

WIELOWYMIAROWA ANALIZA STATYSTYCZNA W BADANIU RYNKU KAPITAŁOWEGO

Katarzyna Budny¹, Jan Tatar²

Wstęp

Zdecydowana większość (jeżeli nie wszystkie) wielkości rozważanych w procesie badania i analizy sytuacji społeczno-demograficzno-gospodarczej konkretnego kraju, grupy krajów bądź konkretnego regionu ma charakter zmiennych losowych. Tzn. są one funkcjami określonymi na przestrzeni probabilistycznej o wartościach należących do znanego zbioru. Jeżeli wartości te są ponadto kwantyfikowalne, to mamy wówczas do czynienia ze zmiennymi losowymi o wartościach liczbowych. Często badaniu poddawane są nie pojedyncze zmienne losowe ale ich stosownie dobrane „zestawy”, czyli zmienne losowe wielowymiarowe nazywane także wektorami losowymi.

Klasycznie rozumiana statystyka matematyczna, w konsekwencji także klasyczna statystyka opisowa, wypracowały stosowną metodologię opisu, badania i analizy wektorów losowych. Charakterystyczną cechą tej metodologii jest wszakże fakt, że badając dany wektor losowy w istocie badaniu poddaje się jednowymiarową zmienną losową będącą odpowiednio zdefiniowaną funkcją zmiennych losowych stanowiących „współrzędne” owego wektora.

W szeregu wcześniejszych prac autorzy niniejszego opracowania przedstawili propozycję odmiennego od klasycznego sposobu opisu oraz analizy wielowymiarowych rozkładów prawdopodobieństwa oraz wykorzystali zaproponowane nowe narzędzia do modelowania i badania tych wielkości, które mają charakter wielowymiarowych wektorów losowych.

W prezentowanej obecnie pracy – należącej także do nurtu badań nad wielowymiarowymi charakterystykami społeczno-gospodarczymi – uzyskane wcześniej „nowe” narzędzia probabilistyczne wykorzystane zostaną do opisu oraz badania wybranych wektorowych charakterystyk polskiego rynku kapitałowego. Zasygnalizujemy już w tym miejscu, że współrzędnymi badanych wektorów będą poziomy indeksów giełdowych: WIG, WIG-20, WIG-Banki oraz WIG-Paliwa a także rentowności tych indeksów. Dane, które wykorzystamy w części merytorycznej artykułu (p.2) pochodzą z okresu od 4 stycznia 2016 r. do 7 lipca 2017 r.

W pierwszej, metodologicznej części opracowania (p.1) przypominamy – wskazując jednocześnie na konkretne pozycje literaturowe – probabilistyczne i statystyczne narzędzia,

¹ Katedra Matematyki UEK

² Katedra Matematyki UEK

czyli definicje oraz postaci zarówno samych parametrów wielowymiarowych rozkładów prawdopodobieństwa jak i ich estymatorów, których wartości, czyli oszacowania (oceny) parametrów zostaną wyznaczone i zinterpretowane w części merytorycznej.

1. Wybrane parametry wielowymiarowych rozkładów prawdopodobieństwa oraz ich estymatory

Podstawowymi charakterystykami rozkładu jednowymiarowej zmiennej losowej są jego momenty zwykłe oraz centralne [por. np. Feller 1969, Shao 2003]. W pracach [Tatar 1996, 1999] – wykorzystując definicję potęgi wektora w przestrzeni z iloczynem skalarnym – zaproponowano wielowymiarowe uogólnienie tych pojęć.

Niech $n, r \in N_o = N \cup \{0\}$ oraz niech

$$L_n^r(\Omega) = \left\{ X : \Omega \rightarrow R^n : X \text{ wektor losowy i } \int_{\Omega} \|X\|^r dP < +\infty \right\}$$

będzie przestrzenią wektorów losowych sumowalnych (całkowalnych) w r -tej potęgde.

Nadmieńmy w tym miejscu, że w literaturze probabilistycznej wielkość $E[\|X\|^r] = \int_{\Omega} \|X\|^r dP$

bywa niekiedy nazywana [por. np. Bilodeau i Brenner 1999] momentem rzędu r wektora losowego X i oznaczana symbolem $E[X^r]$. Natomiast w pracy [Tatar 2002], przez analogię do przypadku jednowymiarowego, wyrażenie to określono jako *moment absolutny rzędu r wektora losowego X* .

Rozważmy zatem wektor $X : \Omega \rightarrow R^n$ należący do przestrzeni $L_n^r(\Omega)$ taki, że jego moment absolutny rzędu r jest skończony. Przyjmijmy ponadto, że

$$\Sigma_X = \begin{bmatrix} \sigma_1^2 & \dots & \rho_{1n}\sigma_1\sigma_n \\ \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \dots & \sigma_n^2 \end{bmatrix}$$

jest klasycznie rozumianą macierzą kowariancji wektora X , zaś

$$R_X = \begin{bmatrix} 1 & \dots & \rho_{1n} \\ \vdots & \ddots & \vdots \\ \rho_{1n} & \dots & 1 \end{bmatrix}$$

macierzą jego korelacji cząstkowych.

Definicja 1. [Tatar 1996, 1999] *Momentem zwykłym rzędu r wektora losowego $X : \Omega \rightarrow R^n$ nazywamy wyrażenie $\alpha_{r,n}(X) = E[X^r]$.*

Zauważmy, że moment zwykły rzędu pierwszego wektora losowego jest wektorem wartości oczekiwanych jego składowych, czyli $\alpha_{1,n}(X) = m(X) = EX$; moment ten będziemy nazywać *wartością oczekiwaną wektora losowego X* .

Definicja 2. [Tatar 1996, 1999] *Momentem centralnym rzędu r wektora losowego $X : \Omega \rightarrow R^n$ nazywamy wielkość $\mu_{r,n}(X) = E[(X - EX)^r]$.*

Szczególne znaczenie ma moment centralny rzędu drugiego wektora X , czyli $\mu_{2,n}(X) = E[(X - EX)^2]$. Moment ten będziemy nazywać *wariancją wektora losowego X* oraz oznaczać także symbolem $D^2 X$.

Za pomocą momentów centralnych wektora losowego definiujemy z kolei takie charakterystyki wielowymiarowego rozkładu prawdopodobieństwa jak *współczynnik asymetrii* czy *kurtoza*.

Definicja 3. [Tatar 2000] *Współczynnik asymetrii wektora losowego X definiujemy jako*

$$\gamma_{1,n}(X) = \text{Skew}X = \frac{\mu_{3,n}(X)}{(\mu_{2,n}(X))^{\frac{3}{2}}} = \frac{E[(X - EX)^3]}{(D^2 X)^{\frac{3}{2}}}.$$

Zwróćmy uwagę, że zdefiniowana powyżej miara asymetrii rozkładu wektora losowego dostarcza także wartości wektorowej. Wskazuje ona zatem kierunek, na którym występuje ewentualna „skośność” (asymetria). Do pomiaru i wyrażenia jej wielkości można z kolei wykorzystać długość lub kwadrat wektora $\gamma_{1,n}(X)$, czyli $\|\gamma_{1,n}(X)\|$ lub $\beta_{1,n}(X) = \gamma_{1,n}^2(X)$.

Definicja 4. [Budny 2009, Budny i Tatar 2009] *Kurtozą wektora losowego X nazywamy wielkość $\beta_{2,n}(X)$ wyrażoną jako*

$$\beta_{2,n}(X) = \text{Kurt}X = \frac{\mu_{4,n}(X)}{(\mu_{2,n}(X))^2} = \frac{E[(X - EX)^4]}{(D^2 X)^2}.$$

W pracy [Budny 2012] wykazano, że kurtoza wektora losowego $N : \Omega \rightarrow R^n$ o wielowymiarowym rozkładzie normalnym z tą samą wektorową wartością oczekiwaną i tą samą macierzą kowariancji jakimi charakteryzuje się wektor losowy X przyjmuje postać

$$\beta_{2,n}(N) = Kurt N = 1 + \frac{2 \sum_{i=1}^n (D^2 N_i)^2 + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^n \rho_{ij}^2 D^2 N_i \cdot D^2 N_j}{\sum_{i,j=1}^n D^2 N_i \cdot D^2 N_j} = 1 + \frac{2 \sum_{i=1}^n \sigma_i^4 + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^n \rho_{ij}^2 \sigma_i^2 \sigma_j^2}{\sum_{i,j=1}^n \sigma_i^2 \sigma_j^2},$$

gdzie ρ_{ij} oznacza współczynnik korelacji składowych N_i oraz N_j wektora N .

Wykorzystując pojęcie kurtozy w pracy [Budny 2014a] zdefiniowano kolejną charakterystykę wielowymiarowego rozkładu prawdopodobieństwa.

Definicja 5. [Budny 2014a] *Współczynnikiem ekscesu (ekscesem)* wektora losowego $X : \Omega \rightarrow R^n$ nazywamy wielkość $\gamma_{2,n}(X)$ określona następująco

$$\gamma_{2,n}(X) = Ekscess X = Kurt X - \left(1 + \frac{2 \sum_{i=1}^n (D^2 X_i)^2 + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^n \rho_{ij}^2 D^2 X_i \cdot D^2 X_j}{\sum_{i,j=1}^n D^2 X_i \cdot D^2 X_j} \right) = \beta_{2,n}(X) - \beta_{2,n}(N).$$

W badaniu wielowymiarowych wielkości losowych istotne znaczenie ma pomiar oraz analiza ich wzajemnych zależności. Jedną z ważniejszych miar w tym zakresie jest *współczynnik korelacji wielowymiarowej*.

Niech $X : \Omega \rightarrow R^m$ oraz $Y : \Omega \rightarrow R^n$ będą dowolnymi wektorami losowymi, przy czym o wektorze Y zakładamy, że $\mu_{2,n}(Y) = D^2 Y \neq 0$.

Definicja 6. [Budny 2017b] *Współczynnik korelacji wielowymiarowej (współczynnik*

korelacji wektorów losowych X oraz Y) definiujemy jako $\rho_{multi}(X, Y) = \sqrt{\frac{D^2(A_{X,Y} X)}{D^2 Y}}$, gdzie

$A_{X,Y}$ jest macierzą realizującą warunek $\min_{\substack{A \in M_{(n \times m)}(R) \\ B \in M_{(n \times 1)}(R)}} E[(Y - (AX + B))^2]$.

Ważnymi własnościami współczynnika korelacji wielowymiarowej są jego nieujemność oraz niezmienniczość względem skali i translacji wektorów. Inną istotną własnością miary ρ_{multi} jest brak jej symetrii; tzn. mogą istnieć pary wektorów X oraz Y , dla których zachodzi nierówność $\rho_{multi}(X, Y) \neq \rho_{multi}(Y, X)$.

W pracy [Budny 2017b] wykazano, że kwadrat zdefiniowanego powyżej współczynnika korelacji wielowymiarowej można – w sposób równoważny – zapisać jako

$$\rho_{multi}^2(X, Y) = \sum_{i=1}^n \frac{D^2 Y_i}{D^2 Y} \rho^2(X, Y_i),$$

gdzie $\rho(X, Y_i)$ jest współczynnikiem korelacji wielokrotnej (wielorakiej) między zmienną losową Y_i a wektorem losowym $X = (X_1, \dots, X_m)$, dla wszystkich $i \in \{1, \dots, n\}$.

W praktyce, w analizie wielowymiarowych danych empirycznych – badając ich charakterystyki oparte na definicji potęgi wektora – wykorzystuje się (podobnie jak to ma miejsce w przypadku jednowymiarowym) ich odpowiednie estymatory, czyli oszacowania uzyskane na podstawie dostępnej próby statystycznej. W pracach [Budny 2014b, 2017a, 2017b] zaproponowano postać estymatorów momentów zwykłych, momentów centralnych, współczynnika asymetrii, kwadratu współczynnika asymetrii, kurtozy i ekscesu wektora losowego oraz wykazano, że są to statystyki zgodne i (co najmniej) asymptotycznie nieobciążone. Dla potrzeb dalszej części pracy przypomnijmy ich definicje.

Niech w dalszym ciągu $X : \Omega \rightarrow R^n$ będzie badanym n -wymiarowym wektorem losowym oraz niech $k \in N$ oznacza liczebność wykorzystywanej próby losowej. Sama zaś próba losowa prosta z rozkładu n -wymiarowego niech będzie ciągiem wektorów

$$X^1 : \Omega \rightarrow R^n, \dots, X^k : \Omega \rightarrow R^n.$$

Definicja 7. [Budny 2014b, 2017b] Estymatorem momentu zwykłego rzędu r wektora losowego (inaczej: momentem zwykłym rzędu r w próbie wielowymiarowej) nazywamy statystykę

$$a_{r,n} = \frac{\sum_{i=1}^k (X^i)^r}{k}. \quad (1)$$

Szczególne znaczenie ma estymator momentu zwykłego rzędu pierwszego (estymator wartości oczekiwanej), czyli $a_{1,n}$ oznaczony także symbolem $\bar{X}$.

Definicja 8. [Budny 2017a, 2017b] Estymatorem momentu centralnego rzędu r wektora losowego (momentem centralnym rzędu r w próbie wielowymiarowej) nazywamy statystykę

$$m_{r,n} = \frac{\sum_{i=1}^k (X^i - \bar{X})^r}{k}. \quad (2)$$

W tym przypadku szczególnie istotny jest estymator momentu centralnego rzędu drugiego (tzn. estymator wariancji) $m_{2,n}$.

W konsekwencji pierwiastek kwadratowy z wariancji, czyli

$$s = \sqrt{m_{2,n}} = \sqrt{\frac{\sum_{i=1}^k (X^i - \bar{X})^2}{k}} \quad (3)$$

jest estymatorem odchylenia standardowego.

Estymatory momentów zwykłych i centralnych posłużyły do zdefiniowania kolejnych ważnych statystyk.

Definicja 9. [Budny 2017b] Estymatorem współczynnika asymetrii wektora losowego (inaczej: współczynnikiem asymetrii w próbie wielowymiarowej) nazywamy wielkość

$$\hat{\gamma}_{1,n} = \frac{m_{3,n}}{(m_{2,n})^{\frac{3}{2}}}. \quad (4)$$

Naturalną konsekwencją powyższej definicji są postacie:

estymatora długości wektora (współczynnika) asymetrii

$$\|\hat{\gamma}_{1,n}\| = \left\| \frac{m_{3,n}}{(m_{2,n})^{\frac{3}{2}}} \right\|, \text{ oraz} \quad (5)$$

estymatora kwadratu współczynnika asymetrii wektora losowego [Budny 2017b]:

$$\hat{\beta}_{1,n} = (\hat{\gamma}_{1,n})^2. \quad (6)$$

Kolejne dwie definicje dostarczają postaci estymatorów współczynników koncentracji i spłaszczenia wielowymiarowego rozkładu prawdopodobieństwa.

Definicja 10. [Budny 2017b] Estymatorem kurtozy wektora losowego (kurtozą w próbie wielowymiarowej) nazywamy statystykę

$$\hat{\beta}_{2,n} = \frac{m_{4,n}}{(m_{2,n})^2}. \quad (7)$$

Definicja 11. [Budny 2017b] Estymatorem współczynnika ekscesu wektora losowego (współczynnikiem ekscesu w próbie wielowymiarowej) nazywamy wielkość

$$\hat{\gamma}_{2,n} = \hat{\beta}_{2,n} - \left(1 + \frac{2 \sum_{i=1}^n (m_{2,n}^i)^2 + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^n r_{ij}^2 m_{2,n}^i m_{2,n}^j}{(m_{2,n})^2} \right), \quad (8)$$

przy czym $m_{2,n}^i$ oznacza „jednowymiarową” wariancję w próbie i – tej współrzędnej wektora losowego, a r_{ij}^2 współczynnik korelacji cząstkowej w próbie między i – tą oraz j – tą współrzędną, dla wszystkich $i, j \in \{1, \dots, n\}$.

Na koniec tej części pracy przypomnijmy postać estymatora zaproponowanej w definicji 6 miary zależności wektorów losowych $X : \Omega \rightarrow R^m$ oraz $Y : \Omega \rightarrow R^n$.

Definicja 12. [Budny 2017b]

Estymatorem kwadratu współczynnika korelacji wielowymiarowej wektorów X oraz Y nazywamy wielkość

$$R_{multi}^2(X, Y) = \sum_{i=1}^n \frac{m_{2,n}^{Y_i}}{m_{2,n}^Y} R_{X, Y_i}^2, \quad (9)$$

gdzie $m_{2,n}^{Y_i}$ jest klasycznym estymatorem „jednowymiarowej” wariancji zmiennej losowej Y_i ,

$m_{2,n}^Y = \sum_{i=1}^n m_{2,n}^{Y_i}$, natomiast R_{X, Y_i}^2 jest estymatorem kwadratu współczynnika korelacji

wielokrotnej między zmienną losową Y_i a zestawem zmiennych losowych $X_1, \dots, X_m$, czyli wektorem losowym $X = (X_1, \dots, X_m)$, dla wszystkich $i \in \{1, \dots, n\}$.

2. Wybrane charakterystyki polskiego rynku kapitałowego – analiza

W tej części pracy wykorzystamy przypomniane w p.1 charakterystyki wielowymiarowych rozkładów prawdopodobieństwa (oraz ich estymatory) w badaniu kilku wektorów losowych wybranych z polskiego rynku finansowego. Współrzędnymi (składowymi) tych wektorów będą wielkości wymienione we wstępie do niniejszego opracowania.

Próbę, którą wykorzystamy do uzyskania „ocen” wybranych parametrów rozkładu badanych wektorów stanowią dane pobrane ze strony www.money.pl. Są to poziomy „zamknięcia” w każdym dniu notowań w okresie od 4 stycznia 2016 r. do 10 lipca 2017 r. W przypadku danych nominalnych wykorzystujemy zatem próbę 380-elementową, zaś w badaniu rentowności wybranych instrumentów (indeksów giełdowych) – z oczywistych względów – próbę 379-elementową. Dla każdego z wyspecyfikowanych wektorów finansowych wyznaczmy estymatory następujących parametrów ich rozkładów:

- a) *wartość oczekiwana* (wektor); wg. wzoru (1)
- b) *wariancja łączna* (skalar); wg. wzoru (2)
- c) *odchylenie standardowe łączne* (skalar); wg. wzoru (3)
- d) *współczynnik asymetrii* (wektor); wg. wzoru (4)
- e) *norma/długość współczynnika asymetrii* (skalar); wg. wzoru (5)
- f) *kwadrat współczynnika asymetrii* (skalar); wg. wzoru (6)
- g) *kurtoza*; wg. wzoru (7)

oraz

- h) *współczynnik ekscesu*; wg. wzoru (8).

W celach poglądowych, poznawczych oraz porównawczych w każdym przypadku wyznaczmy także: *macierz kowariancji*, *macierz współczynników korelacji cząstkowych* oraz klasycznie rozumiane następujące charakterystyki rozkładów brzegowych: *wartość oczekiwana*, *wariancja*, *odchylenie standardowe*, *współczynnik asymetrii*, *kurtoza* oraz *współczynnik ekscesu*.

Dla wybranych par badanych finansowych wektorów losowych wyznaczmy także (wg. wzoru (9)) estymator kwadratu współczynnika korelacji wielowymiarowej jako jedną z miar ich zależności.

2.1.

Niech X_1 oznacza wartość indeksu WIG, X_2 - wartość indeksu WIG-20, X_3 - wartość indeksu WIG-Banki oraz X_4 - wartość indeksu WIG-Paliwa na Warszawskiej Giełdzie Papierów Wartościowych. Jako pierwszy poddamy badaniu wektor postaci $X = (X_1, X_2, X_3, X_4)$. Korzystając z danych, o których mowa we wprowadzeniu do p. 2, tj. z próby 380-elementowej, otrzymujemy następujące wyniki:

a) macierz kowariancji

$$\Sigma_X = \begin{bmatrix} 36590093,7 & 1265614,20 & 3773372,28 & 6361990,66 \\ 1265614,20 & 46583,96 & 134004,30 & 220345,67 \\ 3773372,28 & 134004,30 & 413766,16 & 630121,41 \\ 6361990,66 & 220345,67 & 630121,41 & 1169543,61 \end{bmatrix}$$

b) macierz współczynników korelacji cząstkowej

$$R_X = \begin{bmatrix} 1 & 0,9694 & 0,9698 & 0,9725 \\ 0,9694 & 1 & 0,9652 & 0,9440 \\ 0,9698 & 0,9652 & 1 & 0,9058 \\ 0,9725 & 0,9440 & 0,9058 & 1 \end{bmatrix}$$

c) charakterystyki rozkładów brzegowych wektora X

Tabela 1: Estymatory charakterystyk rozkładów brzegowych

Charakterystyka	WIG	WIG-20	WIG-Banki	WIG-Paliwa
Wartość oczekiwania	50 976,24	1 955,92	6 376,48	5 323,85
Wariancja	36 590 093,47	46 583,96	413 766,16	1 169 543,61
Odchylenie standardowe	6 048,98	215,83	643,25	1 081,45
Współczynnik asymetrii	0,6267	0,6909	0,5168	0,8203
Kurtoza	1,8384	1,9625	1,9341	2,1897
Współczynnik ekscesu	-1,1616	-1,0375	-1,0659	-0,8103

Źródło: Opracowanie własne

d) charakterystyki łączne wektora X

Tabela 2: Estymatory charakterystyk łącznych rozkładu wektora losowego X (opartych na definicji potęgi wektora)

Charakterystyka	Wartość estymatora
EX	(50 976,24; 1 955,92; 6 376,48; 5 323,85)
$D^2 X$	38 219 987,20
Odchylenie standardowe	6 182,23
$\gamma_{1,n}(X)$	(0,6162; 0,0238; 0,0625; 0,1197)
Norma (długość) $\gamma_{1,n}$	0,6313
$\beta_{1,n}(X)$	0,3985
$\beta_{2,n}(X)$	1,8419

$\gamma_{2,n}(X)$	-1,1488
-------------------	---------

Źródło: Opracowanie własne

Interpretacja wyników:

- (i) Wariancja łączna wektora losowego X jest równa sumie wariancji rozkładów brzegowych. Fakt ten nie jest zaskoczeniem ponieważ w pracach (Tatar 1996; Tatar 1999) udowodniono prawdziwość tego związku w przypadku każdego wektora losowego.
- (ii) Największy wpływ na wartość łącznego odchylenia standardowego wektora losowego X ma odchylenie standardowe zmiennej losowej X_1 , czyli poziomu indeksu WIG.
- (iii) Każdy z rozkładów brzegowych wektora X charakteryzuje się wyraźną asymetrią prawostronną („dodatnią”). Najbardziej „wydłużone prawe ramię” ma rozkład zmiennej „WIG-Paliwa” (0,8203) zaś najmniej rozkład indeksu „WIG-Banki”.
- (iv) Łączna asymetria wektora X wyrażona czterowymiarowym wektorem $\gamma_{1,n} = (0,6162; 0,0238; 0,0625; 0,1197)$ również wskazuje na „dodatnią skośność” jego rozkładu na każdym z czterech kierunków. Jednak zdecydowanie największą wartość ma ta współrzędna wektora asymetrii, która odpowiada „kierunkowi” (zmiennej losowej) WIG. Co więcej, długość wektora asymetrii łącznej (0,6313) jest zbliżona do wartości współczynnika asymetrii brzegowej zmiennej losowej WIG. To spostrzeżenie skłania do konkluzji, że próba analizy asymetrii wektora losowego poprzez analizę asymetrii jego rozkładów brzegowych może prowadzić do nieuprawnionych wniosków. W rozważanym przez nas przypadku takie podejście mogłoby – błędnie – sugerować, że wektor X będzie charakteryzował się największą asymetrią (skośnością) na kierunku WIG-Paliwa. Możemy zatem sformułować – skądinąd zupełnie „zdroworozsądkowy” – wniosek, że analizując asymetrię rozkładów wielowymiarowych nie wystarczy zbadać jak bardzo (i w którą stronę) „wydłużone” są ramiona rozkładów brzegowych ale trzeba nadto zmierzyć jak duża „masa prawdopodobieństwa” mieści się pod owymi ramionami, a na to pozwala właśnie współczynnik asymetrii łącznej $\gamma_{1,n}$.
- (v) Ostatnie dwa wskaźniki, tzn. kurtoza oraz (będący jej konsekwencją) współczynnik ekscesu mogą być interpretowane jako miary odpowiednio „spłaszczenia” badanego rozkładu oraz jego różnienia się („odbiegania”) od rozkładu normalnego o tej samej wartości oczekiwanej oraz wariancji.

Interpretacja ta odnosi się zarówno do rozkładów jednowymiarowych jak i wielowymiarowych. W przypadku badanego przez nas wektora indeksów giełdowych największym spłaszczeniem charakteryzuje się rozkład brzegowy WIG-Paliwa (2,1897) chociaż jest on najmniej (spośród wszystkich czterech rozkładów brzegowych) „odbiegający” od rozkładu normalnego (eksces równy - 0,8103).

Kurtoza łącznego rozkładu wektora X przyjmuje wartość (1,84190) zbliżoną do wartości kurtozy głównego indeksu WGPW, czyli WIG-u. Podobnie rzecz się ma w odniesieniu do wskaźnika ekscesu: czterowymiarowy rozkład wektora indeksów giełdowych różni się od odpowiadającego mu rozkładu normalnego w przybliżeniu w tym samym stopniu (-1,1488) w jakim rozkład WIG-u różni się od rozkładu normalnego o wartości oczekiwanej 50 976,24 oraz odchyleniu standardowym 6048,98.

Interpretując wartość estymatora współczynnika ekscesu należy ponadto zwrócić uwagę na jego znak. Ujemna wartość wskazuje na to, że „odbieganie” od rozkładu normalnego ma miejsce przede wszystkim dla realizacji bardziej odległych od wartości oczekiwanej (czyli w przypadku jednowymiarowym na tzw. „ogonach”) zaś dodatni współczynnik ekscesu świadczy o różnicach między badanym rozkładem a rozkładem normalnym dla realizacji bliskich wartości średniej.

2.2.

Przy oznaczeniach z p. 2.1 będziemy teraz rozważać wektor losowy $Y=(X_1, X_2)$, czyli wektor którego brzegowymi zmiennymi losowymi są WIG oraz WIG-20. Korzystając z tej samej próby historycznej (ciąg 380-elementowy) wyznaczmy estymatory charakterystyk wymienionych na początku p. 2. Stosując stosowne formuły otrzymujemy:

- a) macierz kowariancji

$$\Sigma_Y = \begin{bmatrix} 36590093,7 & 1265614,20 \\ 1265614,20 & 46583,96 \end{bmatrix}$$

- b) macierz współczynników korelacji cząstkowej

$$R_Y = \begin{bmatrix} 1 & 0.9694 \\ 0.9694 & 1 \end{bmatrix}$$

- c) charakterystyki rozkładów brzegowych wektora Y ; wartości estymatorów zostały obliczone w p. 2.1 i znajdują się w Tabeli 1.

d) charakterystyki łączne wektora Y

Tabela 3: Estymatory charakterystyk łącznych rozkładu wektora losowego Y
(opartych na definicji potęgi wektora)

Charakterystyka	Wartość
EY	(50 976,24 ; 1 955,92)
D^2Y	36 636 677,43
Odchylenie standardowe	6 052,82
$\gamma_{1,n}(Y)$	(0,6264 ; 0,0242)
Norma (długość) $\gamma_{1,n}$	0,6269
$\beta_{1,n}(Y)$	0,3930
$\beta_{2,n}(Y)$	1,8384
$\gamma_{2,n}(Y)$	-1,1613

Źródło: Opracowanie własne

Interpretacja wyników

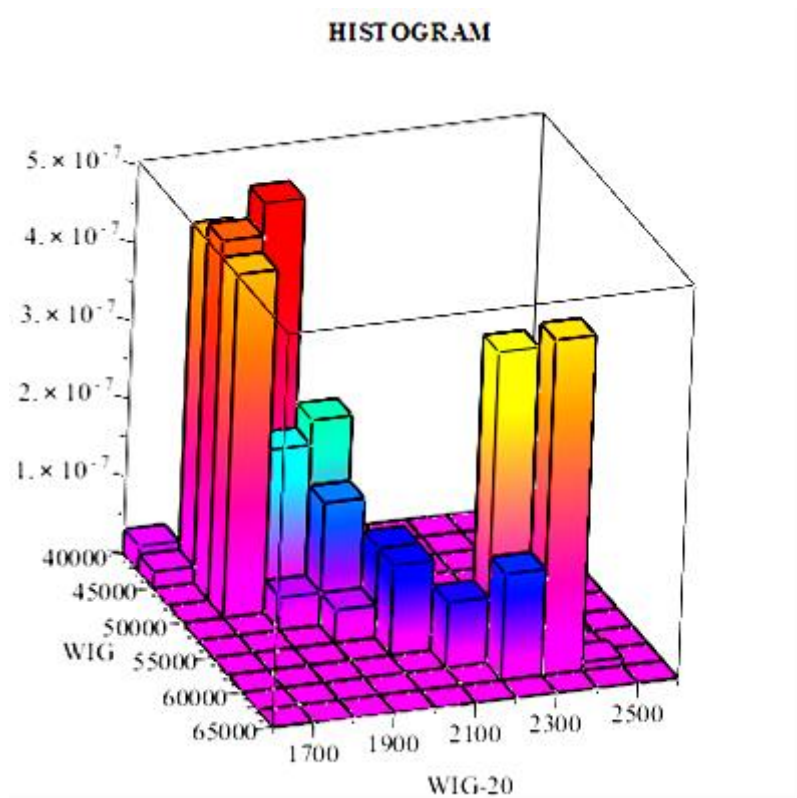
- (i) Podobnie jak w przypadku wektora badanego w p. 2.1 wariancja łączna jest liczbą, czyli skalarzem nie zaś wektorem. Jest to immanentna własność wariancji liczonej według nowej koncepcji charakteryzacji wielowymiarowych rozkładów prawdopodobieństwa. Co więcej, według tej koncepcji wszystkie momenty rzędu parzystego są skalarzami, zaś wszystkie momenty rzędu nieparzystego są wektorami. I znowu: nie powinno to budzić żadnego zaniepokojenia; przeciwnie, skoro momenty rzędu nieparzystego mają służyć (i służą) do definiowania parametrów położenia rozkładu, to w przestrzeni wielowymiarowej powinny być wektorami; skoro zaś momenty rzędu parzystego wykorzystuje się do określania parametrów rozproszenia danego rozkładu, to – także w przestrzeni wielowymiarowej – powinny być skalarzami (liczbami).
- (ii) Również teraz, współczynniki asymetrii obliczone oddzielnie dla obydwu rozkładów brzegowych mogłyby skłaniać do błędnego wniosku, że badany wektor losowy jest „bardziej asymetryczny” na kierunku X_2 , tzn. WIG-20, (0,6909 vs 0,6267). Tymczasem wyznaczony łączny (wektorowy) współczynnik (0,6264 ; 0,0242) wskazuje na zdecydowaną skośność na kierunku WIG. O przyczynach takiego stanu rzeczy pisaliśmy interpretując wyniki uzyskane w p. 2.1.
- (iii) Rozkład prawdopodobieństwa wektora dwóch indeksów giełdowych rozważanego w tym punkcie pracy także charakteryzuje się znacznym spłaszczeniem (kurtoza

równa 1,8384) oraz „odbieganiem” od normalności (-1,1613) przede wszystkim dla tych realizacji, które są dość odległe od wartości oczekiwanej (ujemny wskaźnik ekscesu).

- (iv) Dwuwymiarowość wektora Y stwarza możliwość graficznego przedstawienia jego rozkładu. Poniżej prezentujemy trzy wykresy: histogram rozkładu wektora Y skonstruowany w oparciu o przywoływaną już 380-elementową próbę; dwuwymiarowy rozkład normalny o tej samej wartości oczekiwanej oraz wariancji; a także jednocześnie zestawienie obydwu grafik. Wykresy te potwierdzają słuszność sformułowanych wcześniej wniosków.

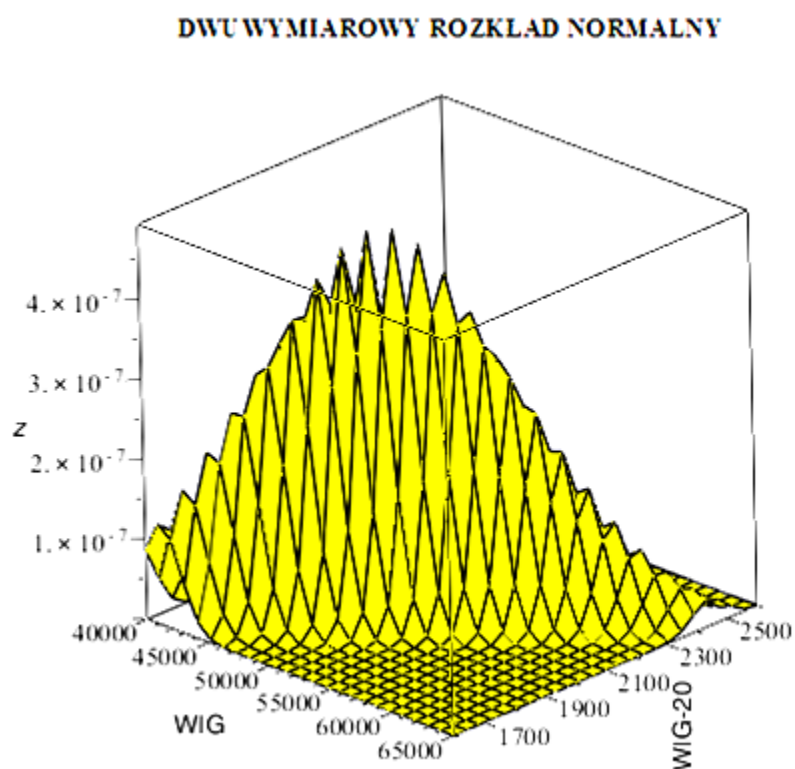
Niestety, z przyczyn oczywistych, nie jest możliwe graficzne zobrazowanie rozkładu (oraz jego charakterystyk) czterowymiarowego wektora badanego w poprzednim punkcie pracy – wymagałoby to sporządzenia wykresu w przestrzeni pięciowymiarowej (!).

Rysunek 1 *Histogram rozkładu wektora Y .*



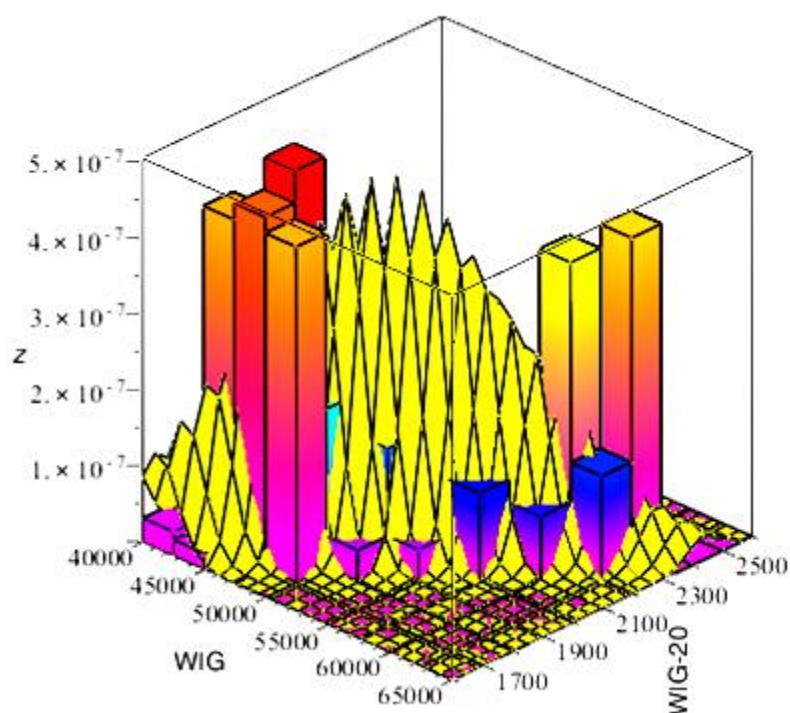
Źródło: opracowanie własne

Rysunek 2 Dwuwymiarowy rozkład normalny o wektorze wartości oczekiwanych EY i macierzy kowariancji Σ_Y .



Źródło: opracowanie własne

Rysunek 3 Zestawienie histogramu oraz dwuwymiarowego rozkładu normalnego.



Źródło: opracowanie własne

2.3.

Niech w dalszym ciągu obowiązują ustalenia z punktu 2.1 oraz niech $Y=(X_1, X_2)$ i $Z=(X_3, X_4)$. Będziemy zatem rozważać dwa dwuwymiarowe wektory: $Y=(WIG, WIG-20)$ oraz $Z=(WIG-Banki, WIG-Paliwa)$. Zbadamy ich wzajemną zależność obliczając – według wzoru (9) – estymator kwadratu wielowymiarowego współczynnika korelacji. Otrzymujemy:

$$\rho_{multi}^2(Y, Z) = 0,9473 \quad \text{oraz} \quad \rho_{multi}^2(Z, Y) = 0,9898 .$$

Powyższy rezultat potwierdza, że wykorzystywany w niniejszej pracy współczynnik korelacji (jako miara zależności wektorów losowych) nie ma własności symetrii. I nie należy w tym upatrywać niczego zaskakującego: pozostając we wzajemnym związku wektory Y i Z pełnią w nim różne role, a zatem „siła” zależności Y od Z nie musi być (i na ogół nie jest) taka sama jak „siła” zależności Z od Y . Uwaga powyższa obowiązuje także w przypadku zmiennych losowych jednowymiarowych: jeżeli – na przykład – zysk danego przedsiębiorstwa w jakimś stopniu zależy od ogólnego poziomu rynkowych stóp procentowych, to przecież wysokość stóp procentowych zapewne w innym stopniu (jeżeli w ogóle) zależy od zysku tego konkretnego przedsiębiorstwa. W rozważanym przez nas przypadku okazuje się, że obie zależności (Y od Z oraz Z od Y) są bardzo silne, co – biorąc pod uwagę składowe oraz strukturę tych wektorów Y i Z – nie jest zaskoczeniem. Nie zaskakuje także fakt, że w tym przypadku zależność między Z a Y jest wyraźnie silniejsza niż zależność między Y a Z .

2.4.

Niech w dalszym ciągu $X=(X_1, X_2, X_3, X_4)$ będzie wektorem ustalonym na początku punktu 2.1, czyli wektorem czterech wybranych indeksów notowanych na WGPW. Określimy ponadto wektor $U=(U_1, U_2, U_3, U_4)$, gdzie U_1 jest dobową (lepiej: jednosesyjną) rentownością indeksu WIG, U_2 - dobową rentownością indeksu WIG-20, U_3 - dobową rentownością indeksu WIG-Banki i wreszcie U_4 - dobową rentownością wskaźnika WIG-Paliwa. Rozważamy zatem czterowymiarowy wektor rentowności (stóp zwrotu) wybranych indeksów. Ponieważ mówimy o klasycznie rozumianych rentownościach jednosesyjnych, więc historyczna próba, którą wykorzystamy do dalszych obliczeń jest próbą 379-elementową.

Postępując podobnie jak dotychczas oraz korzystając z odpowiednich formuł otrzymujemy:

- a) macierz kowariancji

$$\Sigma_U = \begin{bmatrix} 0.000086 & 0.000103 & 0.000118 & 0.000091 \\ 0.000103 & 0.000129 & 0.000145 & 0.000114 \\ 0.000118 & 0.000145 & 0.000205 & 0.000093 \\ 0.000091 & 0.000114 & 0.000093 & 0.000222 \end{bmatrix}$$

b) macierz współczynników korelacji cząstkowej

$$R_U = \begin{bmatrix} 1 & 0.9825 & 0.8892 & 0.6587 \\ 0.9825 & 1 & 0.8937 & 0.6765 \\ 0.8892 & 0.8937 & 1 & 0.4381 \\ 0.6587 & 0.6765 & 0.4381 & 1 \end{bmatrix}$$

c) charakterystyki rozkładów brzegowych wektora U

Tabela 4: Estymatory charakterystyk rozkładów brzegowych

Charakterystyka	rent. WIG	rent. WIG-20	rent. WIG-Banki	rent.WIG-Paliwa
Wartość oczekiwania	0,000818	0,000706	0,000659	0,001463
Wariancja	0,000086	0,000129	0,000205	0,000222
Odchylenie standardowe	0,009274	0,026571	0,014318	0,014900
Współczynnik asymetrii	-0,2429	-0,0012	0,3592	0,0646
Kurtoza	4,6668	3,8779	4,9574	3,5143
Współczynnik ekscesu	1,6668	0,8779	1,9574	0,5143

Źródło: Opracowanie własne

d) charakterystyki łączne wektora U

Tabela 5: Estymatory charakterystyk łącznych rozkładu wektora losowego U (opartych na definicji potęgi wektora)

Charakterystyka	Wartość
EU	(0,000818 ; 0,000706 ; 0,000659 ; 0,001463)
D^2U	0,000641
Odchylenie standardowe	0,025318
$\gamma_{1,n}(U)$	(-0,002488 ; 0,023501 ; 0,072290 ; 0,004629)
Norma (długość) $\gamma_{1,n}$	0,076197
$\beta_{1,n}(U)$	0,005806
$\beta_{2,n}(U)$	3,0710
$\gamma_{2,n}(U)$	0,7748

Źródło: Opracowanie własne

Interpretacja uzyskanych wyników:

- (i) Rozważając odrębnie rozkłady brzegowe wektora U zauważamy, że największą asymetrią (prawostronną) charakteryzuje się zmienna losowa „rent.WIG-Banki” (0,3592); kolejną zmienną jest pod tym względem „rent.WIG” (-0,2429); najmniejszą asymetrię brzegową wykazuje „rent.WIG-20” (-0,0016). Z kolei asymetria łączna (wektorowa) wskazuje, że wprowadzie wektor U charakteryzuje się największą skośnością także na kierunku „rent.WIG-Banki” (0,0723) ale kolejnym jest kierunek „rent.WIG-20” ze skośnością równą 0,0235.
- (ii) Kurtoza łączna wektora U wynosi 3,071 i jest zdecydowanie mniejsza niż kurtoza każdego z rozkładów brzegowych. Oznacza to, że łącznie rozpatrywany wektor rentowności indeksów giełdowych nie wykazuje aż tak dużego „spłaszczenia” jak każdy z jego rozkładów brzegowych.
- (iii) Miara „odbiegania” rozkładu łącznego wektora rentowności od rozkładu normalnego (współczynnik ekscesu) wynosi 0,7748 i jest znacznie niższa od ekscesu wyznaczonego w p.2.1 dla wektora wartości nominalnych wybranych indeksów. Potwierdza to słuszność podejścia stosowanego w badaniu jednowymiarowych cen aktywów oraz ich rentowności, zgodnie z którym o ile często uzasadnionym jest założenie o normalności rozkładu stóp zwrotu, to na ogół błędnym byłoby analogiczne przypuszczenie w odniesieniu do cen nominalnych rozważanych aktywów.

2.5.

Niech $U = (U_1, U_2, U_3, U_4)$ będzie w dalszym ciągu wektorem rentowności wybranych indeksów WGPW określonym w p. 2.4. Rozważmy teraz wektor $V = (U_1, U_2)$, czyli $V = (\text{rent.WIG}, \text{rent.WIG} - 20)$. Wykorzystując w dalszym ciągu 279-elementowy ciąg danych historycznych otrzymujemy następujące estymatory wybranych charakterystyk:

- a) macierz kowariancji

$$\Sigma_V = \begin{bmatrix} 0.000086 & 0.000103 \\ 0.000103 & 0.000129 \end{bmatrix}$$

- b) macierz współczynników korelacji cząstkowej

$$R_V = \begin{bmatrix} 1 & 0,9825 \\ 0,9825 & 1 \end{bmatrix}$$

- c) charakterystyki rozkładów brzegowych wektora V ; wartości ich estymatorów zostały obliczone w p. 2.4 i znajdują się w Tabeli 4.
- d) charakterystyki łączne wektora V

Tabela 6: Estymatory charakterystyk łącznych rozkładu wektora losowego V (opartych na definicji potęgi wektora)

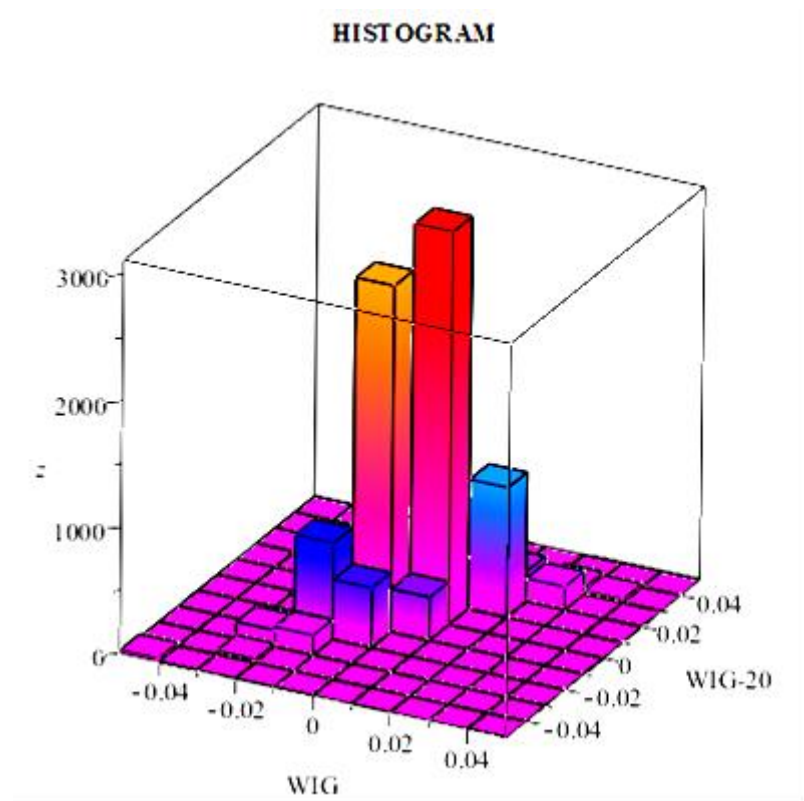
Charakterystyka	Wartość
EV	(0,000818 ; 0,000706)
D^2V	0,000214
Odchylenie standardowe	0,014629
$\gamma_{1,n}(V)$	(-0,088765 ; -0,047867)
Norma (długość) $\gamma_{1,n}$	0,100849
$\beta_{1,n}(V)$	0,010171
$\beta_{2,n}(V)$	4,1276
$\gamma_{2,n}(V)$	1,1609

Źródło: Opracowanie własne

Interpretacja uzyskanych wyników:

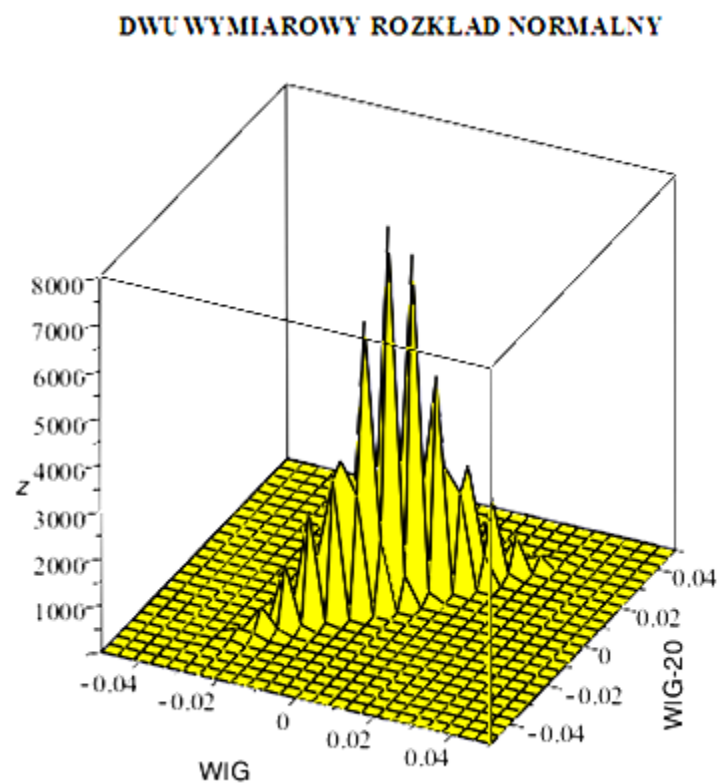
- (i) Kolejny raz zauważamy, że łączny wektor asymetrii $\gamma_{1,n}(V)$ wskazuje na zupełnie inny jej kierunek niż ten, który mógłby być odczytany z współczynników asymetrii rozkładów brzegowych. Także „poziom (wielkość)” asymetrii łącznej jest różna od poziomu asymetrii rozkładów brzegowych.
- (ii) Kurtoza oraz współczynnik ekscesu rozkładu wektora V sugerują, że znacznie bardziej „odbiega” on od rozkładu normalnego niż było w przypadku rozważanego w p. 2.4 wektora U . Fakt ten można także dostrzec na poniższych ilustracjach: zauważamy, że histogram rozkładu wektora V jest zdecydowanie bardziej „spłaszczony” niż rozkład normalny o tej samej wartości oczekiwanej oraz wariancji.

Rysunek 4 Histogram rozkładu wektora V .



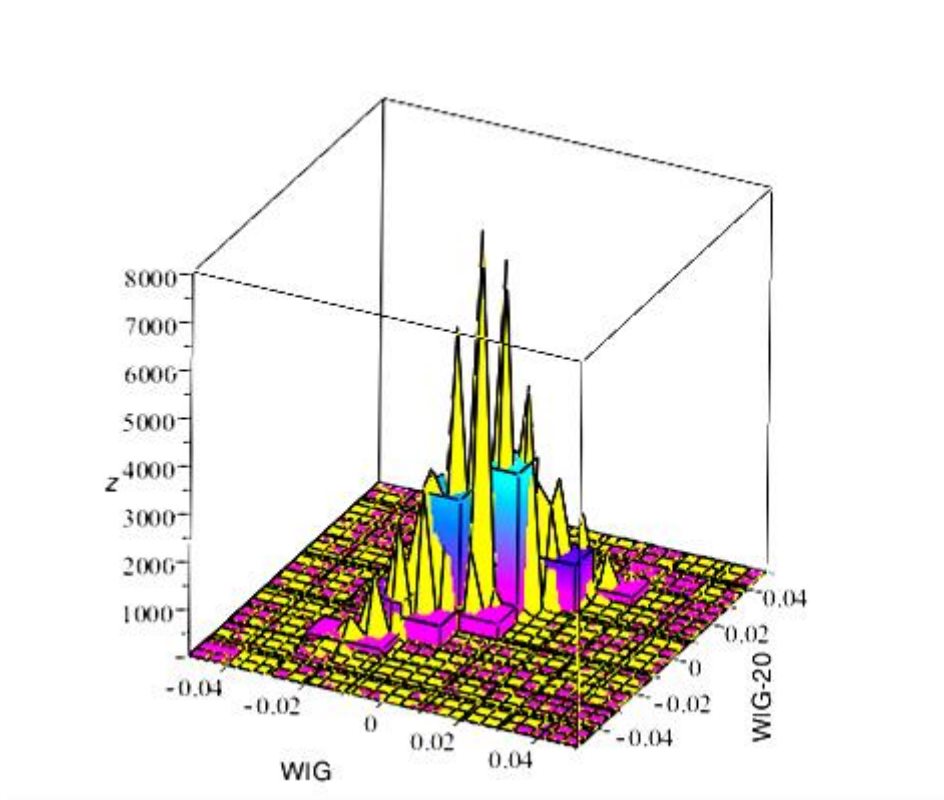
Źródło: opracowanie własne

Rysunek 5 Dwuwymiarowy rozkład normalny o wektorze wartości oczekiwanych EV i macierzy kowariancji Σ_V .



Źródło: opracowanie własne

Rysunek 6 Zestawienie histogramu oraz dwuwymiarowego rozkładu normalnego.



Źródło: opracowanie własne

2.6.

W ostatnim punkcie prowadzonej analizy przyjmijmy, że $V = (U_1, U_2)$ oraz $W = (U_3, U_4)$, gdzie U_1, U_2, U_3, U_4 mają znaczenie nadane im w p. 2.4. Rozważamy zatem dwa dwuwymiarowe wektory:

$$V = (\text{rent.WIG}, \text{rent.WIG} - 20) \quad \text{oraz} \quad W = (\text{rent.WIG} - \text{Banki}, \text{rent.WIG} - \text{Paliwa}).$$

Obliczając estymatory kwadratu współczynników korelacji wielowymiarowe (korzystając z wzoru (9)) otrzymujemy:

$$\rho_{multi}^2(V, W) = 0,6249 \quad \text{oraz} \quad \rho_{multi}^2(W, V) = 0,8919.$$

Komentarz jaki można uczynić do uzyskanych wyników jest analogiczny do tego jaki sformułowaliśmy na zakończenie p. 2.3 .

Uwagi końcowe i perspektywy badawcze

Zaprezentowane w pracy przykłady wykorzystania łącznych charakterystyk wielowymiarowych rozkładów prawdopodobieństwa pozwalają sformułować wniosek, że owo odmienne od klasycznego podejście do analizy rozkładów wektorów losowych pozwala

wierniej opisać badaną rzeczywistość niż badanie ich jednowymiarowych wektorów losowych. Odnosi się to m.in. do takich ich charakterystyk jak asymetria, kurtoza czy współczynnik ekscesu. W szczególności eksces rozkładu wielowymiarowego może być ważną przesłanką na rzecz ewentualnego formułowania hipotezy o jego normalności. Oczywiście, taka hipoteza musiałaby zostać poddana stosownej weryfikacji. Do tego zaś potrzebne będą odpowiednie testy statystyczne. Konstrukcja takich testów jawi się jako kolejne zadanie w procesie rozwijania koncepcji „nowej charakteryzacji wielowymiarowych rozkładów prawdopodobieństwa”. Owszem, zadanie to nie musi być łatwe ale – zdaniem autorów – dogodnym punktem wyjścia do jego realizacji mogą być uzyskane już wcześniej postaci estymatorów szeregu charakterystyk. Estymatory te z powodzeniem zostały wykorzystane także w niniejszym opracowaniu.

Publikacja została sfinansowana ze środków przyznanych Wydziałowi Finansów i Prawa Uniwersytetu Ekonomicznego w Krakowie w ramach dotacji na utrzymanie potencjału badawczego.

Literatura

- Bilodeau M., Brenner D. [1999], *Theory of Multivariate Statistics*, Springer - Verlag, New York.
- Budny K. [2009], *Kurtoza wektora losowego*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, **78**, seria: Ekonometria, **26**, 44 -54.
- Budny K., Tatar J. [2009], *Kurtosis of a random vector – special types of distributions*, „Statistics in Transition - new series”, **10** (3), 445 – 456.
- Budny K. [2012], *Kurtoza wektora losowego o wielowymiarowym rozkładzie normalnym*, [w:] *Zastosowanie metod ilościowych w finansach i ubezpieczeniach*, red. S. Folrlicz, CeDeWu, Warszawa.
- Budny K. [2014a], *Współczynnik ekscesu wektora losowego*, „Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach”, **203**, 28 – 38.
- Budny K. [2014b], *Estymacja momentów zwykłych wektora losowego opartych na definicji potęgi wektora*, „Folia Oeconomica Cracoviensia”, **55**, 81 – 96.
- Budny K. [2017a], *Estimation of the central moments of a random vector based on the definition of the power of a vector*, „Statistics in Transition – new series”, **18**(1), 1 – 20.

- Budny K. [2017b] *Nowe charakterystyki rozkładu i zależności wektorów losowych – konstrukcja, estymacja, zastosowania*, Monografie: Prace doktorskie, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie (w druku) .
- Feller W. [1969], *Wstęp do rachunku prawdopodobieństwa*. Tom II, PWN, Warszawa.
- Shao J. [2003], *Mathematical statistics*, 2nd ed., Springer.
- Tatar J. [1996], *O niektórych miarach rozproszenia rozkładów prawdopodobieństwa*, „Przegląd Statystyczny”, **43** (3 – 4), 267 – 274.
- Tatar J.[1999], *Moments of a random variable in a Hilbert space*, „Przegląd Statystyczny”, **46**(2), 261 – 271.
- Tatar J. [2000], *Asymetria wielowymiarowych rozkładów prawdopodobieństwa*, [w:] Materiały z XXXV Konferencji Statystyków, Ekonometryków i Matematyków Akademii Ekonomicznych Polski Południowej zorganizowanej przez Katedrę Statystyki Akademii Ekonomicznej w Krakowie (Osieczany, 23-25 III 1999 r.), Wydawnictwo AE w Krakowie, Kraków.
- Tatar J. (2002), *Nierówność Lapunowa dla wielowymiarowych rozkładów prawdopodobieństw*, „Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie”, nr 549.