ACTA UNIVERSITATIS NICOLAI COPERNICI
EKONOMIA XLI – NAUKI HUMANISTYCZNO-SPOŁECZNE – ZESZYT 397 – TORUŃ 2010

Uniwersytet Warmińsko-Mazurski w Olsztynie
Katedra Metod Ilościowych

Ewa Wędrowska

OCZEKIWANA ILOŚĆ INFORMACJI O ZMIANIE STRUKTUR JAKO MIARA NIEPODOBIEŃSTWA STRUKTUR

Zarys treści. W artykule skupiono uwagę na wskazaniu możliwości wykorzystania metod teorii informacji do badania zjawisk społeczno-ekonomicznych. W szczególności przedstawiono możliwość wykorzystania entropii rzeczywistej oraz entropii trygonometrycznej do badania własności rozkładów struktur. Zaproponowano również wykorzystanie oczekiwanej ilości informacji o zmianie założonej struktury w strukturę zrealizowaną do badania stopnia niepodobieństwa pomiędzy strukturami.

Słowa kluczowe: entropia rzeczywista, *Entropies of Mixing*, wartość oczekiwana informacji, podobieństwo.

WSTĘP

Pojęcie entropii zostało zdefiniowane przy okazji drugiej zasady termodynamiki, natomiast zastosowanie termodynamiki w teorii informacji wprowadziło pojęcie entropii do systemów komunikowania się. O entropii można mówić wszędzie tam, gdzie istnieje potrzeba zilustrowania różnorodności i losowości, a więc w szczególności w teorii informacji i statystyce. Często różnorodność traktowana jest jako niedogodność, w teorii informacji jednak różnorodność jest źródłem informacji. Dążność do ustalenia stopnia tej różnorodności opiera się na wykorzystaniu entropii.

Obecnie pojawiają się zastosowania pojęcia entropii w analizie zachowania się różnorodnych systemów, w tym systemów ekonomiczno-społecznych.

Poznanie reguł decydujących o ewolucji struktury zjawisk ekonomicznych jest ważnym problemem w osiąganiu celów teoretycznych i praktycznych. Porównywanie struktur, analiza ich zmian w czasie jest przedmiotem wielu badań w zagadnieniach społeczno-ekonomicznych. Zmianom, a zarazem różnicom struktur towarzyszy różnorodność, do badania której możliwe jest wykorzystanie entropii.

W zagadnieniach społeczno-ekonomicznych występuje konieczność oceny stopnia zmian pomiędzy założoną strukturą (*a priori*) a strukturą, która wystąpiła w rzeczywistości (*a posteriori*). Dotyczyć to może chociażby porównania założonej struktury wydatków pewnych obiektów gospodarczych ze strukturą, która została zrealizowana w rzeczywistości. Celem artykułu jest wskazanie możliwości oceny stopnia tych zmian. Zaproponowano wykorzystanie oczekiwanej ilości informacji o zmianie założonej struktury w strukturę zrealizowaną do badania stopnia niepodobieństwa pomiędzy strukturami. Oczekiwana ilość informacji o zmianie apriorycznej struktury w strukturę aposterioryczną znana jest w literaturze jako miara Kullbacka-Leiblera. W artykule przedstawiono również możliwość wykorzystania entropii rzeczywistej struktur oraz trygonometrycznej entropii do badania własności struktur.

1. ENTROPIA EMPIRYCZNA STRUKTUR

W zagadnieniach społeczno-gospodarczych często przedmiotem zainteresowań są zmiany struktur charakteryzujących obiekty. W literaturze z zakresu taksonomii pojęcie struktury występuje w dwojakim znaczeniu (Strahl, 1998):

1. Struktura jako specyficzny obiekt S^n, scharakteryzowany nie wartościami badanej cechy, lecz ciągiem *n* części składowych $\alpha_1, \alpha_2, \ldots, \alpha_n$, będących wskaźnikami struktury bądź też wskaźnikami udziału.
2. Struktura jako zbiór punktów w przestrzeni wielowymiarowej. Tak rozumiana struktura nie jest obiektem, ale zbiorem punktów scharakteryzowanych różnymi cechami.

Wskaźniki struktury oraz wskaźniki udziału będące odpowiednio składnikami struktury S^n spełniają następujące warunki:

1. Normowalność: $0 \le \alpha_i \le 1$ ($i = 1, \ldots, n$),
2. Warunek sumy jednostkowej: $\sum_{i=1}^{n} \alpha_i = 1$ ($i = 1, \ldots, n$).

W niniejszym artykule, zgodnie ze Strahl (1998), struktura będzie interpretowana jako obiekt opisany wektorem wskaźników struktury (lub udziału). Wyznaczanie wektora S^n jest zasadne tylko wtedy, gdy cecha X podlegająca badaniu spełnia własność addytywności. W szczególności w analizach ekonomicznych klasę najczęściej spotykanych struktur stanowią addytywne struktury ekonomiczne, dla których suma wartości przyporządkowanych określonym elementom procesu gospodarczego ma sens ekonomiczny (Kukuła, 2000).

Badanie zmian oraz różnic pomiędzy strukturami może mieć charakter statyczny lub dynamiczny, zatem analiza dotyczyć może różnic pomiędzy strukturami w przestrzeni *n*-wymiarowej bądź też badana jest zmienność struktur w czasie. Rozważmy strukturę S^n scharakteryzowaną przez wektor $S^n = [\alpha_1, \alpha_2, ..., \alpha_n]$ należący do przestrzeni euklidesowej $\Re^n$, dla którego składniki α_i $(i = 1, 2, ..., n)$ spełniają warunki 1 i 2. Zbiór $\Re^n = \Re \times \Re \times ... \times \Re$ jest przestrzenią wektorową nad ciałem liczb rzeczywistych $\Re$, jeżeli określone są w nim dodawanie wektorów i mnożenie wektora przez liczbę, tzn.

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \in \Re^n,$$

$$\lambda\mathbf{x} = (\lambda x_1, \lambda x_2, ..., \lambda x_n) \in \Re^n,$$

gdzie $\mathbf{x} = (x_1, x_2, ..., x_n) \in \Re^n$, $\mathbf{y} = (y_1, y_2, ..., y_n) \in \Re^n$, $\lambda \in \Re$.

Przestrzeń $\Re^n$, w której zdefiniowano iloczyn skalarny wektorów według wzoru:

$$\mathbf{x} \circ \mathbf{y} = (x_1 \cdot y_1, x_2 \cdot y_2, ..., x_n \cdot y_n) \in \Re,$$

jest przestrzenią Euklidesa. Jeżeli ponadto każdemu wektorowi przestrzeni Euklidesa przyporządkowana zostanie liczba $\|\mathbf{x}\|_E = (\sum_{i=1}^{n} x_i^2)^{0,5}$, to przestrzeń $\Re^n$ jest unormowaną przestrzenią Euklidesa, a liczba $\|\mathbf{x}\|_E$ jest normą wektora **x**.

Dysponując pełnymi danymi o składnikach struktury S^n, można wyznaczyć entropię rzeczywistą (empiryczną), zdefiniowaną przez Shannona, struktury S^n, przyjmując za podstawę logarytmu liczbę 2:

$$H(S^n) = -\sum_{i=1}^{n} \alpha_i \log_2 \alpha_i = \sum_{i=1}^{n} \alpha_i \log_2 \frac{1}{\alpha_i}. \quad (1)$$

Entropia $H(S^n)$ spełnia fundamentalne aksjomaty entropii (Przybyszewski, Wędrowska, 2005). Wartość entropii $H(S^n)$ zależy wyłącznie od częstości występowania *i*-tego wariantu cechy X w strukturze S^n, a więc od wskaźników struktury (lub udziału) charakteryzujących badany obiekt O_j. O wartości entropii stanowią wskaźniki struktury α_i, a nie wartości cechy skojarzone z tymi wskaźnikami. Wartość entropii $H(S^n)$ jest wyznacznikiem niepewności i koncentracji rozkładu składników α_i struktury S^n.

Entropia $H(S^n)$ dana wzorem (1) ma następujące własności:

i. jest wielkością nieujemną, $\forall \alpha_i \in [0,1]$ $H(S^n) \geq 0$;
ii. przyjmuje wartość 0, gdy $\alpha_i = 1$ dla pewnego i $(i = 1, 2, ..., n)$;

iii. przyjmuje wartość największą, gdy wszystkie składniki struktury α_i są sobie równe dla $i = 1, 2, \ldots, n$: $H_{\max}(S^n) = \log_2 n$, jeżeli $\alpha_1 = \alpha_2 = \ldots = \alpha_n$;

iv. spełnia własność symetrii: $H(\alpha_1, \alpha_2, \ldots, \alpha_n) = H(\alpha_{(1)}, \alpha_{(2)}, \ldots, \alpha_{(n)})$;

v. spełnia własność wklęsłości: $\forall \alpha_i \in [0, 1] \quad \frac{\delta^2}{\delta \alpha_i^2} H(S^n) \le 0$.

Własności entropii rozkładów i wynikająca z nich możliwość zastosowania entropii do badania struktur powoduje, iż poszukiwane są inne funkcje spełniające aksjomaty entropii (Zoido, Carreno, 2000). Interesującą propozycję trygonometrycznego ujęcia entropii przedstawiono w pracy Lavenda (2006). *Entropies of Mixing* (EOM) zdefiniowano następująco:

$$H^{EOM}(S^n) = \frac{1}{n} \sum_{i=1}^{n} \sin \pi \alpha_i . \tag{2}$$

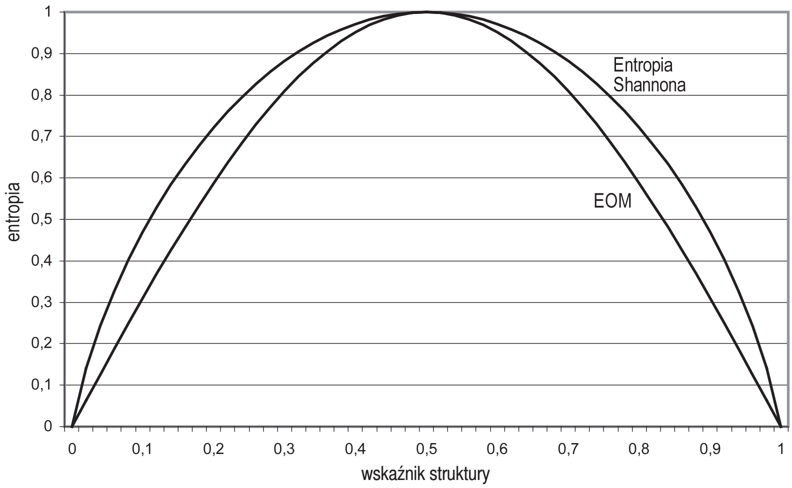
Trygonometryczna entropia $H^{EOM}(S^n)$ spełnia aksjomaty i własności entropii (Lavenda, 2006), zatem spełnia własności (i)–(v).

Znajomość entropii $H(S^n)$ lub $H^{EOM}(S^n)$ pozwala na rozpoznanie koncentracji rozkładu składników struktury S^n. Zgodnie z własnością (iii) entropia $H(S^n)$ osiągnie maksimum dla struktury S_j^n o wskaźnikach α_{ji}, takich, że $\alpha_1 = \alpha_2 = \ldots = \alpha_n$. Maksimum to wynosi $\log_2 n = H_{\max}$. Jeżeli entropia struktury $H(S^n)$ równa jest entropii maksymalnej $H_{\max}$, występuje rozkład równomierny składników struktury. W przypadku EOM wartość maksymalna entropii wynosi $H_{\max}^{EOM} = \sin \frac{\pi}{n}$. Entropia Shannona i trygonometryczna entropia EOM przyjmują wartość zero w przypadku, gdy jeden ze składników równy jest 1, pozostałe natomiast są zerami. Towarzyszy to sytuacji, w której występuje koncentracja rozkładu w jednym, dowolnym składniku struktury. Własność symetrii, jaką ma entropia Shannona oraz EOM, powoduje, że składnik $\alpha_i = 1 \, (i = 1, 2, \ldots, n)$ może być dowolną i-tą współrzędną struktury S^n. Wartości entropii $H(S^n)$ oraz $H^{EOM}(S^n)$ dla struktury S^n zawierają się zatem odpowiednio w przedziałach:

$$H(S^n) \in [0, \log_2 n];$$

$$H^{EOM}(S^n) \in \left[0, \sin \frac{\pi}{n}\right].$$

Entropia EOM wykazuje mniejszą wrażliwość na odbieganie rozkładu wskaźników struktury od rozkładu równomiernego. Dla struktury dwuelementowej $[\alpha, 1-\alpha]$ różnice w wartościach entropii Shannona i EOM ilustruje rysunek 1.



Rysunek 1. Wartości entropii Shannona i entropii EOM struktur dwuelementowych

Źródło: opracowanie własne.

Dzięki znajomości entropii $H(S^n)$ lub $H^{EOM}(S^n)$ możliwe jest rozpoznanie koncentracji rozkładu składników struktury S^n poprzez konstrukcję miar koncentracji będących funkcją entropii (Wędrowska, 2003).

Wśród miar koncentracji odrębną klasę stanowią miary wykorzystujące właśnie formułę entropii. Związek entropii z miarami koncentracji wydaje się bowiem oczywisty, gdyż entropia jest miarą nieokreśloności i niejednorodności rozkładu elementów struktury S^n. Mierniki te stanowią miarę koncentracji, tak jak miara

$$h = 1 - \frac{H(S^n)}{H_{\max}(S^n)} \tag{3}$$

proponowana przez Roeske-Słomka (1994), bądź też miarę dekoncentracji postaci:

$$e = \frac{H(S^n)}{H_{\max}(S^n)} \tag{4}$$

przedstawioną między innymi w pracach Wędrowska, 2003; McBratney, Minasny, 2007. Wartości obu miar są unormowane w przedziale [0, 1], przy czym miara (3) ma charakter miary koncentracji, co oznacza, że występowanie coraz większych wartości mierników towarzyszy silniejszemu zjawisku skupienia wybranych wartości cechy w wąskiej grupie jednostek.

2. STRUKTURA *A PRIORI* ORAZ *A POSTERIORI*

W badaniu zmian zachodzących w strukturach charakteryzujących pewne obiekty czy zjawiska istotny jest stopień zachodzących zmian. Rozważona zostanie sytuacja, w której zakładane jest wystąpienie struktury S_p^n (np. struktura planowanych wydatków). Struktura S_p^n nazywana będzie strukturą *a priori*. Interesujące jest zbadanie stopnia zmian pomiędzy założoną strukturą *a priori* S_p^n a zrealizowaną w rzeczywistości, oznaczoną S_q^n. Struktura S_q^n nazywana będzie strukturą *a posteriori*.

W pracy Theil (1979) rozważana jest oczekiwana ilość informacji niesionej w wiadomości o transformacji prawdopodobieństw p_i ($i = 1, 2, ..., n$) w prawdopodobieństwa q_i dla n wzajemnie wykluczających się zdarzeń $E_1, E_2, \dots, E_n$. Wykorzystując tę koncepcję w analizie struktur, można wyznaczyć ilość informacji o zmianie apriorycznej struktury S_p^n w aposterioryczną S_q^n. Niech struktura *a priori* S_p^n będzie wyrażona wektorem wskaźników struktury (lub udziału) $[p_1, p_2, \dots, p_n]$, a struktura *a posteriori* S_q^n wektorem wskaźników $[q_1, q_2, \dots, q_n]$, przy czym wskaźniki p_i oraz q_i spełniają warunki unormowania oraz sumy jednostkowej $i = 1, 2, ..., n$. Oczekiwana ilość informacji o zmianie struktury *a priori* S_p^n w strukturę *a posteriori* S_q^n dana jest wzorem (Theil, 1979):

$$I(S_q^n : S_p^n) = \sum_{i=1}^{n} q_i \log_2 \frac{q_i}{p_i}. \tag{5}$$

Wartość oczekiwana informacji dana formułą (5) spełnia następujące własności:

a) $I(S_q^n : S_p^n) = 0$, jeżeli $\forall\, i = 1, 2, ..., n \quad p_i = q_i$;

b) $q_i \log_2 \frac{q_i}{p_i} > 0$ dla $q_i > p_i$;

c) $q_i \log_2 \frac{q_i}{p_i} < 0$ dla $q_i < p_i$;

d) $I(S_q^n : S_p^n) > 0$, jeżeli $S_p^n \neq S_q^n$.

Wielkość $I(S_q^n : S_p^n)$ informuje o stopniu zmian pomiędzy założoną strukturą *a priori* S_p^n a zrealizowaną strukturą *a posteriori* S_q^n, a więc o stopniu podobieństwa bądź niepodobieństwa struktur S_p^n oraz S_q^n. Zatem znajomość oczekiwanej ilości informacji o zmianie apriorycznej struktury S_p^n w aposterioryczną strukturę S_q^n może być przydatna w ocenie podobieństwa struktur. Zgodnie z własnością (a) wartość oczekiwana informacji $I(S_q^n : S_p^n)$ przyjmuje wartość równą zero dla dwóch identycznych struktur $S_p^n = S_q^n$, to znaczy struktur, dla których każdy odpowiedni wskaźnik $p_i = q_i$ dla każdego $i = 1, 2, ..., n$. Wraz z występowaniem coraz większych różnic pomiędzy strukturami S_p^n oraz S_q^n wartość oczekiwana informacji $I(S_q^n : S_p^n)$ jest dodatnia (własność (d)) i rośnie do nieskończoności.

Oczekiwana ilość informacji jest niesymetryczna dla struktur $S_p^n \neq S_q^n$, dlatego miary tej nie powinno traktować się jako odległości pomiędzy strukturami, lecz jako rozbieżność przy założeniu, że jedna ze struktur stanowi strukturę bazową. Formuła (5) pozwala na intuicyjne pojmowanie wartości oczekiwanej informacji jako „koszt" rozpoznania nieokreśloności rozkładu struktury S_p^n przy znajomości nieokreśloności rozkładu struktury S_q^n.

Miary wykorzystywane w taksonomii struktur do badania zgodności struktur są najczęściej miarami stanowiącymi funkcję odległości wskaźników struktur, przez co są miarami symetrycznymi. Istnieją jednak nieliczne podejścia, które uwzględniają miary intensywności zmian strukturalnych niespełniające warunku symetrii. Przykładem jest między innymi propozycja Rutkowskiego (1981). Do tych propozycji można dołączyć wartość oczekiwaną informacji $I(S_q^n : S_p^n)$. Przydatność niesymetrycznej miary rozbieżności struktur obiektów występuje w sytuacjach, gdy zakłada się, że transformacja struktury S_p^n w strukturę S_q^n nie jest równoznaczna z transformacją struktury S_q^n w strukturę S_p^n. Dotyczy to w szczególności badania zmian strukturalnych w ujęciu dynamicznym.
Ponadto, jeżeli różnice $(p_i - q_i)$ są bliskie 0, dla $i = 1, 2, ..., n$, to stosując we wzorze (5) logarytm naturalny, oczekiwaną ilość informacji $I(S_q^n : S_p^n)$ można aproksymować wyrażeniem $\frac{1}{2}\sum_{i=1}^{n} q_i \cdot \left(\frac{p_i - q_i}{q_i}\right)^2$.

Zachodzi zatem:

$$I(S_q^n : S_p^n) \approx \frac{1}{2}\sum_{i=1}^{n} q_i \cdot \left(\frac{p_i - q_i}{q_i}\right)^2 = \frac{1}{2}\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{q_i}. \tag{6}$$

Własność tę podano w pracy Theil (1979), w niniejszym artykule wykazana zostanie prawdziwość tej własności. Aby wykazać zatem prawdziwość wyrażenia (6), należy rozwinąć funkcję $\ln\frac{q_i}{p_i}$ w szereg potęgowy. Stąd należy funkcję $\ln\frac{q_i}{p_i}$ zapisać w takiej postaci, dla której rozwinięcie jest oczywiste:

$$\ln\left(\frac{q_i}{p_i}\right) = -\ln\left(1 + \frac{p_i - q_i}{q_i}\right), \tag{7}$$

a więc

$$I(S_q^n : S_p^n) = -\sum_i q_i \cdot \ln\left(1 + \frac{p_i - q_i}{q_i}\right). \tag{8}$$

W równości (7) prawą stronę można rozwinąć w szereg potęgowy zgodnie ze wzorem:

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \ldots \ . \tag{9}$$

Rozwijanie w szeregi potęgowe stanowi prostą metodę aproksymacji funkcji przez wielomiany, a więc funkcje możliwie najdogodniejsze do obliczeń i przekształceń. Wobec tego funkcję $\ln(1+x)$ można zastąpić przez sumę częściową $S_n(x)$ w przedziale zbieżności. Można zatem przyjąć nierówność:

$$x - \frac{1}{2}x^2 < \ln(1+x) < x, \qquad x > 0\,. \tag{10}$$

Jest to osadzenie funkcji $\ln(1+x)$ pomiędzy pierwszą i drugą sumą częściową jej szeregu potęgowego (9). Uwzględniając (9) oraz (10), można przyjąć dla wyrażenia $\ln\left(1+\frac{p_i - q_i}{q_i}\right)$ aproksymację składnikiem kwadratowym $-\frac{1}{2}\left(\frac{p_i - q_i}{q_i}\right)^2$, dla różnicy $(p_i - q_i)$ bliskiej 0.

Stąd zachodzi rzeczywiście:

$$-\sum_{i=1}^{n} q_i \cdot \ln\left(1+\frac{p_i - q_i}{q_i}\right) \approx \frac{1}{2}\sum_{i=1}^{n} q_i \cdot \left(\frac{p_i - q_i}{q_i}\right)^2 = \frac{1}{2}\sum_{i=1}^{n}\frac{(p_i - q_i)^2}{q_i}. \tag{11}$$

Z prawdziwości wyrażenia (6) wynika, że wartość oczekiwana informacji równa jest połowie ważonej średniej kwadratów względnych odchyleń między odpowiadającymi sobie prawdopodobieństwami *a priori* oraz *a posteriori*, z wagami równymi prawdopodobieństwom *a posteriori*. Z wyrażenia (6) widać również, że $I(S_q^n : S_p^n)$ jest w przybliżeniu proporcjonalne do statystyki χ^2 z q_i jako prawdopodobieństwami teoretycznymi oraz p_i jako obserwowanymi częstościami, przy założeniu, że różnica $(p_i - q_i)$ bliska jest 0.

3. ILUSTRACJA EMPIRYCZNA

Metodologia porównywania struktur obejmuje różnorodne metody ustalania stopnia ich odległości, a mierniki podobieństwa struktur zazwyczaj stanowią funkcję miar odległości ich wskaźników cząstkowych (Młodak, 2006). W niniejszym artykule do zbadania braku podobieństwa pomiędzy dwiema strukturami S_p^n oraz S_q^n wykorzystana zostanie formuła zaproponowana przez Chomątowskiego i Sokołowskiego (1978):

$$P(S_q^n, S_p^n) = 1 - \sum_{i=1}^{n} \min(q_i, p_i). \tag{12}$$

Wielkość $P(S_q^n, S_p^n)$ unormowana jest w przedziale [0, 1], przy czym $P(S_q^n, S_p^n)$ równa jest zero dla identycznych struktur $S_p^n = S_q^n$.

W celu zilustrowania możliwości wykorzystania oczekiwanej ilości informacji o zmianie apriorycznej struktury S_p^n w aposterioryczną strukturę S_q^n do badania niepodobieństwa struktur zbadano struktury czteroelementowe S_p^4 oraz S_q^4 (tabela 1). W analizie rozbieżności struktur przyjęto dane umowne, wybór arbitralny wskaźników struktur nie ma jednak wpływu na otrzymane wyniki, a pozwala zilustrować własności wartości oczekiwanej informacji $I(S_q^n : S_p^n)$.

Wzrostowi wartości miary niepodobieństwa struktur $P(S_q^n, S_p^n)$ towarzyszy wzrost wartości oczekiwanej ilości informacji $I(S_q^n : S_p^n)$. Dla struktur całkowicie różnych $P(S_q^n, S_p^n)$ przyjmuje wartość 1, natomiast wartości miary $I(S_q^n : S_p^n)$ dążą wtedy do nieskończoności. W przypadku struktur identycznych obie miary przyjmują wartość zero.

Tabela 1. Wartości entropii rzeczywistej, *Entropies of Mixing*, oczekiwanej ilości informacji oraz miary niepodobieństwa struktur

Struktury S_q^2 i S_p^2	Entropia $H(S^n)$	Wsp. koncentracji h	EOM $H^{EOM}(S^n)$	$I(S_q^n : S_p^n)$	$P(S_q^n, S_p^n)$
S_q^2 = [0,4;0,2;0,1;0,3]	1,84644	0,077	0,66422	0	0
S_p^2 = [0,4;0,2;0,1;0,3]	1,84644	0,077	0,66422		
S_q^2 = [0,4;0,3;0,2;0,1]	1,84644	0,077	0,66422	0,134	0,20
S_p^2 = [0,6;0,2;0,1;0,1]	1,57095	0,215	0,53922		
S_q^2 = [0,1;0,2;0,3;0,4]	1,84644	0,007	0,66422	1,192	0,50
S_p^2 = [0,6;0,2;0,1;0,1]	1,57095	0,215	0,53922		
S_q^2 = [0,1;0,2;0,6;0,1]	1,57095	0,215	0,53922	1,292	0,50
S_p^2 = [0,6;0,2;0,1;0,1]	1,57095	0,215	0,53922		
S_q^2 = [0,25;0,25;0,25;0,25]	2,00000	0	0,70711	1,382	0,65
S_p^2 = [0,02;0,03;0,05;0,9]	0,61754	0,692	0,15559		
S_q^2 = [0,05;0,4;0,45;0,1]	1,59546	0,202	0,60105	1,734	0,55
S_p^2 = [0,6;0,2;0,1;0,1]	1,57095	0,215	0,53922		
S_q^2 = [0,7;0,2;0,05;0,05]	1,25678	0,372	0,42742	3,568	0,85
S_p^2 = [0,02;0,03;0,05;0,9]	0,61754	0,691	0,15559		
S_q^2 = [0,9;0,05;0,03;0,02]	0,61754	0,691	0,15559	4,848	0,90
S_p^2 = [0,02;0,03;0,05;0,9]	0,61754	0,691	0,15559		

Źródło: obliczenia własne.

Dodatkowa znajomość entropii rzeczywistej struktur (lub entropii EOM) pozwala na ocenę stopnia koncentracji struktur. Entropia rzeczywista oraz *Entropies of Mixing* spełniają własność symetrii, czyli dla różnych struktur, dla których współczynniki struktury są takie same, lecz stanowią inne współrzędne wektorów struktur, wartości entropii są takie same. W przypadku równomiernego podziału wariantów cechy, a więc struktur o równych sobie wskaźnikach, entropia przyjmuje wartość maksymalną. Dla entropii rzeczywistej struktur czteroelementowych jest to wartość $\log_2 n = \log_2 4 = 2$, wartość maksymalna entropii EOM jest równa $\sin\frac{\pi}{n} = \sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}$.

Znajomość miary koncentracji danej wzorem (3), bazującej na formule entropii, umożliwia rozpoznanie stopnia skupienia wartości cechy wokół wyróżnionych jednostek. W tabeli 1 zamieszczono wartości współczynnika koncentracji danego formułą (3). Z konstrukcji miary (3) wynika, że w pomiarze stopnia koncentracji rozkładu wskaźników struktury największy wpływ na otrzymane wyniki mają wartości cechy występujące z dużą częstością. Powoduje to, iż znaczna część obszaru zmienności miary (3) przeznaczona jest dla rozkładów bliskich pełnej koncentracji, a miernik charakteryzuje się niewielką awersją do nierównomiernego podziału łącznego funduszu cechy pomiędzy jednostki badania.

PODSUMOWANIE

Szczególnym przypadkiem analiz zjawisk społeczno-ekonomicznych jest badanie podobieństwa bądź niepodobieństwa struktur charakteryzujących zmieniające się zjawiska gospodarcze. Zmiany te mogą mieć zarówno charakter statyczny, jak i dynamiczny. W artykule przedstawiono możliwość wykorzystania oczekiwanej ilości informacji o zmianie struktury *a priori* w strukturę *a posteriori* do rozpoznania stopnia zmian pomiędzy badanymi strukturami. Wartość oczekiwana informacji jest niesymetryczna i nieograniczona. Dlatego też zasadność wykorzystania tej miary uwidacznia się w sytuacjach, gdy zakłada się, że transformacja struktury S_p^n w strukturę S_q^n nie jest równoznaczna z transformacją struktury S_q^n w strukturę S_p^n. Dotyczy to w szczególności badania zmian strukturalnych w ujęciu dynamicznym.

Rozpoznanie własności struktur może być wzbogacone o znajomość entropii empirycznej rozkładu wskaźników struktury lub EOM, co przyczyni się do rozpoznania stopnia koncentracji wskaźników struktury w n-elementowej strukturze S^n.

LITERATURA

Chomątowski S., Sokołowski A. (1978), *Taksonomia struktur*, „Przegląd Statystyczny”, nr 2, 217–222.

Kukuła K. (2000), *Metoda unitaryzacji zerowej*, PWN, Warszawa.

Lavenda B. H. (2006), *Geometric Entropies of Mixing (EOM)*, „Open Systems & Information Dynamics”, No. 13, 91–101.

McBratney A., Minasny B. (2007), *On measuring pedodiversity*, „Geoderma”, No. 141, 149–154.

Młodak A. (2006), *Analiza taksonomiczna w statystyce regionalnej*, Wyd. Dyfin, Warszawa.

Przybyszewski R., Wędrowska E. (2005), *Aksjomatyczna teoria entropii*, „Przegląd Statystyczny”, nr 2, t. 52, 85–101.

Roeske-Słomka I. (1994), *Wielkość i entropia gospodarstw domowych w Polsce: (w świetle powojennych spisów ludności)*, Wydawnictwo Akademii Ekonomicznej w Poznaniu.

Rutkowski J. (1981), *Podobieństwo struktur i zmiany strukturalne – zagadnienia kwantyfikacji*, „Wiadomości Statystyczne”, nr 8, 20–23.

Strahl D. (red.) (1998), *Taksonomia struktur w badaniach regionalnych*, Wyd. Akademii Ekonomicznej we Wrocławiu, Wrocław.

Theil H. (1979), *Zasady ekonometrii*, PWN, Warszawa.

Wędrowska E. (2003), *Datalogiczna miara ilości informacji strukturalnej jako instrument zarządzania zasobami informacji statystycznej*, „Prace Naukowe AE we Wrocławiu”, nr 975, Wrocław, 447–459.

Zoido J. M, Carreno F. (2000), *Geometrical entropies. The extended entropy*, „The European Physical Journal”, B 17, 459–469.

EXPECTED AMOUNT OF INFORMATION ON TRANSFORMATION OF STRUCTURES AS A MEASURE OF DISSIMILARITY OF STRUCTURES

A b s t r a c t . In socio-economic issues there is a need to assess the degree of transformations between the assumed structure (*a priori*) and the structure actually formed (*a posteriori*). This can be applied to, for instance, the comparison of the assumed expenditures structure of certain economic objects with the structure that was actually implemented. The objective of the paper is to present the possibility of assessment of the degree of these transformations. It was proposed to use the expected amount of information on the transformation of the assumed structure into the actually achieved structure for investigating the degree of dissimilarity between these structures. The paper also presents the possibility of applying the actual entropy of structures and trigonometric entropy for studies on characteristics of structures.

K e y w o r d s : entropy, measures of dissimilarity of structures, *Entropies of Mixing*.