Włodzimierz Szkutnik

MODELOWANIE LOSOWEJ SKŁADKI UBEZPIECZENIOWEJ W ASPEKCIE RYZYKA SZACOWANIA ROSZCZEŃ W ZALEŻNOŚCI REGRESYJNEJ

1. Składka netto jako wartość oczekiwana wielkości roszczeń i innych czynników

W opracowaniu dotyczącym relacji między wielkością roszczenia a oszacowaniem składki podstawowe znaczenie ma analiza rozkładu prawdopodobieństwa łącznej wartości roszczeń, która wywiera bezpośredni wpływ na ustalanie składki netto związanej z daną grupą ryzyka.

Przedmiotem wyjściowych rozważań będzie zależność między wartością składki ζ, która wiąże się z daną polisą ubezpieczeniową i jest traktowana w modelu jako wielkość zmieniająca się losowo w sposób ciągły, a wielkością roszczenia Y, które może „pojawić się” w danej polisie. Przyjmujemy, że zachodzi następujący związek liniowy:

$$\zeta = h \cdot Y + \nu \tag{1}$$

gdzie h jest ustaloną liczbą, a ν – zmienną losową charakteryzującą wpływ różnych czynników na poziom składki i dokładność operacji ustalania poziomu składki na podstawie zależności (1). Ponadto będziemy zakładać, że zmienna losowa ν nie zależy od wielkości roszczenia Y. Wielkość Y może reprezentować informacje historyczne, a zmienna ν dane teraźniejsze.

Celem opracowania jest dokonanie analizy wpływu dokładności szacowania roszczeń na dokładność składki ubezpieczeniowej. Pierwsze podejście do tej analizy przeprowadzimy, zakładając, że zmienne ζ i Y mają rozkłady normalne. Składkę netto P będziemy szacować przyjmując poziom średni $\overline{\zeta}$

zmiennej losowej z zależności (1). Równość (1) oznacza, że wielkość roszczeń w danym momencie wpływa bezpośrednio na losowy poziom składki ζ w tym samym lub tuż po nim następującym momencie.

Uzasadnienie dla powyższego sformułowania modelowego dotyczącego szacowania składki netto P wynika stąd, że ryzyko szacowanej składki ma dwa źródła:

1) średni poziom $\overline{\zeta}$ przyjęty jako poziom składki netto może być inny od wielkości roszczeń zrealizowanych w okresie późniejszym,

2) poziom $\overline{\zeta}$ może być różny, i tak na pewno jest, od nieznanej wartości E(Y) i jest to tym bardziej realne, że nie czynimy tu założenia, że $E(\nu) = 0$.

Istnienie przypadkowych zmian wielkości ν powoduje, że przy jednym i tym samym wymiarze roszczeń Y wysokość składki ulega pewnym przypadkowym zmianom. Mamy tu więc zależność stochastyczną między ζ a Y. Ponieważ między zmiennymi ζ i Y istnieje zależność stochastyczna i każda z nich ma rozkład normalny, będziemy więc mówić, że istnieje między nimi **korelacja normalna**.

Wariancję $D^2(\zeta)$ składki, jako miarę szacowania jej ryzyka, można obliczyć bezpośrednio z równości (1):

$$D^2(\zeta) = h^2 \cdot D^2(Y) + D^2(v) \qquad (2)$$

Jeśli znane są wielkości $D^2(Y)$, $D^2(v)$, h, to ocena wpływu dokładności oceny roszczeń na dokładność składki netto P nie nastręcza trudności.

W celu uściślenia dokładności składki P oszacowanej na poziomie:

$$E(\zeta) = h \cdot E(Y) + E(v) \qquad (3)$$

załóżmy, że w procesie szkód scharakteryzowanym tak samo jak wcześniej, wielkości h i $D^2(v)$ nie zmieniają się. Jeśli zmniejszeniu ulega jedynie wariancja roszczenia Y, to wariancja $D^2(\zeta)$ z równości (2) także się zmniejszy. Widzimy zatem, że przy takim podejściu do procesu roszczeń, zmianie ulega jedynie pierwszy składnik ze wzoru (2). Będzie tak jednak tylko w pewnych granicach obserwowania wielkości roszczeń. Jest przy tym oczywiste, że uściślanie wariancji składki na podstawie tego samego rozkładu roszczeń Y nigdy nie będzie mniejsze od wariancji składnika przypadkowego v, wpływającego na losowość składki z modelu (1).

W powyższych rozważaniach założyliśmy, że dane są wielkości $D^2(Y)$, $D^2(v)$ i h. Zauważmy tu, że wariancja $D^2(Y)$ może być określona na podstawie przeszłych danych dotyczących roszczeń lub, jak to przyjęliśmy, w kolejnych etapach na podstawie równości (1) określającej dokładność (mierzoną wariancją) wyznaczania składki, przy założeniu niezmienności warunków (stałe h i $D^2(v)$) procesu roszczeń. Sytuacja taka możliwa jest do przyjęcia jedynie w bardzo małym przedziale czasu Δt, chociaż pojawiają się tu pewne wątpliwości dotyczące przyjmowania zmiennego poziomu przez wariancję $D^2(Y)$, co wpływa na szacowanie wariancji $D^2(\zeta)$ w kolejnych iteracjach oznaczanej przez $D^2(\zeta_i)$. Niezależnie od tego (możemy tu arbitralnie założyć, że kolejne roszczenia mają właśnie taką zmniejszającą się wariancję równą $D^2(\zeta_i)$) należy stwierdzić, że znacznie trudniej jest określić wariancję $D^2(\nu)$ oraz h.

2. Wyznaczenie regresji składki względem wielkości roszczeń

Przeprowadzimy rozważania, będące analogicznymi do znanych rozstrzygnięć w statystyce i rachunku prawdopodobieństwa, dotyczące linii regresji. Z uwagi na kontekst dydaktyczny i dalsze daleko idące uogólnienia, rozważania te będą dość dokładne.

Wydaje się, że przed przystąpieniem do rozwiązania zasygnalizowanych problemów celowe będzie przypomnienie postaci składki obliczanej według zasady czystego ryzyka (*pure risk premium*). Przy przyjętych wyżej oznaczeniach składkę szacuje się jako:

1) $P = E(Y)$ – wartość oczekiwana zmiennej losowej wielkości roszczeń,
2) $P = E(Y) + \beta\sigma(Y)$,
3) $P = E(Y) + \gamma D^2(Y)$,
4) $P = E(Y) + \beta'\sigma(Y) + \gamma' D^2(Y)$ – kompromisowe,
5) $P^* = E(Y) + \sigma \cdot \mathrm{cov}(Y, Y + Y^*)$,

gdzie Y* jest nowym procesem roszczeń w porównaniu z Y będącym podstawą wyznaczania składki według zasad 1-4.

Uwzględniając znaną równość $D^2(\nu) = E(\nu^2) - [E(\nu)]^2$ możemy składkę P, obliczoną na podstawie zasady 2, zapisać w postaci:

$$P = E(\zeta) = h \cdot E(Y) + \sqrt{E(\nu^2) - D^2(\nu)} \tag{4}$$

Otrzymana postać, podobnie jak te z zasad 2-5, uwzględnia zmienność wielkości roszczeń Y wyrażoną tu przez czynnik przypadkowy ν.

W celu ustalenia zależności między wielkościami roszczeń Y a poziomami ζ składki, badacz musi dokonać obserwacji, w których wielkości roszczeń odpowiada poziom składki ubezpieczeniowej (dane z przeszłości). Uzyskane pary liczb (Y, ζ) charakteryzują wielkości poprzednich, ale i następnych operacji szacowania składki netto P na podstawie danej zasady. Możliwe też jest (co nie jest jednak zadaniem prostym ze względu na koszty) obserwowanie wielkości roszczeń i zarazem szacowanie składki netto na podstawie nowych danych bez względu na to, na jakim poziomie była oszacowana składka netto dla celów praktyki ubezpieczeniowej.

Z formalnego punktu widzenia, pary liczb (Y, ζ) można rozpatrywać jako realizacje dwuwymiarowej zmiennej losowej, którą dla prostoty zapisu będziemy oznaczać tak samo, jak te realizacje. Z ilościowego punktu widzenia istnieje pewna zależność między operacyjnymi wartościami zmiennych Y i ζ w poszczególnych okresach. Tę ilościową zależność można scharakteryzować za pomocą współczynnika korelacji i równania korelacyjnego.

Przed przystąpieniem do analizy tych zależności, zbadamy dokładniej własności zmiennych losowych Y i ζ. Wysokość roszczenia przy sformułowanych założeniach (lub bez nich) możemy przedstawić jako sumę większej liczby niezależnych zmiennych losowych – roszczeń z indywidualnych polis lub jednej polisy w ustalonym okresie, z których każda składa się na łączne roszczenie w portfelu lub w danej polisie:

$$Y=Y_1+...+Y_m$$

Zmienna losowa ν ze wzoru (1) jest także sumą kilku, np. m zmiennych losowych, z których każda jest sumą błędów uwzględniających „niepewność" w obliczaniu składek w operacji $Y \to \zeta$. Suma ta wyraża część ryzyka związanego z szacowaniem składki i przyjmuje postać:

$$\nu = \nu_1 + \nu_2 + ... + \nu_m$$

Zgodnie z wcześniejszymi założeniami, roszczenia Y_i są niezależne od składowych ryzyk ν_j.

Z równości (1) wynika, że zmienna ζ z modelu składki losowej wyraża się wzorem:

$$\zeta = h \cdot Y_1 + h \cdot Y_2 + ... + h\cdot Y_m + \nu_1 + \nu_2 + ... + \nu_m$$

a zatem jest także sumą niezależnych zmiennych losowych.

W stochastycznej analizie zależności między dwiema zmiennymi losowymi, z których każdą można zapisać w postaci sumy zmiennych losowych, wykorzystamy następujące twierdzenie Bernsteina:

Niech zmienne losowe:

$$Z_1 = U_1 + U_2 + ... + U_n, \qquad Z_2 = \zeta_1 + \zeta_2 + ... + \zeta_n$$

będą sumami bardzo dużej liczby składników spełniających założenia twierdzenia Lapunowa:

$$\lim_{k\to\infty}\frac{\sum_{i=1}^{k}E\left[X_i-E(X_i)\right]^3}{\left[D^2(Z)\right]^{\frac{3}{2}}}=0$$

gdzie:

$$Z = X_1 + ... + X_k$$

Wtedy (teza twierdzenia Lapunowa):

$$\lim_{k\to\infty}F_k(x)=\Phi\left(\frac{Z-EZ}{D\ Z}\right)$$

a zatem suma Z asymptotycznie ma rozkład normalny.

W rozważanym przez nas przypadku funkcja gęstości zmiennej (Z_1, Z_2) przy n dążącym do nieskończoności jest równa:

$$\varphi(z_1,z_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{z(1-\rho^2)}\left(\frac{(z_1-a_1)^2}{\sigma_1^2}+\frac{(z_2-a_2)^2}{\sigma_2^2}-2\rho\frac{(z_1-a_1)(z_2-a_2)}{\sigma_1\sigma_2}\right)\right\}$$

(5)

gdzie:

$$\sigma_1=\sqrt{D^2(Z_1)}\ ,\ \sigma_2=\sqrt{D^2(Z_2)}\ ,\ a_1=E(Z_1)\ ,\ a_2=E(Z_2)$$

a ρ jest współczynnikiem korelacji między zmiennymi Z_1 i Z_2, który wyraża się znanym wzorem (Fisz, 1969):

$$\rho=\frac{\mathrm{cov}(Z_1,Z_2)}{\sigma_1\sigma_2}$$

gdzie cov (Z_1, Z_2) = E(Z_1 Z_2) – E(Z_1) E(Z_2) jest kowariancją zmiennych losowych Z_1 i Z_2. Funkcja (5) określa oczywiście zmienną dwuwymiarową o rozkładzie normalnym.

Twierdzenie Bernsteina ma takie znaczenie w zastosowaniach praktycznych dwuwymiarowego rozkładu normalnego, jak twierdzenie Lapunowa w zastosowaniach rozkładu jednowymiarowego.

Dla współczynnika korelacji rozpatrywanej zależności operacyjnych wyników wielkości Y i ζ mamy następującą procedurę obliczeniową:

$$\mathrm{cov}(Y, \zeta) = E(Y\cdot\zeta) - E(Y) \cdot E(\zeta) \tag{6}$$

Ponieważ (równanie (1)):

$$Y\cdot\zeta = Y \cdot (h \cdot Y + \nu) = h \cdot Y^2 + Y \cdot \nu$$

więc:

$$E(Y\cdot\zeta) = E(h \cdot Y^2 + Y \cdot \nu) = E(h \cdot Y^2) + E(Y \cdot \nu) = h \cdot E(Y^2) + E(Y) \cdot E(\nu)$$

Z drugiej strony, ponieważ:

$$D^2Y = EY^2 - E^2(Y)$$

więc:

$$E(Y\cdot\zeta) = h \cdot D^2(Y) + h \cdot E(Y^2) + E(Y) \cdot E(\nu) \tag{7}$$

Korzystając ponownie ze wzoru (1) mamy:

$$E(\zeta) = h \cdot E(Y) + E(\nu)$$

i dlatego otrzymujemy równanie:

$$E(Y) \cdot E(\zeta) = h \cdot E^2(Y) + E(Y) \cdot E(\nu) \tag{8}$$

Podstawiając (7) i (8) do wzoru (6) otrzymamy:

$$\begin{aligned} &\mathrm{cov}\,(Y, \zeta) = \\ &= h \cdot D^2(Y) + h \cdot E(Y^2) + E(Y) \cdot E(\nu) - [h \cdot E^2(Y) + E(Y) \cdot E(\nu)] = \\ &= h \cdot D^2(Y) = h \cdot \sigma_Y^2 \end{aligned} \tag{9}$$

Zatem współczynnik korelacji ρ między zmiennymi Y i ζ można obliczyć ze wzoru:

$$\rho = \frac{h \cdot \sigma_Y^2}{\sigma_Y \cdot \sigma_\zeta} = h \cdot \frac{\sigma_Y}{\sigma_\zeta} \tag{10}$$

Ze wzoru (10) wynika więc, że stała h, charakteryzująca tzw. dodatek bezpieczeństwa (Szkutnik, 2003), czyli współczynnik mający uwzględniać ryzyko przy szacowaniu składek, ma postać:

$$h = \rho \cdot \frac{\sigma_\zeta}{\sigma_Y} \tag{11}$$

Wzór (11) jest analogią wzoru określającego współczynnik regresji a_y w równaniu linii regresji:

$$\hat{y} = a_y x + b_y$$

Jest oczywiste, że wzór (5), po dokonaniu odpowiednich, wyżej wyznaczonych podstawień, jest funkcją gęstości rozkładu dwuwymiarowej zmiennej losowej (Y,ζ).

Dalsze rozważania wymagają wprowadzenia rozkładów warunkowych. Dla zmiennych Z_1 i Z_2 o rozkładach ciągłych, gęstość rozkładu warunkowego jednej z tych zmiennych wyraża się wzorem:

$$g_{(z_i/z_j)} = \frac{\varphi(z_i, z_j)}{f(z_j)}, \qquad i = 1,2, \quad j \neq i \tag{12}$$

Samo pojęcie rozkładu warunkowego zmiennej losowej ciągłej ma taki sam sens, jak pojęcie rozkładu warunkowego zmiennej losowej dyskretnej. Łatwo można bowiem stwierdzić, że:

$$g(z_2 / z_1) dz_1 = \frac{P\{z_1 \leq Z_1 \leq z_1 + dz_1, z_2 \leq Z_2 \leq z_2 + dz_2\}}{P\{z_1 \leq Z_1 \leq z_1 + dz_1\}}$$

a zatem $g(z_2 / z_1) dz_1$ określa prawdopodobieństwo tego, że zmienna losowa Z_2 przyjmie wartość z przedziału $[z_2, z_2 + dz_2]$ przy warunku, że zmienna losowa Z_1 przyjmie wartość z przedziału $[z_1, z_1 + dz_1]$.

Należy tu zauważyć, że rozkłady warunkowe są jednocześnie rozkładami jednowymiarowymi, tzn. rozkładami pewnej zmiennej losowej. Można więc mówić o warunkowej wartości oczekiwanej.

W ogólnym przypadku, gdy zmienna losowa $Z_1 = \sum_{i=1}^{n} u_i$ i spełnione są założenia twierdzenia Lapunowa, wtedy zmienna Z_1 ma rozkład normalny. Ze wzoru (12) otrzymamy więc:

$$\varphi(z_1, z_2) = \frac{1}{\sigma_2 \sqrt{2\pi} \sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(\sigma_2)\sqrt{1-\rho^2}} \left[z_2 - a_2 - \rho \frac{\sigma_2}{\sigma_1} (z_1 - a_1) \right]^2 \right\}$$

z której to postaci wynika, że rozkład warunkowy zmiennej Z_2 jest rozkładem normalnym o wartości oczekiwanej:

$$E(Z_2 / Z_1) = a_2 + \rho \frac{\sigma_2}{\sigma_1}(z_1 - a_1) \tag{13}$$

i o wariancji:

$$D^2(Z_2/Z_1) = (1 - \rho^2)\sigma_2^2 \tag{14}$$

Warunkowa wartość oczekiwana zmiennej losowej Z_2 jest funkcją zmiennej Z_1. Wychodząc z tej konstatacji wprowadzimy oznaczenie:

$$E(Z_2 / Z_1) = \overline{Z}_2(Z_1)$$

Wówczas z (13) wynika, że:

$$\overline{Z}_2(Z_1) = a_2 + \rho \frac{\sigma_2}{\sigma_2}(z_1 - a_1) \tag{15}$$

Równanie (15) jest znane jako empiryczna funkcja regresji, inaczej – równanie korelacyjne zmiennej Z_2 względem Z_1. Wykres tej funkcji przechodzi przez punkt (a_1, a_2). Zatem dla regresji składki jako wartości oczekiwanej warunkowej zmiennej ζ wartości roszczeń Y mamy:

$$\overline{\zeta}(Y) = a_2 + \rho \frac{\sigma_\zeta}{\sigma_Y}(Y - a_1) \tag{16}$$

gdzie:

$$a_1 = E(Y), \quad a_2 = E(\zeta) = h \cdot E(Y) + E(\nu)$$

Korzystając ze wzoru (10) określającego współczynnik ρ mamy:

$$\overline{\zeta}(Y) = h \cdot Y + E(\nu) \tag{17}$$

Przeciętna składka $\overline{\zeta}(Y)$ rośnie wobec tego wraz z wartością roszczenia, przy czym wzrost ten jest zintensyfikowany przez wartość h.

Równanie (17) można otrzymać bezpośrednio, gdy założymy, że zmienna Y jest ustalona (w próbie). Wtedy z równości (1) mamy:

$$E(\zeta/Y) = h \cdot Y + E(\nu) = \overline{\zeta}(Y)$$

Zatem otrzymujemy tu uogólniający schemat klasycznego liniowego modelowania ekonometrycznego polegający na osłabieniu założenia dotyczącego składnika losowego modelu.

Zakładając niezmienność wielkości roszczeń Y dla pewnego okresu, mamy składkę netto:

$$P = h \cdot Y + E(\nu)$$

gdzie:

$$h = \rho \cdot \frac{\sigma_\zeta}{\sigma_Y}$$

Składnik ν i odpowiednio $E(\nu)$ może być interpretowany jako dodatkowe roszczenie mogące się pojawić przy nietypowym, w losowy sposób realizowanym procesie szkód.

3. Interpretacje

Równanie linii regresji nie daje nam niestety możliwości odpowiedzi na pytanie, o ile zwiększy się dokładność oszacowanej składki, jeżeli zwiększy się dokładność prognozy dotyczącej wielkości roszczeń. Na to pytanie możemy jednak odpowiedzieć wtedy, gdy na podstawie danych z przeszłości o dwuwymiarowej zmiennej losowej (Y, ζ) obliczymy wielkości ρ, $D^2(\zeta)$, $D^2(Y)$. Kolejność postępowania jest następująca:

Ze wzoru:

$$h = \rho \cdot \frac{\sigma_\zeta}{\sigma_Y}$$

możemy obliczyć stałą h. Natomiast wariancję możemy obliczyć jako różnicę:

$$D^2(\nu) = D^2(\zeta) - h^2 \cdot D^2(Y) \tag{18}$$

Zakładając, że zmienna losowa ν nie zależy od Y, możemy obliczyć wariancję $D^2(\nu)$, która nie zależy od wartości $D^2(Y)$:

$$D^2(\nu) = D^2(\zeta) - \rho^2 \frac{\sigma_\zeta^2}{\sigma_Y^2} D^2(Y) = D^2(\zeta) - \rho^2 \sigma_\zeta^2 = D(\zeta)\left(1 - \rho^2\right) \tag{19}$$

Dlatego znając parametry ρ, $D^2(\zeta)$, $D^2(Y)$ można określić h i $D^2(\nu)$ występujące we wzorze (2), a mając $D^2(Y)$ – określić wpływ dokładności oszacowania wielkości roszczeń na dokładność oszacowanej składki.

Przykład 1

Na podstawie danych i hipotetycznych obliczeń stwierdzono, że wariancja wielkości roszczeń wynosi $D^2(Y) = 1000$, wariancja składki losowej $D^2(\zeta) = 100$, współczynnik korelacji $\rho = 0{,}3$ między wielkością roszczeń a składką „ciągłą" w takim znaczeniu, w jakim określiliśmy to na wstępie, czyli pewnej wartości traktowanej jako hipotetyczna wielkość losowa.

Określić, o ile zwiększy się dokładność oszacowania składki, jeżeli wariancja roszczeń będzie opracowana na podstawie dodatkowych danych roszczeniach i przyjmie wartość:

$$D^2(Y) = 500$$

Na podstawie wzoru (11) mamy:

$$h^2 = (0{,}3)^2 \cdot \frac{100}{1000} = 0{,}009$$

Korzystając ze wzoru (19) odpowiadającemu ryzyku szacowania składki lub dodatkowemu roszczeniu:

$$D^2(\nu) = 100\,(1 - 0{,}09) = 100 \cdot 0{,}91 = 91{,}0$$

lub inaczej:

$$D^2(\nu) = 100 - 0{,}09 \cdot 1000 = 100 - 9 = 91$$

Zgodnie z równaniem (2), zależność między $D^2(\zeta)$ i $D^2(Y)$ przyjmuje postać:

$$D^2(\zeta) = 0{,}009 \cdot D^2(Y) + 91$$

Zatem, jeśli $D^2(Y) = 500$, to $D^2(\zeta) = 95{,}5$, co oznacza, że wariancja szacowania składki jest nieco mniejsza od pierwotnie wyznaczonej.

Wniosek 1

Zmniejszenie wariancji roszczeń wpływa na zmniejszenie ryzyka szacowania składki. Zauważmy ponadto, że wyraźna poprawa dokładności oszacowania składki nastąpi wtedy, gdy zwiększy się stopień skorelowania między wielkościami ζ i Y.

Załóżmy, że $\rho = 0{,}9$, wtedy $h^2 = 0{,}081$, $D^2(\nu) = 19$, a następnie obliczymy wariancję $D^2(\zeta)$ składki losowej:

$$D^2(\zeta) = 0{,}081 \cdot D^2(Y) + 19$$

co przy $D^2(Y) = 500$ daje nam oszacowanie dokładności szacowanej składki na poziomie $D^2(\zeta) = 59{,}5$. Zatem nastąpiło istotne polepszenie dokładności oszacowania.

Wniosek 2

Przy silnym skorelowaniu ζ i Y (ρ ≈ 1) można twierdzić, że zwiększenie dokładności w szacowaniu roszczeń prowadzi do istotnego zwiększenia dokładności oszacowania składki, natomiast przy słabym skorelowaniu, zwiększanie dokładności szacowania wielkości roszczeń (w porównaniu z istniejącą standardową normą tego oszacowania) nie wpływa istotnie na dokładność oszacowania składki.

Trzeba tu podkreślić, że zmniejszenie dokładności szacowania roszczeń, przy danym współczynniku korelacji między ζ i Y, spowoduje znaczne zaniżenie dokładności szacowanej składki, czyli jest niewskazane przy wyznaczonej korelacji między ζ i Y. Przy szacowaniu składki w kolejnych etapach powinniśmy opierać się na danych szacunkowych względem wielkości roszczeń. Ma to istotne znaczenie wtedy, gdy współczynnik ρ uzyskano z dużego agregatu, a następnie wprowadzono go w nielicznym portfelu podobnych ryzyk o niezbyt długiej historii lub w odmiennych warunkach przejawiania się procesu szkód.

Uwaga 1

Odnosząc się do równania (16), tzn. wartości warunkowej $\overline{\zeta}(Y)$ oczekiwanej składki względem Y, należy zauważyć, że jeżeli zmienne ζ i Y są nieskorelowane, tzn. jeżeli ρ = 0, to są one niezależne (ζ i Y). Własność ta wynika ze wzoru (5), wtedy:

$$\varphi(z_1, z_2) = f_1(z_1) \cdot f_2(z_2)$$

Oczywiste jest też, co wynika z własności rozkładu normalnego, że jeżeli przynajmniej jedna ze zmiennych ζ i Y nie ma rozkładu normalnego, to zerowa wartość współczynnika korelacji nie zawsze jest równoznaczna z niezależnością tych zmiennych.

4. Uwagi uzasadniające przyjęte rozstrzygnięcia modelowe i możliwe uogólnienia

Istotnie ograniczającym warunkiem w powyższych rozważaniach jest założenie normalności odpowiednich zmiennych, chociaż z twierdzenia Lapunowa uzyskujemy uzasadnienie mające dla tego założenia praktyczne podstawy w ilości obserwowanych roszczeń. Brakuje tu jednak wyraźnie wyjaśnienia przyczyn powodujących powstanie zależności typu (1). Zwrócić należy tu także uwagę na traktowanie w równaniu (1) poziomu ζ składki jako zmiennej losowej. Wynika to z jednej strony ze względów formalnych, a z drugiej z autentyczności tego założenia jako procesu ciągłego, który w rzeczywistości się upraszcza, bo nie można opłacać składki korygując ją w sposób ciągły. Ustala się ją na pewien okres, korygując ewentualne rozbieżności między jej faktyczną a ustaloną wartością, co wyraża dodany do modelu składnik ν mający niezerową wartość przeciętną.

W uogólnieniu przeanalizowanego w opracowaniu podejścia modelowego można rozpatrzyć przypadek uwzględniający istnienie rozstępu między obserwowanymi wartościami wielkości roszczeń. Umożliwia to analizę związku między rozstępem roszczeń a rozstępem szacowanej składki losowej. W prześledzonym uogólnieniu przejście od rozstępu roszczeń do rozstępu składki losowej jest związane z występowaniem elementów ryzyka ubezpieczeniowego (Szkutnik, 2003), jakim jest szacowanie składki, uwzględnionych i występujących w procesie szkód w sposób stały, które są nazywane czynnikami wpływu zależnymi od stopnia tzw. sztywności założeń i od siły wpływu postaci analitycznej na szacowaną składkę ubezpieczeniową.

Innym podejściem (od przedstawionego w opracowaniu) jest przyjęcie silnego, ale uzasadnionego założenia, że wielkość h, intensyfikująca wpływ losowych roszczeń na wielkość składki, jest wielkością losową.

Możliwa jest także analiza zależności szacowania składki w zależności od wielkości zakładanego roszczenia z uwzględnieniem korelacji. Rozważyć można też model wyceny składki przy zmiennym w czasie sposobie szacowania poziomu roszczeń, a także model szacowania błędu oceny roszczeń przy różnych alternatywnych sposobach oceniania roszczeń, w tym model porównań wyceny roszczeń przy kryterium bezwzględnej różnicy błędu ich oceny.

Literatura

Fisz M.: *Rachunek prawdopodobieństwa i statystyka matematyczna*. PWN, Warszawa 1969.

Ronka-Chmielowiec W.: *Ryzyko ubezpieczeniowe – Metody oceny*. AE, Wrocław 1997.

Szkutnik W.: *Wybrane modele w zarządzaniu ryzykiem ubezpieczeniowym w ujęciu probabilistycznym*. AE, Katowice 2003.

MODELLING OF THE RANDOM RISK PREMIUM ON THE BASIS OF THE ESTIMATION RISK OF THE CLAIMS IN THE REGRESSIVE RELATION

Summary

The model solutions presented in the paper concern the regressive dependence between the pure risk premium – analysed as the random amount – and the sum of the random amount of claims and some other factors, the so called factors of influence. The analysis of this dependence has enabled to draw certain conclusions, which have a precise meaning in the possible generalisations of the considered dependence. Moreover, the directions of the evolution of the analysis carried out in the paper have been shown.