PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1108 — 2006
Statystyka aktuarialna – stan i perspektywy rozwoju w Polsce

Stanisław Heilpern
Akademia Ekonomiczna we Wrocławiu

ANALIZA ZALEŻNEGO RYZYKA UBEZPIECZENIOWEGO – ZASTOSOWANIE FUNKCJI ŁĄCZĄCYCH

1. Wstęp

W klasycznych aktuarialnych modelach (por. [1; 6; 9]) zakłada się zwykle niezależność występujących zmiennych losowych. Jest to założenie wygodne z punktu widzenia badacza rozwiązującego postawiony problem; prostsza jest wtedy matematyczna, czy statystyczna analiza rozpatrywanego zadania, ale na ogół założenie to jest dalekie od rzeczywistości. Z tego też powodu, chcąc uzyskać bardziej realistyczne modele opisujące badany problem, powinno się osłabić w nich bardzo silne założenie niezależności przynajmniej niektórych zmiennych losowych.

W literaturze dotyczącej analizy i modelowania zależności w zagadnieniach ubezpieczeniowych możemy spotkać dwa podejścia. Pierwsze polega na przedstawieniu zmiennych występujących w modelu w postaci sumy niezależnych składników (por. [2; 3; 4; 5]). Niektóre składniki są wspólne dla części zmiennych. Można je interpretować jako czynniki zewnętrzne, np. klimatyczne, makroekonomiczne czy polityczne, wpływające na rozpatrywane zmienne. Wspólne składniki powodują zależność zmiennych losowych występujących w modelu. Na przykład w indywidualnym modelu ryzyka indykatory I_j można przedstawić jako sumę

$$I_j = J_j + J_0,$$

gdzie: J_j – składnik charakteryzujący j-tą polisę,

J_0 – wspólny składnik oddający wpływ czynników zewnętrznych.

Druga metoda wykorzystuje do modelowania zależności tzw. funkcje łączące (*copula*), będące łącznikami między dystrybuantami brzegowymi F_j a dystrybuantą łączną F (por. [2; 3; 5]).

Praca ma charakter przeglądowy. Przeprowadzono w niej analizę zależnego ryzyka ubezpieczeniowego za pomocą funkcji łączących, posługując się głównie dystrybuantami, a nie zmiennymi losowymi, tak jak w przypadku zastosowania pierwszej metody.

W punkcie drugim przedstawiono podstawowe informacje dotyczące funkcji łączących, głównie archimedesowych. Archimedesowe funkcje łączące indukują ukrytą zmienną losową, którą można traktować jako wpływ wspólnego czynnika na rozpatrywane zmienne losowe. Własność ta ułatwia wyznaczanie rozkładu sumy zmiennych losowych. Punkt trzeci poświęcony jest indywidualnemu modelowi ryzyka, w którym zależność indykatorów jest modelowana za pomocą zarówno archimedesowych, jak i ogólnych funkcji łączących. Ostatnia część jest próbą zastosowania funkcji łączących do opisu zależności w wieloklasowym modelu ryzyka kolektywnego.

2. Funkcje łączące

Do opisu zależności między zmiennymi losowymi $Z_1, \ldots, Z_n$ można wykorzystać tzw. funkcje łączące. Stanowią one łącznik między dystrybuantami F_i rozkładów brzegowych Z_i, a dystrybuantą F rozkładu łącznego. Funkcja łącząca $C : [0, 1]^n \to [0, 1]$ spełnia bowiem zależność (zob. [10; 8])

$$F(z_1, \ldots, z_n) = C(F_1(z_1), \ldots, F_n(z_n)).$$

W przypadku ciągłych rozkładów brzegowych funkcja łącząca C jest wyznaczona jednoznacznie. Dla rozkładów dyskretnych istnieje nieskończenie wiele takich funkcji.

Do najprostszych i najczęściej stosowanych w praktyce funkcji łączących należą archimedesowe funkcje łączące. Są one wyznaczone przez jednowymiarowy addytywny generator φ wzorem:

$$C(u_1, \ldots, u_n) = \varphi^{-1}(\varphi(u_1) + \ldots + \varphi(u_n)),$$

gdzie $\varphi : [0, 1] \to R_+$ jest malejącą, całkowicie monotoniczną funkcją, tzn. $(-1)^k \frac{d^k}{du^k} \varphi(u) \geq 0$ dla każdego $k = 1, 2, \ldots$, spełniającą warunki: $\varphi(0) = \infty$ oraz $\varphi(1) = 0$. Archimedesowe funkcje łączące możemy też definiować, korzystając z generatora multiplikatywnego $g(u) = \exp(-\varphi(u))$. Zachodzi wtedy zależność:

$$C(u_1, \ldots, u_n) = g^{-1}(g(u_1) \cdot \ldots \cdot g(u_n)).$$

Addytywny generator φ jest funkcją całkowicie monotoniczną. Jej odwrotność $\varphi^{-1}(t)$ jest wtedy transformatą Laplace'a $L_\Theta(t)$ pewnej nieujemnej zmiennej losowej Θ (por. [7; 2]). Otrzymujemy wtedy

$$\varphi^{-1}(t) = L_\Theta(t) = M_\Theta(-t),$$

gdzie $M_\Theta(s)$ jest funkcją tworzącą momenty zmiennej losowej Θ. Natomiast dla generatora multiplikatywnego g otrzymujemy

$$g(u) = \exp\left(M_\Theta^{-1}(u)\right)$$

oraz

$$g^{-1}(s) = M_\Theta\left(\ln(s)\right).$$

Zmienną losową Θ możemy traktować w zagadnieniach ubezpieczeniowych, np. jako zewnętrzny, często ukryty czynnik wpływający jednocześnie na wszystkie typy ryzyka Z_i. Mogą to być czynniki pogodowe, związane z katastrofami, takimi jak: powodzie, huragany czy trzęsienia ziemi, czynniki makroekonomiczne – cena ropy naftowej, kursy walut lub krachy finansowe – oraz czynniki polityczne, takie jak ewentualne konflikty zbrojne, kryzysy rządowe czy rozruchy. Zmienna losowa Θ stanowi podstawę tzw. modelu słabości (*frailty*) (zob. [7; 11]). Zakłada się w nim, że warunkowe dystrybuanty $F_{i|\theta}(z) = P\left(Z_i \le z \mid \Theta = \theta\right)$ spełniają warunek

$$F_{i|\theta}(z) = \left(H_i(z)\right)^\theta,$$

gdzie $H_i\left(z\right)$ jest dystrybuantą bazowej zmiennej losowej odpowiadającej jednostkowej wartości ukrytej zmiennej Θ (słabości), tzn. gdy $\Theta = 1$. Bezwarunkowa dystrybuanta zmiennej Z_i jest wtedy równa

$$F_i(z) = M_\Theta\left(\ln\left(H_i(z)\right)\right).$$

W modelu tym przyjmuje się warunkową niezależność zmiennych Z_i dla ustalonej wartości słabości $\Theta = \theta$, czyli

$$F_\theta\left(z_1, ..., z_n\right) = F_\theta\left(z_1, ..., z_n\right) = \prod_{i=1}^{n} F_{i|\theta}\left(z_i\right).$$

W przypadku gdy struktura zależności między zmiennymi Z_i opisana jest archimedesową funkcją łączącą C z generatorem multiplikatywnym g, to bazowe zmienne losowe są równe

$$H_i(z) = g\left(F_i(z)\right),$$

a bezwarunkowa dystrybuanta łączna wynosi (por. [5; 7])

$$F(z_1, ..., z_n) = \int_0^\infty \prod_{i=1}^n (H_i(z_i))^\theta dF_\Theta(\theta) = \int_0^\infty F_\theta(z_1, ..., z_n) dF_\Theta(\theta).$$

Jest więc ona mieszanką warunkowych dystrybuant łącznych F_θ, gdzie zmienną mieszającą jest słabość Θ.

Przykład 1

A. W przypadku gdy struktura zależności opisana jest funkcją łączącą należącą do rodziny Claytona

$$C(u_1, ..., u_n) = \left(u_1^{-\alpha} + ... + u_n^{-\alpha} - n + 1\right)^{-1/\alpha},$$

gdzie parametr $\alpha > 0$, z generatorem $g(u) = \exp\left(1 - u^{-\alpha}\right)$, wówczas zmienna ukryta Θ ma rozkład gamma $G(1/\alpha, 1)$.

B. Dla funkcji łączącej z rodziny Gumbela

$$C(u_1, ..., u_n) = \exp\left(-\left((-\ln u_1)^\alpha + ... + (-\ln u_n)^\alpha\right)^{1/\alpha}\right),$$

gdzie $\alpha \geq 1$, z $g(u) = \exp\left(-(-\ln u)^\alpha\right)$, otrzymujemy ukrytą zmienną Θ o rozkładzie stabilnym z transformatą Laplace'a $L_\Theta(t) = \exp\left(-t^{1/\alpha}\right)$.

C. Dla funkcji łączącej Franka

$$C(u_1, ..., u_n) = \frac{1}{\alpha} \ln\left(1 + \prod_{i=1}^n \left(\exp(\alpha u_i) - 1\right) / \left(\exp(\alpha) - 1\right)\right),$$

gdzie $\alpha \neq 0$, z generatorem $g(u) = \dfrac{\exp(\alpha u) - 1}{\exp(\alpha) - 1}$, otrzymujemy zmienną losową Θ o rozkładzie logarytmicznym.

D. Gdy zmienne losowe Z_i są niezależne, to funkcja łącząca jest równa $C(u_1, ..., u_n) = u_1 \cdot ... \cdot u_n$, generator multiplikatywny jest identycznością $g(u) = u$, a odpowiadająca ukryta zmienna Θ ma rozkład skupiony w 1.

Funkcja generująca momenty sumy zmiennych losowych $S = Z_1 + ... + Z_n$ jest mieszanką funkcji generujących momenty dla warunkowej sumy $S_\theta = Z_{1|\theta} + ... + Z_{n|\theta}$ (por. [11]), czyli

$$M_S(s) = \int_0^\infty M_{S_\theta}(s) dF_\Theta(\theta).$$

W podobny sposób możemy przedstawić jako odpowiednie mieszanki dystrybuantę $F_S(z)$ i gęstość $f_S(z)$ sumy S oraz składkę zatrzymania straty (*stop-loss premium*) $\pi_S(d) = E\left((S-d)_+\right)$, często stosowaną w ubezpieczeniach jako miarę całkowitego ryzyka (por. [2; 3]). Otrzymujemy wtedy:

$$F_S(z) = \int_0^\infty F_{S_\theta}(z)dF_\Theta(\theta),$$

$$f_S(z) = \int_0^\infty f_{S_\theta}(z)dF_\Theta(\theta)$$

oraz

$$\pi_S(d) = \int_0^\infty \pi_{S_\theta}(d)dF_\Theta(\theta).$$

Przykład 2

Załóżmy, że struktura zależności zmiennych losowych $X_1, ..., X_n$, przedstawiających wielkości szkód, opisana jest funkcją łączącą Claytona z parametrem α. Wtedy, jak wiadomo z przykładu 1a, ukryta zmienna losowa Θ ma rozkład gamma $G(1/\alpha, 1)$. Zmienna ta może reprezentować wspólny czynnik wpływający na wielkości szkód X_i. Przyjmijmy, że przy ustalonej wartości $\Theta = \theta$ niezależne zmienne losowe $X_{i|\theta}$, mają rozkład wykładniczy $W\left(\lambda_i / \theta\right)$. Dystrybuanta warunkowego, łącznego rozkładu zmiennych X_i jest wtedy równa

$$F_\theta\left(x_1, \dots, x_n\right) = 1 - \exp\left(-\theta\sum_{i=1}^{n}\frac{x_i}{\lambda_i}\right),$$

warunkowa suma szkód S_θ ma rozkład gamma $G(n, \lambda/\theta)$, a bezwarunkowy rozkład zmiennych X_i jest rozkładem Pareto $\mathrm{Pa}\left(1/\alpha, \lambda_1\right)$. Natomiast gęstość $f_S(x)$ bezwzględnej sumy szkód $S = X_1 + ... + X_n$ możemy wyznaczyć jako gęstość mieszanki rozkładu gamma $\mathrm{G}\left(n, \dfrac{\lambda}{\theta}\right)$ z rozkładem $\mathrm{G}\left(\dfrac{1}{\alpha}, 1\right)$. Jest ona równa

$$f_S(x) = \int_0^{\infty} \frac{x^{n-1}\theta^n}{\lambda^n \Gamma(n)} \exp\left(-\frac{\theta}{\lambda}x\right) \frac{\theta^{1/\alpha-1}}{\Gamma(1/\alpha)} \exp(-\theta) d\theta =$$

$$= \frac{x^{n-1}}{\lambda^n \Gamma(n)\Gamma(1/\alpha)} \int_0^{\infty} \theta^{n+1/\alpha-1} \exp\left(-\frac{x+\lambda}{\lambda}\theta\right) d\theta =$$

$$= \frac{\Gamma(n+1/\alpha)}{\Gamma(n)\Gamma(1/\alpha)} \frac{\lambda^{1/\alpha} x^{n-1}}{(x+\lambda)^{n+1/\alpha}}.$$

Jest to gęstość uogólnionego rozkładu Pareto $\text{UPa}\left(\frac{1}{2}, \lambda, n\right)$.

3. Indywidualny model ryzyka

W klasycznym indywidualnym modelu ryzyka (por. [1; 6]) rozpatruje się portfel składający się z n polis. Każda polisa generuje losową stratę X_j, którą możemy zapisać w postaci

$$X_j = I_j B_j,$$

gdzie indykator I_j przyjmuje wartość 1 lub 0, w zależności, czy była lub nie wypłata z j-tej polisy, a B_j jest dodatnią zmienną losową o dystrybuancie $F_{B_j}(x)$, przedstawiającą wielkość wypłaty z j-tej polisy, gdy do niej dojdzie. Oznaczmy symbolem p_j prawdopodobieństwo wystąpienia wypłaty z j-tej polisy, tzn.

$$p_j = P(I_j = 1) = P(X_j > 0).$$

Wtedy prawdopodobieństwo niewystąpienia wypłaty jest równe:

$$q_j = P(I_j = 0) = P(X_j = 0),$$

gdzie $q_j = 1 - p_j$.

W klasycznym modelu indywidualnego ryzyka zakłada się niezależność wszystkich zmiennych losowych: I_j oraz B_j. Założenie to osłabimy, dopuszczając zależność zmiennych losowych I_j opisujących wystąpienie wypłat. Jest to bardziej realistyczne założenie, zgodne z sytuacjami spotykanymi w praktyce, kiedy często występuje wspólny czynnik zewnętrzny, taki jak pogoda czy uwarunkowania gospodarcze, wpływający jednocześnie na większość polis (por. [2; 4; 5]). Zależność zmiennych losowych I_j powoduje jednoczesną zależność zmiennych X_j. Zadaniem naszym będzie wyznaczenie rozkładu łącznego sumy wypłat:

$$S = X_1 + \ldots + X_n,$$

przy założeniu, że struktura zależności wektora zmiennych losowych $\mathbf{I} = (I_1, \ldots, I_n)$ opisana jest funkcją łączącą C, tzn.

$$F_{\mathbf{I}}(i_1, \ldots, i_n) = C(F_1(i_1), \ldots, F_n(i_n)),$$

gdzie $F_{\mathbf{I}}$ jest łączną dystrybuantą zmiennych losowych I_j, a F_j brzegowymi dystrybuantami.

3.1. Archimedesowe funkcje łączące

Załóżmy, że struktura zależności zmiennych losowych I_i opisana jest archimedesową funkcją łączącą C z multiplikatywnym generatorem g. Generator g wyznacza rozkład ukrytej zmiennej losowej Θ. Bazowa dystrybuanta H_j wyznaczona dla j-tej polisy przybiera wtedy postać

$$H_j(i) = g(F_j(i)).$$

Jest to rozkład dwupunktowy jednoznacznie wyznaczony prawdopodobieństwem $r_j = g(q_j)$. Natomiast warunkowy rozkład zmiennej losowej I_j określony jest dystrybuantą $F_{j|\theta}(i) = (H_j(i))^{\theta}$. Jest to również rozkład dwupunktowy opisany prawdopodobieństwem

$$r_j^{\theta} = (g(q_j))^{\theta} = P(I_j = 0 \mid \Theta = \theta).$$

Bezwarunkowy rozkład zmiennej I_j jest jednoznacznie wyznaczony przez prawdopodobieństwo r_i i funkcję tworzącą momenty zmiennej Θ. Zachodzi bowiem zależność

$$P(I_j = 0) = q_j = \int_0^{\infty} r_j^{\theta} dF_{\Theta}(\theta) = M_{\Theta}(\ln(r_j)).$$

Warunkowa dystrybuanta łączna wektora $\mathbf{I}$ jest równa $F_{\mathbf{I}|\theta}(i_1, \ldots, i_n) = \prod_{j=1}^{n} F_{j|\theta}(i_j)$, ponieważ warunkowe zmienne $I_{j|\theta}$ są niezależne. Jest to dystrybuanta zmiennej losowej dyskretnej skupionej na n-wymiarowych punktach o współrzędnych równych 0 lub 1. Niech $\mathbf{c} = (c_1, ..., c_n)$ będzie takim punktem, którego k współrzędnych c_{j_m} dla $m = 1, 2, .., k$, jest równych 0. Wtedy

$$F_{\mathbf{I}|\theta}(\mathbf{c}) = \begin{cases} \prod_{m=1}^{k} r_{j_m}^{\theta} & k > 0 \\ 1 & k = 0. \end{cases}$$

Po skorzystaniu z niezależności możemy wyznaczyć funkcję tworzącą prawdopodobieństwa warunkowego losowego wektora **I**:

$$\mathrm{P}_{\mathbf{I}|\theta}(t_1, ..., t_n) = \prod_{j=1}^{n} \left(r_j^{\theta} + \left(1 - r_j^{\theta}\right) t_j \right).$$

Funkcja tworząca momenty zmiennej losowej X_j, gdy $\Theta = \theta$, jest równa

$$M_{X_j|\theta}(t) = E\left(\exp\left(tX_j\right) \middle| \Theta = \theta\right) = E\left(\exp\left(tB_j\right)^{I_j} \middle| \Theta = \theta\right) =$$
$$= \mathrm{P}_{I_j|\theta}\left(M_{B_j}(t)\right) = r_j^{\theta} + \left(1 - r_j^{\theta}\right) M_{B_j}(t).$$

Opierając się na przedstawionej zależności, możemy wyznaczyć funkcję tworzącą momenty sumy wypłat:

$$M_S(t) = \int_0^{\infty} M_{S|\theta}(t) dF_{\Theta}(\theta),$$

gdzie $M_{S|\theta}(t) = M_{\mathbf{X}|\theta}(t, ..., t) = \prod_{j=1}^{n} \left(r_j^{\theta} + \left(1 - r_j^{\theta}\right) M_{B_j}(t) \right)$.

Dystrybuantę rozkładu zmiennej S wyznaczamy, odwracając funkcję tworzącą momenty. Można to zrobić numerycznie w sposób przybliżony, np. stosując metodę szybkiej transformaty Fouriera (zob. [11; 9]). Inny sposób jej określenia opiera się na zależności

$$F_S(x) = \int_0^{\infty} F_{S|\theta}(x) dF_{\Theta}(\theta). \tag{1}$$

Warunkową dystrybuantę $F_{S|\theta}(x)$ wyznaczamy, stosując numeryczne metody, takie jak algorytm Deprila (zob. [9; 6]) czy przybliżenie złożonym rozkładem Poissona (zob. [2; 5; 6; 9]).

Wartość oczekiwana sumy wypłat jest równa

$$E(S) = \sum_{j=1}^{n} E\left(X_j\right) = \sum_{j=1}^{n} E\left(B_j\right) p_j,$$

jej wariancja wynosi (por. [2])

$$V(S)=\sum_{j=1}^{n}V\left(X_j\right)+2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}Cov\left(X_j,\,X_k\right),$$

gdzie $V\left(X_j\right)=E\left(I_j^2B_j^2\right)-E^2\left(I_j\right)E^2\left(B_j\right)=E\left(B_j^2\right)p_j-E^2\left(B_j\right)p_j^2$, a kowariancja $Cov\left(X_j,\,X_k\right)=E\left(I_jB_jI_kB_k\right)-E\left(I_j\right)E\left(B_j\right)E\left(I_k\right)E\left(B_k\right)=E\left(B_j\right)E\left(B_k\right)\left(E\left(I_jI_k\right)-\right.$ $\left.-p_jp_k\right)$. Natomiast wartość oczekiwaną iloczynu $E\left(I_jI_k\right)$ możemy w tym przypadku obliczyć, korzystając z ukrytej zmiennej Θ:

$$E\left(I_jI_k\right)=\int_0^\infty\left(1-r_j^\theta\right)\left(1-r_k^\theta\right)dF_\Theta(\theta)=1-q_j-q_k+M_\Theta\left(\ln\left(r_jr_k\right)\right)$$

oraz

$$E\left(I_jI_k\right)-p_jp_k=M_\Theta\left(\ln\left(r_jr_k\right)\right)-q_jq_k.$$

Przykład 3

Załóżmy, że zmienne losowe B_j mają rozkład wykładniczy $W\left(\lambda_j\right)$. Wtedy funkcja tworząca momenty sumy wypłat przybiera postać

$$M_S(t)=\int_0^\infty\prod_{j=1}^{n}\left(r_j^\theta+\frac{1-r_j^\theta}{1-\lambda_jt}\right)dF_\Theta(\theta)=\int_0^\infty\prod_{j=1}^{n}\frac{1-r_j^\theta\lambda_jt}{1-\lambda_jt}dF_\Theta(\theta).$$

Wartość oczekiwana i wariancja sumy są równe:

$$E(S)=\sum_{j=1}^{n}\lambda_jp_j,$$

$$V(S)=\sum_{j=1}^{n}\lambda_j^2p_j(2-p_j)+2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}\lambda_j\lambda_k\left(M_\Theta\left(\ln\left(r_jr_k\right)\right)-q_jq_k\right).$$

W sytuacji, gdy zmienne losowe B_j i I_j mają ten sam rozkład, możemy zapisać funkcję tworzącą momenty bezwarunkowej sumy wypłat w następujący sposób:

$$M_S(t)=\int_0^\infty\left(r^\theta+\left(1-r^\theta\right)M_B(t)\right)^n dF_\Theta(\theta)=$$
$$=\int_0^\infty\sum_{k=0}^n\binom{n}{k}r^{(n-k)\theta}\left(1-r^\theta\right)^k M_B^k(t)dF_\Theta(\theta)=$$
$$=\int_0^\infty\sum_{k=0}^n\binom{n}{k}\sum_{j=0}^k(-1)^j\binom{k}{j}r^{(n-k+j)\theta}M_B^k(t)dF_\Theta(\theta)=$$
$$=\sum_{k=0}^n\binom{n}{k}M_B^k(t)\sum_{j=0}^k(-1)^j\binom{k}{j}\int_0^\infty r^{(n-k+j)\theta}dF_\Theta(\theta)=$$
$$=\sum_{k=0}^n\binom{n}{k}M_B^k(t)\sum_{j=0}^k(-1)^j\binom{k}{j}M_\Theta\left((n-k+j)\ln(r)\right),$$

gdzie $M_B(t)$ jest funkcją tworzącą momenty zmiennych losowych B_j, a $r=g(q)$, gdzie q jest prawdopodobieństwem wyznaczającym indykatory I_j. Wykorzystując własności funkcji tworzących momenty, możemy wyznaczyć dystrybuantę sumy wypłat:

$$F_S(t)=\sum_{k=0}^n\binom{n}{k}F_B^{*k}(x)\sum_{j=0}^k(-1)^j\binom{k}{j}M_\Theta\left((n-k+j)\ln(r)\right),$$

gdzie $F_B^{*k}(x)$ jest dla $k\geq 1$ k-krotnym splotem dystrybuanty $F_B(x)$ zmiennych losowych B_j, a $F_B^{*0}(x)$ jest deltą Diraca, czyli dystrybuantą rozkładu skupionego w 0. Przedstawiony wzór możemy otrzymać, korzystając z (1) (por. [2]). Przykładowo przyjmijmy, że $n = 3$. Wtedy

$$F_S(x)=M_3+3\left(M_2-M_3\right)F_B(x)+3\left(M_1-2M_2+M_3\right)F_B^{*2}(x)+$$
$$+\left(1-3M_1+3M_2-M_3\right)F_B^{*3}(x),$$

gdzie $M_i=M_\Theta\left(i\ln(r)\right)=g^{-1}\left(g(q)^i\right)=\varphi^{-1}\left(i\varphi(q)\right)$.

Widzimy, że dystrybuanta sumy wypłat jest kombinacją liniową dystrybuant:

$$F_S(x)=\sum_{k=0}^n a_k F_B^{*k}(x).$$

Jest to kombinacja wypukła, ponieważ można sprawdzić, że $a_k \geq 0$ oraz $\sum_{k=0}^{n} a_k = 1$. Dystrybuantę sumy można wtedy traktować jako mieszankę splotów dystrybuant F_B. Wartości współczynników a_k są kombinacjami liniowymi $M_i =$ $= M_\Theta(i\ln(r))$, zależą więc one jedynie od rozkładu zmiennej Θ, czyli od funkcji łączącej C, oraz od wartości prawdopodobieństwa niewystąpienia wypłaty q.

W przypadku, gdy indykatory I_j są niezależne, funkcja łącząca, opisująca ich zależność, jest równa $C(u_1,...u_n) = u_1 \cdot ... \cdot u_n$ oraz $a_k = \binom{n}{k} q^{n-k}(1-q)^k$. Rozkład mieszający jest wtedy rozkładem dwumianowym.

Wartość oczekiwana sumy wypłat jest w tym wypadku równa

$$E(S) = n\mu p,$$

gdzie $\mu = E(B_j)$, a wariancja

$$V(S) = np(\mu_2 - \mu^2 p) + n(n-1)(\mu^2(M_\Theta(2\ln r) - q^2)),$$

gdzie $\mu_2 = E(B_j^2)$.

Przykład 4 (ciąg dalszy przykładu 3)

1. W przypadku, gdy zmienne losowe B_j o rozkładzie wykładniczym oraz I_j mają ten sam rozkład, to funkcja tworząca momenty sumy wypłat przybiera postać:

$$M_S(t) = \int_0^\infty \left(\frac{1 - r^\theta \lambda t}{1 - \lambda t}\right)^n dF_\Theta(\theta).$$

Natomiast $F_B^{*k}(x)$, dla $k \geq 1$, są dystrybuantami rozkładu gamma G(k, λ).

2. Przyjmijmy teraz, że n = 3, λ = 1, q = 0,9 oraz struktura zależności zmiennych losowych I_j określona jest funkcją łączącą Claytona z parametrem α = 2. Wtedy zmienna Θ ma rozkład gamma G(0,5; 1), $M_\Theta(t) = (1-t)^{-0,5}$, r = 0,791, $M_1 = 0,9$; $M_2 = 0,825$; $M_3 = 0,766$ oraz

$$F_S(x) = 0,766 + 0,177F_B(x) + 0,048F_B^{*2}(x) + 0,09F_B^{*3}(x).$$

Załóżmy, że struktura zależność między zmiennymi I_j opisana jest funkcją łączącą Claytona z parametrem α. Jeśli $\alpha \leq \beta$, to składka zatrzymania straty $\pi_\alpha(d)$

sumy wypłat S_α jest niewiększa niż składka $\pi_\beta(d)$ dla każdego $d \geq 0$ (por. [5]). Wynika to z faktu, że $\varphi_\alpha \circ \varphi_\beta^{-1}$ jest transformatą Laplace'a. Mówimy wtedy, że suma S_α jest zdominowana przez sumę S_β w porządku zatrzymania straty. Własność tę mają również rodziny Gumbela i Franka.

3.2. Funkcje łączące, przypadek ogólny

Załóżmy teraz, że struktura zależności indykatorów I_j opisana jest dowolną n-wymiarową funkcją łączącą, niekoniecznie archimedesową. Można pokazać (zob. [3]), że funkcja tworząca momenty losowego wektora **X** przybiera wtedy postać

$$M_{\mathbf{X}}(t_1, \ldots, t_n) = \mathrm{P}_{\mathbf{I}}\left(\mathrm{M}_{B_1}(t_1), \ldots, \mathrm{M}_{B_n}(t_n)\right).$$

Funkcja tworząca prawdopodobieństwa wektora **I** jest równa:

$$P_{\mathbf{I}}(t_1, \ldots, t_n) = E\left(t_1^{I_1}, \ldots, t_n^{I_n}\right) = \sum_{i_1, \ldots, i_n \in \{0,1\}} f_{\mathbf{I}}(i_1, \ldots, i_n) t_1^{i_1}, \ldots, t_n^{i_n}, \tag{2}$$

gdzie $f_{\mathbf{I}}(i_1, \ldots, i_n) = P(I_1 = i_1, \ldots, I_n = i_n)$.

Prawdopodobieństwa $P(I_1 = i_1, \ldots, I_n = i_n)$ można wyznaczyć, korzystając ze znajomości funkcji łączącej C oraz prawdopodobieństw q_j (por. [3]). Niech $N = \{1, 2, \ldots, n\}$, a symbolem $\mathbf{P}_r$ będziemy oznaczać klasę jego r-elementowych podzbiorów. Dla każdego podzbioru $A \subset N$ możemy określić jego indykator $\mathbf{1}_A = (i_1, \ldots, i_n) \in \{0,1\}^n$, gdzie $i_j = \begin{cases} 1 & \text{gdy } j \in A \\ 0 & \text{gdy } j \notin A \end{cases}$. Dla każdego $A \in \mathbf{P}_r$ otrzymujemy

$$f_{\mathbf{I}}(\mathbf{1}_A) = \sum_{k=0}^{r} (-1)^k \sum_{\substack{D \subset A, \\ D \in \mathbf{P}_{r-k}}} F(\mathbf{1}_D).$$

Przykład 5

Niech $N = \{1, 2, 3, 4, 5\}$ i $A = \{2, 4, 5\}$, wtedy $\mathbf{1}_A = (0, 1, 0, 1, 1)$ oraz $f_{\mathbf{I}}(\mathbf{1}_A) = F(0, 1, 0, 1, 1) - F(0, 1, 0, 1, 0) - F(0, 1, 0, 0, 1) - F(0, 0, 0, 1, 1) + F(0, 1, 0, 0, 0) + F(0, 0, 0, 1, 0) + F(0, 0, 0, 0, 1) - F(0, 0, 0, 0, 0)$.

Gdy zmienne I_j mają ten sam rozkład, to łączną dystrybuantę wektora **I** możemy przedstawić, korzystając z funkcji łączącej C, w postaci:

$$F_{\mathbf{I}}(i_1, \ldots, i_n) = C\left(F_1(i_1), \ldots, F_n(i_n)\right),$$

gdzie $F_j(0)=q$ oraz $F_j(1)=1$. Jeśli wielkości wypłat B_j mają również ten sam rozkład o dystrybuancie F_B, to funkcja tworząca momenty sumy jest równa:

$$M_S(t)=\sum_{i_1,\ldots,i_n\in\{0,1\}} f_I(i_1,\ldots,i_n)(M_B(t))^{\sum_{k=1}^{n} i_k}.$$

Gdy odwrócimy funkcję tworzącą momenty $M_S(t)$, otrzymujemy postać dystrybuanty rozkładu sumy wypłat. Jest ona również kombinacją wypukłą wielokrotnych splotów dystrybuanty F_B:

$$F_S(x)=\sum_{k=0}^{n} a_k F_B^{*k}(x),$$

gdzie współczynniki tej kombinacji $a_k=\sum_{D\in \mathbf{P}_k} f_I(\mathbf{1}_D)$ zależą jedynie od funkcji łączącej C oraz prawdopodobieństwa q.

Załóżmy teraz, że funkcja łącząca C jest symetryczna. Oznaczmy symbolem $c_{j,n}$ n-wymiarowy wektor złożony z j jedynek na dowolnym miejscu. Wtedy

$$f_I\left(c_{j,n}\right)=\sum_{k=0}^{j}(-1)^k\binom{j}{k}F_I\left(c_{j-k,n}\right).$$

W sytuacji, gdy zmienne losowe I_j, tak jak i B_j, mają ten sam rozkład, a funkcja łącząca jest symetryczna to funkcja tworząca momenty sumy wypłat wynosi

$$M_S(t)=\sum_{j=0}^{n}\binom{n}{j}f_I\left(c_{j,n}\right)(M_B(t))^j.$$

Natomiast dystrybuanta tego rozkładu jest równa

$$F_S(x)=\sum_{j=0}^{n}\binom{n}{j}f_I\left(c_{j,n}\right)F_B^{*j}(x).$$

Jest to również kombinacja wypukła splotów dystrybuanty $F_B(x)$. W przypadku łączących funkcji archimedesowych otrzymujemy

$$F_I\left(c_{j-k,n}\right)=\varphi^{-1}\left((j-k)\varphi(q)\right).$$

Przykład 6

Załóżmy, że $n = 3$, prawdopodobieństwo wypłaty $p = 0,1$, a zależność zmiennych I_j opisana jest funkcją łączącą Morgensterna $C(u_1, u_2, u_3) = u_1u_2u_3(1 + 0,5(1-u_1)(1-u_2) - 0,8(1-u_1)(1-u_3) + 0,7(1-u_1)(1-u_2) + 0,2(1-u_1)(1-u_2)(1-u_3))$. Wtedy dystrybuanta rozkładu sumy wypłat jest równa

$$F_S(x) = 0,733 + 0,228F_B(x) + 0,038F_B^{*2}(x) + 0,001F_B^{*3}(x).$$

Przykładowo otrzymujemy

$$a_0 = f_I(0,0,0) = F_I(0,0,0) = C(q,q,q) = 0,733,$$

$$a_1 = f_I(1,0,0) + f_I(0,1,0) + f_I(0,0,1) = F_I(0,0,0) + F_I(0,1,0) + F_I(0,0,1) - \\ -3F_I(0,0,0) = C(1,q,q) + C(q,1,q) + C(q,q,1) = 3C(q,q,q) = 0,228.$$

Funkcja łącząca Morgensterna C nie jest symetryczna, np. $C(1,q,q) = 0,817$, a $C(q,1,q) = 0,802$. Natomiast w przypadku niezależnych zmiennych losowych I_1, I_2, I_3 dystrybuanta sumy przybiera postać

$$F_S(x) = 0,729 + 0,243F_B(x) + 0,027F_B^{*2}(x) + 0,001F_B^{*3}(x).$$

Gdy funkcja łącząca jest równa $C(u_1,\dots,u_n) = \min(u_1,\dots,u_n)$, tzn. gdy zmienne I_j są współmonotoniczne (por. [11]), to dystrybuanta sumy przyjmuje prostą postać:

$$F_S(x) = q + (1-q)F_B^{*n}(x).$$

4. Wieloklasowy model ryzyka kolektywnego

Zakładamy, że szkody są podzielone na m klas uwzględniających np. charakter szkód lub ich regionalizację. W każdej klasie o określonym czasie nastąpiło N_i wypłat X_{ij}, gdzie N_i jest zmienną losową przyjmującą wartości całkowite, a $i = 1,\dots,m$. Łączna wypłata S jest równa sumie wszystkich łącznych wypłat S_i w każdej klasie, czyli

$$S = S_1 + \dots + S_m.$$

Natomiast łączna wypłata dla i-tej klasy jest równa

$$S_i = \sum_{j=1}^{N_i} X_{ij}.$$

Przyjmujemy, że wielkości wypłat X_{ij} są niezależne od liczebności N_i oraz że dla ustalonej klasy i wielkości wypłat $X_{i1}, X_{i2}, ..., X_{iN_i}$ są niezależne i mają ten sam rozkład. Zmienną losową o tym rozkładzie będziemy oznaczać umownie symbolem $X^{(i)}$. Liczebności wypłat N_i mogą być zależne, a strukturę zależności opisano funkcją łączącą C.

Funkcja tworząca momenty sumy S jest wtedy równa (por. [4])

$$M_S(t) = M_S(t, ..., t),$$

gdzie $\mathbf{S} = (S_1, ..., S_m)$, a

$$M_S(t_1, ..., t_m) = E\left(e^{t_1 S_1} ... e^{t_m S_m}\right) = E\left(E\left(e^{t_1 S_1} ... e^{t_m S_m} \mid \mathbf{N}\right)\right),$$

gdzie $\mathbf{N} = (N_1, ..., N_m)$. Zmienne losowe $X_{i1}, X_{i2}, ..., X_{iN_i}$ są dla każdej klasy i niezależne, dlatego też

$$M_S(t_1, ..., t_m) = E\left((M_{X^{(1)}}(t_1))^{N_1}, ..., \left(M_{X^{(m)}}(t_m)\right)^{N_m}\right) =$$

$$= \mathrm{P_N}\left(P_\mathrm{N}\left(M_{X^{(1)}}(t_1), ..., M_{X^{(m)}}(t_m)\right)\right).$$

Aby wyznaczyć funkcję tworzącą momenty sumy S, należy obliczyć funkcję tworzącą prawdopodobieństwa $\mathrm{P_N}(t_1, ..., t_m)$. Jest ona równa

$$P_\mathrm{N}(t_1, ..., t_m) = E\left(t_1^{N_1} ... t_m^{N_m}\right) = \sum_{n_1=0}^{\infty} ... \sum_{n_m=0}^{\infty} t_1^{n_1} ... t_m^{n_m} P(N_1 = n_1, ..., N_m = n_m).$$

W tym przypadku, w odróżnieniu od indywidualnego modelu ryzyka, gdzie indykatory I_j przyjmowały tylko wartości 0 lub 1, zmienne losowe N_j mogą przyjąć dowolną wartość ze zbioru liczb naturalnych. Powoduje to duże kłopoty obliczeniowe, nawet dla małej liczby klas. Z tego też powodu zastosowanie w tej sytuacji funkcji łączących do modelowania zależności jest bardzo ograniczone.

5. Podsumowanie

W pracy przedstawiono klasyczne modele aktuarialne dopuszczające zależność między zmiennymi losowymi, w których zależność ta modelowana jest za pomocą funkcji łączących. W modelu ryzyka indywidualnego metoda ta zdaje egzamin.

Gdy zależność opisujemy archimedesowymi funkcjami łączącymi, korzystamy z warunkowej niezależności ryzyk i podczas wyznaczania rozkładu sumy wypłat możemy skorzystać z klasycznych metod. W przypadku jednakowych rozkładów indykatorów i wypłat otrzymujemy, nawet dla ogólnych funkcji łączących, kombinację wypukłą wielokrotnych splotów dystrybuant. Sytuacja komplikuje się, gdy rozpatrujemy wieloklasowy model ryzyka kolektywnego. Wtedy, stosując do modelowania zależności funkcje łączące, natrafiamy poważne kłopoty obliczeniowe. Lepsza jest wtedy metoda omówiona we wstępie, przedstawiająca zmienne występujące w modelu w postaci sumy niezależnych składników.

Archimedesowej funkcji łączącej odpowiada ukryta zmienna losowa Θ interpretowana jako wpływ czynników zewnętrznych. Podejście to możemy uogólnić, rozpatrując nie jedną, ale kilka zmiennych losowych $\Theta_1, ..., \Theta_p$, co zostało zasygnalizowane w publikacji [5]. Autor zamierza rozwinąć to zagadnienie w następnych pracach.

Literatura

[1] Bowers N.L., Gerber H. U., Hickman J.C., Jones D.A., Nesbitt C.J., *Actuarial Mathematics*, Society of Actuaries, Schaumburg 1986.

[2] Cossette H., Gaillardetz P., Marceau E., *Common Mixture in the Individual Risk Model*, Mitteilungen der Schweiz, Aktuarvereiningung 2, 2002, 131-157.

[3] Cossette H., Gaillardetz P., Marceau E., Rioux J., *On Two Dependent Individual Risk Models*, „Insurance: Mathematics and Economics" 2002, 30, 153-166.

[4] Cossette H., Marceau E., *The Discrete-Time Risk Model with Correlated Classes of Business*, „Insurance: Mathematics and Economics" 2000, 26, 133-149.

[5] Genest G., Marceau E., Mesfioui M., *Compound Poisson Approximations for Individual Models*, „Insurance: Mathematics and Economics" 2003, 32, 73-91.

[6] Marshall A.W., Olkin I., *Families of Multivariate Distributions*, „Journal of the American Statistical Association" 1988, 83, 834-841.

[7] Nelsen R.B., *An Introduction to Copulas*, Springer, New York 1999.

[8] *Modele aktuarialne*, red. W. Ostasiewicz, AE, Wroclaw 2000.

[9] Rolski T., Schmidli H., Schmidt V., Teugels, *Stochastic Processes for Insurance and Finance*, Wiley, New York 1999.

[10] Schweitzer B., Sklar A., *Probabilistic Metric Spaces*, North-Holland, New York 1983.

[11] Wang S., *Aggregation of Correlated Risk Portfolios: Models and Algorithms*, CAS Proceedings, 1998, 848-939.

ANALYSIS OF DEPENDENT ACTUARIAL RISKS - APPLICATION OF COPULAS

Summary

The paper is devoted to the analysis of dependent actuarial risks based on the copulas. The strong assumption about independence of random variables is weaken in the investigated models. The dependence between some random variables is admitted. Such models are more realistic and they better describe the examined problems.

The basic information connected with copulas, which are the main tools for study dependence, is presented in our paper. The Archimedean copulas, which induce the latent random variables treated as the influence of the common shock, are described. The distribution of the sum of dependent random variables is studied. The individual model of risk with dependent indicators is presented. This dependence is modeled by general and Archimedean copulas. The last part of our paper is devoted to the attempt of the application of copulas to the description of dependences in the multi-class collective risk model.