjnego, obliczając w kolejnych iteracjach x_i^k oraz y_j^k przy $k \to \infty$. Za punkt startowy przyjmowane są $x_0 = 1$, $y_0 = 1$. Ponadto w każdej iteracji niezbędne jest założenie poziomu zmiennej bazowej *B*. W pracy przyjmowana jest stała baza *B* obliczana jako średnia ważona $B = \frac{\sum_{i,j} n_{ij} r_{ij}}{\sum_{i,j} n_{ij}}$. Warunkiem końcowym algorytmu iteracyjnego w powyższych modelach jest: $x_i = \lim_{k->\infty} \hat{x}_i^k$, $y_j = \lim_{k->\infty} \hat{y}_j^k$ [Mildenhall 1999].

Oszacowane wartości wskaźników zależności służą do wyznaczenia tzw. tablicy taryfikacyjnej, ozn. $T = [t_{ij}]$. Elementy tablicy T w przypadku założenia modelu multiplikatywnego mają postać $t_{ij} = \exp(\beta_i + \beta_j)$. Jeśli założymy, że ryzyko w modelu klasyfikowane jest pod względem dwóch cech: regionu oraz pojemności silnika, to elementy t_{ij} tablicy pokazują, o ile należy skorygować składkę bazową dla *i*-tego regionu oraz *j*-tej pojemności silnika.

Do oceny uzyskanych szacunków, a także do wyboru modelu MBP niezbędne jest ustalenie kryteriów wyboru. W tym celu wyznaczyć można znane miary statystyczne [Bailey, LeRoy 1960], np. procentowe ważone odchylenie przeciętne $d = \dfrac{\sum_{i,j} n_{ij} \dfrac{|E(\widehat{Y}_{ij}) - \mu_{ij}|}{\mu_{ij}}}{\sum_{i,j} n_{ij}}$. Wybrano tę miarę, gdyż w konkretnym przykładzie można łatwo i intuicyjnie interpretować wyniki.

3. Przykład zastosowania procedur MBP dla danych komunikacyjnych

W dalszej części artykułu przedstawiony został przykład obliczeniowy obrazujący działanie procedur MBP dla każdego opisanego podejścia z wykorzystaniem danych komunikacyjnych. W przykładzie przyjęto następujące założenia: szkodowość opisywana jest przez wartość pojedynczej szkody oraz liczbę szkód, w modelu występują dwie zmienne taryfikacyjne: wiek kierowcy oraz sposób użytkowania samochodu, między zmiennymi taryfikacyjnymi występuje zależność multiplikatywna. Algorytm iteracyjny zaimplementowany został w języku **R**. Dane zaczerpnięto z literatury przedmiotu [Mildenhall 1999] ze względu na to, iż autorowi nie udało się uzyskać żadnych danych od ubezpieczycieli funkcjonujących na rynku polskim. Przyjęto założenie o warunku końcowym algorytmu: < 0,0000001. Dla takiego modelu zbieżność uzyskana została dla *k* = 4. Ponadto przyjęto stałą wartość bazową obliczaną jako średnia ważona $B = \dfrac{\sum_{ij} n_{ij} r_{ij}}{\sum_{ij} n_{ij}}$. W analizowanym przykładzie baza wyniosła więc $B = 241{,}46$. Uzyskane wyniki przedstawiają tab. 1-2.

We wszystkich pięciu zastosowanych modelach widać, że dla kierowców w wieku młodym i średnim od 17 do 34 lat występuje większa średnia szkodowość w stosunku do wartości bazowej. Oznacza to, iż składkę należy odpowiednio zwiększać. Podobna sytuacja występuje w odniesieniu do użytkowania samochodu, kiedy wzrost składki powinien wystąpić dla użytkowników samochodów służbowych oraz dojeżdżających do pracy > 10 km. W pozostałych przypadkach składka powinna zostać zmniejszona.

Tabela 1. Wartości wskaźników wpływu dla zmiennej *X* (wiek)

	*x*1	*x*2	*x*3	*x*4	*x*5	*x*6	*x*7	*x*8
Model 1	1,203546	1,207631	1,15438	1,123666	0,89052	0,970975	0,953408	0,921826
	1,259185	1,222628	1,136752	1,09986	0,878172	0,959818	0,97269	0,954536
	1,260293	1,222945	1,136482	1,099436	0,877956	0,959584	0,972997	0,955178
STOP	1,260314	1,222951	1,136477	1,099428	0,877952	0,959579	0,973003	0,95519
Model 2	1,203546	1,207631	1,154380	1,123666	0,890520	0,970975	0,953408	0,921826
	1,289876	1,209254	1,127879	1,102567	0,871397	0,965814	0,976279	0,961449
	1,292142	1,209284	1,127265	1,102012	0,870950	0,965623	0,976815	0,962458
STOP	1,292200	1,209285	1,127250	1,101997	0,870939	0,965618	0,976829	0,962484
Model 3	1,309298	1,219359	1,162829	1,142823	0,899809	0,991953	0,96854	0,939983
	1,329106	1,249564	1,154683	1,115882	0,894589	0,974865	0,987277	0,969541
	1,329725	1,249977	1,154582	1,115483	0,89444	0,974629	0,987536	0,970037
STOP	1,329735	1,249983	1,154581	1,115477	0,894438	0,974626	0,98754	0,970045
Model 4	1,203546	1,207631	1,15438	1,123666	0,89052	0,970975	0,953408	0,921826
	1,239827	1,234401	1,144724	1,097248	0,88339	0,955742	0,969952	0,948634
	1,240569	1,234749	1,144644	1,096886	0,883228	0,955536	0,970163	0,949075
STOP	1,24058	1,234754	1,144643	1,09688	0,883225	0,955533	0,970166	0,949082
Model 5	1,203546	1,207631	1,15438	1,123666	0,89052	0,970975	0,953408	0,921826
	1,229148	1,244102	1,151596	1,094755	0,886945	0,953368	0,967958	0,943709
	1,229676	1,244391	1,151644	1,094445	0,886801	0,953202	0,968122	0,944017
STOP	1,229683	1,244393	1,151645	1,094441	0,886798	0,9532	0,968124	0,944021

Źródło: obliczenia własne.

Tabela 2. Wartości wskaźników wpływu dla zmiennej *Y* (sposób użytkowania pojazdu)

	*y*1	*y*2	*y*3	*y*4
Model 1	0,854618	0,88785	1,071906	1,390519
	0,850754	0,886292	1,073622	1,396364
	0,85068	0,886265	1,073654	1,396471
STOP	0,850678	0,886264	1,073654	1,396473
Model 2	0,854744	0,88852	1,070131	1,386918
	0,850158	0,885728	1,071178	1,394628
	0,850034	0,885659	1,071205	1,394823
STOP	0,850031	0,885657	1,071205	1,394828
Model 3	0,842359	0,874037	1,056052	1,376687
	0,83892	0,872723	1,057486	1,381844
	0,838863	0,872704	1,057509	1,381921
STOP	0,838862	0,872703	1,05751	1,381922
Model 4	0,854173	0,88745	1,073693	1,393434
	0,85098	0,886537	1,075486	1,398901
	0,850929	0,886525	1,075513	1,39898
STOP	0,850928	0,886525	1,075513	1,398981
Model 5	0,853433	0,887294	1,075481	1,395595
	0,850956	0,886732	1,077103	1,401591
	0,850918	0,886728	1,077124	1,401653
STOP	0,850918	0,886728	1,077124	1,401653

Źródło: obliczenia własne.

Tabela 3. Tablica taryfikacyjna

Model 1	17-20	21-24	25-29	30-34	35-39	40-49	50-59	60+
Prywatnie	1,072	1,040	0,967	0,935	0,747	0,816	0,828	0,813
Do pracy <10km	1,117	1,084	1,007	0,974	0,778	0,850	0,862	0,847
Do pracy >10km	1,353	1,313	1,220	1,180	0,943	1,030	1,045	1,026
Służbowo	1,760	1,708	1,587	1,535	1,226	1,340	1,359	1,334
Model 2								
Prywatnie	1,098	1,028	0,958	0,937	0,740	0,821	0,830	0,818
Do pracy <10km	1,144	1,071	0,998	0,976	0,771	0,855	0,865	0,852
Do pracy >10km	1,384	1,295	1,208	1,180	0,933	1,034	1,046	1,031
Służbowo	1,802	1,687	1,572	1,537	1,215	1,347	1,363	1,343
Model 3								
Prywatnie	1,115	1,049	0,969	0,936	0,750	0,818	0,828	0,814
Do pracy <10km	1,160	1,091	1,008	0,973	0,781	0,851	0,862	0,847
Do pracy >10km	1,406	1,322	1,221	1,180	0,946	1,031	1,044	1,026
Służbowo	1,838	1,727	1,596	1,542	1,236	1,347	1,365	1,341
Model 4								
Prywatnie	1,056	1,051	0,974	0,933	0,752	0,813	0,826	0,808
Do pracy <10km	1,100	1,095	1,015	0,972	0,783	0,847	0,860	0,841
Do pracy >10km	1,334	1,328	1,231	1,180	0,950	1,028	1,043	1,021
Służbowo	1,736	1,727	1,601	1,535	1,236	1,337	1,357	1,328
Model 5								
Prywatnie	1,046	1,059	0,980	0,931	0,755	0,811	0,824	0,803
Do pracy <10km	1,090	1,103	1,021	0,970	0,786	0,845	0,858	0,837
Do pracy >10km	1,325	1,340	1,240	1,179	0,955	1,027	1,043	1,017
Służbowo	1,724	1,744	1,614	1,534	1,243	1,336	1,357	1,323

Źródło: obliczenia własne.

W analizowanym przykładzie (tab. 3) widać, iż w modelu 1, 4 oraz 5 uzyskane zostały zbliżone wartości taryfikatorów. Wyższe wartości, ale również zbliżone do siebie dają oszacowania w modelu 3 i 4. Zatem w celu wyboru ostatecznej wartości wyznaczony został błąd oszacowania *d* (zob. tab. 4).

Tabela 4. Procentowe ważone odchylenie przeciętne

	Model 1	Model 2	Model 3	Model 4	Model 5
$d \to min$	4,4537%	4,7045%	4,4229%	4,2584%	4,1509%

Źródło: obliczenia własne.

Dla wszystkich analizowanych modeli procentowe odchylenie przeciętne wyniosło nieznacznie ponad 4%. Jednak najniższa wartość błędu wystąpiła w modelu 5, zatem dał on najlepsze szacunki w tym przykładzie.

4. Symulacja Monte Carlo dla MBP

Kształtowanie się przebiegu szkodowości w przyszłości będzie odzwierciedlało, na ile trafnie oszacowano wartości wskaźników zależności. Dlatego przydatnym narzędziem do analizy różnych scenariuszy jest symulacja przebiegu szkodowości; w taryfikacji średnią wartość szkody oraz liczbę szkód traktuje się jako zmienne losowe o założonych rozkładach [Fu, Moncher 2004]. Losowanie wartości zmiennej losowej z rozkładu pokazuje w symulacji, jak zmieniają się wskaźniki zależności w różnych scenariuszach. W proponowanej w pracy procedurze symulacyjnej Monte Carlo wyróżnić można następujące etapy (jedna iteracja): założenie postaci rozkładu dla średniej wartości szkody, wygenerowanie macierzy średniej wartości szkody, wyznaczenie poziomu wskaźników wpływu, wyznaczenie tablicy taryfikacyjnej, estymacja średniej wartości szkody, obliczenie błędu szacunku modelu. Procedurę można zastosować w odniesieniu do różnych typów rozkładu dla zmiennych losowych występujących w modelach. Jak już wspomniano, najczęściej liczba szkód modelowana jest rozkładem Poissona, natomiast dla średniej wartości szkody stosowane są rozkłady z dyspersyjnej rodziny rozkładów wykładniczych.

Tabela 5. Rozkłady wskaźników wpływu dla zmiennej X – rozkład Tweedie

Decyle	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$	$x7$	$x8$
Min.	1,0000	1,0000	1,0000	1,0000	0,8769	0,9462	0,9602	0,9348
10%	1,2129	1,2354	1,1460	1,0895	0,8825	0,9501	0,9643	0,9400
20%	1,2190	1,2384	1,1478	1,0911	0,8838	0,9511	0,9655	0,9413
30%	1,2226	1,2407	1,1491	1,0923	0,8847	0,9518	0,9664	0,9421
40%	1,2264	1,2422	1,1502	1,0934	0,8855	0,9523	0,9671	0,9429
50%	1,2298	1,2438	1,1512	1,0943	0,8862	0,9529	0,9677	0,9438
60%	1,2334	1,2460	1,1523	1,0951	0,8870	0,9536	0,9684	0,9445
70%	1,2368	1,2475	1,1535	1,0960	0,8879	0,9543	0,9691	0,9453
80%	1,2413	1,2497	1,1548	1,0972	0,8889	0,9550	0,9700	0,9464
90%	1,2479	1,2529	1,1568	1,0987	0,8904	0,9561	0,9711	0,9478
Max	1,2722	1,2654	1,1640	1,1052	1,0000	1,0000	1,0000	1,0000
Odchylenie standardowe	0,01537	0,01027	0,00634	0,00460	0,00471	0,00276	0,00287	0,00350

Źródło: obliczenia własne.

Tabela 6. Rozkłady wskaźników wpływu dla zmiennej Y – rozkład Tweedie

Decyle	$y1$	$y2$	$y3$	$y4$
Min.	0,8415	0,8804	1,0000	1,0000
10%	0,8463	0,8834	1,0727	1,3931
20%	0,8476	0,8841	1,0737	1,3948
30%	0,8485	0,8846	1,0744	1,3963
40%	0,8492	0,8851	1,0751	1,3974
50%	0,8499	0,8855	1,0756	1,3986
60%	0,8508	0,8860	1,0762	1,3998
70%	0,8516	0,8865	1,0767	1,4010
80%	0,8526	0,8870	1,0775	1,4024
90%	0,8538	0,8877	1,0783	1,4039
Max	1,0000	1,0000	1,0835	1,4139
Odchylenie standardowe	0,00557	0,00399	0,00328	0,01330

Źródło: obliczenia własne.

Tabela 7. Rozkład błędu *d* − rozkład Tweedie

Decyle	*d*
Min.	3,86%
10%	4,03%
20%	4,07%
30%	4,10%
40%	4,13%
50%	4,16%
60%	4,18%
70%	4,21%
80%	4,24%
90%	4,28%
Max.	100,00%
Odchylenie standardowe	0,03031

Źródło: obliczenia własne.

Poniżej przedstawiono przykład proponowanej symulacji Monte Carlo. W pierwszym etapie symulacji założony został rozkład Tweedie z parametrami $T(n_{ij}, 1.2, r_{ij}, 1)$ średniej wartości szkody (liczba szkód nie jest w tym przykładzie modelowana rozkładem, przyjmowane są stały wartości). Do szacowania wskaźników wpływu zastosowano model 5 (dający minimalny błąd *d*). Po wykonaniu 1000 iteracji uzyskano rozkłady zmiennych przedstawione w tab. 5-6.

Widać, że wystąpiło niewielkie zróżnicowanie pomiędzy wartościami wskaźników zależności. Ponadto błąd *d* kształtuje się na relatywnie niskim poziomie (tab. 7).

6. Podsumowanie

W pracy przedstawiono cztery podejścia stosowane w szacowaniu wartości wskaźników wpływu w taryfikacji *a priori* i wykorzystujące zasadę minimalnego obciążenia, tzw. procedury minimalnego obciążenia. W celu prezentacji działania procedur zastosowanych zostało 5 modeli szacowania wskaźników zależności. W analizowanym przykładzie empirycznym wyznaczono dodatkowo błąd oszacowania pozwalający wybrać ostateczne wartości dla tablicy taryfikacyjnej i wskazać optymalny model. Ponieważ modele wymagały zastosowania algorytmu iteracyjnego, wszystkie procedury zaimplementowane zostały w języku **R**. W praktyce konieczność implementacji komputerowej takiego algorytmu postrzegana jest jako główna wada tego typu podejścia do taryfikacji *a priori* (nie dotyczy to np. modeli GLM, które obsługuje większość pakietów statystycznych). Ponadto w pracy zaproponowano procedurę symulacyjną bazującą na symulacji Monte Carlo, pozwalającą analizować zmiany w poziomie wskaźników zależności w zależności od ewentualnego rozwoju szkodowości w przyszłości. Wyznaczając rozkład błędu *d,* możliwe jest również testowanie różnych modeli i dokonanie oceny uzyskiwanych wyników. Wykorzystanie symulacji do wspomagania wyboru modelu taryfikacji będzie przedmiotem dalszej pracy autora.

Literatura

Anderson D., Feldblum S., Modlin C., Schirmacher D., Schirmacher E., Thandi N., *A practitioner's guide to generalized linear models. A foundation for theory, interpretation and application*, www.watsonwyatt.com, 2007.

Bailey R.A., LeRoy J.S., *Two studies in automobile insurance ratemaking*, PCAS 1960, XLVII, s. 1-19.

Fu L., Moncher R. B., *Severity distributions for GLMs: gamma or lognormal? Evidence from Monte Carlo simulations*, Casualty Actuarial Society Discussion Paper Program Casualty Actuarial Society, Arlington, Virginia 2004.

Fu L., Wu Ch.P., *General iteration algorithms for classification ratemaking*, "Variance Journal" 2007, vol. 1, is. 2.

McCullagh P., Nelder J.A., *Generalized linear models*, Chapman & Hall/CRC, New York 1999.

Mildenhall S., *A systematic relationship between minimum bias and generalized linear models*, PCAS 1999, LXXXVI, s. 394-487.

RATING SYSTEMS FOR SETTING FAIR PREMIUMS – AUTOMOBILE INSURANCE EXAMPLE

Summary

Insurance companies specialising in casualty insurance create their own rating systems for setting fair premiums for every risk for different kinds of insurance portfolios. The rating system is mostly based on the data analysis concerning the number and the value of claims for individuals or groups (classes) of insured people within a given portfolio. Based on a given rating system, the premium for a particular risk is calculated in two stages: a priori rating and a posteriori rating. In this paper, the process of a priori rating is analysed with the emphasis on minimum bias methods used for modelling the rating variables.