

Stanisław HEILPERN*

ZALEŻNY ROZKŁAD DWUMIANOWY I JEGO ZASTOSOWANIE W REASEKURACJI I KREDYTACH

Praca jest poświęcona zależnemu rozkładowi dwumianowemu. W odróżnieniu od klasycznego rozkładu dwumianowego odstąpiono od założenia o niezależności zmiennych losowych. Omówiono poszczególne przypadki uwzględniające różne struktury zależności oraz rozszerzenia modelu. Przedstawiono zastosowania w reasekuracji nadwyżki szkody oraz w zarządzaniu ryzykiem kredytowym.

Słowa kluczowe: *zależny rozkład dwumianowy, funkcja łącząca, reasekuracja, kredyt*

1. Wstęp

W klasycznych modelach aktuarialnych, czy też finansowych, rozpatrywane zmienne losowe są na ogół niezależne. Założenie to jest bardzo wygodne z punktu widzenia matematycznego – łatwiej jest wtedy dużo rzeczy udowodnić – ale jest niestety często znacznie odbiegające od rzeczywistości. W praktycznych zastosowaniach występujące procesy i zjawiska są na ogół zależne. Wpływają na nie wspólne czynniki zewnętrzne.

W pracy przedstawiono uogólnienie prostego, klasycznego modelu dwumianowego. Rozpatrzono przypadek, gdy występujące zmienne losowe wyznaczające ten rozkład mogą być zależne. Omówiono też jego rozszerzenia. Zaprezentowany model zastosowano do zagadnień związanych z reasekuracją i z zarządzaniem ryzykiem kredytowym.

Przedstawimy teraz w skrócie te dwa zagadnienia.

* Katedra Statystyki, Akademia Ekonomiczna we Wrocławiu, ul. Komandorska 118/120, 53-345 Wrocław, e-mail: Stanislaw.Heilpern@ae.wroc.pl

1.1. Reasekuracja nadwyżki szkody [7]

Rozpatrzmy portfel składający się z n szkód X_1, \dots, X_n oraz próg retencji d . Interesować nas będzie liczba szkód pokrytych przez reasekuratora, czyli zmienna losowa

$$K = \sum_{j=0}^n Y_j,$$

gdzie dla każdego $i = 1, \dots, n$, Y_i jest zerojedynkową zmienną losową, przyjmującą wartość 1 gdy reasekurator pokrywa szkodę i 0 w przeciwnym razie. Innymi słowy

$$Y_i = \begin{cases} 0 & X_i \leq d, \\ 1 & X_i > d. \end{cases}$$

W klasycznych modelach aktuarialnych zakładamy zwykle niezależność występujących zmiennych losowych. Jednakże na rozpatrywane ryzyka wpływają na ogół w praktyce wspólne czynniki zewnętrzne: klimatyczne, ekonomiczne czy polityczne. Mogą to być powodzie, pożary, trzęsienia ziemi, tornada, kryzysy ekonomiczne czy polityczne, inflacja, hossy lub wojny. Z tego też powodu przyjmujemy w naszym modelu, że rozpatrywane zmienne losowe X_1, \dots, X_n opisujące wielkości szkód mogą być zależne.

1.2. Ryzyko kredytowe [4, 6, 8]

Będziemy zajmować się portfelem dotyczącym n dłużników w ustalonym okresie. Wprowadźmy w tym celu zerojedynkowe zmienne losowe Y_1, \dots, Y_n , przedstawiające status poszczególnych dłużników:

$$Y_j = \begin{cases} \text{nie spłaca,} \\ \text{spłaca.} \end{cases}$$

Przyjmijmy też, że zmienne te są związane z ciągłymi zmiennymi losowymi X_1, \dots, X_n , nazywanymi krytycznymi [8], relacją

$$Y_j = 1 \Leftrightarrow X_j \leq d_j.$$

Krytyczne zmienne przedstawiają ogólną sytuację finansową poszczególnych dłużników, zwykle ich aktywa, a progi d_i ich zdolności kredytowe. W modelu KMV [6], stosowanym powszechnie w przemyśle, X_i jest interpretowana jako zmiana aktywów i -tego dłużnika w określonym czasie. Zmienne krytyczne X_i są traktowane zwykle jako zmienne losowe ukryte, nieobserwowalne [4, 6, 8].

W przykładzie tym interesuje nas zmienna losowa

$$K = \sum_{j=0}^n Y_j,$$

będąca liczbą niewypłacalnych dłużników. Zmienne losowe X_1, \dots, X_n , a także Y_1, \dots, Y_n mogą być również, i na ogół są, w praktyce zależne.

2. Model

Przedstawimy teraz matematyczny model opisujący zaprezentowane przykłady. Z punktu widzenia matematycznego można je opisać jednym modelem. Różnią się bowiem jedynie zwrotem nierówności w relacji wiążącej zmienne X_i oraz Y_i .

Niech $\mathbf{Y} = (Y_1, \dots, Y_n)$ będą zerojedynkowymi zmiennymi losowymi. Prawdopodobieństwa sukcesów, zdarzeń $Y_j = 1$ oraz porażek, gdy $Y_j = 0$, będziemy oznaczać odpowiednio:

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

Rozkład łączny opisujemy funkcją prawdopodobieństwa

$$f_{\mathbf{Y}}(i_1, \dots, i_n) = \Pr(Y_1 = i_1, \dots, Y_n = i_n),$$

gdzie $i_j \in \{0, 1\}$ oraz dystrybuantą

$$F_{\mathbf{Y}}(i_1, \dots, i_n) = \Pr(Y_1 \leq i_1, \dots, Y_n \leq i_n).$$

W przypadku dystrybuanty będą nas interesować jedynie wartości w punktach skoku, tzn. gdy $i_j \in \{0, 1\}$. Funkcja tworząca prawdopodobieństwa jest w tym przypadku równa

$$\mathbf{P}_{\mathbf{Y}}(t_1, \dots, t_n) = \sum_{i_1, \dots, i_n \in \{0, 1\}} f_{\mathbf{Y}}(i_1, \dots, i_n) t_1^{i_1} \dots t_n^{i_n}.$$

Struktura zależności losowego wektora \mathbf{Y} może być opisana tzw. funkcją łączącą (ang. *copula*) $C_{\mathbf{Y}}$. Jest ona n -wymiarową dystrybuantą skupioną na $[0, 1]^n$ o jednostajnych rozkładach brzegowych, będącą łącznikiem między dystrybuantami brzegowymi F_{Y_i} , a dystrybuantą łączną $F_{\mathbf{Y}}$. Funkcja łącząca spełnia zależność [9]

$$F_{\mathbf{Y}}(i_1, \dots, i_n) = C_{\mathbf{Y}}(F_{Y_1}(i_1), \dots, F_{Y_n}(i_n)).$$

Dystrybuanta brzegowa jest określona wzorem

$$F_{Y_j}(i_j) = \Pr(Y_j \leq i_j) = u_j = \begin{cases} 1 & i_j = 1 \\ q_j & i_j = 0 \end{cases}. \quad (1)$$

Funkcja łącząca C_Y wyznacza nam jednoznacznie wartości dystrybuanty łącznej F_Y w punktach skoku i_j . Poza tymi punktami jednoznaczność nie zachodzi. Jednakże w pracy interesować nas będą jedynie punkty skoku. Aby wyznaczyć funkcję prawdopodobieństwa f_Y , musimy znaleźć relację zachodzącą między f_Y a dystrybuantą F_Y . Zauważmy, że dowolny punkt skoku dystrybuanty F_Y możemy przedstawić jako indyktor pewnego podzbioru $A \subset \{1, \dots, n\}$, czyli $(i_1, \dots, i_n) = \mathbf{1}_A$. Wtedy $i_j = 1$, gdy $i_j \in A$ oraz $i_j = 0$ w przeciwnym wypadku. Załóżmy, że $|A| = zk$, gdzie $|A|$ oznacza jego liczebność, wtedy można pokazać [3], że

$$f_Y(\mathbf{1}_A) = \sum_{j=0}^k (-1)^j \sum_{\substack{D \subset A, \\ |D|=k-j}} F_Y(\mathbf{1}_D). \quad (2)$$

Przykład 1. Rozpatrzmy portfel składający się z czterech polis. Obliczymy prawdopodobieństwo zdarzenia, że reasekuranter pokryje tylko pierwszą i trzecią polisę, znając wartości funkcji łączącej C_Y :

$$\begin{aligned} F_Y(1,0,1,0) &= F_Y(1,0,1,0) - F_Y(1,0,0,0) - F_Y(0,0,1,0) + F_Y(0,0,0,0) \\ &= C_Y(1,q_2,1,q_4) - C_Y(1,q_2,q_3,q_4) - C_Y(q_1,q_2,1,q_4) + C_Y(q_1,q_2,q_3,q_4). \end{aligned}$$

Załóżmy, że pomocnicze zmienne losowe X_1, \dots, X_n są ciągłe. Wtedy funkcja łącząca C_X opisująca ich strukturę zależności jednoznacznie określa łączną dystrybuantę

$$F_X(x_1, \dots, x_n) = C_X(F_{X_1}(x_1), \dots, F_{X_n}(x_n)).$$

Dystrybuanty brzegowe F_{X_i} w przypadku modelu dotyczącego reasekuracji spełniają warunek

$$F_{X_j}(d) = \Pr(X_j \leq d) = \Pr(Y_j = 0) = q_j.$$

Ponadto wartości funkcji łączących C_Y oraz C_X są równe w punktach skoku, czyli

$$C_Y(u_1, \dots, u_n) = C_X(u_1, \dots, u_n),$$

gdzie u_j są określone wzorem (1). Zachodzą bowiem zależności

$$C_Y(u_1, \dots, u_n) = \Pr(Y_1 \leq i_1, \dots, Y_n \leq i_n) = \Pr(X_1 \leq r_1, \dots, X_n \leq r_n),$$

gdzie:

$$r_j = \begin{cases} \infty & i_j = 1 \\ d & i_j = 0 \end{cases}$$

oraz

$$\Pr(X_1 \leq r_1, \dots, X_n \leq r_n) = C_{\mathbf{X}}(u_1, \dots, u_n),$$

ponieważ

$$F_{X_j}(r_j) = \begin{cases} q_j & r_j = d \\ 1 & r_j = \infty \end{cases}.$$

Funkcja tworząca prawdopodobieństwa sumy $K = \sum_{j=0}^n Y_j$ jest równa

$$\mathbf{P}_K(t) = \sum_{i_1, \dots, i_n \in \{0,1\}} f_{\mathbf{Y}}(i_1, \dots, i_n) t^{i_1 + \dots + i_n} = \sum_{k=0}^n \sum_{|A|=k} f_{\mathbf{Y}}(\mathbf{1}_A) t^k,$$

a jej rozkład jest określony wzorem

$$\Pr(K = k) = \sum_{|A|=k} f_{\mathbf{Y}}(\mathbf{1}_A) = \sum_{j=0}^k (-1)^j \binom{n-k+j}{j} \sum_{|A|=k-j} F(\mathbf{1}_A).$$

Wynika on bezpośrednio z kombinatorycznego wzoru (2).

3. Szczególne przypadki

Załóżmy teraz, że zmienne losowe X_1, \dots, X_n mają ten sam rozkład oraz że funkcja łącząca $C_{\mathbf{X}}$ jest wymiennalna (ang. *exchangeable*), tzn. dla każdej permutacji π zbioru $\{1, \dots, n\}$ otrzymujemy

$$C_{\mathbf{X}}(u_1, \dots, u_n) = C_{\mathbf{X}}(u_{\pi(1)}, \dots, u_{\pi(n)}).$$

Wtedy dystrybuanta $F_{\mathbf{Y}}$ przyjmuje te same wartości dla ciągów o tej samej liczbie jedynek. Innymi słowy, zachodzi

$$F_{\mathbf{Y}}(\mathbf{1}_A) = F_{\mathbf{Y}}(\mathbf{1}_B),$$

gdy $|A| = |B| = k$. Wartość tę będziemy oznaczać symbolem $F_{k,n}$. Jest ona równa

$$F_{k,n} = \Pr(Y_{k+1} = 0, \dots, Y_n = 0) = C(\underbrace{1, \dots, 1}_k, \underbrace{q, \dots, q}_{n-k}).$$

Korzystając ze wzoru (2), otrzymujemy natomiast w tym przypadku wartość funkcji prawdopodobieństwa $f_{k,n}$:

$$f_{k,n} = f_{\mathbf{Y}}(\mathbf{1}_A) = \Pr(Y_1 = 1, \dots, Y_k = 1, Y_{k+1} = 0, \dots, Y_n = 0) = \sum_{j=0}^k (-1)^j \binom{k}{j} F_{k-j,n}$$

oraz rozkład i funkcję tworzącą prawdopodobieństwa zmiennej losowej K :

$$\Pr(K = k) = \sum_{j=0}^k (-1)^j \frac{n!}{(n-k)! j! (k-j)!} F_{k-j,n},$$

$$\mathbf{P}_K(t) = \sum_{k=0}^n \binom{n}{k} f_{k,n} t^k.$$

Przykład 2. Niech $n = 4$, wtedy

$$\Pr(K = 2) = 6F_{2,4} + 12F_{1,4} + 6F_{0,4}.$$

Wartość oczekiwana zmiennej losowej K oraz kowariancja zmiennych losowych Y_i, Y_j są odpowiednio równe

$$E(K) = \sum_{j=1}^n E(Y_j) = np,$$

$$\text{Cov}(Y_i, Y_j) = E(Y_i Y_j) - E(Y_i)E(Y_j) = f_{2,2} - p^2 = C_{\mathbf{X}}(q, q) - q^2,$$

wariancja zmiennej losowej K , jak i współczynnik korelacji wynoszą natomiast

$$V(K) = \sum_{j=1}^n V(Y_j) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{Cov}(Y_i, Y_j) = npq + (n^2 - n)(C_{\mathbf{X}}(q, q) - q^2),$$

$$\rho(Y_i, Y_j) = \frac{C_{\mathbf{X}}(q, q) - q^2}{pq}. \quad (3)$$

Rozpatrzmy teraz poszczególne przypadki, zależne od postaci funkcji łączącej $C_{\mathbf{X}}$ opisującej strukturę zależności losowego wektora \mathbf{X} .

3.1. Niezależność

Niezależności zmiennych losowych X_1, \dots, X_n odpowiada funkcja łącząca postaci

$$C(u_1, \dots, u_n) = u_1, \dots, u_n.$$

Będziemy ją dalej oznaczać symbolem Π . Zmienna losowa K ma wtedy klasyczny rozkład dwumianowy. Zachodzą w tym przypadku ogólnie znane wzory:

$$\begin{aligned} F_{k,n} &= q^{n-k}, & f_{k,n} &= p^k q^{n-k}. \\ \Pr(K = k) &= \binom{n}{k} p^k q^{n-k}, & \mathbf{P}_K(t) &= (q + pt)^n. \\ V(K) &= npq, & \rho(Y_i, Y_j) &= 0. \end{aligned}$$

3.2. Współmonotoniczność

Przeciwnością niezależności jest *współmonotoniczność* (ang. *comonotonicity*). Jest to ścisła zgodna zależność, opisana funkcją łączącą

$$C(u_1, \dots, u_n) = \min(u_1, \dots, u_n),$$

którą będziemy oznaczać symbolem M . Wtedy otrzymujemy:

$$\begin{aligned} F_{k,n} &= \begin{cases} q & k < n \\ 1 & n \end{cases}, & f_{k,n} &= \begin{cases} q & k = 0 \\ 0 & 0 < k < n, \\ p & n \end{cases} \\ \Pr(K = k) &= \begin{cases} q & k = 0 \\ 0 & 0 < k < n, \\ p & n \end{cases}, & \mathbf{P}_K(t) &= q + pt^n, \\ V(K) &= n^2 pq, & \rho(Y_i, Y_j) &= 1. \end{aligned}$$

3.3. Mieszanka Π oraz M

Tallis [10] oraz następnie Kolev, Paiva [7] rozpatrywali funkcję łączącą postaci

$$(1 - \rho)\Pi + \rho M,$$

gdzie $0 \leq \rho \leq 1$, będącą kombinacją wypukłą niezależności i współmonotoniczności. Oddaje ona zarówno wpływ czynników indywidualnych, charakterystycznych dla każdej polisy, czy dłużnika, jak i wpływ czynników zewnętrznych oddziałujących na wszystkie jednostki. Funkcja tworząca prawdopodobieństwa sumy K jest wtedy równa

$$\mathbf{P}_K(t) = (1 - \rho)(q + pt)^n + \rho(q + pt^n),$$

a współczynnik korelacji jest równy współczynnikowi kombinacji wypukłej, tzn.

$$\rho(Y_i, Y_j) = \rho.$$

Wariancja zmiennej losowej K jest określona wzorem

$$V(K) = npq(1 + \rho(n - 1)).$$

3.4. Archimedesowa funkcja łącząca

W praktycznych zastosowaniach często wykorzystywane są tzw. *archimedesowe* funkcje łączące. Dzieje się tak zwykle z powodu prostej ich postaci. Funkcje te tworzone są przez generator φ , który jest malejącą, wypukłą funkcją, spełniającą warunek: $\varphi(0) = \infty$, $\varphi(1) = 0$. Archimedesowe funkcje łączące przyjmują quasi-addytywną postać, charakteryzującą się rozdzieleniem zmiennych [5, 9]:

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

Wartość dystrybuanty zależy wtedy od generatora φ oraz prawdopodobieństwa porażki q i jest określona wzorem

$$F_{k,n} = \varphi^{-1}((n - k) \varphi(q)).$$

W zagadnieniach praktycznych wykorzystuje się zwykle parametryzowane rodziny archimedesowych funkcji łączących. Parametr oddaje wtedy stopień zależności. Można też przedstawić wzorem zależność między wartością tego parametru, a współczynnikiem korelacji rang Kendala [5, 9]. Dla $n > 2$ dowolna archimedesowa funkcja łącząca C spełnia następujące nierówności:

$$\Pi(u_1, \dots, u_n) \leq C(u_1, \dots, u_n) \leq M(u_1, \dots, u_n).$$

Wynika z nich oraz ze wzoru (3), że zawsze będzie zachodzić nieujemna zależność między zmiennymi Y_j , tzn. $\rho(Y_i, Y_j) \geq 0$.

Scharakteryzujemy teraz w dużym skrócie trzy najczęściej stosowane rodziny archimedesowych funkcji łączących. Funkcja łącząca *Claytona* jest opisana wzorem

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

gdzie $\alpha > 0$, a generator przyjmuje postać

$$\varphi(u) = u^{-\alpha} - 1.$$

Wartość dystrybuanty dla $n - k$ porażek jest wtedy równa

$$F_{k,n} = (1 + (n - k)(q^{-\alpha} - 1))^{-1/\alpha}.$$

Graniczna wartość parametru $\alpha = 0$ odpowiada niezależności, a dla $\alpha = \infty$ mamy współmonotoniczność.

Drugą popularną rodziną archimedesowych funkcji łączących jest rodzina Gumbela. Elementy tej rodziny są określone wzorem

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

gdzie $1 \leq \alpha$, z generatorem

$$\varphi(u) = (-\ln u)^\alpha.$$

Wartość dystrybuanty wynosi

$$F_{k,n} = q^{(n-k)^{1/\alpha}}.$$

Dla $\alpha = 1$ otrzymujemy niezależność, a dla $\alpha = \infty$ współmonotoniczność.

Funkcja łącząca, przedstawiona wzorem

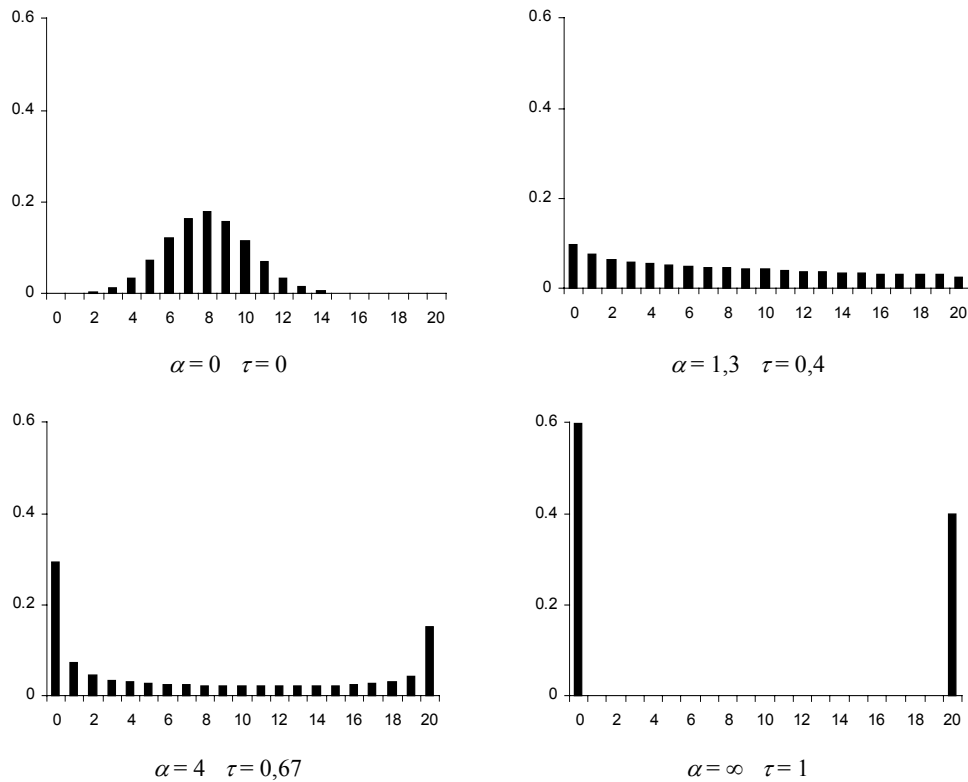
$$C(u_1, \dots, u_n) = -\frac{1}{\alpha} \ln \left(1 + \frac{(e^{-\alpha u_1} - 1) \dots (e^{-\alpha u_n} - 1)}{(e^{-\alpha} - 1)} \right),$$

gdzie $0 \leq \alpha$, należy do archimedesowej rodziny Franka. Jej generator ma postać

$$\varphi(u) = -\ln \frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1}.$$

Podobnie jak dla rodziny Clayтона graniczna wartość parametru $\alpha = 0$ odpowiada niezależności, a jeśli $\alpha = \infty$, to mamy współmonotoniczność. Dystrybuanta jest w tym przypadku określona dość skomplikowanym wzorem.

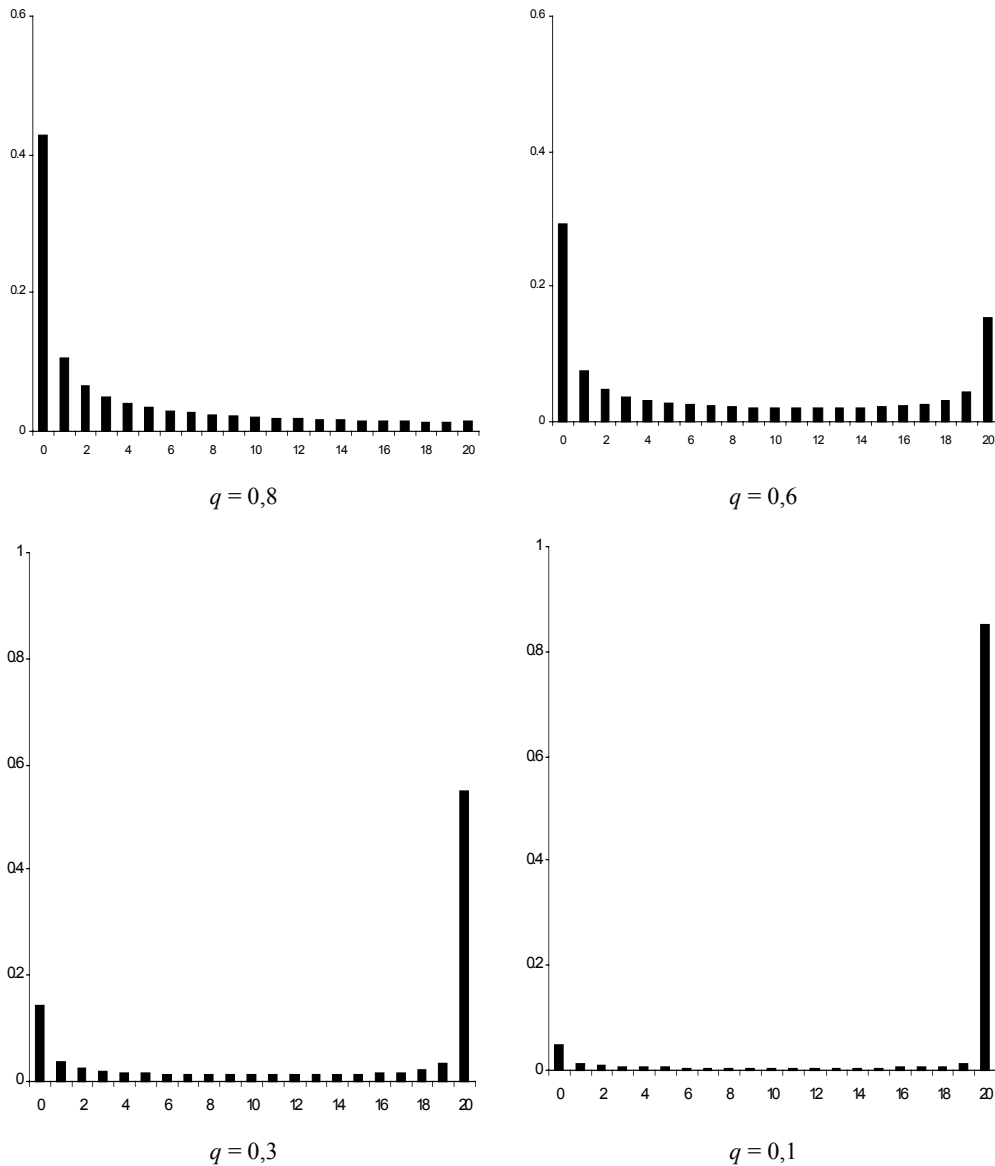
Przykład 3. Rozpatrzmy portfel składający się z $n = 20$ dłużników i przyjmijmy, że prawdopodobieństwo wypłacalności każdego dłużnika jest równe $q = 0,6$ oraz że struktura zależności zmiennych losowych X_1, \dots, X_n jest opisana za pomocą archimedesowej funkcji łączącej Clayтона. Na rysunku 1 przedstawiono rozkłady zmiennej losowej K , przedstawiającej liczbę niewypłacalnych dłużników, dla wartości parametrów rodziny Clayтона α odpowiednio równych 0 (niezależność); 1,3; 4 oraz ∞ (współmonotoniczność). Odpowiadające im wartości współczynnika korelacji rang Kendala τ są równe: 0; 0,4; 0,67 oraz 1.



Rys. 1. Rozkłady zmiennej losowej K (liczba niewypłacalnych dłużników) dla różnych stopni zależności
 Źródło: Opracowanie własne.

Można zauważyć, że w przypadku klasycznym, zakładającym niezależność, rozkład liczby dłużników jest jednomodalny, skupiony wokół oczekiwanej liczby dłużników równej 8. Jednakże w praktyce trudno oczekiwać w tym przypadku całkowitej niezależności. Zwykle działają na dłużników wspólne czynniki zewnętrzne, takie jak zmiany kursów akcji oraz walut lub kryzysy ekonomiczne. Na ogół możemy się spotkać z pewną zależnością sytuacji badanych dłużników. Dla słabych zależności, małych wartości α , wykres rozkładu liczby dłużników staje się bardziej rozciągnięty. Następnie wraz ze wzrostem zależności masa prawdopodobieństwa przesuwa się w lewą stronę. Najbardziej prawdopodobnym staje się brak niewypłacalnych dłużników, a kolejne liczby niewypłacalnych dłużników są coraz mniej prawdopodobne. Gdy współczynnik korelacji Kendalla τ przekracza 0,5, wykres staje się U-kształtny, tzn. najbardziej prawdopodobne stają się skrajne wartości zmiennej K , czyli przypadek gdy nie ma niewypłacalnych dłużników lub wszyscy stają się niewypłacalni. Oczywiście drugie zdarzenie zachodzi z mniejszym prawdopodobieństwem. W przypadku całkowitej zależności, gdy $\tau = 1$, co

odpowiada nieskończonej wartości parametru α , rozkład staje się dwupunktowy. Wszyscy dłużnicy są wypłacalni z prawdopodobieństwem 0,6 lub niewypłacalni z prawdopodobieństwem 0,4. Przypadek całkowitej zależności zachodzi bardzo rzadko, najczęściej należy się spodziewać niezbyt dużej zależności.



Rys. 2. Rozkłady zmiennej losowej K w zależności od prawdopodobieństwa wypłacalności
Źródło: Opracowanie własne.

Na rysunku 2 przedstawiono rozkłady prawdopodobieństwa zmiennej losowej K przy ustalonej wartości parametru $\alpha = 4$, odpowiadającej wartości współczynnika Kendala $\tau = 0,67$, dla różnych prawdopodobieństw wypłacalności q równych odpowiednio: 0,8; 0,6; 0,3 oraz 0,1. Widzimy, że wykresy rozkładów są U-kształtne, a masa prawdopodobieństwa „wędruje” wraz ze spadkiem wartości prawdopodobieństwa q z lewej strony wykresu do prawej. Zwiększa się wtedy prawdopodobieństwo wystąpienia większej liczby niewypłacalnych dłużników, co jest oczywiście zgodne z intuicją.

4. Rozszerzenia

4.1. Wartości

Dotychczas interesowała nas liczba sukcesów, zmienna losowa K , przedstawiająca liczbę polis pokrytych przez reasekuratora lub liczbę niewypłacalnych dłużników. Teraz będziemy się zajmować wartością badanego procesu.

W przypadku reasekuracji będą to wartości szkód pokrytych przez reasekuratora $Z_j = X_j - d$, gdzie $j = 1, \dots, n$, a dla zagadnień dotyczących ryzyka kredytowego, wartość straconego kredytu Z_j . Głównym przedmiotem naszych zainteresowań będzie w tej sytuacji globalna wartość szkód podlegających reasekuracji lub wartość straconych kredytów, czyli zmienna losowa

$$S = \sum_{j=1}^n Z_j .$$

Funkcja tworząca momenty tak określonej zmiennej losowej S jest równa [3]

$$\mathbf{M}_S(t) = \sum_{i_1, \dots, i_n \in \{0,1\}} f_Y(i_1, \dots, i_n) (\mathbf{M}_{Z_1}(t))^{i_1}, \dots, (\mathbf{M}_{Z_n}(t))^{i_n} .$$

Gdy zmienne losowe Z_1, \dots, Z_n mają te same rozkłady, a funkcja łącząca C_Z opisująca ich strukturę zależności, równa oczywiście funkcji łączącej C_X , jest wymienna, wtedy funkcja tworząca momenty zmiennej S ma postać

$$\mathbf{M}_S(t) = \sum_{k=0}^n \binom{n}{k} f_{k,n} (\mathbf{M}_Z(t))^k .$$

Jej dystrybuanta natomiast jest kombinacją wypukłą kolejnych splotów dystrybuanty F_Z zmiennej Z_j :

$$F_S(x) = \sum_{k=0}^n \binom{n}{k} f_{k,n} F_Z^{*k}(x).$$

Jeśli zmienne te są współmonotoniczne, dystrybuanta sumy S zależy jedynie od n -tego splotu:

$$F_S(x) = q + pF_Z^{*n}(x).$$

4.2. Losowa liczba szkód

W zagadnieniach aktuarialnych [1] liczba szkód jest na ogół traktowana jako zmienna losowa N . Wtedy liczba sukcesów, liczba polis pokrytych przez reasekuratora jest losową sumą

$$K = \sum_{j=1}^N Y_j.$$

W przypadku niezależnych zmiennych losowych Y_j , funkcja tworząca prawdopodobieństwa sumy K przyjmuje znaną postać

$$\mathbf{P}_K(t) = \mathbf{P}_M(q + pt).$$

Dla współmonotonicznych zmiennych jest ona równa

$$\mathbf{P}_K(t) = q + p\mathbf{P}_M(t)$$

i możemy w tym przypadku określić rozkład zmiennej losowej K :

$$\Pr(K = k) = \begin{cases} q + p \Pr(N = 0) & k = 0 \\ p \Pr(N = k) & k > 0 \end{cases}.$$

Jeśli funkcja łącząca C_Y jest kombinacją wypukłą $(1 - \rho)II + \rho M$, to funkcja tworząca prawdopodobieństwa liczby sukcesów K jest równa

$$\mathbf{P}_K(t) = (1 - \rho)\mathbf{P}_M(q + pt) + \rho(q + p\mathbf{P}_M(t)).$$

Ciekawsza sytuacja występuje w przypadku archimedesowej funkcji łączącej C_Y . Możemy wtedy skorzystać z tzw. *modelu słabości* (ang. *frailty*) [11]. Istnieje ukryta zmienna Θ , reprezentująca czynnik zewnętrzny oddziałujący jednocześnie na wszystkie składniki, związana z generatorem φ archimedesowej funkcji łączącej relacją

$$\varphi^{-1}(t) = \mathbf{M}_\Theta(-t).$$

Jeśli funkcja łącząca C_Y należy do rodziny Claytona, to ukryta zmienna losowa Θ ma rozkład gamma. Dla funkcji łączącej Gumbela otrzymujemy rozkład stabilny, a dla rodziny Franka rozkład logarytmiczny [11].

Istnieje też w tym przypadku tzw. dystrybuanta bazowa $H_i(y)$, taka, że przy ustalonej wartości θ ukrytej zmiennej losowej Θ , wartość warunkowej dystrybuanty brzegowej jest funkcją potęgową dystrybuanty bazowej, tzn.

$$F_j(y | \Theta = \theta) = (H_j(y))^\theta.$$

Dystrybuanta bazowa jest określona wzorem [2, 5]

$$H_j(y) = e^{-\varphi(F_j(y))},$$

ponadto, co jest w tym przypadku najważniejsze, dla ustalonej θ otrzymujemy niezależność:

$$F(y_1, \dots, y_n | \Theta = \theta) = \prod_{j=1}^n F_j(y_j | \Theta = \theta).$$

Bezwarunkowa łączna dystrybuanta zmiennych Y_1, \dots, Y_n przyjmuje wtedy postać następującej mieszanki:

$$F(y_1, \dots, y_n) = \int_0^\infty F(y_1, \dots, y_n | \Theta = \theta) dF_\Theta(\theta).$$

Zmienne losowe Y_1, \dots, Y_n są zerojedynkowe, podobnie jak rozkład wyznaczony przez dystrybuanty bazowe. Warunkowe rozkłady brzegowe tych zmiennych są określone wzorami:

$$\Pr(Y_j = 0 | \Theta = \theta) = r^\theta,$$

gdzie $j = 1, \dots, n$, a prawdopodobieństwo porażki dla bazowego rozkładu jest równe

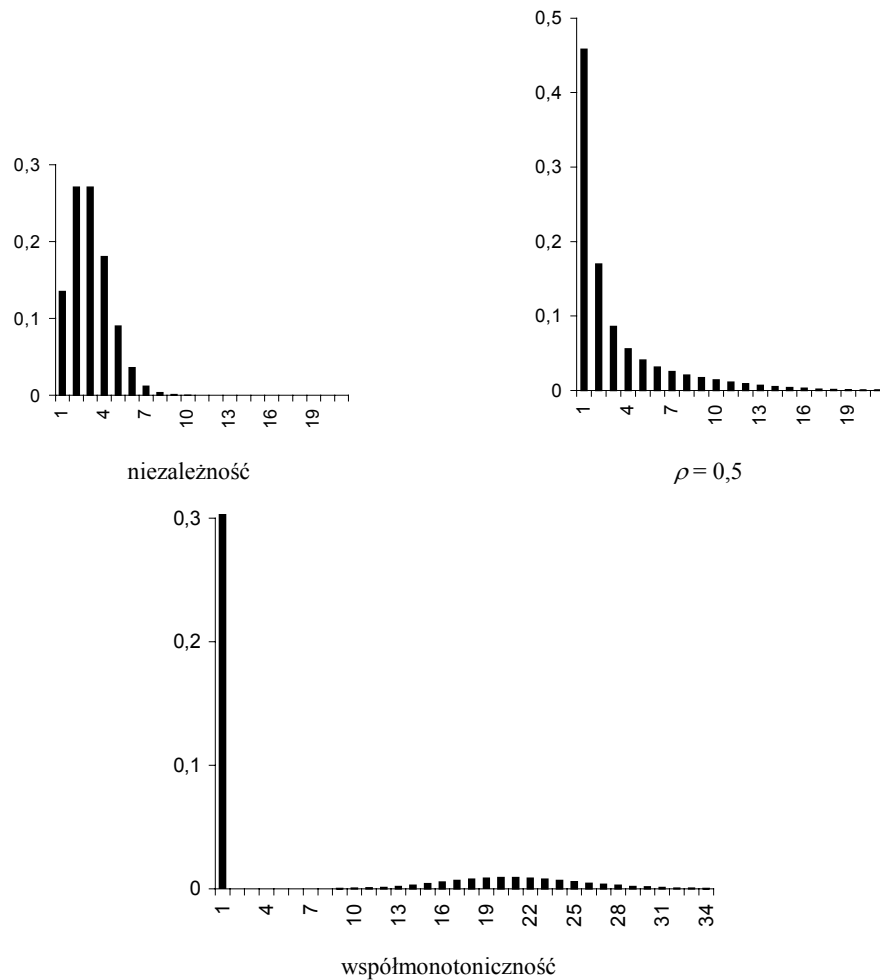
$$r = e^{\mathbf{M}_\Theta^{-1}(q)} = e^{-\varphi(q)}.$$

Warunkowa i bezwarunkowa funkcja tworząca prawdopodobieństwa zmiennej losowej K określone są wzorami:

$$\begin{aligned} \mathbf{P}_{K|\theta}(t) &= \mathbf{P}_N(r^\theta + (1 - r^\theta)t), \\ \mathbf{P}_K(t) &= \int_0^\infty \mathbf{P}_N(r^\theta + (1 - r^\theta)t) dF_\Theta(\theta). \end{aligned}$$

Powyższe rozważania umożliwiają wyznaczenie rozkładu liczby polis pokrytych przez reasekuratora, w przypadku losowej liczby szkód.

Przykład 4. Załóżmy, że liczba szkód jest zmienną losową o rozkładzie Poissona z parametrem $\lambda = 20$, prawdopodobieństwo pokrycia szkody przez reasekuranta wynosi $p = 0,1$, a struktura zależności jest opisana archimedesową rodziną Franka. Na rysunku 3 przedstawiono rozkłady prawdopodobieństwa zmiennej losowej K , liczby szkód pokrytych przez reasekuratora. Rozpatrzono trzy przypadki: niezależności, $\rho = 0,5$ oraz współmonotoniczności.



Rys. 3. Przypadek losowej liczby szkód
Źródło: Opracowanie własne.

W przypadku klasycznym, niezależnych szkód, otrzymujemy jednomodalny rozkład, skupiony wokół oczekiwanej liczby szkód pokrytych przez asekuratora. Oczekiwana liczba tego rodzaju szkód jest równa 2. Gdy współczynnik korelacji Kendalla

$\tau=0,5$, najbardziej prawdopodobny jest brak szkód pokrytych przez asekuratora. Prawdopodobieństwo tego zdarzenia jest bliskie 0,5, a kolejne liczby pokrytych szkód zachodzą z coraz to mniejszym prawdopodobieństwem. W ostatnim, dość skrajnym i rzadko spotykanym przypadku, dopuszczającym ścisłą zależność rozpatrywanych szkód, prawdopodobieństwo braku pokrytych szkód jest natomiast nieco większe niż 0,9. Pozostała, niewielka masa prawdopodobieństwa jest skupiona wokół przeciętnej liczby szkód, wynoszącej 20.

Podsumowanie

W pracy omówiono zależny rozkład dwumianowy. Jest to uogólnienie klasycznego rozkładu, w którym dopuszcza się zależność tworzących go zerojedynkowych zmiennych losowych. Pokazano podstawowe własności zależnego rozkładu dwumianowego, jego szczególne przypadki uwzględniające różne struktury zależności oraz jego uogólnienia.

Przedstawiony w pracy rozkład został zastosowany w zagadnieniach związanych z reasekuracją oraz zarządzaniem ryzykiem kredytowym. W opisanych modelach, reasekuracyjnym i kredytowym, klasyczne założenie o niezależności zostało zastąpione bardziej realistycznym założeniem o zależności występujących zmiennych losowych. Modele te są proste, analizowane problemy zostały w nich jedynie zasygnalizowane, łączy je wspólny model matematyczny oparty na zależnym rozkładzie dwumianowym. W zagadnieniu reasekuracyjnym rozszerzono model zaproponowany przez Koleva i Paivę [7], rozpatrując struktury zależności oparte na funkcjach łączących, głównie archimedesowych. W modelu dotyczącym ryzyka kredytowego poddano analizie zmienną losową K , będącą liczbą niewypłacalnych dłużników. Zbadano jej rozkład zarówno zależności od różnych stopni zależności, jak i od prawdopodobieństwa wypłacalności dłużnika. Otrzymane rozkłady w istotny sposób różnią się od klasycznego, zakładającego niezależność rozkładu.

Bibliografia

- [1] BOWERS N., GERBER H.U., HICKMAN J.C., JONES D.A., NESBITT C.J., *Actuarial Mathematics*, Society of Actuaries, Schaumburg 1997.
- [2] COSSETTE H., GAILLARDETZ P., MARCEAU E., *Common mixture in the individual risk model*, Mitteilungen der Schweiz, Aktuarvereinigung, 2002, Vol. 2, s. 131–157.
- [3] COSSETTE H., GAILLARDETZ P., MARCEAU E., RIOUX J., *On two dependent individual risk models*, Insurance: Mathematics and Economics, 2002, Vol. 30, s. 153–166.

- [4] FREY R., MCNEIL A.J., *Modelling dependent defaults*, ETH Zurich 2001, <http://www.math.ethz.ch/~frey>.
- [5] HEILPERN S., *Funkcje łączące – podstawowe pojęcia i własności*, Prace Naukowe AE Wrocław, 2006, nr 1105, s. 27–52.
- [6] KMV-Corporation, *Modelling Default Risk*, Technical Document 1997, <http://www.kmv.com>.
- [7] KOLEV N., PAIVA D., *Multinomial model for random sums*, Insurance: Mathematics and Economics, 2005, Vol. 37, s. 494–504.
- [8] MCNEIL A.J., FREY R., EMBRECHTS P., *Quantitative Risk Management*, Princeton University Press, Princeton 2005.
- [9] NELSEN R.B., *An Introduction to copulas*, Springer, New York 1999.
- [10] TALLIS G.M., *The use of generalized multinomial distribution in the estimation of correlation in discrete data*, J. R. Stat. Soc., Ser. B, 1962, Vol. 24, s. 530–534.
- [11] WANG S., *Aggregation of correlated risk portfolios: Models and algorithms*, CAS Proceedings, 1998, s. 848–939.

Dependent binomial distribution and its application in reinsurance and credits

The paper is devoted to the dependent binomial distribution. The assumption of independence of the random variables in the classical binomial distribution is omitted, so we obtain a more realistic situation. The definition and basic properties of such distribution are presented. The dependent structure of the random variables is characterized by the copula. The cases, which are dependent on the different copulas: exchangeable, independent, comonotonicity, the mixture of such copulas, and Archimedean, are studied. The two extensions of our model, i.e., the values of process and the random number of variables, are investigated, too. The applications of the dependent binomial distribution to the excess-of-loss reinsurance and the credit risk management are presented.

Keywords: *dependent binomial distribution, copula, reinsurance, credit*