

Stanisław Heilpern

Akademia Ekonomiczna we Wrocławiu

MODELOWANIE ZALEŻNOŚCI – ZASTOSOWANIE W ZARZĄDZANIU RYZYKIEM

1. Wstęp

W klasycznych modelach związanych z zarządzaniem ryzyka, rozpatrując zależności zachodzące między zmiennymi losowymi opisującymi ryzyko, przyjmuje się zwykle ich niezależność. Jest to założenie bardzo wygodne z punktu widzenia rozważań teoretycznych, ale na ogół znacznie odbiegające od rzeczywistości. W praktyce mamy niestety zwykle do czynienia ze zmiennymi zależnymi. Artykuł ma charakter przeglądowy, przedstawione w nim będą wybrane sposoby modelowania zależności i przykłady wykorzystania ich w problemach związanych z zarządzaniem ryzykiem, głównie w zagadnieniach aktuarialnych oraz wypłacalności kredytowej.

Pierwszy sposób modelowania zależności zaprezentowany w pracy polega na przedstawianiu zależnych zmiennych losowych jako kombinacji liniowej zmiennych niezależnych. Drugi opiera się na mieszkankach rozkładów i wykorzystuje warunkową niezależność rozpatrywanych zmiennych losowych. Natomiast trzeci sposób polega na opisanu zależności zmiennych za pomocą funkcji łączących (ang. *copula*).

Przedstawione metody umożliwiają sprowadzenie rozpatrywanych modeli do postaci, w której występują niezależne zmienne losowe. Można wtedy stosować klasyczne metody wykorzystywane w przypadku niezależności. W pracy omawia się też metody przybliżone oraz symulacyjne.

W artykule będziemy się zajmować sumą losowej liczby zmiennych losowych:

$$S = X_1 + X_2 + \dots + X_N.$$

W klasycznym modelu zakładamy niezależność występujących zmiennych losowych, natomiast w naszej pracy dopuszczamy zależność. Powyższy model będzie rozszerzony przez podzielenie składników sumy na klasy lub uproszczony przez ustalenie liczby zmiennych losowych X_i .

2. Kombinacja liniowa zmiennych niezależnych

Założmy, że mamy m klas, a w i -tej klasie mamy N_i wypłat, gdzie $i = 1, \dots, n$. Całkowita suma wszystkich odszkodowań jest wtedy równa

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

gdzie

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

a X_{ij} jest wielkością j -tej wypłaty w i -tej klasie.

W klasycznym modelu kolektywnego ryzyka, bo na tym modelu oparty jest rozważany przypadek, zakłada się, że występujące zmienne losowe są niezależne. W tym przypadku zmienne losowe $N_1, \dots, N_m$ opisujące liczebności poszczególnych klas mogą być zależne. Przyjmiemy w pracy, że powyższe zmienne losowe będą liniowymi kombinacjami niezależnych zmiennych losowych $L_1, \dots, L_k$. Losowy wektor $\mathbf{N} = [N_1, \dots, N_m]^T$ będzie spełniać zależność [1]

$$\mathbf{N} = \mathbf{A}\mathbf{L},$$

gdzie $\mathbf{L} = [L_1, \dots, L_k]^T$, $\mathbf{A} = [a_{ij}]_{m \times k}$ oraz $a_{ij} \in \mathbb{N}$. Zwykle przyjmuje się, że elementy macierzy współczynników $\mathbf{A}$ są stałymi binarnymi.

W modelu tym zakłada się, że w każdej klasie i zmienne losowe X_{ij} opisujące wielkości wypłat są niezależne i mają ten sam rozkład o gęstości f_i . Zmienną losową o tym rozkładzie będziemy oznaczali symbolem X_i . Ponadto przyjmuje się, że zmienne losowe X_{ij} oraz N_i są niezależne. W przypadku identycznościowej macierzy współczynników, tzn. gdy $\mathbf{A} = \mathbf{I}$, otrzymujemy niezależność zmiennych losowych N_i .

W rozpatrywanym modelu będzie nas interesować całkowita suma wypłat S . Funkcja tworząca momenty tej zmiennej losowej określona jest wzorem [1]

$$M_S(t) = \prod_{j=1}^k P_{L_j} \left(\prod_{i=1}^m M_{X_i}(t)^{a_{ij}} \right),$$

gdzie P_L jest funkcją tworzącą prawdopodobieństwa zmiennej losowej L . Z jej postaci wynika, że całkowitą sumę wypłat S możemy przedstawić jako sumę k niezależnych składników:

$$S = S'_1 + \dots + S'_k,$$

gdzie $S'_j = X'_{j1} + \dots + X'_{jL_j}$, a w każdej nowej klasie j zmienne losowe X'_{jl} są niezależne i mają ten sam rozkład. Zmienną losową o tym rozkładzie będziemy oznaczać symbolem X'_j . Jej gęstość jest równa

$$f_{X'_j}(x) = (f_{X_1}^{*(a_{1j})} * \dots * f_{X_m}^{*(a_{mj})})(x),$$

gdzie symbol $*$ oznacza spłot, a $f_X^{*(a)}$ a -krotny spłot gęstości zmiennej losowej X . W przypadku binarnych współczynników a_{ij} wartość 1 powoduje uwzględnienie w tym splocie danego składnika, a wartość 0 brak jego uwzględnienia. Otrzymujemy więc klasyczny, wieloklasowy model kolektywnego ryzyka, gdzie występujące zmienne losowe są niezależne.

Przyjmijmy teraz, że zmienne losowe L_j mają rozkład Poissona z parametrem λ_j , tzn. $L_j \sim \text{Po}(\lambda_j)$. Wtedy funkcja tworząca momenty całkowitej sumy S jest równa [1]

$$M_S(t) = \exp\left(\sum_{j=1}^k \lambda_j \left(\prod_{i=1}^m (M_{X_j}(t))^{a_{ij}} - 1\right)\right).$$

Zmienna losowa S ma więc złożony rozkład Poissona $\text{CP}(\lambda, f_{X'}(x))$, gdzie parametr $\lambda = \sum_{j=1}^k \lambda_j$, a gęstości zmiennych losowych X' są określone wzorem

$$f_{X'}(x) = \frac{1}{\lambda} \sum_{j=1}^k \lambda_j (f_{X_1}^{*(a_{1j})} * \dots * f_{X_m}^{*(a_{mj})})(x).$$

W zastosowaniach [3; 13] często wykorzystuje się model przedstawiający wpływ różnych czynników na poszczególne klasy, tzw. model wspólnego czynnika (ang. *the common shock model*). Przedstawimy teraz prosty przypadek opisujący wpływ różnych czynników na trzy klasy. Przybiera on postać [3]

$$N_1 = L_1 + L_{12} + L_{13} + L_{123},$$

$$N_2 = L_2 + L_{12} + L_{23} + L_{123},$$

$$N_3 = L_3 + L_{13} + L_{23} + L_{123}.$$

W tym przypadku mamy $\mathbf{N} = [N_1, N_2, N_3]^T$, $\mathbf{L} = [L_1, L_2, L_3, L_{12}, L_{13}, L_{23}, L_{123}]^T$ oraz

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Zmienne losowe L_1, L_2, L_3 oddają wpływ wewnętrznych czynników wpływających wyłącznie na pojedyncze klasy, zmienne L_{12}, L_{13}, L_{23} oddają wpływ czynników oddziałujących na pary klas, a zmienna L_{123} reprezentuje wpływ wspólnego czynnika zewnętrznego działającego na wszystkie klasy.

Jeśli zmienne losowe będące składowymi wektora $\mathbf{L}$ mają rozkład Poissona z parametrami odpowiednio równymi: $\lambda_1, \lambda_2, \lambda_3, \lambda_{12}, \lambda_{13}, \lambda_{23}, \lambda_{123}$, to całkowita suma wypłat S ma złożony rozkład Poissona $CP(\lambda, f_{X'}(x))$, gdzie $\lambda = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_{12} + \lambda_{13} + \lambda_{23} + \lambda_{123}$, a zmienna losowa X ma gęstość

$$f_{X'}(x) = \frac{\lambda_1}{\lambda} f_{X_1}(x) + \frac{\lambda_2}{\lambda} f_{X_2}(x) + \frac{\lambda_3}{\lambda} f_{X_3}(x) + \frac{\lambda_{12}}{\lambda} f_{X_1+X_2}(x) + \frac{\lambda_{13}}{\lambda} f_{X_1+X_3}(x) + \frac{\lambda_{23}}{\lambda} f_{X_2+X_3}(x) + \frac{\lambda_{123}}{\lambda} f_{X_1+X_2+X_3}(x).$$

Powyższy model może być zastosowany w zagadnieniach ruiny z czasem dyskretnym [3]. Rozpatruje się wtedy proces ryzyka

$$U_n = u + cn - S_{(n)},$$

gdzie u jest kapitałem początkowym, c – składką otrzymaną w każdym okresie $n = 0, 1, \dots$, a $S_{(n)}$ – sumą wszystkich wypłat dokonanych w pierwszych n okresach. Ponadto zakłada się, że wypłaty podzielone są na m klas, a zależność między liczebnościami każdej klasy jest modelowana za pomocą liniowej kombinacji niezależnych zmiennych losowych. Można też powyższe rozważania zastosować do modelu indywidualnego ryzyka [5; 8].

3. Wspólne mieszanki

W tym przypadku zależność liczebności poszczególnych klas $N_1, \dots, N_m$ będziemy modelować za pomocą dodatkowej dodatniej zmiennej losowej Θ reprezentującej wpływ wspólnego, zewnętrznego czynnika oddziałującego na wszystkie klasy. Zakładamy przy tym warunkową niezależność zmiennych losowych N_i dla ustalonej wartości θ zmiennej Θ .

Najczęściej stosowanym i zarazem najprostszym modelem wykorzystującym wspólne mieszanki jest model oparty na rozkładzie Poissona [19]. Przyjmuje się w nim, że zmienne losowe N_i wyrażające liczebności klas mają dla ustalonej wartości θ warunkowy rozkład Poissona, tzn.

$$N_i | (\Theta = \theta) \sim \text{Po}(\theta \lambda_i),$$

gdzie $\lambda_i > 0$ dla każdego $i = 1, \dots, m$. Ponadto zakłada się niezależność warunkowych zmiennych losowych $N_i | (\Theta = \theta)$. Zmienna losowa mieszająca Θ wpływa na wszystkie liczebności N_i , powodując ich zależność.

Przy tak określonych założeniach modelu, można wyznaczyć funkcję tworzącą prawdopodobieństwa łącznego rozkładu zmiennych $N_1, \dots, N_m$. Jest ona mieszanką funkcji łączącej prawdopodobieństwa łącznego warunkowego rozkładu $N_1, \dots, N_m$ oraz rozkładu zmiennej Θ . Określona jest formułą [19]

$$P_{N_1, \dots, N_m}(t_1, \dots, t_m) = M_{\Theta}(\lambda_1(t_1 - 1) + \dots + \lambda_m(t_m - 1)).$$

Natomiast współczynnik korelacji między zmiennymi N_i , N_j zależy jedynie od wartości oczekiwanej i wariancji zmiennej losowej Θ i wynosi

$$\rho(N_i, N_j) = \frac{\text{Var}(\Theta)}{(E(\Theta))^2}.$$

W rozpatrywanym modelu możemy też w prosty sposób wyznaczyć funkcję tworzącą momenty całkowitej sumy wypłat

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

gdzie $S_i = \sum_{j=1}^{N_i} X_{ij}$ jest sumą wypłat w i -tej klasie. Jest ona równa

$$M_S(t) = M_{\Theta}(\lambda(M_X(t) - 1)),$$

gdzie $M_X(t) = \frac{1}{\lambda} \sum_{i=1}^m \lambda_i M_{X_i}(t)$ oraz $\lambda = \lambda_1 + \dots + \lambda_m$. W przypadku gdy zmienna

losowa Θ przedstawiająca czynnik zewnętrzny ma rozkład gamma, to całkowita suma S ma złożony rozkład ujemny dwumianowy.

W innym modelu [3] rozpatruje się dodatkowo, oprócz mieszającej zmiennej globalnej Θ , wpływ czynników wewnętrznych. Przyjmuje się w nim, że liczebność każdej klasy N_i , gdzie $i = 1, \dots, m$, jest sumą dwóch zmiennych losowych:

$$N_i = N_{ii} + N_{i0}.$$

Zmienne losowe N_{ii} przedstawiają wpływ czynników wewnętrznych działających tylko na daną klasę. Zakłada się w tym modelu, że zmienne te mają rozkład ujemny dwumianowy $NB(\alpha_{ii}, \beta_i)$ oraz że są niezależne. Drugi składnik N_{i0} reprezentuje wpływ wspólnego czynnika zewnętrznego na każdą klasę. Ma on warunkowy rozkład Poissona, tzn.

$$N_{i0} | (\Theta = \theta) \sim P(\theta\beta_i),$$

oraz, podobnie jak w poprzednim modelu, zachodzi, przy ustalonej wartości θ zmiennej mieszającej Θ , warunkowa niezależność zmiennych losowych N_{i0} . Ponadto przyjmuje się, że zmienna losowa Θ ma rozkład gamma $G(\alpha_0, 1)$. Zmienna ta powoduje zależność zmiennych losowych N_i .

Przy tak określonych założeniach funkcja tworząca momenty całkowitej sumy wypłat S jest równa

$$M_S(t) = \prod_{i=1}^m (1 - \beta_i(M_{X_i}(t) - 1))^{-\alpha_{ii}} (1 - \sum_{i=1}^m \beta_i(M_{X_i}(t) - 1))^{-\alpha_0}.$$

4. Funkcje łączące

Do modelowania zależności zachodzących między zmiennymi losowymi można wykorzystać też funkcje łączące (ang. *copula*). Funkcja łącząca $C: [0, 1]^n \rightarrow [0, 1]$ jest łącznikiem między dystrybuantami brzegowymi F_i a dystrybuantą F rozkładu łącznego zmiennych losowych $Z_1, \dots, Z_n$. Spełniona jest bowiem następująca zależność:

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

Funkcja łącząca jest n -wymiarową dystrybuantą, której brzegowe dystrybuanty mają rozkład jednostajny na $[0, 1]$.

Najczęściej w praktyce stosuje się tzw. archimedesowe funkcje łączące. Przybierają one postać

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

gdzie $\varphi: [0, 1] \rightarrow [0, +\infty]$ jest malejącą funkcją nazywaną generatorem. Generator separuje zmienne i jednoznacznie wyznacza archimedesową funkcję łączącą.

W zagadnieniach praktycznych korzysta się głównie z parametryzowanych rodzin archimedesowych funkcji łączących. Przykładem takiej rodziny jest rodzina Claytona, której elementy zadane są wzorem

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

gdzie parametr $\alpha > 0$. Funkcje łączące są ściśle związane z rangowymi współczynnikami korelacji: Kendalla i Spearmana. Przykładowo dla funkcji łączących należących do rodziny Claytona współczynnik Kendalla jest równy

$$\tau = \frac{\alpha}{\alpha + 2}.$$

Więcej informacji dotyczących funkcji łączących czytelnik znajdzie w [15; 6].

Archimedesowa funkcja łącząca C określona generatorem φ indukuje dodatnią zmienną losową Θ oraz dystrybuanty bazowe H_i , takie że, dla ustalonej wartości θ zmiennej Θ , brzegowe dystrybuanty warunkowe $F_{i|\theta}(z) = \Pr(Z_i \leq z | \Theta = \theta)$ spełniają zależność [19; 6]

$$F_{i|\theta}(z) = (H_i(z))^\theta.$$

Zachodzi też wtedy warunkowa niezależność, tzn.

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

gdzie F_θ jest warunkową dystrybuantą rozkładu łącznego zmiennych losowych $Z_1, \dots, Z_n$.

Jest to tzw. model słabości (ang. *frailty*) [14]. Indukowaną zmienną losową Θ interpretuje się zwykle jako ukrytą zmienną reprezentującą oddziaływanie czynnika zewnętrznego. Funkcja odwrotna do generatora φ archimedesowej funkcji łączącej jest równa transformacji Laplace'a zmiennej losowej Θ , tzn. zachodzi zależność

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

Ponadto generator φ wyznacza też dla każdego rozkładu brzegowego odpowiadającą mu dystrybuantę bazową H_i . Jest ona określona wzorem

$$H_i(z) = e^{-\varphi(F_i(z))}.$$

W przypadku gdy zależność zmiennych losowych $Z_1, \dots, Z_n$ opisana jest funkcją łączącą Claytona C_{α} , ukryta zmienna losowa Θ ma rozkład gamma $G\left(\frac{1}{\alpha}, 1\right)$.

Łączną dystrybuantę zmiennych losowych $Z_1, \dots, Z_n$ można przedstawić w postaci mieszanki warunkowej dystrybuanty łącznej F_{θ} oraz dystrybuanty ukrytej zmiennej losowej F_{Θ} :

$$F(z_1, \dots, z_n) = \int_0^{\infty} F_{\theta}(z_1, \dots, z_n) dF_{\Theta}(\theta),$$

a także za pomocą funkcji tworzącej momenty ukrytej zmiennej losowej Θ oraz z dystrybuant bazowych H_i :

$$F(z_1, \dots, z_n) = M_{\Theta}\left(\sum_{i=1}^n \ln(H_i(z_i))\right).$$

4.1. Sumy zależnych szkód

Zadaniem naszym będzie wyznaczenie sumy n szkód

$$S = X_1 + \dots + X_n$$

w przypadku gdy zależność zmiennych losowych $X_1, \dots, X_n$ opisujących wielkości szkód jest przedstawiona za pomocą archimedesowej funkcji łączącej C z generatorem φ . Funkcja łącząca C indukuje nam ukrytą zmienną losową Θ . Korzystając z warunkowej niezależności zmiennych losowych X_i przy ustalonej wartości θ zmiennej losowej Θ oraz z klasycznych metod wyznaczania dystrybuanty i funkcji tworzącej momenty sumy niezależnych zmiennych losowych, możemy obliczyć bezwarunkową dystrybuantę lub funkcję tworzącą momenty sumy szkód S . Dystrybuanta jest określona wzorem

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

Jest to mieszanka warunkowej dystrybuanty sumy oraz dystrybuanty zmiennej ukrytej Θ . W podobny sposób możemy wyznaczyć funkcję tworzącą momenty sumy szkód S :

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

4.2. Indywidualny model ryzyka

Indywidualny model ryzyka należy do klasycznych modeli stosowanych w statystyce aktuarialnej [2; 9; 16]. Rozpatruje się w nim ustaloną liczbę polis, przyjmując zasadę, że z każdej polisy może być dokonana co najwyżej jedna wypłata. Suma wszystkich wypłat jest określona wzorem

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

gdzie $X_j = I_j B_j$, B_j jest wielkością wypłaty z i -tej polisy, gdy do niej dojdzie, a I_i jest indykátorem określonym formułą

$$I_i = \begin{cases} 1, & \text{gdy nastąpiła wypłata z } i\text{-tej polisy,} \\ 0, & \text{gdy nie było wypłaty z } i\text{-tej polisy.} \end{cases}$$

Wystąpienia wypłaty z każdej polisy określone są prawdopodobieństwami

$$p_j = \Pr(I_j = 1), \quad q_j = 1 - p_j = \Pr(I_j = 0).$$

W klasycznym modelu ryzyka indywidualnego przyjmuje się niezależność występujących zmiennych losowych. Założenie to osłabimy, dopuszczając zależność indykatorów I_i . Przyjmiemy, że zależność $I_1, \dots, I_n$ jest opisana archimedesową funkcją łączącą C z generatorem φ , indukującą ukrytą zmienną losową Θ . Wtedy, dla ustalonej wartości θ zmiennej losowej Θ , warunkowe indykatory $I_{i\theta}$ mają również rozkład 0-1 określony prawdopodobieństwem [4]

$$r_j^\theta = \Pr(I_j = 0 \mid \Theta = \theta),$$

gdzie $r_j = \exp(-\varphi(q_j))$. Ponadto zachodzi zależność $q_j = M_\Theta(\ln(r_j))$.

Warunkowa funkcja tworząca momenty całkowitej sumy wypłat S_θ jest określona wzorem

$$M_{S_\theta}(t) = \prod_{j=1}^n (r_j^\theta + (1 - r_j^\theta) M_{B_j}(t)),$$

a bezwarunkowa funkcja tworząca momenty tej sumy jest mieszaną przedstawioną formułą

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

W podobny sposób możemy wyznaczyć bezwarunkową dystrybuantę całkowitej sumy wypłat:

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

gdzie $F_{S_\theta}(x)$ jest warunkową dystrybuantą tej sumy. Korzystając z warunkowej niezależności składników całkowitej sumy, możemy wyznaczyć warunkową dystrybuantę lub funkcję tworzącą momenty sumy S_θ , opierając się na metodach wykorzystywanych w klasycznej teorii ryzyka indywidualnego.

W przypadku gdy zarówno wielkości wypłat B_i , jak i indykatory I_i mają te same rozkłady charakteryzowane dystrybuantą F_B oraz prawdopodobieństwami niewystąpienia wypłaty q , dystrybuanta całkowitej wypłaty jest określona wzorem

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

Jest to kombinacja wypukła kolejnych splotów dystrybuant charakteryzujących wielkości wypłat. Współczynniki tej kombinacji są równe [4; 10]

$$a_k = \binom{n}{k} \sum_{j=0}^k (-1)^j \binom{k}{j} M_{n-k+j}, \quad (1)$$

gdzie $M_k = M_\Theta(k \ln(r)) = \varphi^{-1}(k \varphi(q))$ jest równe prawdopodobieństwu zdarzenia, że będzie co najmniej k polis bez wypłaty.

Funkcja łącząca $C(u_1, \dots, u_n) = u_1 \cdot \dots \cdot u_n$ opisuje niezależność indykatorów, czyli otrzymujemy w tej sytuacji klasyczny model indywidualnego ryzyka. Wtedy współczynniki kombinacji wypukłej opisane są prostszym wzorem

$$a_k = \binom{n}{k} q^{n-k} (1-q)^k.$$

Chcąc wyznaczyć rozkład całkowitej sumy wypłat, natrafiamy w większości przypadków, nawet dla klasycznego modelu ryzyka indywidualnego, na trudności natury obliczeniowej. W tym celu stosuje się aproksymację złożonym rozkładem Poissona. Opis tej aproksymacji w klasycznym przypadku niezależnych indykatorów czytelnik może znaleźć w [12; 18].

W sytuacji gdy zależność indykatorów I_i opisana jest archimedesową funkcją łączącą C z generatorem φ korzystamy z klasycznej metody aproksymacji oraz mieszanki opartej na ukrytej zmiennej losowej Θ indukowanej przez funkcję łączącą C [4; 8]. Wtedy warunkową sumę S_θ przybliżamy zmienną losową T_θ o złożonym rozkładzie Poissona $CP(p_\theta, F_{Z_\theta})$, gdzie $p_\theta = \sum_{i=1}^n p_{\theta i} = \sum_{i=1}^n (1 - r_i^\theta)$ oraz

$$F_{Z_\theta}(x) = \sum_{i=1}^n \frac{p_{\theta i}}{p_\theta} F_{B_i}(x).$$

Natomiast suma bezwarunkowa S jest przybliżana zmienną T o dystrybucie

$$F_T(x) = \int_0^\infty F_{T_\theta}(x) dF_\Theta(\theta).$$

Podobnie jak w przypadku klasycznym można oszacować błąd powyższej aproksymacji. Przykładowo, opierając się na odległości całkowitej zmienności między zmiennymi losowymi

$$d_{TV}(S, T) = \sup_A |\Pr(S \in A) - \Pr(T \in A)|,$$

gdzie A jest borelowskim zdarzeniem, oszacowanie to opiera się następującej nierówności [8]:

$$d_{TV}(S, T) \leq \sum_{i=1}^n (\varphi^{-1}(2\varphi(q_i)) + 1 - 2q_i).$$

W przypadku niezależności otrzymujemy znany wzór pochodzący od Gerbera [18]:

$$d_{TV}(S, T) \leq \sum_{i=1}^n p_i^2.$$

5. Ryzyko kredytowe

Załóżmy, że mamy portfel składający się z n dłużników, a zdolności spłaty zaciągniętych kredytów w danym okresie przez dłużników opisane są indykatorami

$$Y_i = \begin{cases} 1 & \text{nie spłaca,} \\ 0 & \text{spłaca} \end{cases}$$

oraz prawdopodobieństwami niespłacenia długu $p_i = \Pr(Y_i = 1)$. Zwykle zakłada się wymiennność (ang. *exchangeable*) łącznego rozkładu indykatorów $\mathbf{Y} = (Y_1, \dots, Y_n)$, tzn. przyjmuje się, że zachodzi warunek

$$(Y_1, \dots, Y_n) \stackrel{d}{=} (Y_{\Pi(1)}, \dots, Y_{\Pi(n)}),$$

gdzie Π jest permutacją zbioru $\{1, \dots, n\}$. Bada się też rozkład liczby niewypłacalnych dłużników [7]:

$$M = \sum_{i=1}^n Y_i,$$

oraz prawdopodobieństwo π_k zdarzenia, że przynajmniej k dłużników będzie niewypłacalnych

$$\pi_k = \Pr(Y_{i_1} = 1, \dots, Y_{i_k} = 1).$$

Powyższy model można rozszerzyć wprowadzając ukryte zmienne losowe $X_1, \dots, X_n$ przedstawiające aktywa poszczególnych dłużników, oraz ich progi płatności zaciągniętych kredytów $d_1, \dots, d_n$ [7]. Wtedy dłużnik nie będzie w stanie spłacić kredytu, jeśli jego aktywa nie przekraczają jego progu płatności, tzn. zachodzi

$$Y_i = 1 \Leftrightarrow X_i \leq d_i.$$

Jeśli zależność zmiennych losowych charakteryzujących aktywa dłużników $X_1, \dots, X_n$ opisuje funkcja łącząca C , to prawdopodobieństwo π_k jest wyznaczone przez brzegową k -wymiarową funkcję łączącą $C_{1,\dots,k}$ wzorem

$$\pi_k = C_{1,\dots,k}(p, \dots, p).$$

Wtedy prawdopodobieństwo zdarzenia, że będzie dokładnie k niewypłacalnych dłużników, jest równe

$$\Pr(M = k) = \sum_{i=0}^{n-k} (-1)^i \frac{n!}{i! k! (n-k-i)!} \pi_{k+i}.$$

Jest to wersja wzoru (1). Rozpatrując dodatkowo wielkości straconego kredytu $B_1, \dots, B_n$ otrzymujemy model indywidualnego ryzyka.

6. Metody symulacyjne

Jak już wspomniano, chcąc wyznaczyć dokładne rozkłady omawianych wcześniej zmiennych losowych, natrafiamy na ogół na poważne kłopoty obliczeniowe. Jednym ze sposobów rozwiązania tych problemów jest zastosowanie metod symulacyjnych. W przypadku gdy zależność zmiennych losowych $X_1, \dots, X_n$ jest

opisana za pomocą funkcji łączącej C , mamy ułatwioną sytuację, ponieważ istnieją proste procedury umożliwiające generowanie wartości z łącznego rozkładu tych zmiennych. Ogólna zasada generowania tych wielkości polega na wygenerowaniu liczb losowych $u_1, \dots, u_m$ z rozkładu łącznego $(U_1, \dots, U_m)$, gdzie zmienne brzegowe U_i mają rozkład jednostajny na $[0, 1]$, a ich zależność jest opisana funkcją łączącą C . Następnie wyznaczamy losowe wartości zmiennych X_i , odwracając dystrybuanty brzegowe, tzn. obliczamy $x_i = F_i^{-1}(u_i)$.

Liczby losowe $u_1, \dots, u_m$ możemy wyznaczyć, generując u_1 z rozkładu zmiennej U_1 , a następne wartości – generując z rozkładów warunkowych. Przykładowo liczbę u_k generujemy z rozkładu o warunkowej dystrybuancie

$$\Pr(U_k \leq t \mid U_1 = u_1, \dots, U_{k-1} = u_{k-1}) = \frac{\frac{\partial^{k-1}}{\partial u_1, \dots, \partial u_{k-1}} C(u_1, \dots, u_{k-1}, t, 1, \dots, 1)}{\frac{\partial^{k-1}}{\partial u_1, \dots, \partial u_{k-1}} C(u_1, \dots, u_{k-1}, 1, \dots, 1)}.$$

Procedurę tę możemy stosować w przypadku ogólnym. Jeśli zależność jest opisana archimedesową funkcją łączącą, stosuje się zwykle łatwiejszą metodę generowania liczb losowych wykorzystującą specyficzne własności tych funkcji łączących. Również dla każdej z rodzin funkcji łączących istnieją procedury specjalnie dla nich skonstruowane. Więcej na ten temat czytelnik znajdzie w [11; 15; 19; 17].

Dzięki metodom symulacyjnym możemy przykładowo badać rozkłady sumy zależnych zmiennych losowych. W [20] autorzy rozpatrzyli przykład, w którym metodami symulacyjnymi badali rozkład całkowitej sumy 4 klas szkód ubezpieczeniowych, gdy zależność składników była opisana różnymi funkcjami łączącymi. Wyznaczyli też podstawowe miary ryzyka: wartość narażoną na ryzyko VaR (ang. *Value-at-Risk*) i jego warunkową wersję CTE. Miary te, często wykorzystywane w zagadnieniach związanych z finansami i ubezpieczeniami, wyrażone są wzorami:

$$\text{VaR}_p(S) = \inf\{x \in \mathbb{R} : F_S(x) \geq p\}, \quad \text{CTE}_p(S) = E(S : S > \text{VaR}_p(S)).$$

W pracy tej zbadano też przykład związany z rozwojem firmy ubezpieczeniowej, dotyczący wprowadzenia na rynek nowego produktu. Rozpatrywany problem sprowadza się do badania rozkładu nowej zagregowanej sumy szkód

$$S' = S + X,$$

gdzie S jest całkowitą sumą szkód przed rozszerzeniem portfela polis, a X jest wielkością szkód dotyczącą nowego produktu.

Przykład

Przyjmijmy, że zależność zmiennych losowych S oraz X jest opisana funkcją łączącą Clayтона C_α , zmienna losowa S przedstawiająca wielkości szkód ma rozkład logarytmiczno-normalny $\text{LN}(2,7; 0,2)$, a zmienna X dotycząca nowego pro-

duktu ma rozkład Pareto $\text{Pa}(3, 4)$. W celu zbadania rozkładu sumy $S' = S + X$ przeprowadzono 5000 symulacji wartości tej sumy.

Tabela 1. Wartości VaR_p i CTE_p dla starego i nowego portfela

	Stary portfel S	Nowy portfel S'		
		$\tau = 0$	$\tau = 0,5$	$\tau = 0,91$
$\text{VaR}_{0,95}$	20,64	24,67	26,09	27,74
$\text{CTE}_{0,95}$	22,61	29,39	30,93	33,35
$\text{VaR}_{0,99}$	23,86	30,68	33,02	35,95
$\text{CTE}_{0,99}$	25,65	39,83	40,72	43,75

Źródło: obliczenia własne.

Wartości podstawowych miar ryzyka otrzymane dla różnych wartości współczynnika Kendalla dla starego S i nowego S' portfela polis zawarte są w tab. 1. Z tabeli tej wynika, że wartości podstawowych miar ryzyka wzrastają wraz ze wzrostem współczynnika korelacji Kendalla. Wzrost stopnia zależności między starym portfelem a nowym produktem zwiększa ryzyko wystąpienia większej szkody w nowym portfelu.

Literatura

- [1] Ambagaspitiya R.G., *On the Distribution of a Sum of Correlated Aggregate Claims*, „Insurance: Mathematics and Economics” 1998, 23, 15-19.
- [2] Bowers N.L., Gerber H.U., Hickman J.C., Jones D.A., Nesbitt C.J., *Actuarial Mathematics*, Society of Actuaries, Schaumburg 1986.
- [3] Cossette H., Marceau E., *The Discrete-Time Model with Correlated Classes of Business*, „Insurance: Mathematics and Economics” 2000, 26, 133-149.
- [4] Cossette H., Gaillardetz P., Marceau E., *Common Mixture in the Individual Risk Model*, „Mitteilungen der Schweiz, Aktuarvereinigung” 2002, 2, 131-157.
- [5] Cossette H., Gaillardetz P., Marceau E., Rioux J., *On Two Dependent Individual Risk Models*, „Insurance: Mathematics and Economics” 2002, 30, 153-166.
- [6] Frees E.W., Valdez E.A., *Understanding Relationships Using Copulas*, „North American Actuarial Journal” 1998, 2, 1-25.
- [7] Frey R., McNeil A.J., *Modelling Dependent Defaults*, working paper ETH, Zürich 2001, <http://www.math.ethz.ch/~frey>.
- [8] Genest G., Marceau E., Mesfioui M., *Compound Poisson Approximations for Individual Models*, „Insurance: Mathematics and Economics” 2003, 32, 73-91.
- [9] Gerber H.U., *An Introduction to Mathematical Risk Theory*, Huebner Foundation, Philadelphia 1979.
- [10] Heilpern S., *Analiza zaleznego ryzyka ubezpieczeniowego – zastosowanie funkcji łączących*, Prace Naukowe Akademii Ekonomicznej nr 1108, AE, Wrocław 2006, 269-285.
- [11] Heilpern S., *Funkcje łączące – podstawowe pojęcia i własności*, Prace Naukowe Akademii Ekonomicznej nr 1105, AE, Wrocław 2006, 27-52.
- [12] Klugman S.A., Panjer H.H., Willmot G.E., *Loss Models: From Data to Decisions*, J. Wiley, New York 1998.

- [13] Lindskog F., McNeil A.J., *Common Poisson Shock Models: Applications to Insurance and Credit Modelling*, „ASTIN Bulletin” 2003, 33, 209-238.
- [14] Marshall A.W., Olkin I., *Families of Multivariate Distributions*, „Journal of the American Statistical Association” 1988, 83, 834-841.
- [15] Nelsen R.B., *An Introduction to Copulas*, Springer, New York 1999.
- [16] Ostasiewicz W. (red.), *Modele aktuarialne*, AE, Wrocław 2000.
- [17] Romano C., *Calibrating and Simulating Copula Functions: An Application to the Italian Stock Market*, University of Rome, Working Paper 2002.
- [18] Rolski T., Schmidli H., Schmidt V., Teugels J.L., *Stochastic Processes for Insurance and Finance*, J. Wiley, New York 1999.
- [19] Wang S., *Aggregation of Correlated Risk Portfolios: Models and Algorithms*, „CAS Proceedings” 1998, 848-939.
- [20] Wu F., Valdez E.A., Sherris M., *Simulating Exchangeable Multivariate Archimedean Copulas and Its Applications*, working paper, 2006, <http://www.ermii.org/Research/downloads/SimulatingArchimedeanCopulas.pdf>.

MODELLING OF DEPENDENCE – APPLICATION IN THE RISK MANAGEMENT

Summary

Three methods of modelling of dependence of random variables and examples of the application of these methods to risk management are presented. We study the application process in the context of ruin and the credit solvency.

In the first method, the dependent random variables are represented as the linear combination of the independent random variables. The second method is based on the mixtures and uses the conditional independency random of variables. In the third one we describe the dependency of random variables using copulas.

These methods let us reduce the studied models (the random sum of random variables in our case), to the independent form. So, we can use the classical methods connected with independent case.