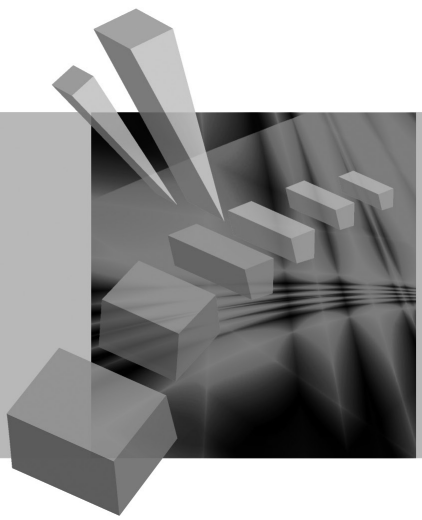
PRACE NAUKOWE
Uniwersytetu Ekonomicznego we Wrocławiu
RESEARCH PAPERS
of Wrocław University of Economics

279

Taksonomia 21

Klasyfikacja i analiza danych – teoria i zastosowania



Redaktorzy naukowi
Krzysztof Jajuga
Marek Walesiak

Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2013

Redaktor Wydawnictwa: Aleksandra Śliwka
Redaktor techniczny: Barbara Łopusiewicz
Korektor: Barbara Cibis
Łamanie: Małgorzata Czupryńska
Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:
www.ibuk.pl, www.ebscohost.com,
The Central and Eastern European Online Library www.ceeol.com,
a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon
http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się
na stronie internetowej Wydawnictwa
www.wydawnictwo.ue.wroc.pl

Tytuł dofinansowany ze środków Narodowego Banku Polskiego
oraz ze środków Sekcji Klasyfikacji i Analizy danych PTS

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana
Druk: Drukarnia TOTEM

Spis treści

Summaries

PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU
RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS nr 279 • 2013

Taksonomia 21. Klasyfikacja i analiza danych – teoria i zastosowania ISSN 1899-3192

Ewa Wędrowska
Uniwersytet Mikołaja Kopernika w Toruniu

WRAŻLIWOŚĆ MIAR DYWERGENCJI JAKO MIERNIKÓW NIEPODOBIEŃSTWA STRUKTUR

Streszczenie: W badaniach zjawisk społeczno-ekonomicznych często podejmowana jest problematyka podobieństwa obiektów gospodarczych scharakteryzowanych wskaźnikami struktury. Zazwyczaj miary wykorzystywane do kwantyfikacji podobieństwa bądź niepodobieństwa struktur są funkcjami metryk odległości ich wskaźników cząstkowych. W badaniu podobieństwa struktur wykorzystać można także miary dywergencji. W artykule wskazana została możliwość wykorzystania do oceny stopnia rozbieżności struktur miar dywergencji klasy Csiszára (*f*-dywergencje), w szczególności takich jak: odległość Hellingera, odległość trójkątną, symetryczną chi-kwadrat dywergencję, dywergencję Kullbacka-Leiblera, dywergencję Jensena-Shannona. Cel artykułu stanowi zbadanie oraz wzajemne porównanie stopnia wrażliwości wskazanych miar na zmiany stopnia rozbieżności struktur.

Słowa kluczowe: miary dywergencji Csiszára, podobieństwo struktur, analiza danych.

1. Wstęp

W analizie porównawczej obiektów scharakteryzowanych wskaźnikami struktury stosowanych jest wiele mierników o różnorodnej konstrukcji. Miary zgodności struktur określają stopień podobieństwa pary struktur. Badanie podobieństwa obiektów opisanych przez wskaźniki struktury bądź udziału może mieć charakter statyczny (przestrzenny) lub dynamiczny. Wśród popularnych miar zgodności wskazać można te, które należą do rodziny miar dywergencji. Przykładem są odległość Hellingera, odległość trójkątna czy też entropia względna Kullbacka-Leiblera. Wykorzystanie tych miar w analizach porównawczych struktur staje się przyczynkiem do zastosowania we wspomnianych analizach innych mierników należących do miar dywergencji.

Celem artykułu jest porównanie popularnych miar niepodobieństwa struktur oraz miar dywergencji pod względem stopnia wrażliwości na rozbieżność w rozkładzie składowych porównywanych struktur.

Miary dywergencji, które odgrywają znaczącą rolę w statystyce i teorii informacji, zaproponowane zostały przez Jeffreysa [1946], Kullbacka i Leiblera [1951], Rényi'ego [1961], Csiszára [1963, 1967, 1974], Havrada i Charváta [1967], Sibsona [1969], Burbea i Rao [1982], Kapura [1984], Lina [1991], Taneja [1995]. W ciągu

ostatnich lat badane są teoretyczne własności miar dywergencji oraz ich wzajemne zależności [Sahoo, Wong 1988; Topsøe 2000; 2001; Kumar, Chhina 2005; Taneja, Kumar 2006; Dragomir 2004; Taneja 2005; 2008; Kumar, Johnson 2005; Anwar, Hussain, Pečarić 2009; Wędrowska 2012].

Jednym z istotnych problemów zastosowania miar dywergencji jest badanie odległości, rozbieżności czy dyskryminacji pomiędzy rozkładami prawdopodobieństwa. Wśród miar dywergencji można wyróżnić mierniki o różnorakich własnościach. Należą tu zarówno miary spełniające własności metryki, jak i miary niespełniające warunku symetrii.

2. Miary dywergencji klasy Csiszára

Koncepcja *f*-dywergencji jako miary rozbieżności pomiędzy dwoma rozkładami prawdopodobieństwa zaproponowana została równocześnie przez Csiszára w 1967 r. oraz Ali'ego i Silveya w roku 1966. W literaturze miary należące do klasy *f*-dywergencji określane są najczęściej mianem *dywergencji Csiszára* lub, rzadziej, *Csiszár–Ali–Silvey dywergencjami*.

Dywergencja Csiszára jest uogólnieniem pewnych miar rozbieżności i stanowi klasę miar zdefiniowanych za pomocą wypukłych funkcji *f* określonych na przedziale $[0,\infty)$.

Miara dywergencji należąca do klasy Csiszára (*f*-dywergencja) pomiędzy strukturami S_r^n oraz S_s^n ze zbioru $\Gamma^n = \left\{ S_j^n = [\omega_{1j}, \omega_{2j}, ..., \omega_{nj}]^T \mid 0 \le \omega_{ij} \le 1, \sum_{i=1}^{n} \omega_{ij} = 1 \right\}$ dla $j = 1, ..., m$ określona jest następująco:

$$C_f(S_r^n, S_s^n) = \sum_{i=1}^{n} \omega_{is} f\left(\frac{\omega_{ir}}{\omega_{is}}\right), \tag{1}$$

gdzie $f : [0,\infty) \to \Re$ jest funkcją różniczkowalną i wypukłą, taką że dla $x = 1$ $f(1) = 0$, $f''(1) \ge 0$ oraz dla $x = 0$ zachodzi: $0 \cdot f\left(\frac{0}{0}\right) = 0$ oraz $0 \cdot f\left(\frac{\omega}{0}\right) = \lim_{x\to\infty} \frac{f(x)}{x}$ [Menéndez i in. 2003].

Dywergencja $C_f(S_r^n, S_s^n)$ dla pary struktur $(S_r^n, S_s^n) \in \Gamma^n \times \Gamma^n$ jest wypukła i przyjmuje wartości nieujemne dla wypukłej funkcji $f : [0,\infty) \to \Re$, takiej że $f(1) = 0$ [Taneja 2005]. Ponadto $C_f(S_r^n, S_r^n) = 0$ dla wszystkich funkcji przyjmujących wartość zero dla argumentu równego jedności [Dragomir, Gluščević, Pearce 2001].

Wiele znanych miar dywergencji należy do uogólnionej klasy zaproponowanej przez Csiszára. Do najczęściej stosowanych *f*-dywergencji należą: odległość miejska, kwadrat odległości Hellingera, odległość trójkątna, χ^2-dywergencja, dywergencja Kullbaca-Leiblera oraz propozycje przedstawione przez Lina, Taneja czy też Kumara.

Tabela 1. Wybrane miary dywergencji Csiszára wraz z wypukłą funkcją *f*

Nazwa	Formuła	Wypukła funkcja $f:[0,\infty)\to\Re$	Przedział wartości	Źródło*
Metryka miejska	$V_{rs}=\sum_{i=1}^{n}\lvert\omega_{ir}-\omega_{is}\rvert$	$f_V(x)=\lvert x-1\rvert$	[0,2]	[Anwar i in. 2009]
Odległość Bray'a-Curtisa	$d_{rs}^{BC}=\frac{1}{2}\sum_{i=1}^{n}\lvert\omega_{ir}-\omega_{is}\rvert$	$f_{BC}(x)=\frac{1}{2}\lvert x-1\rvert$	[0,1]	[Wędrowska 2012]
Odległość trójkątna	$d_{rs}^{\Delta}=\sum_{i=1}^{n}\frac{\lvert\omega_{ir}-\omega_{is}\rvert^2}{\omega_{ir}+\omega_{is}}$	$f_{\Delta}(x)=\frac{(x-1)^2}{x+1}$	[0,2]	[Taneja 2005]
Unormowana odległość trójkątna	$d_{rs}^{\Delta*}=\frac{1}{2}\sum_{i=1}^{n}\frac{\lvert\omega_{ir}-\omega_{is}\rvert^2}{\omega_{ir}+\omega_{is}}$	$f_{\Delta*}(x)=\frac{1}{2}\frac{(x-1)^2}{x+1}$	[0,1]	[Wędrowska 2012]
Kwadrat odległości Hellingera	$\left(d_{rs}^{H}\right)^2=\sum_{i=1}^{n}\left(\sqrt{\omega_{ir}}-\sqrt{\omega_{is}}\right)^2$	$f_H(x)=\left(\sqrt{x}-1\right)^2$	[0,2]	[Simic 2009]
Unormowany kwadrat odległość Hellingera	$\left(d_{rs}^{H*}\right)^2=\frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{\omega_{ir}}-\sqrt{\omega_{is}}\right)^2$	$f_{H*}(x)=\frac{1}{2}\left(\sqrt{x}-1\right)^2$	[0,1]	[Wędrowska 2012]
χ^2-dywergencja	$\chi_{rs}^2=\sum_{i=1}^{n}\frac{(\omega_{ir}-\omega_{is})^2}{\omega_{is}}$	$f_{\chi^2}(x)=\left(x-1\right)^2$	[0,+∞]	[Anwar i in. 2009]
Dywergencja Kullbacka-Leiblera	$D(S_r^n,S_s^n)=\sum_{i=1}^{n}\omega_{ir}\log_2\frac{\omega_{ir}}{\omega_{is}}$	$f_{KL}(x)=x\log_2 x$	[0,+∞]	[Wędrowska 2012]
K-dywergencja	$K_{rs}=\sum_{i=1}^{n}\omega_{ir}\log_2\frac{\omega_{ir}}{\frac{1}{2}\omega_{ir}+\frac{1}{2}\omega_{is}}$	$f_K(x)=x\log_2\frac{2x}{1+x}$	[0,1]	[Lin 1991]
Dywergencja Jensena-Shannona**	$JS(S_r^n,S_s^n)=H_S\left(\frac{S_r^n+S_s^n}{2}\right)-\left(\frac{H_S(S_r^n)+H_S(S_s^n)}{2}\right)$	$f_{JS}(x)=\frac{x}{2}\log_2 x+\frac{x+1}{2}\log_2\left(\frac{2}{x+1}\right)$	[0,1]	[Taneja 2005]

* Źródło dotyczy literatury, w której wskazano wypukłą funkcję $f:[0,\infty)\to\Re$.

** W formule określającej dywergencję Jansena-Shannona $H_s(S)$ oznacza entropię Shannona.

Źródło: opracowanie własne na podstawie cytowanej literatury.

Miary dywergencji kwantyfikują stopień rozbieżności pomiędzy składowymi porównywanych struktur, mają zatem charakter miar niepodobieństwa. Dla struktur identycznych osiągają wartość zero. Z kolei w przypadku całkowitej rozbieżności pomiędzy strukturami nie zawsze występuje górne ograniczenie zbioru wartości. Wśród miar dywergencji Csiszára występują miary o wartościach z przedziału [0, 1] lub z przedziału ograniczonego z góry przez pewną liczbę dodatnią, a także miary o nieograniczonych z góry wartościach [Wędrowska 2012].

Porównywanie struktur w czasie i przestrzeni dokonywane jest za pomocą wielu mierników. Wskazuje się najczęściej, że na wybór miernika mają głównie wpływ: cel badania oraz możliwość oceny i interpretacji rezultatów analiz [Młodak 2006]. Kolejna determinanta wyboru metodologii porównywania struktur wynika z własności algebraicznych i statystycznych metody i dotyczy wrażliwości poszczególnych miar na określony układ strukturalny.

Porównanie stopnia wrażliwości miar niepodobieństwa struktur oraz miar dywergencji na rozbieżność pomiędzy składowymi struktur ograniczono do miar unormowanych w przedziale $[0,1]$, co umożliwiło zachowanie porównywalności uzyskanych wyników. Wybrano następujące miary należące do klasy dywergencji Csiszára: odległość Braya-Curtisa, odległość trójkątną, *K*-dywergencję oraz dywergencję Jensena-Shannona. Pozostałe miary wybrane do analizy to popularne miary wykorzystywane w badaniu zróżnicowań strukturalnych. Są to: unormowana odległość przeciętna $\left(d_{rs}^{E*} = \sqrt{\frac{1}{2}\sum_{i=1}^{n}(\omega_{ir} - \omega_{is})^2},\ (r,s = 1, ..., m) \right)$, unormowana wersja metryki Canberra $\left(d_{rs}^{Can*} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\omega_{ir} - \omega_{is}|}{\omega_{ir} + \omega_{is}}, (r,s = 1, ..., m) \right)$, współczynnik dywergencji Clarka $\left(d_{rs}^{Cl} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\omega_{ir} - \omega_{is}}{\omega_{ir} + \omega_{is}} \right)^2}, (r,s = 1, ..., m) \right)$ oraz unormowana odległość Hellingera.

Za współczynnik wrażliwości przyjęto następujący miernik [Młodak 2006]:

$$\gamma = \left| \frac{d(S_1^n, S_3^n)}{d(S_2^n, S_3^n)} - 1 \right|, \tag{2}$$

gdzie $d(S_1^n, S_3^n)$ oraz $d(S_2^n, S_3^n)$ stanowią wartości odpowiednich miar niepodobieństwa.

W pierwszej kolejności zbadany został stopień wrażliwości miar w przypadku struktur, dla których występowały jednakowe moduły różnic pomiędzy składowymi. Rozważono w tym celu przypadek arbitralnie dobranych składowych trzech struktur (tab. 2).

Tabela 2. Struktura trzech obiektów z występującymi jednakowymi modułami różnicy składników

Obiekty	Elementy struktury obiektów			Moduły różnic	
	S_1^4	S_2^4	S_3^4	$S_1^4 - S_3^4$	$S_2^4 - S_3^4$
X_1	0,30	0,10	0,20	0,10	0,10
X_2	0,35	0,25	0,30	0,05	0,05
X_3	0,20	0,30	0,25	0,05	0,05
X_4	0,15	0,35	0,25	0,10	0,10
Suma	1,00	1,00	1,00	0,30	0,30

Źródło: obliczenia własne.

Rezultaty wartości porównywanych miar dla odpowiednich par struktur wraz z wartościami współczynnika wrażliwości przedstawiono w tab. 3.

Tabela 3. Wartości miar niepodobieństwa struktur oraz miar dywergencji Csiszára wraz ze współczynnikami wrażliwości dla danych z tab. 2.

Miara	Wartość miary dla pary struktur		Współczynnik wrażliwości
	(S_1^4, S_3^4)	(S_2^4, S_3^4)	γ
d_{rs}^{BC}	0,150000	0,150000	0,00000
d_{rs}^{E*}	0,112000	0,112000	0,00000
d_{rs}^{Can*}	0,159509	0,170455	0,06422
d_{rs}^{Cl}	0,173755	0,197115	0,11851
d_{rs}^{H*}	0,117295	0,122687	0,04395
$d_{rs}^{\Delta*}$	0,027201	0,029545	0,07936
$K(S_r^n, S_s^n)$	0,020090	0,022624	0,11202
$K(S_s^n, S_r^n)$	0,019454	0,020532	0,05252
$JS(S_r^n, S_s^n)$	0,019772	0,021578	0,08371

Źródło: obliczenia własne.

Analiza wrażliwości miar dywergencji na rozbieżność w rozkładzie składowych porównywanych struktur w przypadku jednakowych modułów różnic pomiędzy składowymi struktur S_1^4 i S_3^4 oraz S_2^4 i S_3^4 prowadzi do następujących wniosków:

- Pomimo odmienności w rozkładach składowych analizowanych struktur nie występuje zróżnicowanie wartości miar wykorzystujących jedynie różnice składowych (odległość Bray'a-Curtisa oraz unormowana odległość przeciętna). W przypadku tych miar współczynnik wrażliwości przyjmuje wartość zero.

- W przypadku pozostałych miar występuje zróżnicowanie wartości w określaniu stopnia rozbieżności pomiędzy strukturami. Miary, w których różnice pomiędzy składowymi odnoszone są do sumy tych składowych, czyli unormowana metryka Canberra oraz współczynnik dywergencji Clarka, przypisują większe znaczenie różnicom uzyskanym ze składowych o niższych udziałach, stąd ich wartości są największe dla struktur o składowych zaproponowanych w tab. 2.
- Dla struktur o składowych zaproponowanych w tab. 2 wartości współczynnika wrażliwości wskazują, że największe zróżnicowanie pomiędzy stopniem rozbieżności odpowiednio struktur S_1^4 i S_3^4 oraz S_2^4 i S_3^4 wykazują kolejno: współczynnik dywergencji Clarka, jedna z *K*-dywergencji, dywergencja Jansena-Shannona. Najmniejsze zróżnicowanie wartości wystąpiło dla odległości Hellingera.

Kolejnym etapem w badaniu wrażliwości analizowanych miar niepodobieństwa oraz miar dywergencji na określony układ strukturalny jest rozważenie struktur o niejednakowych modułach różnicy ich składowych, lecz o tej samej sumie ich modułów. Ponownie rozważony został przypadek o arbitralnie przyjętych wskaźnikach. Składowe struktur zamieszczone zostały w tab. 4, a wartości miar niepodobieństwa oraz miar dywergencji Csiszára dla porównywanych struktur wraz ze współczynnikiem wrażliwości zamieszczono w tab. 5.

Tabela 4. Struktura trzech obiektów o niejednakowych modułach różnicy składników

Atrybuty cechy	Elementy struktury obiektów			Moduły różnic	
	S_1^4	S_2^4	S_3^4	$S_1^4 - S_3^4$	$S_2^4 - S_3^4$
X_1	0,55	0,60	0,45	0,10	0,15
X_2	0,30	0,20	0,25	0,05	0,05
X_3	0,10	0,15	0,20	0,10	0,05
X_4	0,05	0,05	0,10	0,05	0,05
Suma	1,00	1,00	1,00	0,30	0,30

Źródło: obliczenia własne.

Tabela 5. Współczynniki wrażliwości dla unormowanych miar niepodobieństwa struktur oraz miar dywergencji Csiszára dla danych z tab. 4

Miara Oznaczenie	Wartość miary dla pary struktur		Współczynnik wrażliwości
	(S_1^4, S_3^4)	(S_2^4, S_3^4)	γ
1	2	3	4
d_{rs}^{BC}	0,15000	0,15000	0,00000
d_{rs}^{E*}	0,11180	0,12247	0,08713
d_{rs}^{Can*}	0,21439	0,18254	0,17451
d_{rs}^{Cl}	0,24520	0,20265	0,20994

1	2	3	4
d_{rs}^{H*}	0,12850	0,11341	0,13307
$d_{rs}^{\Delta*}$	0,03227	0,02540	0,27074
$K(S_r^n, S_s^n)$	0,02554	0,01899	0,34466
$K(S_s^n, S_r^n)$	0,02173	0,01796	0,21034
$JS(S_r^n, S_s^n)$	0,02364	0,01848	0,27938

Źródło: obliczenia własne.

Wartości miar niepodobieństwa struktur oraz miar dywergencji są odmienne, co jest oczywiste, gdyż miary te mają odmienne konstrukcje. W przypadku porównywania struktur o niejednakowych modułach różnicy ich składowych jedynie odległość Braya–Curtisa nie wykazuje wrażliwości na rodzaj odmienności pomiędzy strukturami S_1^4 i S_3^4 oraz S_2^4 i S_3^4. Analiza wartości pozostałych miar wskazuje, że występują istotne różnice w wartościach tych miar odpowiednio dla par struktur (S_1^4, S_3^4) oraz (S_2^4, S_3^4), dla których występuje jednakowa suma modułów różnic składowych. Stąd uzyskano znaczne wartości współczynnika wrażliwości dla analizowanych miar niepodobieństwa struktur oraz miar dywergencji Csiszára. Największe rozróżnienie pomiędzy wartościami rozważanych miar uzyskanymi dla porównywanych par struktur wykazują kolejno: jedna z *K*-dywergencji, dywergencja Jensena-Shannona, odległość trójkątna, a zatem miary należące do klasy dywergencji Csiszára. Dotyczy to miar, które osiągały relatywnie najniższe wartości, uzyskując jednocześnie najwyższe wartości współczynnika wrażliwości.

3. Podsumowanie

Porównywane miary niepodobieństwa struktur oraz miary dywergencji Csiszára charakteryzują się różnym stopniem wrażliwości na stopień rozbieżności pomiędzy strukturami. Wpływ na to mają nie tylko bezwzględne różnice składowych struktur, ale też fakt, czy różnice te uzyskane są ze składowych o relatywnie dużych czy też małych wartościach. Zatem dobór odpowiedniej miary niepodobieństwa powinien wynikać ściśle z charakteru i specyfiki badań nad rozbieżnością struktur, a stopień wrażliwości miar powinien być uwzględniany przez badacza. Można jednak uznać, że miary dywergencji Csiszára mogą stanowić poszerzenie aparatu pomiarowego stosowanego w analizie porównawczej struktur ze względu na dużą wrażliwość na rozbieżność w rozkładzie składowych porównywanych struktur. Zastosowanie tych miar zasadne jest w sytuacjach, gdy istnieje potrzeba uwypuklenia tej wrażliwości.

Literatura

Anwar M., Hussain S., Pečarić J., *Some inequalities for Csiszár-divergence measures*, "Int. Journal of Math. Analysis" 2009, vol. 3, no. 26, 1295-1304.

Burbea J., Rao R.C,. *On the convexity of some divergence measures based on entropy functions*, "IEEE Transactions on Information Theory" 1982, vol. 28, no. 3, 489-495

Csiszár I., *Eine inrofmationtheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizit¨at von Markotschen ketten*, „Publ. Math. Inst. Hungar. Acad. Sci." 1963, no. 8, 85-108

Csiszár I., *Information-type measures of difference of probability distributions and indirect observation*, "Studia Scientiarum Mathematicarum Hungarica" 1967, no. 2, 229-318.

Csiszár I., *Information measures: a critical survey*, "Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions" 1974, vol. 2, 73-87.

Dragomir S.S., *A converse inequality for the Csisz´ar Φ−divergence*, "Tamsui Oxford Journal of Mathematical Sciences" 2004, no. 20(1), 35-53

Dragomir S.S., Gluščević V., Pearce C.E.M., *Csiszár f–divergence, Ostrowski's inequality and mutual information*, "Nonlinear Analysis" 2001, no. 47, 2375-2386.

Havrada J., Charvát F., *Quantification methods of classification processes: Concept of structural α-entropy*, "Kybernetica (Prague)" 1967, no. 3, 95-100.

Jeffreys H., *An invariant form for the prior probability in estimating problems*, "Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences" 1946, 453-461.

Kapur J.N., *A comparative assessment of various measures of directed divergence*, "Advances in Management Studies" 1984, vol. 3, no. 1, 1-16.

Kullback S., Leibler R.A., *On information and sufficiency*, "Annals of Mathematical Statistics" 1951, vol. 22, no. 1, 79-86.

Kumar P., Chhina S.A., *Symmetric information divergence measure of the Csiszár's f-divergence class and its bounds*, "Computers and Mathematics with Applications" 2005, vol. 49, 575-588.

Kumar P., Johnson A., *On a symmetric divergence measure and information inequalities*, "Journal of Inequalities in Pure and Applied Mathematics" 2005, vol. 6, Issue 3, Article 65.

Lin J., *Divergence measures based on the Shannon entropy*, "IEEE Transactions on Information Theory" 1991, no. 37, 145-151.

Menéndez M.L., Pardo J.A., Pardo L., Zografos K. *On tests of homogeneity based on minimum φ-divergence estimator with constraints*, "Computational Statistics and Data Analysis" 2003, vol. 43, 215-234.

Młodak A., *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa 2006.

Rényi A., *On measures of entropy and information*, "Proc. Fourth Berkeley Symp. Math. Stat. and Prob.", University of California Press, 1961, 547-561.

Sahoo P.K., Wong A.K.C. *Generalized Jensen difference based on entropy functions*, "Kybernetika" 1988, vol. 24, no. 4, 241-250.

Sibson R., *Information radius*, "Probability Theory and Related Fields" 1969, vol. 12, no. 2, Springer Berlin 1969, 149-160.

Taneja I.J., *New Developments in Generalized Information Measures*, [w:] *Advances in Imaging and Electron Physics*, red. P.W. Hawkes, 1995, 37-135.

Taneja I.J., *On symmetric and non-symmetric divergence measures and their generalizations*, "Advances in Imaging and Electron Physics", vol. 138, 2005, 198-248.

Taneja I.J., *On Mean Divergence Measures*, [w:] *Advances in Inequalities from Probability Theory & Statistics*, red. N.S. Barnett, S.S. Dragomir, Nova Science Publishers, 2008, 169-186.

Taneja I.J., Kumar P., *Relative information of type s, Csiszár's f-divergence, and information inequalities*, "Information Sciences" 2006, no. 166, 105-125.

Topsøe T., *Some inequalities for information divergence and related measures of discrimination*, "IEEE Transactions on Information Theory" 2000, vol. 46, no. 4, 1602-1609.

Topsøe T., *Bounds for entropy and divergence of distributions over a two-element set*, "J. Ineq. Pure Appl. Math." 2001, vol. 2, Article 25, 13 pp.
Wędrowska E., *Miary entropii i dywergencji w analizie struktur*, Wyd. UWM, Olsztyn 2012.

SENSITIVITY OF DIVERGENCE MEASURES AS STRUCTURE DISSIMILARITY MEASUREMENTS

Summary: The analyses of social and economic phenomena often involve the issue of similarity between business objects characterized by structure indicators. Usually, measures used for quantifying similarity or the lack of similarity between structures are a function of the distance metrics of their partial indicators. An examination of the similarity between structures can also apply divergence measures. This article indicates the possibility of using Csiszár class divergence measures (f-divergences), in particular: Hellinger discrimination, triangular discrimination, symmetric Chi-square divergence, arithmetic-geometric mean divergence, Kullback-Leibler divergence and Jensen-Shannon divergence to evaluate the degree of discrepancy between structures. The aim of the article is to examine the sensitivity of the indicated measures to the changes in the degree of discrepancy between structures.

Keywords: Csiszar's divergence, similarity of structure, data analysis.