Zeszyty Naukowe Uniwersytet Ekonomiczny w Krakowie
Metody analizy danych
923
ISSN 1898-6447
Zesz. Nauk. UEK, 2013; 923: 47–58
DOI: 10.15678/ZNUEK.2013.0923.04

Katarzyna Budny
Katedra Matematyki
Uniwersytet Ekonomiczny w Krakowie

Wybrane własności kurtozy wektora losowego

Streszczenie

Prezentowana praca jest kontynuacją rozważań prowadzonych nad kurtozą wektora losowego. Kurtoza wektora losowego rozumiana jest tutaj jako moment centralny czwartego rzędu wektora losowego podzielony przez kwadrat jego wariancji. Pojęcie momentu centralnego wektora losowego, a w szczególności wariancji, opiera się na definicji potęgi wektora i zostało zaproponowane przez J. Tatara. W niniejszym artykule przedstawione zostały wybrane, istotne własności kurtozy wektora losowego. Należą do nich m.in. własność niezmienniczości względem pewnych przekształceń afinicznych. Ponadto ustalony został związek między kurtozą wektora losowego a kwadratem jego współczynnika asymetrii. Spełnienie podanych własności przez tak skonstruowaną miarę może uzasadniać jej wybór na wielowymiarowy odpowiednik kurtozy jednowymiarowej zmiennej losowej.

Słowa kluczowe: kurtoza wektora losowego, potęga wektora, rozkład wielowymiarowy, charakterystyki rozkładu wielowymiarowego.

1. Wprowadzenie

Prezentowana praca jest kontynuacją rozważań prowadzonych nad kurtozą wektora losowego [Budny i Tatar 2009, Budny 2009]. U podstaw tych rozważań leży zaproponowana przez J. Tatara [1996 i 1999] definicja potęgi wektora w przestrzeni z iloczynem skalarnym.

Definicja 1. Dla dowolnego $v \in R^n$ oraz dowolnej liczby $k \in N_\circ = N \cup \{0\}$ k-tą potęgę wektora v definiujemy w następujący sposób:

$$v^o = 1 \in R \text{ oraz } v^k = \begin{cases} v^{k-1} \cdot v & \text{dla } k\text{-nieparzystych} \\ \langle v^{k-1}, v \rangle & \text{dla } k\text{-parzystych} \end{cases}$$

W pracy rozważać będziemy przestrzeń wektorową (R^n, R, +, $\cdot$) w której określono klasyczny (euklidesowy) iloczyn skalarny postaci:

$$\langle v, w \rangle := \sum_{i=1}^{n} v_i w_i, \text{ gdzie } v = (v_1, v_2, \ldots, v_n), w = (w_1, w_2, \ldots, w_n) \in R^n.$$

Dzięki wykorzystaniu pojęcia potęgi wektora zostały zdefiniowane m.in. momenty centralne wielowymiarowego wektora losowego [Tatar 1996 i 1999].

W pracy przyjęto następujące oznaczenie:

– D^2X to moment centralny rzędu 2 wektora losowego $X = (X_1, \ldots, X_k): \Omega \to R^k$ oparty na definicji potęgi wektora czy też równoważnie wariancja całkowita, czyli $D^2 X = \sum_{i=1}^{k} D^2 X_i$ (por. [Bilodeu i Brenner 1999]).

Za pomocą momentów centralnych wektora losowego, na podstawie postaci kurtozy jednowymiarowej zmiennej losowej [Cramer 1958, Jakubowski i Sztencel 2004], zdefiniowana została kurtoza wektora losowego.

Załóżmy zatem, że $X = (X_1, \ldots, X_n): \Omega \to \Re^n$ jest wektorem losowym, dla którego istnieje moment centralny czwartego rzędu.

Definicja 2. Przez kurtozę wektora losowego $X: \Omega \to \Re^n$ rozumiemy liczbę postaci:

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

Tak skonstruowany wskaźnik posiada szereg istotnych, pożądanych własności. Dotychczas podano postać kurtozy dla wybranych typów rozkładów (przypadek wektorów losowych o stochastycznie niezależnych współrzędnych, posiadających rozkłady normalne czy też rozkłady *t*-Studenta) [Budny i Tatar 2009] oraz nieco ogólniej wyznaczono postać kurtozy dla wektora losowego o stochastycznie niezależnych współrzędnych [Budny 2009]. Warto w tym miejscu zwrócić uwagę na wektor losowy, którego wszystkie współrzędne posiadają ten sam rozkład normalny i są stochastycznie niezależne. Kurtoza takiego wektora losowego jest bowiem postaci:

$$\beta_{2,n} = 1 + \frac{2}{n}.$$

W pracy omówione zostaną kolejne ważne własności kurtozy wektora losowego, rozumianej jako iloraz momentu centralnego czwartego rzędu wektora losowego dzielonego przez kwadrat jego wariancji całkowitej. Własności te wraz z przedstawionymi powyżej uzasadniają wybór tak skonstruowanej miary na wielowymiarowy odpowiednik kurtozy zmiennej losowej (jednowymiarowej).

W literaturze przedmiotu pojawiają się propozycje definicji kurtozy wektora losowego [Mardia 1970, Kotz, Johnston i Balakrishnan 2000]. Prezentowane w niniejszym opracowaniu podejście jest inne niż spotykane do tej pory.

2. Kurtoza jako niezmiennik pewnych przekształceń afinicznych

W kolejnych trzech twierdzeniach zaprezentowana zostanie własność niezmienniczości kurtozy względem wybranych typów przekształceń afinicznych.

Twierdzenie 1. Kurtoza jako niezmiennik skali

Przypuśćmy, że $X = (X_1, \ldots, X_n)\colon \Omega \to \Re^n$ jest wektorem losowym, dla którego istnieje kurtoza, natomiast $a \in \Re$ jest dowolną liczbą rzeczywistą.
Wtedy:

$$\mathrm{Kurt}(aX) = \mathrm{Kurt}X. \tag{1}$$

Dowód. Poniższe przekształcenia uzasadniające równość (1) wynikają z definicji wariancji wektora losowego, definicji parzystych potęg wektora losowego oraz własności iloczynu skalarnego:

$$\mathrm{Kurt}(aX) = \frac{E\left[(aX - E(aX))^4\right]}{(D^2(aX))^2} = \frac{E\left(\langle aX - E(aX), aX - E(aX)\rangle^2\right)}{(E(\langle aX - E(aX), aX - E(aX)\rangle))^2} =$$

$$\frac{E\left(a^4 \langle X - EX, X - EX\rangle^2\right)}{(a^2(D^2X))^2} = \frac{a^4 E\left(\langle X - EX, X - EX\rangle^2\right)}{a^4 (D^2X)^2} = \frac{E\left[(X - EX)^4\right]}{(D^2X)^2} = \mathrm{Kurt}X.$$

Twierdzenie 2. Kurtoza jako niezmiennik translacji

Niech $X\colon \Omega \to R^n$ będzie wektorem losowym, dla którego istnieje kurtoza, oraz niech $b \in R^n$ będzie dowolnym wektorem z przestrzeni R^n.
Wówczas:

$$\mathrm{Kurt}(X + b) = \mathrm{Kurt}X. \tag{2}$$

Dowód. Równość (2) wynika wprost z definicji kurtozy wektora losowego oraz z faktu, że wariancja wektora losowego jest niezmiennikiem translacji [Tatar 1999]. Istotnie:

$$\mathrm{Kurt}(X + b) = \frac{E\left[(X + b - E(X + b))^4\right]}{(D^2(X + b))^2} = \frac{E\left[(X + b - E(X) - b)^4\right]}{(D^2X)^2} =$$

$$= \frac{E\left[(X + b - E(X) - b)^4\right]}{(D^2X)^2} = \frac{E\left[(X - E(X))^4\right]}{(D^2X)^2} = \mathrm{Kurt}X.$$

Przed sformułowaniem kolejnej własności kurtozy wektora losowego rozpatrzmy następujący lemat.

Lemat 1. Załóżmy, że $X:\Omega \to R^n$ jest wektorem losowym, dla którego istnieje moment centralny rzędu $2k$ oraz $C_{(n\times n)}$ jest macierzą ortogonalną (tzn. $C^T = C^{-1}$).

Wówczas:

$$E\left[(CX-E(CX))^{2k}\right]=E\left[(X-E(X))^{2k}\right],$$

czyli momenty centralne parzystego rzędu są niezmiennikami przekształceń ortogonalnych.

Dowód. Na początek zauważmy, że macierz CX jest postaci:

$$CX=\begin{bmatrix} c_{11} & \dots & c_{1n} \\ . & . & . \\ . & . & . \\ c_{n1} & \dots & c_{nn} \end{bmatrix}\begin{bmatrix} X_1 \\ . \\ . \\ X_n \end{bmatrix}=\begin{bmatrix} \sum\limits_{j=1}^{n} c_{1j}X_j \\ . \\ . \\ \sum\limits_{j=1}^{n} c_{nj}X_j \end{bmatrix}.$$

Dzięki własnościom momentów centralnych parzystego rzędu otrzymujemy:

$$E\left[(CX-E(CX))^{2k}\right]=E\left[\left((C(X-EX))^2\right)^k\right]=E\left[\left(\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}\left(X_j-E\left(X_j\right)\right)\right)^2\right)^k\right]=$$

$$=E\left[\sum_{i[1],\,\dots,\,i[k]=1}^{n}\left(\left(\sum_{s,\,t=1}^{n}c_{i[1]s}c_{i[1]t}\left(X_s-E\left(X_s\right)\right)\left(X_t-E\left(X_t\right)\right)\right)\cdot\ldots\cdot\left(\sum_{s,\,t=1}^{n}c_{i[k]s}c_{i[k]s}\left(X_s-E\left(X_s\right)\right)\left(X_t-E\left(X_t\right)\right)\right)\right)\right]=$$

$$=E\left[\sum_{i[1],\,\dots,i[k]=1}^{n}\left(\sum_{s[1],\,\dots,s[k],t[1],\,\dots,t[k]=1}^{n}c_{i[1]s[1]}c_{i[1]t[1]}\cdot\ldots\cdot c_{i[k]s[k]}c_{i[k]t[k]}\left(X_{s[1]}-E\left(X_{s[1]}\right)\right)\left(X_{t[1]}-E\left(X_{t[1]}\right)\right)\cdot\ldots.\right.\right.$$

$$\left.\left.\ldots.\cdot\left(X_{s[k]}-E\left(X_{s[k]}\right)\right)\left(X_{t[k]}-E\left(X_{t[k]}\right)\right)\right)\right]=$$

$$=E\left[\sum_{s[1],\,\dots,s[k],t[1],\,\dots,t[k]=1}^{n}\left(X_{s[1]}-E\left(X_{s[1]}\right)\right)\left(X_{t[1]}-E\left(X_{t[1]}\right)\right)\cdot\ldots\cdot\left(X_{s[k]}-E\left(X_{s[k]}\right)\right)\left(X_{t[k]}-E\left(X_{t[k]}\right)\right)\cdot\right.$$

$$\left.\cdot\left(\sum_{i[1],\,\dots,i[k]=1}^{n}c_{i[1]s[1]}c_{i[1]t[1]}\cdot\ldots\cdot c_{i[k]s[k]}c_{i[k]t[k]}\right)\right]=$$

$$=E\left[\sum_{s[1],\,\dots,s[k],t[1],\,\dots,t[k]=1}^{n}\left(X_{s[1]}-E\left(X_{s[1]}\right)\right)\left(X_{t[1]}-E\left(X_{t[1]}\right)\right)\cdot\ldots\cdot\left(X_{s[k]}-E\left(X_{s[k]}\right)\right)\left(X_{t[k]}-E\left(X_{t[k]}\right)\right)\cdot\right.$$

$$\left.\cdot\left(\sum_{i[1]=1}^{n}c_{i[1]s[1]}c_{i[1]t[1]}\right)\cdot\ldots\cdot\left(\sum_{i[k]=1}^{n}c_{i[k]s[k]}c_{i[k]t[k]}\right)\right]=$$

$$= E\left[\sum_{s[1],\ldots,s[k],t[1],\ldots,t[k]=1}^{n}\left(X_{s[1]}-E\left(X_{s[1]}\right)\right)\left(X_{t[1]}-E\left(X_{t[1]}\right)\right)\cdot\ldots\cdot\left(X_{s[k]}-E\left(X_{s[k]}\right)\right)\left(X_{t[k]}-E\left(X_{t[k]}\right)\right)\cdot\right.$$
$$\left.\cdot\left\langle c_{\bullet s[1]},c_{\bullet t[1]}\right\rangle\cdot\ldots\cdot\left\langle c_{\bullet s[k]},c_{\bullet t[k]}\right\rangle\right],$$

gdzie $c_{\bullet j}$ to j-ta kolumna macierzy CX.

Przypomnijmy, że kolumny (wiersze) macierzy ortogonalnej tworzą układ ortonormalny, a zatem:

$$E\left[(CX-E(CX))^{2k}\right]=$$
$$= E\left[\sum_{s[1],\ldots,s[k],t[1],\ldots,t[k]=1}^{n}\left(X_{s[1]}-E\left(X_{s[1]}\right)\right)\left(X_{t[1]}-E\left(X_{t[1]}\right)\right)\cdot\ldots\cdot\left(X_{s[k]}-E\left(X_{s[k]}\right)\right)\left(X_{t[k]}-E\left(X_{t[k]}\right)\right)\cdot\right.$$
$$\left.\cdot I_{s[1]t[1]}\cdot\ldots\cdot I_{s[k]t[k]}\right],\text{ gdzie } I_{st}=\begin{cases}1 & \text{dla} \quad s=t\\ 0 & \text{dla} \quad s\neq t\end{cases}$$

Wobec tego:

$$E\left[\left(CX-E(CX)\right)^{2k}\right]=E\left[\sum_{s[1],\ldots,s[k]=1}^{n}\left(X_{s[1]}-E\left(X_{s[1]}\right)\right)^{2}\cdot\ldots\cdot\left(X_{s[k]}-E\left(X_{s[k]}\right)\right)^{2}\right]=$$
$$=E\left[\left(\sum_{s[1]=1}^{n}\left(X_{s[1]}-E\left(X_{s[1]}\right)\right)^{2}\right)\cdot\ldots\cdot\left(\sum_{s[k]=1}^{n}\left(X_{s[k]}-E\left(X_{s[k]}\right)\right)^{2}\right)\right]=E\left[\left(\sum_{s=1}^{n}\left(X_{s}-E\left(X_{s}\right)\right)^{2}\right)^{k}\right]=$$
$$=E\left[(X-E(X))^{2k}\right].$$

Twierdzenie 3. Kurtoza jako niezmiennik przekształceń ortogonalnych

Jeżeli $X:\Omega\to R^{n}$ jest wektorem losowym, dla którego istnieje kurtoza, oraz $C_{(n\times n)}$ jest macierzą ortogonalną (tzn. $C^T = C^{-1}$) to:

$$\mathrm{Kurt}(CX) = \mathrm{Kurt}X.$$

Dowód. Twierdzenie 3 jest natychmiastowym wnioskiem z lematu 1. Istotnie z lematu 1 wynika, że momenty centralne czwartego oraz drugiego rzędu (wariancja) są niezmiennikami przekształceń ortogonalnych, a zatem kurtoza wektora losowego jako iloraz ich odpowiednich funkcji również spełnia tę własność.

W następnym rozdziale ustalone zostaną ograniczenia dolne dla wartości kurtozy.

3. Kurtoza wektora losowego – ograniczenia dolne

Twierdzenie 4. Niech $X:\Omega \to R^n$ będzie wektorem losowym, dla którego istnieje kurtoza $\beta_{2,n}$. Wówczas:

$$\beta_{2,n} \geq 1.$$

Dowód. Kluczową rolę w dowodzie odgrywać będzie nierówność Jensena. Przypomnijmy zatem jej postać:

$$g(E[\widetilde{X}]) \leq E[g(\widetilde{X})],$$

gdzie $\widetilde{X}$ to zmienna losowa, natomiast g to funkcja wypukła.

Zauważmy, że $g(x) = x^2$ jest funkcją wypukłą. Jeżeli w nierówności Jensena za $\widetilde{X}$ przyjmiemy zmienną losową $(X - m)^2$, to uzyskamy:

$$(E[(X-m)^2])^2 \leq E\left[((X-m)^2)^2\right],$$

czyli:

$$(D^2X)^2 \leq E[(X-m)^4].$$

Mamy więc $\dfrac{E[(X-m)^4]}{(D^2X)^2} \geq 1$, a zatem $\beta_{2,n} \geq 1$.

Uwaga 1. Ograniczenie dolne dla kurtozy wektora losowego w przypadku n-wymiarowym jest realizowane przez rozkład wektora losowego, będącego zestawieniem n niezależnych zmiennych losowych o tym samym rozkładzie dwupunktowym (zero-jedynkowym) z $p = \frac{1}{2}$, czyli przez szczególny przypadek wielowymiarowego rozkładu dwupunktowego.

Dowód. Istotnie dla wektora losowego $X = (X_1, \ldots, X_n): \Omega \to \Re^n$ spełniającego powyższe założenia kurtoza jest postaci:

$$\beta_{2,n} = \frac{\sum_{i=1}^{n} E[(X_i - m_i)^4] + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} E[(X_i - m_i)^2]E[(X_j - m_j)^2]}{(D^2X)^2} =$$

$$= \frac{n(D^2X_i)\mathrm{Kurt}X_i + n(n-1)(D^2X_i)^2}{n^2(D^2X_i)^2} = \frac{\mathrm{Kurt}X_i + n - 1}{n},$$

gdzie X_i to zmienna losowa o rozkładzie zero-jedynkowym z $p = \frac{1}{2}$, dla której kurtoza wynosi $\frac{(3p^2 - 3p + 1)}{p(1-p)}$, czyli $\text{Kurt}X_i = \frac{\left(3\left(\frac{1}{2}\right)^2 - 3\left(\frac{1}{2}\right) + 1\right)}{\frac{1}{2}\left(1 - \frac{1}{2}\right)} = 1$. Wobec tego $\beta_{2,n} = \frac{1 + n - 1}{n} = 1$.

Zauważmy ponadto, że pośród wektorów losowych będących zestawieniem n niezależnych zmiennych losowych o tym samym rozkładzie dwupunktowym (zero-jedynkowym) jedynie wektor z $p = \frac{1}{2}$, ma kurtozę równą 1.

Istotnie z poprzednich obliczeń wynika, iż:

$$\beta_{2,n} = \frac{\frac{(3p^2 - 3p + 1)}{p(1-p)} + n - 1}{n} = \frac{(3p^2 - 3p + 1) + (n-1)p(1-p)}{np(1-p)}. \tag{3}$$

Szukamy $p \in (0,1)$ spełniających równanie $\beta_{2,n} = 1$, a więc:

$$\frac{(3p^2 - 3p + 1) + (n-1)p(1-p)}{np(1-p)} = 1.$$

Otrzymujemy wobec tego równanie kwadratowe zmiennej p postaci: $4p^2 - 4p + 1 = 0$, którego jedynym rozwiązaniem jest $p = \frac{1}{2}$.

Uwaga 2. Nie ma górnego ograniczenia dla wartości kurtozy n-wymiarowego wektora losowego.

Dowód. Dla wektorów losowych będących zestawieniem n niezależnych zmiennych losowych o tym samym rozkładzie dwupunktowym (zero-jedynkowym) z postaci (3) wynika, że jeżeli $p \to 0^+$ lub $p \to 1^-$, to $\beta_{2,n} \to +\infty$.

Przed sformułowaniem kolejnego twierdzenia przypomnijmy następującą definicję.

Definicja 3. Współczynnikiem asymetrii wektora losowego $X\colon \Omega \to R^n$ nazywamy wektor postaci [Tatar 2000]:

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

Twierdzenie 5. Nierówność między kurtozą wektora losowego a kwadratem jego współczynnika asymetrii

Niech $X = (X_1, \dots, X_n) : \Omega \to \Re^n$ będzie wektorem losowym, dla którego istnieje kurtoza. Załóżmy, że wartość oczekiwana tego wektora to m, a macierz kowariancji ma postać $\sigma^2 I$, gdzie $\sigma > 0$. Wówczas prawdziwa jest nierówność:

$$\beta_{2,n} \geq 1 + n(\gamma_{1,n})^2. \tag{4}$$

Dowód. W pierwszej części dowodu wykażemy, że nierówność (4) jest prawdziwa dla wektorów losowych (spełniających założenia twierdzenia 5) o zerowym wektorze wartości oczekiwanych i jednostkowej macierzy kowariancji.

Niech zatem $X \sim d_n(0, I)$. Z tego faktu (w oczywisty sposób) wynikają następujące własności:
– $E(X_i) = 0$ dla każdego $i \in \{1, \dots, n\}$, (5)
– $D^2 X_i = 1$ dla każdego $i \in \{1, \dots, n\} \Rightarrow D^2 X = E(X^2) = n$, (6)
– $E(X_i X_j) = 0$ dla wszystkich $i, j \in \{1, \dots, n\}$ takich, że $i \neq j$. (7)

Zauważmy, że dla wszystkich $a_0 \in R, a_1 \in R^n$ oraz $a_2 \in R$ zachodzi nierówność:

$$E\left[(a_0 + \langle a_1, X \rangle + a_2)^2\right] \geq 0. \tag{8}$$

Nierówność (8) można natomiast przedstawić w postaci:

$$a_0^2 + \langle a_1, a_1 \rangle + 2a_0 a_2 n + 2a_2 E\left[\langle a_1, X \rangle X^2\right] + a_2^2 E[X^4] \geq 0. \tag{9}$$

Istotnie z własności całki uzyskujemy:

$$\begin{aligned} E\left[\left(a_0 + \langle a_1, X \rangle + a_2 X^2\right)^2\right] &= E\left[(a_0 + \langle a_1, X \rangle)^2 + 2(a_0 + \langle a_1, X \rangle) a_2 X^2 + (a_2 X^2)^2\right] = \\ = E\left[a_0^2 + 2a_0 \langle a_1, X \rangle + \langle a_1, X \rangle^2 + 2a_0 a_2 X^2 + 2\langle a_1, X \rangle a_2 X^2 + a_2^2 X^4\right] = \\ = a_0^2 + 2a_0 E[\langle a_1, X \rangle] + E\left[\langle a_1, X \rangle^2\right] + 2a_0 a_2 E[X^2] + 2a_2 E\left[\langle a_1, X \rangle X^2\right] + a_2^2 E[X^4]. \end{aligned}$$

Zwróćmy teraz uwagę na równości:

$$E[\langle a_1, X \rangle] = 0, \tag{10}$$

$$E\left[\langle a_1, X \rangle^2\right] = \langle a_1, a_1 \rangle. \tag{11}$$

Wynikają one bezpośrednio z własności (5)–(7).

Rzeczywiście $E[\langle a_1, X \rangle] = E\left(\sum_{i=1}^{n} a_i^1 X_i\right) = \sum_{i=1}^{n} a_i^1 E(X_i) \overset{(5)}{=} 0$,

natomiast:

$$E\left[\langle a_1, X \rangle^2\right] = E\left[\left(\sum_{i=1}^{n} a_i^1 X_i\right)^2\right] = E\left[\sum_{i,j=1}^{n} a_i^1 a_j^1 X_i X_j\right] = \sum_{i=1}^{n} (a_i^1)^2 E(X_i^2) + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} a_i^1 a_j^1 E(X_i X_j) \overset{(6),(7)}{=}$$

$$= \sum_{i=1}^{n} (a_i^1)^2 = \langle a_1, a_1 \rangle.$$

Dzięki równościom (10) i (11) otrzymujemy zatem postać (9).
Zauważmy również, że zachodzą następujące równości:

$$\left(E\left(X^3\right)\right)^2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} E\left(X_j^2 X_i\right)\right)^2 \tag{12}$$

oraz

$$E\left[\left\langle E\left(X^3\right), X\right\rangle X^2\right] = \left(E\left(X^3\right)\right)^2. \tag{13}$$

Istotnie:

$$\begin{aligned}
&\left(E\left(X^3\right)\right)^2 = \left\langle E\left(X^3\right), E\left(X^3\right)\right\rangle = \\
&= \left\langle \left(E\left(\sum_{j=1}^{n} X_j^2 X_1\right), \ldots, E\left(\sum_{j=1}^{n} X_j^2 X_n\right)\right), \left(E\left(\sum_{j=1}^{n} X_j^2 X_1\right), \ldots, E\left(\sum_{j=1}^{n} X_j^2 X_n\right)\right)\right\rangle = \\
&= \left\langle \left(\sum_{j=1}^{n} E\left(X_j^2 X_1\right), \ldots, \sum_{j=1}^{n} E\left(X_j^2 X_n\right)\right), \left(\sum_{j=1}^{n} E\left(X_j^2 X_1\right), \ldots, \sum_{j=1}^{n} E\left(X_j^2 X_n\right)\right)\right\rangle = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} E\left(X_j^2 X_i\right)\right)^2.
\end{aligned}$$

Do równości (13) natomiast prowadzą poniższe przekształcenia, w których korzystamy z własności (12):

$$\begin{aligned}
&E\left[\left\langle E\left(X^3\right), X\right\rangle X^2\right] = E\left[\left\langle \left(E\left(\sum_{j=1}^{n} X_j^2 X_1\right), \ldots, E\left(\sum_{j=1}^{n} X_j^2 X_n\right)\right), \left(X_1, \ldots, X_n\right)\right\rangle \left(\sum_{s=1}^{n} X_s^2\right)\right] = \\
&= E\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{n} E\left(X_j^2 X_i\right) X_i\right)\left(\sum_{s=1}^{n} X_s^2\right)\right] = E\left[\sum_{i,s=1}^{n}\left(\sum_{j=1}^{n} E\left(X_j^2 X_i\right) X_i\right) X_s^2\right] = \sum_{i,s=1}^{n}\left(\sum_{j=1}^{n} E\left(X_j^2 X_i\right)\right) E\left(X_i X_s^2\right) = \\
&= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} E\left(X_j^2 X_i\right)\right)\left(\sum_{s=1}^{n}\left(X_s^2 X_i\right)\right) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} E\left(X_j^2 X_i\right)\right)^2 \overset{(12)}{=} \left(E\left(X^3\right)\right)^2.
\end{aligned}$$

Nierówność (9) jest prawdziwa dla wszystkich $a_0 \in R, a_1 \in R^n$, oraz $a_2 \in R$, a więc w szczególności dla a_2 spełniającego równość $a_2^2 = \frac{1}{n^2}$ oraz a_0 takiego, że $a_0^2 + 2a_0 a_2 n = -1$.

Przyjmijmy, że $a_2 = \frac{1}{n}$. Wówczas funkcja $f(a_0) = a_0^2 + 2a_0 \frac{1}{n} n = a_0^2 + 2a_0$ osiąga minimum w punkcie $a_0 = -1$ równe $f_{\min}(-1) = -1$. Przyjmijmy zatem $a_0 = -1$. Zauważmy, że (przy powyższych a_0 i a_2) jeśli $a_1 = cE\left(X^3\right)$, gdzie $c \in R$, to na podstawie własności (13) nierówność (9) przybiera postać:

$$-1 + c^2 \left\langle E\left(X^3\right), E\left(X^3\right)\right\rangle + \frac{2c}{n}\left(E\left(X^3\right)\right)^2 + \frac{E\left[X^4\right]}{n^2} \geq 0.$$

Mamy zatem:

$$-1 + \left(\frac{c^2 n + 2c}{n}\right)\left(E\left(X^3\right)\right)^2 + \beta_{2,n} \geq 0.$$

Przy założeniu $X \sim d_n(0, I)$ otrzymujemy $(\gamma_{1,n})^2 = \frac{(E(X^3))^2}{n^3}$. Wówczas

$$-1 + (c^2 n^3 + 2cn^2)(\gamma_{1,n})^2 + \beta_{2,n} \geq 0.$$

Stałą $c \in R$ chcemy tak dobrać, aby $c^2 n^3 + 2cn^2 = -n$. Otrzymujemy więc równanie kwadratowe zmiennej c postaci $n^3 c^2 + 2n^2 c + n = 0$, w którym $\Delta = 0$, czyli posiadające (w zbiorze liczb rzeczywistych) jedyne rozwiązanie postaci $c_0 = \frac{-2n^2}{2n^3} = -\frac{1}{n}$. Za a_1 przyjmujemy zatem wektor postaci $-\frac{1}{n}E(X^3)$.

Podsumowując, dla $a_0 = -1$, $a_1 = -\frac{1}{n}E(X^3)$ oraz $a_2 = \frac{1}{n}$ nierówność (9) przyjmuje postać nierówności (4) (przy założeniu $X \sim d_n(0, I)$).

Weźmy teraz pod uwagę wektor losowy X o wektorze wartości oczekiwanych m oraz macierzy kowariancji postaci $\sigma^2 I$.

Zauważmy, że jeżeli $X \sim d_n(m, \sigma^2 I)$, to:

– $E(X_i - m_i) = 0$ dla każdego $i \in \{1, \ldots, n\}$, (14)

– $D^2 X_i = \sigma^2$ dla każdego $i \in \{1, \ldots, n\}$, (15)

– $E[(X_i - m_i)(X_j - m_j)] = 0$ dla wszystkich $i, j \in \{1, \ldots, n\}$ takich, że $i \neq j$. (16)

Rozważmy wektor losowy $\widetilde{X}$ postaci:

$$\widetilde{X} = \left(\frac{X_1 - m_1}{\sigma}, \ldots, \frac{X_n - m_n}{\sigma}\right).$$

Spełnia on własności (5)–(7).

Istotnie na podstawie własności (14):

$$E(\widetilde{X}_i) = E\left(\frac{X_i - m_i}{\sigma}\right) = \frac{1}{\sigma}E(X_i - m_i) \overset{(14)}{=} 0 \text{ dla każdego } i \in \{1, \ldots, n\}.$$

Dzięki własnościom wariancji zmiennej losowej oraz (15) otrzymujemy:

$$D^2 \widetilde{X}_i = D^2\left(\frac{X_i - m_i}{\sigma}\right) = \frac{1}{\sigma^2}D^2(X_i - m_i) = \frac{1}{\sigma^2}D^2 X_i \overset{(15)}{=} \frac{1}{\sigma^2}\sigma^2 = 1$$

$$\text{dla każdego } i \in \{1, \ldots, n\}.$$

W końcu (16) implikuje własność nieskorelowania zmiennych losowych $\widetilde{X}_i$ oraz $\widetilde{X}_j$ dla wszystkich $i, j \in \{1, \ldots, n\}$, takich że $i \neq j$. Istotnie:

$$E(\widetilde{X}_i \widetilde{X}_j) = E\left[\left(\frac{X_i - m_i}{\sigma}\right)\left(\frac{X_j - m_j}{\sigma}\right)\right] = \frac{1}{\sigma^2}E[(X_i - m_i)(X_j - m_j)] \overset{(16)}{=} 0.$$

Na podstawie powyższych ustaleń można stwierdzić, że wektor losowy $\widetilde{X}$ ma zerowy wektor wartości oczekiwanych oraz jednostkową macierz kowariancji, czyli $\widetilde{X} \sim d_n(0, I)$. Dla wektorów losowych o takich rozkładach nierówność (4) została już wykazana, a więc:

$$\operatorname{Kurt}\widetilde{X} \geq 1 + \left(\operatorname{Skew}\widetilde{X}\right)^2 .$$

Zauważmy również, że $\operatorname{Kurt}\widetilde{X} = \operatorname{Kurt}X$ oraz $\left(\operatorname{Skew}\widetilde{X}\right)^2 = \left(\operatorname{Skew}X\right)^2$. Istotnie:

$$\operatorname{Kurt}\widetilde{X} = \frac{E\left[\left(\widetilde{X} - E\widetilde{X}\right)^4\right]}{\left(D^2\widetilde{X}\right)^2} = \frac{E\left[\widetilde{X}^4\right]}{n^2} = \frac{\sum\limits_{i,\,j=1}^{n} E\left[\widetilde{X}_i^2 \widetilde{X}_j^2\right]}{n^2} = \frac{\sum\limits_{i,\,j=1}^{n} E\left[\left(\frac{X_i - m_i}{\sigma}\right)^2 \left(\frac{X_j - m_j}{\sigma}\right)^2\right]}{n^2} =$$

$$= \frac{\sum\limits_{i,\,j=1}^{n} E\left[(X_i - m_i)^2 (X_j - m_j)^2\right]}{n^2\sigma^4} = \frac{E\left[(X - m)^4\right]}{\left(n\sigma^2\right)^2} = \frac{E\left[(X - m)^4\right]}{\left(D^2X\right)^2} = \operatorname{Kurt}X.$$

Równość dotyczącą kwadratów współczynników asymetrii uzasadniają natomiast następujące przekształcenia:

$$\left(\operatorname{Skew}\widetilde{X}\right)^2 = \frac{\left(E\left[\left(\widetilde{X} - E\widetilde{X}\right)^3\right]\right)^2}{\left(\left(D^2\widetilde{X}\right)^{\frac{3}{2}}\right)^2} = \frac{\left(E\left[\widetilde{X}^3\right]\right)^2}{n^3} =$$

$$= \frac{\left\langle \left(E\left(\sum\limits_{i=1}^{n} \widetilde{X}_i^2 \widetilde{X}_1\right), \ldots, E\left(\sum\limits_{i=1}^{n} \widetilde{X}_i^2 \widetilde{X}_n\right)\right), \left(E\left(\sum\limits_{i=1}^{n} \widetilde{X}_i^2 \widetilde{X}_1\right), \ldots, E\left(\sum\limits_{i=1}^{n} \widetilde{X}_i^2 \widetilde{X}_n\right)\right)\right\rangle}{n^3} =$$

$$= \frac{\left(E\left(\sum\limits_{i=1}^{n} \left(\frac{X_i - m_i}{\sigma}\right)^2 \left(\frac{X_1 - m_1}{\sigma}\right)\right), \ldots, E\left(\sum\limits_{i=1}^{n} \left(\frac{X_i - m_i}{\sigma}\right)^2 \left(\frac{X_n - m_n}{\sigma}\right)\right)\right)^2}{n^3} =$$

$$= \frac{\frac{1}{\sigma^6}\left(E\left(\sum\limits_{i=1}^{n} (X_i - m_i)^2 (X_1 - m_1)\right), \ldots, E\left(\sum\limits_{i=1}^{n} (X_i - m_i)^2 (X_1 - m_1)\right)\right)^2}{n^3} =$$

$$= \frac{\left(E\left[X^3\right]\right)^2}{n^3\sigma^6} = \frac{\left(E\left[X^3\right]\right)^2}{\left(n\sigma^2\right)^3} = \frac{\left(E\left[X^3\right]\right)^2}{\left(D^2X\right)^3} = \left(\operatorname{Skew}X\right)^2 .$$

Nierówność (4) została udowodniona.

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”, nr 76.

Budny K., Tatar J. [2009], *Kurtosis of a Random Vector – Special Types of Distributions*, „Statistics in Transition – New Series”, vol. 10, nr 3.

Cramer H. [1958], *Metody matematyczne w statystyce*, PWN, Warszawa.

Jakubowski J., Sztencel R. [2004], *Wstęp do rachunku prawdopodobieństwa*, Script, Warszawa.

Kotz S., Johnson N.L., Balakrishnan N. [2000], *Continuous Multivariate Distributions: Model and Applications*, 2nd ed., John Wiley and Sons, New York.

Mardia K.V. [1970], *Measures of Multivariate Skewness and Kurtosis with Applications*, „Biometrika", vol. 57, nr 3.

Tatar J. [1996], *O niektórych miarach rozproszenia rozkładów prawdopodobieństwa*, „Przegląd Statystyczny", z. 3/4.

Tatar J. [1999], *Moments of a Random Variable in a Hilbert Space*, „Przegląd Statystyczny", z. 2.

Tatar J. [2000], *Asymetria wielowymiarowych rozkładów prawdopodobieństwa*, Materiały z XXXV Konferencji Statystyków, Ekonometryków i Matematyków Akademii Ekonomicznych Polski Południowej zorganizowanej przez Katedrę Statystyki Akademii Ekonomicznej w Krakowie (Osieczany, 23–25 marca 1999 r.), Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

Some Properties of the Kurtosis of a Random Vector

The paper is a continuation of a discussion of the kurtosis of a random vector. Multivariate kurtosis is defined as the fourth central moment divided by the square of the variance of a random vector. This term is built on the definition of the power of a random vector proposed by J. Tatar. The paper presents selected, essential properties of multivariate kurtosis – among other things the invariance property under a number of affine transformations. Besides that, the relation between kurtosis of the random vector and its skewness is fixed. In view of these properties, the fourth central moment divided by the square of the variance of a random vector may be regarded as a satisfactory measure of multivariate kurtosis.

Keywords: kurtosis of a random vector, power of a vector, multivariate distribution, characteristics of the multivariate distribution.