

Spis treści

Wstęp	7
Danuta Strahl: Dwustopniowa klasyfikacja pozycyjna obiektów hierarchicznych ze względu na strukturę obiektów niższego rzędu	9
Andrzej Dudek: Klasyfikacja spektralna a tradycyjne metody analizy skupień	21
Andrzej Dudek, Izabela Michalska-Dudek: Zastosowanie skalowania wielowymiarowego oraz drzew klasyfikacyjnych do identyfikacji czynników warunkujących wykorzystanie Internetu w działalności promocyjnej dolnośląskich obiektów hotelarskich	35
Aneta Rybicka: Oprogramowanie wspomagające segmentację konsumentów z wykorzystaniem metod wyborów dyskretnych	50
Justyna Wilk: Przegląd metod wielowymiarowej analizy statystycznej wykorzystywanych w badaniach segmentacyjnych	59
Anna Błaczkowska, Alicja Grześkowiak: Analiza porównawcza struktury wieku mieszkańców Polski	71
Dariusz Biskup: Analiza zależności w odniesieniu do danych regionalnych ...	84
Dariusz Biskup: Zastosowanie bayesowskich metod wyboru modelu do identyfikacji czynników wpływających na jakość życia	93
Albert Gardoń: Metody testowania hipotez o liczbie składników mieszanki rozkładów	104
Grzegorz Michalski: Financial effectiveness of investments in operating cash	120
Aleksandra Iwanicka: Wpływ zewnętrznych czynników ryzyka na prawdopodobieństwo ruiny w nieskończonym horyzoncie czasowym w wieloklasowym modelu ryzyka	138
Jacek Welc: Próba oceny efektywności strategii inwestycyjnej opartej na regresji liniowej mnożnika P/R spółek notowanych na GPW	152

Summaries

Danuta Strahl: Two-level positional classification of hierarchical objects with regard to the structure of lower level objects	20
Andrzej Dudek: Spectral clustering vs traditional clustering methods	34

Andrzej Dudek, Izabela Michalska-Dudek: Application of multidimensional scaling and classification trees for identifying factors determining internet usage in promotional activity of Lower Silesian hotels	49
Aneta Rybicka: A review of computer software supporting consumer segmentation with an application of discrete choice methods	58
Justyna Wilk: Multivariate data analysis in market segmentation research: a review article	70
Anna Błaczkowska, Alicja Grześkowiak: Comparative analysis of the population age structure in Poland	83
Dariusz Biskup: Areal data dependence analysis	92
Dariusz Biskup: Application of bayesian model choice procedures to identify factors influencing the quality of life	103
Albert Gardoń: Statistical tests for the number of components in mixed distributions	119
Grzegorz Michalski: Efektywność finansowa inwestycji w gotówkę operacyjną	137
Aleksandra Iwanicka: An impact of some outside risk factors on the infinite-time ruin probability for risk model with n classes of business	151
Jacek Welc: The trial of evaluation of the effectiveness of the investment strategy based on the linear regression of the p/r multiple of Warsaw Stock Exchange listed companies	163

Aneta Rybicka

Uniwersytet Ekonomiczny we Wrocławiu – Wydział w Jeleniej Górze

OPROGRAMOWANIE WSPOMAGAJĄCE SEGMENTACJĘ KONSUMENTÓW Z WYKORZYSTANIEM METOD WYBORÓW DYSKRETNYCH

Streszczenie: W metodach wyborów dyskretnych preferencje są mierzone na skalach słabych, a zgromadzone dane nie wystarczają do estymacji modeli na poziomie indywidualnym. Jest to jedna z podstawowych słabości podejścia opartego na wyborach. Wyjściem z tej sytuacji mogą być modele klas ukrytych, które stwarzają możliwość powiększenia zasobu informacyjnego danych i estymacji użyteczności cząstkowych na poziomie segmentowym i indywidualnym.

W artykule przybliżono modele klas ukrytych oraz oprogramowanie wykorzystywane w segmentacji konsumentów metodami wyborów dyskretnych.

Słowa kluczowe: preferencje wyrażone konsumentów, metody wyborów dyskretnych, segmentacja konsumentów, oprogramowanie komputerowe.

1. Wstęp

Metody wyborów dyskretnych, oparte na teorii użyteczności losowej, charakteryzują się bardzo istotnym ograniczeniem – estymacja użyteczności cząstkowych przeprowadzana jest zazwyczaj na poziomie zagregowanym (a nie na poziomie indywidualnym) [Zwerina 1997, s. 75]. Co za tym idzie, modele wyborów dyskretnych zakładają homogeniczną strukturę preferencji. Estymacja użyteczności cząstkowych na poziomie zagregowanym pozwala na oszacowanie udziałów w rynku poszczególnych produktów lub usług (profilów). Jednakże zastosowanie pewnych modeli pozwala na estymację na poziomie segmentowym bądź też indywidualnym.

Niektórzy badacze proponują wykorzystanie modeli klas ukrytych opartych na wyborach w estymacji użyteczności cząstkowych na poziomie segmentowym [Zwerina 1997, s. 75; Huber, Orme, Miller 1999, s. 6]. W badaniach z wykorzystaniem modeli klas ukrytych zakłada się, że w badanej próbie istnieje skończona liczba grup konsumentów o podobnych preferencjach, natomiast między grupami występują istotne różnice (grupy te nie są znane *a priori*, tylko są „ukryte”) [Bąk 2004, s. 134].

Z kolei jednym ze sposobów przeprowadzenia estymacji na poziomie indywidualnym, w metodach wyborów dyskretnych, jest połączenie danych pochodzących z różnych źródeł [Zwerina 1997, s. 75; Huber, Orme, Miller 1999, s. 6]. W badaniach takich wykorzystujemy modele z parametrami losowymi, przyjmując założenie o heterogeniczności konsumentów [Bąk 2004, s. 142]. W metodach wyborów dyskretnych najczęściej wykorzystywany jest hierarchiczny model Bayesa.

Conjoint analysis różni się od metody wyborów dyskretnych (jak również od większości metod wielowymiarowych – *multivariate methods*) tym, że oszacowuje użyteczności cząstkowe na poziomie indywidualnym, tzn. że generowane są oddzielne modele, by określić preferencje dla każdego respondenta [Hair i in. 1995, s. 563]. W większości wielowymiarowych metod dokonuje się pojedynczego pomiaru preferencji (obserwacji) dla każdego respondenta, a następnie przeprowadza się analizę dla wszystkich respondentów jednocześnie. W rzeczywistości wiele metod wymaga, by respondent dostarczał tylko pojedynczych obserwacji (założenie o niezależności), a dopiero później budowany jest wspólny model dla wszystkich respondentów [Hair i in. 1995, s. 563]. W metodach *conjoint analysis* estymacja może być przeprowadzana albo na poziomie indywidualnym (*disaggregate*), albo na poziomie próby, zagregowanym (*aggregate*). Rezultaty na poziomie indywidualnym mogą być agregowane, by przedstawić również całkowity model [Hair i in. 1995, s. 563]. Przy wyborze podejścia (między podejściem zagregowanym a podejściem indywidualnym) badacz powinien równoważyć korzyści uzyskiwane przy wykorzystaniu metod zagregowanych z wnikliwością otrzymywaną poprzez wykorzystanie modeli indywidualnych, takich jak *conjoint analysis*.

W artykule zaprezentowano modele klas ukrytych oraz modele hierarchiczne Bayesa pozwalające oszacować użyteczności cząstkowe na poziomie segmentowym i indywidualnym oraz przedstawiono przegląd oprogramowania, które wspomaga segmentację konsumentów metodą wyborów dyskretnych.

2. Modele klas ukrytych

Główny rozwój modeli klas ukrytych datuje się na drugą połowę XX wieku, natomiast praktyczne zastosowanie tych modeli, na różnych polach badawczych, w ostatnim trzydziestoleciu. Analiza klas ukrytych została wprowadzona w 1950 r. przez Lazarfelda, który użył tej techniki jako narzędzia do zbudowania pewnej typologii segmentów na podstawie zaobserwowanych dychotomicznych zmiennych. Ponad 20 lat później, w 1974 r. Goodman uczynił ją użyteczną w praktyce dzięki rozwinięciu algorytmu pozwalającego otrzymać estymatory metody największej wiarygodności parametrów modelu.

W przypadku badań, których celem jest segmentacja konsumentów, zakłada się niejednorodność zbioru obserwacji, czyli heterogeniczność konsumentów. Możliwe jest wówczas wykorzystanie dodatkowych informacji o preferencjach pocho-

dzących spoza próby. W takim badaniu prawdopodobieństwo wyboru P_{is} szacuje się na podstawie **modelu klas ukrytych** (*latent class models*) [Bąk 2004, s. 110]. Istotną zaletą modeli segmentowych jest to, że pozwalają na uzyskanie informacji o homogenicznych grupach konsumentów, o których początkowo zakłada się, że stanowią zbiorowość o preferencjach heterogenicznych [Huber 1998, s. 1]. Modele te jednocześnie rozdzielają próbę na segmenty danego rynku (różniące się preferencjami respondentów) oraz oszacowują użyteczności przedstawiające preferencje każdego z segmentów [Orme, Heft 1999, s. 4]. Segmentowe użyteczności częściowe szacowane są z wykorzystaniem metody największej wiarygodności. Oszacowane prawdopodobieństwa przynależności respondentów do segmentów pozwalają na estymację indywidualnych użyteczności częściowych jako średnie ważone użyteczności częściowe na poziomie segmentowym (gdzie wagami są prawdopodobieństwa przynależności respondentów do segmentów) [Huber 1998, s. 1; Johnson 1997, s. 199-200].

W badaniach, w których wykorzystywane są modele klas ukrytych, może być uwzględniany zarówno efekt główny, jak i interakcje między atrybutami [Orme, Heft 1999, s. 4]. Analiza klas ukrytych jest bardzo istotna w segmentacji rynku, jednakże mimo że pozwala na redukcję problemu związanego z własnością IIA, to nie umożliwia estymacji użyteczności częściowych na poziomie indywidualnym [Orme 2000, s. 5-6].

Modele klas ukrytych charakteryzują się m.in. tym, że [Bąk 2004, s. 141; Cameron, Trivedi 2005, s. 621-625]:

- zakładają heterogeniczność preferencji konsumentów,
- znajdują zastosowanie w segmentacji konsumentów,
- wykorzystują modele rozkładów mieszanych,
- pozwalają na „powiększenie” zasobu informacyjnego danych empirycznych,
- wymagają rozwiązania problemu „dużej” liczby segmentów,
- istnieje konieczność rozwiązania problemu ustalenia liczby segmentów,
- estymację przeprowadzamy z wykorzystaniem uniwersalnego algorytmu E-M¹ (estymacji modeli metrycznych i niemetrycznych).

W badaniach, w których przeprowadzana jest estymacja preferencji indywidualnych, wykorzystujemy m.in. **hierarchiczny model Bayesa** [Hair i in. 2005, s. 503-505].

Metody bayesowskie reprezentują nieklasyczne metody wnioskowania statystycznego. Nazwa ich wywodzi się od nazwiska Bayesa, który już w XVIII wieku zajmował się metodami wyznaczania prawdopodobieństwa zaistnienia „przyczyny” wtedy, gdy znamy prawdopodobieństwo wystąpienia określonego „skutku”

¹ E-M – algorytm maksymalizacji wartości oczekiwanej (*expectation-maximization*). Algorytm ten jest często wykorzystywany w programach komputerowych przeznaczonych do estymacji mieszanek rozkładów wielowymiarowych [Bąk 2004, s. 139; Wedel, Kamakura 1998, s. 81].

działania tej „przyczyny”² [Domański, Pruska, Wagner 1998, s. 149]. Natomiast rozwój metod bayesowskich³ rozpoczął się w połowie XX wieku wraz z rozwojem statystycznej teorii decyzji, której podstawowe pojęcia (gry statystycznej i statystycznej funkcji decyzyjnej) zdefiniował Wald [Domański, Pruska, Wagner 1998, s. 149].

Model Bayesa jest bardzo efektywną techniką „pożyczania danych”, która pozwala na estymację indywidualnych użyteczności cząstkowych, wykorzystując informacje pochodzące nie tylko od danego respondenta, lecz również innych respondentów z badanej grupy [Orme, Baker 2000, s. 1; Johnson 1997, s. 207-208; McCullough 2003, s. 2].

Model Bayesa nazywany jest „hierarchicznym”, ponieważ składa się z dwóch poziomów. Na wyższym poziomie (*higher, upper level*) przyjęto założenie, że parametry indywidualne (indywidualne użyteczności cząstkowe) określone są wielowymiarowym rozkładem normalnym [Orme 2000, s. 1; Johnson 2000, s. 8]. Taki rozkład scharakteryzowany jest wektorem średnich użyteczności cząstkowych oraz macierzą wariancji-kowariancji rozkładu użyteczności cząstkowych wśród respondentów. Na niższym poziomie (*lower level*) przyjęte jest założenie, że indywidualne prawdopodobieństwa wyboru określonych profilów opisane są za pomocą wielomianowego modelu logitowego (bądź też za pomocą regresji liniowej) [Orme 2000, s. 1].

W estymacji modeli hierarchicznych Bayesa istotną rolę odgrywają metody *Monte Carlo Markov Chain* (MCMC), szczególnie metoda Gibbsa [Johnson 2000, s. 5; Allenby, Arora, Ginter 1995, s. 153].

Modele hierarchiczne Bayesa (modele z parametrami losowymi⁴) charakteryzują się m.in. tym, że [Bąk 2004, s. 145; Hauser, Rao 2002, s. 14]:

- zakładają heterogeniczność konsumentów,
- pozwalają na przeprowadzenie estymacji indywidualnych użyteczności cząstkowych na podstawie wyborów,
- umożliwiają estymację dużej liczby parametrów (większej niż liczba obserwacji),
- wykorzystują względnie dużą złożoność obliczeniową algorytmów estymacji parametrów, informację *a priori*,
- brakuje popularnego oprogramowania komputerowego.

Inną metodą estymacji użyteczności cząstkowych na poziomie indywidualnym w metodach wyborów dyskretnych jest **estymacja wyborów indywidualnych**

² Teoria Bayesa była opublikowana w 1764 r.: *An Essay toward Solving a Problem in the Doctrine of Chances* przez the Royal Society of Londyn, England (Allenby, Rossi 2003, s. 1).

³ Szerzej twierdzenie Bayesa przedstawiają prace [Bobrowski 1986; Domański, Pruska 2000; Cameron, Trivedi 2005].

⁴ W modelu tym przyjmuje się, że szacowane parametry nie są nieznanymi wielkościami stałymi, lecz zmiennymi losowymi o określonych rozkładach [Bąk 2004, s. 142].

(*Individual Choice Estimation* – ICE) [Orme, Heft 1999, s. 4; Chrzan 1999, s. 237]. Modele te są „rozszerzoną wersją” modeli klas ukrytych. Pozwalają one na oszacowanie użyteczności odzwierciedlających preferencje indywidualne (poszczególnych respondentów) [Orme, Heft 1999, s. 4].

ICE to podejście pragmatyczne. Pozwala na oszacowanie indywidualnych użyteczności cząstkowych, które najbardziej odpowiadają indywidualnym wyborom. Estymację taką można wykonać bardzo szybko, wykorzystując na wstępie wyniki otrzymane po zastosowaniu modeli klas ukrytych (badanie z wykorzystaniem modeli ICE można również przeprowadzić bez wykorzystania w tymże badaniu wyników z modeli klas ukrytych) [Huber 1998, s. 2]. Jednak główną wadą tego podejścia jest to (również jak w badaniach z wykorzystaniem modeli klas ukrytych), że otrzymane wyniki zależą od wyboru segmentów oraz od liczby wykorzystanych segmentów w badaniu.

Dostępne oprogramowanie komputerowe pozwala raczej na przeprowadzenie badania z wykorzystaniem tylko modeli uwzględniających efekty główne pomiędzy atrybutami [Orme, Heft 1999, s. 4].

Zestawienie cech modeli klas ukrytych oraz modeli hierarchicznych Bayesa (modeli z parametrami losowymi) przedstawia tab. 1.

Tabela 1. Cechy modeli klas ukrytych i modeli hierarchicznych Bayesa

Cechy	Modele klas ukrytych	Modele hierarchiczne Bayesa
Nominalna skala pomiaru preferencji	+	+
Estymacja na poziomie indywidualnym	+ –	+
Estymacja na poziomie segmentowym	+	–
Uwzględnienie niejednorodności preferencji	+	+
Parametryczna reprezentacja niejednorodności preferencji	–	+
Możliwość stosowania symulatorów wyboru	+	+
Redukcja problemu IIA	+	+
Wykorzystanie informacji spoza próby	– +	+
Wykorzystanie metod symulacyjnych w estymacji parametrów	+	+
Podejście bayesowskie	– +	+
Oprogramowanie komputerowe	+	– +

Źródło: [Bąk 2005, s. 188].

Badania przeprowadzone przez Hubera [1998] pozwalają wnioskować, że zarówno model hierarchiczny Bayesa, jak i estymacja wyborów indywidualnych dają podobne rezultaty w praktyce⁵. Natomiast modele klas ukrytych słabo „przedstawiają” indywidualne wybory [Huber 1998, s. 3; *CBC Latent...* 2000-2001, s. 2].

⁵ Podobne wnioski prezentują w swojej pracy Andrews, Ainslie i Currim [2002].

3. Oprogramowanie komputerowe

Użytecznym narzędziem analitycznym jest moduł pozwalający na przeprowadzenie segmentacji z wykorzystaniem modeli klas ukrytych: **The Latent Class Segmentation Module**⁶ [*The CBC...* 2004, s. 1]. Moduł ten wykorzystujemy wraz z danymi otrzymanymi np. z badania CBC lub też CBC/Web w celu przypisania respondentów o podobnych preferencjach (uzyskanych z wyborów, jakich dokonali np. w ankiecie CBC) do poszczególnych segmentów. Pozwala on na wykorzystanie analizy klas ukrytych w celu estymacji użyteczności cząstkowych każdego z segmentów oraz prawdopodobieństwa przynależności każdego respondenta do tych segmentów.

Moduł ten spełnia funkcję podobną do modułu Logit w CBC, jednak nie szuka przeciętnej użyteczności cząstkowych dla wszystkich respondentów razem, ale rozpatruje podgrupy respondentów różniące się od siebie i oszacowuje użyteczności cząstkowe dla każdego segmentu. Podgrupy respondentów charakteryzują się tym, że respondenci wewnątrz grup mają podobne preferencje, natomiast preferencje respondentów w poszczególnych podgrupach różnią się.

Modele klas ukrytych przydzielają użyteczności cząstkowe dla każdego z segmentów. Analiza z wykorzystaniem modeli klas ukrytych nie zakłada, że każdy z respondentów całkowicie przynależy do jednej czy też drugiej grupy. Raczej rozpatrywane jest niezerowe prawdopodobieństwo przynależności każdego z respondentów do każdej z grup. Jeśli rozwiązanie dopasowane jest do danych bardzo dobrze, to te prawdopodobieństwa zbliżają się do 0 bądź też do 1.

Sawtooth Software oferuje również inny moduł pozwalający na estymację indywidualnych wyborów: **Individual Choice Estimation (ICE) Module for choice-based Conjoint**.

ICE było skonstruowane jako rozbudowa analizy modeli klas ukrytych, by pozwolić na lepszą prognozę przez bardziej efektywniejsze rozpoznanie heterogeniczności preferencji [*Individual Choice...* 2001, s. 2].

Moduł ten charakteryzuje się kilkoma zaletami. Jedną z nich jest jego „szybkość”. Jeśli modele klas ukrytych są wybrane jako „punkt startowy”, wysiłek potrzebny do obliczenia indywidualnych użyteczności jest minimalny. Nawet jeśli badacz zdecyduje, że nie wykorzysta modeli klas ukrytych jako „punktu startowego”, ICE może oszacować użyteczności, zaczynając od początku, w rozsądnym czasie. Dla danych otrzymanych od 300 respondentów oraz z 25 poziomami atrybutów, rozwiązanie, które mogłoby być zaakceptowane, można uzyskać w kilka minut. ICE jest modułem szybszym niż pozostałe dwa rozwiązania stosowane w oszaco-

⁶ Modele klas ukrytych stały się popularne w połowie lat 90. XX wieku. Jedną z zalet tychże modeli jest to, że pozwalają na redukcję negatywnego efektu własności modelu logitowego, tzn. *IIA* [*The CBC Latent...*, s. 1].

waniach heterogeniczności respondentów: jest zdecydowanie szybszy niż analiza klas ukrytych oraz jest nieporównywalnie szybszy niż analiza hierarchiczna Bayesa.

Inną zaletą modułu ICE jest to, że jest on lepszy w „uchwyceniu” heterogeniczności aniżeli modele klas ukrytych. Doświadczenia wskazują, że również metody hierarchiczne Bayesa są lepsze w oszacowywaniu heterogeniczności aniżeli modeli klas ukrytych.

Trzecią zaletą tego modułu jest to, że pozwala badaczom przechodzić z zagregowanej analizy danych pochodzących z wyborów do analizy danych na poziomie indywidualnym. Jednym z problemów badaczy rynkowych jest zaś umiejętność przewidywania reakcji rynku na złożone kombinacje iteracji, zróżnicowane efekty krzyżowe oraz zmieniające się podobieństwa między produktami. Wydaje się, że wszystkie te problemy mogą być zredukowane z wykorzystaniem modeli na poziomie indywidualnym [*Individual Choice...* 2001, s. 23].

Również bardzo popularnymi oprogramowaniami wykorzystywanymi w badaniach z wykorzystaniem modeli klas ukrytych są LatentGOLD oraz GLIMMIX. Porównując oba oprogramowania, możemy stwierdzić, że [Deal 2000, s. 38-40]:

- LatentGOLD pozwala na przeprowadzenie analizy regresji czynnikowej, klasyfikacji. W GLIMMIX można wybrać jedynie pomiędzy regresją i klasyfikacją.
- Program GLIMMIX wymaga zdefiniowania rozkładu zmiennej zależnej, w LatentGOLD rozkład jest oceniany na podstawie realizacji zmiennej zależnej.
- Oba programy pozwalają badaczowi wybierać różne funkcje analizy oraz konfigurację danych wyjściowych.
- Informacje o klasyfikacji są dostępne w obydwu programach. LatentGOLD zapisuje automatycznie klasyfikacje w pliku w standardach SPSS, GLIMMIX wyświetla klasyfikację w odrębnym oknie i zapisuje je w pliku.
- Zarówno GLIMMIX, jak i LatentGOLD dostarczają wyników klasyfikacji. W LatentGOLD informacje te są dodatkowo przedstawiane na wykresach, co może pomóc w podejmowaniu decyzji.
- W oprogramowaniu LatentGOLD możliwe jest importowanie danych z SPSS, rozpoznawane są nazwy zmiennych, można również importować dane z plików tekstowych czy tablicowych. Natomiast w GLIMMIX dane muszą być specjalnie przygotowane lub mieć formę drzewiastą, spacje nie są akceptowane jako braki danych.

Inne programy możliwe do wykorzystania przy estymacji modeli ze zmiennymi ukrytymi to: PANMARK, LEMWin, MLLSA, GLIMMIX, WINLTA, GAUSS, MATLAB, MAPLE V, SPSS, SAS.

Do praktycznego stosowania modeli klas ukrytych w badaniach preferencji niezbędne jest odpowiednie oprogramowanie komputerowe. Procedury szacujące modele mieszanek rozkładów są dostępne zarówno w uniwersalnych pakietach statystycznych (np. SPSS, SAS/STAT, R), jak i w programach przeznaczonych wyłącz-

nie do estymacji modeli klas ukrytych (np. GLIMMIX, LatenGOLD – współautorstwa Jeroena Vermunta z Uniwersytetu w Tilburgu).

Procedury metod wyborów dyskretnych są realizowane w podobny sposób w ramach różnych programów statystycznych. Mogą się różnić np. prezentacją respondentom profilów w badaniu, trybem, w jakim się pracuje (niektóre programy pozwalają na pracę w trybie okienek dialogowych, inne zaś wymagają programowania). Niektóre z nich pozwalają tylko na oszacowanie preferencji na poziomie zagregowanym (np. CBC, CBC Advanced Design Module), inne tylko na poziomie segmentowym (The Latent Class Segmentation Module, GLIMMIX Latent-GOLD, a jeszcze inne na poziomie indywidualnym (CBC/HB, Individual Choice Estimation Module). Wybór konkretnego programu zależy od celu badań. Należy też podkreślić, że większość tych pakietów jest dostępna na zasadach komercyjnych.

Literatura

- Allenby G.M., Arora N., Ginter J.L., *Incorporating prior knowledge into the analysis of conjoint studies*, „Journal of Marketing Research” 1995, vol. XXXII, May, s. 152-162.
- Allenby G.M., Rossi P.E., *Perspectives Based on 10 Years of HB in Marketing Research*, www.sawtoothsoftware.com/download/techpap/allenby.pdf, 2003.
- Andrews R.L., Ainslie A., Currim I.S., *An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity*, „Journal of Marketing Research” 2002 nr 34 (November), s. 479-487.
- Bąk A., *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1013, Seria: Monografie i Opracowania nr 157, AE, Wrocław 2004.
- Bąk A., *Problemy estymacji parametrów w modelach dekompozycyjnych z dyskretną zmienną objaśnianą*, [w:] *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia 12, Sekcja Klasyfikacji i Analizy Danych PTS, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1076, AE, Wrocław 2005, s. 180-189.
- Bobrowski D., *Probabilistyka w zastosowaniach technicznych*, Wydawnictwo Naukowo-Techniczne, Warszawa 1986.
- Cameron A.C., Trivedi P.K., *Microeconometrics. Methods and applications*, Cambridge University Press, New York 2005.
- CBC Latent Class Analysis Technical Paper*, www.sawtoothsoftware.com/download/techpap/lctech.pdf, 2000-2001
- Chrzan K., *Full versus Partial Profile Choice Experiments: Aggregate and Disaggregate Comparison*, Sawtooth Software Conference Proceedings, Sequim, February 1999.
- Deal K., *Latent class analysis: Latent GOLD 2.0 vs. GLIMMIX 2.0*, „Marketing Research” 2000, Winter, vol. 12, Issue 4, s. 38-40.
- Domański C., Pruska K., *Nieklasyczne metody statystyczne*, PWE, Warszawa 2000.
- Domański C., Pruska K., Wagner W., *Wnioskowanie statystyczne przy nieklasycznych założeniach*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 1998.

- Hair J.F., Anderson R.E., Tatham R.L., Blach W.C., *Multivariate Data Analysis with Readings*, Englewood Cliffs, Prentice-Hall 1995.
- Hair J.F., Black W.C., Babin B.J., Anderson R.E., Tatham R.L., *Multivariate Data Analysis*, Pearson Prentice-Hall, Upper Saddle River, New Jersey 2005.
- Hauser J.R., Rao V.R., *Conjoint Analysis, Related Modeling, and Application*, http://web.mit.edu/hauser/www/Papers/GreenTributeConjoint_092302.pdf, 2002.
- Huber J., *Achieving Individual-Level Predictions from CBC Data: Comparing ICE and Hierarchical Bayes*, Sawtooth Software Research Paper Series, www.sawtoothsoftware.com/download/techpap/indlvcbc.pdf, 1998.
- Huber J., Orme B.K., Miller R., *Dealing with Product Similarity in Conjoint Simulations*, www.sawtoothsoftware.com/download/techpap/prodsim.pdf, 1999.
- Individual Choice Estimation (ICE) Module for Choice-based Conjoint*, www.sawtoothsoftware.com/download/techpap/icetech.pdf, 2001.
- Johnson R.M., *Individual Utilities from Choice Data: A New Method*, Sawtooth Software Research Paper Series, Proceedings of the Sawtooth Software Conference, www.sawtoothsoftware.com/download/techpap/1997Proceedings.pdf, 1997, s. 199-208.
- Johnson R.M., *Understanding HB: An Intuitive Approach*, www.sawtoothsoftware.com/download/techpap/undhb.pdf, 2000.
- McCullough D., *Abbreviated Task Sets: Estimating Disaggregate Choice Models with Extremely Few Tasks per Respondent*, www.macroinc.com/html/art/abbreviated.html, 2003.
- Orme B.K., *Hierarchical Bayes: Why All the Attention?* Sawtooth Software Research Paper Series, www.sawtoothsoftware.com/download/techpap/hbwhy.pdf, 2000.
- Orme B.K., Baker G.C., *Comparing Hierarchical Bayes Draws and Randomized First Choice for Conjoint Simulations*, Sawtooth Software Research Paper Series, www.sawtoothsoftware.com/download/techpap/rfcdrw.pdf, 2000.
- Orme B.K., Heft M.A., *Predicting Actual Sales with CBC: How Capturing Heterogeneity Improves Results*, Sawtooth Software Research Paper Series, www.sawtoothsoftware.com/download/techpap/predict.pdf, 1999.
- The CBC Latent Class Technical Paper (Version 3)*, www.sawtoothsoftware.com/download/techpap/lctech.pdf, 2004.
- Wedel M., Kamakura W.A., *Market Segmentation. Conceptual and Methodological Foundation*, Kluwer Academic Publishers, Boston-Dordrecht-London 1998.
- Zwerina K., *Discrete Choice Experiments in Marketing*, Physica-Verlag, Heidelberg-New York 1997.

A REVIEW OF COMPUTER SOFTWARE SUPPORTING CONSUMER SEGMENTATION WITH AN APPLICATION OF DISCRETE CHOICE METHODS

Summary: In discrete choice methods preferences are measured on weak scales of measurement and all collected data is not enough to estimate individual level models. It is one of the most common weaknesses of choice methods approach. One of solutions for this situation can be latent class models which allow to expand information volume from data by segment and individual part-worth utilities estimation.

The article presents latent class models and computer software that was applied in client segmentation with an application of discrete choice methods.