

Aneta Rybicka

Akademia Ekonomiczna we Wrocławiu

BADANIE PREFERENCJI KLIENTÓW BIUR PODRÓŻY Z WYKORZYSTANIEM METOD WYBORÓW DYSKRETNÝCH

1. Wstęþ

Badaniem zachowań konsumenta oraz poszczególnymi jego elementami zainteresowano się już w XIX w. To właśnie w tym czasie po raz pierwszy sformułowano pojęcie użyteczności, która oznacza subiektywnie odczuwaną satysfakcję z powodu realizacji określonej struktury konsumpcji. Zajmowano się również problematyką pomiaru użyteczności.

W związku z tym, iż bezpośredni pomiar użyteczności jest bardzo trudny, wprowadzono do teorii użyteczności pojęcie preferencji, które umożliwia pewną kwantyfikację użyteczności.

W pomiarze preferencji wykorzystywane są dane historyczne oraz dane o charakterze antycypacyjnym (które opisują intencje konsumentów). Dlatego też rozróżnia się metody analizy preferencji ujawnionych oraz metody analizy preferencji wyrażonych.

W artykule przedstawiono badanie preferencji klientów biur podróży z wykorzystaniem metod wyborów dyskretnych (reprezentujących dekompozycyjne podejście badania preferencji wyrażonych). Badanie to zostało przeprowadzone na przełomie lipca i sierpnia 2005 r. na terenie Jeleniej Góry. W badaniu do analizy preferencji konsumentów mierzonych na skali nominalnej wykorzystano warunkowy model logitowy (*conditional logit model*).

2. Metody wyborów dyskretnych

Proces wyboru między profilami ma charakter probabilistyczny, ponieważ nabywcy nie zawsze postępują w sposób konsekwentny. Stanowi to podstawę teorii

użyteczności losowej (model ocen porównawczych) (*random utility theory*), która została przedstawiona przez Thurstone'a w 1927 r. [Bąk 2004a, s. 106]. Do literatury marketingu modele te zostały wprowadzone przez Louviere'a i Woodwortha w 1983 r. Teoria użyteczności losowej jest podstawą teoretyczną metod wyborów dyskretnych (*choice-based conjoint analysis, discrete choice modeling, discrete choice experiment*).

Gdy w badaniach wykorzystujemy metody wyborów dyskretnych, to [Bąk 2004b, s. 156]:

- liczba uwzględnianych w badaniu atrybutów zawiera się w przedziale 6-8, a liczba poziomów każdego z atrybutu nie przekracza 9-15,
- modelem zależności zmiennych może być model efektów głównych bądź model uwzględniający interakcje pomiędzy atrybutami,
- modelem preferencji jest zazwyczaj probabilistyczny model nieliniowy,
- metodą prezentacji danych jest metoda wyboru ze zbiorów profili pełnych, częściowych lub blokowych (metoda wyborów dyskretnych),
- stosujemy zazwyczaj częściowy eksperyment czynnikowy,
- skalą pomiaru jest zwykle skala nominalna (niemetryczna),
- stosujemy jednostopniowe i dwustopniowe procedury generowania układów czynnikowych.

Metody estymacji parametrów w metodach wyborów dyskretnych to metody bazujące na prawdopodobieństwie wyboru. Do tej grupy należą metody LOGIT i PROBIT. Najczęściej stosowanymi modelami są [Bąk 2004b, s. 157-160]: wielomianowy model logitowy (*MultiNominal Logit model* – MNL) – jeżeli zmienne objaśniające charakteryzują konsumentów, lub warunkowy model logitowy (*Conditional Logit model* – CLM) – jeżeli zmienne objaśniające opisują obiekty będące przedmiotem wyboru. Nieznane parametry modelu szacuje się uogólnioną metodą najmniejszych kwadratów lub metodą największej wiarygodności.

Użyteczności częściowe szacowane są na poziomie indywidualnym (estymacja na poziomie segmentowym oraz poziomie zagregowanym jest możliwa w przypadku zastosowania modeli klas ukrytych – w przypadku estymacji segmentowej, bądź modeli hierarchicznych Bayesa – w przypadku estymacji na poziomie zagregowanym) [Bąk 2004b, s. 110, s. 142-145; Huber 1998, s. 1; Orme i Hefft 1999, s. 4; Orme i Baker 2000, s. 1; Johnson 1997, s. 207-208].

Do najważniejszych zalet metod wyborów dyskretnych należą [Bąk 2002, s. 395-396]:

- symulacja bardziej realistycznej sytuacji wyboru,
- możliwość rezygnacji z wyboru,
- łączna ocena wszystkich atrybutów,
- wymagają od respondentów mniejszej liczby rozstrzygnięć niż metody *conjoint analysis*,

- wartości estymatorów parametrów modelu są obliczane na poziomie zagregowanym (pozwala to na oszacowanie udziałów w rynku oraz prognozowanie udziałów nowo wprowadzanych profili).
- Metody wyborów dyskretnych mają również wady [Bąk 2002, s. 396]:
- procedura wyborów prowadzi do zgromadzenia danych o mniejszej zawartości informacyjnej (w porównaniu z procedurą ocen rankingowych lub pozycyjnych),
- eksperyment oparty na wyborach wymaga zgromadzenia dużej liczby obserwacji (w celu uzyskania rzetelnych estymatorów użyteczności cząstkowych),
- nie umożliwiają oszacowania odrębnych modeli użyteczności dla każdego respondenta (nie jest zatem możliwa segmentacja).

Oprogramowanie komputerowe wykorzystywane w badaniach preferencji konsumentów z wykorzystaniem metod wyborów dyskretnych to przede wszystkim: SPSS, SAS/STAT, Statistica, S-Plus; Glimmix, Latent Gold (modele klas ukrytych); Sawtooth Software CBC/HB, autorskie programy źródłowe w języku: C, GAUSS, Fortran (modele hierarchiczne Bayesa).

3. Badanie preferencji klientów biur podróży

Przykładem zastosowania metod wyborów dyskretnych w badaniach preferencji konsumentów jest przeprowadzone na przełomie lipca i sierpnia 2005 r. na terenie Jeleniej Góry badanie preferencji klientów biur podróży.

Zgromadzone dane zostały wykorzystane do:

- identyfikacji czynników, które wywierały decydujący wpływ na wybór określonego biura podróży oraz określonej formy i rodzaju wypoczynku,
- analizy struktury wyborów formy i rodzaju wypoczynku w zakresie ilościowym i wartościowym, przestrzennym, rodzajowym,
- analizy częstotliwości i okoliczności wyborów formy i rodzaju wypoczynku.
- Ankieta składała się z czterech części:
- pytania w pierwszej części dotyczyły wyboru biura podróży,
- pytania w drugiej części dotyczyły wyboru formy i rodzaju wypoczynku,
- trzecia część ankiety zawierała profile, które charakteryzowały formę i rodzaj wypoczynku, spośród których respondent wybierał ten, który najlepiej spełniał oczekiwania, lub rezygnował z wyboru,
- pytania w czwartej części dotyczyły podstawowych cech socjologicznych badanych respondentów.

W badaniu zostały wykorzystane atrybuty i ich poziomy przedstawione w tab. 1.

Pełny eksperyment czynnikowy zawierałby 3^{45} zbiorów (5 profili opisanych 4 atrybutami zawierającymi po 3 poziomy). Zatem liczba zbiorów w pełnym eksperymencie wyniosłaby 3 486 784 401. Minimalny rozmiar eksperymentu zawierałby

$20 \cdot (3-1) + 1 = 41$ zbiorów, a rozmiar eksperymentu wykorzystanego w badaniu wyniósł 45 zbiorów. Efektywność układu czynnikowego wyniosła $D = 85,96^1$.

Tabela 1. Atrybuty i ich poziomy charakteryzujące formę i rodzaj wypoczynku

Atrybuty	Poziomy
Miejsce	morze, jezioro, góry, zwiedzanie, obojętnie
Zakwaterowanie	camping (bez wyżywienia), apartament (bez wyżywienia), hotel (śniadanie)
Transport	samochód, samolot, autokar
Cena	do 1500 zł (za tydzień), od 1500 do 3000 zł (za 2 tygodnie), powyżej 3000 zł (za 2 tygodnie)
Kraj	w Polsce, w Europie, poza Europą

Źródło: opracowanie własne.

W badaniu wykorzystano 3 bloki. Liczba zbiorów w każdym z bloków wyniosła 15. Liczba profilów w pełnym eksperymencie równała się 405 (5.3.3.3.3). Każdy ze zbiorów zawierał 6 profilów (5 + opcja „żaden z tych profilów”).

Tabela 2. Zbiór profilów charakteryzujących formę i rodzaj wypoczynku

1. Wybierz preferowany profil formy i rodzaj wypoczynku lub zrezygnuj z wyboru (zaznacz jedną z 6 opcji):					
Miejsce	Zakwaterowanie	Transport	Cena	Kraj	Wybieram opcję
Morze	Camping (bez wyżywienia)	Samochód	Powyżej 3000 zł (za 2 tygodnie)	Poza Europą	1
Jezioro	Camping (bez wyżywienia)	Samolot	Od 1500 do 3000 zł (za 2 tygodnie)	W Europie	2
Góry	Apartament (bez wyżywienia)	Samolot	Od 1500 do 3000 zł (za 2 tygodnie)	W Europie	3
Zwiedzanie	Apartament (bez wyżywienia)	Autokar	Do 1500 zł (za tydzień)	W Europie	4
Obojętnie	Camping (bez wyżywienia)	Samolot	Powyżej 3000 zł (za 2 tygodnie)	W Europie	5
Żaden z tych profilów					6

Źródło: opracowanie własne.

Liczba profilów wykorzystanych w badaniu równała się zatem 270 (225 profili charakteryzujących formę wypoczynku oraz 45 profili opcji rezygnacji z wyboru). Rozprowadzono 440 ankiet, uzyskano zaś 394 prawidłowo wypełnione (wykorzystane w badaniu). Liczba ankiet w poszczególnych blokach przedstawiała

¹ Miarą D -efektywności jest funkcja średniej geometrycznej wektora wartości własnych macierzy informacyjnej układu czynnikowego, gdzie kryterium wyboru układu czynnikowego jest maksymalna wartość wyznacznika macierzy informacyjnej, gdyż im bliższy zera jest wyznacznik, tym silniej skorelowane są czynniki planu [Bąk 2004a, s. 86-87]. Wartości miar efektywności dla odpowiednio zakodowanych planów czynnikowych są wyskalowane w przedziale (0; 100), przy czym wyższe wartości wskazują na efektywniejszy układ czynnikowy oraz informują o względnej optymalności układu w odniesieniu do hipotetycznego układu ortogonalnego.

się następująco: blok 1 – 130, blok 2 – 130, blok 3 – 134 ankiety. Łącznie zgromadzano 35 460 obserwacji (15 zbiorów · 6 profilów · 394 respondentów).

W wyniku estymacji warunkowego modelu logitowego uzyskano oszacowania parametrów, które zostały przedstawione w tab. 3.

Tabela 3. Estymacja parametrów modelu

Variable		DF	Parametr Estimate	Standard Error	Chi-Square	Pr > Chi-Square	Hazard Ratio
Miejsce	morze	1	0,49417	0,06153	64,5034	<0,0001	1,639
	jezioro	1	0,22887	0,06435	12,6484	0,0004	1,257
	góry	1	-0,04748	0,06488	0,5355	0,4643	0,954
	zwiedzanie	1	-0,13182	0,06522	4,0856	0,0433	0,876
	obojętnie	1	-0,05339	0,06519	0,6561	0,4179	0,948
	brak	0	0				
Zakwaterowanie	brak	0	0				
	apartament (bez wyżywienia)	1	-0,08058	0,03836	4,4135	0,0357	0,923
	camping (bez wyżywienia)	1	-0,04234	0,03896	1,1813	0,2771	0,959
	hotel (śniadanie)	0	0				
Transport	brak	0	0				
	autokar	1	-0,03180	0,03738	0,7239	0,3949	0,969
	samochód	1	-0,12988	0,03938	10,8751	0,0010	0,878
	samolot	0	0				
Cena	brak	0	0				
	do 1500 zł (za tydzień)	1	0,04791	0,03924	1,4909	0,2221	1,049
	od 1500 do 3000 zł (za 2 tygodnie)	1	0,05903	0,03991	2,1882	0,1391	1,061
	powyżej 3000 zł (za 2 tygodnie)	0	0				
Kraj	brak	0	0				
	poza Europą	1	-0,10509	0,03928	7,1568	0,0075	0,900
	w Europie	1	0,30420	0,03692	67,8793	<0,0001	1,356
	w Polsce	0	0				

Źródło: opracowanie własne.

W wyniku estymacji oszacowano użyteczności cząstkowe poziomów atrybutów. Użyteczności te zostały wykorzystane do obliczenia użyteczności całkowitych każdego z prezentowanych profilów. Następnym krokiem było obliczenie prawdopodobieństwa wyboru profilów w każdym ze zbiorów oraz prawdopodobieństwa wyboru każdego profilu z całego ocenianego zestawu. Tabela 4 przedstawia 5 profilów o największych prawdopodobieństwach wyboru i 5 profilów o najmniejszych prawdopodobieństwach wyboru spośród wszystkich 270 ocenianych profilów.

W związku z tym, iż wartości oszacowanych współczynników regresji logistycznej (β_j) nie są interpretowane, możliwości interpretacyjne stwarza prze-

kształcenie oszacowanego równania regresji logistycznej w iloraz szans (*hazard ratio*) [Rószkiewicz 2002, s. 177].

Iloraz szans przedstawia relatywną możliwość wystąpienia zdarzenia, a w przypadku regresji logistycznej poziom szans można oszacować jako funkcję zmiennych niezależnych, upraszczając iloraz szans do postaci [Rószkiewicz 2002, s. 178]:

$$\psi = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}.$$

Wyrażenie e^{β_j} określa relatywną zmianę możliwości wystąpienia zdarzenia w wyniku działania czynnika opisanego przez zmienną X_j (zakładając kontrolowanie, inaczej stabilność, pozostałych zmiennych uwzględnionych w równaniu).

Tabela 4. Profile o największym i najmniejszym prawdopodobieństwie wyboru

Miejsce	Zakwaterowanie	Transport	Cena	Kraj	Prawdopodobieństwo
największe prawdopodobieństwo wyboru					
Morze	hotel (śniadanie)	samolot	do 1500 zł (za 1 tydzień)	w Europie	0,007562115
Morze	hotel (śniadanie)	autokar	od 1500 do 3000 zł (za 2 tygodnie)	w Europie	0,007407292
Morze	camping (bez wyżywienia)	samolot	od 1500 do 3000 zł (za 2 tygodnie)	w Europie	0,007329619
Morze	hotel (śniadanie)	samolot	powyżej 3000 zł (za 2 tygodnie)	w Europie	0,007208337
Morze	camping (bez wyżywienia)	autokar	od 1500 do 3000 zł (za 1 tydzień)	w Europie	0,007021681
najmniejsze prawdopodobieństwo wyboru					
Obojętnie	apartament (bez wyżywienia)	samochód	od 1500 zł (za 1 tydzień)	poza Europą	0,002353367
Zwiedzanie	apartament (bez wyżywienia)	autokar	powyżej 3000 zł (za 2 tygodnie)	poza Europą	0,002287765
Zwiedzanie	camping (bez wyżywienia)	samochód	do 1500 zł (za 1 tydzień)	poza Europą	0,002260641
Góry	apartament (bez wyżywienia)	samochód	powyżej 3000 zł (za 2 tygodnie)	poza Europą	0,002256560
Zwiedzanie	apartament (bez wyżywienia)	samochód	do 1500 zł (za 1 tydzień)	poza Europą	0,002175834

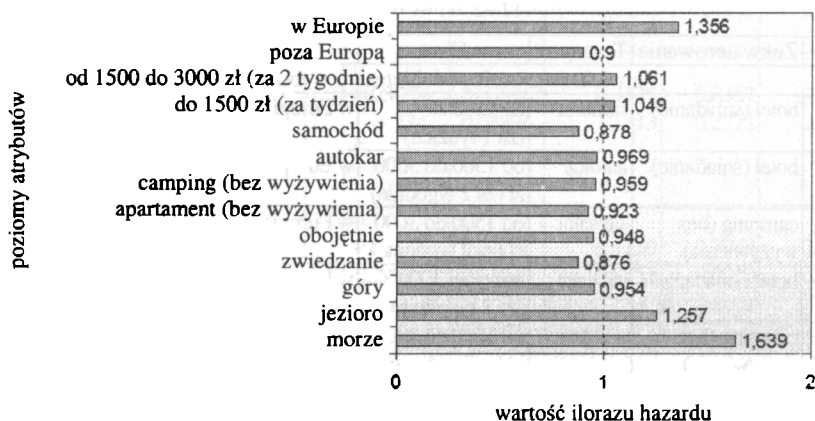
Źródło: opracowanie własne.

Otrzymany iloraz szans porównuje się z wartością 1 i interpretuje następująco [Rószkiewicz 2002, s. 178]:

- jeśli $e^{\beta_j} > 1$, to przyjmuje się, że czynnik, który jest opisywany przez zmienną niezależną X_j , działa stymulująco na prawdopodobieństwo wystąpienia badanego zjawiska (przy kontrolowanym wpływie pozostałych zmiennych uwzględnionych w równaniu);

- jeśli $e^{\beta_j} < 1$, to przyjmuje się, że czynnik, który jest opisywany przez zmienną niezależną X_j , działa destymulująco na prawdopodobieństwo wystąpienia badanego zjawiska (przy kontrolowanym wpływie pozostałych zmiennych uwzględnionych w równaniu);
- jeśli $e^{\beta_j} = 1$, to przyjmuje się, że czynnik, który jest opisywany przez zmienną niezależną X_j , nie ma wpływu na prawdopodobieństwo wystąpienia badanego zjawiska (przy kontrolowanym wpływie pozostałych zmiennych uwzględnionych w równaniu).

Na rysunku 3 przedstawione zostały ilorazy szans dla parametrów, których wartości są różne od zera.



Rys. 1. Wartości ilorazu hazardu dla $\beta \neq 0$

Źródło: opracowanie własne.

Z przeprowadzonego badania wynika, iż:

- o wyborze profilu formy i rodzaju wypoczynku decydują w kolejności cechy: miejsce, kraj, cena, zakwaterowanie i transport;
- wpływ stymulujący na prawdopodobieństwo wyboru mają: miejsce wypoczynku nad morzem, kraj wypoczynku w Europie, miejsce wypoczynku nad jeziorem, cena od 1,5 do 3 tys. zł (za 2 tygodnie), cena do 1,5 tys. zł (za 1 tydzień).

Z wyników pierwszej części badania ankietowego dotyczącej wyboru biura podróży wynika, iż:

- 36,3% ankietowanych korzysta rzadziej niż raz w roku z usług biura podróży;
- najbardziej znanym i najczęściej wybieranym biurem podróży jest biuro podróży York (21,1% wskazań);
- według 28,4% ankietowanych reklama i promocja raczej mają wpływ na wybór biura podróży.

Z wyników z drugiej części badania ankietowego dotyczącej wyboru rodzaju i formy wypoczynku wynika, iż:

- dla większości respondentów (38,4%) nie ma znaczenia rodzaj wypoczynku (krajowy bądź zagraniczny);
- najbardziej preferowanym miejscem wypoczynku jest morze (33,2%);
- najczęściej wybieranym środkiem transportu podczas wypoczynku jest transport własny (35%);
- objazdowo-pobytowa forma wypoczynku była wskazywana najczęściej (36,8%);
- większość ankietowanych wybiera czas trwania wypoczynku powyżej 7 dni (39,3%);
- najczęściej wybieraną formą wyżywienia jest zapewnione wyżywienie częściowe (36,3%);
- kwatery prywatne są najbardziej preferowanym rodzajem zakwaterowania (36,4%);
- ankietowani wypoczywają zazwyczaj z rodziną (48,2%);
- na wypoczynek respondenci przeznaczają rocznie przeciętnie od 1,5 do 3 tys. zł (37,8%);
- 51,7% ankietowanych organizuje swój wypoczynek zarówno za pośrednictwem biura podróży, jak i we własnym zakresie;
- 47,2% ankietowanych wskazało, że raz w roku korzysta z wypoczynku.

Z wyników czwartej części badania ankietowego dotyczącej charakterystyki respondenta wynika, iż;

- ankietowani to zazwyczaj kobiety (55%) w wieku 33,5 lat, z zakończoną edukacją i wykształceniem średnim;
- ankietowani to najczęściej pracownicy fizyczni (28,9%) z miesięcznym dochodem na 1 osobę w rodzinie do 2000 zł (52,6%) i źródłem dochodu pracy najemnej u pracodawcy państwowego (36,3%);
- ankietowani to zazwyczaj osoby pozostające w związku małżeńskim (50,3%) mające dzieci (54,3%), najczęściej 1 dziecko (29,4%), i zamieszkujące w mieście do 50 tys. mieszkańców (35,3%).

4. Podsumowanie

Podsumowując, można stwierdzić, że metody wyborów dyskretnych znajdują zastosowanie w analizie pomiaru preferencji wyrażonych. Umożliwiają one symulację rzeczywistych wyborów rynkowych. Na etapie przygotowania danych do pomiaru preferencji można wykorzystać efektywne układy czynnikowe, a na etapie estymacji użyteczności cząstkowych wykorzystuje się warunkowy model logitowy. Zarówno etap przygotowania danych, jak i etap estymacji modelu wyborów dyskretnych wspomagany jest przez oprogramowanie komputerowe.

Literatura

Bąk A. (2002), *Pomiar preferencji metodą conjoint analysis opartą na wyborach*, [w]: *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia 9,

- red. K. Jajuga, M. Walesiak, *Prace Naukowe Akademii Ekonomicznej we Wrocławiu* nr 942, AE, Wrocław, s. 386-399.
- Bąk A. (2004a), *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, *Prace Naukowe Akademii Ekonomicznej we Wrocławiu* nr 1013, Monografie i Opracowania nr 157, AE, Wrocław.
- Bąk A. (2004b), *Metody dyskretnych wyborów w badaniach zachowań konsumentów*, [w]: *Zastosowania metod ilościowych*, red. J. Dziechciarz, *Prace Naukowe Akademii Ekonomicznej we Wrocławiu* nr 1010, AE, Wrocław, s. 150-163.
- Huber J. (1998), *Achieving Individual-Level Predictions from CBC Data: Comparing ICE and Hierarchical Bayes*, Sawtooth Software Research Paper Series, artykuł dostępny w Internecie na stronie: www.sawtoothsoftware.com/download/techpap/indlvcbc.pdf.
- Johnson R.M. (1997), *Individual Utilities from Choice Data: A New Method*, Sawtooth Software Research Paper Series, Proceedings of the Sawtooth Software Conference Artykuł dostępny w Internecie na stronie: www.sawtoothsoftware.com/download/techpap/1977Proceedings.pdf, s. 199-208.
- Orme B.K., Hefft M.A. (1999), *Predicting Actual Sales with CBC: How Capturing Heterogeneity Improves Results*, Sawtooth Software Research Paper Series, artykuł dostępny w Internecie na stronie: www.sawtoothsoftware.com/download/techpap/predict.pdf.
- Orme B.K., Baker G.C. (2000), *Comparing Hierarchical Bayes Draws and Randomized First Choice for Conjoint Simulations*, Sawtooth Software Research Paper Series, Artykuł dostępny w Internecie na stronie: www.sawtoothsoftware.com/download/techpap/rfcdrw.pdf.
- Rószkiewicz M. (2002), *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa.

PREFERENCES ANALYSIS OF TRAVEL AGENCY CUSTOMERS WITH DISCRETE CHOICE METHODS

Summary

In measurement of stated preferences we used the decompositional methods: conjoint analysis and discrete choice methods. In this paper presented an example of discrete choice method application in the study on preferences of travel agency consumers. This research carried out in July and August 2005 on a sample of population of Jelenia Góra. Data were collected especially for this work using questionnaires (394 correctly filled in forms).

In the paper a conditional logit model is used to analyse consumer preferences measured against a nominal scale.