

Bartłomiej Jefmański

PRZEGLĄD ZASTOSOWAŃ ANALIZY DYSKRYMINACYJNEJ W BADANIACH MARKETINGOWYCH

1. Wstęp

Spośród wielowymiarowych metod statystycznej analizy danych w literaturze dotyczącej badań marketingowych najczęściej wymieniane są: metody klasyfikacji i porządkowania liniowego, skalowanie wielowymiarowe, analiza czynnikowa, drzewa klasyfikacyjne, analiza conjoint, analiza regresji wielorakiej, wielowymiarowa analiza wariancji i kowariancji, analiza korelacji kanonicznej, analiza dyskryminacyjna (zob. [43, s. 74]). W polskiej literaturze przedmiotu można znaleźć stosunkowo niewiele opracowań dotyczących praktycznych zastosowań analizy dyskryminacyjnej. W literaturze zagranicznej jest ona znacznie częściej wskazywana oraz wykorzystywana jako użyteczna metoda w rozwiązywaniu wybranych problemów marketingowych.

Celem niniejszego opracowania jest przegląd i analiza stanu badań w zakresie aplikacji liniowych funkcji dyskryminacyjnych w badaniach marketingowych.

2. Idea i ogólna charakterystyka analizy dyskryminacyjnej

Zagadnienie analizy dyskryminacyjnej jest stosunkowo dobrze rozpoznane szczególnie w literaturze anglojęzycznej. Na uwagę zasługują prace m.in. takich autorów, jak: Cacoullos (1973), Eisenbeis, Avery [6], Huberty [13], Johnson [18], Klecka [19], Lachenbruch [24], McLachlan [28], Webb (2002). W literaturze polskiej spośród prac poświęconych analizie dyskryminacyjnej na szczególną uwagę zasługuje monografia Krzyżski [21] oraz prace Kolonki [20] i Perkala [32].

Termin analiza dyskryminacyjna związany jest z kilkoma bliskimi sobie zagadnieniami statystycznymi. Jak podkreśla Klecka [19], nie wszystkie muszą być wy-

korzystywane w ramach konkretnych badań, dlatego proponuje on w swoim opracowaniu podzielenie analizy dyskryminacyjnej na dwa zagadnienia. Podział ten znajduje potwierdzenie także w innych opracowaniach dotyczących metod analizy dyskryminacyjnej. Zazwyczaj wskazuje się w nich na jej dwa cele: opisowy (*descriptive*) i predyktywny (*predictive*) (por. [13, s. 28-31; 17, s. 427-428; 25, s. 427]).

Na opisowy cel analizy dyskryminacyjnej (*descriptive discriminant analysis*) zwracają uwagę w swoich opracowaniach m.in.: Hair i in. [11], Huberty [13], Klecka [19], Lattin, Carroll, Green [25]. Opisowa analiza dyskryminacyjna umożliwia interpretację różnic między populacjami, a konkretnie znalezienie zmiennych różnicujących populacje, określenie ich mocy dyskryminacyjnej oraz wybranie tych, które dyskryminują populacje w sposób najbardziej istotny (zob. [19, s. 8-9]). Ponadto w ramach opisowego nurtu analizy dyskryminacyjnej bada się, czy różnice między populacjami są istotne statystycznie (zob. [25, s. 427]). W takim podejściu można więc mówić o analizie dyskryminacyjnej jako statystycznej metodzie wielowymiarowej umożliwiającej testowanie hipotez statystycznych o równości wektorów wartości średnich w poszczególnych populacjach (zob. [11, s. 182]). Statystyki wykorzystywane do weryfikacji tego typu hipotez statystycznych zostały zaczerpnięte z metod wykorzystywanych na potrzeby wielowymiarowej analizy wariancji (*MANOVA*).

Predyktywny cel analizy dyskryminacyjnej (*predictive discriminant analysis*) podkreślany jest m.in. w pracach takich autorów, jak: Everitt, Dunn [5], Grabiński [9], Hadasik [10], Kolonko [20], Krzyśko [21], Lachenbruch [24], Sitgreaves [40], Weiss, Kulikowski [44]. Predyktywna analiza dyskryminacyjna dotyczy problemu decyzyjnego, którego podstawowym zadaniem jest przyporządkowanie nieznaney dotąd obserwacji (obiektu) do jednej z kilku rozłącznych populacji (kategorii, grup, klas) przy jak najmniejszym błędzie klasyfikacji (por. [10, s. 100; 24, s. 1]). Inaczej mówiąc, poszukuje się formuły matematycznej umożliwiającej identyfikację przynależności niesklasyfikowanej obserwacji do jednej z wcześniej wyróżnionych populacji. Formuła przynależności konstruowana jest na podstawie tzw. próby uczącej zawierającej wielowymiarowe obserwacje, których przynależność do odpowiednich populacji jest znana *a priori* (zob. [5, s. 248; 7, s. 4; 27, s. 170]).

Należy podkreślić, że w badaniach empirycznych analiza dyskryminacyjna wykorzystywana jest często jednocześnie do celów zarówno interpretacji, jak i klasyfikacji.

Analiza dyskryminacyjna wymaga znalezienia liniowych kombinacji zmiennych niezależnych, które najlepiej rozróżnią ustalone *a priori* grupy obiektów. Liniowa kombinacja zmiennych niezależnych określana jest mianem funkcji dyskryminacyjnej (zob. [11, s. 182]). Jej postać uogólnioną dla przypadku p zmiennych niezależnych zaprezentowano poniżej:

$$Y = a_1X_1 + a_2X_2 + \dots + a_pX_p, \quad (1)$$

gdzie: $X_1, X_2, \dots, X_p$ – zmienne niezależne, dla których dokonywana jest dyskryminacja,

$a_1, a_2, \dots, a_p$ – współczynniki dyskryminacyjne.

Funkcje dyskryminacyjne wyznaczane są w sposób umożliwiający maksymalizację zróżnicowania ich wartości między grupami w stosunku do zróżnicowania wewnątrz grup. Inaczej mówiąc, poszukuje się takich współczynników wektora $\mathbf{a}$, które zapewnią maksymalizację następującego ilorazu (zob. np. [27, s. 575; 15, s. 130]):

$$\lambda = \frac{\mathbf{a}^T \mathbf{B} \mathbf{a}}{\mathbf{a}^T \mathbf{W} \mathbf{a}}, \quad (2)$$

gdzie: $\mathbf{a}$ – wektor kolumnowy współczynników dyskryminacyjnych,

$$\mathbf{B} = \sum_{g=1}^k n_g (\bar{\mathbf{X}}_{(g)} - \bar{\mathbf{X}})(\bar{\mathbf{X}}_{(g)} - \bar{\mathbf{X}})^T - \text{międzygrupowa macierz sum kwadratów i iloczynów mieszanych (macierz rozrzutu międzygrupowego),}$$

$$\mathbf{W} = \sum_{g=1}^k \sum_{i=1}^{n_g} (\mathbf{X}_i - \bar{\mathbf{X}}_{(g)})(\mathbf{X}_i - \bar{\mathbf{X}}_{(g)})^T - \text{wewnątrzgrupowa macierz sum kwadratów i iloczynów mieszanych (macierz rozrzutu wewnątrzgrupowego).}$$

Budowa liniowej funkcji dyskryminacyjnej wymaga poczynienia i weryfikacji kilku istotnych założeń (zob. [19, s. 8-10]):

1. Obiekty należą do jednej z dwóch lub większej liczby grup. Grupy winny być zdefiniowane w taki sposób, aby każdy z obiektów należał wyłącznie do jednej z nich.
2. Pomiar zmiennych niezależnych wybranych do dyskryminowania grup (zmiennych dyskryminacyjnych) dokonywany jest na skali przedziałowej bądź ilorazowej (równania matematyczne wykorzystywane w ramach analizy dyskryminacyjnej wymagają zmiennych mierzonych na silnych skalach pomiaru celem oszacowania np. średniej arytmetycznej, wariancji itp.). W wielu problemach marketingowych bazuje się na zmiennych mierzonych za pomocą słabych skal pomiaru. W takiej sytuacji skale pomiarowe są sztucznie wzmacniane, ale, jak podkreśla Walesiak [43], sposób ten jest nie do przyjęcia ze względów metodologicznych.
3. Żadna ze zmiennych niezależnych nie może stanowić liniowej kombinacji pozostałych zmiennych.
4. Macierz wariancji-kowariancji dla każdej z grup jest taka sama (takie założenie pozwala uprościć formuły matematyczne wykorzystywane na potrzeby oszacowania współczynników dyskryminacyjnych). Witkowska [46] wskazuje, że założenie to może zostać pominięte w zastosowaniach praktycznych analizy

dyskryminacyjnej przy odpowiednio dużej liczebności próby losowej pobranej z badanej populacji.

5. Każda z grup obiektów charakteryzuje się wielowymiarowym rozkładem normalnym (założenie to pozwala na przeprowadzanie testów statystycznych za pomocą statystyk wykorzystywanych na potrzeby analizy dyskryminacyjnej). W pracach m.in.: Morrisona [30], Sokołowskiego [41], Handa [12] podkreśla się, że funkcja dyskryminacyjna jest dobrą statystyką klasyfikacyjną, pomimo niespełnienia tego założenia (zob. [46, s. 86]).

Należy podkreślić, że w badaniach empirycznych rzadko spełnione są wszystkie z wyszczególnionych warunków stawianych zgromadzonym danym. Jak opisać, często jednak w literaturze przedmiotu wskazuje się na odporność analizy dyskryminacyjnej na niektóre spośród stawianych jej wymogów.

Analiza dyskryminacyjna jest wykorzystywana w wielu dziedzinach nauki. Na gruncie ekonomii najczęściej kojarzona jest z tzw. modelami scoringowymi umożliwiającymi prognozowanie upadłości przedsiębiorstw. Nie bez przyczyny, gdyż jednym z pierwszych i najbardziej znanych zastosowań analizy dyskryminacyjnej był tzw. model Altmanna, który cechował się bardzo wysoką trafnością uzyskiwanych na jego podstawie prognoz. Jak dowodzi w swoim opracowaniu Hadasik [10], pomimo wielu zarzutów formalnych stawianych analizie dyskryminacyjnej, trafność uzyskiwanych na podstawie funkcji dyskryminacyjnych prognoz jest nie gorsza, a w wielu przypadkach lepsza niż otrzymana na podstawie innych bardziej wyrafinowanych i skomplikowanych metod.

Wydaje się, że drugą w kolejności dziedziną ekonomii ze względu na częstość praktycznego wykorzystywania analizy dyskryminacyjnej są badania marketingowe.

3. Zastosowania analizy dyskryminacyjnej w badaniach marketingowych

W celu charakterystyki wybranych zastosowań analizy dyskryminacyjnej w badaniach marketingowych zaprezentowanych w tab. 1 wykorzystano opracowania publikowane w takich czasopismach, jak: „European Journal of Marketing”, „Journal of Business”, „Journal of Marketing Research”, „Journal of Real Estate Research”, „Construction Management and Economics”, „The Service Industries Journal” oraz zamieszczone w Internecie.

Tabela 1. Wybrane zastosowania analizy dyskryminacyjnej w badaniach marketingowych

Autor	Cel badania	Charakterystyka badania
Evans (1959)	Identyfikacja czynników decydujących o posiadaniu jednej z dwóch marek samochodów osobowych: Forda i Chevroleta	Próba badawcza: 140 respondentów, z czego 71 posiadało Forda, a 69 – Chevroleta. Zmienne: Posiadacze analizowanych marek samochodów zostali scharakteryzowani za pomocą zmiennych o charakterze demograficznym i socjoekonomicznym
Robertson, Kennedy (1968)	Identyfikacja czynników różnicujących konsumentów chętnie nabywających nowe produkty pojawiające się na rynku oraz konsumentów, którzy nowych produktów nie nabywają	Próba badawcza: 100 respondentów, z czego 60 stanowiło grupę osób nabywających nowe produkty pojawiające się na rynku a pozostałe 40 grupę konsumentów, którzy nie nabywają nowych produktów Zmienne: Konsumentki poproszeni zostali o ustosunkowanie się do pytań dotyczących m.in. ich skłonności do zakupu produktu bezpośrednio po jego pojawieniu się na rynku. Ponadto w badaniu wykorzystano zmienne dotyczące okresu zamieszkiwania pod aktualnym adresem czy też częstotliwości zmiany miejsca pracy
Shuchman (1975)	Identyfikacja czynników różnicujących gospodarstwa domowe lojalne oraz nielojalne w nabywaniu pasty do zębów marki Crest	Próba badawcza: Badaniem objęto 1674 gospodarstwa domowe, z czego połowę stanowiły gospodarstwa lojalne wobec badanej marki, a drugą połowę gospodarstwa nielojalne Zmienne: Gospodarstwa zostały scharakteryzowane za pomocą 6 zmiennych demograficznych, 4 zmiennych socjoekonomicznych (np. dochód, wykształcenie) oraz 6 zmiennych dotyczących m.in. liczby marek past wybieranych przez gospodarstwo czy też liczby punktów sprzedaży, w których gospodarstwo nabywa pastę do zębów
Montgomery (1975)	Identyfikacja zmiennych mających istotny wpływ na podejmowanie decyzji o zaopatrzeniu supermarketu w nowy produkt	Próba badawcza: Badaniem objęto 124 nowe produkty, które zostały poddane ocenie przez pracowników supermarketów Zmienne: Produkty scharakteryzowano za pomocą 18 zmiennych niezależnych, takich jak np.: jakość, cena, opakowanie, innowacyjność, marka
Villani (1975)	Dostarczenie informacji reklamodawcom na temat wybranych cech osobowości i stylu życia widzów określonych audycji telewizyjnych	Próba badawcza: Badaniem objęto 955 losowo wybranych kobiet przy czym do oszacowania funkcji dyskryminacyjnej wykorzystano 655 respondentów. Informacje na temat pozostałych uczestniczek badania wykorzystano w próbie walidacyjnej Zmienne: Respondentki scharakteryzowano za pomocą zmiennych niezależnych, w skład których weszły m.in. zmienne demograficzne oraz zmienne dotyczące stylu życia

Tab. I cd.

Riordan (1977)	Celem badania było znalezienie zmiennych różnicujących następujące zestawienia: (1) nabywców i osób, które nie wykupiły polisy ubezpieczeniowej, (2) agentów ubezpieczeniowych i nabywców polis ubezpieczeniowej na życie, (3) agentów ubezpieczeniowych i osób, które nie nabyły polisy	Próba badawcza: W badaniu uczestniczyło 186 respondentów: 90 osób, które nabyły polisę; 96 osób, które polisy nie nabyły; 19 agentów ubezpieczeniowych Zmienne: Respondenci musieli ustosunkować się do 21 stwierdzeń zawartych w kwestionariuszu ankiety. Stwierdzenia dotyczyły np. ogólnej oceny agentów ubezpieczeniowych, produktu ubezpieczeniowego, jakim jest polisa ubezpieczeniowa na życie, sprzedaży polisy ubezpieczeniowej wśród studentów, oceny oferty firmy ubezpieczeniowej na tle ofert innych firm
Calantone (1978)	Identyfikacja zmiennych różnicujących 5 segmentów klientów bankowych (segmenty wydzielone w wyniku zastosowania analizy czynnikowej i jednej z metod klasyfikacji)	Próba badawcza: Badaniem objęto 344 klientów bankowych Zmienne: Respondenci zostali scharakteryzowani na podstawie ich opinii o ważności wybranych cech w podejmowaniu decyzji o wyborze banku. Były to następujące zmienne: rozmiar banku, zróżnicowanie oferty bankowej, reklama, dostępność placówek bankowych, opinia o banku, oprocentowanie oszczędności, nowoczesność, jakość obsługi, wypłacalność finansowa banku, godziny otwarcia banku, zaangażowanie w życie lokalnej społeczności, liczba miejsc parkingowych, atmosfera panująca w banku, gotowość do udzielenia pożyczki, szybka obsługa, wysokość oprocentowania pożyczek, dostępność banku (w rozumieniu, że oferta banku ma charakter uniwersalny i nie jest kierowana do ściśle określonej grupy klientów)
Westbrook (1979)	Identyfikacja zmiennych różnicujących 4 segmenty nabywców dóbr trwałego użytku	Próba badawcza: Badaniem objęto 236 nabywców detalicznych dóbr trwałego użytkowania (artykułów gospodarstwa domowego, takich jak lodówka, zmywarka np.) Zmienne: Respondenci zostali scharakteryzowani za pomocą takich zmiennych niezależnych, jak: wiek, wykształcenie, satysfakcja z poprzedniego produktu tej samej marki, aktualny stan użytkowanego dobra (np. w pełni sprawny lub zepsuty), sposób podjęcia decyzji zakupu, liczba marek branych pod uwagę w trakcie dokonywania zakupu
Priscilla (1983)	Odpowiedź na pytanie, czy klienci lojalni wobec danej marki (czyli tacy, którzy w okresie badania dokonywali powtórnego zakupu produktu tej samej marki) różnią się ze względu na przyjęte zmienne niezależne od klientów nielojalnych, czyli takich, którzy w okresie badania dokonali zakupu produktu innej marki	Próba badawcza: Z grupy 500 losowo dobranych osób, 190 zadeklarowało chęć uczestnictwa w badaniu. Ostatecznie wyniki badań oparto na podstawie kwestionariuszy ankiet wypełnionych przez 87 ankietowanych Zmienne: Respondenci zostali scharakteryzowani za pomocą takich zmiennych niezależnych, jak: poziom satysfakcji z użytkowania produktu danej marki, prawdopodobieństwo zakupu poprzednio nabywanej marki produktu, prawdopodobieństwo ponownego zakupu aktualnie preferowanej marki produktu

Tab. 1 cd.

Pessemier, Handelsman (1984)	Znalezienie i zbadanie wpływu wybranych czynników na zmienność zachowań konsumentów w nabywaniu trzech grup produktów. Autorzy opracowania zastosowali indeks ITV (<i>index of temporal variety</i>) jako miarę ogólnej zmienności zachowań nabywczych konsumentów	Próba badawcza: Na podstawie wartości ITV dla każdej z trzech grup produktów otrzymano dwie klasy respondentów o niskiej i wysokiej zmienności zachowań nabywczych Zmienne: Respondentów scharakteryzowano za pomocą 118 zmiennych niezależnych: 26 zmiennych demograficznych, 62 zmiennych charakteryzujących postawy wobec marki produktów oraz 30 zmiennych dotyczących zakupu i użytkowania produktu
Sinclair, Stalling (1990)	Identyfikacja czynników mających wpływ na wybór materiałów budowlanych przez firmy budowlane zgrupowane w jednym z czterech segmentów	Próba badawcza: 412 przedsiębiorstw budowlanych Zmienne: Każda z firm poproszona została o określenie ważności każdego z 23 czynników (pomiaru dokonano na 5-stopniowej skali pomiaru) w wyborze takich produktów budowlanych, jak np. płyta pilśniowa, płyta winylowa, aluminium itp. Oceniane czynniki dotyczyły m.in. wytrzymałości, łatwości montażu, konkurencyjności cenowej
Crook (1992)	Identyfikacja czynników, które różnicują dwie grupy posiadaczy kart kredytowych (użytkowników i tych, którzy kart nie wykorzystują)	Próba badawcza: Badaniem objęto 825 respondentów (679 użytkowników kart kredytowych i 146 respondentów, którzy mają kartę, ale jej nie używają) Zmienne: Respondentów scharakteryzowano za pomocą 24 zmiennych niezależnych (zmiennych socjodemograficznych i ekonomicznych)
Hair i in. (1995)	Identyfikacja zmiennych różnicujących dwie grupy przedsiębiorstw wyróżnione ze względu na sposób ich zaopatrzenia	Próba badawcza: Przedmiotem badania było 100 przedsiębiorstw Zmienne: Przedsiębiorstwa scharakteryzowano za pomocą 7 zmiennych niezależnych takich, jak np. cena produktu, jakość produktu, szybkość dostawy produktu
Malhotra, Birks (2002)	Identyfikacja cech najsilniej różnicujących rodziny uczestniczące w urlopie spędzonym w miejscowości wypoczynkowej (w ciągu ostatnich 2 lat) od pozostałych rodzin	Próba badawcza: Badaniem została objęta próba składająca się z 30 rodzin Zmienne: Rodziny scharakteryzowane zostały za pomocą 5 zmiennych niezależnych. Były to: roczny dochód gospodarstwa domowego, stosunek do podróżowania, waga przywiązana do rodzinnych wakacji, liczba członków gospodarstwa domowego, wiek głowy rodziny

Tab. 1 cd.

Kim (2000)	Ustalenie profilu trzech segmentów: nabywców domów jednorodzinnych, nabywców apartamentów oraz osób, które nie nabyły żadnego lokalu mieszkalnego	Próba badawcza: W badaniu uczestniczyli właściciele nieruchomości. Przebadano 2000 respondentów z czego 1986 ankiet zostało wypełnionych poprawnie Zmienne: Respondentów scharakteryzowano za pomocą 2 zmiennych demograficznych (wiek, liczba członków rodziny), 4 zmiennych socjoekonomicznych (zawód, wykształcenie, dochód, wartość miesięcznych wydatków na życie) oraz zmiennych charakteryzujących liczbę przeprowadzek od momentu zawarcia związku małżeńskiego, warunki, na których użytkowana jest nieruchomość, liczbę przeprowadzek w okresie po zawarciu związku małżeńskiego
Gallagher, Mansour (2000)	Identyfikacja czynników różnicujących 5 segmentów rynku usług hotelarskich wydzielonych przy zastosowaniu jednej z metod klasyfikacji	Próba badawcza: Obiektami badania było 58 metropolii zlokalizowanych na terenie USA. Metropolie zostały przyporządkowane do jednego z 5 segmentów Zmienne: Metropolie zostały scharakteryzowane za pomocą zmiennych o charakterze ekonomicznym, czyli np. wzrost zatrudnienia lub liczba turystów zagranicznych na danym obszarze w przyjętym okresie
Churchill (2002)	Identyfikacja czynników różnicujących dobrych, średnich i słabych przedstawicieli handlowych	Próba badawcza: Sprzedawców firmy przemysłowej podzielono na 3 grupy w zależności od liczby pozyskanych nowych klientów w okresie trzech miesięcy. Grupę pierwszą stanowili sprzedawcy, którzy pozyskali więcej niż 15 nowych klientów, natomiast grupę drugą – sprzedawcy, którzy pozyskali przynajmniej 5 nowych klientów. W grupie trzeciej znaleźli się sprzedawcy, którym udało się pozyskać najmniej klientów. Liczebność wszystkich grup była jednakowa i wynosiła 15 Zmienne: Sprzedawcy zostali scharakteryzowani za pomocą zmiennych niezależnych takich, jak: liczba wizyt u nowych klientów, odsetek wizyt wcześniej umówionych, liczba rozmów telefonicznych z przyszłymi klientami, liczba odwiedzonych nowych klientów
Rószkiewicz (2002)	Identyfikacja cech profilowych rozstrzygających o przynależności klientów do wyróżnionych segmentów	Próba badawcza: 155 losowo wybranych klientów firmy zostało zaklasyfikowanych do jednego z dwóch segmentów odróżniających stałych nabywców danego produktu od nabywców okazjonalnych Zmienne: Respondenci zostali scharakteryzowani za pomocą trzech zmiennych niezależnych: wykształcenia, miejsca zamieszkania, stanu cywilnego

Tab. 1 cd.

Lattin, Carroll, Green (2003)	Identyfikacja czynników decydujących o kupnie lub rezygnacji z zakupu książki oferowanej przez firmę zajmującą się dystrybucją książek przez Internet	Próba badawcza: Spośród osób, które nie zamówiły książki, wylosowano 83 osoby, tworząc w ten sposób grupę pierwszą. W grupie drugiej znalazły się 83 osoby, które książkę zamówiły Zmienne: Klienci przyporządkowani do jednej z dwóch grup scharakteryzowani zostali za pomocą takich zmiennych, jak np.: wiek, płeć, całkowita kwota wydana na zamówione książki w trakcie korzystania z usług firmy, liczba miesięcy, które upłynęły od ostatniego zamówienia
Lattin, Carroll, Green (2003)	Celem badania było zdiagnozowanie, czy domy wystawione na sprzedaż w jednej z trzech dzielnic San Francisco (Los Altos, Menlo Park, Palo Alto) różnią się od siebie pod względem zmiennych je charakteryzujących	Próba badawcza: Próba liczyła 35 nieruchomości, z czego 9 sprzedawanych było w Los Altos, 13 – w Menlo Park oraz 13 – w Palo Alto Zmienne: Nieruchomości scharakteryzowano z uwzględnieniem ich ceny, liczby sypialni oraz powierzchni, jaką zajmują
Kvetny (2003)	Oszacowanie funkcji klasyfikującej umożliwiającej zaklasyfikowanie nowego produktu (na podstawie wyszczególnionych parametrów technicznych) do jednej z wyróżnionych grup	Próba badawcza: 162 rodzaje klimatyzatorów asygnowanych marką jednego z producentów (dystrybutorów) zestawiono w ramach 5 grup (marka identyfikowała grupę) Zmienne: Klimatyzatory scharakteryzowano za pomocą zmiennych o charakterze technicznym (np. waga, energochłonność) i jednej zmiennej o charakterze ekonomicznym, którą była cena
Ogunmoku, Wong (2004)	Identyfikacja czynników różnicujących przedsiębiorstwa o wysokim i niskim stopniu przystosowania do praktyk marketingowych obowiązujących na zagranicznych rynkach zbytu	Próba badawcza: 80 australijskich przedsiębiorstw eksportujących swoje towary i usługi na rynki zagraniczne Zmienne: Przedsiębiorstwa opisano za pomocą zmiennych charakteryzujących ich organizacyjny i ekonomiczny aspekt funkcjonowania (np. liczba zatrudnionych, okres funkcjonowania na rynkach zagranicznych, wykształcenie pracowników, faza cyklu życia produktu)

Źródło: opracowanie własne.

Zaprezenetowane przykłady dowodzą dużej użyteczności analizy dyskryminacyjnej na gruncie badań marketingowych. Metoda ta cieszy się uznaniem wśród badaczy już od blisko pół wieku. Należy zauważyć, że jest wykorzystywana w badaniach marketingowych prowadzonych na różnych rynkach, m.in.: motoryzacyjnym, nieruchomości, bankowym, ubezpieczeniowym, medialnym, dóbr trwałego użytku. Spektrum wykorzystania jest więc bardzo szerokie.

Cel zdecydowanej większości badań dotyczył bądź identyfikacji zmiennych profilowych, bądź ustalenia profilu segmentu w oparciu na uprzednio wybranych zmiennych profilowych. Inaczej mówiąc, każdy z przypadków dotyczył któregoś z etapów profilowania segmentów rynkowych (por. [16, s. 14]). Trudno zatem nie zgodzić się ze stwierdzeniem Lawsona [26], według którego na gruncie badań marketingowych analiza dyskryminacyjna jest najbardziej użyteczna jako metoda profilowania segmentów rynkowych.

Należy podkreślić, że w zdecydowanej większości przytoczonych badań wykorzystano sposób budowy modelu dyskryminacyjnego polegający na iteracyjnym doborze potencjalnych zmiennych profilowych. Procedura ta określana jest mianem krokowej procedury konstrukcji modelu dyskryminacyjnego i pozwala na włączenie do modelu jedynie tych zmiennych, które charakteryzują się odpowiednimi właściwościami dyskryminacyjnymi. Stosowanie takiego sposobu budowy modelu wynika z pewnością z faktu, że, jak zauważono, bardzo często celem tych opracowań był wybór zmiennych profilowych.

Cel badań w naturalny sposób implikował ich charakter w kontekście wykorzystanego nurtu analizy dyskryminacyjnej, dlatego w opracowaniach dominował nurt deskryptywny. Autorzy badań koncentrowali się głównie na interpretacji otrzymanych wyników w kontekście różnicowania segmentów. Wkład zmiennych profilowych w różnicowaniu segmentów zazwyczaj oceniano na podstawie standaryzowanych współczynników funkcji dyskryminacyjnych. Nie oznacza to jednak, że nie wykorzystywano skonstruowanego modelu do celów identyfikacji przynależności obiektów do analizowanych segmentów. Jednak można było zauważyć, że cel predyktywny dotyczył zazwyczaj jedynie weryfikacji jakości modelu i poświęcano mu mniej miejsca w opracowaniach.

W wielu opracowaniach, oprócz analizy dyskryminacyjnej, autorzy wykorzystywali również inne metody statystycznej analizy wielowymiarowej. Metody klasyfikacji pozwalały wyodrębniać segmenty rynkowe, które następnie poddawano zabiegowi profilowania. Z kolei za pomocą analizy czynnikowej redukowano zbiór potencjalnych zmiennych profilowych.

W przytoczonych przykładach zaobserwowano duże zróżnicowanie liczebności prób badawczych. Niektóre z nich liczyły nawet blisko 2000 obiektów, podczas gdy w innych wykorzystano zaledwie 30. Uogólniając, można stwierdzić, że około jedna trzecia z zaprezentowanych badań oparta została na stosunkowo mało licznych próbach, w których liczba obiektów nie przekroczyła 100.

Należy również podkreślić częste niespełnianie założeń stawianych materiałowi empirycznemu w ramach analizy dyskryminacyjnej. Ponadto w niektórych przypadkach założenia te nie zostały szczegółowo zweryfikowane.

Całkowita trafność klasyfikacji konstruowanych modeli dyskryminacyjnych w niektórych przypadkach przekraczała 80%. Jednak autorzy badań uznawali za

zadowalającą trafność klasyfikacji na poziomie już ok. 70%, szczególnie w przypadkach, gdy liczba analizowanych segmentów przekraczała 2.

4. Podsumowanie

W opracowaniu krótko scharakteryzowano analizę dyskryminacyjną oraz wskazano na jej wybrane zastosowania w badaniach marketingowych. Analiza przytoczonych opracowań upoważnia do sformułowania następujących wniosków:

- analiza dyskryminacyjna jest stosunkowo często wykorzystywaną metodą na gruncie badań marketingowych;
- pomimo różnych problemów badawczych formułowanych w przytoczonych opracowaniach, rola analizy dyskryminacyjnej sprowadza się zazwyczaj do identyfikacji zmiennych niezależnych najsilniej różnicujących badane grupy obiektów (konsumentów, produktów, przedsiębiorstw). Można zatem stwierdzić, że stosowanie tej metody jest szczególnie użyteczne w zagadnieniu profilowania segmentów rynkowych;
- w opracowaniach zazwyczaj wykorzystywany jest zarówno predyktywny, jak i opisowy wymiar analizy dyskryminacyjnej;
- funkcje dyskryminacyjne otrzymywane w zaprezentowanych przykładach charakteryzowały się wysoką trafnością klasyfikacji, pomimo częstych sytuacji niespełnienia niektórych założeń m.in. dotyczących skali pomiaru zmiennych niezależnych.

Literatura

- [1] Calantone R.J., Sawyer A.G., *The Stability of Benefit Segments*, „Journal of Marketing Research” 1978, vol. XV.
- [2] Churchill G.A., *Badania marketingowe. Podstawy metodologiczne*, Wydawnictwo Naukowe PWN, Warszawa 2002.
- [3] Crook J.N., Hamilton R., Thomas L.C., *Credit Card Holders: Characteristic of Users and Non-Users*, „The Service Industries Journal” 1992, vol. 12, nr 2.
- [4] *Discriminant Analysis and Applications*, red. T. Cacoullous, Academic Press, New York 1973.
- [5] Dunn G., Everitt B.S., *Applied Multivariate Data Analysis*, Arnold, London: 2001.
- [6] Eisenbeis R.A., Avery R.B., *Discriminant Analysis and Classification Procedures*, MA: D.C. Heath and Company, Lexington 1972.
- [7] Fernandez G.C.J., *Discriminant Analysis. A Powerful Classification Technique in Data Mining*, SAS Institute Inc., Orlando 2002.
- [8] Gallagher M., Mansour A., *An Analysis of Hotel Real Estate Market Dynamics*, „Journal of Real Estate Research” 2000, vol. 19, nr 1/2.
- [9] Grabiński T., *Metody taksonometrii*, AE, Kraków 1992.
- [10] Hadasik D., *Upadłość przedsiębiorstw w Polsce i metody jej prognozowania*, AE, Poznań 1998.

- [11] Hair J.F., Anderson R.E., Tatham R.L., Black W.C., *Multivariate Data Analysis*, Prentice-Hall, Englewood Cliffs 1995.
- [12] Hand D.J., *Discrimination and Classification*, John Wiley & Sons, Chichester 1981.
- [13] Huberty C.J., *Applied Discriminant Analysis*, John Wiley & Sons, New York 1994.
- [14] Jae-Chan K., Dae-Ho K., Jae-Jun K., Jong-Suk Y., Hyun-Soo L., *Segmenting the Korean Housing Market Using Multiple Discriminant Analysis*, „Construction Management and Economics” 2000, vol. 18.
- [15] Jajuga K., *Statystyczna teoria rozpoznawania obrazów*, Państwowe Wydawnictwo Naukowe, Warszawa 1990.
- [16] Jefmański B., *Poznaj profil swojego klienta*, „Marketing i Rynek” 2006, nr 1.
- [17] Joachimsthaler E.A., Stam A., *Mathematical Programming Approaches for the Classification Problem in Two-Group Discriminant Analysis*, „Multivariate Behavioral Research” 1990, nr 25.
- [18] Johnson R.M., *Multiple Discriminant Analysis: Applications to Marketing Research*, Market Facts, Chicago 1970.
- [19] Klecka W.R., *Discriminant Analysis*, Sage Publication, Inc., London 1980.
- [20] Kolonko J., *Analiza dyskryminacyjna i jej zastosowania w ekonomii*, Państwowe Wydawnictwo Naukowe, Warszawa 1980.
- [21] Krzyśko M., *Analiza dyskryminacyjna*, Wydawnictwo Naukowe Uniwersytetu w Poznaniu, Poznań 1982.
- [22] Kvetny M., *Using Discriminant Analysis for Classification of Home Air Conditioners Manufacturers*, „Proceedings of Student Research Works” 2003, vol. 4.
- [23] LaBarbera P.A., Mazursky D., *A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction: The Dynamic Aspect of the Cognitive Process*, „Journal of Marketing Research” 1983, vol. XX.
- [24] Lachenbruch P.A., *Discriminant Analysis*, Hafner Press, London 1975.
- [25] Lattin J., Carroll J.D., Green P.E., *Analyzing Multivariate Data*, Thomson Learning, Pacific Grove 2003.
- [26] Lawson R., *Discriminant Analysis – An Aid to Market Segment Description*, „European Journal of Marketing” 1980, vol. 14, nr 7.
- [27] Malhotra N.K., Birks D.F., *Marketing Research. An Applied Approach*, Prentice-Hall, Inc., Harlow 1999.
- [28] McLachlan G.J., *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley & Sons, New York 1992.
- [29] Montgomery D.B., *New Product Distribution: An Analysis of Supermarket Buyer Decisions*, „Journal of Marketing Research” 1975, vol. XII.
- [30] Morrison D.F., *Wielowymiarowa analiza statystyczna*, Państwowe Wydawnictwo Naukowe, Warszawa 1990.
- [31] Ogunmokun G., Wong J., *Determinants of Marketing Adaptation/Globalisation Practices of Australian Exporting Firms*, „World Review of Science, Technology and Sustainable Development” 2004, vol. 1, nr 1.
- [32] Perkal J., *Matematyka dla przyrodników i rolników. Część II*, Państwowe Wydawnictwo Naukowe, Warszawa 1963.
- [33] Pessemier E., Handelsman M., *Temporal Variety in Consumer Behavior*, „Journal of Marketing Research” 1984, vol. XXI.
- [34] Rencher A.C., *Interpretation of Canonical Discriminant Functions, Canonical Variates and Principal Components*, „American Statistical Association” 1992, vol. 46, nr 3.
- [35] Riordan E.A., Oliver R.L., Donnelly J.H., *The Unsold Prospect: Dyadic and Attitudinal Determinants*, „Journal of Marketing Research” 1977, vol. XIV.

-
- [36] Robertson T.S., Kennedy J.N., *Prediction of Consumer Innovation: Application of Multiple Discriminant Analysis*, „Journal of Marketing Research” 1968, vol. 14.
 - [37] Rószkiewicz M., *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 2002.
 - [38] Shuchman A., Riesz P.C., *Correlates of Persuasibility: The Crest Case*, „Journal of Marketing Research” 1975, vol. XII.
 - [39] Sinclair S.A., Stalling E.C., *How to Identify Differences between Market Segments with Attribute Analysis*, „Industrial Marketing Management” 1990, vol. 19, nr 1.
 - [40] Sitgreaves R., *Some Operating Characteristics of Linear Discriminant Functions*, [w:] *Discriminant Analysis and Applications*, red. T. Cacoullos, Academic Press, New York 1973.
 - [41] Sokołowski A., *Analizy wielowymiarowe*, StatSoft Polska, Kraków 1999.
 - [42] Villani K.E.A., *Personality/Life Style and Television Viewing Behavior*, „Journal of Marketing Research” 1975, vol. XII.
 - [43] Walesiak M., *Metody analizy danych marketingowych*, Wydawnictwo Naukowe PWN Warszawa 1996.
 - [44] Weiss S.M., Kulikowski C.A., *Computer System That Learn. Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems*, Morgan Kaufmann Publishers, Inc., California 1991.
 - [45] Westbrook R.A., Fornell C., *Patterns of Information Source Usage Among Durable Goods Buyers*, „Journal of Marketing Research” 1979, vol. XVI.
 - [46] Witkowska D., *Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe*, Wydawnictwo C.H. Beck, Warszawa 2002.

THE REVIEW OF APPLICATIONS OF DISCRIMINANT ANALYSIS IN MARKETING RESEARCH

Summary

Discriminant analysis is one of multivariate statistical data analysis methods used in marketing research. The aim of this paper is to present and analyse selected applications of discriminant analysis in marketing research.

Mgr Bartłomiej Jefmański jest pracownikiem Katedry Ekonometrii i Informatyki w Akademii Ekonomicznej we Wrocławiu.