

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

236

Badania marketingowe – metody, nowe podejścia i konteksty badawcze



pod redakcją

Krystyny Mazurek-Łopacińskiej

Magdaleny Sobocińskiej



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2011

Recenzenci: Anna Dąbrowska, Lechosław Garbarski, Józef Garczarczyk

Redaktor Wydawnictwa: Agnieszka Flasińska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Publikacja jest dostępna na stronie www.ibuk.pl

Streszczenia opublikowanych artykułów są dostępne w międzynarodowej bazie danych The Central European Journal of Social Sciences and Humanities <http://cejsh.icm.edu.pl> oraz w 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

Kopiowanie i powielanie w jakiegokolwiek formie wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2011

ISSN 1899-3192

ISBN 978-83-7695-248-2

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp.....	9
------------	---

Część 1. Rozwój badań marketingowych – kierunki, koncepcje, wymiary

Krystyna Mazurek-Łopacińska, Magdalena Sobocińska: Rozwój badań marketingowych – w kierunku nowych podejść i kontekstów badawczych związanych z funkcjonowaniem przedsiębiorstwa.....	13
Dariusz Oczachowski: Kultura kognitywna organizacji a prowadzenie badań marketingowych	23
Tomasz Heryszek: Myślenie kognitywne czy afektywne? Między danymi twardymi a przeczuciem – dylematy współczesnego przedsiębiorcy	33
Jan W. Wiktor: Koncepcja i sposób pomiaru internacjonalizacji przedsiębiorstwa w świetle doświadczeń projektu „Strategie marketingowe przedsiębiorstw na rynkach międzynarodowych”	42
Aleksandra Nizielska: Dystans kulturowy w procesie internacjonalizacji przedsiębiorstw – metodyczny aspekt badań.....	51
Magdalena Soboń, Iga Rudawska, Sylwia Bąkowska: Metodyka badania zachowań konsumenckich w międzynarodowej sieci badawczej.....	61
Bogusław Bembenek: Rola wywiadu gospodarczego w zarządzaniu wiedzą w klastrze	71

Część 2. Nowe podejścia badawcze wynikające z rozwoju technologii informacyjnych i komunikacyjnych

Mariusz Kuziak: Wyzwania wobec badań użytkowników Internetu.....	83
Karol Łopaciński: Model tematycznej hurtowni danych na potrzeby badania przebiegu i efektów kampanii e-mailingowych.....	95
Magdalena Jaciow: Fora internetowe jako źródło informacji – możliwości i ograniczenia wykorzystania w badaniach zachowań nabywczych	113
Beata Kolny: Wykorzystanie netnografii do badania usług zagospodarowujących czas wolny	123
Radosław Szulc, Piotr Ciszewski: Wykorzystanie technologii łączności bezprzewodowej w badaniach marketingowych na rynku produktów <i>beauty care</i>	132

Część 3. Metody ilościowe – nowe podejścia i zastosowania oraz triangulacja metod

Adam Sagan: Asymetryczne metody wielowymiarowe w badaniach marketingowych	145
Mariusz Łapczyński: Łączenie metod i narzędzi w budowie modeli predycyjnych.....	155
Józef Garczarczyk, Robert Skikiewicz: Zastosowanie metody grupowania dwustopniowego w segmentacji klientów indywidualnych na rynku usług bankowych.....	164
Anna Bryja: Zmienne jakościowe w segmentacji rynku: miary powiązań a wyniki grupowania.....	175
Bartłomiej Jefmański: Nowe podejście w pomiarze opinii respondentów z zastosowaniem skal porządkowych i elementów teorii zbiorów rozmytych – charakterystyka wybranych aspektów metodologicznych.....	184
Grzegorz Maciejewski: Wykorzystanie analizy czynnikowej w badaniach konsumenckiego ryzyka	192
Paweł Chlipała: Zastosowanie eksperymentu w badaniach społecznie odpowiedzialnej konsumpcji – refleksje nad metodą, wyniki	203

Część 4. Badania jakościowe i ich wykorzystanie w rozwiązywaniu problemów badawczych i decyzyjnych

Marcin Komor: Znaczenie i rozwój metod jakościowych w badaniach empirycznych w marketingu.....	215
Sylwia Wrona: Dobór próby w jakościowych badaniach marketingowych – problemy prawidłowej selekcji i rekrutacji uczestników	225
Mateusz Rak, Joanna Nogiec: Wykorzystanie wyników badań jakościowych do identyfikacji populacji w badaniach ilościowych.....	234
Iwona Olejnik: Metoda obserwacji – zastosowania w badaniach marketingowych.....	242
Zbigniew Piskorz: Rozpoznawanie okazji przedsiębiorczych – rezultaty badań jakościowych.....	250
Zbigniew Spyra: Zastosowanie podejścia etnograficznego we współczesnych badaniach marketingowych w sferze kultury	260
Wanda Patrzalek: Przebieg i zaburzenia procesów wymiany informacji oraz komunikacji wewnątrz gospodarstwa domowego w ogólnopolskich badaniach fokusowych	271

Joanna Wardzała-Kordyś: Wiedza i opinie o procesach reklamacji produktów wadliwych w kontekście badań gospodarstw domowych	280
Agnieszka Dejnaka: Komunikacja pomiędzy członkami gospodarstwa domowego przy użyciu nowoczesnych narzędzi wymiany informacji	290
Jolanta Tkaczyk: Rola opowieści w badaniach marketingowych	301

Summaries

Part 1. Development of marketing research – trends, concepts, dimensions

Krystyna Mazurek-Łopacińska, Magdalena Sobocińska: Development of marketing research – towards new approaches and contexts of research related to the functioning of a company	22
Dariusz Oczachowski: Cognitive culture of an organization and conducting marketing research	32
Tomasz Heryszek: Cognitive or affective thinking? Between hard data and foreboding – modern business dilemmas	41
Jan W. Wiktor: The concept and measurement method of the companies' internationalization in the light of the research project "Marketing strategies of companies on international markets"	50
Aleksandra Nizielska: Cultural distance in the process of companies' internationalization – methodological aspect of research	60
Magdalena Soboń, Iga Rudawska, Sylwia Bąkowska: Consumer behaviour research methods in the International Research Network	70
Bogusław Bembenek: The role of economic intelligence in knowledge management of cluster	80

Part 2. New research approaches arising from the development of information and communication technologies

Mariusz Kuziak: Challenges to Internet audience measurement	94
Karol Łopaciński: Model of data mart prepared due to research regarding proceeding of e-mail campaigns and its effects	112
Magdalena Jaciow: Online forums as a source of information – possibilities and limitations of use in purchasing behaviour research	122
Beata Kolny: The application of netnographic surveys to research on leisure time services	131
Radosław Szulc, Piotr Ciszewski: The implementation of wireless technology in marketing research within the beauty industry	142

Part 3. Quantitative methods – new approaches, applications and triangulation methods

Adam Sagan: Asymmetric multivariate methods in marketing research	154
Mariusz Łapczyński: Combining methods and tools in building predictive models.....	163
Józef Garczarczyk, Robert Skikiewicz: Applying a Two Step Cluster method in the segmentation of individual customers for the banking services market	174
Anna Bryja: Qualitative variables in market segmentation: Similarity coefficients and clustering results.....	183
Bartłomiej Jefmański: A new approach in respondents' opinion measurement using ordinal scales and elements of fuzzy sets theory – characteristics of selected methodological aspects	191
Grzegorz Maciejewski: The use of factor analysis in consumer risk research	202
Paweł Chlipała: Using an experiment in the research of socially responsible consumption – reflections on the method, the results.....	212

Part 4. Qualitative research and its use in problems solving research and decision

Marcin Komor: The importance and development of quality methods in empirical research in marketing	224
Sylvia Wrona: Selection of a sample in qualitative marketing research – issues related to the accurate selection and recruitment of participants.....	233
Mateusz Rak, Joanna Nogiec: Using the results of qualitative research to the identification of the population in quantitative research.....	241
Iwona Olejnik: The method of observation – application in marketing research	249
Zbigniew Piskorz: Entrepreneurial opportunity recognition – results of qualitative research	259
Zbigniew Spyra: Application of ethnographic approach to contemporary marketing research in culture.....	270
Wanda Patrzalek: Proceeding and abnormal processes of information exchange and communication within the household in nationwide focus studies	279
Joanna Wardzała-Kordyś: Knowledge and opinions about complaints to the defective products in the context of household surveys	289
Agnieszka Dejnaka: Communication between members of the household by using modern information exchange tools.....	300
Jolanta Tkaczyk: The role of storytelling in marketing research	310

Adam Sagan

Uniwersytet Ekonomiczny w Krakowie

ASYMETRYCZNE METODY WIELOWYMIAROWE W BADANIACH MARKETINGOWYCH

Streszczenie: Celem referatu jest prezentacja marketingowych zastosowań asymetrycznych metod wielowymiarowych, a w szczególności asymetrycznej analizy korespondencji oraz związanej z nią analizy wariancji dla danych nominalnych (CATANOVA), pozwalającej na poprawną ocenę układów zależności przyczynowych w tabelach na podstawie dekompozycji asymetrycznych współczynników siły związku w tabeli kontyngencji (współczynników τ Godmana-Kruskala, τ_M Marcotorchina, Δ Simonettiego lub indeksu Graya-Williamsa). Zastosowanie tych metod pozwala na poprawną analizę zależności w tabelach danych budowanych na podstawie stosunkowo małych prób.

Słowa kluczowe: asymetryczna analiza korespondencji, analiza wariancji zmiennych nominalnych (CATANOVA).

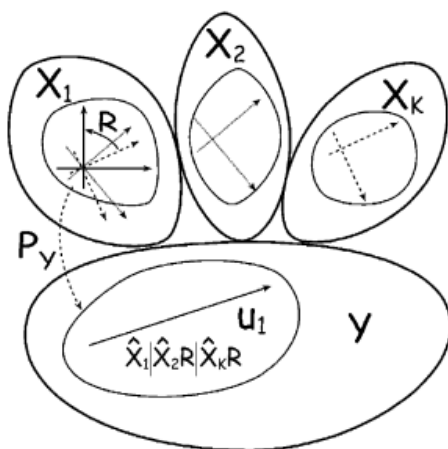
1. Metody wielowymiarowe w badaniach marketingowych

Wielowymiarowa analiza danych ma już długą tradycję w badaniach marketingowych. W analizie tej wykorzystywane są takie metody, jak analiza głównych składowych, analiza głównych współrzędnych (metryczne skalowanie wielowymiarowe) i analiza korespondencji. Powszechność stosowania tych metod wiąże się jednakże z dosyć częstą praktyką asymetrycznej interpretacji wyników analiz (z wyróżnieniem zmiennych zależnych i niezależnych). Jest to szczególnie widoczne w przypadku analizy korespondencji, w której asymetrycznej tabeli danych ($n \times m$) towarzyszy (błędnie) asymetryczna interpretacja wyników. Jest to widoczne szczególnie w badaniach marketingowych, w których badaczka mniej interesuje dotarcie do ukrytej struktury wymiarów, a bardziej określenie zależności między grupami predyktorów (zmiennych niezależnych) i wywoływanych przez nie efektów (zmiennych zależnych). W tego typu sytuacjach (a stanowią one zdecydowaną większość zastosowań analiz wielowymiarowych w badaniach marketingowych), stosowanie metod klasycznych może prowadzić do błędnej interpretacji wyników. W tego typu sytuacjach bardziej poprawnymi metodami wielowymiarowymi są asymetryczne metody analizy korespondencji, głównych składowych lub skalowania wielowymiarowego.

2. Asymetryczne metody wielowymiarowe

Do głównych metod wielowymiarowych stosowanych w badaniach marketingowych należą: analiza głównych składowych (dekompozycja macierzy korelacji), skalowanie wielowymiarowe (dekompozycja macierzy dystansów) oraz analiza korespondencji (dekompozycja współczynnika phi). Wszystkie te metody mają również swoje asymetryczne odpowiedniki w analizie danych.

W obszarze zastosowań analizy głównych składowych do tego typu metod asymetrycznych zalicza się analizę głównych składowych ze zmiennymi instrumentalnymi (*Principal Component Analysis with Instrumental Variables* – PCAIV) [Rao 1964] oraz analizę głównych składowych z projekcją na podprzestrzeń referencyjną (*Principal Component Analysis onto a Reference Subspace* – PCAR) [Esposito, Balbi 1999]. Szczególne zastosowania w badaniach marketingowych reprezentuje druga metoda analizy. Celem PCAR jest odzwierciedlenie w układzie zredukowanej przestrzeni osi głównych relacji zachodzących nie w ramach jednego, ale między dwoma zbiorami zmiennych, z których jeden jest zbiorem zmiennych niezależnych (X), a drugi zbiorem zmiennych zależnych (Y). Na podstawie informacji o kowariancjach między blokami zmiennych (XX, XY, YX, YY) analiza pozwala na wyodrębnienie głównych składowych uzyskiwanych poprzez projekcję zbioru zmiennych zależnych (Y) na przestrzeń zbudowaną w oparciu o zbiór zmiennych niezależnych, maksymalizując zakres wyjaśnianej wariancji dla każdej ze zmiennych zależnych przez zbiór predyktorów (wyjaśnianiu podlega suma współczynników korelacji wielorakich między daną zmienną zależną a zbiorem predyktorów). Idea PCAR jest przedstawiona na rys. 1.



Rys. 1. Idea metody PCAR

Źródło: [Vinzi 2001].

W pierwszym etapie analizy dokonywana jest ortogonalna rotacja procrustowa (R) macierzy \mathbf{X}_2 oraz $\mathbf{X}_1, \dots, \mathbf{X}_k$ zmiennych niezależnych. Następnie dokonywana jest projekcja X na wektor Y zmiennej zależnej. W celu identyfikacji struktury zmienności wykonywana jest analiza głównych składowych (u_1 reprezentuje tu pierwszy wektor własny).

Skalowanie wielowymiarowe jest zwykle związane z analizą danych symetrycznych, jakimi są miary odległości (np. euklidesowa), w których odległość A do B między punktami reprezentującymi obiekty (np. marki) jest taka sama jak odległość B do A . W badaniach marketingowych danymi wejściowymi są często dane o podobieństwach między obiektami (*proximities*). W tego typu danych ich przekształcenie na miary odległości wiąże się z utratą informacji. Cechą charakterystyczną miar bliskości jest ich asymetryczność wywoływana naturą relacji między badanymi obiektami. Typowymi przykładami danych asymetrycznych relacji w macierzach podobieństw są dane dotyczące pozycjonowania produktów, gdzie relacje typu prototyp–kopia w badaniach percepcji produktów (naśladowca jest bardziej „podobny” do lidera niż lider do naśladowcy) czy dane dotyczące zmian preferencji marek (*brand switching*) w tabelach przepływów mają charakter asymetryczny.

W asymetrycznym skalowaniu wielowymiarowym tabela asymetrycznych podobieństw \mathbf{X} jest dekomponowana na dwie macierze: symetryczną macierz odległości (\mathbf{M}) oraz jej asymetryczną część (\mathbf{N}) (*skew-symmetric*), tak że $\mathbf{X} = \mathbf{M} + \mathbf{N}$ [Zielman, Heiser 1996]. W procesie dekompozycji stosowane są różne metody (dekompozycja według wartości osobliwej, model odległościowy, metoda wektorów dryfujących (*drift vectors*), model odległości promieniowych (*radius distance model*) itp. [Borg, Groenen 2005]. W metodzie Gowera część symetryczna jest przedmiotem klasycznej analizy MDS, a część asymetryczna jest reprezentowana w przestrzeni wielowymiarowej dekomponowana z wykorzystaniem analizy głównych składowych [Constantine, Gower 1978].

3. Asymetryczna analiza korespondencji

Celem zastosowanie asymetrycznej analizy korespondencji jest ocena zależności (a nie współzależności) między kategoriami wierszy i kolumn tabeli kontyngencji. W jej ocenie nie można stosować tradycyjnej miary współzależności, jaką jest statystyka χ^2 . W analizie asymetrycznej zamiast współczynnika χ^2 stosowane są asymetryczne miary zależności w tabelach kontyngencji, takie jak współczynnik τ Goodmana-Kruskala, współczynnik τ_M dla danych kubicznych Marcotorchina, indeks Graya-Williamsa, wskaźnik Δ Simonettiego, zagregowany wskaźnik predykcyjny (API) i wskaźnik Tallura [D’Ambra, Lauro 1992].

$$\tau = \frac{\sum \sum_{ij} p_{\cdot j} \left(\frac{p_{ij}}{p_{\cdot j}} - p_i \right)^2}{1 - \sum p_{i\cdot}^2} = \sum_j \tau_j \quad (1)$$

Współczynnik τ Goodmana-Kruskala (1) lub jego rozszerzona wersja dla kostek danych τ_M Marcotorchina są najczęściej wykorzystywanymi miarami oceny siły predykcji w tabelach kontyngencji. Odzwierciedla on różnicę między predykcją poszczególnych kategorii zmiennej zależnej dokonywanej na podstawie rozkładów brzegowych (p_i) a predykcją dokonaną na podstawie rozkładów warunkowych ($p_{ij}/p_{j\cdot}$). Jeżeli rozkłady warunkowe są identyczne jak rozkłady brzegowe, to znajomość rozkładów warunkowych nie poprawia predykcji na podstawie rozkładów brzegowych i współczynnik $\tau = 0$. W sytuacji przeciwnej występuje doskonała przewidywalność rozkładów zmiennej zależnej na podstawie znajomości rozkładów warunkowych i współczynnik $\tau = 1$. Z tego powodu takie współczynniki, jak τ Goodmana-Kruskala, λ Kruskala lub współczynnik niepewności są miarami proporcjonalnej redukcji błędu predykcji (PRE).

Bardzo często jednak współczynnik τ przybiera niskie wartości pomimo istniejących zależności w tabeli kontyngencji, dlatego jest on uzupełniany o test hipotezy dotyczącej równości brzegowych i warunkowych rozkładów w tabeli danych i tym samym istotności relacji predykcyjnych wskazywanych przez współczynnik τ . W asymetrycznej analizie korespondencji najczęściej stosowana jest w tym celu statystyka C kategoryjnej analizy wariancji (CATANOVA) Margolina–Lighta [Beh, D’Ambra 2009]:

$$C = (n-1)(I-1)\tau = \frac{(n-1)(I-1)}{1 - \sum p_{i\cdot}^2} \sum_{m=1} \lambda_m^2. \quad (2)$$

Jest ona pochodną klasycznej analizy wariancji i współczynnika Giniego. Wynika to stąd, że suma kwadratów odchyłeń od średniej jest równa kwadratowi różnicy wartości między wszystkimi parami obserwacji. Analiza ta polega na dekompozycji całkowitej zmienności obserwacji w tabeli danych na zmienność międzygrupową i wewnątrzgrupową. Zmienność całkowita (TSS) jest funkcją kwadratów obserwacji brzegowych zmiennej niezależnej i jest dana jako:

$$TSS = \frac{n}{2} - \frac{1}{2n} \sum_{i=1}^I n_i^2. \quad (3)$$

Zmienność wewnątrzgrupowa (BSS) jest funkcją odchyłeń rozkładów warunkowych od odpowiednich rozkładów brzegowych:

$$BSS = \frac{n}{2} - \frac{1}{2} \sum_{j=1}^J \frac{1}{n_{\cdot j}} \sum_{i=1}^I n_{ij}^2. \quad (4)$$

Zmienność wewnątrzgrupowa jest różnicą między zmiennością całkowitą a zmiennością międzygrupową ($BSS = TSS - WSS$). Podobnie jak w analizie wariancji, odnosząc zmienność całkowitą do liczebności próby, a zmienność międzygrupową do liczby kategorii zmiennej niezależnej (grup), otrzymujemy średnią zmienność

całkowitą (MS_T), międzygrupową (MS_B) i wewnątrzgrupową (MS_W). Statystyka C jest również funkcją stosunku zmienności międzygrupowej do zmienności całkowitej:

$$C = (n-1)(I-1) \left(\frac{BSS}{TSS} \right)$$

Na tej podstawie można również obliczyć współczynnik $\tau = \frac{MS_B}{MS_T}$. W porównaniu do klasycznego testu χ^2 , statystyka C jest bardziej odporna i powinna być wybierana w przypadku małej liczebności próby [Sarnacchiaro, D'Ambra 2007].

Celem asymetrycznej analizy korespondencji jest dekompozycja współczynnika τ , a dokładniej tabeli reszt predykcji (reprezentowanych przez licznik tego współczynnika) $\pi_{ij} = \frac{p_{ij}}{p_{\bullet j}} - p_{i\bullet}$ w zredukowanym układzie o mniejszej liczbie wymiarów i ukazanie predykcyjnych zależności między wierszami i kolumnami tabeli kontyngencji. Jest to analiza asymetryczna, więc w zależności od tego, czy zmienna w wierszach czy kolumnach jest traktowana jako zmienna zależna, obliczany jest współczynnik τ_a lub τ_b (zależny wierszowo, lub zależny kolumnowo)¹. Macierz reszt predykcji (różnice między predykcją brzegową i warunkową) jest przedmiotem dekompozycji SVD:

$$\pi_{ij} = \frac{p_{ij}}{p_{\bullet j}} - p_{i\bullet} = \sum_{s=1}^S \lambda_s a_{is} b_{js}. \quad (5)$$

Tabelą wejściową do analizy asymetrycznej jest tabela scentrowanych profili kolumnowych (lub wierszowych w zależności od tego, gdzie znajduje się zmienna niezależna). Zawiera ona informacje, które komórki w tabeli charakteryzują się wyższą (lub niższą) proporcją w porównaniu do rozkładów brzegowych i tym samym wskazuje na wzrost (lub spadek) predykcji zmiennej zależnej dla określonych kategorii zmiennej niezależnej. W efekcie biplot asymetrycznej analizy korespondencji prezentuje kategorie zmiennej niezależnej i zależnej w zredukowanej przestrzeni, maksymalizując moc predykcyjną zmiennych niezależnych.

Porównując dekompozycję w symetrycznej i asymetrycznej analizie korespondencji, należy również zauważyć, że w obliczeniu bezwładności układu (inercji) w symetrycznej analizie korespondencji uwzględnione są rozkłady brzegowe ($1/p_{i\bullet}$) (por. wzór (6)).

$$I = \sum_{i=1}^I \frac{1}{p_{i\bullet}} \sum_{j=1}^J \left(\frac{p_{ij}}{\sqrt{p_{\bullet j}}} - \frac{p_{i\bullet}}{\sqrt{p_{\bullet j}}} \right)^2. \quad (6)$$

¹ Współczynniki te nie mają związku z symetrycznymi współczynnikami korelacji (zgodności) rang τ_a , τ_b oraz τ_c dla danych porządkowych. Jest to tylko zbieżność nazw.

Natomiast w asymetrycznej analizie korespondencji rozkłady te nie są brane pod uwagę (por. wzór (7)), co w przypadku tabel zbudowanych na podstawie mniejszych prób (tabel z wieloma wartościami zerowymi w komórkach) czyni asymetryczną analizę korespondencji mniej wrażliwą na małe proporcje brzegowe w tabeli danych (na przypadki odstające).

$$\tau_{num} = \sum_{ij} p_{\bullet j} \left(\frac{p_{ij}}{p_{\bullet j}} - p_{i\bullet} \right)^2 = \sum_{ij} \left(\frac{p_{ij}}{\sqrt{p_{\bullet j}}} - \frac{p_{i\bullet}}{\sqrt{p_{\bullet j}}} \right)^2. \quad (7)$$

Diagnoza mocy predykcyjnej w NSCA jest dokonywana na podstawie wskaźników jakości predykcji, do których należą: wskaźnik względnego wkładu zmiennej zależnej w dekompozycję τ_{num} (*Relative Response Contribution*), wskaźnik wkładu zmiennej niezależnej w dekompozycję τ_{num} (*Relative Predictor Contribution*) oraz wskaźnik względnego wkładu osi głównej w dekompozycję τ_{num} (*Relative Axis Contribution*).

$$RA_m = \frac{\lambda_m^2}{\tau_{num}} \quad RPC_j = \frac{p_{\bullet j} \sum_{i=1}^I \pi_{ij}^2}{\tau_{num}} = \frac{p_{\bullet j} \sum_{m=1}^M g_{jm}^2}{\tau_{num}} \quad RRC_i = \frac{\sum_{j=1}^J p_{\bullet j} \pi_{ij}^2}{\tau_{num}} = \frac{\sum_{m=1}^M f_{im}^2}{\tau_{num}}. \quad (8)$$

Graficzną prezentacją zależności asymetrycznych jest graficzny biplot w zredukowanej przestrzeni osi głównych. Celem biplotu jest prezentacja asymetrycznych zależności między zmiennymi, więc współrzędne punktów reprezentujących kategorie zmiennej zależnej są współrzędnymi standardowymi, a punkty reprezentujące kategorie zmiennej niezależnej mają współrzędne główne. Informacja o statystycznej istotności powiązań jest przedstawiona za pomocą okręgów lub elips ufności. Dla przykładu 95-procentowy okrąg ufności dla danej kategorii zmiennej zależnej jest obliczany za pomocą wzoru:

$$r_j^J = \sqrt{\frac{5.99 \left(1 - \sum_{i=1}^I p_{i\bullet}^2 \right)}{p_{\bullet j} (n-1)(I-1)}}. \quad (9)$$

4. Porównanie wyników asymetrycznej i symetrycznej analizy korespondencji

Porównanie wyników symetrycznej i asymetrycznej analizy korespondencji jest przedstawione na podstawie danych dotyczących wyboru czasu antenowego stacji telewizyjnych zawartych w opracowaniu A. Stanimir [2006, s. 200]. Badaniem objęto 109 studentów uczelni wrocławskich. Tabela 1 przedstawia zależności zachodzące między wyborem stacji a przyciągającymi odbiorców cechami tych stacji telewizyjnych.

Tabela 1. Liczebności wyborów stacji z punktu widzenia ich cech

Cecha	TVP1	TVP2	TVN	Polsat	TeDe
Prezenterzy	1	2	2	3	3
Obraz	1	2	5	2	3
Reklamy	4	2	9	3	6
Wizerunek	2	5	6	2	5
Popularność	2	5	3	3	9
Przyzwyczajenie	9	3	4	1	2

Źródło: [Stanimir 2006].

Analiza tabeli danych została przeprowadzona na podstawie symetrycznej i asymetrycznej analizy korespondencji. Celem analizy klasycznej jest przedstawienie współzależności między cechami w układzie wielowymiarowym (dla uproszczenia wybrane zostały dwa pierwsze wymiary). Analiza asymetryczna dotyczy oceny zależności wyboru konkretnej stacji od cech poszczególnych stacji. Zmienna wybór stacji jest w tym układzie zmienną zależną, a cechy stacji telewizyjnych jej predyktorem. Tabela 2 zawiera wyniki kategoryjnej analizy wariancji dla danych nominalnych (CATANOVA).

Tabela 2. Analiza wariancji dla danych nominalnych

	MS_T	MS_B	MS_W	Tau
Zależne kolumny	0,393	0,022	0,371	0,057
Zależne wiersze	0,411	0,019	0,391	0,048

Źródło: opracowanie własne na podstawie wyników analizy Asymtab 6.

Dla 20 stopni swobody poziom p dla tej statystyki wynosi 0,21, dlatego nie ma podstaw do odrzucenia hipotezy, że rozkłady warunkowe są równe rozkładom brzegowym.

Tabela 3 zawiera porównanie wyjaśnianej zmienności tabeli danych w obu metodach. W CA jest są to odchylenia standaryzowane statystyki χ^2 , a w NSCA tabelą wejściową są centrowane profile zmiennej niezależnej (wierszy). Obliczony na podstawie tabeli danych współczynnik $\tau_a = 0,057$ wskazuje na niewielką zależność między zmiennymi.

Struktura wartości własnych w obu analizach jest bardzo podobna. Należy jednak zwrócić uwagę, że w analizie korespondencji łączna bezwładność tabeli wynosi 0,23, a wartość wskaźnika związku predykcyjnego analizie asymetrycznej to zaledwie 0,057. Wnioskować należy, że cechy stacji w bardzo niewielkim stopniu przewidyują wybory poszczególnych stacji.

Tabele 4 i 5 przedstawiają porównanie wskaźników jakości odwzorowania kategorii wierszy i kolumn analizy symetrycznej i asymetrycznej.

Tabela 3. Struktura wartości własnych

Analiza korespondencji		Asymetryczna analiza korespondencji	
Łączna bezwładność = 0,23		Współczynnik $\tau = 0,057$, łączna bezwładność = 0,04 (= $2 * MS_B$)	
Wartości własne	Wyjaśniana bezwładność (\square)	Wartości własne	Wyjaśniana predykcja (\square)
0,39	65%	0,027	62%
0,22	22%	0,012	27%
0,14	8%	0,002	6%
0,10	5%	0,002	5%

Źródło: opracowanie własne.

Tabela 4. Struktura wkładu wymiarów w wyjaśnianiu inercji i τ w CA i NSCA dla kolumn [inercja]

Analiza kolumn	Analiza korespondencji			Asymetryczna analiza korespondencji		
	względny wkład kolumny w inercję układu (χ^2)	względny wkład wymiaru 1 w bezwładność	względny wkład wymiaru 2 w bezwładność	względny wkład kolumny w inercję układu (τ)	względny wkład wymiaru 1 w predykcję zmiennej zależnej	względny wkład wymiaru 2 w predykcję zmiennej zależnej
TVP 1	0,4972	0,742	0,04	0,445	0,956	0,038
TVP 2	0,089	0,016	0,193	0,079	0,137	0,380
TVN	0,150	0,000	0,670	0,205	0,002	0,982
Polsat	0,114	0,081	0,000	0,075	0,427	0,000
TeDe	0,150	0,162	0,095	0,197	0,772	0,113

Źródło: opracowanie własne.

Z porównania analiz korespondencji wynika, że struktura wkładu w wyjaśnienie bezwładności tabeli i współczynnika τ jest podobna. Analiza asymetryczna nieco silniej różnicuje wkład poszczególnych wymiarów w predykcję kolumn. Pierwszy wymiar najlepiej przewiduje wybory TVP 1 i TeDe, drugi natomiast – wybory TVN.

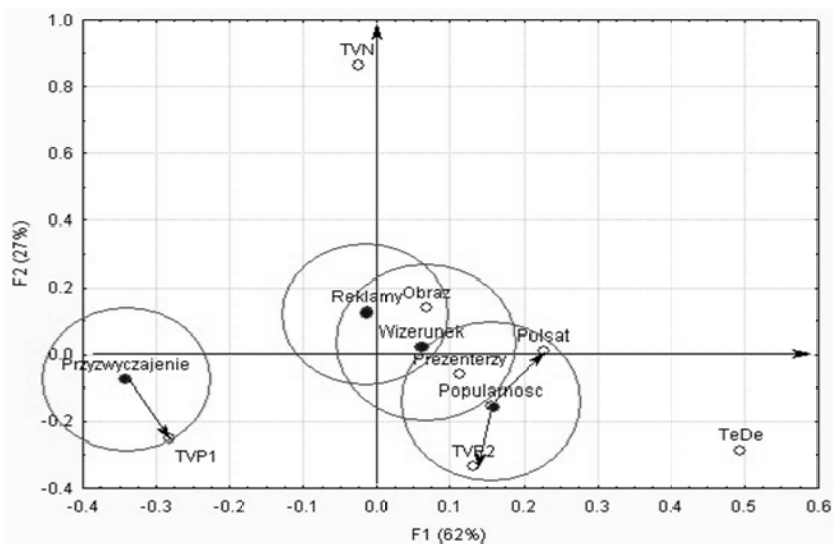
Dla punktów wierszowych (zmiennych niezależnych) największy wkład w wyjaśnienie bezwładności (lub τ) tabeli ma popularność (moda) danego kanału i przyzwyczajenie widzów.

Finalnym efektem analizy jest graficzna prezentacja zależności między wierszami a kolumnami tabeli danych. Rysunek 2 przedstawia wykres analizy korespondencji z uwzględnionymi okręgami ufności dla predyktorów. Wykres potwierdza wyniki wcześniejszych analiz. Większość przedziałów ufności przecina początek układu

Tabela 5. Struktura wkładu wymiarów w wyjaśnianie inercji i τ w CA i NSCA dla wierszy

Analiza wierszy	Analiza korespondencji	Asymetryczna analiza korespondencji
	względny wkład wierszy w inercję układu (χ^2)	względny wkład wierszy w inercję układu (τ)
Prezenterzy	0,101	0,079
Obraz	0,061	0,067
Reklamy	0,089	0,100
Wizerunek	0,060	0,054
Popularność	0,185	0,225
Przyzwyczajenie	0,503	0,474

Źródło: opracowanie własne.



Rys. 2. Graficzna prezentacja wyników asymetrycznej analizy korespondencji wraz z okręgami ufności dla predyktorów

Źródło: opracowanie własne.

współrzędnych, co wskazuje, że profile warunkowe tych zmiennych nie różnią się istotnie od profili brzegowych. Jedynie kategorie „przyzwyczajenia” i „popularność” istotnie przewidują wybory TVP1 oraz TVP2 i Polsatu².

² Dla czytelności rysunku przedstawione zostały okręgi ufności tylko dla wybranych zmiennych niezależnych. Wszystkie pozostałe zależności również okazały się nieistotne statystycznie. Bardziej poprawne w tej analizie byłoby przedstawienie elips ufności z powodu dużych różnic w zakresie wyjaśnianej predykcji (i bezwładności) przez wymiar 1 i 2.

5. Podsumowanie

Asymetryczna analiza korespondencji jest metodą mającą zastosowanie do wizualizacji zależności między zmienną niezależną a zmienną zależną w tabelach danych. Pozwala na bardziej jednoznaczne, w porównaniu do klasycznej analizy korespondencji, określenie zależności predykcyjnych w tabeli kontyngencji i jest bardziej właściwą metodą do analizy tabel rozproszonych, zbudowanych na podstawie mniej licznych prób. Zastosowanie NSCA w badaniach marketingowych, które w większości przypadków dotyczą układów zależnościowych i predykcyjnych między zmiennymi, pozwala więc na istotną poprawę interpretacji uzyskiwanych wyników.

Literatura

- Beh E.J., D'Ambrà L., *Some interpretative tools for non-symmetrical correspondence analysis*, „Journal of Classification” 2009, vol. 26, s. 55–76.
- Borg I., Groenen P.J.F., *Modern Multidimensional Scaling*, Springer, New York 2005.
- Constantine A.G., Gower J. C., *Graphic representation of asymmetric matrices*, „Applied Statistics” 1978, vol. 27, s. 297–304.
- D'Ambrà L., Lauro N.C., *Non-symmetrical exploratory data analysis*, „Statistica Applicata” 1992, vol. 4, s. 511–529.
- Esposito V., Balbi S., *Simultaneous non-symmetrical principal component analysis with group structure*, „Applied Stochastic Models in Business and Industry” 1999, vol. 15, s. 301–309.
- Rao C.R., *The use and interpretation of principal components in applied research*, „Sankhya” 1964, vol. 26, s. 329–357.
- Sarnacchiaro P., D'Ambrà A., *Explorative data analysis and CATANOVA for ordinal variables: An integrated approach*, „Journal of Applied Statistics” 2007, vol. 34, no. 9, s. 1035–1050.
- Stanimir A., *Analiza korespondencji*, [w:] A. Stanimir (red.), *Analiza danych marketingowych. Problemy, metody, przykłady*, AE, Wrocław 2006.
- Vinzi E., *Explanatory methods for comparative analyses*, „Chemometrics and Intelligent Laboratory Systems” 2001, vol. 58, s. 275–286.
- Zielman B., Heiser W.J., *Models of asymmetric proximities*, „British Journal of Mathematical and Statistical Psychology” 1996, vol. 49, s. 127–146.

ASYMMETRIC MULTIVARIATE METHODS IN MARKETING RESEARCH

Summary: The aim of the paper is to present the marketing applications of asymmetric multivariate methods, particularly **asymmetric correspondence analysis and the analysis of variance for nominal data (CATANOVA)**. Those methods allow for a correct analysis of dependency structures between variables, based on the decomposition of asymmetrical association coefficients in the contingency tables (τ Goodman-Kruskal and Marcotorchino, Simonetti Δ or Gray-Williams Index). The application of these methods may also help in a proper analysis of dependence relations based on relatively small samples.

Keywords: asymmetric correspondence analysis, analysis of variance for nominal data (CATANOVA).