

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

Helena Jasiulewicz

Uniwersytet Przyrodniczy we Wrocławiu

Wojciech Kordecki

Wyższa Szkoła Handlowa we Wrocławiu

SKŁADKI ZAUFANIA Z ZASTOSOWANIEM NIESYMETRYCZNYCH FUNKCJI STRAT

Streszczenie: W pracy przedstawione są niesymetryczne funkcje strat (LINEX, Stein, Entropy) w kontekście ubezpieczeń. Następnie omówione zostały wyznaczone estymatory składki netto z zastosowaniem niesymetrycznej funkcji straty LINEX. Przeprowadzono analizę otrzymanych estymatorów oraz porównano je z estymatorami składki wyznaczonymi przy użyciu kwadratowej funkcji straty.

Słowa kluczowe: składka zaufania, funkcja straty LINEX, estymator bayesowski.

1. Wstęp

Teoria zaufania dostarcza technik wyznaczania przyszłych składek ubezpieczeniowych dla kontraktów, które należą do bardziej lub mniej niejednorodnego portfela. W takich portfelach wiedza o kontrakcie jest ograniczona (skromna), ale o całym portfelu jest duża. Teoria zaufania jest jedną z teorii dostarczających ilościowych narzędzi, które pozwalają dopasować przyszłe składki, opierając się na wiedzy historycznej polisy i wiedzy o całym portfelu. Składka taka, zwana składką zaufania, jest wypukłą kombinacją średniej arytmetycznej \bar{X} z bieżącej obserwacji polisy X_1, \dots, X_n i średniej *a priori* portfela m , tzn.

$$z\bar{X} + (1-z)m,$$

gdzie współczynnik zaufania $z \in (0, 1)$.

Do wyznaczania składki zaufania korzysta się z osiągnięć statystyki bayesowskiej. Teorii zaufania z wykorzystaniem kwadratowej funkcji straty poświęcone są monografie Bühlmana i Gislera [2005] i Jasiulewicz [2005]. W artykule zajmujemy się składkami zaufania przy niesymetrycznej funkcji straty.

W pkt 2 omówiono najczęściej spotykane niesymetryczne funkcje straty. W pkt 3 podano postać estymatora bayesowskiego przy funkcji straty LINEX oraz przykład liniowego estymatora bayesowskiego dla zadanego rozkładu.

Główne wyniki zawarte są w pkt 4. W modelu Bühlmana przy założeniu, że zmienna losowa $(m(\Theta), \bar{X})$ ma rozkład normalny, podano postać liniowego estymatora bayesowskiego składki zaufania.

W dodatku przedstawiono własności warunkowej wartości oczekiwanej, warunkowej wariancji i warunkowej kowariancji, które wraz z dowodami można znaleźć np. w pracy Jasiulewicz [2005].

2. Niesymetryczne funkcje straty

W teorii podejmowania decyzji najczęściej przyjmowana jest symetryczna kwadratowa funkcja straty $L(x) = L_{SOR}(x) = x^2$, gdzie x jest wartością błędu. Oznacza to, że przeszacowanie i niedoszacowanie nieznanego parametru o tę samą wartość dla podejmowania decyzji ma takie samo znaczenie. Jednak od wielu lat pojawiają się prace naukowe wskazujące na potrzebę stosowania niesymetrycznej funkcji straty w odniesieniu do podejmowania decyzji w świecie realnym.

Po raz pierwszy ten problem zauważył Varian [1974], badając straty ponoszone przy niewłaściwej wycenie nieruchomości. Wykazał, że niedoszacowanie realnego majątku daje w przybliżeniu liniową stratę dochodu, podczas gdy przeszacowanie majątku prowadzi do strat w przybliżeniu wykładniczych, spowodowanych różnego rodzaju sporami sądowymi i odwołaniami. Przeszacowanie jest poważniejsze niż niedoszacowanie majątku.

Innym ważnym i aktualnym przykładem jest niedoszacowanie i przeszacowanie poziomu lustra wody zbiorników retencyjnych o tę samą wartość, które prowadzi do różnych konsekwencji ekonomicznych i społecznych (zob. [Zellner 1986]). W konstrukcji zbiornika wodnego niedoszacowanie poziomu lustra wody jest zwykle bardziej poważne niż przeszacowanie górnego poziomu wody.

Również w ubezpieczeniach niedoszacowanie lub przeszacowanie składki o tę samą wielkość może prowadzić do różnych konsekwencji finansowych dla ubezpieczyciela. Gdy ubezpieczony jest niedoszacowany, ubezpieczyciel traci swoje pieniądze, podczas gdy przeszacowana jest składka, ubezpieczyciel może stracić polisę. Pierwszy zwrócił uwagę na ten problem w ubezpieczeniach Ferreira [1977]. Niesymetryczną funkcję straty dla systemu Bonus-Malus stosowali Lemaire [1979], De-nuit i Dhaene [2001], Bermúdez i in. [2001], Pérez i in. [2002]. Z polskich autorów należy wspomnieć o pracach Niemirowicza [2006] i Boratyńskiej [2008].

Wszystkie powyższe przykłady przemawiają za tym, żeby funkcje straty związane z estymacją parametrów lub prognozą opisanych w nich zjawisk były takiej postaci, aby stosować bardziej surowe kary dla przeszacowania niż dla niedoszacowania lub odwrotnie. Więcej egzemplifikacji lub odsyłaczy do przykładów wskazujących na zastosowanie niesymetrycznych funkcji straty jako alternatywy dla symetrycznej kwadratowej funkcji straty można znaleźć w pracy [Parsian, Kirmani 2002].

Rodzinę niesymetrycznych funkcji strat zdefiniowali Thomson i Basu [1996]. Funkcja straty $L(x)$ z tej rodziny, gdzie x jest wartością błędu, ma własności:

- 1) $L(0) = 0$,
- 2) $L(x) > 0$ dla $x \neq 0$,
- 3) $L(x)$ jest wypukła,
- 4) $L(x)$ jest rosnąca z jednej strony zera i malejąca z drugiej strony zera,
- 5) szybkości zmian wartości funkcji $L(x)$ po obu stronach zera są różne.

Zwykle jako alternatywę do kwadratowej funkcji straty bierze się niesymetryczną funkcję straty

$$L(\theta, d) = b(e^{-c(\theta-d)} + c(\theta-d) - 1), \quad (1)$$

gdzie θ jest nieznanym parametrem, a d – wartością estymatora tego parametru. Funkcję tę nazywa się LINEX (Linear-Exponential), bo z jednej strony zera jest prawie liniowa, a z drugiej prawie wykładnicza. Parametr $c \neq 0$ jest parametrem kształtu, a parametr $b > 0$ jest parametrem skali i na ogół zakłada się, że $b = 1$. Znak parametru c odzwierciedla kierunek asymetrii. Jeżeli $c > 0$, to przeszacowanie jest bardziej poważne w skutkach niż niedoszacowanie. Jeśli $c < 0$, to niedoszacowanie rodzi większe konsekwencje niż przeszacowanie. Im większe $|c|$, tym bardziej funkcja LINEX jest asymetryczna.

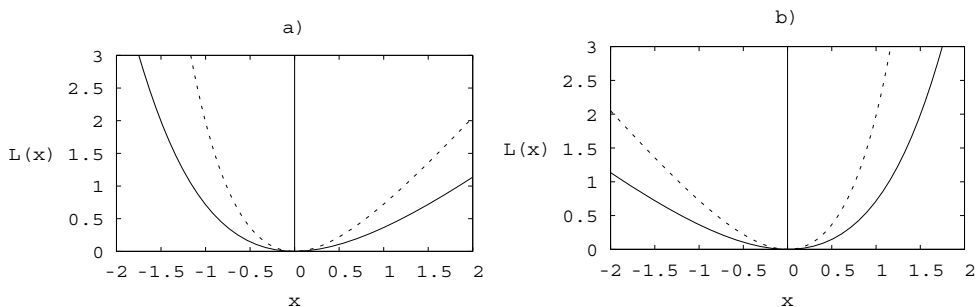
Niech $x = \theta - d$. Wówczas $L(x) = L_{LINEX}(x) = e^{-cx} + cx - 1$. Dla małych c funkcja LINEX jest prawie symetryczna, przy czym dla $c \rightarrow 0$

$$2L_{LINEX}(x) \sim L_{SQR}(x),$$

tzn.

$$\frac{2L_{LINEX}(x)}{L_{SQR}(x)} \rightarrow 1,$$

gdzie $L_{SQR}(x)$ oznacza kwadratową funkcję straty.



Rys. 1. Wykresy funkcji $L_{LINEX}(x)$: a) $c = 1$ – linia ciągła, $c = 1.5$ – linia przerywana, b) $c = -1$ – linia ciągła, $c = -1.5$ – linia przerywana

Źródło: opracowanie własne.

Na rysunku 1 przedstawione są wykresy $L(x)$ dla $c=1$, $c=1.5$, $c=-1$, $c=-1.5$. Szczegółowe badanie własności funkcji LINEX przy ich zastosowaniu do estymacji można znaleźć np. w pracy [Parsian, Kirmani 2002].

Błąd estymacji można mierzyć nie tylko odległością różnicy $x = \theta - d$ od zera, ale również odległością ilorazu $x = \theta/d$ od jedynki. W tym drugim przypadku należy określić rodzinę funkcji strat nieco inaczej, niż to zrobili Thomson i Basu [1996]. Mianowicie funkcja straty $L(x)$, $x > 0$ musi spełniać następujące własności:

- 1) $L(1) = 0$,
- 2) $L(x) > 0$ dla $x \neq 1$,
- 3) $L(x)$ jest rosnąca z jednej strony jedynki i malejąca z drugiej strony jedynki,
- 4) szybkości zmian wartości funkcji $L(x)$ po obu stronach jedynki są różne.

Przykłady takich funkcji strat są podane poniżej.

Funkcja straty ENTROPY jest określona wzorem:

$$L(\theta, d) = \frac{d}{\theta} - \ln \frac{d}{\theta} - 1 = x^{-1} - \ln x^{-1} - 1, \quad (2)$$

gdzie $x = \theta/d$. Funkcja ta nie jest funkcją wypukłą i ma punkt przegięcia dla $x = 2$.

Funkcja straty STEIN jest określona wzorem:

$$L(\theta, d) = \frac{\theta}{d} - \ln \frac{\theta}{d} - 1 = x - \ln x - 1, \quad (3)$$

gdzie $x = \theta/d$. Funkcja ta jest wypukła.

Funkcje straty określone wzorami (2) i (3) można zapisać ogólniej jednym wzorem

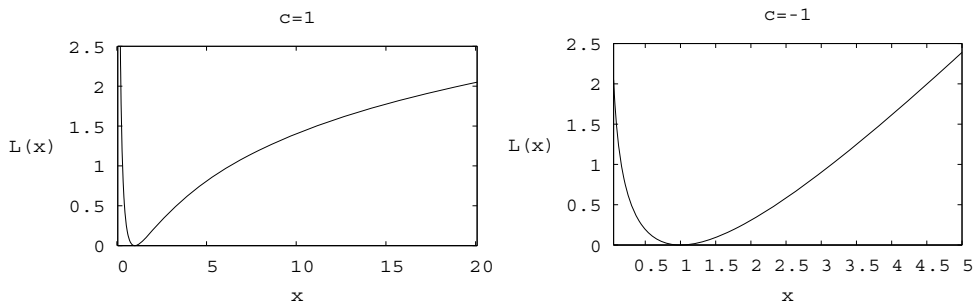
$$L(\theta, d) = L_{ENT}(x) = x^{-c} - \ln x^{-c} - 1, \quad (4)$$

gdzie $x = \theta/d$, $c \neq 0$. Dla $c=1$ otrzymujemy funkcję ENTROPY określoną przez (2), a dla $c=-1$ funkcję STEIN określoną przez (3). Funkcja straty ENTROPY jest szczególnym przypadkiem funkcji ilości informacji Kullbacka-Leiblera (zob. np. [Parsian, Nematollahi 1996]). Zauważmy też, że dla $x > 0$

$$L_{ENT}(x) = L_{LINEX}(\ln x). \quad (5)$$

Tak samo jak dla funkcji LINEX, jeżeli $c > 0$, to przeszacowanie jest bardziej poważne w skutkach niż niedoszacowanie. Jeśli $c < 0$, to niedoszacowanie rodzi większe konsekwencje niż przeszacowanie. Na rysunku 2 znajdują się wykresy funkcji dla $c=1$ i $c=-1$.

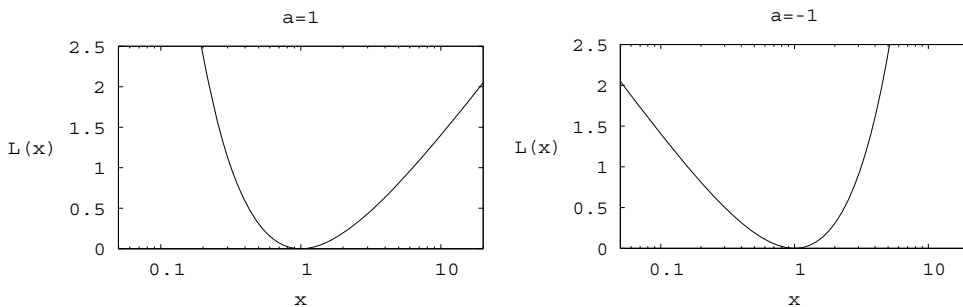
Symetrię analogiczną jak dla funkcji straty LINEX, ale nie dla różnicy $\theta - d$, tylko dla ilorazu θ/d , widać wyraźniej, gdy wykresy są przedstawione w skali logarytmicznej dla x . Wykresy takie są przedstawione na rys. 3.



Rys. 2. Wykresy funkcji $L_{ENT}(x)$

Źródło: opracowanie własne.

Rys. 3. Wykresy funkcji $L_{ENT}(x)$ w skali logarymicznej dla x



Źródło: opracowanie własne.

3. Estymator bayesowski przy funkcji straty LINEX

Niech zmienna losowa Θ opisuje strukturę ryzyka w badanym portfelu. Parametr θ będący realizacją zmiennej losowej Θ reprezentuje cechy ubezpieczonego wpływające na jego liczbę roszczeń i kwotę roszczeń. Oczywiście parametr θ jest nieznan i właściwy danemu ubezpieczonemu. Informacje o θ zawarte są w obserwacjach ubezpieczonego. Niech X_1, \dots, X_n oznaczają całkowitą liczbę roszczeń lub całkowitą kwotę roszczeń w kolejnych latach ubezpieczenia dla ubezpieczonego o parametrze ryzyka θ . Zmienne losowe X_1, \dots, X_n są zależne przez wspólny parametr θ , ale ich rozkłady warunkowe przy danym θ są niezależne, tzn.

$$\Pr(X_1 < x_1, \dots, X_n < x_n | \Theta = \theta) = \Pr(X_1 < x_1 | \theta) \dots \Pr(X_n < x_n | \theta)$$

dla każdego $x_1, \dots, x_n \in R$. Warunkowa wartość oczekiwana $m_i(\theta) = E(X_i | \Theta = \theta)$ jest składką netto w i -tym roku dla ubezpieczonego o parametrze ryzyka θ . Składka ta jest nieznaną, bo nieznanym jest parametr θ .

Niech $\delta = \delta(X_1, \dots, X_n)$ będzie estymatorem parametru $m(\theta)$. Szukamy takiego estymatora δ_{exp} , żeby wartość oczekiwana funkcji straty LINEX była najmniejsza przy ograniczeniu takim, aby równowaga finansowa ubezpieczyciela była zachowana.

Zatem

$$E L_{LINEX}(m_{n+1}(\Theta), \delta_{exp}) = \min_{\delta} E L_{LINEX}(m_{n+1}(\Theta), \delta), \quad (6)$$

przy ograniczeniu

$$E m_i(\Theta) = E X_i \quad (7)$$

dla $i = 1, 2, \dots$, gdzie

$$L_{LINEX}(m(\theta), \delta) = e^{-c(m(\theta) - \delta)} + c(m(\theta) - \delta) - 1. \quad (8)$$

Wartość oczekiwana w (6) jest określona dla rozkładu łącznego $(\Theta, X_1, \dots, X_n)$. Przy ograniczeniu (7) zagadnienie optymalizacji (6) sprowadza się do szukania δ_{exp} takiego, żeby

$$E e^{-c(m_{n+1}(\Theta) - \delta_{exp})} = \min_{\delta} E e^{-c(m_{n+1}(\Theta) - \delta)}. \quad (9)$$

Twierdzenie 1. Dla funkcji straty LINEX minimum wartości oczekiwanej straty

$$E L_{LINEX}(m_{n+1}(\Theta), \delta) = E e^{-c(m_{n+1}(\Theta) - \delta)} + c(m_{n+1}(\Theta) - \delta) - 1$$

w zbiorze funkcji mierzalnych $\delta: R^n \rightarrow R$ spełniających ograniczenie

$$E m_i(\Theta) = E X_i = m_i \quad (10)$$

dla $i = 1, 2, \dots$ jest osiągnięte dla

$$\begin{aligned} \delta_{exp}(X_1, \dots, X_n) = m_{n+1} + \frac{1}{c} E \left[\ln E \left(e^{-c m_{n+1}(\Theta)} \mid X_1, \dots, X_n \right) \right] \\ - \frac{1}{c} \ln E \left(e^{-c m_{n+1}(\Theta)} \mid X_1, \dots, X_n \right). \end{aligned} \quad (11)$$

Dowód. Zauważmy, że przy ograniczeniu (10) oczekiwana strata

$$\begin{aligned} E \left(e^{-c(m_{n+1}(\Theta) - \delta(X_1, \dots, X_n))} + c(m_{n+1}(\Theta) - \delta(X_1, \dots, X_n)) - 1 \right) \\ = E e^{-c(m_{n+1}(\Theta) - \delta(X_1, \dots, X_n))} - 1, \end{aligned}$$

osiąga minimum w tym samym punkcie, w którym minimum osiąga

$$E e^{-c(m_{n+1}(\Theta) - \delta(X_1, \dots, X_n))}. \quad (12)$$

W tym miejscu powołamy się na wynik z pracy [Bermúdez i in. 2001]. Pokazano w nim, że minimum w (12) jest osiągnięte dla $\delta_{exp}(X_1, \dots, X_n)$ danego wzorem (11). \square

Jeżeli estymator bayesowski $\delta_{exp}(X_1, \dots, X_n)$ składki netto jest wypukłą kombinacją średniej z obserwacji i średniej portfela, czyli

$$\delta_{exp}(X_1, \dots, X_n) = z_{exp} \bar{X} + (1 - z_{exp}) m, \quad (13)$$

gdzie $z_{exp} \in (0, 1)$, to mówimy, że $\delta_{exp}(X_1, \dots, X_n)$ jest składką zaufania przy funkcji straty LINEX.

Estymator bayesowski dany wzorem (11) zostanie wyznaczony dla zmiennej losowej Θ o rozkładzie gamma ze średnią α/β i wariancją α/β^2 oraz zmiennej losowej X pod warunkiem $\Theta = \theta$ o rozkładzie Poissona z parametrem θ . Oczywiście $m(\theta) = E(X | \theta) = \theta$.

Estymator składki netto Θ wyznaczony ze wzoru (11) jest postaci:

$$\delta_{exp}(X_1, \dots, X_n) = z_{exp} \bar{X} + (1 - z_{exp}) \frac{\alpha}{\beta}, \quad (14)$$

gdzie współczynnik zaufania jest równy

$$z_{exp} = \frac{n}{c} \ln \left(1 + \frac{c}{\beta + n} \right) \quad (15)$$

(zob. [Bermúdez i in. 2001]).

Dobrze jest znany fakt, że minimalna strata przy funkcji kwadratowej jest osiągnięta dla estymatora δ_{sqr} będącego wartością oczekiwaną rozkładu *a posteriori*, tzn.

$$\delta_{sqr}(X_1, \dots, X_n) = E(\Theta | X_1, \dots, X_n). \quad (16)$$

Przy założeniach, że zmienna Θ ma rozkład gamma z parametrami α i β , a zmienna X przy warunku $\Theta = \theta$ ma rozkład Poissona z parametrem θ , wzór (16) ma postać

$$\delta_{sqr}(X_1, \dots, X_n) = z_{sqr} \bar{X} + (1 - z_{sqr}) \frac{\alpha}{\beta}, \quad (17)$$

gdzie współczynnik zaufania wyraża się wzorem

$$z_{sqr} = \frac{n}{n + \beta}. \quad (18)$$

Łatwo zauważyć, że (zob. [Bermúdez i in. 2001])

$$z_{exp} = \frac{n}{c} \ln \left(1 + \frac{c}{n + \beta} \right) \leq \frac{n}{n + \beta} = z_{sqr}.$$

Oznacza to, że przy funkcji straty LINEX mniejszą wagę przykładają się do historii niż przy kwadratowej funkcji straty. Warto wspomnieć o wynikach tej pracy dotyczących rozważanego przykładu. Mianowicie

$$\lim_{c \rightarrow 0} \delta_{exp} = \delta_{sqr}$$

oraz

$$\lim_{c \rightarrow \infty} z_{exp} = 0$$

tak, że

$$\delta_{exp} \rightarrow \frac{\alpha}{\beta}.$$

Ponadto pochodna

$$\frac{dz_{exp}}{dc} < 0.$$

Oznacza to, że gdy parametr c rośnie, to waga przykładana do historii ubezpieczonego maleje przy funkcji straty LINEX.

Korzystając ze wzorów (14) i (18), podamy przedziały $(x_L(\theta), x_U(\theta))$ takie, że

$$\Pr(x_L < \theta - \delta < x_U) \geq 1 - \gamma = 0.95, \quad (19)$$

dla danych realizacji θ . Ponieważ rozkład nie jest symetryczny, to przyjmujemy

$$\Pr(\theta - \delta > x_L) \geq 1 - \frac{\gamma}{2}, \quad (20)$$

$$\Pr(\theta - \delta < x_U) \geq 1 - \frac{\gamma}{2}, \quad (21)$$

dla estymatorów $\delta = \delta_{sqr}$ i $\delta = \delta_{exp}$.

Niech z oznacza z_{exp} lub z_{sqr} . Wtedy wzór (19) ma postać

$$\Pr\left(x_L < \theta - \left(z\bar{X} + (1-z)\frac{\alpha}{\beta}\right) < x_U\right) = 1 - \gamma. \quad (22)$$

Najpierw wyznaczmy x_L . Ze wzoru (20) otrzymujemy

$$\begin{aligned} & \Pr\left(x_L - \theta + (1-z)\frac{\alpha}{\beta} < -z\bar{X}\right) \\ &= \Pr\left(n\bar{X} < \frac{n\theta - (1-z)\frac{\alpha}{\beta} - x_L}{z}\right) \geq 1 - \frac{\gamma}{2}. \end{aligned}$$

Ponieważ $n\bar{X}$ ma rozkład Poissona z parametrem $\lambda = n\theta$, to przyjmując

$$v_L = \frac{n\theta - (1-z)\frac{\alpha}{\beta} - x_L}{z},$$

wyznamy v_L ze związku

$$\sum_{k=0}^{\lfloor v_L \rfloor} \frac{\lambda^k}{k!} \geq \left(1 - \frac{\gamma}{2}\right) e^\lambda,$$

gdzie $\lfloor x \rfloor$ oznacza część całkowitą liczby x . Stąd

$$x_L = \theta - (1-z) \frac{\alpha}{\beta} - \frac{z v_L}{n}.$$

Analogicznie wyznaczmy v_U ze związku

$$\sum_{k=0}^{\lfloor v_U \rfloor} \frac{\lambda^k}{k!} \geq \frac{\gamma}{2} e^\lambda,$$

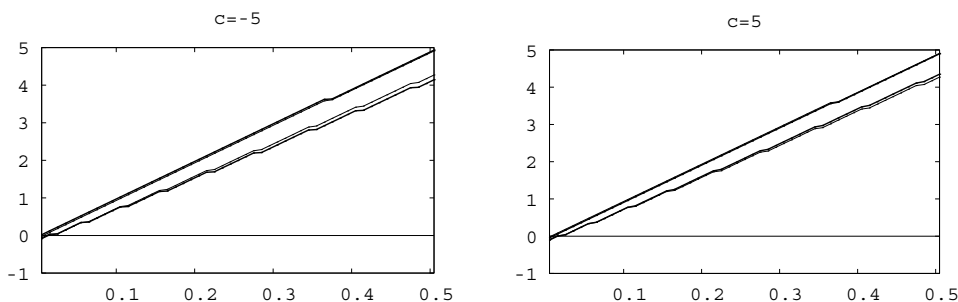
gdzie

$$v_U = \frac{n\theta - (1-z) \frac{\alpha}{\beta} - x_U}{z}.$$

Stąd

$$x_U = \theta - (1-z) \frac{\alpha}{\beta} - \frac{z v_U}{n}.$$

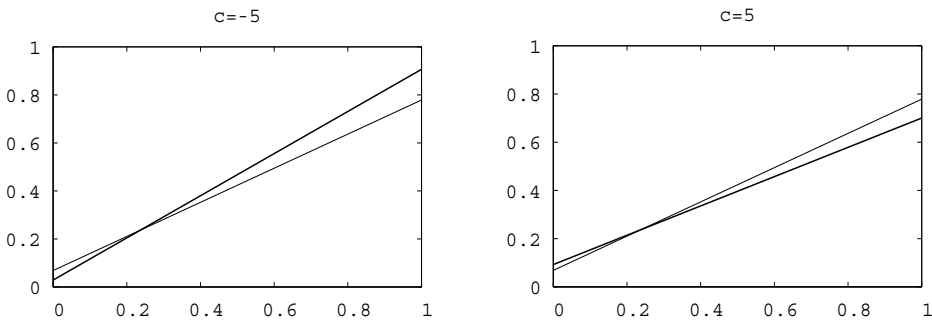
Przyjmijmy $n=10$, $\alpha=0.962$, $\beta=4.076$, skąd $E\Theta = \alpha/\beta = 0.2360$, $\text{Var}\Theta = \alpha/\beta^2 = 0.0579$. Dla δ_{exp} przyjmijmy $c=-5$ i $c=5$. Wykresy $x_L(\theta)$ i $x_U(\theta)$ dla podanych wyżej parametrów są przedstawione na rys. 4. Na każdym rysunku cieńszą linią narysowany jest wykres dla δ_{sqr} (taki sam na każdym rysunku), a grubszą linią narysowany jest wykres dla δ_{exp} .



Rys. 4. Wykresy $x_L(\theta)$ i $x_U(\theta)$ dla $c=-5$ i $c=5$

Źródło: opracowanie własne.

Jak widać, przedziały (x_L, x_U) dla funkcji strat L_{SQR} i L_{LINEX} różnią się nieznacznie nawet dla realizacji θ zmiennej losowej Θ znacznie większych od $E\Theta$. Niemniej dla danych realizacji \bar{x} średniej obserwowanej \bar{X} wartości estymatorów δ_{sqr} i δ_{exp} różnią się znacznie. Pokazują to na rys. 5 wykresy estymatorów δ jako funkcji obserwowanej średniej \bar{x} . Przyjęto te same parametry i oznaczenia co poprzednio.



Rys. 5. Wykresy δ_{sqr} i δ_{exp} dla $c = -5$ i $c = 5$

Źródło: opracowanie własne.

Dla wartości $\bar{x} = 1$ przekraczającej $E\Theta = 0.2360$ przeszło czterokrotnie różnica między δ_{sqr} i δ_{exp} osiąga ponad 10%, co jest z praktycznego punktu widzenia różnicą bardzo istotną.

4. Empiryczny bayesowski estymator zaufania przy funkcji LINEX

Estymator bayerowski $\delta_{exp}(X_1, \dots, X_n)$ nie dla każdej pary rozkładów zmiennej losowej Θ i zmiennej losowej X przy danym $\Theta = \theta$ jest liniowy. Gdy nie jest on liniowy, to nie jest składką zaufania. Z tego powodu poszukiwania δ_{exp} ograniczymy do klasy estymatorów liniowych

$$\delta(X_1, \dots, X_n) = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n.$$

Przy założeniach modelu Bühlmana obserwacje (X_1, \dots, X_n) są symetryczne (równoprawne), zatem estymatory liniowe dają się zapisać w postaci:

$$\delta(X_1, \dots, X_n) = a + b\bar{X},$$

gdzie

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Zagadnienie (9) sprowadza się do szukania stałych c i b takich, aby

$$\min_{a,b} E e^{-c(m(\Theta) - (a+b)\bar{X})} = E e^{-c(m(\Theta) - \delta_{exp})}$$

przy ograniczeniu $Em(\Theta) = EX = m$. Ponieważ $Em(\Theta) = E(c + b\bar{X})$, więc $m = c + bm$. Stąd

$$c = (1 - b)m.$$

Szukamy minimum funkcji jednej zmiennej

$$f(b) = E e^{-c((m(\Theta)-m)-b(\bar{X}-m))} = E \left(e^{-c(m(\Theta)-m)} e^{d(\bar{X}-m)} \right), \quad (23)$$

gdzie $d = cb$.

Wprowadzimy następujące oznaczenia na zmienne losowe w (23):

$$Z = e^{-c(m(\Theta)-m)}, \quad (24)$$

$$W = e^{d(\bar{X}-m)}. \quad (25)$$

Przypomnijmy wykorzystywane dalej własności rozkładu lognormalnego. Zmienna losowa X ma rozkład lognormalny, gdy $Y = \ln X$ ma rozkład normalny, $Y \sim N(m, \sigma)$, gdzie $EY = m$, $\text{Var}Y = \sigma^2$. Zmienna losowa o rozkładzie lognormalnym ma wartość oczekiwaną

$$E(X) = e^{m+\sigma^2/2} \quad (26)$$

i wariancję

$$\text{Var}(X) = e^{2m+2\sigma^2} (e^{\sigma^2} - 1). \quad (27)$$

Zmienna losowa dwuwymiarowa (X, Y) ma rozkład łączny lognormalny, gdy zmienna losowa $(\ln X, \ln Y)$ ma dwuwymiarowy rozkład normalny.

Założmy, że zmienna losowa dwuwymiarowa (Z, W) ma rozkład łączny lognormalny.

Twierdzenie 2. Dwuwymiarowa zmienna losowa o łącznym rozkładzie normalnym $(\ln Z, \ln W)$, gdzie Z i W są określone wzorami (24) i (25), ma rozkład z parametrami:

$$E \ln Z = E \ln W = 0, \quad (28)$$

$$\text{Var} \ln Z = c^2 \psi, \quad (29)$$

gdzie

$$\psi = \text{Var} E(X | \Theta), \quad (30)$$

$$\text{Var} \ln W = c^2 b^2 \left(\psi + \frac{\varphi}{n} \right), \quad (31)$$

gdzie

$$\varphi = E \text{Var}(X | \Theta) \quad (32)$$

oraz

$$\rho = \rho(\ln Z, \ln W) = \begin{cases} \frac{\sqrt{n\psi}}{\sqrt{\varphi + n\psi}} & \text{dla } b > 0, \\ -\frac{\sqrt{n\psi}}{\sqrt{\varphi + n\psi}} & \text{dla } b < 0. \end{cases} \quad (33)$$

Dowód. Najpierw udowodnimy wzór (28). Ponieważ

$$\ln Z = \ln e^{-c(m(\Theta)-m)} = -c(m(\Theta)-m),$$

to ze względu na (7) otrzymujemy

$$E \ln Z = -c(E m(\Theta) - m) = 0.$$

Analogicznie

$$\ln W = \ln e^{d(\bar{X}-m)} = d(\bar{X} - m),$$

skąd

$$E \ln W = 0.$$

Następnie udowodnimy wzory (29) i (31).

$$\text{Var}(\ln Z) = \text{Var}(c(m(\Theta) - m)) = c^2 \text{Var} m(\Theta) = c^2 \text{Var} E(X | \Theta),$$

co po skorzystaniu ze wzoru (30) daje (29). Podobnie

$$\text{Var} \ln W = \text{Var} d(\bar{X} - m) = d^2 \text{Var} \bar{X}.$$

Ponieważ

$$\begin{aligned} \text{Var} \bar{X} &= E(\text{Var}(\bar{X} | \Theta)) + \text{Var}(E(\bar{X} | \Theta)) \\ &= \frac{1}{n} \text{Var}(X | \Theta) + \text{Var} E(X | \Theta) = \frac{\varphi}{n} + \psi, \end{aligned}$$

gdzie $\varphi = E\text{Var}(X | \Theta)$, to otrzymujemy (31).

Na końcu udowodnimy wzór (33). Współczynnik korelacji

$$\rho = \rho(\ln Z, \ln W) = \frac{\text{Cov}(\ln Z, \ln W)}{\sqrt{\text{Var} \ln Z} \sqrt{\text{Var} \ln W}}. \quad (34)$$

Ze wzorów (29) i (31) otrzymujemy mianownik wzoru (34), równy

$$\sqrt{c^2 \psi d^2 \left(\psi + \frac{\varphi}{n} \right)} = |cd| \psi \sqrt{1 + \frac{\varphi}{n\psi}}.$$

Licznik wzoru (34):

$$\begin{aligned} \text{Cov}(\ln Z, \ln W) &= \text{Cov}(-c(m(\Theta) - m), d(\bar{X} - m)) \\ &= E(-c(m(\Theta) - m)d(\bar{X} - m)) + E c(m(\Theta) - m) E(d(\bar{X} - m)) \\ &= -cd E((m(\Theta) - m)(\bar{X} - m)) = -cd \text{Cov}(m(\Theta), \bar{X}). \end{aligned}$$

Ponieważ

$$\text{Cov}(m(\Theta), \bar{X} | \Theta) = E\left(\left(m(\Theta) - E(m(\Theta) | \Theta)\right)\left(\bar{X} - E(\bar{X} | \Theta)\right)\right) = 0,$$

gdź $E(m(\Theta) | \Theta) = m(\Theta)$ oraz $E(\bar{X} | \Theta) = E(X | \Theta)$, to

$$\begin{aligned} \text{Cov}(m(\Theta), \bar{X}) &= E(\text{Cov}(m(\Theta), \bar{X} | \Theta)) + \text{Cov}(E(m(\Theta) | \Theta), E(\bar{X} | \Theta)) \\ &= \text{Cov}(m(\Theta), E(X | \Theta)) = \text{Cov}(m(\Theta), m(\Theta)) \\ &= \text{Var } m(\Theta) = \text{Var } E(X | \Theta) = \psi. \end{aligned}$$

Zatem

$$\rho(\ln Z, \ln W) = \frac{-cd\psi}{|cd|\psi\sqrt{1+\frac{\varphi}{m\psi}}} = \frac{-c^2b\sqrt{n\psi}}{c^2|b|\sqrt{\varphi+n\psi}},$$

co dowodzi wzoru (33). \square

Lemat 1. Zmienna losowa $U = ZW$, gdzie Z i W są określone wzorami (24) i (25), ma rozkład lognormalny $\text{LN}(0, \sigma)$, gdzie

$$\sigma^2 = c^2(1-b^2)\psi + c^2b^2\frac{\varphi}{n}. \quad (35)$$

Dowód. Ponieważ $(\ln Z, \ln W)$ ma dwuwymiarowy rozkład normalny oraz $\ln U = \ln Z + \ln W$ ma rozkład normalny, to U ma rozkład lognormalny. Ze wzoru (28) otrzymujemy $E \ln U = \ln Z + \ln W = 0$. Ze wzorów (24) i (25) otrzymujemy

$$\begin{aligned} \ln U &= c(m(\Theta) - m) + d(\bar{X} - m) = cm(\Theta) - cm + d\bar{X} - dm \\ &= cm(\Theta) + d\bar{X} + g, \end{aligned}$$

gdzie stała $g = -cm - dm = -m(c - cb)$. Zatem

$$\begin{aligned} \sigma^2 &= \text{Var } \ln U = \text{Var}(cm(\Theta) + d\bar{X} + g) = \text{Var}(cm(\Theta) + d\bar{X}) \\ &= \text{Var } E(cm(\Theta) + d\bar{X} | \Theta) + E \text{Var}(cm(\Theta) + d\bar{X} | \Theta). \end{aligned}$$

Ponieważ

$$E(cm(\Theta) + d\bar{X} | \Theta) = cm(\Theta) + dE(\bar{X} | \Theta) = (c + d)m(\Theta)$$

oraz

$$\text{Var}(cm(\Theta) + d\bar{X} | \Theta) = c^2 \text{Var}(m(\Theta) | \Theta) + d^2 \text{Var}(\bar{X} | \Theta) = \frac{d^2}{n} \text{Var}(X | \Theta),$$

więc

$$\sigma^2 = (c + d)^2 \text{VarE}(X | \Theta) + \frac{d^2}{n\varphi} = c^2(1 - b^2)\psi + c^2b^2\frac{\varphi}{n}. \quad \square$$

Twierdzenie 3. Przy założeniu, że rozkład zmiennej losowej $(m(\Theta), \bar{X})$ jest normalny, składka zaufania przy funkcji straty LINEX jest postaci

$$\delta_{exp} = \frac{n\psi}{n\psi + \varphi} \bar{X} + \frac{\varphi}{n\psi + \varphi} m, \quad (36)$$

gdzie

$$\varphi = \text{EVar}(X | \Theta), \quad \psi = \text{VarE}(X | \Theta), \quad m = EX.$$

Dowód. Z lematu 1 mamy, że zmienna losowa U ma rozkład $\text{LN}(0, \sigma)$. Zatem jej wartość oczekiwana wynosi

$$EU = e^{\sigma^2/2} = \exp\left(\frac{1}{2}c^2\left(b^2\left(\psi + \frac{\varphi}{n}\right) - 2\psi b + \psi\right)\right). \quad (37)$$

Funkcja $f(b) = EU$ określona wzorem (37) osiąga minimum w punkcie

$$b_{opt} = \frac{\psi}{\psi + \frac{\varphi}{n}}.$$

Ponieważ

$$a_{opt} = (1 - b_{opt})m,$$

więc dla estymatora liniowego $\delta_{expo} = a_{opt} + b_{opt}\bar{X}$ przeciętna strata funkcji LINEX jest najmniejsza w klasie estymatorów liniowych. \square

Estymatory nieobciążone i zgodne parametrów struktury φ , ψ , m są w literaturze dobrze znane (zob. [Bühlmann, Gisler 2005, s. 93] lub [Jasiulewicz 2005, s. 71]). Jeśli we wzorze (36) za parametry struktury wstawi się ich estymatory, to wzór (36) nazywa się empirycznym bayesowskim estymatorem zaufania.

Na koniec należy wspomnieć pracę [Najafabadi 2010]. Dla pewnej klasy rozkładów zmiennych Θ i X przy ustalonym $\Theta = \theta$ A.P. Najafabadi przybliży estymatory bayesowskie składki netto przy niesymetrycznych funkcjach straty przez wypukłą kombinację średniej z obserwacji polisy i średniej *a priori* portfela. Tę kombinację nazywa przybliżoną formułą zaufania.

A. Dodatek – parametry warunkowe

Warunkową wartość oczekiwaną zmiennej losowej Y względem zmiennej losowej X oznacza się symbolem $E(Y | X)$. Warunkowa wartość oczekiwana $E(Y | X)$ jest zmienną losową Z będącą pewną funkcją zmiennej losowej X . Wartości zmiennej Z określa się następującym wzorem

$$z = z(x) = E(Y | X = x) = \int_{-\infty}^{\infty} y dF(y | x),$$

gdzie $F(y | x)$ jest dystrybuantą warunkową zmiennej losowej Y przy danym $X = x$. Poniżej zostaną podane własności warunkowej wartości oczekiwanej.

Własność 1. Jeśli istnieje $Eg(Y)$ dla funkcji g przedziałami ciąglej, to

$$E(g(Y) | Y) = g(Y). \quad (38)$$

Szczególnie jeśli istnieje EY , to

$$E(Y | Y) = Y. \quad (39)$$

Własność 2.

$$E(E(Y | X)) = EY, \quad (40)$$

o ile wszystkie wartości oczekiwane istnieją.

Własność tę nazywa się własnością iteracyjną wartości oczekiwanej.

Własność 3. Niech Y_1, \dots, Y_k będą zmiennymi losowymi, a_0, a_1, \dots, a_k zaś będą dowolnymi liczbami rzeczywistymi. Wówczas

$$E\left(\sum_{j=1}^k a_j Y_j + a_0 | X\right) = \sum_{j=1}^k a_j E(Y_j | X) + a_0, \quad (41)$$

pod warunkiem że wszystkie wartości oczekiwane istnieją.

Własność 4. Niech Y i W będą zmiennymi losowymi. Jeżeli istnieje $E(YW)$, to

$$E(YW | W) = W E(Y | W). \quad (42)$$

Własność 5. Jeżeli zmienne losowe Y i W są warunkowo niezależne względem zmiennej losowej X , to

$$E(YW | X) = E(Y | X)E(W | X). \quad (43)$$

Własność 6. Jeżeli X i V są wektorami losowymi, to

$$E(Y | X) = E(E(Y | X, V) | X),$$

oraz

$$E(Y | X) = E(E(Y | X) | X, V). \quad (44)$$

Własność 7. Niech $m^*(X) = E(Y | X)$. Jeżeli $EY^2 < \infty$, to $m^*(X)$ jest najlepszym estymatorem zmiennej losowej Y , w sensie najmniejszego przeciętnego błędu kwadratowego, bazującym na informacji zawartej w X . Innymi słowy zachodzi równość

$$E(Y - m^*(X))^2 = \min_{m \in M} E((Y - m(X))^2), \quad (45)$$

gdzie M jest zbiorem funkcji $m: R^n \rightarrow R$ takich, że $E(m(X))^2 < \infty$.

Wariancję warunkową zmiennej losowej Y względem wektora losowego X określa się wzorem

$$\text{Var}(Y|X) = E\left(\left(Y - E(Y|X)\right)^2 | X\right). \quad (46)$$

Wariancję warunkową można obliczyć również w inny, często wygodniejszy, sposób:

$$\text{Var}(Y|X) = E(Y^2|X) - (E(Y|X))^2. \quad (47)$$

Wzór (47) otrzymuje się z definicji wariancji warunkowej, rozpisując w niej kwadrat różnicy i stosując własności wartości oczekiwanej.

Własność 8. Dla dowolnych liczb rzeczywistych a i b zachodzi równość

$$\text{Var}(aY + b|X) = a^2 \text{Var}(Y|X). \quad (48)$$

Własność 9. Wariancję zmiennej losowej Y można przedstawić jako sumę dwóch nieujemnych składników w postaci

$$\text{Var}Y = E(\text{Var}(Y|X)) + \text{Var}(E(Y|X)). \quad (49)$$

Własność 10. Jeżeli zmienne losowe W i Y są warunkowo niezależne względem wektora losowego X , to

$$\text{Var}(aY + bW|X) = a^2 \text{Var}(Y|X) + b^2 \text{Var}(W|X) \quad (50)$$

dla dowolnych liczb rzeczywistych a i b .

Kowariancję warunkową zmiennych losowych W i Y względem zmiennej losowej Z określa się wzorem

$$\text{Cov}(W, Y|Z) = E\left(\left(W - E(W|Z)\right)\left(Y - E(Y|Z)\right) | Z\right), \quad (51)$$

gdzie zewnętrzna wartość oczekiwana jest wartością oczekiwaną rozkładu warunkowego zmiennej dwuwymiarowej (W, Y) pod warunkiem Z . Jeżeli $f(w, y|z)$ jest gęstością warunkową zmiennej (W, Y) pod warunkiem $Z = z$, to

$$\begin{aligned} \text{Cov}(W, Y|Z = z) \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (w - E(W|Z = z))(y - E(Y|Z = z))f(w, y|z)dw dy. \end{aligned}$$

Własność 11. Kowariancja warunkowa wyraża się wzorem:

$$\text{Cov}(W, Y|Z) = E(WY|Z) - E(W|Z)E(Y|Z). \quad (52)$$

Własność 12. Kowariancję zmiennych losowych W i Y można przedstawić jako sumę dwóch nieujemnych składników w następującej postaci:

$$\text{Cov}(W, Y) = E(\text{Cov}(W, Y|Z)) + \text{Cov}(E(W|Z), E(Y|Z)). \quad (53)$$

Literatura

- Bermúdez L., Denuit M., Dhaene J. (2001), *Exponential bonus-malus systems integrating a priori risk classification*, „Journal of Actuarial Practice”, no 9, <http://ssrn.com/abstract=884468>.
- Boratyńska A. (2008), *Posterior regret γ -minimax estimation of insurance premium in collective risk model*, „ASTIN Bulletin” no 38.
- Bühlmann H., Gisler A. (2005), *A Course in Credibility Theory and its Applications*, Springer-Verlag, Berlin.
- Denuit M., Dhaene J. (2001), *Bonus-malus scales using exponential loss functions*, „Blätter der DGVMF” no 25.
- Ferreira J. (1977), *Identifying Equitable Insurance Premiums for Risk Classes: an Alternative to the Classic Approach*, XXIIIth International Meeting of the Institute of Management Sciences, Athens.
- Jasiulewicz H. (2005), *Teoria zaufania. Modele aktuarialne*, AE, Wrocław.
- Lemaire J. (1979), *How to define a Bonus-Malus system with an exponential utility function*, „Astin Bull.” no 10.
- Najafabadi A.P. (2010), *A new approach to the credibility formula*, „Insurance Math. Econom.”, no 46.
- Niemiro W. (2006), *Bayesian prediction with an asymmetric criterion in a nonparametric model of insurance risk*, „Statistics” no 40.
- Parsian A., Kirmani S.N.U.A. (2002), *Estimation under LINEX Loss Function*, [w:] *Handbook of Applied Econometrics and Statistical Inference*, red. A. Ullah, A.T.K. Wan, A. Chaturvedi, Marcel Dekker, Inc, New York.
- Parsian A., Nematollahi N. (1996), *Estimation of scale parameter under entropy loss function*, „J. Stat. Plann. Inference” no 52.
- Pérez J., Gómez E., Vázquez F. (2002), *An Alternative Solution to the Problem of Overcharges in the Bonus-Malus System*, [w:] *6th International Congress on Insurance: Mathematics & Economics*, <http://pascal.iseg.utl.pt/cemapre/ime2002/>.
- Thomson R.D., Basu A.P. (1996), *Asymmetric Loss Functions for Estimating System Reliability*, [w:] *Bayesian Analysis in Statistics and Econometrics in Honor of Arnold Zellner*, red. D.A. Berry, K.M. Chaloner, J.K. Geweke, John Wiley & Sons, New York.
- Varian H.R. (1974), *A Bayesian Approach to Real Estate Assessment*, [w:] *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, red. S.E. Fienberg, A. Zellner, North Holland, Amsterdam.
- Zellner A. (1986), *Bayesian estimation and prediction using asymmetric loss functions*, „J. Amer. Statist. Assoc.” nr 81.

CREDIBILITY PREMIUMS USING ASYMMETRIC LOSS FUNCTIONS

Summary: The paper presents asymmetric loss functions (Linex, Stein, Entropy) in the context of insurance. The obtained estimators of insurance premium net which use asymmetric loss function Linex are discussed. The analysis of obtained estimators and their comparison with estimators of premium net determined by the quadratic loss function is carried out.

Key words: credibility premium, Linex loss function, Bayes estimator.