

MAREK WALESIAK

GROMADZENIE DANYCH W PROCEDURZE *CONJOINT ANALYSIS*

1. WPROWADZENIE

W metodologii *conjoint analysis*, w pomiarze struktury preferencji, wyróżnia się podejście dekompozycyjne, kompozycyjne i kompozycyjno-dekompozycyjne (por. [6]). W podejściu dekompozycyjnym, na podstawie dokonanych przez respondentów ocen zbioru profilów (produktów, usług) opisanych wybranymi atrybutami, otrzymuje się całkowite preferencje respondentów. Następnie, wykorzystując metody estymacji, dokonuje się podziału całkowitych preferencji respondentów przez obliczenie udziału każdego z atrybutów w oszacowanej całkowitej wartości użyteczności profilu. W podejściu kompozycyjnym użyteczność całkowita profilu jest ważoną sumą ocen poziomów atrybutów, gdzie wagi wyrażają ważność poszczególnych atrybutów (zob. [24], s. 3). W podejściu kompozycyjno-dekompozycyjnym łączy się cechy obu podejść. W pierwszej fazie stosuje się podejście kompozycyjne, a następnie podejście dekompozycyjne (np. w metodzie *ACA*).

Podstawowe kroki procedury *conjoint analysis* przedstawia tab. 1. W artykule scharakteryzowany zostanie czwarty krok tej procedury, tj. gromadzenie danych.

2. WYBÓR METODY PREZENTACJI DANYCH

W *conjoint analysis* do prezentacji danych stosuje się metody ([2], [9], s. 412–414, 424–425, [21], s. 28–29, [22], s. 302):

- pełnych profilów wyboru (*full-profile approach*),
- prezentacji dwóch atrybutów jednocześnie nazywana również metodą korzystającą z macierzy kompromisów (*two-attributes-at-a-time approach* lub *trade-off matrix approach*),
- porównywania profilów parami (*pairwise comparison method*),
- wyboru spośród zbiorów profilów (*the experimental choice approach*),
- oceny poziomów i atrybutów (*self-explicated data approach*).

Metoda pełnych profilów wyboru obejmuje zbiór wszystkich możliwych profilów będących kombinacją atrybutów i ich poziomów. Wadą tej metody jest ograniczenie liczby atrybutów i poziomów uwzględnianych w projektowanym eksperymencie, wraz ze wzrostem ich liczby bowiem znacznie rośnie liczba możliwych kombinacji przedstawionych respondentowi do oceny. Pojawia się więc problem zdolności percepcji i rzetelności wskazań dokonywanych przez respondentów. Zastosowanie technik reduk-

Tabela 1

Procedura *conjoint analysis*

Krok	Nazwa kroku w procedurze	Metody
1	Podejście stosowane do pomiaru struktury preferencji	dekompozycyjne, kompozycyjne, kompozycyjno-dekompozycyjne
2	Specyfikacja problemu badawczego	
3	Wybór postaci modelu zależności zmiennych i modelu preferencji	
	3.1. Wybór postaci modelu zależności zmiennych	modele: efektów głównych, z interakcjami
	3.2. Wybór postaci modelu preferencji	modele: liniowy, kwadratowy, odrębny użyteczności cząstkowych, mieszany
4	Gromadzenie danych	
	4.1. Wybór metody prezentacji danych	pełnych profilów wyboru, prezentacji par atrybutów, porównywania parami, oceny poziomów i atrybutów wyboru spośród zbiorów profilów
	4.2. Określenie skali pomiaru preferencji respondentów	skale metryczne (ilorazowa, przedziałowa), skale niemetryczne (porządkowa, nominalna)
	4.3. Generowanie zbioru profilów	układ czynnikowy, układ częściowo czynnikowy, próba losowa rozkładu wielowymiarowego
	4.4. Określenie formy i sposobu prezentacji profilów	opis słowny, forma kart z opisami słownymi poziomów atrybutów, forma graficzna (rysunki, fotografie), model trójwymiarowy, produkty fizyczne
	4.5. Wybór formy ankiety	wywiad osobisty, wywiad telefoniczny, kwestionariusz wysyłany pocztą, metoda komputerowo-interakcyjna, kombinacja kilku metod
	4.6. Ustalenie niezbędnej liczebności próby	
5	Wybór metody estymacji parametrów (użyteczności cząstkowych) modelu <i>conjoint analysis</i>	metryczne (OLS); niemetryczne: NONANOVA, PREFMAP, LINMAP, bazujące na prawdopodobieństwie wyboru (LOGIT, PROBIT)
6	Ocena wiarygodności pomiaru	ocena trafności i rzetelności pomiaru
7	Interpretacja wyników pomiaru metodą <i>conjoint analysis</i> (w przekroju indywidualnych respondentów oraz dla danych zagregowanych)	– analiza użyteczności cząstkowych, – określenie relatywnej ważności każdej zmiennej w procesie wyboru produktu przez nabywcę.
8	Obszary wykorzystania wyników	– segmentacja rynku, – symulacja udziałów w rynku

Źródło: opracowanie własne na podstawie prac: [5], [7], s. 9, [9], s. 401–402, 419, [21], s. 22.

cji rozmiaru eksperymentu rynkowego wykorzystujących układy częściowo czynnikowe pozwala znacznie zmniejszyć liczbę ocenianych przez respondentów profilów i jednocześnie zwiększyć liczbę atrybutów do 9 (w praktyce liczba atrybutów nie przekracza 6).

Do zalet tej metody należy zaliczyć (zob. [9], s. 412–413):

- przedstawienie respondentom do oceny profilów opisanych wszystkimi wybranymi atrybutami jednocześnie (z taką sytuacją konsument styka się na rynku),
- uwzględnienie interakcji występujących między atrybutami, które mogą generować efekty synergiczne,
- możliwość wyboru skali pomiaru preferencji respondentów (porządkowa, przedziałowa lub ilorazowa).

Metoda prezentacji dwóch atrybutów jednocześnie polega na prezentowaniu respondentom do oceny par atrybutów w formie macierzy. Liczba kolumn (wierszy) macierzy jest równa liczbie poziomów pierwszego (drugiego) z atrybutów. W metodzie tej porównuje się wszystkie możliwe pary atrybutów, co oznacza, że dla m atrybutów otrzymuje się $m(m-1)/2$ macierzy. Zadaniem respondenta jest porangowanie wszystkich kombinacji poziomów wyróżnionych dla dwóch porównywanych atrybutów. Zaletą tej metody jest możliwość uwzględnienia dużej liczby atrybutów, ponieważ respondent ranguje w danym momencie kombinacje poziomów tylko dwóch atrybutów. Metoda ta w praktyce nie jest stosowana zbyt często z uwagi na jej liczne wady (por. [9], s. 412, [19], s. 17–19; [21], s. 30):

- cechuje ją sztuczność (respondent ocenia profile w poszczególnych krokach ankiety na podstawie fragmentarycznych informacji),
- jest uciążliwa dla respondenta w przypadku dużej liczby atrybutów i ich poziomów,
- metoda ta prowadzi do przeceniania przez respondentów znaczenia atrybutów mniej ważnych i niedoceniań rangi atrybutów istotnych,
- nie daje możliwości zastosowania technik redukcji rozmiaru eksperymentu rynkowego wykorzystującego układy częściowo czynnikowe,
- oceny respondentów mogą być tylko prezentowane na skali porządkowej (rangowanie),
- do szacowania parametrów modelu (użyteczności cząstkowych) wymagane są tylko specyficzne metody estymacji (np. MONANOVA).

Metoda porównywania profilów parami łączy w sobie metodę pełnych profilów wyboru z metodą korzystającą z macierzy kompromisów. Respondent w poszczególnych krokach porównuje profile parami (każdy profil może, ale nie musi być opisany pełnym zestawem atrybutów). Respondent nie ocenia tutaj wszystkich profilów jednocześnie (jak w metodzie pełnych profilów wyboru), lecz w każdym kroku określania swoich preferencji wskazuje na jeden z dwóch prezentowanych profilów. Zatem dla n profilów dokonuje $n(n-1)/2$ wskazań.

Przyjmuje się, że respondent jest w stanie ocenić nie więcej niż 25 par profilów (zob. [19], s. 21). Dla ograniczenia liczby porównywanych par korzysta się w tej metodzie z technik redukcji rozmiaru eksperymentu rynkowego. Ponadto dla uzyskania jednoznaczności ocen respondentów postuluje się, aby była przestrzegana zasada przechodniości preferencji (jeżeli $A > B$ i $B > C$, to $A > C$, gdzie symbol $>$ oznacza preferencję lewego argumentu względem prawego).

Metoda wyboru spośród zbiorów profilów polega na tym, że:

- tworzy się zbiory profilów (zob. punkt 4), z których każdy zawiera dwa lub więcej profilów (każdy profil jest opisany wybranymi poziomami wszystkich atrybutów),

– respondenci proszeni są o wybór najbardziej preferowanego profilu w ramach każdego zbioru profilów. Respondent może zrezygnować z wyboru profilu w danym zbiorze, jeżeli żaden z oferowanych nie spełnia jego oczekiwań.

Liczba atrybutów w metodzie wyboru spośród zbiorów profilów w zasadzie nie powinna przekraczać 10 (podobnie jak w metodzie pełnych profilów), a liczba poziomów dla każdego atrybutu 15 (por. [16]).

Do zalet podejścia bazującego na wyborach zalicza się ([24], s. 6, [9], s. 425–429, [1]):

– stwarza możliwość bardziej realistycznej symulacji decyzji zakupu zachodzących na rynku w porównaniu do metody pełnych profilów wyboru,

– wprowadzenie opcji umożliwiającej respondentowi rezygnację z wyboru profilu pozwala oszacować nie tylko udziały poszczególnych profili w rynku, ale również wielkość popytu,

– w porównaniu do innych metod prezentacji danych jest dla respondenta mniej skomplikowana.

Metoda prezentacji danych bazująca na wyborach charakteryzuje się również wadami ([24], s. 6, [9], s. 428–429, [1]):

– w porównaniu z innymi metodami prezentacji danych prowadzi do zgromadzenia danych o mniejszej zawartości informacyjnej (pomiar preferencji respondentów dokonany jest na skali nominalnej),

– dla oszacowania parametrów modelu bazującego na wyborach wymaga się zgromadzenia dużej liczby obserwacji.

Metoda oceny poziomów i atrybutów składa się z dwóch etapów. Najpierw respondent ocenia poziomy atrybutów, np. w 10 punktowej skali pozycyjnej. Następnie respondent, wykorzystując skalę stałych sum, określa relatywną ważność poszczególnych atrybutów opisujących badane obiekty. Iloczyn oceny poziomu i oceny atrybutu daje w rezultacie użyteczność cząstkową, natomiast użyteczność całkowita profilu jest sumą tych iloczynów obliczoną dla wszystkich atrybutów. Model danych samowyjaśniających ma zatem charakter kompozycyjny w odróżnieniu od wcześniejszych metod prezentacji danych, które wykorzystywane są w podejściach dekompozycyjnych *conjoint analysis*. Zaletą tego podejścia jest możliwość uwzględnienia przy opisie profilów do 30 atrybutów. Do istotnych wad tej metody zalicza się to, że ([9], s. 424, [24], s. 4):

– sztuczność ocen respondentów, nie otrzymują oni bowiem do oceny pełnego obrazu danego produktu,

– przy ocenie ważności atrybutów prowadzi do niedocenia ważnych atrybutów oraz – przeceniania znaczenia mniej ważnych atrybutów.

Przykłady metod prezentacji danych obrazuje rys. 1. Wprowadzenie na rynek nowego notebooka wymaga rozpoznania preferencji nabywców. Każdy notebook opisany jest atrybutami:

Poziom atrybutu	Producent	Pamięć RAM	HDD	Cena	Napęd optyczny
1	IBM	128 MB	10 GB	12 tys. zł	CD-ROM
2	Toshiba	192 MB	16 GB	15 tys. zł	DVD-ROM
3	California Access	256 MB	22 GB	18 tys. zł	nagrywarka

3. OKREŚLENIE SKALI POMIARU PREFERENCJI RESPONDENTÓW

W metodologii *conjoint analysis* respondenci wyrażają swoje preferencje, oceniając poszczególne profile. Obserwacjami na zmiennej zależnej są oceny (preferencje) przypisane przez danego respondenta poszczególnym profilom (rys. 1). Otrzymana w ten sposób zmienna zależna mierzona jest na skali (por. [19], s. 36–38, [20], s. 91):

a) ilorazowej (istnieje tutaj naturalny punkt zerowy, który oznacza zupełny brak wielkości mierzonej zmiennej). Respondenci oceniają profile:

– przez podanie prawdopodobieństwa subiektywnego ich wyboru. Prawdopodobieństwo subiektywne „to pewna miara siły (stopnia) przekonania o tym, że zajdzie pewne zdarzenie” ([4], s. 77). Zakłada się, że człowiek jest w stanie przypisywać liczbowe miary sile swoich przekonań oraz miary siły przekonań, czyli subiektywne prawdopodobieństwa, podlegają takim samym prawom formalnym jak prawdopodobieństwa klasyczne lub częstościowe (chodzi o zasady dodawania i mnożenia prawdopodobieństw) ([4], s. 77),

– na skali stałych sum (respondent dokonuje podziału procentów lub stałej kwoty pieniędzy zgodnie z jego preferencjami wobec ocenianych profili);

b) przedziałowej, gdy respondenci oceniają poszczególne profile na skali pozycyjnej (*rating scale*), gdzie wartości ekstremalne oznaczają odpowiednio profil najmniej atrakcyjny i najbardziej atrakcyjny. Pomiar taki traktuje się jako przedziałowy. Nie jest to pomiar na skali przedziałowej *sensu stricto*, ponieważ nie można tutaj określić stałej jednostki;

c) porządkowej, gdy respondenci porządkują poszczególne profile np. przez nadanie im rang będących kolejnymi liczbami naturalnymi (rangę 1 przyporządkowuje się profilowi, który był wybierany w pierwszej kolejności, rangę n profilowi, który był wybierany w ostatniej kolejności);

d) nominalnej dwumianowej (respondenci wybierają jeden spośród dwóch profili) lub wielomianowej (respondenci wybierają jeden spośród więcej niż dwóch profili).

4. GENEROWANIE ZBIORU PROFILÓW

Eksperyment, w którym uwzględnia się wszystkie atrybuty i ich poziomy nazywa się pełnym eksperymencie czynnikowym. Liczba profili uzyskanych w tym układzie jest iloczynem poziomów poszczególnych atrybutów. Celem eksperymencie częściowo czynnikowego jest redukcja liczby profili przedstawianym respondentom do oceny.

Efektywne liniowe układy (*linear designs*) eksperymencie częściowo czynnikowych projektowane dla potrzeb tradycyjnej metodologii *conjoint analysis* (do prezentacji danych stosuje się metody pełnych profili wyboru oraz porównywania profili parami) charakteryzują się dwiema właściwościami ([10], [24], s. 55);

A. Metoda pełnych profili wyboru (*full-profile method*)

Profil	Producent	Pamięć RAM	HDD	Cena	Napęd optyczny	Preferencja
1	IBM	256 MB	22 GB	18 tys. zł	DVD-ROM	
2	Toshiba	192 MB	16 GB	12 tys. zł	DVD-ROM	

B. Metoda prezentacji dwóch atrybutów jednocześnie (*two-attributes-at-a-time approach*)

		Producent			Cena		
		IBM	Toshiba	California Access	12 tys. zł	15 tys. zł	18 tys. zł
Pamięć RAM	128 MB						
	192 MB						
	256 MB						
					Napęd optyczny	CD ROM	
						DVD ROM	
						nagrywarka	

C. Metoda porównywania profili parami (*pairwise comparison method*)

PROFIL 1 Producent: IBM RAM: 256 MB HDD: 22 GB Cena: 18 tys. zł Napęd optyczny: DVD-ROM		lub	PROFIL 2 Producent: Toshiba RAM: 192 MB HDD: 16 GB Cena: 12 tys. zł Napęd optyczny: DVD-ROM		
Jednakowo atrakcyjne					
Zdecydowanie preferuje lewy	1 — 2 — 3 — 4 — 5 — 6 — 7				Zdecydowanie preferuje prawy

D. Metoda wyboru spośród zbiorów profili (*the experimental choice approach*)

NOTEBOOK I	NOTEBOOK II	NOTEBOOK III	Bez wyboru
Producent: IBM	Producent: Toshiba	Producent: California Access	
RAM: 256 MB	RAM: 128 MB	RAM: 192 MB	
HDD: 22 GB	HDD: 10 GB	HDD: 16 GB	
Cena: 18 tys. zł	Cena: 12 tys. zł	Cena: 15 tys. zł	
Napęd optyczny: DVD-ROM	Napęd optyczny: nagrywarka	Napęd optyczny: CD-ROM	

Który profil preferujesz?

Rys. 1. Przykłady metod prezentacji danych

Źródło: opracowanie własne.

– ortogonalność (*orthogonality*), czyli niezależność kolumn macierzy reprezentującej układ eksperymentu (otrzymuje się zerowe wartości wszystkich iloczynów skalarnych obliczonych między kolumnami macierzy),

– zrównoważenie poziomów atrybutów (*level balance*) – każdy poziom danego atrybutu występuje z taką samą częstością w układzie eksperymentu (np. dla atrybutu z trzema poziomami każdy poziom występuje w układzie w 1/3 profilów).

W rzeczywistych problemach badawczych nie zawsze jest możliwe uzyskanie ortogonalnego układu czynnikowego spełniającego dodatkowo zrównoważenie poziomów atrybutów (por. [8], s. 327). Uzależnione to jest od liczby atrybutów, liczby ich poziomów oraz liczby generowanych przez nie profilów. Proponuje się wykorzystanie programów komputerowych, które optymalizują oba kryteria (zob. [24], s. 23). Kryteria optymalizacyjne w tworzeniu efektywnych liniowych układów eksperymentów częściowo czynnikowych omówiono m.in. w pracach: [10], [24], s. 19–22, 51–54; [14].

W pakiecie statystycznym *SPSS for Windows* dostępna jest procedura generująca układ ortogonalny uwzględniający tylko efekty główne. W procedurze tej każdy poziom jednego atrybutu występuje z każdym poziomem innego atrybutu z taką samą lub proporcjonalną częstością, co zapewnia niezależność oszacowanych efektów głównych ([18], s. 9).

Efektywne układy eksperymentów częściowo czynnikowych projektowane dla potrzeb dyskretnych modeli wyboru (do prezentacji danych stosuje się metody wyboru spośród zbiorów profilów) charakteryzują się dodatkowymi dwiema właściwościami ([10], s. 55):

– minimalna powtarzalność poziomów atrybutów w ramach poszczególnych zbiorów profilów wyboru (*minimal level overlap within choice sets*),

– zrównoważenie użyteczności profilów w ramach poszczególnych zbiorów profilów wyboru (*utility balance within choice sets*).

Prostą metodą wygenerowania zbiorów profilów dla potrzeb dyskretnych modeli wyboru jest metoda cykliczna (*cyclic design*) ([24], s. 50). W pierwszej fazie tej metody generuje się układ ortogonalny. Każdy profil stanowi załączek zbioru profili. Kolejne profile w każdym zbiorze tworzy się w sposób cykliczny. Tworzony profil zawiera numer poziomu każdego atrybutu o jeden większy od profilu poprzedniego (w przypadku, gdy profil opisany jest najwyższym numerem poziomu atrybutu numeracja rozpoczyna się od poziomu najniższego). W metodzie tej liczba profili w każdym zbiorze jest ograniczona do maksymalnej liczby poziomów atrybutów. Metoda cykliczna zapewnia optymalne spełnienie warunku ortogonalności, zrównoważenia poziomów atrybutów i minimalnej powtarzalności poziomów atrybutów w ramach poszczególnych zbiorów profilów wyboru.

W tabeli 2 zaprezentowano 18 zbiorów profilów dla potrzeb dyskretnych modeli wyboru otrzymanych w wyniku zastosowania metody cyklicznej. Pierwsze profile poszczególnych zbiorów przedstawiają przykład cząstkowego eksperymentu czynnikowego (dla modelu addytywnego tylko z efektami głównymi), dla notebooków opisanych 5 atrybutami o trzech poziomach każdy, wygenerowane w pakiecie statystycznym *SPSS for Windows* (pełny eksperyment daje 243 profile).

Tabela 2

Przykład zbiorów profilów dla potrzeb dyskretnych modeli wyboru otrzymany w wyniku zastosowania metody cyklicznej

Nr zbioru	Nr profilu w zbiorze	Atrybuty				
		1	2	3	4	5
1	2	3	4	5	6	7
1	I	1	3	3	3	2
	II	2	1	1	1	3
	III	3	2	2	2	1
2	I	2	2	2	1	2
	II	3	3	3	2	3
	III	1	1	1	3	1
3	I	3	2	3	1	3
	II	1	3	1	2	1
	III	2	1	2	3	2
4	I	1	3	2	3	3
	II	2	1	3	1	1
	III	3	2	1	2	2
5	I	3	2	1	3	2
	II	1	3	2	1	3
	III	2	1	3	2	1
6	I	2	3	3	2	1
	II	3	1	1	3	2
	III	1	2	2	1	3
7	I	3	1	3	2	3
	II	1	2	1	3	1
	III	2	3	2	1	2
8	I	3	3	2	1	1
	II	1	1	3	2	2
	III	2	2	1	3	3
9	I	3	1	2	3	1
	II	1	2	3	1	2
	III	2	3	1	2	3
10	I	2	1	2	2	2
	II	3	2	3	3	3
	III	1	3	1	1	1
11	I	1	1	3	1	2
	II	2	2	1	2	3
	III	3	3	2	3	1
12	I	2	2	3	3	1
	II	3	3	1	1	2
	III	1	1	2	2	3

c.d. tabeli 2

1	2	3	4	5	6	7
13	I	3	3	1	2	2
	II	1	1	2	3	3
	III	2	2	3	1	1
14	I	2	3	1	1	3
	II	3	1	2	2	1
	III	1	2	3	3	2
15	I	1	2	1	2	1
	II	2	3	2	3	2
	III	3	1	3	1	3
16	I	1	2	2	2	3
	II	2	3	3	3	1
	III	3	1	1	1	2
17	I	2	1	1	3	3
	II	3	2	2	1	1
	III	1	3	3	2	2
18	I	1	1	1	1	1
	II	2	2	2	2	2
	III	3	3	3	3	3

Cyfry 1, 2 i 3 reprezentują poziomy odpowiednich zmiennych.

Źródło: opracowanie własne.

5. USTALENIE NIEZBĘDNEJ LICZEBNOŚCI PRÓBY

W każdym badaniu za pomocą metod statystycznej analizy wielowymiarowej należy zwracać uwagę na liczebność próby, ponieważ zbyt mała próba powoduje, że (zob. [11]):

1. Za pomocą testów trudno jest zidentyfikować rezultaty statystycznie istotne (ze względu na małą moc testów). Mocą testu statystycznego dla pewnej hipotezy nazywa się prawdopodobieństwo odrzucenia hipotezy alternatywnej, gdy jest ona fałszywa (por. [13], s. 101). Moc testu wyraża się wzorem $1 - \beta$ (błąd II rodzaju). Przy danym poziomie istotności moc testu wzrasta, gdy zwiększa się liczebność próby (por. [9], s. 11).

2. Stosunkowo łatwo można otrzymać rezultaty, które są zbyt dobrze (w sposób sztuczny) dopasowane do danych z próby. W tej sytuacji nie można uogólniać otrzymanych wyników.

Liczebność próby uzależniona jest od:

a) zastosowanego wariantu metody *conjoint analysis*. Najczęściej stosuje się następujące warianty metody *conjoint analysis*:

- klasyczny bazujący na pełnych profilach wyboru (CVA – *Conjoint Value Analysis*),
- adaptacyjny wykorzystujący przy prezentacji danych do oceny przez respondentów metodę ocen poziomów i atrybutów oraz metodę porównywania parami (ACA – *Adaptive Conjoint Analysis*),

– bazujący na wyborach (*CBC – Choice-Based Conjoint analysis*),

b) obszaru wykorzystania wyników (segmentacja rynku, symulacja udziałów w rynku).

W metodzie klasycznej *CVA* i adaptacyjnej *ACA* estymuje się indywidualne użyteczności cząstkowe dla każdego respondenta z osobna.

W metodzie klasycznej *CVA*, wykorzystującej do estymacji użyteczności cząstkowych klasyczną metodę najmniejszych kwadratów ze zmiennymi sztucznymi, wyznaczyć należy minimalną liczbę profilów ocenianych przez poszczególnych respondentów. Niezbędna liczba profilów dana jest nierównością:

$$n \geq p - m + 1, \quad (3)$$

gdzie:

n – liczba profilów ocenianych przez respondentów,

p – liczba poziomów dla wszystkich atrybutów,

m – liczba, atrybutów.

Nierówność (3) nawiązuje do założenia metody najmniejszych kwadratów, które mówi, że liczba obserwacji musi być co najmniej równa liczbie szacowanych parametrów strukturalnych. Orme [16] sugeruje, że dla otrzymania statystycznie istotnych rezultatów liczba obserwacji (profilów ocenianych przez respondentów) powinna trzykrotnie przewyższać liczbę szacowanych parametrów ($p - m + 1$). Należy jednak pamiętać, że respondent nie jest jednak w stanie ocenić jednocześnie zbyt wielu profilów, zwykle więc ogranicza się ich liczbę do pewnego podzbioru. Liczba ocenianych profilów uzależniona jest od typu respondentów (wiek, wykształcenie itp.) i nie powinna przekraczać 30 profilów (zob. [22], s. 299).

W metodzie adaptacyjnej *ACA* należy zdecydować ile par profilów należy przedstawić respondentowi do oceny. Orme [16] sugeruje, że dla otrzymania statystycznie istotnych rezultatów liczba porównywanych par profilów powinna być równa:

$$3(p - m - 1) - p, \quad (4)$$

gdzie:

p – liczba poziomów dla wszystkich atrybutów,

m – liczba atrybutów.

W tej sytuacji liczba obserwacji będzie trzykrotnie przewyższać liczbę szacowanych parametrów.

Z kolei w metodach bazujących na wyborach szacuje się zagregowane użyteczności cząstkowe dla zbiorowości respondentów (w przypadku wykorzystania wielomianowych modeli logitowych) lub indywidualne użyteczności cząstkowe w przekroju respondentów (metoda hierarchiczna Bayesa lub metoda *ICE – Individual Choice Estimation*).

Wielkość próby w przypadku metod bazujących na wyborach, w których szacuje się zagregowane użyteczności cząstkowe dla zbiorowości respondentów wyznacza się z nierówności (zob. [12], Orme [16]):

$$S \cdot t \cdot a/c \geq 500, \quad (5)$$

gdzie:

S – liczba respondentów,

t – liczba zbiorów profilów,

a – liczba profili w zbiorze (z wykluczeniem opcji „bez wyboru”)

c – liczba cel. W przypadku uwzględniania tylko efektów głównych c równa się maksymalnej liczbie poziomów atrybutów. Uwzględnienie dwuczynnikowych interakcji powoduje, że c równa się iloczynowi poziomów dwóch atrybutów o największej liczbie poziomów.

Przykładowe projekty spełniające regułę (5) zawiera tabela:

S	t	a	Największa liczba poziomów dwóch atrybutów	Liczba cel (c)		S · t · a / c	
				A	B	A	B
100	20	4	4,4	4	16	2,000	500
150	15	5	4,5	5	20	2,250	562,5
300	15	5	5,5	5	25	4,500	900
400	10	3	4,5	5	20	2,400	600
500	15	3	6,6	6	36	3,750	625

A – efekty główne, B – interakcje dwuczynnikowe.

Wielkość próby w przypadku metod bazujących na wyborach, w których szacuje się indywidualne użyteczności cząstkowe w przekroju respondentów (np. metoda hierarchiczna Bayesa) nie może być wyznaczona analogicznie jak w metodzie klasycznej CVA i adaptacyjnej ACA. W tym przypadku estymacja indywidualnych użyteczności cząstkowych wymaga bowiem informacji pochodzących od wielu respondentów (zob. [16]).

Obszar wykorzystania wyników conjoint analysis (segmentacja rynku, symulacja udziałów w rynku) powoduje, że dla uzyskania reprezentatywnych wyników zachodzi potrzeba ustalenia niezbędnej liczebności zbioru obserwacji (liczby respondentów w badaniu). W literaturze poświęconej conjoint analysis nie ma formalnych propozycji w tym zakresie. Liczebność zbioru obserwacji może zostać wyznaczona arbitralnie przez badacza, który kieruje się własnym osądem, wiedzą i doświadczeniem, lub wielkość ta zostaje zdeterminowana przez budżet przeznaczony na zebranie danych (zob. [15], s. 35). W praktyce do określenia liczebności zbioru obserwacji korzysta się z doświadczeń historycznych. Według szacunków Wittinka i Cattina [23] próba powinna obejmować od 100 do 1000 respondentów, przy czym za typową należy uznać próbę od 300 do 550. Na podstawie komercyjnych zastosowań tej metody na rynku amerykańskim i zachodnioeuropejskim Wittink i Cattin [23] ustalili, że mediana wielkości próby wynosiła 300 elementów. O ile wielkość tę można uznać za zasadną dla rynku dóbr i usług konsumpcyjnych, o tyle dla rynku dóbr i usług produkcyjnych dopuszczalne są próby mniejsze od 100 elementów.

Zwiększonej liczebności próby wymagają studia segmentacyjne rynku w porównaniu do wykorzystania wyników conjoint analysis do symulacji udziałów w rynku.

6. PODSUMOWANIE

W artykule w syntetycznym ujęciu zaprezentowano zagadnienie gromadzenia danych w procedurze *conjoint analysis*. Szczegółowej analizie poddano zagadnienia, które wymagają zasadniczych rozstrzygnięć decyzyjnych z punktu widzenia metodologii *conjoint analysis*:

- wybór metody prezentacji danych,
- wybór skali pomiaru preferencji respondentów,
- generowanie zbioru profilów,
- ustalenie niezbędnej liczebności próby.

Akademia Ekonomiczna we Wrocławiu

LITERATURA

- [1] Bąk A., *Dyskretne modele wyborów conjoint analysis w badaniach marketingowych*, W: *Ekonometria, Prace Naukowe AE we Wrocławiu* (w redakcji), 2000.
- [2] Carroll J. D., Green P. E., *Psychometric Methods in Marketing Research: Part I, Conjoint Analysis*, „*Journal of Marketing Research*”, November, 32, 1995, s. 385–391.
- [3] Cattin P., Wittink D. R., *Commercial Use of Conjoint Analysis: a Survey*, „*Journal of Marketing*”, Summer 1982, s. 44–53.
- [4] Czerwiński Z., *Granice stosowalności teorii prawdopodobieństwa*, w: *Przestrzennoczasowe modelowanie i prognozowanie zjawisk gospodarczych 1999*, Materiały z XX Ogólnopolskiego Seminarium Naukowego (Zakopane 27–30.04.1998 r.). Praca zbiorowa pod red. A. Zeliasia, Kraków, Wydawnictwo Akademii Ekonomicznej w Krakowie, s. 73–83.
- [5] Green P. E., Srinivasan V., *Conjoint Analysis in Consumer Research: Issues and Outlook*, *Journal of Consumer Research*, September 5, 1978, s. 103–123.
- [6] Green P. E., Srinivasan V., *Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice*, *Journal of Marketing* October, 54, 1990, s. 3–19.
- [7] Gustafsson A., Herrmann A., Huber F. (red.), *Conjoint Measurement: Methods and Applications*, Berlin 2000, Springer.
- [8] Haaijer R., Wedel M., *Conjoint Choice Experiments: General Characteristics and Alternative Model Specifications*. W: Gustafsson A., Herrmann A., Huber F. (red.), *Conjoint Measurement: Methods and Applications*, Berlin 2000, Springer, s. 319–360.
- [9] Hair J. F., Anderson R. E., Tatham R. L., Black W. C., *Multivariate Data Analysis*, Englewood Cliffs, Prentice-Hall, 1998.
- [10] Huber J., Zwerina K., *The Importance of Utility Balance in Efficient Choice Designs*, *Journal of Marketing Research*, August 33, 1996, s. 307–317.
- [11] Jajuga K., Walesiak M., *Uwagi o badaniach niewyczerpujących przy zastosowaniu metod statystycznej analizy wielowymiarowej*, Konferencja naukowa nt. *Statystyka regionalna. Sondaż i integracja baz danych* (Baranowo k. Poznania, 25–27 września 1996 r.), Akademia Ekonomiczna w Poznaniu, 1997, s. 33–41.
- [12] Johnson R., *Getting the most from CBC – part I*, 1996, Artykuł dostępny w Internecie pod adresem: <http://www.sawtoothsoftware.com/TechPap.htm>, Sawtooth Software.
- [13] Kendall M. G., Buckland W. R., *Słownik terminów statystycznych*, Warszawa 1986, PWE.

- [14] Kuhfeld W. F., *Efficient Experimental Designs Using Computerized Searches* 1998, Artykuł dostępny w Internecie pod adresem: <http://www.sawtoothsoftware.com/TechPap.htm>, Sawtooth Software.
- [15] Mynarski S., *Praktyczne metody analizy danych rynkowych i marketingowych*, Kraków, Kantor Wydawniczy Zakamycze 2000.
- [16] Orme B., *Sample Size Issues for Conjoint Analysis Studies* 1998, Artykuł dostępny w Internecie pod adresem: <http://www.sawtoothsoftware.com/TechPap.htm>, Sawtooth Software.
- [17] Orme B., *Which Conjoint Method Should I Use? 2000*, Artykuł dostępny w Internecie pod adresem: <http://www.sawtoothsoftware.com/TechPap.htm>, Sawtooth Software.
- [18] *SPSS Conjoint™ 8.0*, Chicago 1997, SPSS Inc.
- [19] Vriens M., Wittink D. R., *Conjoint Analysis in Marketing*, 1994, (maszynopis powielony).
- [20] Walesiak M., *Metody analizy danych marketingowych*, Warszawa 1996, PWN.
- [21] Walesiak M., Bąk A., *Conjoint analysis w badaniach marketingowych*, Wrocław 2000, Wydawnictwo Akademii Ekonomicznej we Wrocławiu (w druku).
- [22] Wedel M., Kamakura W. A., *Market Segmentation. Conceptual and Methodological Foundations*, Boston-Dordrecht-London, Kluwer Academic Publishers 1998.
- [23] Wittink D. R., Cattin P., *Commercial Use of Conjoint Analysis: an Update*, Journal of Marketing, July, 1989, s. 91–96.
- [24] Zwerina K., *Discrete Choice Experiments in Marketing*, Heidelberg-New York 1997, Physica-Verlag.

Praca wpłynęła do Redakcji w październiku 2000 r.

DATA COLLECTION IN THE CONJOINT ANALYSIS PROCEDURE

Summary

A synthetic presentation of the general steps followed in the design and execution of the conjoint analysis experiment. The author characterised the fourth step, i.e. the data collection procedure, especially those issues which required a number of key decisions as regards the methodology of conjoint analysis:

- choosing a method of data presentation,
- selecting a measurement scale of consumer preference,
- selecting the data collection design,
- sample-size determination.