PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1180 — 2007
Ekonometria 19

Aneta Rybicka

PRZYGOTOWANIE DANYCH W BADANIACH PREFERENCJI KONSUMENTÓW METODAMI WYBORÓW DYSKRETNYCH

1. Wstęp

Metody wyborów dyskretnych obejmują różnorakie techniki projektów eksperymentów, procedury gromadzenia danych oraz statystyczne procedury, które wykorzystywane są w celu określenia wyborów, jakich dokonują konsumenci (respondenci) między alternatywami (profilami). Tę grupę metod dekompozycyjnych stosujemy, gdy respondenci mają możliwość wyboru między różnymi alternatywami [14, s. 1]. Zatem dane otrzymane w wyniku zastosowań metod wyborów dyskretnych to dane o wyborach respondentów, a nie dane o ich opiniach i ocenach o przedstawianych im profilach [11, s. 94]. Projekt eksperymentu dla modelu wyboru składa się ze zbiorów profilów, z których każdy zawiera profile wyboru opisane kombinacją poszczególnych poziomów każdego z atrybutów. Respondenci proszeni są o wybranie najbardziej preferowanego profilu z każdego zbioru. Metoda ta pozwala również, w odróżnieniu od metod *Conjoint Analysis*, na rezygnację z wyboru w sytuacji, gdy żaden z oferowanych profilów nie spełnia oczekiwań respondenta.

Metody wyborów dyskretnych mają kilka zalet. Pierwszą zaletą jest wspomniana już technika badania preferencji konsumentów, która w sposób rzeczywisty przedstawia procesy zachodzące na rynku (wybór jednego ze zbioru profilów bądź rezygnacja z wyboru). Drugą zaletą jest to, że metody te pozwalają na bezpośredni pomiar udziałów w rynku. Z zaletami tymi wiążą się dwie największe wady tychże metod. Pierwszą z nich jest to, że w związku z tym, że w trakcie badania respondenci dokonują wyboru jednego z profilów bądź też rezygnują z wyboru, nie uzyskujemy informacji o preferencjach co do pozostałych profilów. Drugą zaś istotną wadą jest to, że parametry modelu szacowane są na poziomie zagregowanym

(w przekroju całej próby), co nie pozwala na bezpośrednie oszacowanie preferencji każdego z konsumentów (preferencji indywidualnych) [15, s. 6].

2. Eksperyment czynnikowy

Proces eksperymentu wyboru (*Choice Experiment*) można przedstawić w kilku etapach [6; 7; 8; 12, s. 111]:

1. Najpierw określamy, czy nasze badanie jest badaniem markowym[1] (*Branded Study, Alternative-specific, Labelled Choice Experiment*) czy też badaniem rodzajowym, ogólnym[2] (*Generic Study, Unlabelled Choince Experiment*).

2. Następnie precyzujemy atrybuty i ich poziomy. Atrybuty wykorzystane w badaniu mogą być symetryczne (o takiej samej liczbie poziomów) lub też asymetryczne (o różnej liczbie poziomów).

3. Podejmujemy decyzję o tym, czy w zbiorze profilów będzie zawarty profil stały (bazowy)[3], np. opcja rezygnacji z wyboru, bądź też profil, który w każdym zbiorze będzie opisany tymi samymi poziomami atrybutów.

4. Określamy, czy profile w zbiorach będą opisane poziomami wszystkich atrybutów (metoda wyboru ze zbiorów profilów pełnych) lub też – w przypadku dużej liczby atrybutów – poziomami tylko wybranych atrybutów (metoda wyboru ze zbiorów profilów częściowych). Decydujemy również o tym, czy zbiory będą zawierały taką samą (*Fixed Choice Set*) lub zróżnicowaną (*Variable Choice Set*) liczbę profilów, oraz o tym, czy zbiory profilów będą dzielone na bloki (w przypadku dużej liczby zbiorów).

5. Precyzujemy rozmiar badania: obliczamy liczbę zbiorów w pełnym eksperymencie, minimalną liczbę zbiorów oraz optymalną liczbę zbiorów w badaniu.

6. Przeprowadzamy eksperyment. Układy czynnikowe generujemy z wykorzystaniem jednostopniowych bądź dwustopniowych procedur (w przypadku dużej liczby zbiorów profilów, dokonujemy redukcji rozmiaru eksperymentu pełnego, wykorzystując częściowy eksperyment czynnikowy).

7. Sprawdzamy własności eksperymentu (testujemy eksperyment przed gromadzeniem danych).

8. Gromadzimy dane (przeprowadzamy badanie wśród respondentów), następnie przetwarzamy je z wykorzystaniem odpowiedniego oprogramowania komputerowego.

---

[1] Badanie (eksperyment) markowe charakteryzuje się tym, że składa się np. z nazwy marki bądź innej etykiety [8, s. 96]. Marką bądź też etykietą mogą być również inne atrybuty, np. „przeznaczenie" w przypadku badania dotyczącego wypoczynku.

[2] Badanie przedstawia eksperyment, w którym profile nie są opisane atrybutem określającym markę [8, s. 71].

[3] Jest to tzw. profil odniesienia (*Reference Profile, Based Profile*), w stosunku do którego są szacowane prawdopodobieństwa wyboru pozostałych profilów [1, s. 110].

9. Szacujemy parametry modelu (zazwyczaj modelu logitowego z wykorzystaniem metody największej wiarygodności).

10. Szacujemy prawdopodobieństwa wyboru profilów w każdym ze zbiorów oraz prawdopodobieństwa wyboru każdego profilu z całego ocenianego zestawu (szacujemy udziały w rynku). Określamy również, z wykorzystaniem ilorazu hazardu, który z parametrów (z poziomów atrybutów) wpływa stymulująco bądź też destymulująco na prawdopodobieństwo wyboru danego profilu.

3. Przygotowanie danych

Jednym z ważniejszych etapów procedury badawczej w metodach wyborów dyskretnych jest przygotowanie danych, które zostaną wykorzystane w dalszych etapach badania: estymacji parametrów modelu oraz interpretacji otrzymanych wyników.

Na etapie przygotowania danych wykorzystywane są metody wyboru ze zbiorów profilów pełnych oraz metody wyboru ze zbiorów profilów częściowych. W celu zaś wygenerowania zbiorów profilów korzysta się z jedno- i dwustopniowych procedur generowania układów czynnikowych.

Pierwsza grupa metod wyboru – **metody wyboru ze zbiorów profilów pełnych** (*Full-profiles Choice Experiments, Experimental Choice Approach*) – jest najczęściej wykorzystywaną metodą gromadzenia danych w badaniach preferencji konsumentów metodami wyborów dyskretnych. Metoda ta polega na prezentacji respondentom zbioru profilów, które opisane są wszystkimi atrybutami (po jednym z poziomów każdego atrybutu) (tab. 1). Zbiór zawiera profile charakteryzujące produkt lub usługę oraz jeden profil bazowy (stały), np. pozwalający na rezygnację z wyboru, jeśli żaden z prezentowanych profilów w zbiorze nie spełnia oczekiwań respondenta.

Tabela 1. Zbiór profilów opisanych wszystkimi atrybutami

	Marka	Kawa Jacobs	Kawa Tchibo	Kawa Elite
Atrybuty	rodzaj	kawa mielona	kawa rozpuszczalna	kawa mielona
	opakowanie	opakowanie próżniowe	opakowanie szklane	opakowanie miękkie
	rodzaj	kawa kofeinowa	kawa bezkofeinowa	kawa kofeinowa
	waga	opakowanie 500 g	opakowanie 200 g	opakowanie 250 g
	cena	15,50 zł	23,20 zł	6,50 zł
	smak	kawa lekko palona	kawa mocno palona	kawa mocno palona

Źródło: opracowanie własne.

Druga grupa metod wyboru to **metody wyboru ze zbiorów profilów częściowych** (*Partial Profiles Choice Experiments*). Polegają one na prezentacji respondentom zbioru profilów, które są opisane tylko wybranymi atrybutami (po

jednym z poziomów z każdego wybranego atrybutu). Do badania zazwyczaj wybiera się od 2 do 5 atrybutów spośród całego zestawu (tab. 2).

Niektóre badania wymagają wykorzystania większej liczby atrybutów (czasami nawet więcej niż 20). Badania takie mogłyby być uciążliwe dla respondentów, ponieważ byliby zmuszeni (w trakcie dokonywania wyboru) do rozpatrzenia równocześnie wielu atrybutów. Dlatego też zaproponowano metodę wyborów ze zbiorów profilów częściowych, która pozwala na wykorzystanie w badaniu tylko wybranych atrybutów. Stosowanie tej metody tłumaczy się również tym, że w trakcie dokonywania wyborów na rynku ostateczny zbiór kilku produktów lub usług rzadko różni się wszystkimi atrybutami. Bardzo często produkty lub usługi pod względem wielu atrybutów są tak samo satysfakcjonujące, a różnią się w rzeczywistości tylko kilkoma najważniejszymi atrybutami [2, s. 20].

Tabela 2. Zbiór profilów opisanych wybranymi atrybutami

Marki		Kawa Jacobs	Kawa Tchibo	Kawa Elite
Atrybuty	rodzaj	kawa mielona	kawa rozpuszczalna	kawa mielona
	waga	opakowanie 500 g	opakowanie 200 g	opakowanie 250 g

Źródło: opracowanie własne.

Metoda wyborów ze zbiorów profilów częściowych zachowuje wiele cech i zalet metody wyborów ze zbiorów profilów pełnych. Metoda ta charakteryzuje się również innymi, pożądanymi cechami [2, s. 20]:

- każdemu respondentowi wszystkie atrybuty prezentowane są z taką samą częstotliwością (tak samo często);
- atrybuty wykorzystywane w badaniu są tak samo często przydzielane do pierwszego produktu lub usługi co do drugiego itd. Pozwala to na automatyczną kontrolę porządku w przeprowadzanej prezentacji wyboru;
- poszczególne wersje ankiet są generowane w prosty sposób, a respondent otrzymuje tylko jedną wersję ankiety;
- w całej próbie respondentów każda możliwość kombinacji atrybutów w profilach jest prezentowana tak samo często. To gwarantuje, że wyniki badania nie będą niesymetryczne;
- wybory dokonywane przez respondentów (pytania o preferencje) nie pociągają za sobą zdominowanych alternatyw. Oznacza to, że w każdym zbiorze profilów jeden z profilów wyboru jest lepszy od pozostałych przynajmniej pod względem jednego z atrybutów.

Eksperyment czynnikowy, wykorzystujący metodę wyborów ze zbiorów profilów częściowych, można generować na kilka sposobów. W swoich pracach Chrzan i Elrod [2] oraz Kuhfeld [7] przedstawiają metody wyboru jednego z dwóch (w zbiorze) profilów częściowych (*Pair-wise Partial Profile Choice Design,*

Partial Profile with Pairwise Choice Question), metody wyboru z liniowych zbiorów profilów częściowych (*Linear Partial Profile Design*), metody wyboru jednego z trzech (w zbiorze) profilów częściowych (*Choice from Triples Partial Profile*) oraz metody wyboru spośród sześciu alternatyw. Jednak najbardziej popularną metodą wyboru spośród zbiorów profilów częściowych jest metoda wyboru spośród dwóch profilów [5, s. 8].

Przykład eksperymentu z wykorzystaniem profilów częściowych, w którym wykorzystano 20 atrybutów dwupoziomowych, przy 5 atrybutach zmieniających się cyklicznie, przedstawia tab. 3 W poniższym układzie czynnikowym nie został zawarty profil stały.

Tabela 3. Projekt układu czynnikowego do metod wyboru ze zbiorów profilów częściowych

Profile	Atrybuty																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	.	.	.	.	.	.	.	.	.	2	.	2	.	2	.	1	.	.	2	.
2	.	.	.	2	1	.	.	.	.	.	.	.	.	2	.	.	2	2	.	.
3	.	.	1	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	2	1
4	.	.	2	.	2	.	.	.	.	.	.	1	.	.	.	.	1	1	.	.
5	1	.	.	.	2	.	.	.	.	.	.	.	.	1	.	2	.	.	.	1
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
36	.	.	.	.	.	.	.	2	.	.	1	.	.	.	.	1	.	2	1	.
37	.	.	.	1	.	.	.	1	.	.	2	1	.	1	.	.	.	.	.	.
38	.	.	.	.	.	.	.	.	.	1	2	.	.	.	1	2	1	.	.	.
39	.	.	.	.	.	2	.	.	.	.	.	.	.	.	2	.	.	1	2	2
40	.	.	.	2	.	.	.	2	.	2	.	.	1	.	2	.	.	.	.	.

Źródło: [7, s. 331-332].

Zbiory profilów wyboru (układy czynnikowe) w metodach wyborów dyskretnych można generować na dwa sposoby: z wykorzystaniem procedur jedno- oraz dwustopniowych.

Procedury **jednostopniowe**[4] pozwalają na wygenerowanie cząstkowego układu czynnikowego. Procedura ta nazywana jest strategią L^{MN} [3, s. 165]. W planie tym N oznacza liczbę profilów w każdym z podzbiorów, M liczbę wszystkich atrybutów, L zaś oznacza liczbę poziomów każdego z atrybutów. Po zastosowaniu tej procedury jednostopniowej otrzymamy układ czynnikowy o $N \times M$ kolumn oraz L wierszy. Przyjmując, że liczba profilów w każdym z podzbiorów wynosi 3 ($N = 3$), liczba poziomów każdego atrybutu wynosi 3 ($L = 3$), natomiast liczba wszystkich

[4] Po raz pierwszy procedurę tę zastosowano w Australii w latach 1978-1980, następnie zaś na Uniwersytecie Iowa w latach 1980-1981 [13, s. 203].

atrybutów wynosi 4 ($M = 4$), otrzymujemy $3^{4\times3}$ podzbiorów w układzie czynnikowym (12 kolumn i 3 wiersze – w jednym zbiorze). W tak przyjętym układzie czynnikowym minimalna liczba podzbiorów wynosi 27 [3, s. 165]. Tabela 4 przedstawia tak przygotowany przykładowy układ czynnikowy[5].

Tabela 4. Cząstkowy układ czynnikowy typu L^{MN} dla $3^{4\times3}$

Zbiór	Profil 1				Profil 2				Profil 3			
	1	2	3	4	5	6	7	8	9	10	11	12
1	1	2	1	3	2	2	3	1	1	3	3	2
2	3	2	3	1	3	1	3	1	2	3	1	2
3	1	2	3	3	2	3	2	2	2	1	2	3
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
27	3	3	2	1	3	2	1	2	1	1	3	3

Źródło: [3, s. 165].

Procedura typu L^{MN} jest procedurą jednostopniową, ponieważ wszystkie profile w poszczególnych zbiorach otrzymujemy bezpośrednio z każdego rzędu cząstkowego układu czynnikowego. Pierwsze cztery kolumny prezentują profil 1 w zbiorze, następne cztery kolumny (5-8) przedstawiają profil 2, a ostatnie cztery kolumny (9-12) opisują ostatni, 3 profil w każdym ze zbiorów.

Procedury dwustopniowe pozwalają na wygenerowanie zbiorów profilów w dwóch etapach. W pierwszym etapie generowany jest tradycyjny cząstkowy eksperyment czynnikowy (taki sam jak w przypadku badania z wykorzystaniem *Conjoint Analysis* metodą pełnych profilów). Przyjmując, jak w procedurze jednostopniowej, że w badaniu wykorzystujemy 4 atrybuty ($M = 4$), a każdy z atrybutów reprezentowany jest przez 3 poziomy ($L = 3$), to otrzymujemy pełny eksperyment czynnikowy o rozmiarze $3^4 = 81$ profilów. Z pełnego eksperymentu czynnikowego generowany jest cząstkowy eksperyment o rozmiarze 9 profilów, przedstawiony w tab. 5.

W drugim etapie, przy wykorzystaniu profilów z tab. 5, generowane są zbiory profilów m.in. metodami [3, s. 163-164]: cykliczną (inaczej przesuwaną) oraz metodą „wymieszaj i podobieraj".

Pierwsza **metoda cykliczna** (**przesuwana**) (*Shifted Design, Cyclic Design*) przebiega w następujących etapach:

1. Wygenerowany cząstkowy układ czynnikowy, np. ortogonalny (tab. 5), wykorzystywany jest w celu wygenerowania przesuwanego układu profilów, zawiera-

[5] Sposób obliczania minimalnej liczby zbiorów wykorzystanych w badaniach będzie przedstawiony w dalszej części artykułu.

Tabela 5. Układ ortogonalny dla eksperymentu 3^4

Profil	Atrybut 1	Atrybut 2	Atrybut 3	Atrybut 4
1	1	1	1	1
2	1	2	2	3
3	1	3	3	2
4	2	1	2	2
5	2	2	3	1
6	2	3	1	3
7	3	1	3	3
8	3	2	1	2
9	3	3	2	1

Źródło: [3, s. 163].

jącego zbiory profilów. Każdy z profilów z tab. 5 jest pierwszym profilem generowanych zbiorów (zatem profil 1 z tab. 5 jest pierwszym profilem w zbiorze 1, profil 2 z tab. 5 jest pierwszym profilem w zbiorze 2 itd., a 9 profil z tab. 5 jest pierwszym profilem w 9 zbiorze). Zatem liczba wygenerowanych zbiorów jest równa liczbie profilów w układzie, który uzyskujemy po pierwszym etapie procedury dwustopniowej.

2. Otrzymujemy zatem układ z 4 kolumnami (każdy profil opisany jest 4 atrybutami) oraz 9 wierszami (9 zbiorów profilów). Do tak otrzymanego układu dodajemy kolejne cztery kolumny (które będą zawierać kolejny 2 profil w każdym ze zbiorów). Profile te otrzymuje się w wyniku „przesuwania" numeru poziomu atrybutu o 1 w górę (zatem 1 przechodzi w 2, 2 w 3, z poziomu 3 zaś wracamy do 1).

3. Powtarzamy krok 2, uzyskując w ten sposób trzeci profil w każdym z 9 zbiorów[6]. W ten sposób uzyskujemy przesuwany układ zbiorów profilów, przedstawiony w tab. 6.

Tabela 6. Układ przesuwany zbiorów profilów dla układu z tabeli 5

Zbiór	Profil 1				Profil 2				Profil 3			
	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	2	2	2	2	3	3	3	3
2	1	2	2	3	2	3	3	1	3	1	1	2
3	1	3	3	2	2	1	1	3	3	2	2	1
4	2	1	2	2	3	2	3	3	1	3	1	1
5	2	2	3	1	3	3	1	2	1	1	2	3
6	2	3	1	3	3	1	2	1	1	2	3	2
7	3	1	3	3	1	2	1	1	2	3	2	2
8	3	2	1	2	1	3	2	3	2	1	3	1
9	3	3	2	1	1	1	3	2	2	2	1	3

Źródło: [3, s. 164].

[6] Liczba profilów w każdym zbiorze nie może być większa od liczby poziomów atrybutów [1, s. 112-113].

Drugim ze sposobów wygenerowania zbiorów profilów jest **metoda „wymieszaj i podobieraj"** („*Mix and Match*") [3, s. 164-165]:

1. Wygenerowany układ cząstkowy, np. ortogonalny z tab. 5, wykorzystujemy do zbudowania 1 „słupka" profilów (o 4 kolumnach i 9 wierszach, podobnie jak w metodzie cyklicznej).

2. Pierwszy „słupek" (pierwsze 4 kolumny) wykorzystujemy ponownie, tylko tym razem zamieniamy 3 poziom atrybutu na 1 poziom w jednej (bądź więcej) kolumnie, a 1 poziom na 3 itd. W ten sposób otrzymujemy drugi „słupek" (kolumny 5-8).

3. Powtarzamy krok drugi, w wyniku którego otrzymujemy słupek 3.

4. Następnym krokiem jest „przetasowanie" każdego ze „słupków" oddzielnie.

5. W kolejnym kroku należy wybrać po jednym profilu z każdego „słupka". W ten sposób otrzymujemy pierwszy zbiór profilów (zawierający 3 profile).

6. Powtarzamy wybór profilów z poszczególnych „słupków" (wybór bez powtarzania) aż do momentu uzyskania 9 zbiorów profilów.

Dwie inne metody wykorzystywane w drugim etapie dwustopniowych procedur generowania układów czynnikowych przedstawił w swojej pracy Zwerina [15, s. 59-68]. Metody te to: wymiana poziomów atrybutów (*Swapping Attribute Levels, Swapped Choice Design*) oraz przeetykietowanie poziomów atrybutów (*Relabeling Attribute Levels*).

Metoda wymiany poziomów atrybutów polega na wymianie jednego poziomu atrybutu wewnątrz zbioru profilów, by m.in. zredukować dominację jednej z alternatyw w zbiorze (aby prawdopodobieństwa wyboru poszczególnych profilów w zbiorze nie różniły się tak znacznie). Program komputerowy generuje projekt (układ czynnikowy) przez wymianę poziomów atrybutów do momentu, aż miara *D*-efektywności[7] będzie możliwie najlepsza.

Metoda przeetykietowania poziomów atrybutów polega na przeznakowaniu etykiet poziomów układu czynnikowego. Metoda ta stosowana jest raczej do większych układów czynnikowych.

Po wybraniu atrybutów, które będą charakteryzowały badany produkt lub usługę, należy określić rozmiar eksperymentu. Należy zatem sprecyzować liczbę zbiorów w pełnym eksperymencie, liczbie profilów (obserwacji) kandydujących (jest to wstępnie przyjęty zbiór do redukcji z pełnego eksperymentu)[8], minimalną liczbę profilów oraz optymalną liczbę profilów, jaka powinna być zawarta w badaniu.

---

[7] Miary efektywności (optymalności) oparte na macierzy informacyjnej lub bazujące na odległości układu czynnikowego od wstępnie wytypowanego zbioru (tzw. obserwacji kandydujących) stosuje się do oceny jakości statystycznej generowanych układów [1, s. 86].

[8] Zbiór ten wybierany jest z wykorzystaniem oprogramowania komputerowego. Zbiór kandydujący redukowany jest do rozmiaru optymalnego (ostatecznie przyjętego do badania) za pomocą iteracyjnego zmodyfikowanego algorytmu Fedorova (*Modified Fedorov Algorithm*) [6, s. 110]. Zbiór kandydujący z liczbą profilów do 2000 określany jest jako mały, od 2000 do 5000 jako rozsądny, a od 5000 do 10 000 jako duży [6, s. 108].

W zależności od tego, czy w badaniu wykorzystamy projekt markowy czy też rodzajowy (ogólny) oraz czy atrybuty wybrane do badania będą symetryczne bądź też asymetryczne, rozmiar eksperymentu będzie szacowany na różne sposoby.

W przypadku eksperymentu symetrycznego, np. dla 5 marek opisanych 4 atrybutami po 3 poziomy każdy, całkowity rozmiar eksperymentu równałby się $3^{5\times4} = 3^{20}$, a minimalny rozmiar eksperymentu wyniósłby $5 \times 2 \times (3-1) + 1 = 41$ zbiorów. Natomiast w przypadku eksperymentu niesymetrycznego, np. dla 5 marek opisanych 2 atrybutami po 4 poziomy, 2 atrybutami po 2 poziomy oraz 1 atrybutem z 3 poziomami, całkowity rozmiar eksperymentu równałby się $4^{5\times2} \times 2^{5\times2} \times 3^{5\times1}$, a minimalny rozmiar eksperymentu wyniósłby $5 \times 2(4-1) + 5 \times 2(2-1) + 5 \times 1(3-1) + 1 = 41$ zbiorów.

Tak przyjęty rozmiar eksperymentu wykorzystujemy do badania preferencji konsumentów. Z każdego zbioru profilów respondent dokonuje wyboru jednego z profilów, bądź też rezygnuje z wyboru.

Plan eksperymentu cząstkowego jest macierzą **Q**, która została utworzona z wybranych wierszy planu pełnego eksperymentu (który odnosi się do tych samych czynników i ich replikacji) [10, s. 210]. Jednak wybór wierszy planu **Q** nie może być zupełnie dowolny. Powinien on podlegać formalnym ograniczeniom, które wynikają z dążenia do zwiększenia efektywności eksperymentu.

Najważniejszym z tych ograniczeń jest ortogonalność wektorów kolumnowych, dzięki której macierz kowariancji jest diagonalna (a co za tym idzie – oszacowania poszczególnych parametrów są nieskorelowane) [4, s. 108]. Ortogonalność kolumn macierzy prowadzi do uproszczenia obliczeń współczynników funkcji regresji.

Kolejnym wymogiem jest to, by w serii prób, z jakich składa się eksperyment cząstkowy, każdy z czynników był testowany symetrycznie (tzn. z jednakową liczbą razy na poziomie niższym i wyższym względem poziomu średniego) [10, s. 210]. Założenie to prowadzi do formalnego warunku, by sumy elementów w poszczególnych kolumnach macierzy **Q** równe były średniemu poziomowi czynnika[9]. Układy, w których warunek ten nie jest zachowany, nazywane są asymetrycznymi schematami czynnikowymi [1, s. 85; 15, s. 12].

Jeśli mamy do czynienia z eksperymentem wielopoziomowym jednorodnego typu, to dodatkowym wymogiem może być to, żeby badanie każdego czynnika było takie samo, czyli żeby każdy poziom danego czynnika występował tyle samo razy z każdym poziomem innego czynnika; dotyczy to wszystkich par czynników (co formalnie oznacza, że suma kwadratów elementów każdej kolumny macierzy **Q** powinna być taka sama) [1, s. 85]. Układ jest zrównoważony, jeżeli wszystkie pozadiagonalne elementy macierzy kowariancji odpowiadające wyrazowi wolnemu

[9] Jeżeli poziomy czynnika wyrażone są w wartościach zmiennych standardowych, to sumy elementów w poszczególnych kolumnach macierzy **Q** powinny być równe zeru [10, s. 210].

(z reguły w pierwszym wierszu i pierwszej kolumnie) są zerowe. Jeśli macierz ta jest diagonalna, to układ eksperymentu jest ortogonalny i zrównoważony.

Powyższe warunki charakteryzują eksperymenty liniowe, eksperymenty wyboru zaś charakteryzują się jeszcze dwoma warunkami. A mianowicie [15, s. 54; 13, s. 206]:

1. Minimalna część wspólna zbioru (*Minimal Level Overlap*), który zakłada, że prawdopodobieństwo tego, że poziom atrybutu powtórzy się w każdym zbiorze profilów, powinno być jak najmniejsze.

2. Zrównoważenie użyteczności (*Utility Balance*). Warunek ten jest zachowany, jeśli użyteczności poszczególnych alternatyw (profilów) w zbiorach są jak najbardziej równe sobie.

Jeśli wymagany, ortogonalny i zrównoważony układ czynnikowy nie istnieje, to w badaniu wykorzystujemy efektywny nieortogonalny układ czynnikowy. Eksperyment czynnikowy może być wybierany nie pod względem tego, czy jest ortogonalny czy też nieortogonalny, lecz pod względem tego, czy jest efektywny.

Do najczęściej spotykanych w literaturze kryteriów optymalności planu eksperymentu należą miary: *A*, *D* oraz *G*-efektywności [9; 15, s. 21-22].

Miara *A*-efektywności jest to funkcja średniej arytmetycznej wariancji, które są określone przez ślad macierzy informacyjnej $\mathbf{M} = \left(\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\right)^{-1}$. Maksymalizacja śladu macierzy informacyjnej jest tutaj kryterium wyboru układu punktów, a miara ta jest obliczana na podstawie wzoru [7, s. 42]:

$$A = 100 \times \frac{1}{n \times \mathrm{tr}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} / m}, \tag{1}$$

gdzie: n – liczba układów (punktów, profilów, serii);
m – liczba czynników (atrybutów).

Druga z miar, miara *D*-efektywności, jest funkcją średniej geometrycznej wektora wartości własnych macierzy informacyjnej $\mathbf{M} = \left(\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\right)^{-1}$. Maksymalna wartość wyznacznika macierzy informacyjnej jest tutaj kryterium wyboru (im bliższy zera jest wyznacznik, tym silniej skorelowane są czynniki eksperymentu) [1, s. 86]. Miara ta obliczana jest na podstawie wzoru [7, s. 42]:

$$D = 100 \times \frac{1}{n \times \left|\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\right|^{1/m}}. \tag{2}$$

Natomiast miara *G*-efektywności jest oparta na maksymalnym standardowym błędzie predykcji [7, s. 42]. Błąd ten wyznaczony jest ze zbioru wszystkich punktów kandydujących układu. Minimalizacja maksymalnej wartości błędu standardo-

wego obliczonego dla wybranych do układu punktów jest tutaj kryterium wyboru układu [1, s. 87]. Miara ta wyznaczana jest z następującego wzoru [7, s. 42]:

$$G = 100 \times \frac{\sqrt{\frac{m}{n}}}{\sigma_{\max}}, \tag{3}$$

gdzie: $\sigma_{\max}$ – maksymalny standardowy błąd predykcji wyznaczony ze zbioru punktów kandydujących.

Uzyskane wartości dla miar efektywności mieszczą się w przedziale (0; 100) dla odpowiednio zakodowanych planów czynnikowych. Przy czym wyższe wartości wskazują na efektywniejszy układ czynnikowy, informując o względnej optymalności układu w stosunku do hipotetycznego układu ortogonalnego [1, s. 87].

W związku z tym, że pełny eksperyment czynnikowy bardzo często zawiera zbyt dużą do ocenienia liczbę profilów, badacze wykorzystują cząstkowy eksperyment czynnikowy. Pozwala on na precyzyjniejsze oszacowania jednych (istotnych) efektów, przy jednoczesnej rezygnacji z oszacowań innych efektów. Efekty, które nie mogą być oszacowane niezależnie od siebie, są określane jako „pomieszane" bądź „uwikłane" (*Confounded, Aliased Effects*) [15, s. 24].

Poziom rozdzielczości cząstkowego układu czynnikowego (*Resolutions*) oznacza klasyfikację ortogonalnych cząstkowych układów czynnikowych. Poziom ten wskazuje, które efekty są oszacowane (efekty nieuwikłane)[10]. Generalnie, wyższy poziom rozdzielczości wymaga większego układu czynnikowego.

Dla poziomu rozdzielczości III wszystkie efekty główne, nieuwikłane z innymi efektami głównymi, są możliwe do oszacowania (jednak istnieje możliwość uwikłania efektów głównych z efektami iteracji pierwszego rzędu) [8, s. 50].

Poziom rozdzielczości IV cząstkowego układu czynnikowego oznacza, że możliwe do oszacowania są wszystkie efekty główne, nieuwikłane z innymi efektami głównymi oraz z efektami iteracji pierwszego rzędu (istnieje jednak możliwości wzajemnego uwikłania efektów iteracji pierwszego rzędu).

Natomiast poziom rozdzielczości V oznacza, że do oszacowania możliwe są wszystkie efekty główne, nieuwikłane z innymi efektami głównymi oraz wszystkie efekty iteracji pierwszego rzędu, również nieuwikłane z innymi efektami [8, s. 50].

Jeśli poziom rozdzielczości (r) jest nieparzysty, to efekt (interakcje) rzędu $e = (r-1)/2$ lub mniejszego jest możliwy do oszacowania (efekty główne wolne od uwikłań z innymi efektami głównymi; jednakże istnieje możliwość wzajemnego uwikłania efektów rzędu e z efektami iteracji rzędu $e+1$). Jeśli natomiast poziom (r) rozdzielczości jest parzysty, to efekty (interakcje) rzędu $e = (r-2)/2$ są możliwe do oszacowania (efekty główne nieuwikłane z innymi efektami głównymi oraz z efektami iteracji rzędu $e+1$) [8, s. 50].

[10] Inaczej poziom rozdzielczości informuje o zakresie efektów zarówno głównych, jak i interakcyjnych, które są uwzględnione w eksperymencie [1, s. 85].

W badaniach marketingowych poziom rozdzielczości III cząstkowego układu czynnikowego jest wykorzystywany najczęściej.

Literatura

[1] Bąk A., *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1013, Seria: Monografie i Opracowania nr 157, AE, Wrocław 2004.

[2] Chrzan K., Elrod T., *Choice-based Approach for Large Number of Attributes*, „Marketing News" 1995, Vol. 29, No. 1, American Marketing Association.

[3] Chrzan K., Orme B., *An Overview and Comparision of Design Strategies for Choice-based Conjoint Analysis*, Proceedings of the Sawtooth Software Conference, artykuł dostępny w Internecie na stronie: www.sawtoothsoftware.com/download/techpap/2000Proseedings.pdf, 161-177, 2000.

[4] Gajek L., Kaluszka M., *Wnioskowanie statystyczne – modele i metody*, Wydawnictwa Naukowo-Techniczne, Warszawa 1994.

[5] Hauser J.R., Rao V.R., *Conjoint Analysis, Related Modeling, and Applications*, artykuł dostępny w Internecie na stronie: http://web.mit.edu/hauser/www/Papers/GreenTributeConjoint 092302.pdf, 2002.

[6] Kuhfeld W.F., *Multinominal Logit, Discrete Choince Modeleng. An Introdution to Designing Choince Experiments, Collecting, Proccesing, and Analyzing Choice Data with the SAS® System*, artykul dostępny w Internecie na stronie: http://ftp.sas.com/techsup/download/ technote/ts643/ts643.pdf, Cary, SAS Institute, 2001.

[7] Kuhfeld W.F., *Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint, and Graphical Techniques*, artykul dostępny w Internecie na stronie: http://support.sas.com/ techsup/technote/ts689.pdf, Cary, SAS Institute, 2003.

[8] Kuhfeld W.F., *Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint, and Graphical Techniques*, artykul dostępny w Internecie na stronie: http://http://support.sas. com/techsup/technote/ts722.pdf, Cary, SAS Institute, 2005.

[9] Kuhfeld W.F., Tobias R.D., Garratt M., *Efficient Experimental Design with Marketing Research Applications*, „Journal of Marketing Research" 1994, 31 (November), 545-557.

[10] Kulikowski J.L., *Komputery w badaniach doświadczalnych*, Wydawnictwo Naukowe PWN, Warszawa 1993.

[11] Louviere J.J., *Conjoint Analysis Modelling of Statek Preferences*, „Journal of Transport Economics and Policy" 1988, January, 93-119.

[12] Louviere J.J., Hensher D.A., Swait J.D., *Stated Choice Methods. Analysis and Application*, Cambridge University Press, 2000.

[13] *Marketing Research and Modeling: Progress and Prospects*, eds. Y.J. Wind, P.E. Green, Springer. New York 2004.

[14] Riedesel P., *A Brief Introducion to Discrete Choince Analysis In Marketing Research*, artykul dostępny w Internecie na stronie: www.action-research. com/discrete.htm. 2001.

[15] Zwerina K., *Discrete Choice Experiments in Marketing*, Physica-Verlag, Heidelberg-New York 1997.

DATA PREPARATION IN CONSUMER PREFERENCE ANALYSIS WITH DISCRETE CHOICE METHODS

Summary

Discrete choice methods include varied experiments designs, varied procedures of generating factorial designs.

The paper presents the process of choice experiments and data preparation to be used in next steps of research: full profiles choice experiments, partial profiles choice experiments and one-step and two-step procedures of generating factorial designs.

Aneta Rybicka – dr, asystentka w Katedrze Ekonometrii i Informatyki Akademii Ekonomicznej we Wrocławiu – Wydział w Jeleniej Górze.