

MARCIN PEŁKA, ANETA RYBICKA

POMIAR I ANALIZA PREFERENCJI WYRAŻONYCH Z WYKORZYSTANIEM PAKIETU CONJOINT PROGRAMU R¹

1. WSTĘP

W badaniach preferencji konsumentów wykorzystujemy dane historyczne oraz dane wyrażające intencje, inaczej deklaracje konsumentów. Opierając się na drugim typie danych, możemy przeprowadzić badania preferencji wyrażonych. W tego typu badaniach wykorzystujemy m.in. metody dekompozycyjne, a tu przede wszystkim *conjoint analysis*.

Głównym założeniem metod *conjoint analysis* jest zaprezentowanie respondentom zbiorów obiektów – rzeczywistych lub hipotetycznych produktów lub usług – opisanych za pomocą zmiennych objaśniających, które przyjmują określone wartości (Bąk, 2004b, s. 48).

W trakcie badania respondenci oceniają prezentowane obiekty. Celem takiego badania jest pomiar preferencji konsumentów względem przedstawionych obiektów, a wynikiem pomiaru jest zbiór wartości zmiennej objaśnianej.

Model *conjoint analysis* może być szacowany:

- na poziomie indywidualnym (w przekroju respondenta),
- na poziomie zagregowanym (w przekroju całej badanej próby).

Natomiast wyniki estymacji modelu *conjoint* wykorzystywane są do (Zwerina, 1997, s. 5):

- identyfikacji preferencji,
- analizy udziałów w rynku,
- segmentacji konsumentów.

Identyfikacja preferencji wyrażonych, analiza udziałów w rynku, segmentacja konsumentów są niezwykle istotnym elementem badań ekonomicznych. W metodach tych wykorzystujemy różne oprogramowanie komputerowe. W przypadku metod *conjoint analysis* zastosowanie znajdują (Bąk 2004a, s. 220): SAS/STAT, SPSS, SYSTAT, Conjoint Analysis; Sawtooth Software (CVA), PC-MDS (PREFMAP, MONANOVA, TRADEOFF, CONJOINT), Bretton-Clark (Conjoint Analyzer, Conjoint Designer), R.

¹ Praca naukowa finansowana ze środków na naukę w latach 2009-2012 jako projekt badawczy nr N N111 446037.

Program R, ze względu na dostępność na zasadach licencji publicznej GPL, zdobywa coraz więcej zwolenników, zarówno wśród osób zajmujących się badaniami preferencji, jak i osób korzystających z metod analizy wielowymiarowej. Z uwagi na fakt, że dotychczas brakowało jednolitego pakietu pozwalającego na przeprowadzenie całej procedury *conjoint analysis* z wykorzystaniem programu R opracowany został pakiet *conjoint*² programu R. Pakiet ten stanowi odpowiedź na dotychczasowe braki i pozwala przeprowadzić całą procedurę *conjoint analysis* z wykorzystaniem programu R.

W artykule przedstawiono, wykorzystującą mikrodane, metodę badania preferencji wyrażonych, tj. *conjoint analysis*, opisano pokrótce pakiet *conjoint* programu R oraz zaprezentowano badanie konsumentów wina, które zostało przeprowadzone z wykorzystaniem pakietu *conjoint*.

2. MIKROEKONOMETRIA I MIKRODANE

Pojęcie preferencji wykorzystywane jest w celu kwantyfikacji użyteczności, której nie można bezpośrednio zmierzyć. Teorie użyteczności mieszczą się w obrębie mikroekonomii, natomiast metody badania preferencji są narzędziem badawczym mikroekonometrii. Termin „mikroekonometria” pojawia się w literaturze ekonometrycznej od połowy lat osiemdziesiątych XX w. i jest odpowiedzią na potrzebę wyodrębnienia tej części ekonometrii, która obejmuje metody wykorzystywania mikrodanych w analizie zagadnień ekonomicznych (por. np. Gruszczyński, 2010, s. 16).

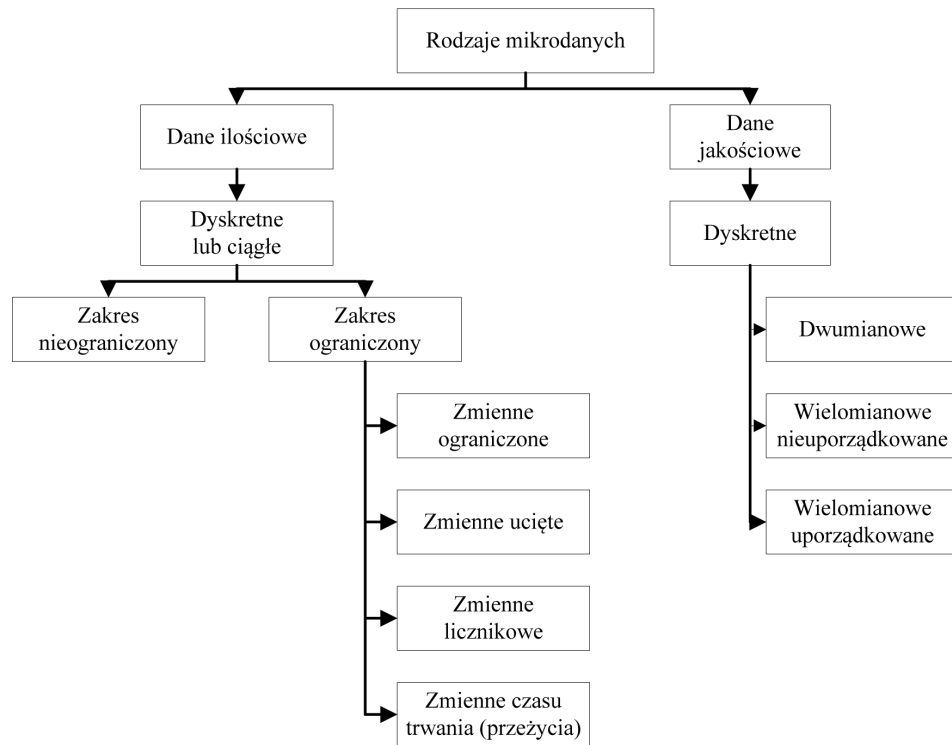
Mikroekonometria wyróżnia się kilkoma cechami (Gatnar, Walesiak, 2011, s. 112-113):

- badanie zachowań ekonomicznych jednostek,
- analiza mikrodanych na poziomie indywidualnym,
- mikrodane wykorzystywane w analizie to dane szczegółowe,
- możliwość zaobserwowania zjawisk lub zdarzeń, które są niewidoczne w danych zagregowanych (np. metody stosowane do pomiaru preferencji umożliwiają „wydobycie” ukrytych w mikrodanych informacji),
- nieliniowy rozkład obserwacji,
- wykorzystywanie nieliniowych modeli i metod estymacji parametrów,
- niejednorodność obserwacji (niejednorodność badanych jednostek),
- masowość mikrodanych,
- przekrojowy charakter mikrodanych.

Mikrodane to zbiór danych liczbowych o pojedynczych jednostkach (por. Gruszczyński, 2010, s. 13). Mikrodane są podstawą modelowania w mikroekonometrii i można klasyfikować je na różne sposoby. Jedną z możliwych klasyfikacji mikrodanych przedstawia rys. 1.

Podstawą klasyfikacji mikrodanych jest model, który objaśnia zmienną o określonej postaci (tab. 1) (Gruszczyński, 2010, s. 17-18).

² Pakiet ten opracowano w Katedrze Ekonometrii i Informatyki Uniwersytetu Ekonomicznego we Wrocławiu pod kierownictwem dr hab. Andrzeja Bąka, prof. UE.



Rysunek 1. Rodzaje mikrodanych

Źródło: Gruszczyński (2010, s. 15).

Tabela 1.

Systematyka modeli zmiennych jakościowych i ograniczonych

1. Modele dwumianowe	2. Modele wielomianowe
Liniowy model prawdopodobieństwa (LMP) Probitowy Logitowy Logarytmiczno-logitowy Gompitowy (krzywej Gomperta) Komplementarny log-log Burrítowy (rozkładu Burra) Ucięty LMP Krzywej Urbana	Modele kategorii uporządkowanych Uporządkowany model logitowy i probitowy Uogólniony model uporządkowany Modele danych sekwencyjnych Modele kategorii nieuporządkowanych Wielomianowy model logitowy i probitowy Warunkowy model logitowy (McFadden) Zagnieżdżony model logitowy Mieszany model logitowy
3. Modele zmiennych ograniczonych	4. Modele licznikowe i czasu trwania
Regresja ucięta Modele tobitowe (klasyczne) Dwugraniczny model tobitowy Model doboru próby (Heckman) Modele efektów oddziaływania	Regresja Poissonowska Model rozkładu ujemnego dwumianowego Model czasu trwania Model licznikowy ucięty Model płotkowy

Źródło: Gruszczyński (2010, s. 18).

W najszerszym ujęciu model mikroekonometryczny to model typu regresyjnego oparty na mikrodanych.

Należałoby wspomnieć, że za ważniejszych przedstawicieli mikroekonometrii uważa się laureatów Nagrody im. Alfreda Nobla z dziedziny ekonomii w roku 2000: Jamesa Heckmana i Daniela MacFaddena. Jamesa Heckman został wyróżniony nagrodą za: „rozwój teorii i metod analizy prób selektywnych”, natomiast MacFadden za: „rozwój teorii i metod analizy wyboru dyskretnego”.

3. CONJOINT ANALYSIS

Conjoint analysis to zespół numerycznych metod analizy danych dotyczących preferencji, które są oparte na założeniach wynikających z teorii pomiaru łącznego (addytywnego pomiaru łącznego – *conjoint measurement*) (zob. Bąk, 2003, s. 214).

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).

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 (zob. Green, Wind, 1975).

Następnym krokiem jest dekompozycja całkowitych preferencji i obliczenie udziału każdej ze zmiennej objaśniającej w oszacowanej całkowitej wartości użyteczności obiektu (zob. Walesiak, Bąk, 2000, s. 14). Uzyskujemy oszacowane użyteczności cząstkowe, związane z poziomami atrybutów w postaci macierzy (Bąk, 2000, s. 217).

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).

Poszczególne etapy i kroki procedury *conjoint analysis* przedstawia tab. 2.

³ Termin atrybut jest używany w statystyce w odniesieniu do zmiennych niemetrycznych, najczęściej nominalnych (zob. Kendall i 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 *conjoint analysis*.

Tabela 2.

Etapy i kroki procedury *conjoint analysis*

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

Źródło: Walesiak, Gatnar (2009, s. 292).

Do zalet metod *conjoint analysis* należą (por. *Understanding Conjoint Analysis*, 1997, s. 6-7):

- możliwość wyboru skali pomiaru preferencji,
- w miarę proste projektowanie eksperymentu czynnikowego,
- możliwość oszacowania użyteczności na poziomie indywidualnym respondenta,
- oszacowane użyteczności na poziomie indywidualnym pozwalają na łatwiejsze przeprowadzenie segmentacji rynku.

Wady tychże metod to przede wszystkim (Hair i in., 1995, s. 572):

- ograniczona liczba atrybutów wykorzystywanych w badaniu,
- sposoby oceny profilów są dalekie od rzeczywistych wyborów rynkowych konsumenta (porządkowanie obiektów w kolejności od najbardziej do najmniej satysfakcjonującego bądź też odwrotnie lub ocenę względnej atrakcyjności przedstawionych obiektów),
- nie zawsze istnieje możliwość wykorzystania częściowego eksperymentu czynnikowego,
- nie zawsze możliwa jest łączna ocena wszystkich atrybutów jednocześnie (w przypadku metod opartych na macierzach kompromisów).

Metody *conjoint analysis* nie są precyzyjnie zdefiniowaną metodą badań, ale złożoną procedurą badawczą, która pozwala na wybór różnorodnych ścieżek analizy danych. Nie można jednak wskazać wyborów optymalnych. Niektóre z możliwości mogą zależeć

od doświadczenia badacza, intuicji, posiadanego oprogramowania komputerowego bądź funduszy.

Szerzej metodę *conjoint analysis* prezentują m.in. prace Bąka (2000, 2002, 2003, 2004b), Bąka oraz Bartłomowicza (2011), Rybickiej i Pełki (2009), Walesiaka i Bąka (2000), Zweriny (1997).

4. PAKIET CONJOINT PROGRAMU R⁵

Pakiet *conjoint*, opracowany w Katedrze Ekonometrii i Informatyki Uniwersytetu Ekonomicznego we Wrocławiu pod kierownictwem prof. UE dr hab. Andrzeja Bąka, jest odpowiedzią na fakt, że nie wszystkie kroki procedury *conjoint analysis* znalazły swe odzwierciedlenie w programie R. W szczególności dotyczy to oceny ważności atrybutów oraz symulacji udziałów w rynku na etapie analizy i interpretacji wyników. Pakiet *conjoint* stanowi implementację dobrze znanej, klasycznej metody *conjoint analysis*. Pakiet ten w wersji zawiera 12 funkcji pogrupowanych w 5 kategorii (zob. tab. 3).

Tabela 3.

Etapy procedury *conjoint analysis* w programie R

Etapy procedury	Krok procedury	Wybrane pakiety i funkcje programu R
Gromadzenie danych	– dane symulacyjne	Pakiet <i>base</i> (funkcja <i>sample</i>) Pakiet <i>stats</i> (funkcja <i>runif</i>)
	– układy czynnikowe	Pakiet <i>AlgDesign</i> (funkcje: <i>gen.factorial</i> , <i>optFederov</i>)
	– próba losowa	Pakiet <i>poLCA</i> (funkcja <i>poLCA.simdata</i>)
Estymacja modelu	– model metryczny – MNK	Pakiet <i>conjoint</i> (funkcje: <i>caModel</i> , <i>caPartUtilities</i> , <i>caTotalUtilities</i>) Pakiet <i>stats</i> (funkcja <i>contrasts</i> , <i>lm</i>) Pakiet <i>base</i> (funkcja <i>factor</i>)
	– ocena ważności atrybutów	Pakiet <i>conjoint</i> (funkcje: <i>caImportance</i> , <i>caUtility</i>)
Analiza i interpretacja wyników	– symulacja udziałów w rynku	Pakiet <i>conjoint</i> (funkcje: <i>caBTL</i> , <i>caLogit</i> , <i>caMaxUtility</i> , <i>ShowAllSimulations</i>)
	– segmentacja	Pakiet <i>conjoint</i> (funkcja <i>caSegmentation</i>) Pakiet <i>stats</i> (funkcja <i>kmeans</i>)
	– wykresy ważności atrybutów i ich poziomów	Pakiet <i>conjoint</i> (funkcje: <i>Conjoint</i> , <i>caUtilities</i> , <i>ShowAllUtilities</i>) Pakiet <i>graphics</i> (funkcje: <i>barplot</i> , <i>title</i>)

Źródło: Bąk, Bartłomowicz (2011, s. 97).

⁵ Opracowano na podstawie prac Bąka i Bartłomowicza (2011, s. 96-100) oraz Rybickiej i Pełki (2009).

Funkcje pakietu *conjoint*⁶ umożliwiają: oszacowanie parametrów oraz segmentację respondentów, szacowanie użyteczności cząstkowych oraz teoretycznych użyteczności całkowitych, pomiar ważności atrybutów oraz szacowanie udziału w rynku profilów symulacyjnych. Inną kategorią funkcji są te umożliwiające uzyskanie zbiorczych wyników wybranych pomiarów.

Pakiet *conjoint* pozwala w szczególności estymować model w zakresie szacowania użyteczności cząstkowych na poziomie indywidualnym oraz zagregowanym. Standardowa funkcja programu R – *lm()* szacuje parametry modelu regresji na podstawie obserwacji zmiennej objaśnianej i zmiennych objaśniających. Natomiast w analizie *conjoint* parametry są szacowane na poziomie indywidualnym oraz parametry jednego modelu na poziomie zagregowanym. Dodatkowo pakiet ten pozwala na estymację użyteczności całkowitych profilów w przekroju respondentów, które mogą być wykorzystane w celu segmentacji respondentów. Inną cechą tego pakietu jest możliwość analizowania i interpretacji wyników w zakresie oceny istotności atrybutów oraz udziałów poszczególnych poziomów w ocenie całkowitej profilów. Pakiet ten pozwala także analizować i interpretować wyniki w zakresie analizy symulacyjnej udziałów w rynku profilów, które nie były uwzględniane w badaniu.

5. POMIAR I ANALIZA PREFERENCJI KONSUMENTÓW WINA

W okresie od marca 2011 do października 2011 roku przeprowadzono badanie dotyczące preferencji konsumentów wina z wykorzystaniem kwestionariusza ankiety. Łącznie zebrano 707 prawidłowo wypełnionych ankiet.

Próbę badawczą dobrano według wygody (w sposób przypadkowy). Dobór taki, co prawda, nie pozwala na weryfikację hipotez statystycznych, ale pozwala w pewnym, choć ograniczonym stopniu, poznać preferencje konsumentów wina (zob. Szreder, 2004, s. 48-50).

W ankiecie respondenci oceniali 17 profilów wina scharakteryzowanych trzema atrybutami: barwa wina, zawartość cukru i pochodzenie (kraj lub miejsce pochodzenia). Każda ze zmiennych przyjmuje określone poziomy (atrybuty):

- a) barwa wina – białe, czerwone, różowe,
- b) zawartość cukru – słodkie, półsłodkie, półwytrawne, wytrawne,
- c) pochodzenie – Francja, Włochy, Mołdawia, Grecja, Kalifornia (USA).

Różne kombinacje tych poziomów atrybutów pozwalają uzyskać łącznie 60 różnych profilów – jest to pełny układ czynnikowy. Z uwagi na fakt, że ocena takiej liczby profilów byłaby zadaniem trudnym dla respondenta w badaniu wykorzystano cząstkowy układ czynnikowy otrzymany z zastosowaniem funkcji *gen.factorial* z pakietu *AlgDesign*. W efekcie otrzymano 17 profilów (zob. tab. 4).

⁶ Pakiet *conjoint* można pobrać bezpłatnie ze strony internetowej: <http://keii.ue.wroc.pl/conjoint> lub ze stron internetowych CRAN (www.r-project.org).

Tabela 4.

Profile oceniane przez respondentów

Numer profilu	Barwa wina	Zawartość cukru	Pochodzenie
1	białe	półwytrawne	Francja
2	czerwone	słodkie	Francja
3	różowe	słodkie	Francja
4	czerwone	wytrawne	Włochy
5	różowe	półwytrawne	Włochy
6	czerwone	półsłodkie	Włochy
7	białe	słodkie	Włochy
8	białe	wytrawne	Mołdawia
9	różowe	półwytrawne	Mołdawia
10	czerwone	słodkie	Mołdawia
11	różowe	wytrawne	Grecja
12	czerwone	półwytrawne	Grecja
13	białe	półsłodkie	Grecja
14	czerwone	wytrawne	Kalifornia (USA)
15	białe	półwytrawne	Kalifornia (USA)
16	różowe	półsłodkie	Kalifornia (USA)
17	czerwone	słodkie	Kalifornia (USA)

Źródło: opracowanie własne.

Respondenci oceniali prezentowane im profile w skali 0 – najmniej preferuję do 10 – najbardziej preferuję, zgodnie z własnymi odczuciami.

Znajomość ocen respondentów (użyteczności całkowitych) pozwala na oszacowanie użyteczności częściowych, które określają względne znaczenie, jakie mają poszczególne poziomy atrybutów w użyteczności całkowitej. Użyteczności częściowe są szacowane w wyniku dekompozycji empirycznych użyteczności całkowitych. W tym celu szacuje się parametry modelu regresji, w którym zmienną objaśnianą jest użyteczność całkowita przypisana przez respondenta poszczególnym profilom w badaniu. W badaniu preferencji konsumentów wina wykorzystano jedynie model efektów głównych *conjoint analysis*.

Do oszacowania użyteczności częściowych wykorzystano funkcję `caPartUtilities()` z pakietu `conjoint`. Fragment wyników dla czterech respondentów przedstawia się następująco:


```
>caPartUtilities(prefmatrix, profile, levels)
```

	intercept	białe	czerwone	różowe	słodkie	półsłodkie		
[1,]	2,705	0,266	0,269	-0,534	5,331	-1,547		
[2,]	4,599	0,062	-1,524	1,462	-0,678	1,71		
[3,]	3,211	0,834	-0,068	-0,766	3,335	1,174		
[4,]	4,22	0,513	0,774	-1,287	2,699	-0,504		
	półwytrawne	wytrawne	Francja	Włochy	Mołdawia	Grecja	Kalifornia	
[1,]	-1,851	-1,932	-0,641	0,978	-0,887	0,072	0,478	
[2,]	0,697	-1,729	0,955	-0,217	-0,028	-0,491	-0,217	
[3,]	-3,225	-1,284	-0,359	1,056	-1,486	0,234	0,556	
[4,]	-0,665	-1,53	-1,464	0,086	0,612	0,68	0,086	

Uzyskane wyniki oznaczają, że według respondenta 1 najważniejszymi kryteriami (poziomami atrybutów) decydującymi o wyborze tego, a nie innego wina są: wino białe lub czerwone, najchętniej słodkie, pochodzące z Włoch (ewentualnie z Kalifornii lub ostatecznie z Grecji).

Oszacowanie parametrów modelu w przekroju całej zbiorowości respondentów przeprowadzić można za pomocą funkcji Conjoint() pakietu conjoint:

```
>Conjoint(preferencje, profile, poziomy)
```

```
[1] Part worths (utilities) of levels (model parameters for whole sample):
```

	levnms	utls
1	intercept	3,9498
2	białe	0,1148
3	czerwone	0,1724
4	różowe	-0,2872
5	słodkie	0,6728
6	półsłodkie	0,6278
7	półwytrawne	-0,3753
8	wytrawne	-0,9254
9	Francja	0,2307
10	Włochy	0,1276
11	Mołdawia	-0,2799
12	Grecja	-0,0155
13	Kalifornia	-0,063

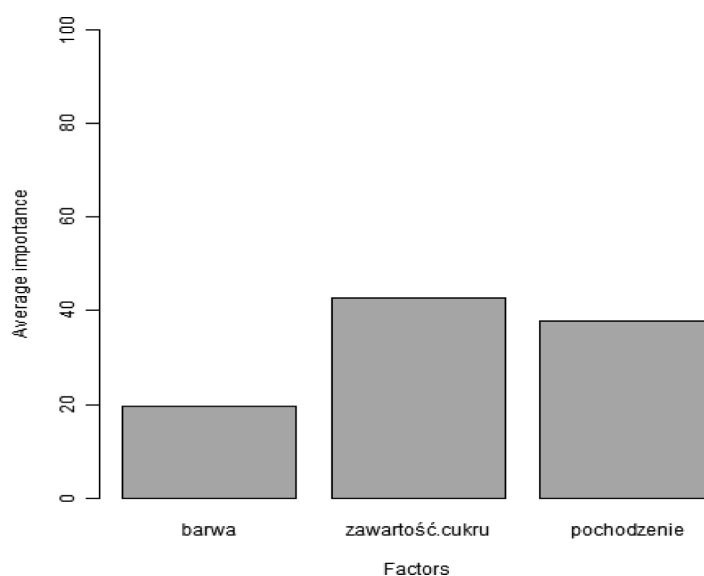
```
[1] Average importance of factors (attributes):
```

```
[1] 19,53 42,68 37,79
```

[1] Sum of average importance: 100

[1] Chart of average factors importance

Analizując wyniki na poziomie zagregowanym można wskazać, że najbardziej preferowane jest wino czerwone, słodkie pochodzące z Francji. Wykres przeciętnej ważności poszczególnych atrybutów w przekroju wszystkich respondentów przedstawia rys. 2.



Rysunek 2. Wykres przeciętnej ważności poszczególnych atrybutów dla respondentów

źródło: obliczenia własne z wykorzystaniem pakietu conjoint programu R.

Najważniejszym atrybutem decydującym o wyborze wina jest zawartość cukru (42,68%), następnie pochodzenie wina (37,79%) a na końcu jego barwa (19,53%)

Na potrzeby badania przygotowano pięć profili symulacyjnych, które nie były oceniane przez respondentów (zob. tab. 5).

Tabela 5.

Profile symulacyjne

Lp.	Barwa wina	Zawartość cukru	Pochodzenie
1	białe	słodkie	Grecja
2	różowe	półwytrawne	Francja
3	różowe	wytrawne	Włochy
4	białe	wytrawne	Kalifornia (USA)
5	białe	słodkie	Mołdawia

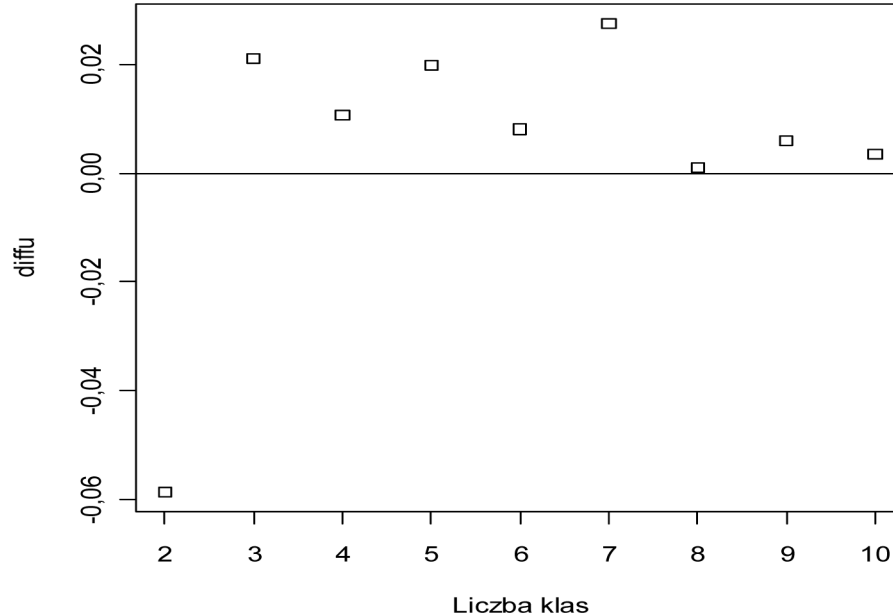
Źródło: opracowanie własne.

Następnie z wykorzystaniem funkcji `ShowAllSimulations()` pakietu `conjoint` obliczono udziały w rynku profilów symulacyjnych:

```
>ShowAllSimulations(sim, y, x)
      TotalUtility  MaxUtility  BTLmodel  LogitModel
1          4,72          26,45      22,29      25,65
2          3,98          21,07      19,17      17,56
3          4,42          21,50      21,69      22,75
4          3,08          13,58      15,67      12,98
5          4,46          17,40      21,19      21,07
```

Najwyższy udział w rynku ma profil 1 – wino białe, słodkie pochodzące z Grecji, nieco niższy udział w rynku ma profil 3 (według modelu maksymalnej użyteczności, modelu BTL i logitowego) – wino różowe, półwytrawne pochodzące z Francji.

Ostatnim krokiem w analizie *conjoint* jest segmentacja. W tym celu przygotowano skrypt w programie R obliczający podział zbioru danych od 2 do 10 klas z wykorzystaniem metody klasyfikacji *k-medoidów*. Do wyboru liczby klas wykorzystano indeks *gap* (zob. rys. 3) oprogramowany w pakiecie `clusterSim` (zob. Gatnar, Walesiak, 2004, s. 418).



Rysunek 3. Wykres wartości indeksu *gap* dla różnej liczby klas

Źródło: obliczenia własne z wykorzystaniem programu R.

Na podstawie analizy wykresu (por. rys. 3) i wartości *diffu* okazało się, że minimalna liczba klas dla której $diffu \geq 0$ wynosi 3 dla $diffu = 0,006$. Oznacza to, że zbiór danych należy podzielić na trzy klasy.

Następnie z wykorzystaniem funkcji `replication.Mod()` z pakietu `clusterSim` obliczono skorygowany indeks Randa wykorzystując 20 symulacji. Skorygowany indeks Randa wyniósł 0,5897614. Oznacza to, że mamy do czynienia z relatywnie stabilnym podziałem 707 respondentów na trzy klasy. Charakterystyki klas względem zmiennych są następujące:

a) klasa 1 jest to klasa relatywnie najbardziej jednorodna (biorąc pod uwagę odchylenia standardowe wszystkich zmiennych). W klasie tej znaleźli się respondenci którzy najbardziej preferują wina zarówno białe, czerwone jak i różowe, słodkie i półsłodkie pochodzące z Francji, Włoch, Mołdawii lub Kalifornii (profile 2, 3, 6, 7, 10, 17 z tab. 4).

Respondenci z tej klasy najmniej preferują wina wytrawne i półwytrawne, białe i czerwone, pochodzące z Francji, Włoch, Mołdawii oraz Grecji (profile 1, 4, 8, 11, 14 z tab. 4).

b) klasa 2 – jest to klasa względnie najmniej jednorodna (biorąc pod uwagę odchylenia standardowe wszystkich zmiennych). Respondenci z tej klasy preferują generalnie wina pół wytrawne, wytrawne oraz półsłodkie, zarówno białe, czerwone jak i różowe, pochodzące z Francji, Włoch, Mołdawii, Grecji oraz Kalifornii (profile 1, 4, 5, 8, 9, 11, 12, 13, 14, 15, 16 z tab. 2).

c) klasa 3 – to klasa względnie jednorodna, jeżeli weźmiemy pod uwagę jedynie profile 3 i 9. Respondenci z tej klasy najczęściej wskazywali odpowiedzi wskazujące o niewielkich preferencjach w stosunku do prezentowanych profilów. Respondenci z tej klasy najniżej oceniali zarówno wina półsłodkie, słodkie i półwytrawne, zarówno białe, jak i czerwone i różowe, pochodzące z Francji, Włoch, Mołdawii, Grecji i Kalifornii.

6. PODSUMOWANIE

W przedstawionym przykładzie empirycznym dokonano analizy preferencji 707 konsumentów wina. Najważniejszym atrybutem decydującym o zakupie win okazała się zawartość cukru. Z tego też poniekąd wynika fakt, że najbardziej preferowane są wina słodkie. Potwierdza to także udział w rynku profilów symulacyjnych oraz charakterystyka wyodrębnionych trzech klas.

Pakiet *conjoint* programu R pozwala na przygotowanie i przeprowadzenie badań preferencji konsumentów z zastosowaniem metody *conjoint analysis*. Pakiet ten jest nowym pakietem dostępnym dla programu R. Zaprojektowany został z myślą o statystykach, ekonometrykach, ekonomistach oraz studentach kierunków ekonomicznych. Zawarte są w nim funkcje pozwalające szacowanie użyteczności, oszacowanie ważności atrybutów, oszacowanie udziałów rynkowych nowych profilów (które nie brały udziału w badaniu). Dodatkowo istnieje możliwość przygotowania wielu modyfikacji tego pakietu, które poszerzą możliwości zastosowania tego pakietu – np. jest możliwe uwzględnienie interakcji dowolnego rzędu.

LITERATURA

- [1] Bąk A., (2000), *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 nr 874 Akademii Ekonomicznej we Wrocławiu, 217-226.
- [2] Bąk A., (2002), *Pomiar preferencji metodą conjoint analysis opartą na wyborach*, [w:] Jajuga K., Walesiak M. (red.) Taksonomia 9 – Klasyfikacja i analiza danych – teoria i zastosowania. Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 942, 386-399.
- [3] Bąk A., (2003), *Algorytmy conjoint analysis w pakiecie statystycznym SAS/STAT*. [w:] Jajuga K., Walesiak M. (red.) Taksonomia 10 – Klasyfikacja i analiza danych – teoria i zastosowania. Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 988, 211-221.
- [4] Bąk A., (2004a), *Metody dyskretnych wyborów w badaniach zachowań konsumentów*. [w:] Dziechciarz J. (red.) Ekonometria nr 13 – Zastosowania metod ilościowych, Prace Naukowe Akademii Ekonomicznej we Wrocławiu Nr 1010, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław, 150-163.
- [5] Bąk A., (2004b), *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1013 z serii Monografie i Opracowania nr 157, Wydawnictwo AE we Wrocławiu, Wrocław.
- [6] Bąk A., Bartłomowicz T., (2011), *Implementacja klasycznej metody conjoint analysis w pakiecie conjoint programu R*, Prace Naukowe UE we Wrocławiu nr 176, 94-104.
- [7] Gatnar E., Walesiak M., (red.) (2004), *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
- [8] Gatnar E., Walesiak M., (red.) (2011), *Analiza danych jakościowych i symbolicznych z wykorzystaniem programu R*, C.H. Beck, Warszawa.
- [9] Green P.E., Wind Y., (1975), *New way to measure consumers' judgments*, „Harvard Business Review”, July-August, 53, 107-117.
- [10] Gruszczyński M., (red.) (2010), *Mikroekonometria. Modele i metody analizy danych indywidualnych*. Wolters Kluwer Polska Sp. z o.o., Warszawa.
- [11] Hair J.F., Anderson R.E., Tatham R.L., Blach W.C., (1995), *Multivariate Data Analysis with Readings*, Prentice-Hall, Englewood Cliffs.
- [12] Rybicka A., Pełka M., (2009), *Analiza i interpretacja wyników conjoint analysis*, Prace Naukowe UE we Wrocławiu nr 86, 185-193.
- [13] Szreder M., (2004), *Metody i techniki sondażowych badań opinii*, PWE, Warszawa.
- [14] *Understanding Conjoint Analysis (1997), A Review of Conjoint Analysis*, Technical Paper from DSS Research, [URL:] www.dssresearch.com/marketresearch/Library/Conjoint/conjoint.asp.
- [15] Walesiak M., Bąk A., (2000), *Conjoint analysis w badaniach marketingowych*. Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
- [16] Walesiak M., Gatnar E., (red.) (2009), *Statystyczna analiza danych z wykorzystaniem programu R*, Wydawnictwo Naukowe PWN, Warszawa.
- [17] Zwerina K., (1997), *Discrete Choice Experiments in Marketing*, Physica-Verlag, Heidelberg-New York.

POMIAR I ANALIZA PREFERENCJI WYRAŻONYCH Z WYKORZYSTANIEM PAKIETU
CONJOINT PROGRAMU R

Streszczenie

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, względnego znaczenia każdego z nich, analizy udziałów w rynku oraz segmentacji konsumentów.

Na rynku obecnie oferowane są różne oprogramowania komputerowe pozwalające na przeprowadzenie badań preferencji konsumentów z wykorzystaniem metod *conjoint analysis*.

W artykule przedstawiono pakiet *conjoint* programu R oraz opracowane pakiety i funkcje programu R, niezbędne w prowadzeniu empirycznych badań preferencji.

Program R, ze względu na dostępność na zasadach licencji GPL, zdobywa coraz więcej zwolenników, zarówno wśród osób zajmujących się badaniami preferencji, jak i osób korzystających z metod analizy wielowymiarowej.

Natomiast pakiet *conjoint* jest odpowiedzią na fakt, że nie wszystkie kroki procedury *conjoint analysis* znalazły swe odzwierciedlenie w programie R. W szczególności dotyczy to oceny ważności atrybutów oraz symulacji udziałów w rynku na etapie analizy i interpretacji wyników.

W artykule przedstawiono również wyniki badania preferencji konsumentów wina z wykorzystaniem pakietu *conjoint*, analizę udziałów w rynku wybranych profili symulacyjnych oraz segmentację konsumentów wina.

Słowa kluczowe: pomiar i analiza preferencji, *conjoint analysis*, program R

THE EVALUATION AND ANALYSIS OF REVEALED PREFERENCES WITH APPLICATION OF R
PACKAGE CONJOINT

Abstract

Conjoint analysis is a statistical method in which consumer preferences are decomposed in order to evaluate: utility function for each attribute, importance of each attribute, market shares simulations and segmentation of consumers.

There are many different computer programs that can be applied for *conjoint analysis* research.

The paper presents *conjoint* package of R software which are useful to evaluate empirical preferences. The R program is more and more popular and many researchers are applying it.

The *conjoint* package of R software is a response to a fact, that not all steps of *conjoint analysis* were programmed in R. In particular it concerns evaluation of attributes' importance, market share simulations and interpretation of results.

The article presents also results of the evaluation of wine consumers' preferences with application of *conjoint* package, market share simulations and segmentation of consumers.

Key words: evaluation and analysis of preference, *conjoint analysis*, R software