Uniwersytet Ekonomiczny w Katowicach
Studia Ekonomiczne. Zeszyty Naukowe
Uniwersytetu Ekonomicznego w Katowicach
ISSN 2083-8611 Nr 246 · 2015

Współczesne Finanse 3

Tomasz Kulpa

Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie
Wydział Matematyczno-Przyrodniczy. Szkoła Nauk Ścisłych
tomasz.kulpa@uksw.edu.pl

MIARY ZALEŻNOŚCI OPARTE NA KOPULACH

Streszczenie: Celem artykułu jest przedstawienie miar zależności opartych na kopulach. Omówione zostały następujące miary: tau Kendalla, rho Spearmana, sigma Schweizera i Wolffa oraz gamma Giniego. Przedstawiono też ogólne aksjomaty miar zgodności i zależności.

Słowa kluczowe: kopula, miara zależności, miara zgodności.

Wprowadzenie

Funkcją łącznikową, inaczej kopulą, nazywamy dowolną funkcję C: $[0, 1]^2 \to [0, 1]$ spełniającą następujące warunki:

$$C(x, 0) = C(0, y) = 0, \; C(x, 1) = x, \; C(1, y) = y \tag{1}$$

dla $x, y \in [0, 1]$ oraz:

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0 \tag{2}$$

dla wszystkich $x_1, y_1, x_2, y_2 \in [0, 1]$ takich, że $x_1 \leq x_2$ i $y_1 \leq y_2$.

Pojęcie kopuli (ang. copula) pierwszy raz pojawiło się w pracy Sklara [1959, s. 229-231]. Wykazał on, że dla każdej pary zmiennych losowych X_1, X_2 określonych na tej samej przestrzeni probabilistycznej, o dystrybuantach F_1 i F_2 odpowiednio, istnieje funkcja łącznikowa C taka, że:

$$F_{12}(x_1, x_2) = C(F_1(x_1), F_2(x_2)), \tag{3}$$

gdzie F_{12} oznacza dystrybuantę wektora losowego (X_1,X_2). Funkcja C jest określona jednoznacznie na zbiorze $R_1 \times R_2$, gdzie R_1 i R_2 oznaczają zbiór wartości dystrybuant F_1 i F_2 odpowiednio. W szczególności dla zmiennych losowych o rozkładzie ciągłym funkcja C w powyższym twierdzeniu jest wyznaczona jednoznacznie. Z drugiej strony dla danej funkcji łącznikowej C istnieją zmienne losowe o wspólnym rozkładzie zadanym funkcją C.

Do podstawowych kopuli zaliczamy funkcję:

$$\Pi(x, y) = xy,$$

która odpowiada parze niezależnych zmiennych losowych X i Y, oraz funkcje:

$$M(x, y) = \min\{x, y\},$$

$$W(x, y) = \max\{x + y - 1, 0\},$$

które odpowiadają parze zmiennych losowych X i Y takich, że Y jest prawie na pewno deterministyczną funkcją X, silnie rosnącą w przypadku kopuli M i silnie malejącą w przypadku kopuli W.

Kopule W i M są też nazywane granicami Frecheta-Hoeffdinga, ponieważ dla dowolnej kopuli C i wszystkich $(u, v) \in \mathrm{I}^2$ zachodzi nierówność:

$$W(u, v) \leq C(u, v) \leq M(u, v), \tag{4}$$

gdzie $\mathrm{I}^2 = [0, 1]^2$ oznacza kwadrat jednostkowy.

Badanie stopnia zależności zmiennych losowych to od dawna domena statystyki. Wiele miar zależności zmiennych losowych jest niezmienniczych ze względu na skale. Analogiczną własność posiadają kopule, mianowicie jeżeli α i β są ściśle rosnącymi funkcjami, a X i Y zmiennymi losowymi, to kopula X i Y jest taka sama, jak kopula $\alpha(X)$ i $\beta(Y)$, tzn.:

$$C_{\alpha(X),\, \beta(Y)} = C_{X,Y}\,.$$

Oznacza to, że te miary zależności, które są niezmiennicze ze względu na ściśle rosnące transformacje zmiennych losowych, mogą być wyrażone za pomocą kopul.

Miary zależności i zgodności oparte na kopulach znalazły zastosowanie między innymi przy konstrukcji narzędzi do zarządzania ryzykiem w finansach i ubezpieczeniach [Denuit i in., 2005; Cherubini, Luciano i Vecchiato, 2004; Cherubini, Della Lunga, 2007].

Celem opracowania jest przedstawienie wybranych miar zależności zdefiniowanych za pomocą kopul.

1. Miara tau Kendalla

Załóżmy, że (x_i, y_i) oraz (x_j, y_j) są dwiema niezależnymi obserwacjami wektora losowego (X,Y) o ciągłych rozkładach brzegowych. Obserwacje (x_i, y_i) oraz (x_j, y_j) są **zgodne**, jeżeli $x_i < x_j$ i $y_i < y_j$ lub $x_i > x_j$ i $y_i > y_j$. Podobnie obserwacje (x_i, y_i) oraz (x_j, y_j) są **niezgodne**, jeżeli $x_i < x_j$ i $y_i > y_j$ lub $x_i > x_j$ i $y_i < y_j$.

Można zatem powiedzieć, że obserwacje wektora (X,Y) są zgodne, gdy obserwowanemu wzrostowi X towarzyszy wzrost Y, natomiast niezgodne w przeciwnym wypadku. Powyższą definicję można też sformułować w następującej postaci:

Definicja 1. *Dwie niezależne obserwacje* (x_i,y_i) *i* (x_j,y_j) *wektora losowego* (X, Y) *o ciągłych rozkładach brzegowych są* **zgodne**, *gdy* $(x_i - x_j)(y_i - y_j) > 0$, *oraz* **niezgodne**, *gdy* $(x_i - x_j)(y_i - y_j) < 0$.

Pojęcie zgodnych i niezgodnych obserwacji pozwala definiować statystyczne miary zgodności. Jedną z nich jest liczba tau Kendalla. Niech $(x_1, y_1), (x_2, y_2),...,(x_n, y_n)$ będzie losową próbą z rozkładu (X, Y). Daje to nam $n(n-1)/2$ różnych par obserwacji wektora (X, Y). Oznaczmy przez c ilość par obserwacji zgodnych, a przez d ilość par obserwacji niezgodnych. Liczbę tau definiujemy wzorem:

$$\tau = \frac{c-d}{c+d} = \frac{2(c-d)}{n(n-1)}. \tag{5}$$

Probabilistycznym odpowiednikiem liczby τ jest miara τ zdefiniowana w poniższy sposób:

Definicja 2. *Załóżmy, że X i Y są zmiennymi losowymi o rozkładzie ciągłym. Niech* (X_1, Y_1) *oraz* (X_2, Y_2) *będą niezależnymi wektorami losowymi o tym samym rozkładzie, co* (X, Y). *Miarę zgodności tau zmiennych losowych X i Y definiujemy jako różnice prawdopodobieństw zgodności i niezgodności wektorów losowych* (X_1, Y_1) *i* (X_2, Y_2), *tzn.* $\tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$. (6)

Pokażemy teraz, że miarę tau Kendalla można zdefiniować w języku kopul. W tym celu wprowadzimy najpierw pojęcie funkcji zgodności Q zdefiniowaną jako różnice prawdopodobieństw zgodności i niezgodności wektorów losowych (X_1, Y_1) i (X_2, Y_2) o tych samych, ciągłych rozkładach brzegowych z dystrybuantami F i G, ale niekoniecznie identycznymi rozkładami łącznymi, z dwuwymiarowymi dystrybuantami H_1 i H_2 odpowiednio.

Definicja 3. *Załóżmy, że* (X_1, Y_1) *i* (X_2, Y_2) *są niezależnymi wektorami losowymi z kopulami* C_1 *i* C_2 *odpowiednio. Załóżmy ponadto, że zmienne losowe* X_1 *i* X_2

mają rozkłady ciągłe z tą samą dystrybuantą F, a Y_1 i Y_2 mają rozkłady ciągłe z tą samą dystrybuantą G. Funkcje zgodności Q definiujemy wzorem:

$$Q = Q(C_1, C_2) = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \quad (7)$$

Uwaga. Przy oznaczeniach z powyższej definicji, jeśli oznaczymy przez H_i dystrybuantę rozkładu łącznego wektora losowego (X_i, Y_i), $i = 1, 2$, to:

$$H_i(x, y) = C_i(F(x), G(y)),\ i = 1, 2$$

dla wszystkich $(x, y) \epsilon \mathrm{R}^2$.

W pierwszej kolejności pokażemy, że wartość funkcji zgodności Q faktycznie zależy od kopul C_1 i C_2, a nie zależy od rozkładów brzegowych F i G. Zachodzi następujące twierdzenie:

Twierdzenie 4. *Przy założeniach definicji 3 zachodzi następujący wzór:*

$$Q(C_1, C_2) = 4 \int\int_{\mathbb{I}^2} C_2(u, v) dC_1(u, v) - 1, \quad (8)$$

gdzie $\mathrm{I}^2 = [0, 1]^2$ *oznacza kwadrat jednostkowy.*

Dowód: Ponieważ rozważane zmienne losowe są ciągłe, zatem:

$$P[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1 - P[(X_1 - X_2)(Y_1 - Y_2) > 0],$$

a stąd :

$$Q = 2P[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1. \quad (9)$$

Zauważmy, że :

$P[(X_1 - X_2)(Y_1 - Y_2) > 0] = P[X_1 > X_2,\ Y_1 > Y_2] + P[X_1 < X_2,\ Y_1 < Y_2]$,
oraz $P[X_1 > X_2,\ Y_1 > Y_2] = P[X_2 < X_1,\ Y_2 < Y_1] =$

$$= \int\int_{\mathbb{R}^2} P[X_2 \leqslant x,\ Y_2 \leqslant y]\, dH_1(x, y) =$$

$$= \int\int_{\mathbb{R}^2} H_2(x, y)\, dH_1(x, y) =$$

$$= \int\int_{\mathbb{R}^2} C_2(F(x), G(y))\, dC_1(F(x), G(y)),$$

co po podstawieniu $u = F(x)$ i $v = G(y)$ daje wzór:

$$P[(X_1 - X_2)(Y_1 - Y_2) > 0] = \int\int_{\mathbb{I}^2} C_2(u, v)\, dC_1(u, v).$$

Analogicznie:

$$P[X_1 < X_2, Y_1 < Y_2] = \iint_{\mathbb{R}^2} P[X_2 > x, Y_2 > y]\, dC_1(F(x), G(y)) =$$

$$= \iint_{\mathbb{R}^2} (1 - F(x) - G(y) + H_2(x, y))\, dH_1(x, y) =$$

$$= \iint_{\mathbb{R}^2} (1 - F(x) - G(y) + C_2(F(x), G(y)))\, dC_1(F(x), G(y)) =$$

$$= \iint_{\mathbb{I}^2} (1 - u - v + C_2(u, v))\, dC_1(u, v) =$$

$$= 1 - 0,5 - 0,5 + \iint_{\mathbb{I}^2} C_2(u, v)\, dC_1(u, v),$$

co łącznie z poprzednimi wzorami daje tezę twierdzenia.

Zauważmy, że bezpośrednio z definicji 3 wynika poniższy wniosek:

Wniosek 5. *Funkcja zgodności Q jest symetryczna, tzn.* $Q(C_1, C_2) = Q(C_2, C_1)$. (10)

Pokażemy teraz, że Q jest rosnąca. Zaczniemy od następującej definicji:

Definicja 6. *Kopula C_1 nie przekracza kopuli C_2, tzn. $C_1 < C_2$ wtedy i tylko wtedy, gdy:*

$$C_1(u, v) \leq C_2(u, v)$$

dla wszystkich $(u, v) \in \mathrm{I}^2$.

Możemy teraz sformułować wniosek:

Wniosek 7. *Funkcja zgodności Q jest rosnąca, tzn. jeżeli $C_1 < C'_1$ oraz $C_2 < C'_2$, to:*

$$Q(C_1, C_2) \leq Q(C'_1, C'_2).$$

Dowód: Zauważmy, że z symetrii Q (wniosek 5) oraz wzoru (8) wynika, że:

$$\iint_{\mathbb{I}^2} C_2(u, v)\, dC_1(u, v) = \iint_{\mathbb{I}^2} C_1(u, v)\, dC_2(u, v). \tag{11}$$

Korzystając z powyższego faktu oraz monotoniczności całki, otrzymujemy:

$$Q(C_1, C_2) = 4 \iint_{\mathbb{I}^2} C_2\, dC_1 - 1 \leqslant 4 \iint_{\mathbb{I}^2} C'_2\, dC_1 - 1 =$$

$$= 4 \iint_{\mathbb{I}^2} C_1\, dC'_2 - 1 \leqslant 4 \iint_{\mathbb{I}^2} C'_1\, dC'_2 - 1 = Q(C'_1, C'_2).$$

Przykład 1. Wyznaczymy wartości funkcji zgodności Q dla kopul M, W i Π. Skorzystamy z faktu, że miary generowane przez M i W są rozłożone jednostajnie na odpowiednich przekątnych kwadratu jednostkowego I^2, a miara generowana przez Π ma rozkład jednostajny na całym kwadracie jednostkowym I^2. Mamy:

$$Q(M,M) = 4\int\int_{\mathbb{I}^2} M(u,v)\,dM(u,v) - 1 = 4\int_0^1 u\,du - 1 = 1,$$

$$Q(M,\Pi) = 4\int\int_{\mathbb{I}^2} \Pi(u,v)\,dM(u,v) - 1 = 4\int_0^1 u^2\,du - 1 = \frac{1}{3},$$

$$Q(M,W) = 4\int\int_{\mathbb{I}^2} W(u,v)\,dM(u,v) - 1 = 4\int_{0,5}^1 (2u-1)\,du - 1 = 0,$$

$$Q(W,W) = 4\int\int_{\mathbb{I}^2} W(u,v)\,dW(u,v) - 1 = 4\int_0^1 0\,du - 1 = -1,$$

$$Q(W,\Pi) = 4\int\int_{\mathbb{I}^2} \Pi(u,v)\,dW(u,v) - 1 = 4\int_0^1 u(1-u)\,du - 1 = -\frac{1}{3},$$

$$Q(\Pi,\Pi) = 4\int\int_{\mathbb{I}^2} \Pi(u,v)\,d\Pi(u,v) - 1 = 4\int\int_{\mathbb{I}^2} uv\,du\,dv - 1 = 0.$$

Wracając do miary zgodności tau, z przedstawionych rozważań wynika następujące twierdzenie:

Twierdzenie 8. *Załóżmy, że X i Y są ciągłymi zmiennymi losowymi z rozkładem łącznym zadanym kopula C. Wówczas ich miara zgodności tau wyraża się wzorem:*

$$\tau_{X,Y} = \tau_C = Q(C,C) = 4\int\int_{\mathbb{I}^2} C(u,v)\,dC(u,v) - 1. \qquad (12)$$

Z poprzedniego przykładu wynika, że:

$$\tau_M = 1, \quad \tau_\Pi = 0, \qquad \tau_W = -1.$$

W szczególności dla pary niezależnych zmiennych losowych X i Y miara tau wynosi 0. W przypadku gdy Y jest silnie rosnącą, deterministyczną funkcją X, miara tau przyjmuje wartość 1, a w sytuacji gdy Y jest silnie malejącą, deterministyczną funkcją X, miara tau przyjmuje wartość -1.

Kopule W i M są odpowiednio dolną i górną granicą Frecheta-Hoefdinga (4), zatem dla dowolnej kopuli C zachodzą relacje $W < C < M$. Z monotoniczności funkcji zgodności Q otrzymujemy:

$$\tau_W \le \tau_C \le \tau_M,$$

co daje następujący wniosek:

Wniosek 9. *Dla dowolnej kopuli C zachodzi:*

$$-1 \leq \tau_C \leq 1.$$

Ważną klasę kopul stanowią kopule archimedesowe. Są one między innymi wykorzystywane do analizy zależnego ryzyka ubezpieczeniowego [Kulpa, 2011]. Wprowadźmy najpierw pojęcie funkcji całkowicie monotonicznej.

Definicja 10. *Funkcje $\varphi(t)$ nazywamy* **całkowicie monotoniczną** *na przedziale $J \subset \mathrm{R}$, jeżeli jest ona ciągła i nieskończenie wiele razy różniczkowalna oraz jej kolejne pochodne są przeciwnych znaków, tzn. spełniają warunek:*

$$(-1)^k \frac{d^k}{dt^k}\varphi(t) \geqslant 0$$

dla wszystkich $t \in int\, J$, $k = 0, 1, 2, \ldots$.

Możemy teraz zdefiniować kopulę archimedesową.

Definicja 11. *Niech φ: $(0, 1] \to [0,\infty)$ będzie funkcją całkowicie monotoniczną spełniającą warunek $\varphi(1) = 0$ i niech funkcja C_φ $[0, 1]^n \to [0, 1]$ będzie zadana wzorem*:

$$C_\varphi (u_1, u_2, \ldots , u_n) = \varphi^{-1}(\varphi (u_1) + \varphi (u_2) + \ldots + \varphi (u_n)).$$

Funkcje C_φ nazywamy **kopulą archimedesową** *z generatorem φ.*

W przypadku kopul archimedesowych miarę zależności tau można wyznaczyć korzystając z poniższego twierdzenia [Nelsen, 1999].

Twierdzenie 12. *Załóżmy, że X i Y są ciągłymi zmiennymi losowymi z rozkładem wspólnym zadanym kopulą archimedesową C z generatorem φ. Wówczas miara zależności tau dla tych zmiennych losowych zadana jest wzorem:*

$$\tau_C = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt. \tag{13}$$

Korzystając z powyższego twierdzenia, wyznaczymy wartość miary tau dla wybranych rodzin kopul archimedesowych.

Przykład 2. (Rodzina Claytona) Niech $\varphi_\theta (t) = t^{-\theta} - 1$ dla pewnego $\theta > 0$. Wówczas $\varphi'_\theta (t) = -\theta\, t^{-(\theta+1)}$ i

$$\tau_\theta = 1 + 4\int_0^1 \frac{t^{\theta+1} - t}{\theta}\, dt = \frac{\theta}{\theta + 2}.$$

Przykład 3. (Rodzina Gumbela) Niech $\varphi_\theta(t) = (-\ln t)^\theta$ dla pewnego $\theta \geq 1$. Wówczas:

$$\varphi'_\theta(t) = \theta\frac{(-\ln t)^\theta}{t\ln t} \text{ i}$$

$$\tau_\theta = 1 + 4\int_0^1 \frac{t\ln t}{\theta}\,dt = \frac{\theta-1}{\theta}.$$

2. Miara rho Spearmana

Innym przykładem miary zgodności bazującej na pojęciu prawdopodobieństwa zgodności i niezgodności obserwacji jest miara rho Spearmana. Zaczniemy od przedstawienia definicji tej miary:

Definicja 13. *Załóżmy, że X i Y są zmiennymi losowymi o rozkładzie ciągłym. Niech (X_1, Y_1), (X_2, Y_2) oraz (X_3, Y_3) będą niezależnymi wektorami losowymi o tym samym rozkładzie, co (X, Y). Miarę zgodności rho zmiennych losowych X i Y definiujemy jako różnicę prawdopodobieństw zgodności i niezgodności wektorów losowych (X_1, Y_1) i (X_2, Y_3), tzn.:*

$$\rho_{X,Y} = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]). \tag{14}$$

W powyższej definicji, ze względu na symetrię, możemy zastąpić wektor (X_2, Y_3) przez wektor (X_3, Y_2), uzyskując równoważną definicję:

$$\rho_{X,Y} = 3(P[(X_1 - X_3)(Y_1 - Y_2) > 0] - P[(X_1 - X_3)(Y_1 - Y_2) < 0]).$$

Różnica w nawiasie po prawej stronie w równaniu (14) to różnica prawdopodobieństw zgodności i niezgodności dla wektorów losowych (X_1, Y_1) i (X_2, Y_3). Zgodnie z definicją zmienne losowe X_2 i Y_3 są niezależne, zatem kopula wektora (X_2, Y_3) to Π. Jeżeli kopule wektora (X, Y) (a zatem i (X_1, Y_1)) oznaczymy przez C, to otrzymujemy natychmiastowy wniosek:

Wniosek 14. *Załóżmy, że X i Y są ciągłymi zmiennymi losowymi z kopulą C. Wówczas miara rho Spearmana wyraża się wzorem:*

$$\rho_{X,Y} = \rho_C = 3Q(C, \Pi). \tag{15}$$

Biorąc pod uwagę wzór (8) z twierdzenia 4, otrzymujemy kolejny wniosek:

Wniosek 15. *Załóżmy, że X i Y są ciągłymi zmiennymi losowymi z kopulą C. Wówczas miara rho Spearmana wyraża się wzorem:*

$$\rho_{X,Y} = \rho_C = 12\iint_{\mathbb{I}^2} C(u,v)\,du\,dv - 3, \tag{16}$$

oraz:

$$\rho_{X,Y} = \rho_C = 12 \int\int_{\mathbb{I}^2} uv\, dC(u,v) - 3. \tag{17}$$

Przykład 4. Na podstawie obliczeń wykonanych w przykładzie 1 oraz wzoru (15) z wniosku 14 otrzymujemy:

$$\rho_M = 1, \quad \rho_\Pi = 0, \quad \rho_W = -1.$$

Korzystając z wzoru (15), monotoniczności funkcji zgodności Q oraz faktu, że $W < C < M$ dla dowolnej kopuli C, otrzymujemy nierówność:

$$\rho_W \leq \rho_C \leq \rho_M,$$

co można zapisać w postaci poniższego wniosku:

Wniosek 16. *Dla dowolnej kopuli C zachodzi oszacowanie:*

$$-1 \leq \rho_C \leq 1. \tag{18}$$

Przykład 5. (Rodzina Farlie-Gumbela-Morgensterna) Niech:

$$C_\theta(u, v) = uv + \theta uv\,(1 - u)((1 - v)$$

dla pewnego $\theta \epsilon [-1, 1]$. Korzystając z wzoru (16), mamy:

$$\rho_\theta = 12 \int\int_{\mathbb{I}^2} (uv + \theta uv(1-u)(1-v))\, du\, dv - 3 = \frac{\theta}{3}.$$

3. Miara Giniego

Innym przykładem miary zależności jest miara Giniego, którą można zdefiniować w następujący sposób:

Definicja 17. *Załóżmy, że X i Y są ciągłymi zmiennymi losowymi z kopulą C. Miarę Giniego dla tych zmiennych losowych definiujemy wzorem:*

$$\gamma_{X,Y} = \gamma_C = 2 \int\int_{\mathbb{I}^2} (|u+v-1| - |u-v|)\, dC(u,v). \tag{19}$$

Podobnie jak miary tau Kendalla i rho Spearmana, również miarę Giniego można zapisać za pomocą funkcji zgodności Q. Mówi o tym poniższe twierdzenie.

Twierdzenie 18. *Załóżmy, że X i Y są ciągłymi zmiennymi losowymi z kopulą C. Miara Giniego dla tych zmiennych losowych wyraża się wzorem:*

$$\gamma_{X,Y} = \gamma_C = Q(W,C) + Q(C,M). \tag{20}$$

Dowód. Zauważmy, że $M(u, v) = ½ (u + v - |u - v|)$. Mamy z wzoru (8):

$$Q(C,M) = 2\int\int_{\mathbb{I}^2}(u+v-|u-v|)\,dC(u,v)-1 = 1-2\int\int_{\mathbb{I}^2}|u-v|\,dC(u,v).$$

Podobnie korzystając z faktu, że $W(u, v) = ½ (u + v - 1 + |u + v - 1|)$ oraz (8), otrzymujemy:

$$Q(C,W) = 2\int\int_{\mathbb{I}^2}(u+v-1-|u+v-1|)\,dC(u,v)-1 =$$

$$= 2\int\int_{\mathbb{I}^2}|u+v-1|\,dC(u,v)-1,$$

co razem z poprzednim wzorem oraz (19) daje tezę twierdzenia.

Przykład 6. Na podstawie obliczeń z przykładu 1 oraz wzoru (20) mamy:

$$\gamma_M = 1, \quad \gamma_\Pi = 0, \qquad \gamma_W = -1.$$

Korzystając z wzoru (20), monotoniczności funkcji zgodności Q oraz faktu, że $W < C < M$ dla dowolnej kopuli C, otrzymujemy nierówność:

$$\gamma_W \le \gamma_C \le \gamma_M\,,$$

co można zapisać w postaci wniosku:

Wniosek 19. *Dla dowolnej kopuli C zachodzi oszacowanie:*

$$-1 \le \gamma_C \le 1. \tag{21}$$

4. Miary zgodności i zależności

Rozważane poprzednio miary zależności tau Kendalla, rho Spearmana oraz gamma Giniego mają pewne wspólne własności, które można sformułować jako aksjomaty **miary zgodności**.

Definicja 20. *Miara κ, zdefiniowana dla każdej pary ciągłych zmiennych losowych, jest miarą zgodności, jeżeli jest ona funkcją kopuli C tych zmiennych losowych oraz:*

1) $-1 \leq \kappa_{X,Y} \leq 1$,

2) $\kappa_{X,-X} = -1$ *i* $\kappa_{X,X} = 1$,

3) $\kappa_{X,Y} = 0$ *dla niezależnych zmiennych losowych X i Y*,

4) $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$,

5) jeżeli $C_1 < C_2$*, to* $\kappa_{C1} \leq \kappa_{C2}$,

6) jeżeli $\lim_{n\to\infty} C_n(u, v) = C(u, v)$ *dla każdego* $(u, v) \in \mathrm{I}^2$*, to* $\lim_{n\to\infty} \kappa_{Cn} = \kappa_C$.

Z powyższej definicji wynika, że miara zgodności κ ma następujące własności:

Wniosek 21. *Załóżmy, że κ jest miarą zgodności ciągłych zmiennych losowych X i Y. Wówczas:*

1) jeżeli Y jest ściśle rosnącą, deterministyczną funkcją X, to $\kappa_{X,Y} = \kappa_M = 1$,

2) jeżeli Y jest ściśle malejącą, deterministyczną funkcją X, to $\kappa_{X,Y} = \kappa_W = -1$,

3) jeżeli α i β są funkcjami ściśle rosnącymi, to $\kappa_{\alpha(X)}, \kappa_{\beta(Y)} = \kappa_{X,Y}$.

W podobny sposób można sformułować postulaty dla **miary zależności.**

Definicja 22. *Miara δ zdefiniowana dla każdej pary ciągłych zmiennych losowych X i Y jest miarą zależności, jeżeli jest ona funkcją kopuli C tych zmiennych losowych oraz:*

1) $0 \leq \delta_{X,Y} \leq 1$,

2) $\delta_{X,Y} = \delta_{Y,X}$,

3) $\delta_{X,Y} = 0$ *wtedy i tylko wtedy, gdy X i Y są niezależne,*

4) $\delta_{X,Y} = 1$ *wtedy i tylko wtedy, gdy Y jest prawie na pewno ściśle monotoniczną, deterministyczną funkcją X,*

5) jeżeli α i β są funkcjami ściśle rosnącymi, to $\delta_{\alpha(X),\, \beta(Y)} = \delta_{X,Y}$,

6) jeżeli $\lim_{n\to\infty} C_n(u, v) = C(u, v)$ *dla każdego* $(u, v) \in \mathrm{I}^2$*, to* $\lim_{n\to\infty} \delta_{Cn} = \delta_C$.

Pokażemy teraz przykłady miar zależności bazujące na wcześniejszych miarach zgodności. Zaczniemy od miary rho Spearmana, którą, korzystając z wzoru (16), można przedstawić w postaci:

$$\rho_C = 12 \int\!\!\int_{\mathbb{I}^2} (C(u,v) - uv)\, du\, dv.$$

Zmieniając nawias pod całką na wartość bezwzględną, otrzymujemy wzór:

$$\sigma_C = 12 \int\!\!\int_{\mathbb{I}^2} |C(u,v) - uv|\, du\, dv, \tag{22}$$

który można interpretować jako odległość kopul C i Π w przestrzeni $L^1(\mathrm{I}^2)$.

Można pokazać, że δ jest miarą zależności. Pierwszy raz miara ta była rozważana przez Schweizera i Wolffa [Schweizer i Wolff, 1981, s. 870-885; Wolf, 1981, s. 175-188] w 1981 roku. W podobny sposób, bazujący na odległości kopul C i Π w przestrzeniach $L^p(\mathrm{I}^2)$, $p \in [1,\infty)$, można konstruować miary zależności, uzyskując wzory:

$$K_p(C) = k_p \left(\int\int_{\mathbb{I}^2} |C(u,v) - uv|^p \, du \, dv \right)^{1/p}, \tag{23}$$

gdzie stała k_p jest tak dobrana, by $K_p(M) = K_p(W) = 1$. W szczególności dla $p = 2$ mamy $k_2 = 90$ i otrzymujemy miarę:

$$\Phi_C = 90 \left(\int\int_{\mathbb{I}^2} |C(u,v) - uv|^2 \, du \, dv \right)^{1/2}. \tag{24}$$

Kwadrat tej miary $\Phi^2_{X,Y}$ jest nazywany indeksem zależności zmiennych losowych X i Y.

W przypadku gdy $p = \infty$, otrzymujemy miarę:

$$\lambda_C = 4 \sup\{C(u,v) : \ (u,v) \in \mathbb{I}^2\}, \tag{25}$$

która spełnia wszystkie punkty definicji 22 z wyjątkiem punktu 4.

Literatura

Cherubini U., Della Lunga G. (2007), *Structured Finance. The Object Oriented Approach*, John Wiley & Sons, Chichester.

Cherubini U., Luciano E., Vecchiato W. (2004), *Copula Methods in Finance*, John Wiley & Sons, Chichester.

Denuit M., Dhaene J., Goovaerts M., Kaas R. (2005), *Actuarial Theory for Dependent Risks. Measures, Orders and Models*, John Wiley & Sons, Chichester.

Kulpa T. (2011), *Zastosowanie funkcji łącznikowych do modelowania zależnego ryzyka ubezpieczeniowego*, UKSW, Warszawa.

Nelsen R.B. (1999), *An Introduction to Copulas*, Lecture Notes in Statistics 139, Springer--Verlag, New York.

Schweizer B., Wolff E.F. (1981), *On Nonparametric Measures of Dependence for Random Variables*, "Ann. Statist.", No. 9.

Sklar A. (1959), *Fonctions de répartition à n dimensions et leur marges*, Publ. Inst. Statist. Univ. Paris, No. 8.

Wolff E.F. (1981), *N-dimensional Measures of Dependence*, "Stochastica", No. 4.

MEASURES OF DEPENDENCE BASED ON COPULAS

Summary: Measures of dependence and measures of association based on copulas are presented, including Kendall's Tau, Spearman's rho, Schweizer & Wolff's sigma and Gini's gamma. General axioms for measures of dependence and measures of association are discussed.

Keywords: Copula, measure of dependence, measure of association.