

EKONOMETRIA

26

Zastosowanie matematyki w ekonomii

Redaktor naukowy Janusz Łyko



**Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2009**

Spis treści

Wstęp	7
Beata Bal-Domańska , Ekonometryczna analiza sigma i beta konwergencji regionów Unii Europejskiej	9
Andrzej Bąk, Aneta Rybicka, Marcin Pelka , Modele efektów głównych i modele z interakcjami w <i>conjoint analysis</i> z zastosowaniem programu R	25
Katarzyna Budny , Kurtoza wektora losowego	44
Wiktor Ejsmont , Optymalna liczebność grupy studentów	55
Kamil Fijorek , Model regresji dla cechy przyjmującej wartości z przedziału $(0,1)$ – ujęcie bayesowskie	66
Paweł Hanczar , Wyznaczanie zapasu bezpieczeństwa w sieci logistycznej ...	77
Roman Huptas , Metody szacowania wewnątrzdziennej sezonowości w analizie danych finansowych pochodzących z pojedynczych transakcji	83
Aleksandra Iwanicka , Wpływ zewnętrznych czynników ryzyka na prawdopodobieństwo ruiny w skończonym horyzoncie czasowym w wieloklasowym modelu ryzyka.....	97
Agnieszka Lipieta , Stany równowagi na rynkach warunkowych	110
Krystyna Melich-Iwanek , Polski rynek pracy w świetle teorii histerezy.....	122
Rafał Piszczek , Zastosowanie modelu logit w modelowaniu upadłości	133
Marcin Salamaga , Próba weryfikacji teorii parytetu siły nabywczej na przykładzie kursów wybranych walut	149
Antoni Smoluk , O zasadzie dualności w programowaniu liniowym	160
Małgorzata Szulc-Janek , Influence of recommendations announcements on stock prices of fuel market	170
Jacek Welc , Regresja liniowa w szacowaniu fundamentalnych współczynników Beta na przykładzie spółek giełdowych z sektorów: budownictwa, informatyki oraz spożywczego	180
Andrzej Wilkowski , O współczynniku korelacji	191
Mirosław Wójciak , Klasyfikacja nowych technologii energetycznych ze względu na determinanty ich rozwoju.....	199
Andrzej Wójcik , Wykorzystanie modeli wektorowo-autoregresyjnych do modelowania gospodarki Polski	209
Katarzyna Zeug-Żebro , Rekonstrukcja przestrzeni stanów na podstawie wielowymiarowych szeregów czasowych.....	219

Summaries

Beata Bal-Domańska , Econometric analysis of sigma and beta convergence in the European Union regions	24
Andrzej Bąk, Aneta Rybicka, Marcin Pelka , Main effects models and main and interactions models in <i>conjoint analysis</i> with application of R software.....	43
Katarzyna Budny , Kurtosis of a random vector	53
Wiktor Ejsmont , Optimal class size of students	65
Kamil Fijorek , Regression model for data restricted to the interval (0,1) – Bayesian approach	76
Paweł Hanczar , Safety stock level calculation in a supply chain network.....	82
Roman Huptas , Estimation methods of intraday seasonality in transaction financial data analysis	96
Aleksandra Iwanicka , An impact of some outside risk factors on the finite-time ruin probability for a multi-classes risk model.....	109
Agnieszka Lipieta , States of contingent market equilibrium	121
Krystyna Melich-Iwanek , The Polish labour market in light of the hysteresis theory	132
Rafał Piszczek , Logit model applications for bankruptcy modelling.....	148
Marcin Salamaga , Attempt to verify the purchasing power parity theory in the case of some foreign currencies.....	159
Antoni Smoluk , On dual principle of linear programming	168
Małgorzata Szulc-Janek , Analiza wpływu rekomendacji analityków na ceny akcji branży paliwowej (Analiza wpływu rekomendacji analityków na ceny akcji branży paliwowej)	178
Jacek Welc , A linear regression in estimating fundamental betas in the case of the stock market companies from construction, it and food industries	190
Andrzej Wilkowski , About the coefficient of correlation	198
Mirosław Wójciak , Classification of new energy related technologies based on the determinants of their development	208
Andrzej Wójcik , Using vector-autoregressive models to modelling economy of Poland.....	218
Katarzyna Zeug-Żebro , State space reconstruction from multivariate time series	227

Andrzej Bąk, Aneta Rybicka, Marcin Pełka

Uniwersytet Ekonomiczny we Wrocławiu

MODELE EFEKTÓW GŁÓWNYCH I MODELE Z INTERAKCJAMI W *CONJOINT ANALYSIS* Z ZASTOSOWANIEM PROGRAMU R

Streszczenie: *Conjoint analysis* jest metodą statystyczną, w której preferencje empiryczne wobec różnych ofert są poddawane dekompozycji w celu określenia funkcji użyteczności każdego atrybutu oraz względnego znaczenia każdego z nich. Najczęściej wykorzystywane są modele efektów głównych *conjoint analysis*.

W artykule zaprezentowano procedurę badawczą *conjoint analysis* dla modeli efektów głównych oraz modeli uwzględniających interakcje między atrybutami wraz z przedstawieniem możliwości zastosowania programu R w poszczególnych etapach badania. Przedstawione zostały podstawowe pojęcia z zakresu *conjoint analysis* oraz omówiono pojęcie i rodzaje interakcji. Artykuł zawiera fragmenty kodów źródłowych w języku R przygotowanych na użytek badań empirycznych. Część empiryczna opisuje procedurę badania *conjoint analysis* z wykorzystaniem programu R „krok po kroku”.

Słowa kluczowe: *conjoint analysis*, interakcje, modele efektów głównych, modele uwzględniające interakcje.

1. Wstęp

Conjoint analysis jest metodą statystyczną, w której preferencje empiryczne respondentów wobec różnych ofert (rzeczywistych lub hipotetycznych) są poddawane dekompozycji w celu określenia funkcji użyteczności każdego atrybutu oraz względnego znaczenia każdego z nich. Najczęściej w badaniach empirycznych wykorzystywane są modele efektów głównych *conjoint analysis*.

W artykule zaprezentowano procedurę badawczą *conjoint analysis* dla modeli efektów głównych oraz modeli uwzględniających interakcje między atrybutami wraz z przedstawieniem możliwości zastosowania programu R w poszczególnych etapach badania. Przedstawione zostały podstawowe pojęcia z zakresu *conjoint analysis* oraz omówiono pojęcie i rodzaje interakcji. Artykuł przedstawia również fragmenty kodów źródłowych procedur w programie R przygotowanych przez autorów na użytek badań empirycznych. W tym celu wykorzystano dane zebrane przez studentów na potrzeby ich prac magisterskich.

2. Metody pomiaru preferencji wyrażonych

Celem istnienia i działania każdego przedsiębiorstwa jest przede wszystkim pozyskiwanie nabywców na wytwarzane oraz oferowane produkty lub usługi. Warunkiem pomyślności takiego działania jest zorientowanie na konsumenta, czyli m.in. rozpoznanie jego potrzeb i sposobów ich zaspokajania. Istotne jest również to, w jaki sposób przedsiębiorstwa dostosują do tych potrzeb swoje oferty towarowe. Dlatego też kierownictwo zleca badania postępowania konsumentów na rynku. Badaniom tym podlegają zachowania, opinie i postawy, preferencje i upodobania nabywców, plany i zamiary zakupów oraz motywacje postępowania.

Badania postaw i preferencji konsumentów umożliwiają opisanie ich zachowań względem oferowanych na rynku produktów lub usług. Postawy te mogą być kształtowane przez różne czynniki zewnętrzne, jak również mogą zmieniać się w czasie [Duliniec 1997, s. 135].

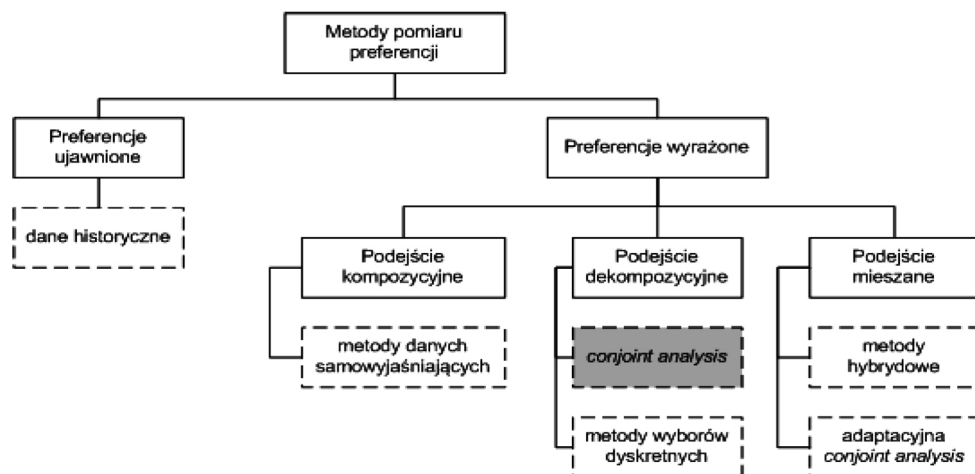
Jeśli oceny postaw dotyczą różnych produktów lub usług należących do tej samej grupy, to można określić relacje, jakie zachodzą między tymi produktami lub usługami w sposób wymierny [Bąk 2000a, s. 69]. Relacje te nazywamy w teorii ekonomii preferencjami, gdyż informują one nas o stosunku konsumentów do określonych produktów lub usług, co umożliwi nam ich uszeregowanie od najbardziej do najmniej pożądanym.

W **podjęciu kompozycyjnym** (*compositional approach*) wykorzystywana jest idea modelu postaw Fishbeina¹ oraz założenia, które są związane z modelem wartości oczekiwanej, gdzie użyteczność całkowita wielowymiarowego profilu jest ważoną sumą ocen poziomów zmiennych, a wagi wyrażają ważność poszczególnych zmiennych [Walesiak, Bąk 1997, s. 14; Zwerina 1997, s. 3]. Modele kompozycyjne jest to klasa modeli wielu zmiennych, których przykładem są m.in. modele regresji oraz analiza dyskryminacyjna [Hair i in. 1995, s. 562-563]. Analitycy, stosując modele kompozycyjne, zbierają oceny respondentów na temat wielu cech produktu lub usługi, a następnie wiążą te oceny w całkowite preferencje. Analitycy zatem określają (inaczej „komponują, składają”) preferencje respondentów z ocen dokonywanych przez respondentów każdego z atrybutów produktu lub usługi.

W **podjęciu dekompozycyjnym** (*decompositional approach*) w celu przeprowadzenia analizy preferencji konsumentów wykorzystujemy metody *conjoint analysis* oraz metody oparte na wyborach [Bąk 2000a, s. 76]. Modele dekompozy-

¹ Model ten znajduje się w nurcie teorii poznawczych, których przedstawicielami w marketingu są również Lutz, Bettman, Rosenberg, Wilkie i Pessemier [Sagan 2004]. Fishbein twierdził, że postawa wobec produktu zależy od stopnia przekonania o istnieniu danej cechy w produkcie (w skali od 1 do 7) oraz od ocen wartościujących daną cechę (w skali od -3 do +3). W późniejszym modelu Fishbein i Ajzen przyjęli, że lepszym predykatorem wyborów produktów są nie tyle postawy wobec marek, lecz postawy wobec konsekwencji posiadania tych marek i stopień akceptacji wymuszeń otoczenia w zakresie ich zakupu [Sagan 2004]. Z tego modelu wynika, że wybory produktów są funkcją wyrażonych intencji zakupów, które zależą m.in. od siły postaw.

cyjne to klasa modeli, które „rozkładają” preferencje całkowite konsumentów. Wykorzystując modele dekompozycyjne, prezentuje się respondentom zbiór profili, zazwyczaj w formie hipotetycznych lub rzeczywistych produktów lub usług [Hair i in. 1995, s. 558]. Za pomocą metod statystycznych i algorytmów komputerowych przeprowadza się dekompozycję preferencji całkowitych i oblicza się użyteczności cząstkowe² [Bąk 2004, s. 42].



Rys. 1. Metody pomiaru preferencji

Źródło: opracowanie własne na podstawie prac [Bąk 2004; Walesiak, Bąk 2000].

W **podjęciu mieszanym** formułowane są modele łączące cechy obu opisanych wyżej ujęć. Zastosowanie znajdują tu przede wszystkim modele hybrydowe *conjoint analysis* oraz adaptacyjna *conjoint analysis*. W obu tych metodach stosuje się dwufazowe procedury pomiaru preferencji [Bąk 2004, s. 44]. I etap procedury to bezpośrednie oceny atrybutów i ich poziomów, II etap zaś to ocena wybranych par lub podzbiorów profili produktów lub usług.

Conjoint analysis prezentuje podejście dekompozycyjne i jest najbardziej popularną metodą pomiaru preferencji. Przeprowadzając badania z wykorzystaniem *conjoint analysis*, respondentowi przedstawia się do oceny, w formie ankiety, zbiór profili, czyli produktów lub usług. Profile te, rzeczywiste bądź hipotetyczne, opisane są wybranymi zmiennymi objaśniającymi (atrybutami).

² W metodach dekompozycyjnych użyteczności całkowite są to oceny profili, które stanowią podstawę dalszej analizy polegającej na dekompozycji użyteczności całkowitych profili na użyteczności cząstkowe poziomów atrybutów oraz na oszacowaniu udziałów poszczególnych atrybutów w kształtowaniu użyteczności całkowitej każdego profilu [Bąk 2004, s. 48-49].

Zgodnie z terminologią stosowaną w literaturze przedmiotu zmienne objaśniające opisujące dobra lub usługi nazywa się atrybutami (*attributes*)³ lub czynnikami (*factors*), natomiast ich realizacje są nazywane **poziomami** (*levels*)⁴. Atrybuty i ich poziomy generują różne warianty dóbr lub usług, nazywane **profilami** (*profiles, stimuli, treatments, runs*). Liczba wszystkich możliwych do wygenerowania profili zależy od liczby atrybutów i liczby poziomów (jest to iloczyn liczby poziomów wszystkich atrybutów). Zależności między atrybutami, poziomami i profilami są zilustrowane na rys. 2, przedstawiającym jeden atrybut (A) o 3 poziomach i dwa atrybuty (B) i (C) o 2 poziomach.

Respondenci oceniają profile produktów lub usług, wyrażając w ten sposób swoje preferencje. Oceny profili są nazywane **użytecznościami całkowitymi** i stanowią podstawę dalszej analizy, która polega na **dekompozycji** użyteczności całkowitych profili na **użyteczności cząstkowe** poziomów atrybutów oraz na oszacowaniu udziałów poszczególnych atrybutów w kształtowaniu użyteczności całkowitej każdego profilu [Green, Wind 1975].

		Atrybuty			
		A	B	C	
		1	1	1	Poziomy
		2	2	2	
		3			
		↓	↓	↓	
Profile		A1	B1	C1	
1	→				
2	→	A1	B1	C2	
:	:	:	:	:	
3 x 2 x 2 = 12	→	A3	B2	C2	

Rys. 2. Zależność między atrybutami, poziomami i profilami

Źródło: [Walesiak, Gatnar 2009, s. 288].

Następnym krokiem jest dekompozycja całkowitych preferencji i obliczenie udziału każdej zmiennej objaśniającej w oszacowanej całkowitej wartości użytecz-

³ Termin „atrybut” jest używany w statystyce w odniesieniu do zmiennych niemetrycznych, najczęściej nominalnych [Kendall, Buckland 1986, s. 13].

⁴ Terminy te są stosowane w statystycznym planowaniu doświadczeń. Układy doświadczalne odgrywają w metodach dekompozycyjnych bardzo ważną rolę i stanowią jeden z najważniejszych etapów procedury badawczej realizowanej za pomocą analizy *conjoint*.

ności obiektu [Walesiak, Bąk 2000, s. 14]. Uzyskujemy oszacowane użyteczności cząstkowe, związane z poziomami atrybutów w postaci macierzy [Bąk 2000b, s. 217]. Kolumny macierzy odpowiadają poziomom wyróżnionym dla poszczególnych zmiennych, a liczba kolumn – liczbie poziomów. Natomiast wiersze w tej macierzy odpowiadają respondentom biorącym udział w badaniu, a ich liczba – liczbie respondentów. Tak otrzymana macierz użyteczności cząstkowych podlega w dalszych etapach analizie i interpretacji.

W badaniu preferencji z wykorzystaniem *conjoint analysis* stosuje się najczęściej dwa sposoby pomiaru preferencji respondentów [Bąk 2002, s. 394]:

- porządkowanie obiektów w kolejności od najbardziej do najmniej satysfakcjonującego bądź też odwrotnie (ranking, na skali porządkowej);
- ocenę względnej atrakcyjności przedstawionych obiektów (na skali pozycyjnej).

Tabela 1. Etapy i kroki procedury *conjoint analysis*

Lp.	Etap procedury	Krok procedury
1	Specyfikacja zadania badawczego	<ul style="list-style-type: none"> • zmienna objaśniana • zmienne objaśniające (atrybuty)
2	Określenie postaci modelu	<ul style="list-style-type: none"> • model zależności zmiennych objaśniających (efektów głównych lub z interakcjami) • model preferencji (liniowy, kwadratowy, użyteczności cząstkowych)
3	Gromadzenie danych	<ul style="list-style-type: none"> • metody gromadzenia danych (pełne profile, porównywanie profilów parami, prezentacja par atrybutów, dane symulacyjne) • metody generowania profilów (układy czynnikowe, próba losowa)
4	Prezentacja profilów	<ul style="list-style-type: none"> • forma prezentacji (opis słowny, rysunek, model, produkt fizyczny) • forma badań (wywiad bezpośredni, poczta, telefon, komputer, Internet)
5	Wybór skali pomiaru preferencji	<ul style="list-style-type: none"> • skale niemetryczne (nominalna, porządkowa) • skale metryczne (przedziałowa, ilorazowa)
6	Estymacja modelu	<ul style="list-style-type: none"> • modele niemetryczne (MONANOVA, PREFMAP) • modele metryczne (MNK) • modele probabilistyczne (LOGIT, PROBIT)
7	Analiza i interpretacja wyników	<ul style="list-style-type: none"> • analiza preferencji (ocena ważności atrybutów) • symulacja udziałów w rynku • segmentacja

Źródło: [Walesiak, Gatnar 2009, s. 292].

Tradycyjna *conjoint analysis* (jak również metody wyborów dyskretnych) może być wykorzystywana w celu analizy preferencji, segmentacji rynku czy analizy symulacyjnej oraz prognozowania rynku (zob. [Zwerina 1997, s. 5]).

Najważniejsze cechy *conjoint analysis* opartej na metodzie profilów pełnych:

- 1) liczba atrybutów uwzględnionych w badaniu jest ograniczona zwykle do 6,

- 2) profile przedstawiane respondentom do oceny są opisane wszystkimi atrybutami,
- 3) profile są generowane na podstawie ortogonalnych układów czynnikowych,
- 4) profile wygenerowane na podstawie układów ortogonalnych są wzajemnie maksymalnie zróżnicowane,
- 5) w modelu *conjoint analysis* można uwzględnić, oprócz efektów głównych, również efekty interakcji atrybutów,
- 6) wszyscy respondenci oceniają ten sam zbiór profili,
- 7) model *conjoint analysis* reprezentuje tzw. podejście dekompozycyjne, tzn. na podstawie empirycznych użyteczności całkowitych profili szacuje się użyteczności cząstkowe poziomów atrybutów,
- 8) można wykorzystać różne metody gromadzenia danych ze źródeł pierwotnych,
- 9) poszczególne etapy procedury *conjoint analysis* (tj. przygotowanie profili, gromadzenie danych, estymacja parametrów, symulacja udziałów w rynku) są rozdzielone.

3. Modele efektów głównych i modele z interakcjami

W procedurze modelowania *conjoint analysis* konstruujemy modele formalne, m.in. dotyczące reguł określających sposób powiązania zmiennych, czyli charakteru zależności zachodzących między zmiennymi [Walesiak, Bąk 2000, s. 24; Hair i in. 1995].

Reguły, które określają sposób powiązania zmiennych, dotyczą sposobu, w jaki respondent agreguje użyteczności cząstkowe poszczególnych zmiennych w celu oszacowania użyteczności całkowitej danego profilu. Możemy wyróżnić dwa typy modeli zależności użyteczności całkowitej od użyteczności cząstkowych: **model efektów głównych** (addytywny) oraz **model uwzględniający interakcje między zmiennymi** (efekty główne i współdziałania) [Akaah, Korgaonkar 1983, s. 190; Carmone, Green 1981, s. 90].

Każdy z poziomów zmiennej objaśniającej wywiera określony wpływ na zmienną objaśnianą, nazywany efektem oddziaływania danego poziomu. Efekt ten mierzy się różnicą między wartością parametru odpowiadającego zmiennej objaśnianej wynikającą z wpływu wszystkich pozostałych poziomów [Steczkowski, Zeliaś 1982, s. 18-19].

Wpływy poszczególnych zmiennych wynikające z niezależnego oddziaływania poziomów poszczególnych zmiennych nazywa się efektami głównymi [Żuk 1989, s. 108].

Model addytywny (efektów głównych) możemy przedstawić następująco [Walesiak, Bąk 2000 s. 24-25; Carmone, Green, Jain 1978, s. 301; Green, Devita 1975, s. 146]:

$$\widehat{U}_{is} = b_s + \sum_{j=1}^m \sum_{p=1}^{m_j} u_{jp(s)} x_{jp(i)}, \quad (1)$$

- gdzie: \widehat{U}_{is} – oszacowana użyteczność całkowita i -tego profilu dla s -tego respondenta;
- $u_{jp(s)}$ – oszacowana użyteczność cząstkowa p -tego poziomu j -tej zmiennej objaśniającej dla s -tego respondenta (przedstawia efekt główny p -tego poziomu j -tej zmiennej objaśniającej);
- $x_{jp(i)}$ – zmienna sztuczna, która reprezentuje poziomy zmiennej objaśniającej, np. w przypadku kodowania zero-jedynkowego zmienna ta przyjmuje wartości 1 lub 0, zgodnie z zasadą: $x_{jp(i)} = 1$, jeśli p -ty poziom j -tej zmiennej występuje w i -tym profilu oraz $x_{jp(i)} = 0$ w przeciwnym przypadku;
- m – liczba zmiennych;
- m_j – liczba poziomów j -tej zmiennej objaśniającej
- $j = 1, \dots, m$ – numer zmiennej objaśnianej (atrybutu);
- b_s – wyraz wolny modelu.

Zgodnie z modelem addytywnym sumowane są wartości dla każdego atrybutu (wartości cząstkowe), by uzyskać wartość całkowitą dla kombinacji atrybutów (produktu lub usługi). Model ten znajduje zastosowanie w większości (80-90%) przeprowadzanych badań preferencji, w prawie wszystkich sytuacjach i w większości zastosowań jest wystarczający [Hair i in. 1995, s. 570].

W badaniu wielu zjawisk zachodzi jednak sytuacja, w której oprócz niezależnego wpływu zmiennych objaśniających na zmienną objaśnianą występuje dodatkowy efekt łącznego oddziaływania obu zmiennych objaśniających wyrażający się ich iloczynem. Zależność występującą między zmiennymi objaśniającymi, która przejawia się w taki sposób, że wpływ jednej zmiennej objaśniającej zależy od wartości poziomów drugiej zmiennej objaśniającej, nazywamy interakcją lub współdziałaniem efektów tych czynników [Żuk 1989, s. 108].

Model uwzględniający interakcje, które występują między atrybutami, może przyjąć postać [Walesiak i Bąk 2000, s. 25; Carmone, Green, Jain 1978, s. 301]:

$$\widehat{U}_{is} = b_s + \sum_{j=1}^m \sum_{p=1}^{m_j} u_{jp(s)} x_{jp(i)} + \sum_{j=1}^{m-1} \sum_{k=j+1}^m \sum_{p=1}^{m_j} u_{jkp(s)} x_{jkp(i)}, \quad (2)$$

- gdzie: \widehat{U}_{is} – oszacowana użyteczność całkowita i -tego obiektu dla s -tego respondenta;
- $u_{jkp(s)}$ – szacowana użyteczność cząstkowa p -tego poziomu j -tej zmiennej uwzględniająca efekt interakcji między zmiennymi $j \times k$ (użyteczność cząstkowa wynikająca z efektu dwuczynnikowych

- interakcji między zmiennymi objaśniającymi $j \times k$ dla s -tego respondenta, która reprezentuje efekt j -tej oraz k -tej zmiennej objaśniającej);
- $x_{jkp(i)}$ – zmienna sztuczna, która reprezentuje efekty dwuczynnikowych interakcji pomiędzy zmiennymi objaśniającymi $j \times k$, np. w przypadku kodowania zero-jedynkowego zmienna sztuczna przyjmuje wartości 1 lub 0 zgodnie z zasadą: $x_{jkp(i)} = 1$ jeśli efekt interakcji $j \times k$ występuje w i -tym profilu, oraz $x_{jkp(i)} = 0$ w przeciwnym przypadku;
- m – liczba zmiennych objaśniających (atrybutów);
- m_j – liczba poziomów j -tej zmiennej objaśniającej;
- $j = 1, \dots, m$ – numer zmiennej objaśnianej (atrybutu);
- b_s – wyraz wolny modelu.

Liczba możliwych interakcji zależy od liczby uwzględnionych w badaniu zmiennych. Jeżeli w badaniu uwzględniono więcej niż dwie ($l > 2$), np. A, B, C, D , to mówi się o: efektach głównych A, B, C, D , których liczba wynosi l ; interakcjach pierwszego rzędu, np.: AB, AC, BC itd., których liczba wynosi $\frac{l(l-1)}{2}$; interakcjach drugiego rzędu ABC, ABD, BCD itd., których liczba wynosi $\frac{l(l-1)(l-2)}{3!}$, oraz interakcjach trzeciego rzędu $ABCD$, których liczba wynosi $\frac{l(l-1)(l-2)(l-3)}{4!}$.

[Steczkowski, Zeliaś 1982, s. 47-48].

Ze względu na trudności interpretacji, koszty oraz pracochłonność porównań zazwyczaj nie wychodzi się poza interakcje czteroczynnikowe, czyli interakcje trzeciego rzędu. Często w badaniach pewne interakcje z góry uznaje się za mało ważne i pomija się je, a równocześnie inne uznaje się za szczególnie ważne.

W niekompletnych eksperymentach czynnikowych efekty interakcji rzędów wyższych niż rzędu drugiego są pomijane. W sytuacji takiej niektóre efekty uwzględnione w modelu są szacowane na podstawie innych efektów nieuwzględnionych w modelu, których wpływu nie znamy. Efekty te są więc zależne (są to liniowe kombinacje) od efektów niewłączonych do modelu [Walesiak, Bąk 2000, s. 38]. Mamy do czynienia wtedy z tzw. interakcjami uwikłanymi, a układ odpowiadający temu nosi nazwę układu z uwikłaną interakcją [Steczkowski, Zeliaś 1982, s. 48]. Uwikłanie interakcji oznacza występowanie korelacji między efektami uwzględnionymi w modelu a efektami nieuwzględnionymi – jest to zjawisko współliniowości zmiennych objaśniających [Bąk 2004, s. 83].

Efekty, które nie mogą być oszacowane niezależnie od siebie, są określane jako „pomieszane” bądź „uwikłane” (*confounded, aliased effects*) [Zwerina 1997, s. 24].

Poziom rozdzielczości cząstkowego układu czynnikowego (*resolutions*) oznacza klasyfikację ortogonalnych cząstkowych układów czynnikowych. Poziom ten informuje o zakresie efektów zarówno głównych, jak i interakcyjnych uwzględnionych w eksperymencie [Bąk 2004, s. 85]. Wskazuje zatem, które efekty są oszacowane (efekty niewikłane). Najczęściej stosowanymi układami czynnikowymi są układy o rozdzielczości III, IV i V [Zwerina 1997, s. 14-15]. Generalnie, wyższy poziom rozdzielczości wymaga większego układu czynnikowego.

Dla poziomu rozdzielczości III wszystkie efekty główne, niewikłane z innymi efektami głównymi, są możliwe do oszacowania (jednak istnieje możliwość uwikłania efektów głównych z efektami iteracji pierwszego rzędu) [Kuhfeld 2005, s. 50].

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

Natomiast poziom rozdzielczości V oznacza, że możliwe są do oszacowania wszystkie efekty główne, niewikłane z innymi efektami głównymi, oraz wszystkie efekty iteracji pierwszego rzędu, również niewikłane z innymi efektami [Kuhfeld 2005, s. 50].

Generalnie, jeśli poziom rozdzielczości (r) jest nieparzysty, wtedy 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 niewikłane z innymi efektami głównymi oraz z efektami iteracji rzędu $e + 1$) [Kuhfeld 2005, s. 50].

W badaniach marketingowych poziom rozdzielczości III cząstkowego układu czynnikowego jest wykorzystywany najczęściej, ponieważ pozwala uwzględnić stosunkowo dużo czynników przy minimalnej liczbie wygenerowanych profili [Bąk 2004, s. 85]. Kosztem tego podejścia jest możliwość występowania efektów interakcji, których wpływ nie będzie oszacowany (wyższy poziom rozdzielczości wymaga eksperymentu czynnikowego o większych rozmiarach, ale pozwala uwzględnić więcej niewikłanych efektów) [Zwerina 1997, s. 14-15].

Wybór jednego z modeli zależności użyteczności całkowitej od użyteczności cząstkowych przesądza o liczbie profili (produktów) ocenianych przez respondentów oraz o metodzie estymacji użytej do oszacowania wartości użyteczności cząstkowych [Walesiak, Bąk 2000, s. 25]. Pierwszy z modeli, model efektów głównych, warunkuje mniejszą liczbę profili do oceny. Model ten gwarantuje również łatwiejsze uzyskanie estymatorów użyteczności cząstkowych. W związku z tym wybór modelu zależności zmiennych decyduje o tym, w jaki sposób zmienne są powiązane wzajemnie z punktu widzenia respondenta, który ocenia dany profil charakteryzowany tymi zmiennymi.

4. Pakiety i funkcje programu R

Program R jest narzędziem stworzonym głównie do obliczeń statystycznych i graficznej prezentacji danych. Jest to program darmowy dostępny na warunkach licencji GPL 2 (<http://gnu.org.pl/text/licencja-gnu.html>).

Aktualnie nie jest dostępny pakiet programu R realizujący w sposób kompleksowy procedurę *conjoint analysis*. Jednakże zawiera on wiele pakietów, które mogą być wykorzystane w poszczególnych jej etapach. W tabeli 2 zawarto funkcje i pakiety programu R, które mogą być wykorzystane w procedurze *conjoint analysis*.

Tabela 2. Pakiety i funkcje przydatne w procedurze *conjoint analysis*

Etap procedury	Krok procedury	Wybrane pakiety i funkcje programu R
Gromadzenie danych	Metody gromadzenia danych – dane symulacyjne	Pakiet base (funkcja <code>sample</code>) Pakiet stats (funkcja <code>runif</code>)
Gromadzenie danych	Metody generowania profilów – układy czynnikowe	Pakiet AlgDesign (funkcje: <code>gen.factorial</code> , <code>optFederov</code>)
Gromadzenie danych	Metody generowania profilów – próba losowa	Pakiet poLCA (funkcja <code>poLCA.simdata</code>)
Estymacja modelu	Model metryczny – MNK	Pakiet stats (funkcje: <code>contrasts</code> , <code>lm</code>) Pakiet base (funkcja <code>factor</code>)
Analiza i interpretacja wyników	Ocena ważności atrybutów	brak
Analiza i interpretacja wyników	Symulacja udziałów w rynku	brak
Analiza i interpretacja wyników	Segmentacja	Pakiet stats (funkcja <code>kmeans</code>)
Analiza i interpretacja wyników	Wykres ważności atrybutów	Pakiet graphics (funkcje: <code>barplot</code> , <code>title</code>)

Źródło: [Walesiak, Gatnar 2009, s. 304].

W celu realizacji szacowania użyteczności cząstkowych na poziomie indywidualnym i zagregowanym oraz estymacji macierzy teoretycznych użyteczności całkowitych przygotowano następujące funkcje [Walesiak, Gatnar 2009, s. 305-306]:

- `partutilities(xfrm, y, x, n, p, S)` – funkcja oblicza macierz użyteczności cząstkowych poziomów atrybutów w przekroju respondentów (na poziomie indywidualnym); `xfrm` – wyrażenie reprezentujące w modelu atrybuty (model *conjoint analysis*); `y` – wektor preferencji empirycznych $y[n \times S]$; `x` – macierz reprezentująca profile; `n` – liczba profilów; `p` – liczba poziomów wszystkich atrybutów z wyrazem wolnym (bez poziomów odniesienia), tj. liczba zmiennych sztucznych plus wyraz wolny; `S` – liczba respondentów. Funkcja zwraca `usl[S, p]` – macierz użyteczności cząstkowych poziomów

- atrybutów w przekroju respondentów; sposób użycia: `uls <- partutilities(xfrm, y, x, n, p, S)`.
- `matexpand(m, n, S, x)` – funkcja rozwija macierz układu czynnikowego reprezentującą profile na zbiór respondentów w celu estymacji modelu na poziomie zagregowanym; m – liczba atrybutów; n – liczba profilów; S – liczba respondentów; x – macierz reprezentująca profile. Funkcja zwraca $X[n \times S, m]$ – rozwiniętą macierz układu czynnikowego; sposób użycia: `x <- as.data.frame(matexpand(m, n, S, x))`.
 - `totalutilities(xfrm, y, x, n, p, S)` – funkcja oblicza macierz teoretycznych użyteczności całkowitych profilów w przekroju respondentów. Znaczenie argumentów `xfrm, y, x, n, p, S` – jak w funkcji `partutilities()`. Funkcja zwraca $Usi[S, n]$ – macierz użyteczności całkowitych n profilów w przekroju respondentów; sposób użycia: `Usi <- totalutilities(xfrm, y, x, n, p, S)`.

Dodatkowo w celu oszacowania interakcji przygotowano kod źródłowy programu R:

```
subsets <- function(n, r) {
  if(is.numeric(n) & length(n) == 1) v <- 1:n else {
    v <- n
    n <- length(v)
  }
  subs <- function(n, r, v)
  if(r <= 0) NULL else
  if(r >= n) matrix(v[1:n], nrow = 1) else
  rbind(cbind(v[1], subs(n - 1, r - 1, v[-1])),
  subs(n - 1, r, v[-1]))
  subs(n, r, v)
}
p<-c(1,2,3)
przedzial<-function(x,poczatki){
  wynik<-length(poczatki)
  for(i in 1:(length(poczatki)-1)){
    if((poczatki[i]<=x) && (x<poczatki[i+1])){
      wynik<-i
    }
  }
  wynik
}
dozwolone <-function(indeksy,poczatki){
  resul<-TRUE
  for(i in 1:(length(indeksy)-1))
  for(j in (i+1):length(indeksy)){
if(przedzial(indeksy[i],poczatki)==przedzial(indeksy[j],poczatki)){
  resul<-FALSE }
  }
  resul
}
```

```

x<-as.matrix(x)
iloscZmiennych<-ncol(x)
poziomInterakcji<-2
maxKombinacje<-poziomInterakcji+1
poziomy<-c(2,1,1)
poczatki<-seq(along.with=poziomy)
poczatki[1]<-1
for(i in 2:length(poziomy)){
  poczatki[i]<-poczatki[i-1]+poziomy[i-1]
}
xInterakcje<-x
for(i in 1:poziomInterakcji){
  for(j in 1:nrow(subsets(p,i+1))){
    indeksy<-subsets(p,i+1)[j,]
    print(paste(paste(indeksy,collapse=","),dozwolone(indeksy,po-
czatki)))
    if(dozwolone(indeksy,poczatki)){
      wynik<-x[,indeksy[1]]
      for(k in 2:(i+1)){
        wynik<-wynik*x[,indeksy[k]]
      }
      xInterakcje<-cbind(xInterakcje,wynik)
    }
  }
}

```

Funkcja `subsets` została przygotowana przez Billa Venablesa z CASIRO Laboratories z Cleveland w Australii (www.r-project.org) i zamieszczona w Internecie na prawach licencji GPL 2. Funkcja ta generuje `r`-elementowe zbiory ze zbioru `n`-elementowego bez powtórzeń. Zbiory `r`-elementowe są następnie wykorzystywane przy mnożeniu odpowiednich kolumn przekodowanej macierzy profilów.

Funkcje `przedzialy` i `dozwolone` sprawdzają, które z kolumn w przekodowanej macierzy profilów mogą zostać przemnożone, ponieważ nie może mieć miejsca sytuacja, w której mnożone są kolumny zmiennych sztucznych kodujące poziomy tego samego atrybutu.

Element `poziomInterakcji` wskazuje, którego rzędu interakcje mają być oszacowane. Kod został opracowany w ten sposób, że wskazanie interakcji wyższego (np. drugiego) rzędu powoduje automatyczne obliczanie interakcji niższego rzędu (w tym przypadku pierwszego). Element `p<-c(1,2,3)` wskazuje numery zmiennych, które mają zostać uwzględnione przy szacowaniu interakcji (w tym przypadku wybrano wszystkie zmienne).

Wynikiem obliczeń przeprowadzanych w kodzie źródłowym jest przekodowana macierz profilów wraz z iloczynami dla poszczególnych kolumn (interakcjami) – `xInterakcje`. Macierz ta następnie jest wykorzystywana w ten sam sposób jak macierz `x` w przygotowanych funkcjach programu R.

5. Przykład zastosowania

W badaniach empirycznych celem zobrazowania szacowania modeli *conjoint analysis* bez interakcji i modeli uwzględniających interakcje z wykorzystaniem programu R wykorzystano ankietę przygotowaną przez Dorotę Łakomą. Przy czym do szacowania w obu przypadkach wykorzystano model liniowy.

Ankieta dotyczyła preferencji respondentów dotyczących wody mineralnej. Dla tego produktu określono trzy atrybuty oraz ich poziomy:

- 1) mineralizacja – niska (1), średnia (2), wysoka (3),
- 2) smak – bezsmakowa (1), smakowa (2),
- 3) zawartość CO₂ – niegazowana (1), gazowana (2).

Aby przeprowadzić badanie, konieczne stało się zaprojektowanie pełnego eksperymentu czynnikowego. W tym przykładzie pełny eksperyment czynnikowy miałby $3 \times 2 \times 2 = 12$ profilów. Generowanie takiego układu w programie R wymaga załadowania pakietu `AlgDesign` i zastosowania funkcji `gen.factorial`. Funkcja ta wymaga podania dwóch parametrów: wektora o długości równej liczbie atrybutów, gdzie kolejne liczby oznaczają poziomy dla poszczególnych atrybutów, oraz informacji, które zmienne mają być traktowane jako czynniki, lub informacji o wybraniu wszystkich zmiennych, np.: `full_profile<-gen.factorial(c(3,2,2), factors="all")`.

W badaniu można uwzględnić wszystkie możliwe do utworzenia profile (wykorzystać kompletny układ czynnikowy) lub tylko ich podzbiór (wykorzystać cząstkowy układ czynnikowy). W praktyce kompletny układ czynnikowy można wykorzystać jedynie przy bardzo małej liczbie atrybutów i poziomów (por. [Walesiak, Gatnar 2009, s. 298]). Dlatego też często korzysta się z cząstkowego układu czynnikowego. W tym przykładzie, mimo niewielkiej liczby profilów, w pełnym eksperymencie zdecydowano się na wygenerowanie cząstkowego układu czynnikowego. Cząstkowy układ czynnikowy może zostać wygenerowany w programie R z wykorzystaniem pakietu `AlgDesign` i funkcji `optFedorov`. Funkcja ta realizuje algorytm wymiany Fedorova [Fedorov 1972]. Do uruchomienia tej funkcji w jej podstawowym kształcie potrzebne są trzy informacje – formuła powiązania zmiennych, lista zmiennych kandydatek oraz opcjonalnie liczba profilów, którą chcemy otrzymać w wyniku działania tej funkcji, np.: `partial_profile<-optFedorov(~., full_profile, 10)`.

Wynikiem działania tej funkcji jest cząstkowy układ czynnikowy oraz mierniki efektywności układu.

Wygenerowany układ cząstkowy został następnie zawarty w kwestionariuszu ankiety. Respondenci oceniali profile w skali od 0 – profil najgorszy do 10 – profil najlepszy. Respondenci zostali dobrani do badania w sposób przypadkowy. W sumie zebrano dane o preferencjach 96 respondentów.

Pierwszym krokiem było oszacowanie użyteczności cząstkowych dla modelu efektów głównych z wykorzystaniem funkcji `partutilities` programu R. W efekcie otrzymano oszacowania parametrów modelu (zob. tab. 3).

Tabela 3. Oszacowane parametry modelu efektów głównych (fragment tabeli danych)

Numer respondenta	b0	X1.1 b1	X1.2 b2	X2 b3	X3 b4
1	3,845	-2,095	-0,702	0,321	0,75
2	2,512	1,238	-0,083	-0,679	1,607
3	3,286	-0,286	0,143	-3,286	1
4	6,893	-1,143	1,536	-0,893	-0,607
5	1,857	0,143	0,5	-1,857	0,214
6	0,857	1,143	-0,286	-0,857	0,857
7	2,298	3,952	-0,94	0,036	-0,893
8	2,024	1,476	-0,167	-1,357	1,214
9	3,452	-0,952	2,976	-1,786	-2,5
10	4,655	0,095	0,202	0,179	1,75
11	4,881	-2,381	0,976	-0,214	-0,643
12	1,071	1,429	-0,357	-1,071	1,071
13	2,452	-0,952	0,905	-0,286	-2,214
14	4,095	-0,095	0,476	-1,929	1,786
15	5,214	0,286	-0,286	-3,214	-0,429
...

Źródło: obliczenia własne w programie R.

W kolejnym etapie z wykorzystaniem kodu źródłowego z pliku `interakcje.r` obliczono macierz `xInterakcje` (zob. tab. 4).

Tabela 4. Macierz `xInterakcje`

X1.1	X1.2	X2	X3	X1.1_X2	X1.1_X3	X1.2_X2	X1.2_X3	X2_X3	X1.1_X2_X3	X1.2_X2_X3
1	0	1	1	1	1	0	0	1	1	0
1	0	-1	-1	-1	-1	0	0	1	1	0
0	1	1	1	0	0	1	1	1	0	1
1	0	1	-1	1	-1	0	0	-1	-1	0
-1	-1	1	-1	-1	1	-1	1	-1	1	1
1	0	-1	1	-1	1	0	0	-1	-1	0
-1	-1	-1	1	1	-1	1	-1	-1	1	1
0	1	-1	-1	0	0	-1	-1	1	0	1
-1	-1	1	1	-1	-1	-1	-1	1	-1	-1
0	1	1	-1	0	0	1	-1	-1	0	-1

Źródło: obliczenia własne.

W macierzy zaprezentowanej w tab. 3 elementy kolumn prezentujących interakcje są wynikiem mnożenia odpowiednich kolumn z przekodowanej macierzy profilów, np.: kolumna X1.1_X2 jest wynikiem mnożenia kolumn X1.1 i X2. Jak już wspomniano, macierz ta jest następnie wykorzystywana podobnie jak macierz profilów w funkcji *partutilities*.

Wyniki oszacowań parametrów modelu uwzględniającego interakcje zawarto w tab. 5.

Tabela 5. Oszacowane parametry modelu uwzględniającego interakcje (fragment tabeli danych)

Numer respondenta	b0	X1.1 b1	X1.2 b2	X2 b3	X3 b4	X1.1_X2 b5	X1.1_X2 b6	X1.2_X2 b7	X1.2_X3 b8	X2_X3 b9
1	3,75	-2,00	0,50	0,42	1,58	0,33	-1,33	-1,67	0,17	-0,75
2	2,42	1,33	1,33	-0,58	2,58	-1,17	0,17	-1,17	0,17	-1,75
3	3,33	-0,33	1,67	-3,33	2,00	0,33	0,00	-1,67	0,00	-2,00
4	6,92	-1,17	-0,17	-0,92	-1,75	-0,83	1,00	2,17	0,00	1,75
5	1,83	0,17	1,67	-1,83	1,00	-0,17	0,00	-1,67	0,00	-1,00
6	0,67	1,33	1,33	-0,67	2,00	-1,33	0,00	-1,33	0,00	-2,00
7	2,42	3,83	-2,17	-0,08	-1,08	-0,17	-1,67	0,83	1,33	0,75
8	2,00	1,50	1,50	-1,33	2,33	-1,17	-0,33	-1,17	0,67	-2,00
9	2,50	0,00	0,00	-0,83	-4,17	-1,67	1,67	3,33	-3,33	2,50
10	4,75	0,00	0,00	0,08	1,58	0,17	0,17	0,17	0,17	0,25
11	4,33	-1,83	0,17	0,33	-1,00	1,17	1,00	-0,83	-3,00	1,00
12	0,83	1,67	1,67	-0,83	2,50	-1,67	0,00	-1,67	0,00	-2,50
13	2,50	-1,00	1,00	-0,33	-2,17	0,33	0,67	-0,67	-0,33	0,00
14	4,00	0,00	0,50	-1,83	1,83	-0,17	0,17	-0,17	-0,33	0,00
15	5,50	0,00	0,00	-3,50	-0,33	0,00	-0,67	0,00	1,33	0,00
...

Źródło: obliczenia własne z wykorzystaniem programu R.

W modelu *conjoint analysis* uwzględniającym interakcje ostatecznie oszacowano jedynie interakcje pierwszego rzędu. Interakcje drugiego rzędu nie mogły zostać oszacowane ze względu na zbyt małą liczbę profilów.

Oszacowane użyteczności cząstkowe dla modelu efektów głównych i modelu uwzględniającego interakcje wykorzystano do oceny atrakcyjności (użyteczności całkowitej) każdego profilu (wzór 3) i wskazania średniej ważności zmiennych (wzory 4 i 5) (zob. [Walesiak, Bąk 2000, s. 74; Walesiak, Gatnar 2009, s. 303]). Atrakcyjność *i*-tego profilu wyraża się wzorem:

$$U_i = \frac{1}{S} \sum_{s=1}^S \left(\sum_{j=1}^m U_{jl_j^i}^s + b_{0s} \right), \quad (3)$$

gdzie: $U_{jl_j^i}^s$ – użyteczność cząstkowa i -tego poziomu j -tej zmiennej i -tego profilu dla respondenta s ($s = 1, \dots, S$);

l_j^i – numer poziomu dla j -tej zmiennej i i -tego profilu;

$i = 1, \dots, n$ – numer profilu;

$j = 1, \dots, m$ – numer zmiennej;

b_{0s} – wyraz wolny dla respondenta s .

Średnia ważność każdej zmiennej wyraża się wzorem:

$$W_j = \frac{1}{S} \sum_{s=1}^S W_j^s, \quad (4)$$

gdzie:

$$W_j^s = \frac{\max_{l_j} \{U_{jl_j^i}^s\} - \min_{l_j} \{U_{jl_j^i}^s\}}{\sum_{j=1}^m \left(\max_{l_j} \{U_{jl_j^i}^s\} - \min_{l_j} \{U_{jl_j^i}^s\} \right)}, \quad (5)$$

pozostałe oznaczenia jak we wzorze 3.

W tabeli 6 zestawiono atrakcyjność profili zarówno dla modelu efektów głównych, jak i dla modelu uwzględniającego interakcje.

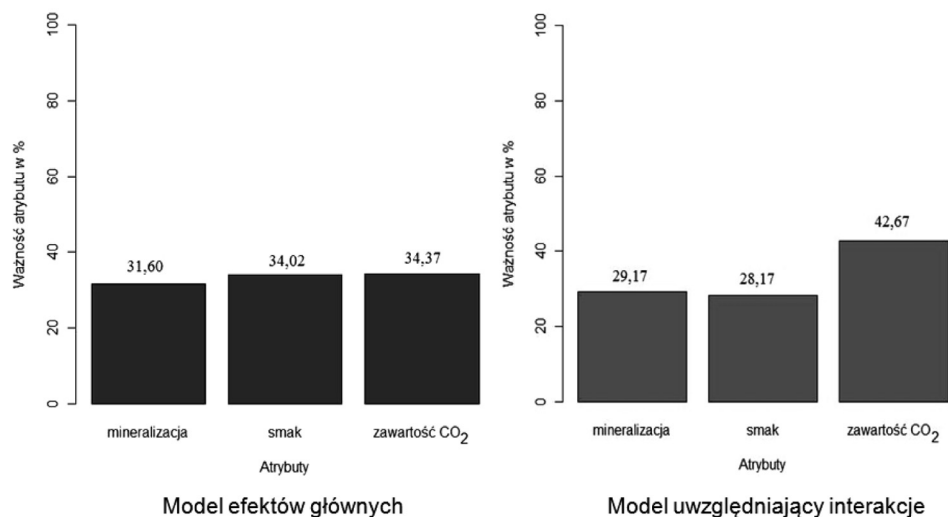
Tabela 6. Atrakcyjność (użyteczność całkowita) profili

Numer profilu	Mineralizacja	Smak	Zawartość CO ₂	Model efektów głównych		Model uwzględniający interakcje	
				użyteczność całkowita	pozycja	użyteczność całkowita	pozycja
				1	2	3	4
1	niska	bezsmałowa	niegazowana	2,903	8	2,583	10
2	niska	smakowa	gazowana	5,450	2	5,542	2
3	średnia	bezsmałowa	niegazowana	3,133	7	3,250	7
4	niska	bezsmałowa	gazowana	3,242	6	3,427	5
5	wysoka	bezsmałowa	gazowana	2,872	9	2,938	9
6	niska	smakowa	niegazowana	5,109	3	5,427	3
7	wysoka	smakowa	niegazowana	4,743	4	4,333	4
8	średnia	smakowa	gazowana	5,676	1	5,969	1
9	wysoka	bezsmałowa	niegazowana	2,533	10	2,990	8
10	średnia	bezsmałowa	gazowana	3,474	5	3,260	6

Źródło: obliczenia własne w programie R.

Profil numer 8 (smakowa, gazowana woda mineralna o średniej mineralizacji) okazał się profilem najbardziej atrakcyjnym w przypadku zarówno modelu efektów głównych, jak i modelu uwzględniającego interakcje. Pozostałe profile były w obu modelach oceniane podobnie; pewne zmiany pojawiły się na dalszych pozycjach.

Na rysunku 3 przedstawiono średnią ważność zmiennych w przypadku obu modeli.



Rys. 3. Średnia ważność zmiennych

Źródło: obliczenia własne z wykorzystaniem programu R.

W przypadku modelu efektów głównych najważniejszym atrybutem (zmienną) jest zawartość CO₂. Na drugim miejscu jest smak, a na trzecim mineralizacja. Należy tu jednak dodać, że wszystkie te zmienne mają zbliżony poziom ważności. W przypadku modelu uwzględniającego interakcje zawartość CO₂ jest znacznie ważniejsza niż smak czy mineralizacja. Drugą zmienną pod względem ważności jest mineralizacja, a trzecią smak.

6. Wnioski

W programie R przygotowano kody źródłowe i funkcje dla niektórych etapów procedury *conjoint analysis* (generowanie profili, szacowanie użyteczności cząstkowych, estymacja modelu uwzględniającego interakcje, kodowanie *quasi*-eksperymentalne oraz zero-jedynkowe, obliczanie interakcji).

W przeciwieństwie do płatnych programów statystycznych (np. SAS), które zazwyczaj albo nie pozwalają szacować interakcji, albo szacują tylko interakcje pierwszego rzędu, przygotowane funkcje i kody źródłowe programu R pozwalają

szacować interakcje dowolnego rzędu, dodatkowo możliwy jest tu wybór zmiennych, które mają wziąć udział w szacowaniu interakcji. Dodatkowym argumentem przemawiającym za stosowaniem programu R jest fakt, że program ten jest darmowy, dostępny na zasadach licencji GPL 2.

W części empirycznej przedstawiono przykład szacowania interakcji na podstawie danych o wodzie mineralnej. Przykład ten ma charakter jedynie obrazowy i służy do prezentacji przygotowanych funkcji i kodów źródłowych programu R. Najbardziej atrakcyjnym profilem w przypadku modelu efektów głównych oraz modelu uwzględniającego interakcje okazał się profil 8 (gazowana, smakowa woda mineralna o średniej mineralizacji).

Kierunkiem dalszych prac powinno stać się oprogramowanie dalszych etapów *conjoint analysis* (segmentacja rynku, modelowanie symulacyjne) oraz oprogramowanie drugiej metody reprezentującej podejście dekompozycyjne, tj. metod wyborów dyskretnych.

Literatura

- Akaah I.P., Korgaonkar P.K., *An Empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint Models*, „Journal of Marketing Research” 1983 vol. XX (May), s. 187-197.
- Bąk A., *Conjoint analysis jako metoda pomiaru postaw i preferencji konsumentów*, [w:] M. Walesiak (red.), *Pomiar w badaniach rynkowych i marketingowych*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 856, AE, Wrocław 2000a, s. 69-81.
- Bąk A., *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, AE, Wrocław 2004.
- Bąk A., *Możliwości wykorzystania alternatywnych algorytmów conjoint analysis w badaniach marketingowych*, [w:] *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia, Zeszyt 7, Sekcja Klasyfikacji i Analizy Danych PTS, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 874, AE, Wrocław 2000b, s. 217-226.
- Bąk A., *Pomiar preferencji metodą conjoint analysis opartą na wyborach*, [w:] *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia 9, Sekcja Klasyfikacji i Analizy Danych PTS, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 942, AE, Wrocław 2002, s. 386-399.
- Carmone F.J., Green P.E., Jain A.K., *Robustness of Conjoint Analysis: Some Monté Carlo Results*, „Journal of Marketing Research” 1978 vol. XV (May), s. 300-303.
- Carmone F.J., Green P.E., *Model Misspecification in Multiattribute Parameter Estimation*, „Journal of Marketing Research” 1981 vol. XVIII (February), s. 87-93.
- Duliniec E., *Badania marketingowe w zarządzaniu przedsiębiorstwem*, Wydawnictwo Naukowe PWN, Warszawa 1997.
- Fedorov V.V., *Theory of optimal experiments*, Academic Press, N.Y. 1972.
- Green P.E., Devita M.T., *An Interaction Model of Consumer Utility*, „Journal of Consumer Research” 1975 vol. 2 (September), s. 146-153.
- Green P.E., Wind Y., *New Way to Measure Consumers' Judgments*, „Harvard Business Review” 1975 no 53 (July-August), s. 107-117.
- Hair J.F., Anderson R.E., Tatham R.L., Black W.C., *Multivariate Data Analysis with Readings*, Prentice-Hall, Englewood Cliffs 1995.

- Kendall M.G., Buckland W.R., *Słownik terminów statystycznych*, PWE, Warszawa 1986.
- Kuhfeld W.F., *Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint, and Graphical Techniques*, <http://support.sas.com/techsup/technote/ts722.pdf>, Cary, SAS Institute 2005.
- Sagan A., *Modele zachowań konsumentów*, www.cem.pl/?a=pages&id=42, 2004.
- Steczkowski J., Zeliaś A., *Analiza wariancyjna i kowariancyjna w badaniach ekonomicznych*, PWN, Warszawa 1982.
- Walesiak M., Bąk A., *Realizacja badań marketingowych metodą conjoint analysis z wykorzystaniem pakietu statystycznego SPSS for Windows*, AE, Wrocław 1997.
- Walesiak M., Bąk A., *Conjoint analysis w badaniach marketingowych*, AE, Wrocław 2000.
- Walesiak M., Gatnar E. (red.), *Statystyczna analiza danych z wykorzystaniem programu R*, Wydawnictwo Naukowe PWN, Warszawa 2009.
- Zwerina K., *Discrete Choice Experiments in Marketing*, Physica-Verlag, Heidelberg-New York 1997.
- Żuk B., *Biometria stosowana*, PWN, Warszawa 1989.

MAIN EFFECTS MODELS AND MAIN AND INTERACTIONS MODELS IN *CONJOINT ANALYSIS* WITH APPLICATION OF R SOFTWARE

Summary: *Conjoint analysis* represents a decomposition approach. This is a statistical technique used in marketing research to determine how people value different features that make up an individual product or service. This method presents a set of different profiles of goods or services (real or not) described by attributes to the respondents. On the basis of respondents preferences a decomposition approach is conducted to extract the share of each attribute in whole profile's utility.

In the paper *conjoint analysis* procedure is shown for main effects models and interactions effects models with presenting the possibility of applying R software in steps of *conjoint analysis* procedure. The basic terms of *conjoint analysis* are presented as well as the term "interactions" and types of interactions are explained. Besides that the article presents the parts of source codes in R language that are prepared by the authors for empirical part. In the empirical part of the article *conjoint analysis* procedure is shown with the application of R software "step by step". To obtain such a goal data collected by students are used.