

Zagadnienia aktuarialne – teoria i praktyka

pod redakcją
Walentego Ostasiewicza



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2011

Recenzenci

Helena Jasiulewicz, Wojciech Kordecki, Zbigniew Palmowski

Redaktor Wydawnictwa

Aleksandra Śliwka

Redakcja techniczna

Barbara Łopusiewicz

Korektor

Barbara Cibis

Łamanie

Beata Mazur

Projekt okładki

Beata Dębska

Publikacja jest dostępna na stronie www.ibuk.pl

Streszczenia opublikowanych artykułów są dostępne w międzynarodowej bazie danych The Central European Journal of Social Sciences and Humanities <http://cejsh.icm.edu.pl> oraz w The Central and Eastern European Online Library www.ceeol.com

Informacje o naborze artykułów i zasadach recenzowania znajdują się na stronie internetowej Wydawnictwa www.wydawnictwo.ue.wroc.pl

Kopiowanie i powielanie w jakiegokolwiek formie wymaga pisemnej zgody Wydawnictwa

© Copyright Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2011

ISSN 1899-3192

ISBN 978-83-7695-186-7

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp	7
Irmina Czarna, Zbigniew Palmowski , Porównanie prawdopodobieństw paryskiej i klasycznej ruiny dla procesu ryzyka typu Lévy'ego	9
Irmina Czarna, Zbigniew Palmowski , Problem wyboru optymalnej paryskiej dywidendy dla procesu ryzyka typu Lévy'ego – numeryczna analiza	22
Joanna Dębicka , Składki netto dla ubezpieczeń wielostanowych obciążone kosztami zawarcia i prowadzenia umowy	38
Monika Dyduch , Niekonwencjonalna metoda prognozy wartości jednostek funduszy emerytalnych	69
Stanisław Heilpern , Niestandardowe modele ryzyka – badanie wpływu stopnia zależności na prawdopodobieństwo ruiny	79
Aleksandra Iwanicka , Wpływ zewnętrznych czynników ryzyka na prawdopodobieństwo ruiny w dwuwymiarowym modelu ryzyka z lekkoogonowymi rozkładami wypłat	92
Helena Jasiulewicz, Wojciech Kordecki , Składki zaufania z zastosowaniem niesymetrycznych funkcji strat	101
Kamil Jodź , Składka w modelu ryzyka indywidualnego z zależnymi roszczeniami opisanymi funkcjami łączącymi	118
Marek Kałuszka, Michał Krzeszowiec , Własności składki <i>mean-value</i> przy zniekształconym prawdopodobieństwie	136
Zbigniew Michna , Procesy Lévy'ego w modelach ubezpieczeniowych	149
Agnieszka Mruklik , Ubezpieczenia na życie ze stochastyczną techniczną stopą oprocentowania – zastosowanie modelu Hulla i White'a	157
Agnieszka Pobłocka , Rezerwa IBNR w ubezpieczeniach majątkowych – praktyczne metody jej szacowania	173
Agata de Sas Stupnicka , Równowaga na rynku ubezpieczeń zdrowotnych w zależności od przyjętego sposobu rozliczania świadczeń medycznych	190
Joanna Sawicka , Zagadnienia kalkulacji składki zaufania na podstawie łącznej wartości i liczby szkód	202
Alicja Wolny-Dominiak , Analiza porównawcza modeli mieszanych szacowania stóp taryf w ubezpieczeniach majątkowych z wykorzystaniem krosvalidacji	229
Walenty Ostasiewicz , Polacy nie gęsi, iż swój język mają!	238

Summaries

Irmina Czarna, Zbigniew Palmowski , Comparison of Parisian and classical ruin probabilities for a Lévy risk process	21
Irmina Czarna, Zbigniew Palmowski , Numerical analysis of dividend problem with Parisian delay for a spectrally negative Lévy risk process	37
Joanna Dębicka , Expense-loaded premiums for multistate insurance contracts	68
Monika Dyduch , Alternative method of forecast of pension funds units value	78
Stanisław Heilpern , Nonstandard risk models – study of influence of the degree of dependence on the probability of ruin	91
Aleksandra Iwanicka , The influence of some outside risk factors on a ruin probability in a two-dimensional risk model with light-tailed claim sizes	100
Helena Jasiulewicz, Wojciech Kordecki , Credibility premiums using asymmetric loss functions	117
Kamil Jodź , Insurance premium in individual risk model with dependent claims described by copulas functions	135
Marek Kaluszka, Michał Krzeszowiec , Properties of mean-value principle under rank-dependent utility model	148
Zbigniew Michna , Lévy processes in insurance models	156
Agnieszka Mruklik , Life insurance with stochastic interest rate – an application of the Hull and White model	172
Agnieszka Pobłocka , IBNR reserve in non-life insurance. Practical methods of its estimation	189
Agata de Sas Stupnicka , Balance on the health insurance market – the impact of payment system	201
Joanna Sawicka , Calculation of credibility premium on the basis of number and total amount of claims	228
Alicja Wolny-Dominiak , Comparative analysis of mixed models for ratemaking in non-life insurance with k-fold cross-validation	237

Joanna Sawicka

Uniwersytet Warszawski

ZAGADNIENIA KALKULACJI SKŁADKI ZAUFANIA NA PODSTAWIE ŁĄCZNEJ WARTOŚCI I LICZBY SZKÓD

Streszczenie: W artykule zostało rozważone zagadnienie obliczania predyktora zaufania dla łącznej wartości szkód. Przedstawiono postać składki obliczoną na podstawie łącznej wartości szkód, na podstawie liczby szkód oraz na podstawie łącznej wartości i liczby szkód jednocześnie. Obliczone zostały także błędy średniokwadratowe zaproponowanych składek oraz rozważone zostały konsekwencje przyjęcia różnych założeń dotyczących parametrów ryzyka rozkładów wartości pojedynczej szkody i liczby szkód dla postaci składek i wielkości ich błędów średniokwadratowych. W artykule zostały także zaprezentowane przykłady numeryczne ilustrujące zagadnienie kalkulacji składki w tym modelu.

Słowa kluczowe: składka zaufania, liczba szkód, wartość szkód.

1. Wstęp

Niniejszy artykuł dotyczy zagadnienia optymalnej liniowej predykcji łącznej wartości szkód w kolejnym okresie dla niejednorodnego portfela ubezpieczonych. Klasyyczny predyktor łącznej wartości szkód (przedstawiony np. w artykule Bühlmanna [1967]) oparty jest na informacji na temat łącznych wartości szkód danego ubezpieczonego z wcześniejszych okresów. Można jednak dokonać predykcji łącznej wartości szkód także na podstawie liczby szkód lub na podstawie liczby i wartości szkód jednocześnie. W niniejszym artykule zostaną zatem wyprowadzone trzy postacie predyktorów obliczone na podstawie liczby, wartości szkód oraz obu tych informacji. Podobną tematykę podjęli także m.in. Goulet i in. [2006], lecz pominieli oni w swojej pracy zagadnienie predykcji na podstawie liczby szkód. W artykule zostaną ponadto obliczone błędy średniokwadratowe trzech wymienionych predyktorów, dzięki czemu będzie można wskazać najlepszy pod tym względem sposób predykcji. Zostanie także rozważona postać predyktorów możliwa do uzyskania po przyjęciu pewnych upraszczających założeń dotyczących rozkładów wartości pojedynczej szkody i liczby szkód. Na zakończenie w celu zilustrowania uzyskanych wyników zostanie zaprezentowany prosty przykład numeryczny.

2. Założenia modelu, podstawowe oznaczenia oraz parametry rozkładów

Mamy zbiór danych zawierający informacje o liczbie i wartości szkód poniesionych przez M ubezpieczonych podczas T okresów. Oznaczmy przez $N_{j,t}$ liczbę szkód j -tego ubezpieczonego w t -tym okresie, przez $Y_{j,t,k}$ wartość k -tej szkody j -tego ubezpieczonego w t -tym okresie; $X_{j,t} = \sum_{k=1}^{N_{j,t}} Y_{j,t,k}$ niech oznacza łączną wartość szkód, jakie poniósł j -ty ubezpieczony w t -tym okresie. Jak widać zatem, łączna wartość szkód j -tego ubezpieczonego w t -tym okresie zależy od dwóch składników, które najczęściej uznaje się za losowe, a mianowicie od liczby szkód i od wartości pojedynczej szkody.

Podstawowym założeniem stosowanym w teorii zaufania jest przyjęcie, że istnieją pewne nieobserwowalne parametry ryzyka opisujące rozkład liczby i wartości pojedynczej szkody j -tego ubezpieczonego. Ponieważ populacja ubezpieczonych jest niejednorodna, zakłada się, że owe parametry ryzyka są także realizacjami zmiennych losowych. Można zatem powiedzieć, że szkody z portfela ubezpieczonych powstają w dwóch etapach – w pierwszym etapie losuje się wartości parametrów ryzyka dla ubezpieczonych; w drugim etapie ubezpieczony o ustalonych wartościach parametrów ryzyka ponosi w kolejnych latach różne liczby szkód o różnych wartościach. Aby opisać to w bardziej sformalizowany sposób, przyjmijmy następujące oznaczenia i założenia:

A) Założenia dotyczące parametrów ryzyka:

A1) Rozkład łącznej wartości szkód j -tego ubezpieczonego zależy od dwóch parametrów ryzyka: λ_j i θ_j , $j = 1, \dots, M$, gdzie λ_j jest parametrem ryzyka rozkładu liczby szkód, a θ_j jest parametrem ryzyka rozkładu wartości pojedynczej szkody. Przyjmujemy, że λ_j jest realizacją zmiennej losowej Λ_j , natomiast θ_j jest realizacją zmiennej losowej Θ_j . Co istotne, zakładamy, że zmienne losowe Λ_j i Θ_j są niezależne, $j = 1, \dots, M$.

A2) Zmienne losowe $\Lambda_1, \Lambda_2, \dots, \Lambda_M$ dla M ubezpieczonych są wzajemnie niezależne i mają ten sam rozkład.

A3) Zmienne losowe $\Theta_1, \Theta_2, \dots, \Theta_M$ dla M ubezpieczonych są wzajemnie niezależne i mają ten sam rozkład.

B) Założenia dotyczące warunkowych rozkładów liczby szkód, wartości pojedynczej szkody oraz łącznej wartości szkód:

B1) Przy ustalonej wartości parametru ryzyka λ_j liczby szkód dla j -tego ubezpieczonego w kolejnych okresach $(N_{j,1}, N_{j,2}, \dots, N_{j,T})$ są warunkowo niezależne i mają ten sam rozkład.

B2) Przy ustalonej wartości parametru ryzyka θ_j wartości kolejnych szkód j -tego ubezpieczonego w kolejnych okresach $(Y_{j,1,1}, Y_{j,1,2}, \dots, Y_{j,1,N_{j,1}}, Y_{j,2,1}, \dots, Y_{j,2,N_{j,2}}, \dots, Y_{j,T,N_{j,T}})$ są warunkowo niezależne i mają ten sam rozkład.

B3) Przy ustalonej wartości λ_j i θ_j łączne wartości szkód j -tego ubezpieczonego w kolejnych okresach $(X_{j,1}, X_{j,2}, \dots, X_{j,T})$ są warunkowo niezależne i mają ten sam rozkład, $j = 1, \dots, M$.

B4) Ponadto warunkowo niezależne (przy danych λ_j i θ_j) są także liczba szkód i łączna wartość szkód dla j -tego ubezpieczonego dla różnych okresów (warunkowa niezależność $N_{j,k}$ i $X_{j,l}$ dla $k = 1, \dots, T$, $l = 1, \dots, T$ i $k \neq l$) oraz wartość pojedynczej szkody i łączna wartość szkód dla j -tego ubezpieczonego dla różnych okresów (warunkowa niezależność $Y_{j,k,s}$ i $X_{j,l}$ dla $s = 1, \dots, N_{j,k}$, $k = 1, \dots, T$, $l = 1, \dots, T$ i $k \neq l$).

Co istotne, ponieważ parametry ryzyka są losowane niezależnie dla wszystkich ubezpieczonych w populacji, można przyjąć, że łączna wartość szkód, liczba szkód i wartości pojedynczych szkód są niezależne dla różnych ubezpieczonych we wszystkich okresach.

Zanim przejdziemy do zagadnienia poszukiwania najlepszego predyktora łącznej wartości szkód w okresie $T + 1$, przyjmijmy pewne założenia dotyczące momentów rozkładów interesujących nas zmiennych losowych, które będą wykorzystywane w dalszej analizie:

$$C1) E(N_{j,t} | \Lambda_j) = \Lambda_j,$$

$$C2) Var(N_{j,t} | \Lambda_j) = \sigma_N^2(\Lambda_j),$$

$$C3) E(Var(N_{j,t} | \Lambda_j)) = E(\sigma_N^2(\Lambda_j)) = s_N^2,$$

$$C4) E(N_{j,t}) = E(\Lambda_j) = \Lambda,$$

$$C5) Var(\Lambda_j) = a_\Lambda^2,$$

$$D1) E(Y_{j,t,k} | \Theta_j) = \Theta_j,$$

$$D2) Var(Y_{j,t,k} | \Theta_j) = \sigma_Y^2(\Theta_j),$$

$$D3) E(Var(Y_{j,t,k} | \Theta_j)) = E(\sigma_Y^2(\Theta_j)) = s_Y^2,$$

$$D4) E(Y_{j,t,k}) = E(\Theta_j) = \Theta,$$

$$D5) Var(\Theta_j) = a_\Theta^2.$$

Powyżej przyjęte założenia i oznaczenia wystarczą do wyprowadzenia liniowych predyktorów, co zostanie wykonane w dalszych częściach artykułu.

Najpierw jednak w lemacie 1 podane zostaną parametry warunkowych i bezwarunkowych rozkładów łącznej wartości szkód, a także potrzebne przy dalszych rozważaniach wartości wariancji i kowariancji (wszystkie wartości będą liczone dla j -tego ubezpieczonego).

Lemat 1.

Przy założeniach A)-D) zachodzą następujące równości:

$$E(X_{j,t} | \Lambda_j, \Theta_j) = \Lambda_j \Theta_j, \quad (1)$$

$$E(X_{j,t}) = \Lambda \Theta, \quad (2)$$

$$Var(X_{j,t} | \Lambda_j, \Theta_j) = \sigma_Y^2(\Theta_j) \Lambda_j + \sigma_N^2(\Lambda_j) \Theta_j^2, \quad (3)$$

$$E\left[Var\left(X_{j,t} \mid \Lambda_j, \Theta_j\right)\right] = s_Y^2 \Lambda + s_N^2 \left(a_\Theta^2 + \Theta^2\right), \quad (4)$$

$$Var\left(E\left(X_{j,t} \mid \Lambda_j, \Theta_j\right)\right) = a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2, \quad (5)$$

$$Var\left(X_{j,t}\right) = s_Y^2 \Lambda + s_N^2 \left(a_\Theta^2 + \Theta^2\right) + a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2, \quad (6)$$

$$Cov\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right), X_{j,k}\right) = a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2, \quad (7)$$

$$Cov\left(X_{j,t}, X_{j,k}\right) = I_{t=k} \left[s_Y^2 \Lambda + s_N^2 \left(a_\Theta^2 + \Theta^2\right)\right] + a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2, \quad (8)$$

$$Cov\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right), \bar{X}_j\right) = a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2, \quad (9)$$

$$Var\left(\bar{X}_j\right) = \frac{1}{T} \left[s_Y^2 \Lambda + s_N^2 \left(a_\Theta^2 + \Theta^2\right)\right] + a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2, \quad (10)$$

$$Cov\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right), N_{j,k}\right) = \Theta a_\Lambda^2, \quad (11)$$

$$Cov\left(N_{j,t}, N_{j,k}\right) = I_{t=k} s_N^2 + a_\Lambda^2, \quad (12)$$

$$Cov\left(N_{j,t}, X_{j,k}\right) = I_{t=k} \Theta s_N^2 + \Theta a_\Lambda^2, \quad (13)$$

$$Cov\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right), \bar{N}_j\right) = \Theta a_\Lambda^2, \quad (14)$$

$$Var\left(\bar{N}_j\right) = \frac{1}{T} s_N^2 + a_\Lambda^2, \quad (15)$$

$$Cov\left(\bar{X}_j, \bar{N}_j\right) = \frac{1}{T} \Theta s_N^2 + \Theta a_\Lambda^2, \quad (16)$$

gdzie $I_{t=k}$ przyjmuje wartość 1, gdy $t = k$, oraz 0, gdy $t \neq k$, $\bar{X}_j = \frac{1}{T} \sum_{t=1}^T X_{j,t}$ oraz $\bar{N}_j = \frac{1}{T} \sum_{t=1}^T N_{j,t}$.

Dowód:

Równość (1) wynika z własności rozkładu złożonego oraz założeń C1) i D1), natomiast równość (2) jest bezpośrednią konsekwencją wzoru (1) w połączeniu z założeniem A1) o niezależności Λ_j i Θ_j , założeniami C4, D4) i własnością iteracyjnej wartości oczekiwanej. Równość (3) ponownie wynika z własności rozkładu złożonego oraz z założeń C1)-C2) i D1)-D2); wzór (4) wynika bezpośrednio ze wzoru (3) oraz z założeń A1), C3)-C4), D3)-D5). Równość (5) można uzyskać w następujący sposób, korzystając z założeń A1), C4)-C5) oraz D4)-D5):

$$\begin{aligned} Var\left(E\left(X_{j,t} \mid \Lambda_j, \Theta_j\right)\right) &= Var\left(\Lambda_j \Theta_j\right) = \left[Var\left(\Lambda_j\right) + \Lambda^2\right] \left[Var\left(\Theta_j\right) + \Theta^2\right] - \Lambda^2 \Theta^2 = \\ &= a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2. \end{aligned}$$

Bezpośrednią konsekwencją wzorów (4) i (5) oraz własności bezwarunkowej wariancji jest równość (6). Aby uzyskać równość (7), należy skorzystać z własności bezwarunkowej kowariancji oraz z wzorów (1) i (5). Równość (8) otrzymujemy, korzystając z założenia B3) oraz wzorów (1), (4) i (5) oraz z własności bezwarunkowej kowariancji:

$$\begin{aligned} \text{Cov}(X_{j,t}, X_{j,k}) &= I_{t=k} \left[E \left(\text{Var}(X_{j,k} | \Lambda_j, \Theta_j) \right) \right] + \text{Var}(\Lambda_j \Theta_j) = \\ &= I_{t=k} \left[s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2) \right] + a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2. \end{aligned}$$

Równość (9) można uzyskać analogicznie jak równość (5), korzystając ponadto z własności warunkowej kowariancji i wzoru (1). Wzór (10) uzyskujemy analogicznie jak równość (6), korzystając z założenia z założenia B3) oraz wzorów (1) i (5):

$$\begin{aligned} \text{Var}(\bar{X}_j) &= E \left[\frac{1}{T^2} \sum_{t=1}^T \text{Var}(X_{j,t} | \Lambda_j, \Theta_j) \right] + \text{Var}(\Lambda_j \Theta_j) = \\ &= \frac{1}{T} \left[s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2) \right] + a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2. \end{aligned}$$

Równości (11)-(13) są analogiczne do wzorów (7)-(8) i wynikają z własności kowariancji bezwarunkowej oraz założeń A1), B1), B4), C1)-C5), D1) i D4), natomiast wzór (14) można uzyskać analogicznie jak wzór (9) i (11), z wykorzystaniem własności bezwarunkowej kowariancji oraz założeń C5) i D4). Wzór (15) to ponownie zastosowanie własności wariancji bezwarunkowej oraz założeń B1), C1)-C3) oraz C5). I wreszcie równość (16) uzyskujemy na podstawie własności kowariancji bezwarunkowej, założenia B4) oraz wzorów (13) i (14). Zestawienie własności iteracyjnej wartości oczekiwanej oraz bezwarunkowej wariancji i kowariancji można znaleźć np. u Jasiulewicz [2005] w dodatku A.

□

Po obliczeniu parametrów warunkowych i bezwarunkowych rozkładów łącznej wartości szkód, a także wybranych wartości wariancji i kowariancji przejdziemy do obliczenia najlepszych liniowych predyktorów na podstawie liczby szkód, łącznej wartości szkód oraz obu tych zmiennych.

3. Najlepszy liniowy predyktor oparty na łącznej wartości szkód

Rozpocznijmy od zagadnienia liniowej predykcji łącznej wartości szkód w okresie $T + 1$ na podstawie dotychczasowych łącznych wartości szkód. Ponieważ zakładamy, że każdy z ubezpieczonych ponosi szkody w sposób niezależny od pozostałych, predyktor dla j -tego ubezpieczonego będzie oparty tylko na obserwacjach dotyczących tegoż ubezpieczonego. Co więcej, można łatwo pokazać, że przy przyjętych założeniach (chodzi tu zwłaszcza o założenie B3)) problem predykcji łącznej wartości szkód w okresie $T + 1$ sprowadza się do problemu predykcji $E(X_{j,T+1} | \Lambda_j, \Theta_j)$,

czyli najlepszy liniowy predyktor będzie rozwiązaniem następującego problemu minimalizacji:

$$\begin{aligned} & BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j) = \\ & = \arg \min_{b_{j,0}, \dots, b_{j,T}} E \left[\left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - b_{j,0} - \sum_{t=1}^T b_{j,t} X_{j,t} \right)^2 \right] = \arg \min_{b_{j,0}, \dots, b_{j,T}} S(b_{j,0}, \dots, b_{j,T}), \end{aligned}$$

gdzie $\mathbf{X}_j = (X_{j,1}, \dots, X_{j,T})$, a $S(\cdot)$ to funkcja, której minimum szukamy.

Twierdzenie 1.

Przy założeniach A-D) najlepszy liniowy predyktor łącznej wartości szkód w okresie $T + 1$ oparty na dotychczasowych łącznych wartościach szkód jest dany następującym wzorem:

$$BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j) = (1 - z_X) \Lambda \Theta + z_X \bar{X}_j, \quad (17)$$

gdzie:

$$z_X = \frac{Ta_{\Lambda}^2 a_{\Theta}^2 + Ta_{\Lambda}^2 \Theta^2 + Ta_{\Theta}^2 \Lambda^2}{Ta_{\Lambda}^2 a_{\Theta}^2 + Ta_{\Lambda}^2 \Theta^2 + Ta_{\Theta}^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_{\Theta}^2 + \Theta^2)}, \quad (18)$$

$$1 - z_X = \frac{s_Y^2 \Lambda + s_N^2 (a_{\Theta}^2 + \Theta^2)}{Ta_{\Lambda}^2 a_{\Theta}^2 + Ta_{\Lambda}^2 \Theta^2 + Ta_{\Theta}^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_{\Theta}^2 + \Theta^2)}. \quad (19)$$

Dowód:

Aby znaleźć rozwiązanie zadanego problemu minimalizacji, policzmy pierwsze pochodne z funkcji $S(b_{j,0}, \dots, b_{j,T})$ po interesujących nas parametrach $b_{j,0}, \dots, b_{j,T}$:

$$\frac{\partial S(b_{j,0}, \dots, b_{j,T})}{\partial b_{j,0}} = -E \left[2 \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - b_{j,0} - \sum_{t=1}^T b_{j,t} X_{j,t} \right) \right],$$

$$\frac{\partial S(b_{j,0}, \dots, b_{j,T})}{\partial b_{j,k}} = -E \left[2 \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - b_{j,0} - \sum_{t=1}^T b_{j,t} X_{j,t} \right) X_{j,k} \right], \quad k = 1, \dots, T.$$

Po przyrównaniu powyższych pochodnych do 0 otrzymamy następujący układ równań:

$$E(X_{j,T+1}) - b_{j,0} - \sum_{t=1}^T b_{j,t} E(X_{j,t}) = 0,$$

$$E(E(X_{j,T+1} | \Lambda_j, \Theta_j) X_{j,k}) - b_{j,0} E(X_{j,k}) - \sum_{t=1}^T b_{j,t} E(X_{j,t} X_{j,k}) = 0, \quad k = 1, \dots, T.$$

Jeśli pomnożymy pierwsze z równań przez $E(X_{j,k})$ i odejmiemy je od drugiego równania, powyższy układ równań będzie można zapisać następująco:

$$E(X_{j,T+1}) - b_{j,0} - \sum_{t=1}^T b_{j,t} E(X_{j,t}) = 0,$$

$$\text{Cov}(E(X_{j,T+1} | \Lambda_j, \Theta_j), X_{j,k}) = \sum_{t=1}^T b_{j,t} \text{Cov}(X_{j,t}, X_{j,k}), \quad k = 1, \dots, T.$$

Wykorzystując obliczenia z wzorów (2), (7) i (8), możemy powyższy układ zapisać równoważnie:

$$\Lambda\Theta - b_{j,0} - \Lambda\Theta \sum_{t=1}^T b_{j,t} = 0,$$

$$a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 = \sum_{t=1}^T b_{j,t} (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) + b_{j,k} (s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)),$$

$k = 1, \dots, T.$

Jak widać, dla każdego k mamy taką samą postać równania, co oznacza, że parametry $b_{j,1} = \dots = b_{j,T}$. Wobec tego powyższy układ równań można zapisać w następującej postaci:

$$\Lambda\Theta - b_{j,0} - \Lambda\Theta T b_{j,k} = 0,$$

$$a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 = b_{j,k} (T a_\Lambda^2 a_\Theta^2 + T a_\Lambda^2 \Theta^2 + T a_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)),$$

$k = 1, \dots, T,$

co prowadzi do następującego rozwiązania:

$$b_{j,k} = \frac{a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2}{T a_\Lambda^2 a_\Theta^2 + T a_\Lambda^2 \Theta^2 + T a_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)}, \quad k = 1, \dots, T$$

oraz:

$$\begin{aligned} b_{j,0} &= \Lambda\Theta - \Lambda\Theta T b_{j,k} = \Lambda\Theta \left[1 - \frac{T a_\Lambda^2 a_\Theta^2 + T a_\Lambda^2 \Theta^2 + T a_\Theta^2 \Lambda^2}{T a_\Lambda^2 a_\Theta^2 + T a_\Lambda^2 \Theta^2 + T a_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)} \right] = \\ &= \frac{s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)}{T a_\Lambda^2 a_\Theta^2 + T a_\Lambda^2 \Theta^2 + T a_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)} \Lambda\Theta. \end{aligned}$$

□

Z wzoru (17) wynika zatem, że najlepszy liniowy predyktor jest ważoną sumą bezwarunkowej wartości oczekiwanej łącznej wartości szkód ($\Lambda\Theta$) oraz elementu przybliżającego wartość oczekiwaną łącznej wartości szkód j -tego ubezpieczonego (\bar{X}_j). Taka postać predyktora to klasyczna postać predyktora liniowego, wyprowadzona np. przez Bühlmana [1967], oczywiście z uwzględnieniem faktu, że rozkład

łącznej wartości szkód zależy od dwóch parametrów ryzyka. Na podstawie wzoru (18) warto także zauważyć, że gdy historia ubezpieczonego się wydłuża, waga przypisywana średniej wartości ponoszonych przez niego szkód (z_X) zbiega do 1. A zatem w długim okresie składka zaufania będzie równa średniej wartości szkód danego ubezpieczonego.

W kolejnym twierdzeniu podany zostanie błąd średniokwadratowy predyktora.

Twierdzenie 2.

Błąd średniokwadratowy predyktora opartego na łącznej wartości szkód jest dany następującym równaniem:

$$MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j\right)\right) = (1 - z_X) \left[a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 \right]. \quad (20)$$

Dowód:

Wyjdźmy od następującego sposobu zapisu błędu średniokwadratowego predyktora:

$$\begin{aligned} & MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j\right)\right) = \\ & = E\left[\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) - BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j\right)\right)^2\right] = \\ & = E\left[\left(\Lambda_j \Theta_j - (1 - z_X) \Lambda \Theta - z_X \bar{X}_j\right)^2\right] = E\left[\left(\left(\Lambda_j \Theta_j - \Lambda \Theta\right) - z_X \left(\bar{X}_j - \Lambda \Theta\right)\right)^2\right] = \\ & = E\left[\left(\Lambda_j \Theta_j - \Lambda \Theta\right)^2 - 2z_X \left(\Lambda_j \Theta_j - \Lambda \Theta\right) \left(\bar{X}_j - \Lambda \Theta\right) + z_X^2 \left(\bar{X}_j - \Lambda \Theta\right)^2\right] = \\ & = Var\left(\Lambda_j \Theta_j\right) - 2z_X Cov\left(\Lambda_j \Theta_j, \bar{X}_j\right) + z_X^2 Var\left(\bar{X}_j\right). \end{aligned}$$

Korzystając z wzorów (5), (9) i (10), równanie na błąd średniokwadratowy możemy zapisać następująco:

$$\begin{aligned} & MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j\right)\right) = \\ & = a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - 2z_X \left(a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2\right) + \\ & + z_X^2 \left(\frac{1}{T} \left[s_Y^2 \Lambda + s_N^2 \left(a_\Theta^2 + \Theta^2\right)\right] + a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2\right). \quad (21) \end{aligned}$$

Na podstawie równania (18) można jednak zauważyć, że:

$$\frac{a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2}{z_X} = a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 + \frac{1}{T} \left[s_Y^2 \Lambda + s_N^2 \left(a_\Theta^2 + \Theta^2\right)\right]. \quad (22)$$

Po wstawieniu informacji z równania (22) do wzoru (21) i uproszczeniu otrzymujemy tezę twierdzenia. □

Na zakończenie tej części artykułu warto zauważyć, że dla wydłużającej się historii danego ubezpieczonego błąd średniokwadratowy predyktora będzie zbiegał do 0, ponieważ waga z_x zbiega do 1.

4. Najlepszy liniowy predyktor oparty na liczbie szkód

Zajmijmy się teraz zagadnieniem poszukiwania najlepszego liniowego predyktora łącznej wartości szkód na podstawie informacji na temat liczby szkód ubezpieczonego. W praktyce, jeśli dane dotyczące wartości szkód spływają z opóźnieniem, obliczenie predyktora na podstawie liczby szkód może być wygodnym rozwiązaniem. Tak jak w poprzedniej części artykułu interesuje nas znalezienie najlepszego liniowego predyktora łącznej wartości szkód w okresie $T + 1$; tak jak wcześniej dzięki założeniu, że każdy z ubezpieczonych ponosi szkody w sposób niezależny od pozostałych, do konstrukcji predyktora dla j -tego ubezpieczonego wystarczą obserwacje dotyczące liczby szkód tegoż ubezpieczonego; także tym razem przyjęte założenia (a szczególnie założenie B4)) sprowadzają problem liniowej predykcji łącznej wartości szkód w okresie $T + 1$ do problemu liniowej predykcji warunkowej wartości oczekiwanej $E(X_{j,T+1} | \Lambda_j, \Theta_j)$, który można zapisać w następujący sposób:

$$\begin{aligned} & BLP\left(E\left(X_{j,T+1} | \Lambda_j, \Theta_j\right) | \mathbf{N}_j\right) = \\ & = \arg \min_{c_{j,0}, \dots, c_{j,T}} E \left[\left(E\left(X_{j,T+1} | \Lambda_j, \Theta_j\right) - c_{j,0} - \sum_{t=1}^T c_{j,t} N_{j,t} \right)^2 \right] = \\ & = \arg \min_{c_{j,0}, \dots, c_{j,T}} S(c_{j,0}, \dots, c_{j,T}), \end{aligned}$$

gdzie $\mathbf{N}_j = (N_{j,1}, \dots, N_{j,T})$, a $S(\cdot)$ to funkcja, której minimum szukamy.

Twierdzenie 3.

Przy założeniach A)-D) najlepszy liniowy predyktor łącznej wartości szkód w okresie $T + 1$ na podstawie dotychczasowej liczby szkód jest dany wzorem:

$$BLP\left(E\left(X_{j,T+1} | \Lambda_j, \Theta_j\right) | \mathbf{N}_j\right) = (1 - z_N) \Lambda \Theta + z_N \bar{N}_j \Theta, \quad (23)$$

gdzie:

$$z_N = \frac{Ta_\Lambda^2}{Ta_\Lambda^2 + s_N^2}, \quad (24)$$

$$1 - z_N = \frac{s_N^2}{Ta_\Lambda^2 + s_N^2}. \quad (25)$$

Dowód:

Aby znaleźć rozwiązanie postawionego problemu minimalizacji, policzmy pierwsze pochodne z funkcji $S(c_{j,0}, \dots, c_{j,T})$ po szukanych parametrach $c_{j,0}, \dots, c_{j,T}$:

$$\frac{\partial S(c_{j,0}, \dots, c_{j,T})}{\partial c_{j,0}} = -E \left[2 \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - c_{j,0} - \sum_{t=1}^T c_{j,t} N_{j,t} \right) \right],$$

$$\frac{\partial S(c_{j,0}, \dots, c_{j,T})}{\partial c_{j,k}} = -E \left[2 \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - c_{j,0} - \sum_{t=1}^T c_{j,t} N_{j,t} \right) N_{j,k} \right], \quad k = 1, \dots, T.$$

Po przyrównaniu powyższych pochodnych do 0 otrzymujemy następujący układ równań:

$$E(X_{j,T+1}) - c_{j,0} - \sum_{t=1}^T c_{j,t} E(N_{j,t}) = 0,$$

$$E(E(X_{j,T+1} | \Lambda_j, \Theta_j) N_{j,k}) - c_{j,0} E(N_{j,k}) - \sum_{t=1}^T c_{j,t} E(N_{j,t} N_{j,k}) = 0, \quad k = 1, \dots, T.$$

Jeśli analogicznie jak w poprzedniej części artykułu pomnożymy pierwsze z równań przez $E(N_{j,k})$ i odejmiemy je od drugiego równania, otrzymamy następujący układ równań:

$$E(X_{j,T+1}) - c_{j,0} - \sum_{t=1}^T c_{j,t} E(N_{j,t}) = 0,$$

$$\text{Cov}(E(X_{j,T+1} | \Lambda_j, \Theta_j), N_{j,k}) = \sum_{t=1}^T c_{j,t} \text{Cov}(N_{j,t}, N_{j,k}), \quad k = 1, \dots, T.$$

Na podstawie przyjętych oznaczeń oraz wzorów (2), (11) i (12) możemy zapisać powyższy układ równań w następującej postaci:

$$\Lambda \Theta - c_{j,0} - \Lambda \sum_{t=1}^T c_{j,t} = 0,$$

$$\Theta a_\Lambda^2 = a_\Lambda^2 \sum_{t=1}^T c_{j,t} + c_{j,k} s_N^2, \quad k = 1, \dots, T.$$

Tak jak poprzednio dla każdego k mamy taką samą postać równania, co oznacza, że $c_{j,1} = \dots = c_{j,T}$. Wobec tego powyższy układ równań można zapisać w następującej postaci:

$$\Lambda \Theta - c_{j,0} - \Lambda T c_{j,k} = 0$$

$$\Theta a_\Lambda^2 = c_{j,k} (T a_\Lambda^2 + s_N^2), \quad k = 1, \dots, T,$$

co da nam następujący wynik:

$$c_{j,k} = \frac{a_\Lambda^2}{Ta_\Lambda^2 + s_N^2} \Theta, \quad k = 1, \dots, T,$$

oraz:

$$c_{j,0} = \Lambda \Theta \left(1 - \frac{Ta_\Lambda^2}{Ta_\Lambda^2 + s_N^2} \right) = \frac{s_N^2}{Ta_\Lambda^2 + s_N^2} \Lambda \Theta.$$

□

Warto porównać postać predyktora z wzoru (23) z predyktorem uzyskanym w poprzedniej części artykułu (wzór (17)). Uzyskany predyktor jest ponownie ważoną sumą dwóch składników: bezwarunkowej wartości oczekiwanej łącznej wartości szkód ($\Lambda \Theta$) oraz elementu $\bar{N}_j \Theta$, który podobnie jak \bar{X}_j ma przybliżać wartość oczekiwaną łącznej wartości szkód j -tego ubezpieczonego. Tak jak w przypadku predyktora opartego na łącznej wartości szkód w miarę wzrostu T waga przypisywana indywidualnej historii ubezpieczonego (z_N) zbiega do 1, a zatem w długim okresie predyktor będzie równy średniej liczbie szkód danego ubezpieczonego pomnożonej przez populacyjną wartość oczekiwaną wartości pojedynczej szkody ($\bar{N}_j \Theta$). Warto także zauważyć, że uzyskany predyktor jest równy predyktorowi liczby szkód w okresie $T + 1$ (obliczonemu na podstawie dotychczasowych liczb szkód) pomnożonemu przez populacyjną wartość oczekiwaną wartości pojedynczej szkody Θ .

Kolejne twierdzenie podaje wielkość błędu średniokwadratowego predyktora opartego na liczbie szkód.

Twierdzenie 4.

Błąd średniokwadratowy predyktora opartego na liczbie szkód jest równy:

$$MSE \left(BLP \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right) \mid \mathbf{N}_j \right) \right) = a_\Lambda^2 a_\Theta^2 + (1 - z_N) a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2. \quad (26)$$

Dowód:

Zacznijmy od zapisania błędu średniokwadratowego w następującej postaci:

$$\begin{aligned} MSE \left(BLP \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right) \mid \mathbf{N}_j \right) \right) &= E \left[\left(\Lambda_j \Theta_j - (1 - z_N) \Lambda \Theta - z_N \bar{N}_j \Theta \right)^2 \right] = \\ &= E \left[\left((\Lambda_j \Theta_j - \Lambda \Theta) - z_N (\bar{N}_j \Theta - \Lambda \Theta) \right)^2 \right] = \\ &= E \left[(\Lambda_j \Theta_j - \Lambda \Theta)^2 - 2z_N \Theta (\Lambda_j \Theta_j - \Lambda \Theta) (\bar{N}_j - \Lambda) + z_N^2 (\bar{N}_j \Theta - \Lambda \Theta)^2 \right] = \\ &= Var(\Lambda_j \Theta_j) - 2z_N \Theta Cov(\Lambda_j \Theta_j, \bar{N}_j) + (z_N \Theta)^2 Var(\bar{N}_j). \end{aligned}$$

Wstawiając informacje z wzorów (5), (14) i (15), otrzymamy, że:

$$\begin{aligned} &MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{N}_j\right)\right)= \\ &= a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - 2z_N a_\Lambda^2 \Theta^2 + z_N^2 \Theta^2 \left(\frac{1}{T} s_N^2 + a_\Lambda^2\right). \end{aligned}$$

Po wstawieniu do powyższego równania wzoru (24) zapisanego w następujący sposób:

$$\frac{a_\Lambda^2}{z_N} = a_\Lambda^2 + \frac{1}{T} s_N^2 \quad (27)$$

i uproszczeniu otrzymamy tezę twierdzenia. \square

Warto zauważyć, że wraz z wydłużaniem się historii ubezpieczonego błąd średniokwadratowy predyktora rozpatrywanego w tej części artykułu nie będzie zbiegał do 0, lecz utrzyma się na stałym poziomie równym $a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2$. Przemawiałoby to zatem na niekorzyść tego sposobu predykcji w porównaniu z predykcją na podstawie łącznej wartości szkód, z drugiej jednak strony historia ubezpieczonych rzadko jest na tyle długa, żeby waga z_x znalazła się odpowiednio blisko 1, a błąd średniokwadratowy w równaniu (20) stał się faktycznie bliski 0. Dokładniejsze porównanie predyktorów pod kątem błędów średniokwadratowych zostanie przeprowadzone w szóstym i siódmym punkcie artykułu, po wyprowadzeniu predyktora opartego na liczbie i wartości szkód oraz jego błędu średniokwadratowego.

5. Najlepszy liniowy predyktor oparty na łącznej wartości i liczbie szkód

Ostatnim wyprowadzonym w niniejszym artykule predyktorem łącznej wartości szkód w okresie $T + 1$ będzie predyktor oparty na liczbie i łącznej wartości szkód. Tak jak w poprzednich częściach, ponieważ zakładamy, że każdy z ubezpieczonych ponosi szkody w sposób niezależny od pozostałych ubezpieczonych z portfela, predyktor dla j -tego ubezpieczonego będzie oparty tylko na obserwacjach dotyczących tegoż ubezpieczonego. Ponadto, tak jak wcześniej, przyjęte założenia sprowadzają problem predykcji łącznej wartości szkód w okresie $T + 1$ do problemu predykcji warunkowej wartości oczekiwanej łącznej wartości szkód $E(X_{j,T+1} \mid \Lambda_j, \Theta_j)$. Problem predykcji można zatem zapisać następująco:

$$\begin{aligned} &BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j, \mathbf{N}_j\right)= \\ &= \arg \min_{\mathbf{d}_j, \mathbf{e}_j} E\left[\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) - d_{j,0} - \sum_{t=1}^T d_{j,t} X_{j,t} - \sum_{t=1}^T e_{j,t} N_{j,t}\right)^2\right]= \\ &= \arg \min_{\mathbf{d}_j, \mathbf{e}_j} S\left(\mathbf{d}_j, \mathbf{e}_j\right), \end{aligned}$$

gdzie $\mathbf{d}_j = (d_{j,0}, \dots, d_{j,T})$, $\mathbf{e}_j = (e_{j,1}, \dots, e_{j,T})$ oraz $S(\cdot)$ to funkcja, której minimum szukamy.

Twierdzenie 5.

Przy założeniach A)-D) najlepszy liniowy predyktor łącznej wartości szkód w okresie $T + 1$ oparty na liczbie i łącznej wartości szkód jest dany wzorem:

$$BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j, \mathbf{N}_j) = (1 - z_X^{X,N} - z_N^{X,N})\Lambda\Theta + z_X^{X,N}\bar{X}_j + z_N^{X,N}\bar{N}_j\Theta, \quad (28)$$

gdzie:

$$z_X^{X,N} = \frac{Ta_\Lambda^2 a_\Theta^2 + Ta_\Theta^2 \Lambda^2}{Ta_\Lambda^2 a_\Theta^2 + Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_\Theta^2}, \quad (29)$$

$$z_N^{X,N} = \frac{Ta_\Lambda^2 s_Y^2 \Lambda - Ta_\Theta^2 s_N^2 \Lambda^2}{(Ta_\Lambda^2 + s_N^2)(Ta_\Lambda^2 a_\Theta^2 + Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_\Theta^2)}, \quad (30)$$

$$1 - z_X^{X,N} - z_N^{X,N} = 1 - z_N = \frac{s_N^2}{Ta_\Lambda^2 + s_N^2}, \quad (31)$$

a z_N zostało zdefiniowane w równaniu (24).

Dowód:

Aby znaleźć rozwiązanie postawionego powyżej problemu minimalizacji, policzmy pierwsze pochodne z funkcji $S(\mathbf{d}_j, \mathbf{e}_j)$ po interesujących nas parametrach \mathbf{d}_j i \mathbf{e}_j :

$$\frac{\partial S(\mathbf{d}_j, \mathbf{e}_j)}{\partial d_{j,0}} = -E \left[2 \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - d_{j,0} - \sum_{t=1}^T d_{j,t} X_{j,t} - \sum_{t=1}^T e_{j,t} N_{j,t} \right) \right],$$

$$\frac{\partial S(\mathbf{d}_j, \mathbf{e}_j)}{\partial d_{j,k}} = -E \left[2 \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - d_{j,0} - \sum_{t=1}^T d_{j,t} X_{j,t} - \sum_{t=1}^T e_{j,t} N_{j,t} \right) X_{j,k} \right],$$

$k = 1, \dots, T,$

$$\frac{\partial S(\mathbf{d}_j, \mathbf{e}_j)}{\partial e_{j,l}} = -E \left[2 \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) - d_{j,0} - \sum_{t=1}^T d_{j,t} X_{j,t} - \sum_{t=1}^T e_{j,t} N_{j,t} \right) N_{j,l} \right],$$

$l = 1, \dots, T.$

Po przyrównaniu powyższych pochodnych do 0 otrzymujemy następujący układ równań:

$$E(X_{j,T+1}) - d_{j,0} - \sum_{t=1}^T d_{j,t} E(X_{j,t}) - \sum_{t=1}^T e_{j,t} E(N_{j,t}) = 0,$$

$$\begin{aligned}
& E\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) X_{j,k}\right) \\
& -d_{j,0}E\left(X_{j,k}\right) - \sum_{t=1}^T d_{j,t}E\left(X_{j,t} X_{j,k}\right) - \sum_{t=1}^T e_{j,t}E\left(N_{j,t} X_{j,k}\right) = 0, \quad k = 1, \dots, T, \\
& E\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) N_{j,l}\right) - d_{j,0}E\left(N_{j,l}\right) \\
& - \sum_{t=1}^T d_{j,t}E\left(X_{j,t} N_{j,l}\right) - \sum_{t=1}^T e_{j,t}E\left(N_{j,t} N_{j,l}\right) = 0, \quad l = 1, \dots, T.
\end{aligned}$$

Jeśli pomnożymy pierwsze równanie przez $E\left(X_{j,k}\right)$ i odejmiemy je od drugiego równania oraz jeśli pomnożymy pierwsze równanie przez $E\left(N_{j,l}\right)$ i odejmiemy je od równania trzeciego, otrzymamy następujący układ równań:

$$\begin{aligned}
& E\left(X_{j,T+1}\right) - d_{j,0} - \sum_{t=1}^T d_{j,t}E\left(X_{j,t}\right) - \sum_{t=1}^T e_{j,t}E\left(N_{j,t}\right) = 0, \\
& \text{Cov}\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right), X_{j,k}\right) = \\
& = \sum_{t=1}^T d_{j,t} \text{Cov}\left(X_{j,t}, X_{j,k}\right) + \sum_{t=1}^T e_{j,t} \text{Cov}\left(N_{j,t}, X_{j,k}\right), \quad k = 1, \dots, T, \\
& \text{Cov}\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right), N_{j,l}\right) = \\
& = \sum_{t=1}^T d_{j,t} \text{Cov}\left(X_{j,t}, N_{j,l}\right) + \sum_{t=1}^T e_{j,t} \text{Cov}\left(N_{j,t}, N_{j,l}\right), \quad l = 1, \dots, T.
\end{aligned}$$

Po wstawieniu odpowiednich wartości z wzorów (2), (7), (8), (11)-(13) otrzymamy:

$$\begin{aligned}
& \Lambda\Theta - d_{j,0} - \Lambda\Theta \sum_{t=1}^T d_{j,t} - \Lambda \sum_{t=1}^T e_{j,t} = 0, \\
& a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 = \sum_{t=1}^T d_{j,t} \left(a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 \right) + d_{j,k} \left[s_Y^2 \Lambda + s_N^2 \left(a_\Theta^2 + \Theta^2 \right) \right] + \\
& \quad + \sum_{t=1}^T e_{j,t} \Theta a_\Lambda^2 + e_{j,k} \Theta s_N^2, \quad k = 1, \dots, T, \\
& \Theta a_\Lambda^2 = \sum_{t=1}^T d_{j,t} \Theta a_\Lambda^2 + d_{j,l} \Theta s_N^2 + \sum_{t=1}^T e_{j,t} a_\Lambda^2 + e_{j,t} s_N^2, \quad l = 1, \dots, T.
\end{aligned}$$

Jak widać, dla każdego k i dla każdego l mamy taką samą postać równania, co oznacza, że parametry $d_{j,1} = \dots = d_{j,T}$ oraz $e_{j,1} = \dots = e_{j,T}$. Wobec tego powyższy układ równań można zapisać w następującej postaci:

$$\Lambda\Theta - d_{j,0} - \Lambda\Theta T d_{j,k} - \Lambda T e_{j,k} = 0, \quad (32)$$

$$a_{\Lambda}^2 a_{\Theta}^2 + a_{\Lambda}^2 \Theta^2 + a_{\Theta}^2 \Lambda^2 = d_{j,k} \left(T a_{\Lambda}^2 a_{\Theta}^2 + T a_{\Lambda}^2 \Theta^2 + T a_{\Theta}^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_{\Theta}^2 + \Theta^2) \right) + e_{j,k} \left(T \Theta a_{\Lambda}^2 + \Theta s_N^2 \right), \quad (33)$$

$$\Theta a_{\Lambda}^2 = d_{j,k} \left(T \Theta a_{\Lambda}^2 + \Theta s_N^2 \right) + e_{j,k} \left(T a_{\Lambda}^2 + s_N^2 \right), \quad (34)$$

$k = 1, \dots, T$.

Jeśli przekształcimy równanie (34) w następujący sposób:

$$e_{j,k} \left(T \Theta a_{\Lambda}^2 + \Theta s_N^2 \right) = \Theta^2 a_{\Lambda}^2 - d_{j,k} \left(T \Theta^2 a_{\Lambda}^2 + \Theta^2 s_N^2 \right) \quad (35)$$

i wstawimy do równania (33), otrzymamy następującą postać równania:

$$a_{\Lambda}^2 a_{\Theta}^2 + a_{\Lambda}^2 \Theta^2 + a_{\Theta}^2 \Lambda^2 = d_{j,k} \left(T a_{\Lambda}^2 a_{\Theta}^2 + T a_{\Lambda}^2 \Theta^2 + T a_{\Theta}^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_{\Theta}^2 + \Theta^2) \right) + \Theta^2 a_{\Lambda}^2 - d_{j,k} \left(T \Theta^2 a_{\Lambda}^2 + \Theta^2 s_N^2 \right),$$

co prowadzi do następującego rozwiązania dla $d_{j,k}$:

$$d_{j,k} = \frac{a_{\Lambda}^2 a_{\Theta}^2 + a_{\Theta}^2 \Lambda^2}{T a_{\Lambda}^2 a_{\Theta}^2 + T a_{\Theta}^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_{\Theta}^2}.$$

Po wstawieniu powyżej obliczonej wartości $d_{j,k}$ do równania (34) i przekształceniu otrzymamy rozwiązanie:

$$e_{j,k} = \frac{a_{\Lambda}^2 s_Y^2 \Lambda - a_{\Theta}^2 s_N^2 \Lambda^2}{\left(T a_{\Lambda}^2 + s_N^2 \right) \left(T a_{\Lambda}^2 a_{\Theta}^2 + T a_{\Theta}^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_{\Theta}^2 \right)} \Theta.$$

Po wstawieniu powyższych rozwiązań dla $d_{j,k}$ i $e_{j,k}$ do równania (32) otrzymamy rozwiązanie dla stałej $d_{j,0}$:

$$d_{j,0} = \Lambda \Theta \left[1 - T d_{j,k} - T \frac{e_{j,k}}{\Theta} \right] = \frac{s_N^2}{T a_{\Lambda}^2 + s_N^2} \Lambda \Theta.$$

□

Z wzoru (28) wynika, że najlepszy liniowy predyktor jest ważoną sumą bezwarunkowej wartości oczekiwanej łącznej wartości szkód ($\Lambda \Theta$), średniej wartości szkód oraz średniej liczby szkód pomnożonej przez populacyjną wartość oczekiwaną wartości pojedynczej szkody. Nietrudno dostrzec, że waga $z_N^{X,N}$ może przyjmować wartości ujemne, a dokładniej dzieje się tak, gdy:

$$a_{\Theta}^2 s_N^2 \Lambda \geq a_{\Lambda}^2 s_Y^2$$

lub równoważnie:

$$E \left(E(N_{j,t} | \Lambda_j) \right) \frac{Var \left(E(Y_{j,t,k} | \Theta_j) \right)}{E \left(Var(Y_{j,t,k} | \Theta_j) \right)} \geq \frac{Var \left(E(N_{j,t} | \Lambda_j) \right)}{E \left(Var(N_{j,t} | \Lambda_j) \right)}.$$

Co więcej, dla rosnącej liczby okresów T waga $z_N^{X,N}$ zbiega do 0, podczas gdy waga przy średniej wartości szkód $z_X^{X,N}$ zbiega do 1. Warto jeszcze zauważyć, że z równania (34) wynika, że:

$$z_N^{X,N} = z_N - z_X^{X,N}, \quad (36)$$

co z kolei prowadzi do wniosku, że predyktor oparty na liczbie i wartości szkód można zapisać w następującej postaci:

$$\begin{aligned} BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j, \mathbf{N}_j\right) &= (1 - z_N) \Lambda \Theta + z_N \bar{N}_j \Theta + z_X^{X,N} (\bar{X}_j - \bar{N}_j \Theta) = \\ &= BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{N}_j\right) + z_X^{X,N} (\bar{X}_j - \bar{N}_j \Theta). \end{aligned} \quad (37)$$

Jak zatem widać, predyktor oparty na liczbie i łącznej wartości szkód to predyktor oparty na liczbie szkód skorygowany o ważoną różnicę między dwoma sposobami przybliżania wartości oczekiwanej łącznej wartości szkód j -tego ubezpieczonego, czyli o element $z_X^{X,N} (\bar{X}_j - \bar{N}_j \Theta)$.

Podajmy jeszcze wzór na błąd średniokwadratowy predyktora dwuczynnikowego:

Twierdzenie 6.

Błąd średniokwadratowy predyktora opartego na liczbie i łącznej wartości szkód wynosi:

$$\begin{aligned} MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j, \mathbf{N}_j\right)\right) &= \\ &= (1 - z_X^{X,N}) (a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2) + (1 - z_N) a_\Lambda^2 \Theta^2 = \end{aligned} \quad (38)$$

$$= (1 - z_X^{X,N}) (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) - z_N^{X,N} a_\Lambda^2 \Theta^2. \quad (39)$$

Dowód:

Obliczenie błędu średniokwadratowego w przypadku predyktora opartego na liczbie i wartości szkód będzie nieco bardziej żmudne niż w przypadku pozostałych predyktorów:

$$\begin{aligned} MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j, \mathbf{N}_j\right)\right) &= \\ &= E\left[\left(\Lambda_j \Theta_j - (1 - z_X^{X,N} - z_N^{X,N}) \Lambda \Theta - z_X^{X,N} \bar{X}_j - z_N^{X,N} \bar{N}_j \Theta\right)^2\right] = \\ &= E\left[\left((\Lambda_j \Theta_j - \Lambda \Theta) - z_X^{X,N} (\bar{X}_j - \Lambda \Theta) - z_N^{X,N} (\bar{N}_j \Theta - \Lambda \Theta)\right)^2\right] = \\ &= E\left[\left(\Lambda_j \Theta_j - \Lambda \Theta\right)^2 - 2z_X^{X,N} (\Lambda_j \Theta_j - \Lambda \Theta)(\bar{X}_j - \Lambda \Theta)\right] \end{aligned}$$

$$\begin{aligned}
& -2z_N^{X,N} (\Lambda_j \Theta_j - \Lambda \Theta) (\bar{N}_j \Theta - \Lambda \Theta) + (z_X^{X,N})^2 (\bar{X}_j - \Lambda \Theta)^2 + \\
& + 2z_X^{X,N} z_N^{X,N} (\bar{X}_j - \Lambda \Theta) (\bar{N}_j \Theta - \Lambda \Theta) + (z_N^{X,N})^2 (\bar{N}_j \Theta - \Lambda \Theta)^2 \Big] = \\
& = \text{Var}(\Lambda_j \Theta_j) - 2z_X^{X,N} \text{Cov}(\Lambda_j \Theta_j, \bar{X}_j) - 2z_N^{X,N} \Theta \text{Cov}(\Lambda_j \Theta_j, \bar{N}_j) + (z_X^{X,N})^2 \text{Var}(\bar{X}_j) + \\
& + 2z_X^{X,N} z_N^{X,N} \Theta \text{Cov}(\bar{X}_j, \bar{N}_j) + (z_N^{X,N} \Theta)^2 \text{Var}(\bar{N}_j).
\end{aligned}$$

Po wstawieniu wartości podanych we wzorach (5), (9), (10), (14)-(16) otrzymamy, że błąd średniokwadratowy predyktora wynosi:

$$\begin{aligned}
& \text{MSE} \left(\text{BLP} \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j, \mathbf{N}_j \right) \right) = \\
& = a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - 2z_X^{X,N} (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) - 2z_N^{X,N} a_\Lambda^2 \Theta^2 + \\
& + (z_X^{X,N})^2 \left(\frac{1}{T} [s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)] + a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 \right) + \\
& + 2z_X^{X,N} z_N^{X,N} \Theta^2 \left(\frac{1}{T} s_N^2 + a_\Lambda^2 \right) + (z_N^{X,N} \Theta)^2 \left(\frac{1}{T} s_N^2 + a_\Lambda^2 \right). \tag{40}
\end{aligned}$$

Dla uproszczenia obliczeń możemy do równania (40) wstawić informację z równania (36) – po przekształceniu otrzymamy wtedy następującą postać równania:

$$\begin{aligned}
& \text{MSE} \left(\text{BLP} \left(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j, \mathbf{N}_j \right) \right) = \\
& = a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - 2z_X^{X,N} (a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2) - 2z_N^{X,N} a_\Lambda^2 \Theta^2 + \\
& + (z_X^{X,N})^2 \left(\frac{1}{T} [s_Y^2 \Lambda + s_N^2 a_\Theta^2] + a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2 \right) + z_N^2 \Theta^2 \left(\frac{1}{T} s_N^2 + a_\Lambda^2 \right).
\end{aligned}$$

Po wstawieniu do powyższego równania wzoru (29) zapisanego w następującej postaci:

$$\frac{a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2}{z_X^{X,N}} = a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2 + \frac{1}{T} [s_Y^2 \Lambda + s_N^2 a_\Theta^2] \tag{41}$$

oraz po skorzystaniu z informacji zawartych w równaniu (27) i uproszczeniu otrzymamy tezę twierdzenia wyrażoną za pomocą wzoru (38). Aby uzyskać postać wzoru z równania (39), trzeba dodatkowo skorzystać z równości (36).

□

Na zakończenie tej części artykułu warto zauważyć, że z równania (38) wynika, iż błąd średniokwadratowy predyktora dwuczynnikowego będzie zbiegał do 0, ponieważ wagi $z_X^{X,N}$ i z_N zbiegają do 1.

6. Porównanie błędów średniokwadratowych predyktorów

Dokonajmy teraz porównania błędów średniokwadratowych obliczonych predyktorów, aby wiedzieć, który z predyktorów jest najbardziej precyzyjny.

Twierdzenie 7.

Dla uzyskanych w twierdzeniach 1, 3 i 5 predyktorów zachodzą następujące nierówności:

$$\begin{aligned} & MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{N}_j\right)\right) \geq \\ & \geq MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j, \mathbf{N}_j\right)\right), \end{aligned} \quad (42)$$

$$\begin{aligned} & MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j\right)\right) \geq \\ & \geq MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j, \mathbf{N}_j\right)\right). \end{aligned} \quad (43)$$

Ponadto jeśli spełniona jest nierówność:

$$\begin{aligned} & \left(a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2\right) \left[a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2 + a_\Lambda^2 \Theta^2\right] \left(T a_\Lambda^2 + s_N^2\right) \geq \\ & \geq a_\Lambda^2 \Theta^2 \Lambda \left(a_\Lambda^2 s_Y^2 - a_\Theta^2 s_N^2 \Lambda\right), \end{aligned} \quad (44)$$

to zachodzi:

$$\begin{aligned} & MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{N}_j\right)\right) \geq \\ & \geq MSE\left(BLP\left(E\left(X_{j,T+1} \mid \Lambda_j, \Theta_j\right) \mid \mathbf{X}_j\right)\right) \end{aligned} \quad (45)$$

Dowód:

Nierówność (42) wynika z bezpośredniego porównania błędu średniokwadratowego predyktora opartego na liczbie szkód i predyktora opartego na liczbie i łącznej wartości szkód (równania (26) i (38)):

$$a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - z_N a_\Lambda^2 \Theta^2 \geq a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - z_X^{X,N} \left(a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2\right) - z_N a_\Lambda^2 \Theta^2.$$

Nieco trudniej pokazać zachodzenie nierówności (43), którą zgodnie z równaniami (20) i (39) można zapisać w następującej postaci:

$$\begin{aligned} & a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - z_X \left(a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2\right) \geq \\ & \geq a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - z_X^{X,N} \left(a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2\right) - z_N^{X,N} a_\Lambda^2 \Theta^2 \end{aligned}$$

lub równoważnie:

$$z_X^{X,N} (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) + z_N^{X,N} a_\Lambda^2 \Theta^2 \geq z_X (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2). \quad (46)$$

Na początku warto zauważyć, że po odpowiednim przekształceniu równania (33) z wykorzystaniem definicji odpowiednich wag oraz informacji z równań (22), (27) i (36) wagę $z_X^{X,N}$ można zapisać jako następującą funkcję wagi z_X :

$$z_X^{X,N} (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) = z_X (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) - z_N^{X,N} z_X \Theta^2 \frac{a_\Lambda^2}{z_N}.$$

Wykorzystując powyższy zapis, nierówność (46) można sprowadzić do następującej postaci:

$$z_X (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) - z_N^{X,N} z_X \Theta^2 \frac{a_\Lambda^2}{z_N} + z_N^{X,N} a_\Lambda^2 \Theta^2 \geq z_X (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2),$$

co z kolei upraszcza się do:

$$z_N^{X,N} (z_N - z_X) \geq 0.$$

Należy teraz sprawdzić, kiedy powyższa nierówność zachodzi. Z wcześniejszych rozważań wiemy już, że $z_N^{X,N} \geq 0$, jeśli zachodzi:

$$a_\Lambda^2 s_Y^2 \geq a_\Theta^2 s_N^2 \Lambda. \quad (47)$$

Zastanówmy się jeszcze, kiedy $z_N - z_X \geq 0$. Wstawiając definicję wag z równań (18) i (24), otrzymamy, że musi zachodzić:

$$\begin{aligned} Ta_\Lambda^2 (Ta_\Lambda^2 a_\Theta^2 + Ta_\Lambda^2 \Theta^2 + Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)) &\geq \\ &\geq (Ta_\Lambda^2 + s_N^2) (Ta_\Lambda^2 a_\Theta^2 + Ta_\Lambda^2 \Theta^2 + Ta_\Theta^2 \Lambda^2), \end{aligned}$$

co upraszcza się ponownie do:

$$a_\Lambda^2 s_Y^2 \geq a_\Theta^2 s_N^2 \Lambda.$$

To kończy dowód drugiej części twierdzenia.

Postać warunku (44) dotyczącego nierówności (45) wynika z porównania błędów średniokwadratowych predyktora opartego na liczbie szkód oraz predyktora opartego na łącznej wartości szkód podanych w równaniach (20) i (26):

$$a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - z_N a_\Lambda^2 \Theta^2 \geq a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2 - z_X (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2),$$

co można zapisać równoważnie w następującej postaci:

$$z_X (a_\Lambda^2 a_\Theta^2 + a_\Lambda^2 \Theta^2 + a_\Theta^2 \Lambda^2) \geq z_N a_\Lambda^2 \Theta^2. \quad (48)$$

Gdy wykorzystamy naszą wiedzę o postaci wag z_X i z_N z równań (18) i (24), nierówność (48) możemy zapisać następująco:

$$\begin{aligned} & \left(a_{\Lambda}^2 a_{\Theta}^2 + a_{\Lambda}^2 \Theta^2 + a_{\Theta}^2 \Lambda^2 \right)^2 \left(a_{\Lambda}^2 + \frac{1}{T} s_N^2 \right) \geq \\ & \geq \left(\Theta a_{\Lambda}^2 \right)^2 \left(a_{\Lambda}^2 a_{\Theta}^2 + a_{\Lambda}^2 \Theta^2 + a_{\Theta}^2 \Lambda^2 + \frac{1}{T} \left[s_Y^2 \Lambda + s_N^2 (a_{\Theta}^2 + \Theta^2) \right] \right), \end{aligned} \quad (49)$$

co po skróceniu pewnych elementów i uporządkowaniu daje nierówność (44). \square

Na podstawie twierdzenia 7 możemy zatem stwierdzić, że predyktor dwuczynnikowy jest zawsze nie gorszy niż predyktor oparty na łącznej wartości szkód. Predyktor dwuczynnikowy jest także nie gorszy niż predyktor oparty na liczbie szkód, a różnica w błędach średniokwadratowych tych dwóch predyktorów wraz ze wzrostem liczby okresów T będzie rosła, aż ustabilizuje się na poziomie $a_{\Lambda}^2 a_{\Theta}^2 + a_{\Theta}^2 \Lambda^2$.

W przypadku predyktorów jednoczynnikowych nie można wskazać jednoznacznie lepszego predyktora. Warto jednak spróbować zapisać warunek (44) w nieco innej postaci, która pozwoliłaby w bardziej intuicyjny sposób odpowiedzieć na pytanie, w jakich sytuacjach lepiej posługiwać się predyktorem opartym na łącznej wartości szkód. Na wstępie zauważmy, że korzystając z równań (9), (10), (14) i (15), warunek (49) można zapisać w następującej postaci:

$$\frac{\left(\text{Cov} \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right), \bar{X}_j \right) \right)^2}{\text{Var} \left(\bar{X}_j \right)} \geq \frac{\left(\text{Cov} \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right), \bar{N}_j \right) \right)^2}{\text{Var} \left(\bar{N}_j \right)}.$$

Jeśli podzielimy obie strony nierówności przez dodatnią $\text{Var} \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right) \right)$ i spierwiastkujemy, otrzymamy:

$$\frac{\text{Cov} \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right), \bar{X}_j \right)}{\sqrt{\text{Var} \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right) \right) \text{Var} \left(\bar{X}_j \right)}} \geq \frac{\text{Cov} \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right), \bar{N}_j \right)}{\sqrt{\text{Var} \left(E \left(X_{j,T+1} \mid \Lambda_j, \Theta_j \right) \right) \text{Var} \left(\bar{N}_j \right)}},$$

co oznacza, że jeśli korelacja wartości, której predykcji dokonujemy, ze średnią wartością łącznej wartości szkód jest większa niż korelacja między wartością, której predykcji dokonujemy, a średnią wartością liczby szkód, powinniśmy zastosować predyktor oparty na łącznej wartości szkód. Nietrudno także zauważyć, że warunek (44) będzie spełniony, gdy:

$$\Lambda \frac{s_N^2}{s_Y^2} \geq \frac{a_{\Lambda}^2}{a_{\Theta}^2}, \quad (50)$$

ponieważ lewa strona nierówności (44) jest zawsze większa od 0. Warunek (50) pojawił się już wcześniej w odwrotnej postaci jako warunek (47) – zachodzenie warunku w postaci z równania (50) oznacza, że waga $z_N^{X,N} \leq 0$ oraz że $z_X \geq z_N$.

Co ciekawe, pewnej wskazówki odnośnie do zachodzenia nierówności (44) może także dostarczyć porównanie wag $z_X^{X,N}$ i $z_N^{X,N}$. Sprawdźmy najpierw, kiedy zachodzi $z_X^{X,N} \geq z_N^{X,N}$:

$$\begin{aligned} z_X^{X,N} &= \frac{Ta_\Lambda^2 a_\Theta^2 + Ta_\Theta^2 \Lambda^2}{Ta_\Lambda^2 a_\Theta^2 + Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_\Theta^2} \geq \\ &\geq \frac{Ta_\Lambda^2 s_Y^2 \Lambda - Ta_\Theta^2 s_N^2 \Lambda^2}{(Ta_\Lambda^2 + s_N^2)(Ta_\Lambda^2 a_\Theta^2 + Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_\Theta^2)} = z_N^{X,N}. \end{aligned} \quad (51)$$

Nierówność (51) można zapisać równoważnie w następujący sposób:

$$(a_\Lambda^2 a_\Theta^2 + a_\Theta^2 \Lambda^2)(Ta_\Lambda^2 + s_N^2) \geq \Lambda(a_\Lambda^2 s_Y^2 - a_\Theta^2 s_N^2 \Lambda). \quad (52)$$

Jak zatem widać, zachodzenie nierówności (52) oznacza, że jest spełniona także nierówność (44), lub, równoważnie, jeżeli $z_X^{X,N} \geq z_N^{X,N}$, wtedy na pewno predyktor oparty na łącznej wartości szkód ma mniejszy błąd średniokwadratowy niż predyktor oparty na liczbie szkód. Oczywiście istnieją jeszcze sytuacje, w których nierówności (50) i (52) nie są spełnione, a nierówność (44) zachodzi, ciężko jest jednak sformułować jakieś ogólne wnioski (o ile założymy, że wszystkie parametry są większe od zera).

Na zakończenie warto jeszcze zauważyć, że nierówność w postaci (44) potwierdza wcześniejszy wniosek mówiący, że dla odpowiednio dużego T predyktor oparty na łącznej wartości szkód będzie na pewno lepszy niż predyktor oparty na liczbie szkód. Jak już jednak zostało wspomniane, wniosek ten może w praktyce mieć mniejsze znaczenie ze względu na fakt, że zazwyczaj obserwujemy ubezpieczonych przez krótki okres.

7. Przypadki szczególne

Przyjęte w drugiej części artykułu założenia mówią o istnieniu dwóch niezależnych parametrów ryzyka mających decydujący wpływ na proces powstawania szkód. Założenia te prowadzą do uzależnienia postaci predyktorów od dużej liczby parametrów, ze względów praktycznych warto zatem rozważyć możliwe uproszczenia postaci predyktorów wynikające z ograniczenia analizy do tylko jednego parametru ryzyka. Być może w praktyce w pewnych sytuacjach możliwe jest przyjęcie założenia, że rozkład wartości pojedynczej szkody lub liczby szkód jest taki sam dla wszystkich ubezpieczonych. Zastanówmy się zatem, jak przyjęcie owych upraszczających założeń wpłynie na postać predyktorów oraz na wielkość ich błędów średniokwadratowych.

Na początek przyjmijmy, że rozkład wartości pojedynczej szkody jest taki sam dla wszystkich ubezpieczonych – innymi słowy sprowadza się to do założenia, że:

$$D1') E(Y_{j,t,k} | \Theta_j) = E(Y_{j,t,k}) = \Theta,$$

$$D2') \text{Var}(Y_{j,t,k} | \Theta_j) = \text{Var}(Y_{j,t,k}) = s_Y^2,$$

$$D3') E(Y_{j,t,k}) = E(\Theta_j) = \Theta,$$

$$D4') \text{Var}(\Theta_j) = 0.$$

Sprawdźmy, jak przyjęcie powyższych założeń wpłynie na postaci predyktorów uzyskane w twierdzeniach 1, 3 i 5.

Twierdzenie 8.

Przy założeniach A)-C) oraz D') predyktor oparty na liczbie szkód i predyktor oparty na liczbie i łącznej wartości szkód mają taką samą postać daną równaniem:

$$\begin{aligned} BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j, \mathbf{N}_j) &= (1 - z_X^{X,N} - z_N^{X,N}) \Lambda \Theta + z_X^{X,N} \bar{X}_j + z_N^{X,N} \bar{N}_j \Theta = \\ &= BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{N}_j), \end{aligned}$$

gdzie $z_X^{X,N} = 0$, $z_N^{X,N} = z_N = \frac{Ta_\Lambda^2}{Ta_\Lambda^2 + s_N^2}$ oraz $1 - z_X^{X,N} - z_N^{X,N} = 1 - z_N = \frac{s_N^2}{Ta_\Lambda^2 + s_N^2}$.

Predyktor oparty na łącznej wartości szkód jest natomiast równy:

$$BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j) = (1 - z_X) \Lambda \Theta + z_X \bar{X}_j,$$

gdzie $z_X = \frac{Ta_\Lambda^2 \Theta^2}{Ta_\Lambda^2 \Theta^2 + s_Y^2 \Lambda + s_N^2 \Theta^2} = \frac{Ta_\Lambda^2}{Ta_\Lambda^2 + s_N^2 + s_Y^2 \Lambda / \Theta^2}$

oraz $1 - z_X = \frac{s_Y^2 \Lambda + s_N^2 \Theta^2}{Ta_\Lambda^2 \Theta^2 + s_Y^2 \Lambda + s_N^2 \Theta^2}$.

Dowód:

Twierdzenie 8 można uzyskać, wstawiając oznaczenia z założeń D') do wzorów podanych w twierdzeniach 1, 3 i 5. □

A zatem, jeżeli rozkład wartości pojedynczej szkody nie jest zróżnicowany między ubezpieczonymi, nie ma różnicy w sposobach przybliżania wartości oczekiwanej łącznej wartości szkód j -tego ubezpieczonego w predykcji na podstawie liczby szkód i w predykcji na podstawie liczby i wartości szkód, a element $(\bar{X}_j - \bar{N}_j \Theta)$ z równania (37) ma zerową wagę. Co istotne, predyktor dwuczynnikowy jest nadal lepszy niż predyktor oparty tylko na łącznej wartości szkód, co oznacza, że w tym przypadku predyktor oparty na liczbie szkód będzie zawsze lepszy niż predyktor

oparty na łącznej wartości szkód – można tu dodać, że zmodyfikowany do obecnych założeń warunek (44) ma postać: $\Lambda s_Y^2 (a_\Lambda^2 \Theta)^2 \leq 0$, co nigdy nie jest spełnione, jeśli przyjmiemy, że rozkład liczby szkód jest zróżnicowany między ubezpieczonymi. Przewaga predykcji opartej na liczbie szkód nad predykcją opartą na łącznej wartości szkód wynika przede wszystkim z faktu, iż waga z_X jest zbudowana na podstawie wariancji średniej wartości łącznej wartości szkód, która jest większa niż wariancja średniej liczby szkód uwzględniona w wadze z_N .

W analogiczny sposób możemy rozpatrzyć sytuację, gdy rozkład liczby szkód jest taki sam dla wszystkich ubezpieczonych. Założenie to prowadzi do następującej postaci momentów rozkładu liczby szkód i parametru ryzyka Λ_j :

$$C1') E(N_{j,t} | \Lambda_j) = E(N_{j,t}) = \Lambda,$$

$$C2') Var(N_{j,t} | \Lambda_j) = Var(N_{j,t}) = s_N^2,$$

$$C3') E(\Lambda_j) = \Lambda,$$

$$C4') Var(\Lambda_j) = 0.$$

W takim przypadku predykcja na podstawie liczby szkód właściwie traci sens, a predyktor upraszcza się do populacyjnej wartości oczekiwanej łącznej wartości szkód. Co ciekawe, predyktor dwuczynnikowy można nadal przedstawić w postaci takiej jak w równaniu (37), czyli jako sumę predyktora opartego na liczbie szkód i ważonej poprawki $z_X^{X,N} (\bar{X}_j - \bar{N}_j \Theta)$. W przypadku predyktora opartego na łącznej wartości szkód ponownie zmieni się tylko postać wagi z_X . Twierdzenie 9 stanowi podsumowanie powyższych wniosków.

Twierdzenie 9.

Przy założeniach A)-B), C') oraz D) predyktor oparty na liczbie szkód ma postać daną równaniem:

$$BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{N}_j) = (1 - z_N) \Lambda \Theta + z_N \bar{N}_j \Theta = \Lambda \Theta,$$

gdzie $z_N = 0$, $1 - z_N = 1$.

Predyktor dwuczynnikowy będzie równy:

$$\begin{aligned} BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j, \mathbf{N}_j) &= (1 - z_X^{X,N} - z_N^{X,N}) \Lambda \Theta + z_X^{X,N} \bar{X}_j + z_N^{X,N} \bar{N}_j \Theta = \\ &= BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{N}_j) + z_X^{X,N} (\bar{X}_j - \bar{N}_j \Theta), \end{aligned}$$

gdzie $z_X^{X,N} = \frac{Ta_\Theta^2 \Lambda^2}{Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_\Theta^2}$, $z_N^{X,N} = -\frac{Ta_\Theta^2 \Lambda^2}{Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 a_\Theta^2} = -z_X^{X,N}$,

$$1 - z_X^{X,N} - z_N^{X,N} = 1.$$

Predyktor oparty na łącznej wartości będzie natomiast miał postać:

$$BLP(E(X_{j,T+1} | \Lambda_j, \Theta_j) | \mathbf{X}_j) = (1 - z_X) \Lambda \Theta + z_X \bar{X}_j,$$

$$\text{gdzie } z_X = \frac{Ta_\Theta^2 \Lambda^2}{Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)}, \quad 1 - z_X = \frac{s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)}{Ta_\Theta^2 \Lambda^2 + s_Y^2 \Lambda + s_N^2 (a_\Theta^2 + \Theta^2)}.$$

Dowód:

Analogicznie jak w przypadku twierdzenia 8 dowód można przeprowadzić, wstawiając do twierdzeń 1, 3 i 5 oznaczenia z założeń C').

□

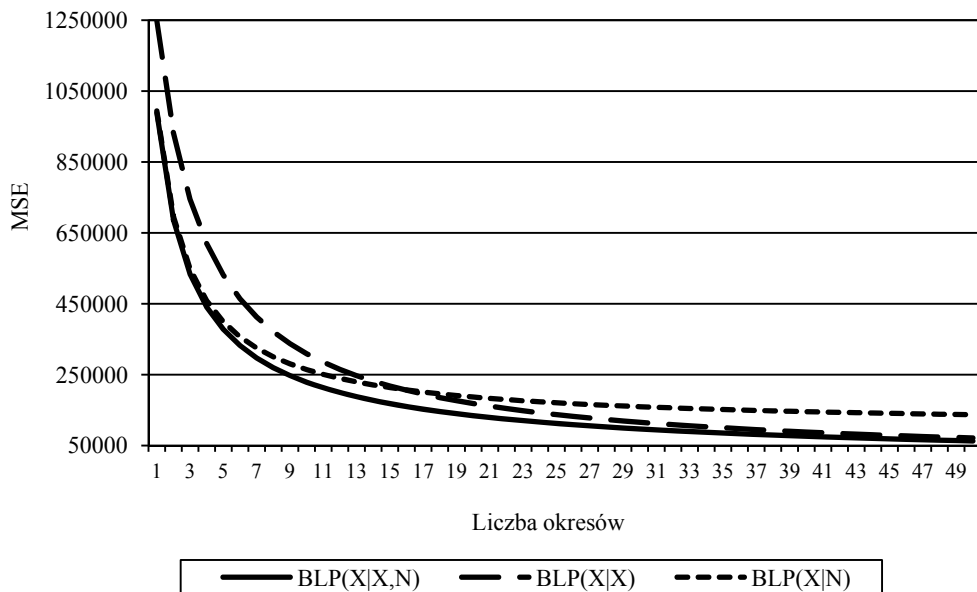
Ujemna wartość wagi $z_N^{X,N}$ w predyktorze dwuczynnikowym może wydać się zaskakująca – taka jej wartość wynika z faktu, że przy predykcji łącznej wartości szkód w elemencie \bar{X}_j uwzględniamy nie tylko informację o wartościach pojedynczych szkód, lecz także informację na temat liczby szkód, która w przypadku gdy rozkłady liczby szkód są takie same dla wszystkich ubezpieczonych, stanowi zbędny szum informacyjny. Uwzględnienie w predyktorze elementu $\bar{N}_j \Theta$ z ujemną wagą powoduje zatem, że do predykcji wykorzystujemy tylko informacje faktycznie świadczące o zróżnicowaniu parametrów ryzyka między ubezpieczonymi.

Na zakończenie warto dodać, że przy obecnie przyjętych założeniach najgorszym predyktorem będzie oczywiście predyktor oparty na liczbie szkód, a predyktorem najlepszym będzie nadal predyktor dwuczynnikowy.

8. Przykład numeryczny

Aby zilustrować zagadnienie optymalnej predykcji wartości szkód na podstawie liczby i łącznej wartości szkód, w niniejszej części artykułu zostanie zaprezentowany przykład numeryczny. Dla pewnych założonych wartości parametrów rozkładów liczby szkód i wartości pojedynczej szkody zostaną obliczone wagi predyktorów opartych na liczbie szkód i łącznej wartości szkód dla rosnącej liczby okresów T , a także błędy średniokwadratowe tych predyktorów.

Przyjmijmy następujące wartości parametrów rozkładów: $\Lambda = s_N^2 = a_\Lambda^2 = 0,6$, $\Theta = 1724,14$, $a_\Theta^2 = 106166,13$, $s_Y^2 = 3078817,7$. Parametry zostały dobrane w taki sposób, aby obserwowane rozkłady liczby i łącznej wartości szkód miały momenty możliwe do zaobserwowania w praktyce oraz jednocześnie, aby w początkowych latach historii ubezpieczonego predykcja na podstawie liczby szkód charakteryzowała się mniejszym błędem średniokwadratowym niż predykcja oparta na łącznej wartości szkód. Ponadto przyjęte wartości Λ , s_N^2 oraz a_Λ^2 odpowiadają sytuacji, gdy parametr ryzyka rozkładu liczby szkód ma rozkład gamma z parametrem kształtu



Rys. 1. Wartości błędów średniokwadratowych dla predyktora dwuczynnikowego, predyktora opartego na łącznej wartości szkód oraz predyktora opartego na liczbie szkód dla rosnącej liczby okresów

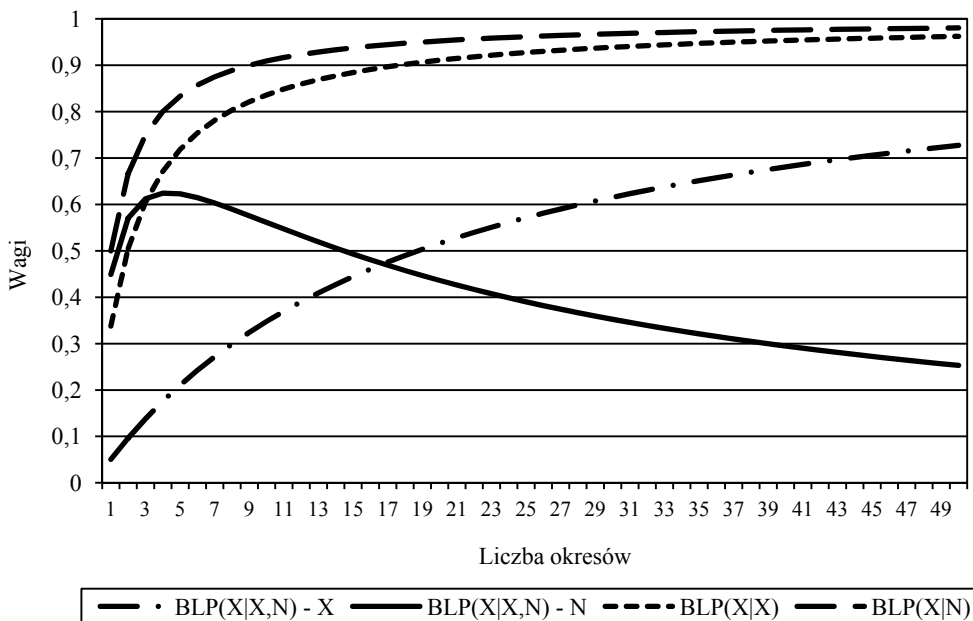
Źródło: obliczenia własne.

równym 1 i parametrem skali równym 0,6, a warunkowy rozkład liczby szkód przy znanej wartości parametru ryzyka to rozkład Poissona¹.

Na rysunku 1 znajdują się wartości błędów średniokwadratowych dla trzech analizowanych predyktorów dla rosnącej liczby okresów: $T = 1, \dots, 50$. Jak już zostało wspomniane, dobór parametrów rozkładów spowodował, że predyktor oparty na liczbie szkód ma początkowo niższy błąd średniokwadratowy niż predyktor oparty na łącznej wartości szkód. W szesnastym roku historii ubezpieczonego jednakże predyktor oparty na łącznej wartości szkód zaczyna mieć przewagę nad predyktorem opartym na liczbie szkód. Ponadto zgodnie z wcześniejszymi rozważaniami predyktor dwuczynnikowy jest we wszystkich okresach lepszy od pozostałych predyktorów.

Na kolejnym wykresie (rys. 2) znajdują się wartości wag predyktorów obliczane dla rosnącej liczby okresów. Zgodnie z wyciągniętymi wcześniej wnioskami wagi z_N , z_X oraz $z_N^{X,N}$ zbiegają do jedynki, natomiast waga $z_N^{X,N}$ zbiega do zera. Można ponadto zauważyć, że waga $z_N^{X,N}$ jest większa od zera, a także $z_N > z_X$, co oznacza, że nierówność (50) nie jest spełniona i że nie można na tej podstawie wyciągać wniosków na temat relacji między błędami średniokwadratowymi predyktora opar-

¹ Złożony rozkład Poissona-gamma jest często stosowany do modelowania liczby szkód ubezpieczonych, por. np. Denuit i in. [2007].



Rys. 2. Wartości wag stosowanych w predyktorze dwuczynnikowym, predyktorze opartym na łącznej wartości szkód oraz predyktorze opartym na liczbie szkód dla rosnącej liczby okresów

Źródło: obliczenia własne.

tego na łącznej wartości szkód i predyktora opartego na liczbie szkód. Od siedemnastego okresu historii ubezpieczonego zachodzi natomiast $z_X^{X,N} \geq z_N^{X,N}$, co oznacza, że od tego momentu spełniona jest nierówność (52) lub, równoważnie, że błąd średniokwadratowy predyktora opartego na liczbie szkód jest większy niż błąd predyktora opartego na łącznej wartości szkód (co znajduje także potwierdzenie na rys. 1).

9. Podsumowanie

Niniejszy artykuł poświęcony został zagadnieniu liniowej predykcji łącznej wartości szkód ubezpieczonego na podstawie liczby i łącznej wartości przeszłych szkód. Z przeprowadzonych obliczeń wynika, że predyktorem o najmniejszym błędzie średniokwadratowym jest predyktor uwzględniający informacje zarówno o liczbie, jak i o łącznej wartości szkód. W pewnych sytuacjach jednakże stosowanie predyktorów jednoczynnikowych może być uzasadnione – dotyczy to np. sytuacji, gdy rozkład wartości pojedynczej szkody nie jest zróżnicowany między ubezpieczonymi.

Literatura

- Bühlmann H. (1967), *Experience rating and credibility*, „ASTIN Bulletin“ vol. 4, no 3.
- Denuit M., Marechal X., Pitrebois S., Walhin J.-F. (2007), *Actuarial Modelling of Claim Counts*, John Wiley & Sons Ltd., Chichester.
- Goulet V., Forgues A., Lu J. (2006), *Credibility for severity revisited*, „North American Actuarial Journal“ vol. 10, no 1.
- Jasiulewicz H. (2005), *Teoria zaufania. Modele aktuarialne*, AE, Wrocław.

CALCULATION OF CREDIBILITY PREMIUM ON THE BASIS OF NUMBER AND TOTAL AMOUNT OF CLAIMS

Summary: In this article we consider the issue of calculation of best linear predictors for claim amount in the next period. We present premiums calculated on the basis of the number of claims, on the basis of claim amount or based on both of these variables. We also calculate and compare mean square errors of predictors and show which predictor is best. Moreover, we consider consequences of certain simplifying assumptions on risk parameters for the premiums and their mean square errors. Finally we present simple numerical example illustrating main ideas of this article.

Key words: credibility premium, claim frequency and severity.