

Ewa GengeUniwersytet Ekonomiczny w Katowicach
e-mail: ewa.genge@ue.katowice.pl

**TEORIA REAKCJI NA POZYCJĘ
W PODEJŚCIU MODELOWYM W TAKSONOMII**

**ITEM RESPONSE THEORY
IN MODEL-BASED CLUSTERING**

DOI: 10.15611/ekt.2016.1.01

JEL Classification: C55

Streszczenie: Teoria reakcji na pozycję (*item response theory*) zaliczana jest do jednego z dwóch nurtów metodologicznych w ocenie rzetelności skali. Z kolei analizę klas ukrytych (*latent class analysis*) można wpisać w nurt podejścia modelowego w taksonomii, wykorzystującego ideę mieszanek rozkładów. Modele te wykorzystywane są do analizy jakościowych zbiorów danych o niejednorodnej strukturze, w których liczba klas jest nieznana (tzw. zmienna ukryta). W ostatnim czasie na popularności zyskuje podejście modelowe w taksonomii, łączące teorię reakcji na pozycje z modelami klas ukrytych. Celem pracy jest przedstawienie propozycji wykorzystania podejścia modelowego w taksonomii, wykorzystującego teorię IRT w analizie zdolności do oszczędzania w polskim społeczeństwie. Badania przeprowadzone będą z zastosowaniem pakietu `MultiLCIRT` programu R.

Słowa kluczowe: teoria IRT, analiza klas ukrytych, podejście modelowe w taksonomii.

Summary: Item response theory is considered to be one of the two trends in methodological assessment of the reliability scale. In turn, latent class models can be viewed as a special case of model-based clustering, for heterogenous multivariate discrete data. We used the approach combining item response theory and latent class models to find groups of Polish households' with similar saving ability levels. We analyzed data collected as part of the Polish Social Diagnosis using `MultiLCIRT` package of R.

Keywords: item response theory, latent class analysis, model-based clustering.

1. Teoria reakcji na pozycję

Teoria reakcji na pozycję (*item resposne theory*) zaliczana jest do jednego z dwóch nurtów metodologicznych w ocenie rzetelności skali. Zdecydowanie większą popularnością cieszy się klasyczna teoria testu (*classical test theory*), która jest punktem odniesienia większości metod rzetelności skali. Najbardziej popularnym mierni-

kiem rzetelności (homogeniczności skali) jest alpha Cronbacha. Należy zauważyć, że współczynnik ten należy interpretować jako wygodną miarę oceny rzetelności skali, a nie jej wymiarowości. Warto podkreślić, że metody budowane na klasycznej teorii testu (KTT) wymagają przyjęcia licznych założeń, np. skala jest jednowymiarowa, czyli wszystkie pozycje skali są odzwierciedleniem jednego i tego samego czynnika.

Teoria reakcji na pozycję (TRP) reprezentuje alternatywny model pomiaru w stosunku do klasycznej teorii testu i przełamuje zarazem podstawowe ograniczenia KTT. Na gruncie tej teorii, będącej teorią probabilistyczną, powstały modele IRT (*Item Response Theory Models*). Modele te wyrażają prawdopodobieństwo określonej reakcji na pozycję skali jako funkcję „zdolności” respondenta oraz „trudności” danej pozycji. Według tej teorii, odpowiedzi na skali mogą być wyjaśnione poprzez czynniki określające ich reakcje poznawcze lub emocjonalne. Wzajemne relacje pomiędzy nieobserwowanymi zdolnościami respondentów a zmiennymi obserwowalnymi (odpowiedziami) mogą być opisane za pomocą rosnącej funkcji monotonicznej, która jest nazywana krzywą charakterystyczną pozycji (*item characteristic function*). Funkcja ta ma charakter probabilistyczny. W modelu Rascha [1960] czy modelu Birnbauma [1968] jest to parametryczna funkcja logistyczna. W ostatnich latach zbudowano wiele modeli pomiaru opartych na podejściu Rascha. Należą do nich m.in. modele cząstkowego zaufania (*partial credit model*) Mastersa [1982], model stopniowanej reakcji (*graded response model*) Samejimy [1969], model skal ocen (*rating scale model*) Andricha [Andrich 1978].

W większości klasycznych modeli IRT wykorzystywany jest jeden parametr opisujący położenie respondentów na nieznanym kontinuum badanej cechy, zwanej zmienną ukrytą (np. zdolności, postawy, preferencje itp. – zakłada się, że zmienna ukryta jest ciągła), natomiast różna jest liczba parametrów opisujących położenie pozycji na tym kontinuum. W zależności od typu modelu występują parametry trudności pozycji, parametr mocy dyskryminacyjnej pozycji czy parametr zgadywania wartości pozycji, zwany też parametrem określającym tzw. pseudoszanse respondenta odgadywania trudnych pozycji skali (*guessing parameter*)¹ [Sagan 2002].

2. Modele IRT o dyskretnej cesze ukrytej

Zwykle klasyczne modele IRT, np. model Rascha czy model Birnbauma, wykorzystywane są w analizach jednej zmiennej obserwowanej. W przypadku badań ankietowych o większej liczbie pytań (badania sondażowe, testy edukacyjne) zakłada się,

¹ Najbardziej ogólny jest trójparametryczny model Birnbauma, w którym prawdopodobieństwo prawidłowej odpowiedzi na pozycję skali jest funkcją trudności pozycji, mocy dyskryminacyjnej pozycji i zgadywania odpowiedzi na pozycję skali. W modelu dwuparametrycznym współczynnik zgadywania jest ustalony na poziomie zerowym, a w modelu jednoparametrycznym (którego szczególnym przypadkiem jest model Rascha) dodatkowo współczynnik mocy dyskryminacyjnej jest stały dla wszystkich pozycji skali (w modelu Rascha wynosi on 1).

że badana cecha ukryta jest jednowymiarowa (zakłada się, że skala jest jednowymiarowa), co oznacza, że wszystkie obserwowane zmienne opisują tę samą cechę ukrytą. Ponadto czasami przyjmuje się również założenie normalności rozkładu dla cechy ukrytej. Niestety, w wielu praktycznych zastosowaniach założenia te są zbyt restrykcyjne. Dlatego też wielu autorów dokonało różnych modyfikacji klasycznych modeli IRT (zob. np. [Bacci i in. 2014; Bartolucci i in. 2014]).

Bartolucci [2007] zaproponował rodzinę wielowymiarowych ukrytych modeli IRT, w których:

- możliwe jest uwzględnienie wielowymiarowości cechy ukrytej (wymiar cechy ukrytej zwany będzie również ukrytym czynnikiem). Parametry tej zmiennej szacowane są równocześnie z pozostałymi parametrami modelu. Każda z obserwowanych zmiennych przypisana jest tylko do jednego z czynników (*between-item multidimensionality*) (zob. [Adams i in. 1997; Zhang 2004]),
- (wielowymiarowa) cecha ukryta charakteryzowana jest za pomocą wektora liczb losowych o rozkładzie dyskretnym, wspólnym dla wszystkich obiektów. Każda z wartości tego rozkładu (*support point*) opisuje inną klasę ukrytą, tj. wartości cechy ukrytej (poziomy zdolności respondentów) wyznaczone są dla skończonej liczby klas (liczba klas ukrytych jest zmienną nominalną).

3. Modele IRT a analiza klas ukrytych

W wielowymiarowym modelu IRT dany jest wektor m jakościowych zmiennych \mathbf{X} , $\mathbf{X} = X_1, \dots, X_m$ obserwowanych. Niech X_j oznacza j -tą ($j = 1, \dots, m$) zmienną obserwowaną (j -tą pozycję skali) o liczbie kategorii l_j , przyjmującą wartości $x = 0, \dots, l_j - 1$. Dla zmiennej binarnej liczba kategorii jest równa dwa ($x = 0, 1$). Niech q oznacza liczbę ukrytych czynników (wymiar zmiennej ukrytej) wyodrębnionych z analizowanych zmiennych obserwowanych, $\Theta = (\Theta_1, \dots, \Theta_q)'$ to wektor ukrytej zmiennej (czynników), a $\theta = (\theta_1, \dots, \theta_q)'$ będzie jedną z jego możliwych realizacji. Zakłada się, że Θ pochodzi z rozkładu dyskretnego i przyjmuje wartości ξ_1, \dots, ξ_u z prawdopodobieństwem π_1, \dots, π_u ($\pi_s = p(\Theta = \xi_s)$).

Ponadto niech δ_{jd} oznacza zmienną binarną, przyjmującą wartość 1, w przypadku gdy j -tą zmienną obserwowaną można przypisać do d -tego czynnika, i 0 w przeciwnym razie ($j = 1, \dots, m$, $d = 1, \dots, q$). Prawdopodobieństwo warunkowe tego, że respondent o cesze ukrytej θ (wyrażającej np. poziom zdolności, umiejętności, skłonności do oszczędzania) udzieli odpowiedzi x dla zmiennej X_j (na tzw. j -tą pozycję skali), wynosi:

$$\phi_{jx}(\theta) = p(X_j = x \mid \Theta = \theta), x = 0, \dots, l_j - 1, \quad (1)$$

gdzie suma wszystkich $\phi_j(\theta) = (\phi_{j0}(\theta), \dots, \phi_{jl_j-1}(\theta))'$ jest równa 1.

Wielowymiarowy model IRT² można przedstawić jako:

$$g_x(\boldsymbol{\varphi}_j(\boldsymbol{\Theta})) = \alpha_j \left(\sum_{d=1}^q \delta_{jd} \phi_d - \mathcal{G}_{jx} \right), \quad j=1, \dots, m, \quad s=1, \dots, u, \quad (2)$$

gdzie g_x jest funkcją połączeń dla kategorii x i tzw. parametrach skali, tj. parametr mocy dyskryminacyjnej pozycji (α_j) oraz parametr trudności pozycji (\mathcal{G}_{jx}). W dalszej części pracy autor będzie posługiwał się nazwami skrótowymi – parametr dyskryminacji i trudności.

Uwzględniając różne możliwości ograniczeń nakładanych na parametry modelu IRT oraz różne rodzaje funkcji połączeń³, Bacci i in. [2014], Bartolucci i in. [2014] rozszerzyli grupę klasycznych, jednowymiarowych modeli IRT. Na przykład wielowymiarowy ukryty model stopniowanej reakcji (*multidimensional latent class graded response model*) dany równaniem (3) jest rozszerzeniem modelu GRM stworzonego przez Semejima w 1969 roku. Z kolei ukryty model skal ocen (*multidimensional LC rating scale model*), dany równaniem (4), jest rozszerzeniem modelu RSM [Andrich 1978]:

$$g_{LC_GRM}(\boldsymbol{\varphi}_j(\boldsymbol{\Theta})) = \log \frac{p(X_j \geq x | \boldsymbol{\Theta} = \boldsymbol{\Theta})}{p(X_j < x - 1 | \boldsymbol{\Theta} = \boldsymbol{\Theta})} = \alpha_j \left(\sum_{d=1}^q \delta_{jd} \varphi_d - \mathcal{G}_{jx} \right), \quad x=0, \dots, l_j - 1, \quad (3)$$

$$g_{LC_RSM}(\boldsymbol{\varphi}_j(\boldsymbol{\Theta})) = \log \frac{p(X_j = x | \boldsymbol{\Theta} = \boldsymbol{\Theta})}{p(X_j = x - 1 | \boldsymbol{\Theta} = \boldsymbol{\Theta})} = \sum_{d=1}^q \delta_{jd} \varphi_d - (\mathcal{G}_{jx} + \alpha_x), \quad x=0, \dots, l_j - 1, \quad (4)$$

gdzie α_x to parametr dyskryminacji dla kategorii x i wszystkich zmiennych obserwowanych. W przypadku gdy $l_j = 2$, $j=1, \dots, m$, równanie (3) odpowiada wielowymiarowemu ukrytemu modelowi 2PL⁴ (*multidimensional LC 2PL model*), równanie (4) zaś – wielowymiarowemu ukrytemu modelowi Rascha (*multidimensional LC Rasch model*) (zob. [Bartolucci 2007]).

Ze względu na to, że rozkład wektora $\boldsymbol{\Theta}$ jest dyskretny, rozkład zmiennej obserwowanej $\mathbf{X} = (X_1, \dots, X_m)'$ dla wszystkich obiektów w s -tej klasie ukrytej można zapisać jako (por. [Bacci i in. 2014]):

² Wielowymiarowy model IRT to taki, w którym ukryta zmienna, wyrażająca zdolności respondentów, może przyjmować więcej niż jeden wymiar (w odróżnieniu od klasycznych modeli IRT, np. modelu Rascha), można tu wyróżnić np. zdolności humanistyczne, matematyczne-przyrodnicze.

³ Wyróżnić tu można tzw. lokalną (*local logit link function*) oraz globalną czy też skumulowaną logitową funkcję połączeń (*global logit link function*) (zob. [Forcina, Bartolucci 2004]). W przypadku zmiennych binarnych funkcje te są sobie równe [Bartolucci i in. 2014].

⁴ Pełna nazwa to wielowymiarowy ukryty dwuparametryczny model regresji logistycznej Birnbauma.

$$p(\mathbf{X} = \mathbf{x}) = \sum_{s=1}^u p(\mathbf{X} = \mathbf{x} | \Theta = \xi_s) \pi_s, \quad (5)$$

uwzględniając klasyczne założenie lokalnej niezależności:

$$\begin{aligned} p(\mathbf{x} | s) = p(\mathbf{X} = \mathbf{x} | \Theta = \xi_s) &= \prod_{j=1}^m p(X_j = x_j | \Theta = \xi_s) \\ &= \prod_{d=1}^q \prod_{j \in I_d} p(X_j = x_j | \Theta_d = \xi_{sd}), \end{aligned} \quad (6)$$

gdzie: I_d to podzbiór $I = 1, \dots, m$ zawierający zmienne obserwowane przypisane do d -tego czynnika ukrytego (d -tego wymiaru cechy ukrytej, $d = 1, \dots, q$).

4. Wybór modelu optymalnego

Wybór optymalnego modelu pociąga więc za sobą podjęcie kilku decyzji dotyczących:

- liczby klas $s = 1, \dots, u$,
- funkcji połączeń g_x ,
- ograniczeń przyjmowanych na parametry funkcji połączeń α_j , θ_{jx} ,
- liczby ukrytych czynników (wymiaru skali) (ang. *latent dimension*) $d = 1, \dots, q$.

Parametry ukrytych modeli IRT najczęściej szacowane są za pomocą algorytmu EM [Dempster 1977, s. 1-38], a wybór liczby klas oraz funkcji połączeń wybierany jest na podstawie kryteriów informacyjnych BIC i AIC [Akaike 1974; Schwarz 1978]. Z kolei wybór liczby ukrytych czynników (wymiaru zmiennej ukrytej), jak również ograniczeń przyjętych na parametry trudności czy dyskryminacji dokonywany jest na podstawie testu ilorazu wiarygodności.

W teście ilorazu wiarygodności (w badaniu wielowymiarowości skali, tj. cechy ukrytej) hipoteza zerowa głosi, że zmienne w podzbiorze I_{d_1} można przypisać do tego samego wymiaru co zmienne w podzbiorze I_{d_2} , gdzie $d = 1, \dots, q$, a I_d to podzbiór $I = 1, \dots, m$. Zakładając, że wszystkie pozostałe kryteria wyboru w analizowanym modelu (liczba klas, ograniczenia parametrów skali, tj. zdolności i dyskryminacji oraz rodzaj funkcji logitowej) pozostają niezmiennie, porównywany jest model ogólny o q -wymiarowej cesze ukrytej z tzw. modelem ograniczonym o liczbie wymiarów zmiennej ukrytej równej $q - 1$. W przypadku testowania hipotezy o dwuwymiarowości skali (dzwuwymiarowej zmiennej ukrytej) testowany jest model jednowymiarowy (wszystkie zmienne są przypisane do jednego podzbioru) z modelem, w którym zmienne należą do podzbiorów I_{d_1} oraz I_{d_2} . Statystyka ilorazu największej wiarygodności jest określona jako $-2(\hat{l}_0 - \hat{l}_1)$, gdzie \hat{l}_0 i \hat{l}_1 oznaczają maksimum funkcji największej wiarygodności modelu ograniczonego i modelu ogólnego. Statystyka ilorazu największej wiarygodności ma asymptotyczny rozkład χ^2_b , gdzie b jest różnicą liczby parametrów dwóch porównywanych modeli (zagnieżdżonych).

W podobny sposób (na podstawie testu ilorazu wiarygodności) podejmowana jest również decyzja o wyborze stosownych ograniczeń przyjmowanych na parametry funkcji połączeń.

5. Analiza empiryczna

Analizę empiryczną przeprowadzono na podstawie danych dotyczących polskich gospodarstw domowych, pochodzących z projektu Diagnoza Społeczna [Czapiński, Panek (red.) 2016]. Analizie podano 7399 gospodarstw, które wzięły udział w ostatnim badaniu, tj. przeprowadzonym w roku 2015. Analiza została dokonana z uwzględnieniem dwunastu zmiennych obserwowanych, charakteryzujących cele oszczędności gospodarstw domowych⁵.

W nawiasie podano oryginalne nazwy zmiennych analizowanego zbioru, udostępnianego na stronie internetowej www.diagnoza.com.

X_1 (HF8_01): rezerwa na bieżące wydatki konsumpcyjne (np. żywność, odzież i ubranie); X_2 (HF8_02): stałe opłaty (np. mieszkaniowe); X_3 (HF8_03): zakup dóbr trwałego użytku; X_4 (HF8_04): zakup domu, zakup mieszkania, wkład do spółdzielni mieszkaniowej; X_5 (HF8_05): remont domu, mieszkania; X_6 (HF8_06): leczenie; X_7 (HF8_07): rehabilitacja; X_8 (HF8_08): wypoczynek; X_9 (HF8_09): rezerwa na sytuacje losowe, tj. na „czarną godzinę”; X_{10} (HF8_10): zabezpieczenie przyszłości dzieci; X_{11} (HF8_11): zabezpieczenie na starość; X_{12} (HF8_12): rozwój działalności gospodarczej.

W badaniach wykorzystano pakiet `MultiLCIRT` [Bartolucci i in. 2015] programu R. Na początku na podstawie wartości kryteriów informacyjnych AIC oraz BIC dokonano analizy jednorodności badanego zbioru obserwacji⁶. Kryterium BIC oraz AIC wskazało minimalną wartość dla trzech ukrytych grup respondentów. W kolejnym kroku dokonano wyboru funkcji połączeń modeli IRT (zob. (3) i (4)). Z racji tego, że analizowane zmienne mierzone są na skali dychotomicznej, wybrano lokalną funkcję połączeń. Następnie zweryfikowano hipotezę głoszącą, że badana cecha ukryta (zdolność do gromadzenia oszczędności) jest jednowymiarowa, wobec hipotezy alternatywnej: umiejętność gromadzenia oszczędności jest inna na potrzeby bieżące oraz na przyszłość⁷. Hipotezę zweryfikowano na podstawie testu wielowymiarowości, którego wyniki przedstawiono na rys. 1.

⁵ Pytanie w kwestionariuszu brzmiało: „W jakim celu Pana(i) gospodarstwo domowe gromadzi oszczędności?”. Należało udzielić odpowiedzi (tak lub nie) dla każdego z wymienionych na rys. 1 celów oszczędzania (zmienne X_1 - X_{12}).

⁶ Założono, że umiejętność gromadzenia oszczędności nie jest taka sama w badanej próbie 7399 respondentów.

⁷ Do pierwszego wymiaru („potrzeby bieżące”) przypisano zmienne: $X_1, X_2, X_5, X_6, X_7, X_8$, do wymiaru zaś drugiego („przyszłość”): $X_3, X_4, X_9, X_{10}, X_{11}, X_{12}$.

```
Call:
test_dim(S = S, yv = yv, k = 3, link = 1, disc = 0, difl = 0,
        multil = dim2)
```

Testing dimension output:

```
Log-likelihood of the constrained model    -40165.153
AIC of the constrained model                80362.305
BIC of the constrained model                80472.851
N.parameters of the constrained model       16.000
Log-likelihood of the unconstrained model  -40103.096
AIC of the unconstrained model             80242.192
BIC of the unconstrained model             80366.556
N.parameters of the unconstrained model     18.000
Deviance                                    124.114
Degrees of freedom                          2.000
p-value                                     0.000
```

Rys. 1. Wyniki testu wielowymiarowości – wydruk z programu R

Źródło: opracowanie własne.

W kolejnym kroku na podstawie testu ilorazu wiarygodności dokonano wyboru modelu o nieograniczonych parametrach dyskryminacji, tj. wielowymiarowego ukrytego modelu 2PL.

Oszacowane prawdopodobieństwa *a priori* oraz parametry zdolności dla dwóch wymiarów zmiennej ukrytej oraz trzech wyodrębnionych klas przedstawione zostały w tab. 1.

Tabela 1. Prawdopodobieństwa *a priori*, parametry zdolności dla klas

Symbol/klasa	Klasa 1	Klasa 2	Klasa 3
π_s	0,26	0,55	0,19
ξ_{s1}	-0,53	-0,76	0,29
ξ_{s2}	-3,62	-1,78	0,46

Źródło: opracowanie własne.

Do klasy pierwszej, stanowiącej 26% wszystkich gospodarstw, należą gospodarstwa, które cechują się najmniejszą zdolnością do oszczędzania na przyszłość ($\xi_{s2} = -3,62$). Gospodarstwa te kierują się głównie przezornościowym motywem oszczędzania (oszczędzają na czarną godzinę 34%, zob. rys. 2). Natomiast ich zdolność do oszczędzania na potrzeby bieżące ($\xi_{s1} = -0,53$) jest wyższa aniżeli w przypadku respondentów klasy drugiej.

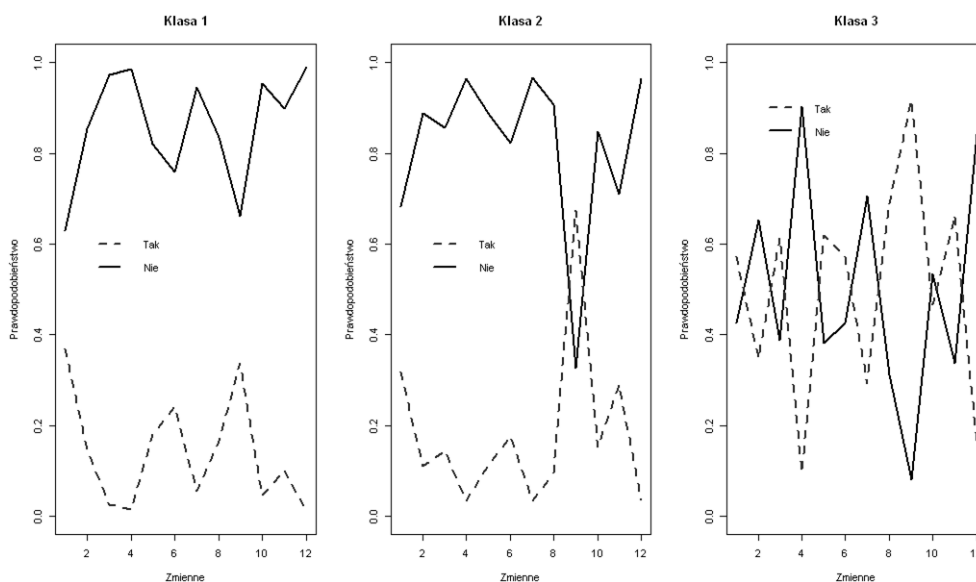
Klasa druga jest klasą najliczniejszą – należy do niej 55% gospodarstw domowych. Respondentów tej klasy wyróżnia umiarkowana zdolność oszczędzania na przyszłość ($\xi_{s2} = -1,78$) oraz najniższa zdolność do oszczędzania na potrzeby bieżące ($\xi_{s1} = -0,76$).

Największą zdolność do oszczędzania na przyszłość ($\xi_{s,2} = 0,46$) można zaobserwować w klasie trzeciej, zawierającej 19% wszystkich gospodarstw. Tam też obserwuje się największą skłonność do oszczędzania na potrzeby bieżące ($\xi_{s,1} = 0,29$).

Podobne wnioski można wyciągnąć również na podstawie prawdopodobieństw warunkowych, oszacowanych dla poszczególnych zmiennych w wyodrębnionych klasach (rys. 2). W klasie pierwszej największy odsetek respondentów (37%) gromadzi środki na bieżące wydatki, następnie na sytuacje losowe (34%) czy też na leczenie (24%). Znikomy odsetek gospodarstw tej klasy oszczędza na rozwój działalności gospodarczej (1%), zakup domu (1,4%), zakup dóbr trwałego użytku (2,6%) czy też zabezpieczenie przyszłości dzieci (5%).

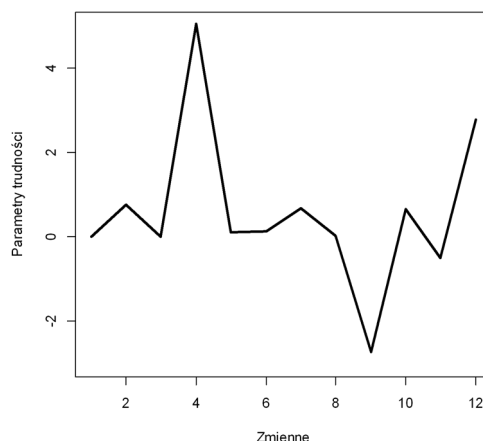
W klasie drugiej największy odsetek gospodarstw (67%) oszczędza na „czarną godzinę”, na starość (29%) oraz na zakup dóbr trwałego użytku (ponad 14%). Tak jak już wspomniano, klasę tę wyróżnia najniższa skłonność do oszczędzania na potrzeby bieżące, tj. 32% gospodarstw oszczędza na żywność, odzież czy ubrania, 11% na opłaty stałe, 9% gospodarstw oszczędza na wypoczynek i tylko 3% gospodarstw oszczędza na rehabilitację.

Klasę trzecią cechuje najwyższa skłonność do oszczędzania zarówno na przyszłość (92% gospodarstw oszczędza na „czarną godzinę”, 66% na starość, 61% na zakup dóbr trwałego użytku, 45% na przyszłość dzieci), jak i na potrzeby bieżące (57% na żywność, 62% na remont domu, 57% na leczenie, a 35% na opłaty stałe).



Rys. 2. Prawdopodobieństwa warunkowe w poszczególnych klasach

Źródło: opracowanie własne.



Rys. 3. Parametry trudności dla zmiennych X_1 - X_{12}

Źródło: opracowanie własne.

Analizie poddano również oszacowane parametry skali, tj. parametry trudności i dyskryminacji. Najmniejsze wartości parametrów trudności oszacowano dla zmiennych X_9 i X_{11} , a najwyższe dla zmiennej X_4 . Z kolei dla zmiennych: X_8 , X_5 i X_7 zaobserwowano najwyższe wartości parametrów dyskryminacji, a wartość najniższą dla zmiennej X_4 . Parametry trudności dla wszystkich zmiennych zamieszczono na rys. 3.

6. Zakończenie

W artykule przedstawiono podejście łączące ze sobą dwa rodzaje modeli wykorzystywanych w badaniach cechy ukrytej: modelowanie w ramach teorii odpowiedzi na pozycje testowe oraz analizę klas ukrytych (*Latent Class Analysis, LCA*), będącą szczególnym przypadkiem skończonych modeli mieszanek (*finite mixture models*). W modelach mieszanek zakłada się, że zbiór badanych respondentów jest niejednorodny pod względem badanych zmiennych (zmienną ukrytą jest klasa, *latent class*) (zob. np. [Bąk 2011; Genge 2013, 2014; Lazarsfeld, Henry 1968]), w modelach IRT zaś (w modelach teorii reakcji na pozycję) ukryta jest tzw. zdolność reakcji respondenta na pozycje skali (cechą ukrytą (*latent trait*) jest zdolność respondenta).

W analizowanym przykładzie empirycznym w celu uzyskania odpowiedzi, czy badany zbiór respondentów można potraktować jako jednorodny (biorąc pod uwagę styl odpowiadania na pytania testowe), porównano miary jakości dopasowania klasycznego modelu IRT oraz ukrytych modeli IRT. W kolejnym kroku zweryfikowano hipotezę o dwuwymiarowości cechy ukrytej. Na podstawie kryteriów informacyj-

nych oraz testów ilorazu wiarygodności wybrano 3 grupy gospodarstw. Najlepiej dopasowanym okazał się ukryty model 2PL, którego parametry poddano interpretacji.

Badania literaturowe oraz otrzymane rezultaty wskazują, że modele te są użyteczną metodą szacowania parametrów w różnych analizach testów egzaminacyjnych, ale również w badaniu zachowań finansowych gospodarstw domowych. Jednak zasadność ich zastosowania jest uzależniona od specyfiki danego badania. Głównym ograniczeniem prezentowanego podejścia badawczego jest wybór ukrytej liczby czynników oraz przypisanie im stosownych zmiennych. Ponadto wraz ze wzrostem liczby ukrytych wymiarów wydłuża się czas obliczeń algorytmu. Na tym etapie można wprawdzie posłużyć się testem ilorazu wiarygodności czy też dendrogramem (graficznym wynikiem hierarchicznej analizy taksonomicznej [Bartolucci 2007]), należy jednak podkreślić, że trzeba pamiętać o znaczeniu i interpretacji oszacowanych parametrów.

W dalszych badaniach analizie poddane zostaną również wielopoziomowe i wielowymiarowe ukryte modele IRT (*multidimensional and multilevel LC IRT models*) [Bacci i in. 2014; Gnaldi i in. 2015].

Literatura

- Adams R., Wilson M., Wang W., 1997, *The multidimensional random coefficients logit*, Applied Psychological Measurement, 21, s. 1-24.
- Akaike H., 1974, *A new look at statistical model identification*, IEEE Transactions on Automatic Control, 19, s. 716-723.
- Andrich D., 1978, *A rating formulation for ordered response categories*, Psychometrika, 43, s. 561-573.
- Bacci S., Bartolucci F., Gnaldi M., 2014, *A class of Multidimensional Latent Class IRT models for ordinal polytomous item responses*, Communication in Statistics – Theory and Methods, 43, s. 787-800.
- Bartolucci F., 2007, *A class of multidimensional IRT models for testing unidimensionality and clustering items*, Psychometrika, 72, s. 141-157.
- Bartolucci F., Bacci S., Gnaldi M., 2014, *MultiLCIRT: An R package for multidimensional latent class item response models*, Computational Statistics and Data Analysis, 71, s. 971-985
- Bartolucci F., Bacci S., Gnaldi M., 2015, *MultiLCIRT: Multidimensional latent class Item Response Theory models. R package version 2.9*, URL <http://CRAN.R-project.org/package=MultiLCIRT>.
- Bąk A., 2011, *Modele klas ukrytych dla danych jakościowych*, [w:] E. Gatnar, M. Walesiak, *Analiza danych jakościowych i symbolicznych z wykorzystaniem programu R*, C.H. Beck, Warszawa, s. 204-222.
- Birnbaum A., 1968, *Some Latent Trait Models and Their Use in Inferring an Examinee's Ability*, [w:] F.M. Lord, M.R. Novick, *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, MA, s. 395-479.
- Czapiński J., Panek T. (red.), 2016, *Diagnoza społeczna 2015. Warunki i jakość życia Polaków (raport)*, Warszawa, Rada Monitoringu Społecznego (29.01.2016).
- Dempster A.P., Laird N.P., Rubin D.B., 1977, *Maximum likelihood for incomplete data via the EM algorithm (with discussion)*, Journal of the Royal Statistical Society, 39, ser. B, s. 1-38.
- Forcina A., Bartolucci F., 2004, *Modelling quality of life variables with non-parametric mixtures*, Environmetrics, 15, s. 519-528.

- Genge E., 2013, *Poczucie śląskości wśród Ślązaków – analiza empiryczna z wykorzystaniem modeli klas ukrytych*, *Ekonometria* 4(42), Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław, s. 48-60.
- Genge E., 2014, *A latent class analysis of the public attitude towards the euro adoption in Poland*, *Advances in Data Analysis and Classification* 8(4), Springer, s. 427-442.
- Gnaldi M., Bacci S., Bartolucci F., 2015, *A multilevel finite mixture item response model to cluster examinees and schools*, *Advances in Data Analysis and Classification*.
- Lazarsfeld P.F., Henry N.W., 1968, *Latent Structure Analysis*, Houghton Mill, Boston, MA.
- Masters G., 1982, *A Rasch model for partial credit scoring*, *Psychometrika*, 47, s. 149-174.
- Rasch G., 1960, *Probabilistic Models for Some Intelligence and Attainment Tests*, Danish Institute for Educational Reserch, Copenhagen.
- Sagan A., 2002, *Zastosowanie wielowymiarowych skal czynnikowych i skal Rascha w badaniach marketingowych (na przykładzie oceny efektów komunikacyjnych reklamy)*, *Zeszyty Naukowe Akademii Ekonomicznej w Krakowie*, 605, s. 73-92.
- Samejima F., 1969, *Estimation of ability using a response pattern of graded scores*, *Psychometrika Monograph*, 17.
- Schwarz G., 1978, *Estimating the dimension of a model*, *Annals of Statistics*, 6, s. 461-464.
- Zhang J., 2004, *Comparison of unidimensional and multidimensional approaches to IRT parameter estimation*, *ETS Research Report Series*, RR-04-44, s. 1-42.