

Aneta Rybicka, Marcin Pełka

Uniwersytet Ekonomiczny we Wrocławiu

ANALIZA I INTERPRETACJA WYNIKÓW *CONJOINT ANALYSIS*

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* ze szczególnym uwzględnieniem etapu analizy i interpretacji wyników z zastosowaniem programu R. W części empirycznej artykułu przedstawiono analizę i interpretację wyników badania *conjoint analysis* z wykorzystaniem programu R na podstawie danych zebranych przez studentów na użytek prac magisterskich.

2. Procedura badawcza *conjoint analysis*

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 motywy 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].

W badaniach preferencji wyrażonych wykorzystuje się dwa główne podejścia:

1. **Podejście kompozycyjne** (*compositional approach*) wykorzystujące ideę 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 2000, s. 14; Zwerina 1997, s. 3];

2. **Podejście dekompozycyjne** (*decompositional approach*), gdzie w celu przeprowadzenia analizy preferencji konsumentów wykorzystujemy metody *conjoint analysis* oraz metody oparte na wyborach [Bąk 2000, s. 76]. Modele dekompozycyjne to klasa modeli, które „rozkładają” preferencje całkowite konsumentów.

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).

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

Rys. 1. Zależność między atrybutami, poziomami i profilami

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

¹ Model ten znajduje się w nurcie teorii poznawczych, których przedstawicielami w marketingu są również Lutz, Bettman, Rosenberg, Wilkie i Pessemier. 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. Z tego modelu wynika, że wybory produktów są funkcją wyrażonych intencji zakupów, które zależą m.in. od siły postaw.

Zgodnie z terminologią stosowaną w literaturze przedmiotu zmienne objaśniające opisujące dobra lub usługi nazywa się atrybutami (*attributes*)² lub czynnikami (*factors*), natomiast ich realizacje są nazywane **poziomami** (*levels*)³. 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 (zob. [Walesiak, Gatnar 2009, s. 287]).

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 profili parami, prezentacja par atrybutów, dane symulacyjne) • metody generowania profili (układy czynnikowe, próba losowa)
4	Prezentacja profili	<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].

W procedurze badawczej *conjoint analysis* można wyróżnić siedem głównych etapów (por. tab. 1).

² Termin „atrybut” jest używany w statystyce w odniesieniu do zmiennych niemetrycznych, najczęściej nominalnych [Kendall, Buckland 1986, s. 13].

³ 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*.

3. Etap analizy i interpretacji wyników *conjoint analysis*

Ogólna procedura modelowania symulacyjnego rynku w zakresie analizy udziałów w rynku na podstawie wyników uzyskanych metodą *conjoint analysis* obejmuje następujące etapy (zob. [Walesiak, Bąk 2000, s. 73; Bąk 2004]):

- oszacowanie indywidualnych użyteczności cząstkowych poziomów atrybutów w przekroju respondentów,
- wygenerowanie zbiorów testowych profili, będących przedmiotem modelowania symulacyjnego,
- modelowanie symulacyjne wyborów respondentów odnośnie do wygenerowanych profili testowych i wyborów optymalnych konfiguracji.

Atrakcyjność (użyteczność całkowita) (por. [Walesiak, Bąk 2000, s. 73-74; Walesiak, Gatnar 2009, s. 302-303]) dla poszczególnych respondentów i dla całego zbioru można obliczyć, stosując wzory:

$$U_{is} = \sum_{j=1}^m U_{jl_j}^s + b_{0s}, \quad (1)$$

gdzie: $U_{jl_j}^s$ – użyteczność cząstkowa l -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

lub dla całego zbioru respondentów:

$$U_i = \frac{1}{S} \sum_{s=1}^S \left(\sum_{j=1}^m U_{jl_j}^s + b_{0s} \right). \quad (2)$$

Wskaźniki relatywnego (względego) znaczenia atrybutów (zmiennych) łącznie oceniają cechę produktu względem pozostałych. Wskaźniki te konstruuje się na podstawie użyteczności charakteryzujących respondenta, natomiast uzyskane na ich podstawie wartości średnie odnoszą się do poszczególnych atrybutów produktu łącznie.

Relatywną ważność każdej zmiennej dla respondenta s wyznacza się za pomocą formuły (zob. [Walesiak, Bąk 2000, s. 74; Walesiak, Gatnar 2009, s. 303]):

$$W_j^s = \frac{\max_{l_j} \{U_{jl_j}^s\} - \min_{l_j} \{U_{jl_j}^s\}}{\sum_{j=1}^m \left(\max_{l_j} \{U_{jl_j}^s\} - \min_{l_j} \{U_{jl_j}^s\} \right)}. \quad (3)$$

Za najważniejszy atrybut dla danego respondenta uznaje się tę cechę produktu, której wskaźnik relatywnej ważności jest największy. Wartości średnie wskaźnika wyznaczone dla wszystkich konsumentów pozwalają określić ranking poszczególnych atrybutów w całej grupie badanej. Średnią ważność zmiennych oblicza się na podstawie wzoru (por. [Walesiak, Bąk 2000, s. 74; Walesiak, Gatnar 2009, s. 303]):

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

Prognozowanie udziałów w rynku zaprojektowanych profili szacuje się, wykorzystując do tego modele decyzyjne, nazywane również regułami wyboru, do których należą (zob. [Walesiak, Bąk 2000, s. 86; Walesiak, Gatnar 2009, s. 303-304]):

1. Reguła maksymalnej użyteczności („pierwszej pozycji”), w której oblicza się odsetek respondentów, dla których dany produkt otrzymał najwyższą ocenę użyteczności całkowitej wśród profili będących przedmiotem symulacji. Model ten ma charakter deterministyczny, ponieważ zakłada, że konsument zawsze wybierze profil o największej użyteczności. Model ten przedstawia się wzorem:

$$P_{is} = \begin{cases} 1, & \text{gdy } \hat{U}_{is} = \max(\hat{U}_{is}), \\ 0, & \text{w przeciwnym razie} \end{cases}, \quad (5)$$

gdzie: P_{is} – prawdopodobieństwo wyboru i -tego profilu przez s -tego respondenta,
 \hat{U}_{is} – oszacowana użyteczność całkowita i -tego profilu dla s -tego respondenta.

2. Probabilistyczna reguła BTL (Bradleya-Terry’ego-Luce’a), w której użyteczność całkowitą odpowiadającą danemu produktowi dzieli się przez sumę użyteczności całkowitych profili będących przedmiotem symulacji (obliczenia wykonuje się osobno dla każdego respondenta, a następnie oblicza się ich wartość przeciętną):

$$P_{is} = \hat{U}_{is} / \sum_{i=1}^n \hat{U}_{is}. \quad (6)$$

3. Probabilistyczna reguła logitowa. Model logitowy różni się od modelu BTL tym, że w obliczeniach stosuje się logarytmy naturalne wartości użyteczności całkowitych zamiast samych użyteczności:

$$P_{is} = \exp(\hat{U}_{is}) / \sum_{i=1}^n \exp(\hat{U}_{is}). \quad (7)$$

Probabilistyczna reguła BTL oraz probabilistyczna reguła logitowa bazują na prawdopodobieństwie wyboru określonego profilu wyliczonym na podstawie użyteczności całkowitych otrzymanych w wyniku zastosowania procedury *conjoint analysis*.

Dodatkowo można oszacować użyteczności całkowite oraz symulowane udziały w rynku profili, które nie były oceniane przez respondentów. Profile te, nazywane **profilami symulacyjnymi**, mogą być utworzone już po oszacowaniu użyteczności częściowych. Profile symulacyjne mogą być wygenerowane w planie eksperymentu

czynnikowego albo mogą być zdefiniowane przez badacza na podstawie przesłanek merytorycznych. Nie są one przedstawiane respondentom do oceny. Użyteczności całkowite profilów symulacyjnych są oszacowane na podstawie użyteczności cząstkowych uzyskanych dla profilów uwzględnionych w badaniu. Tak obliczone użyteczności całkowite profilów symulacyjnych stanowią podstawę do wyznaczenia ich udziałów w rynku.

W programie R można wykorzystać różne pakiety i funkcje w celu przeprowadzenia badania *conjoint analysis* (zob. [Walesiak, Gatnar 2009, s. 304-305]). Dodatkowo na użytek pracy [Walesiak, Gatnar 2009] przygotowane zostały dodatkowe funkcje wykorzystywane w badaniach *conjoint analysis* (zob. [Walesiak, Gatnar 2009, s. 306-311]).

4. Przykład empiryczny

Przedstawiony w artykule przykład ma jedynie charakter ilustracyjny. Wykorzystano dane zebrane przez Małgorzatę Baran na użytek pracy magisterskiej. Dane dotyczą preferencji konsumentów podejmujących decyzje o zakupie herbaty. Na potrzeby badania określono następujące atrybuty i ich poziomy:

1. Cena – niska, średnia, wysoka.
2. Gatunek – czarna, zielona, czerwona.
3. Rodzaj – liściasta, granulowana, ekspresowa.
4. Aromatyzowana – tak, nie.

W ankiecie przedstawiono do oceny 13 profilów stanowiących cząstkowy (por. tab. 2) układ czynnikowy wygenerowany w programie R. Dane zebrano wśród 100 osób, które dobrano w sposób przypadkowy.

Tabela 2. Zbiór profilów przedstawiony do oceny respondentom

Lp.	Cena	Gatunek	Rodzaj	Aromatyzowana
1	wysoka	czarna	liściasta	tak
2	niska	zielona	liściasta	tak
3	średnia	zielona	granulowana	tak
4	średnia	czarna	ekspresowa	tak
5	wysoka	czerwona	ekspresowa	tak
6	średnia	czarna	liściasta	nie
7	wysoka	zielona	liściasta	nie
8	średnia	czerwona	liściasta	nie
9	wysoka	czarna	granulowana	nie
10	niska	czerwona	granulowana	nie
11	niska	czarna	ekspresowa	nie
12	średnia	zielona	ekspresowa	nie
13	wysoka	zielona	ekspresowa	nie

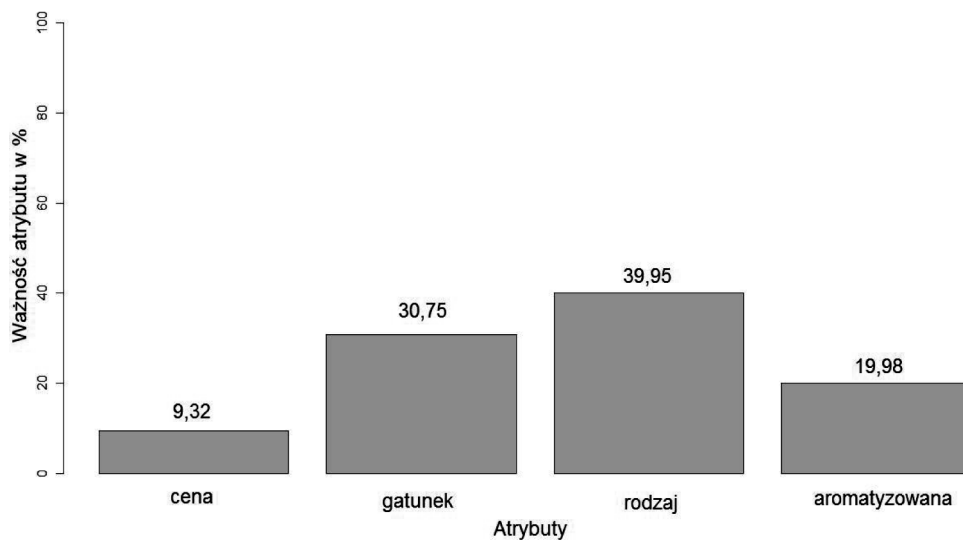
Źródło: obliczenia własne z wykorzystaniem programu R.

Z wykorzystaniem procedur i funkcji programu R oszacowano użyteczności całkowite (atrakcyjność) profili (zob. wzory 1 i 2) liniowego addytywnego modelu *conjoint analysis* efektów głównych. Wyniki oszacowania zestawiono w tab. 3.

Tabela 3. Użyteczność całkowita profili

Lp.	Cena	Gatunek	Rodzaj	Aromatyzowana	Użyteczność całkowita
1	wysoka	czarna	liściasta	tak	4,62
2	niska	zielona	liściasta	tak	4,37
3	średnia	zielona	granulowana	tak	2,96
4	średnia	czarna	ekspresowa	tak	5,19
5	wysoka	czerwona	ekspresowa	tak	3,96
6	średnia	czarna	liściasta	nie	3,76
7	wysoka	zielona	liściasta	nie	3,22
8	średnia	czerwona	liściasta	nie	2,49
9	wysoka	czarna	granulowana	nie	2,77
10	niska	czerwona	granulowana	nie	1,85
11	niska	czarna	ekspresowa	nie	4,75
12	średnia	zielona	ekspresowa	nie	3,79
13	wysoka	zielona	ekspresowa	nie	3,84

Źródło: obliczenia własne z wykorzystaniem programu R.



Rys. 2. Wykres ważności atrybutów

Źródło: obliczenia własne z wykorzystaniem programu R.

Na podstawie danych zestawionych w tab. 3 można stwierdzić, że najbardziej atrakcyjnym profilem wśród wszystkich respondentów jest profil 4 (herbata aromatyzowana, czarna, ekspresowa, o średniej cenie), natomiast profilem najmniej atrakcyjnym jest profil 10 (herbata niearomatyzowana, czerwona, granulowana, o niskiej cenie).

Na podstawie wzorów 3 i 4 z wykorzystaniem programu R obliczono ważność atrybutów (zob. rys. 2).

Z rysunku 2 wynika, że respondenci, dokonując wyboru o zakupie herbaty, na pierwszym miejscu biorą pod uwagę rodzaj herbaty, następnie rozważają jej gatunek, aromatyzowanie, a na końcu cenę. Być może wynika to z faktu, że w ankiecie atrybut ceny miał tylko trzy poziomy (niska, średnia, wysoka), które mogły zostać zinterpretowane w różny sposób przez różnych respondentów.

Na użytek modeli decyzyjnych przygotowano pięć modeli symulacyjnych, które wygenerowano w planie eksperymentu czynnikowego. Z wykorzystaniem modeli maksymalnej użyteczności, BTL i logitowego (por. wzory 5-7) oszacowano prawdopodobieństwa wyboru (zob. tab. 4).

Tabela 4. Profile symulacyjne oraz prawdopodobieństwo ich wyboru

Lp.	Cena	Gatunek	Rodzaj	Aromatyzowana	Modele		
					maksymalnej użyteczności	BTL	logitowy
1	niska	czarna	liściasta	tak	39%	34,16%	36,19%
2	niska	czerwona	liściasta	tak	18%	19,38%	13,67%
3	średnia	zielona	ekspresowa	tak	33%	21,67%	29,90%
4	niska	czerwona	ekspresowa	nie	6%	14,35%	13,87%
5	wysoka	zielona	granulowana	nie	4%	10,44%	6,37%

Źródło: obliczenia własne z wykorzystaniem programu R.

We wszystkich modelach największe prawdopodobieństwo wyboru ma profil pierwszy symulacyjny (herbata aromatyzowana, czarna, liściasta o niskiej cenie). Natomiast profil 5 jest profilem o najmniejszym prawdopodobieństwie wyboru.

5. Wnioski końcowe

Na użytek pracy [Walesiak, Gatnar 2009] przygotowano kody źródłowe i funkcje dla programu R dla niektórych etapów procedury *conjoint analysis* (generowanie profili, szacowanie użyteczności cząstkowych, estymacja modelu uwzględniającego interakcje, kodowanie *quasi*-eksperymentalne oraz zero-jedynkowe, modelowanie symulacyjne).

Najbardziej atrakcyjnym profilem okazał się profil 4 (herbata czarna, ekspresowa, aromatyzowana w średniej cenie), a najmniej atrakcyjnym profil 10 (herbata czerwona, granulowana, niearomatyzowana, w niskiej cenie).

Przy podejmowaniu decyzji o zakupie herbaty przez respondentów najważniejszym atrybutem okazał się rodzaj herbaty (następnie: gatunek, aromatyzacja i cena). Spośród wybranych profilów symulacyjnych największym prawdopodobieństwem zakupu charakteryzowała się herbata czarna, liściasta, aromatyzowana, w niskiej cenie.

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

Literatura

- Bąk A., *Conjoint analysis jako metoda pomiaru postaw i preferencji konsumentów*, [w:] *Pomiar w badaniach rynkowych i marketingowych*, red. M. Walesiak, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 856, AE, Wrocław 2000, s. 69-81.
- Bąk A., *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, AE, Wrocław 2004.
- 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.
- Kendall M.G., Buckland W.R., *Słownik terminów statystycznych*, PWE, Warszawa 1986.
- 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.

ANALYSIS AND INTERPRETATION OF *CONJOINT ANALYSIS* RESULTS

Summary

The paper presents an analysis and interpretation of *conjoint analysis* results with a possibility of applying R software. The basic terms of *conjoint analysis* are presented. In the empirical part of the article, *conjoint analysis* procedure with application of R software is applied. To obtain such a goal data collected by students was used.