PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1126 2006
TAKSONOMIA 13
Klasyfikacja i analiza danych – teoria i zastosowania

Robert Kapłon
Politechnika Wrocławska

ZINTEGROWANE A DWUETAPOWE PODEJŚCIE DO SEGMENTACJI W ANALIZIE *CONJOINT*

1. Wstęp

Segmentacja rynku, a więc podział na grupy, które pod pewnym względem są do siebie podobne, zajmuje szczególne miejsce w marketingu. Powodem tego są różne potrzeby i wymagania konsumentów, których zaspokojenie „uniwersalnym" produktem jest niemożliwe. Nakłada to konieczność uwzględnienia heterogeniczności konsumentów w procesie kreowania oferty rynkowej. Podstawowe zatem pytanie w tym obszarze dotyczy sposobu podziału konsumentów na homogeniczne grupy.

Istnieje kilka koncepcji i kryteriów takiego podziału. Jednym z nich mogą być preferencje. Nakłada to jednak konieczność wcześniejszego ich zbadania. W tym względzie analiza *conjoint* jest użyteczną i często wykorzystywaną metodą. Przyjmując za podstawę segmentacji użyteczności cząstkowe, które zostały wcześniej oszacowane zgodnie z procedurą analizy *conjoint*, dokonuje się podziału badanej zbiorowości, np. metodą *k*-średnich [Walesiak, Bąk 2000]. Jak można zauważyć, proces szacowania użyteczności i proces segmentacji są rozdzielone. W literaturze proponuje się również inne podejście, tzw. zintegrowane, które łączy w sobie procesy estymacji użyteczności i segmentacji. Wyniki badań symulacyjnych sugerują, że podejście to zdaje się być pod pewnymi względami lepsze od podejścia dwuetapowego [Vriens, Wedel, Wilms 1996].

W pracy porównano podejście dwuetapowe z podejściem zintegrowanym opartym na mieszankach rozkładów. Oprócz teoretycznych rozważań wykorzystano wyniki badań ankietowych (dotyczących preferencji na rynku telefonii komórkowej), które stały się również podstawą do wniosków.

2. Model analizy *conjoint* ze skalarną macierzą kowariancji – ujęcie dwuetapowe

W pierwszym etapie metodą największej wiarygodności zostaną oszacowane parametry na poziomie indywidualnym. W tym celu przyjęto następującą notację:

$\mathbf{X} = [x_{ja}]_{J \times A}$ – macierz eksperymentu,

$\mathbf{Y} = [y_{ji}]_{J \times N}$ – macierz preferencji (oceny, jaką konsument i dał produktowi j),

$\boldsymbol{\beta}_i = [\beta_{i1}, \ldots, \beta_{iA}]^T$ – wektor parametrów (istotności atrybutu a (a=1,..., A)),

$\sigma^2\mathbf{I}$ – macierz kowariancji wymiaru J,

$\mathbf{u}_i = [u_1, \ldots, u_J]^T$ – wektor składników losowych o rozkładzie $\mathcal{N}_J(\mathbf{0}, \sigma^2\mathbf{I})$.

Niech użyteczność, jaką dla konsumenta i mają produkty, będzie wyrażona w następujący sposób

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \mathbf{u}_i,$$

wtedy funkcja wiarygodności przybiera postać:

$$L(\boldsymbol{\beta}_i, \sigma_i^2 \mid \mathbf{y}_i) \propto \frac{1}{2}\log|\sigma_i^{-2}\mathbf{I}| - \frac{1}{2}\sigma_i^{-2}(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta}_i)^T(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta}_i).$$

Wykorzystując ją, otrzymuje się następujące estymatory dla konsumenta i:

$$\hat{\boldsymbol{\beta}}_i = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}_i, \tag{1}$$

$$\hat{\sigma}_i^2 = \frac{1}{J}(\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i)^T(\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i). \tag{2}$$

Drugi etap to procedura segmentacji, której podstawą są wyznaczone estymatory (użyteczności cząstkowe). Wykorzystując metodę k-średnich, przy założeniu G klas, ustala się przynależność każdego konsument do jednej z C_g klas (g = 1, ..., G). Można pokazać, że wyznaczone dla segmentu g estymatory będą przybierały postać:

$$\hat{\boldsymbol{\beta}}_g = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\sum_{i \in C_g}\mathbf{y}_i / n_g, \tag{3}$$

$$\hat{\sigma}_g^2 = \frac{1}{n_g J}\sum_{i \in C_g}(\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_g)^T(\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_g), \tag{4}$$

gdzie n_g jest liczbą konsumentów przydzielonych do segmentu g. Powyższe oszacowanie użyteczności cząstkowych pokazuje, że estymator ten jest równy średniej arytmetycznej z wartości estymatorów na poziomie indywidualnym, tzn. $\hat{\boldsymbol{\beta}}_g = \sum_{i \in C_g}\hat{\boldsymbol{\beta}}_i / n_g$.

3. Model analizy *conjoint* oparty na mieszaninie rozkładów

U podstaw podejścia zintegrowanego leży założenie, zgodnie z którym konsumenci pochodzą z populacji składającej się z G podpopulacji. Przynależność konsumentów do konkretnego segmentu g ($g = 1, \ldots, G$) nie jest znana *a priori*, jednak określa się ją z prawdopodobieństwem π_g (oczywiście $\sum_g \pi_g = 1$).

Użyteczność, jaką dla konsumenta i mają produkty (ale pod warunkiem, że konsument przynależy do segmentu C_g), przybiera postać[1]:

$$\mathbf{y}_{i|C_g} = \mathbf{X}\boldsymbol{\beta}_g + \mathbf{u}_{i|C_g}.$$

Podobnie jak w wypadku ujęcia dwuetapowego należy założyć rozkład składnika losowego. Tutaj będzie to warunkowy rozkład normalny:

$$\left(\mathbf{u}_i \,|\, i \in C_g\right) \sim \mathcal{N}_J\left(\mathbf{0}, \sigma_g^2 \mathbf{I}\right).$$

Dla tak sformułowanego modelu można zapisać funkcję wiarygodności i w ten sposób oszacować parametry. Aczkolwiek trudności natury numerycznej przemawiają na niekorzyść takiego podejścia. Dlatego zostanie wykorzystany algorytm **EM** [Dempster, Laird, Rubin 1977; McLachlan, Krishnan 1997]. Algorytm ten jest iteracyjną procedurą; w iteracji (k+1) maksymalizuje się funkcję zdefiniowaną następująco:

$$Q(\boldsymbol{\Psi} \,|\, \boldsymbol{\Psi}^{(k)}) = \mathbb{E}_{\boldsymbol{\Psi}^{(k)}}\left[\log L_c\left(\boldsymbol{\Psi} \,|\, \mathbf{Y}, \mathbf{Z}; \mathbf{X}\right) \middle| \mathbf{Y}, \boldsymbol{\Psi}^{(k)}\right],$$

gdzie: $\boldsymbol{\Psi}$ – wektor parametrów, $\mathbf{Z}$ – macierz dodatkowo wprowadzonych zmiennych ukrytych pozwalająca zdefiniować logarytm funkcji wiarygodności dla kompletnego zbioru danych.

Dla iteracji (k+1) postać funkcji Q jest następująca:

$$Q\left(\boldsymbol{\Psi} \,|\, \boldsymbol{\Psi}^{(k)}\right) \propto \sum_{i=1}^{N}\sum_{g=1}^{G} \tau_{ig}^{(k)} \log \pi_g + \frac{1}{2}\sum_{i=1}^{N}\sum_{g=1}^{G} \tau_{ig}^{(k)} \log \left|\sigma_g^{-2}\mathbf{I}\right| -$$

$$-\frac{1}{2}\sum_{i=1}^{N}\sum_{g=1}^{G} \tau_{ig}^{(k)} \sigma_g^{-2}\left(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta}_g\right)^T\left(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta}_g\right),$$

w której τ_{ig} jest prawdopodobieństwem *a posteriori* przynależności konsumenta i do klasy C_g.

Maksymalizując tę funkcję ze względu na parametr π_g, wprowadzono mnożnik Lagrange'a. Po przekształceniach otrzymano w iteracji (k+1) estymator

$$\hat{\pi}_g^{(k+1)} = \frac{1}{N}\sum_{i=1}^{N} \tau_{ig}^{(k)}, \tag{5}$$

[1] Taka reprezentacja użyteczności wskazuje, że estymowany jest wektor użyteczności cząstkowych na poziomie segmentu.

gdzie:
$$\hat{\tau}_{ig}^{(k)} = \frac{\pi_g^{(k)} f_g(\mathbf{y}_i \mid \boldsymbol{\beta}_g^{(k)}, \sigma_g^{(k)2})}{\sum_{g=1}^{G} \pi_g^{(k)} f_g(\mathbf{y}_i \mid \boldsymbol{\beta}_g^{(k)}, \sigma_g^{(k)2})}$$

$$f_g(\mathbf{y}_i \mid \boldsymbol{\beta}_g^{(k)}, \sigma_g^{(k)2}) = (2\pi)^{-J/2} \mid \sigma_g^{(k)2}\mathbf{I} \mid^{-1/2} \exp\left[-\frac{1}{2\sigma_g^{(k)2}}\left(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta}_g^{(k)}\right)^T\left(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta}_g^{(k)}\right)\right].$$

Odpowiednio estymatorem użyteczności cząstkowych jest

$$\hat{\boldsymbol{\beta}}_g^{(k+1)} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T \frac{\sum_{i=1}^{N} \hat{\tau}_{ig}^{(k)} \mathbf{y}_i}{\sum_{i=1}^{N} \hat{\tau}_{ig}^{(k)}}, \tag{6}$$

wariancji zaś:

$$\hat{\sigma}_g^{(k+1)2} = \frac{1}{J} \frac{\sum_{i=1}^{N} \hat{\tau}_{ig}^{(k)} \left(\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_g^{(k+1)}\right)^T \left(\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_g^{(k+1)}\right)}{\sum_{i=1}^{N} \hat{\tau}_{ig}^{(k)}}. \tag{7}$$

Całą procedurę powtarza się, aż do osiągnięcia zbieżności. Za kryterium stopu można przyjąć dowolnie małą, bezwzględną różnicę między wartościami estymatorów w dwóch następujących po sobie iteracjach.

Analiza porównawcza estymatorów otrzymanych w ujęciach dwuetapowym i zintegrowanym skłania do kilku wniosków. We wzorach (3),(4),(6) i (7) występują średnie ważone. Przy czym w ujęciu zintegrowanym wagi są równe prawdopodobieństwu *a posteriori*, koncepcja dwuetapowa zaś zakłada wagi 0 lub 1 (konsument nie należy do segmentu bądź do niego należy). Wydaje się, że w wypadku trudności zaklasyfikowania konsumentów do konkretnego segmentu lepiej zrezygnować z wag binarnych na rzecz tych drugich. Wtedy wpływ konsumenta na szacowane użyteczności jest proporcjonalny do prawdopodobieństwa przynależności do klasy – podobnie jest z wariancją.

W obu ujęciach należy przyjąć *explicite* liczbę segmentów. Jednak w odróżnieniu od mieszanek rozkładów procedura klasyfikacji oparta na *k*-średnich nie ma kryteriów pozwalających na wybór liczby segmentów.

W przedstawionych modelach założono skalarną macierz kowariancji w każdym segmencie. Gdyby złagodzić to założenie, co z praktycznego punktu widzenia może być uzasadnione, i przyjąć macierz diagonalną (z różnymi wariancjami na przekątnej), wtedy problem z liczbą stopni swobody dyskwalifikuje ujęcie dwuetapowe. Wolna od tego ograniczenia jest koncepcja oparta na mieszankach rozkładów.

Intrygujące wydaje się pytanie o możliwość doboru najlepszej kombinacji wag. M.R. Hagerty [1985] rozważał to zagadnienie dla wag binarnych. Pokazał, że jeśli wagi mają być dobrane w taki sposób, by minimalizować przeciętny średniokwadratowy błąd prognozy, to przy klasyfikacji należy wykorzystać analizę czynnikową. Niestety takiego wniosku nie można wysunąć w wypadku wag z przedziału (0,1).

4. Porównanie metod na podstawie danych empirycznych

Przeprowadzono badanie preferencji na rynku telekomunikacyjnym[2], w którym wzięto pod uwagę następujące atrybuty (w nawiasach podano ich poziomy):

X_1 – sekundowe naliczanie (tak, nie),
X_2 – możliwość doładowania w inny sposób (tak, nie),
X_3 – możliwości wykonywania połączeń po doładowaniu (1 miesiąc, 2 miesiące),
X_4 – poczta głosowa (bezpłatna, płatna),
X_5 – dodatkowe opcje (tak, nie),
X_6 – minimalna kwota doładowania (20 zł, 30 zł, 50 zł).

Każdy z respondentów oceniał w skali od 1 do 100 hipotetycznych 16 wyrobów. Dane te stały się podstawą do estymacji parametrów rozważanych modeli. Aby ustalić liczbę klas w modelu opartym na mieszankach rozkładów, obliczono odpowiednie kryteria informacyjne.

Tabela 1. Mierniki jakości dopasowania modelu

	Liczba segmentów				
	II	III	IV	V	VI
Liczba param.	19	29	39	49	59
LogL	-7379,13	-7348,55	-7331,23	-7312,07	-7297,96
AIC	14796,27	14755,11	14740,46	14722,13	14713,91
BIC	14846,33	14831,51	14843,22	14851,23	14869,36
CAIC	14865,33	14860,51	14882,22	14900,23	14928,36

Źródło: opracowanie własne.

Z tabeli 1 wynika, że wyboru należy dokonać między trzema a sześcioma klasami. Jednak dwa powody przemawiają na korzyść modelu z większą liczbą klas. Pierwszy dotyczy preferowania przez kryterium AIC modeli bardziej rozbudowanych. Wolne od tego mankamentu są dwa pozostałe kryteria. Drugi powód związany jest z przyjętą strategią porównawczą, zgodnie z którą porównane modele powinny mieć taką samą liczbę klas. Przyjęcie sześciu klas w metodzie *k*-średnich prowadzi do wyodrębnienia segmentu, w którym jest tylko jeden respondent, co z punktu widzenia segmentacji nie ma uzasadnienia.

Przyjmując trzy klasy, oszacowano wartości parametrów i zamieszczono je w tab. 2 i 3.

Tabela 2. Oceny parametrów modelu z 3 klasami opartego na podejściu dwuetapowym

	Estymatory									
Klasa	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\sigma}^2$	n_s / N
I	44,03	5,91	-2,50	0,30	0,78	2,77	4,37	-6,30	493,56	0,204
II	43,03	2,87	0,49	5,73	2,65	9,36	-3,05	-6.37	308,30	0,281
III	51,30	4,07	3,53	7,66	11,95	5,02	-9,13	4,52	417,40	0,515

Źródło: opracowanie własne.

[2] Autorowi niniejszego opracowania dane zostały udostępnione przez Panią Helenę Hetmańską i Panią Ewę Suligę.

Tabela 3. Oceny parametrów modelu z 3 klasami opartego na podejściu zintegrowanym

Klasa	Estymatory									
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\sigma}^2$	$\hat{\pi}$
I	43,59	4,39	-1,02	2,91	2,61	4,98	1,75	-5,74	478,65	0,719
II	51,51	5,74	4,29	8,40	13,58	1,54	-12,15	11,41	372,03	0,136
III	55,48	3,77	3,95	7,43	10,62	10,02	-13,53	1,05	171,37	0,145

Źródło: opracowanie własne.

Analizę porównawczą otrzymanych wyników warto rozpocząć od stwierdzenia, że prawdopodobieństwa *a posteriori* $\hat{\tau}_{ig}$ dla zdecydowanej większości respondentów kształtują się na poziomie 0,9 i więcej w jednej z klas[3]. To by sugerowało, że rozmiary segmentów powinny być zbliżone. Jednak tak nie jest, bowiem różnica między najliczniejszymi segmentami wynosi ok. 0,2. Dla pozostałych segmentów, w przybliżeniu, wynosi 0,15 i 0,05. Daje to podstawę do przypuszczeń, że mechanizm klasyfikacji w obu podejściach jest inny. Należy również stwierdzić, że przy problemach zaklasyfikowania respondenta do jednego segmentu (gdy prawdopodobieństwa *a posteriori* byłyby na podobnym poziomie, np. 0,3 0,36 0,34) różnice byłyby znacznie większe. Czy wobec tego można zarekomendować któreś z tych podejść i powiedzieć, że jest lepsze?

Próba odpowiedzi na to pytanie prowadzi do wyboru kryterium rozstrzygającego. Przyjęto, że jakość odzwierciedlenia struktury zaobserwowanych preferencji przez rozważane modele będzie rozstrzygać o wyborze modelu. Miernikiem tej jakości będą współczynniki korelacji. Ponieważ o preferencjach mówi się w kategoriach porządkowych, dlatego obok współczynnika Pearsona obliczono również współczynnik korelacji rang Spearmana i Kendalla.

Warto zwrócić uwagę na sposób szacowania preferencji w modelu zintegrowanym. Pierwszym krokiem do tych szacunków jest przypisanie respondenta do jednej z klas. Wydaje się, że w takim ujęciu, wynikającym z istoty przyjętego modelu, nie uwzględnia się pełnej informacji o przynależności do klas (prawdopodobieństwa *a posteriori*). Na przykład, jeśli prawdopodobieństwa przynależności do klas określonego respondenta wynoszą: 0,35, 0,34, 0,31, to obliczenie preferencji w jednym segmencie (gdzie prawdopodobieństwo jest najwyższe) budzi wątpliwości. Być może lepszym rozwiązaniem będzie obliczenie preferencji jako wypadkowej – suma ważona preferencji w każdym segmencie, gdzie wagami są prawdopodobieństwa *a posteriori*. Obliczone w ten sposób preferencje włączono również do analizy; ujęciu temu nadano nazwę „zintegrowane2".

[3] Wartości tych prawdopodobieństw nie zamieszczono w artykule ze względu na duży rozmiar macierzy (103x3).

Tabela 4. Współczynniki korelacji między preferencjami

Ujęcie	Współcz.	Liczba segmentów				
		II	III	IV	V	VI
Zintegrowane	Pearsona	0,523	0,542	0,588	0,604	0,607
	Spearmana	0,433	0,460	0,567	0,574	0,582
	Kendalla	0,306	0,326	0,397	0,413	0,415
Zintegrowane2	Pearsona	0,528	0,545	0,594	0,610	0,613
	Spearmana	0,445	0,471	0,563	0,579	0,587
	Kendalla	0,310	0,329	0,397	0,412	0,415
Dwuetapowe	Pearsona	0,504	0,552	0,535	0,541	0,597
	Spearmana	0,475	0,528	0,510	0,531	0,576
	Kendalla	0,339	0,377	0,371	0,380	0,412

Źródło: opracowanie własne.

Zaciemnione wartości w tab. 4 wskazują na najwyższe wartości współczynników korelacji. Okazuje się, że modele z IV, V i VI segmentami w ujęciu zintegrowanym lepiej odzwierciedlają strukturę preferencji. Niestety, tak nie jest w modelu z III klasami, chociaż kryteria informacyjne skłaniały do wyboru właśnie tylu klas. Zastanawia również to, że po porównaniu modeli z ujęcia dwuetapowego okazuje się, że model o III klasach ma wyższe współczynniki korelacji niż model z IV, a nawet V (jeśli wziąć pod uwagę współczynnik Pearsona) klasami.

Warto skupić się również na porównaniu jednorodności segmentów w obu ujęciach, której wyznacznikiem może być wariancja.

Tabela 5. Estymatory wariancji

Oszacowane wariancje w odpowiednich segmentach									
Ujęcie dwuetapowe					Ujęcie zintegrowane				
II	III	IV	V	VI	II	III	IV	V	VI
308,3	493,6	483,3	308,3	307,7	480,5	478,6	430,1	372,6	372,5
477,9	308,3	404,4	411,1	331,4	311,4	372,0	372,7	158,7	158,7
	417,4	315,8	284,9	387,1		171,4	170,2	483,5	483,6
		356,3	350,2	365,7			471,4	169,5	57,5
			499,4	219,4				416,1	416,0
				475,9					176,8

Źródło: opracowanie własne.

Począwszy od modelu z III segmentami, w ujęciu zintegrowanym występuje klasa lub klasy charakteryzujące się względnie niską wariancją (porównując z ujęciem dwuetapowym). Przy sześciu segmentach ta różnica jeszcze bardziej się pogłębia – w jednej klasie wynosi nawet 57,5. Czy można wobec tego mówić o jakiejś własności ujęcia zintegrowanego, której konsekwencją będzie odkrywanie segmentów bardziej jednorodnych niż miałoby to miejsce w wypadku ujęcia dwuetapowego? Dla praktyki marketingowej jest to istotne pytanie. Jeśli w wyniku takiej analizy strategia, firmy miałaby być ukierunkowana na jeden segment, to z pewnością taki, który jest najbardziej jednorodny (musi być również odpowiednio

liczny). Oczywiście generalizacja na podstawie badania jednostkowego jest niedopuszczalna. Dlatego dodatkowe analizy oparte na symulacjach i danych rzeczywistych byłyby wskazane.

5. Podsumowanie

Teoretyczne rozważania dotyczące porównania różnych ujęć pokazują, że model oparty na mieszkankach rozkładów jest bardziej elastyczny, gdyż po pierwsze – pozwala zrezygnować z założenia o skalarnej macierzy kowariancji (w ujęciu dwuetapowym jest to niemożliwe), a po drugie – we wzorach na estymatory występują średnie ważone o wagach należących do przedziału (0,1) (w ujęciu konkurencyjnym waga ma wartość 0 lub 1). Druga uwaga skłania do przypuszczeń, że ujęcia powinny dawać takie same wyniki, jeśli wagi w ujęciu zintegrowanym są bliskie 1. Empiryczna weryfikacja nie potwierdziła tego. Mimo że dla zdecydowanej większości respondentów wagi były wyższe od 0,9, to rozmiary segmentów różniły się. Wydaje się więc, że mechanizm klasyfikacji w obu ujęciach jest trochę inny.

Próba odpowiedzi na pytanie o wyższość jednego ujęcia nad drugim polegała również na zbadaniu zależności między preferencjami zaobserwowanymi a oszacowanymi na podstawie modelu. Tutaj nie było jednoznacznych rozstrzygnięć: modele z IV, V i VI klasami w ujęciu zintegrowanym lepiej odzwierciedliły strukturę preferencji, a modele z II i III klasami okazały się lepsze dla ujęcia dwuetapowego.

Ostatnim, istotnym elementem prowadzonego porównania była jednorodność segmentów mierzona wariancją. W ujęciu zintegrowanym można było zaobserwować tendencję do tworzenia się segmentów o zdecydowanie większej jednorodności. Z punktu widzenia marketingu jest to cenna informacja. Aczkolwiek wskazane byłyby w tym obszarze pogłębione analizy, pozwalające rozstrzygnąć, czy można mówić o pewnej własności modelu opartego na mieszankach rozkładów.

Literatura

Dempster A.P., Laird N.M., Rubin D.B. (1977), *Maximum Likelihood From Incomplete Data via the EM Algorithm*, „Journal of the Royal Statistical Society", Ser. B, nr 1(39), s. 1-22.

Hagerty M.R., (1985), *Improving the Predictive Power of Conjoint Analysis: The Use of Factor Analysis and Cluster Analysis*, „Journal of Marketing" vol. 23, nr 2, str. 168-184.

McLachlan G.J., Krishnan T. (1997), *The EM Algorithm and Extensions*, Wiley, New York.

Vriens M., Wedel M., Wilms T. (1996), *Metric Conjoint Segmentation Methods: A Monte Carlo Comparison*, „Journal of Marketing Research" vol. 33, nr 1, s. 73-85.

Walesiak M., Bąk A. (2000), *Conjoint analysis w badaniach marketingowych*, AE, Wrocław.

INTEGRATED AND TWO-STAGE APPROACH TO SEGMENTATION IN CONJOINT ANALYSIS

Summary

The segmentation of markets with conjoint analysis involves two-stage or integrated approach. In two-stage approach the estimation and the segmentation are performed separately. Conversely, if these two stages are combined then one calls it integrated approach.

The aim of this paper is to compare these approaches. Except theoretical considerations the results of questionnaire investigation (concerning preferences in the market of cellular phones) are used in this comparison too.