EKONOMETRIA ECONOMETRICS 4(58) • 2017

ISSN 1507-3866
e-ISSN 2449-9994

Ewa Genge
University of Economics in Katowice
e-mail: ewa.genge@ue.katowice.pl

IRT AND LC-IRT MODELS FOR ITEMS WITH ORDINAL POLYTOMOUS RESPONSES

MODELE IRT ORAZ LC-IRT DLA ZMIENNYCH POLITOMICZNYCH (O KATEGORIACH UPRZĄDKOWANYCH)

DOI: 10.15611/ekt.2017.4.04
JEL Classification: C1, C83, J80

Summary: Item response theory (IRT) is a model-based theory in which the responses to test items depend on some person and item characteristics, according to specific probabilistic relations. The simplest and most popular are dichotomous IRT models that specify a single (i.e. unidimensional) latent trait under the assumption of normal distribution. This article reviews the latent class ordinal polytomous item response models (LC-IRT) and presents the comparison with the well-known traditional IRT models (based on the assumption of a normally distributed latent trait). The main goal of the article is to compare the estimation results of different kinds of ordinal polytomous IRT models with continuous and discrete latent variable in measuring job satisfaction in Poland. We analyzed data collected as part of the International Social Survey Programme using the `ltm` and `MultiLCIRT` packages of `R`.

Keywords: IRT theory, polytomous IRT models, latent trait, latent class models.

Streszczenie: Modele IRT oparte są na probabilistycznej teorii testu i wyrażają prawdopodobieństwo określonej reakcji na pozycję skali jako funkcję zdolności respondenta oraz trudności danej pozycji. W różnego rodzaju analizach najczęściej stosowane są dychotomiczne modele IRT, w których zakłada się, że zmienne obserwowane są dyskretne, a zdolność respondenta (tzw. zmienna ukryta) jest zmienną ciągłą. Celem pracy będzie zestawienie klasycznych politomicznych (o kategoriach uporządkowanych) modeli IRT, tj. o zmiennej ukrytej pochodzącej z rozkładu normalnego, z zyskującymi na popularności modelami IRT o dyskretnej cesze ukrytej (LC-IRT) w badaniu satysfakcji z pracy polskich pracowników. Badania przeprowadzone będą z zastosowaniem pakietów `ltm` oraz `MultiLCIRT` programu `R` dla danych pochodzących z Międzynarodowego Programu Sondaży Społecznych *ISSP 2015*.

Słowa kluczowe: teoria IRT, politomiczne modele IRT, zmienna ukryta, modele klas ukrytych.

1. Introduction

Item response theory (IRT) is widely used in educational and psychological research to model how participants respond to test items in isolation and in bundles [Thissen and Wainer 2001]. The simplest and most popular in the analysis of questionnaires are dichotomous IRT models such as Rasch [1960] or a two-parameter logistic (2PL) [Birnbaum 1968] model. This article reviews the latent class ordinal polytomous item response (LC-IRT) models and presents the comparison with the well-known traditional IRT models (based on the assumption of a normally distributed latent trait).

Polytomous responses include both nominal and ordinal responses. In the nominal responses, there is no natural ordering in the item response categories. In the ordinal responses (presented in the empirical part of this paper) each item has responses corresponding to a number of ordered categories. Examples of ordinal polytomous variables are social class indicator variables (low, middle, high), proficiency level variables (low, moderate, moderately high, high) and Likert scale variables that measure agreement with an item (strongly disagree, disagree, agree, and strongly agree). Ordinal polytomous items are widespread in several contexts such as education, marketing and psychology (see [Wright and Masters 1982; Hambleton and Swaminathan 1985; Verhelst et al. 2005]). Depending on the complexity of the adopted item parametrization, different types of IRT models are defined, among the most well known models for polytomous responses are: the Graded Response Model (GRM) [Semejima 1969], the Partial Credit Model (PCM) [Masters 1982], the Rating Scale Model (RSM) [Andrich 1978] and the Generalized Partial Credit Model (GPCM) [Muraki 1992]. These models are based on the unidimensionality assumption and, most of the time, the normality assumption of this latent trait is explicitly introduced.

In the extended theory for items with ordinal polytomous responses, the random vector used to represent the latent traits (more than one) is assumed to have a discrete distribution with support points corresponding to different latent classes in the population. This extension allows also for different parameterizations for the conditional distribution of the response variables given the latent traits[1] – such as those adopted in the Graded Response Model, in the Partial Credit Model, and in the Rating Scale Model – depending on both the type of link function and the constraints imposed on the item parameters (see [Bacci et al. 2014]).

The main goal of the article is to compare the estimation results of different kinds of ordinal polytomous IRT and LC-IRT models (with continuous and discrete latent trait) in measuring job satisfaction in Poland.

[1] Because the context of the empirical study (the comparison with IRT models under continuous approach of the (unidimensional) latent trait) and because the polytomous items considered in the empirical part of the study cannot be a priori divided between more than one dimension we will consider the unidimensional latent trait for the LC approach as well compare [Bacci et al. 2014; Bartolucci et al. 2016b; Genge 2016]).

2. IRT models for ordinal polytomous responses

To observe responses of *n* individuals to *j* ($j = 1,\dots,m$) items of the questionnaire, X_j denote the response variable for the *j*-th item. This variable has l_j categories denoted by $x = 0,\dots,l_j - 1$.

The conditional probability that a subject with latent trait (or ability) level θ responds by category x to this item is denoted by:

$$p_{jx}(\theta) = p(X_j = x \mid \Theta = \theta). \tag{1}$$

Let also $\mathbf{p}_j(\theta)$ denote the probability vector $(p_{j0}(\theta),\dots, p_{jl_{j-1}}(\theta))'$ that elements of which add up to 1.

The general formulation of IRT models for polytomous responses may be expressed as:

$$g_x(\mathbf{p}_j(\theta)) = \alpha_j(\theta - \vartheta_{jx}),\ j = 1,\dots,m,\ x = 1,\dots,l_j - 1, \tag{2}$$

where g_x is a link function specific for category x and α_j and ϑ_{jx} are item parameters which are usually identified as discrimination indices and difficulty levels[2] and on which suitable constraints may be assumed.

There are many different classification of IRT models for polytomous items i.e. based on the function assumed for $p_{jx}(\theta)$ [Van der Linden and Hambleton 1997; Nering and Ostini 2010]. We will present the criterion of the polytomous IRT models classification based on the link function specification [Agresti 2002; Bartolucci 2007].

On the basis of the specification of the link function in (2) and on the basis of the adopted constraints on the item parameters, different IRT models for polytomous responses' results. The formulation of each of these models depends particularly on:

- Type of link function:

The link function based on global (3), local (4) or continuation ratio logits (5) can be considered. IRT models based on global logits are also known as graded response models, those based on local logits are known as partial credit models and models based on continuation ratio logits are also known as sequential models.

$$g_x(\mathbf{p}_j(\theta)) = \log\frac{p(X_j \ge x \mid \theta)}{p(X_j < x \mid \theta)},\ x = 1,\dots,l_j - 1, \tag{3}$$

$$g_x(\mathbf{p}_j(\theta)) = \log\frac{p(X_j = x \mid \theta)}{p(X_j = x - 1 \mid \theta)},\ x = 1,\dots,l_j - 1, \tag{4}$$

[2] The item difficulty parameter (for ordinal response variable) is also called a threshold (difficulty) parameter.

$$g_x(\mathbf{p}_j(\theta)) = \log \frac{p(X_j > x \mid \theta)}{p(X_j = x \mid \theta)}, \quad x = 1, \ldots, l_j - 1. \tag{5}$$

- Formulation of the item parameters:

  a) Constraints on the discrimination parameters: each item may discriminate differently from the others or all the items discriminate in the same way ($\alpha_j = 1$, $j = 1, \ldots, m$)[3].

  b) Constraints of the item difficulty parameters ϑ_{jx}: the item difficulty parameters may be *unconstrained* (there are as many difficulty parameters as the number of response categories l_j minus 1) or *constrained* (the distance between difficulty levels from category to category is the same for each item ($\tau_{jx} = \tau_x$)[4]. These constraints may be expressed as:

$$\vartheta_{jx} = \vartheta_j + \tau_x, \; j = 1, \ldots, m, \; x = 0, \ldots, l_j - 1, \tag{6}$$

where ϑ_j indicates the difficulty of ite*m j* and τ_x is the difficulty of response category x for all items.

By combining the mentioned constraints, four different specifications of the item parameterization for each of the link functions are obtained. All the possible combinations of the item parameters constraints for local and global as well as continuation logit link functions are given in Table 1.

Table 1. Unidimensional IRT models for Ordinal Polytomous Responses

			Link function		
α_j	ϑ_{jx}	Parameters of the model	Global	Local	Continuation
free	free	$\alpha_j(\theta - \vartheta_{jx})$	GRM	GPCM	2P-SM
free	constrained	$\alpha_j[\theta - (\vartheta_j + \tau_x)]$	RS-GRM	RS-GPCM	2P-RS-SM
constrained	free	$\theta - \vartheta_{jx}$	1P-GRM	PCM	SRM
constrained	constrained	$\theta - (\vartheta_j + \tau_x)$	1P-RS-GRM	RSM	SRSM

Source: see [Bartolucci et al. 2016b, p. 127].

[3] In both cases it is assumed that all response categories share the same α_j ($j = 1, \ldots, m$) instead of a category-specific discrimination parameter.

[4] This formulation is known as *rating scale parameterization* and only makes sense when all items have the same number of response categories. τ_{jx} is known as the threshold or cut-off point between categories whose interpretation depends on the specific model.

The RS-GRM abbreviation indicates the rating scale version of the GRM introduced by Muraki (1990), RS-GPCM and 2P-RS-SM are the rating scale versions of GPCM and 2P-SM[5] respectively [Muraki 1997]; 1P-GRM and 1P-RS-GRM [Van der Ark 2001] are the equally discriminating versions of GRM and RS-GRM, respectively. Finally, SRM and SRSM denote the Sequential Rasch Model and the Sequential Rating Scale Model of Tutz [1990].

As an illustration the selected models are considered below. The choice of the global logit link function based on the probability that an item response is observed in category *x* or higher and the least restrictive parameterization for the item parameters the Semejima's [1969] GRM model results as follows:

$$p_{jx}(\theta) = \frac{e^{\alpha_j(\theta - \vartheta_{jx})}}{1+e^{\alpha_j(\theta - \vartheta_{jx})}} = \frac{1}{1+e^{-\alpha_j(\theta - \vartheta_{jx})'}}. \tag{7}$$

According to the generalization given by equation (2), GRM can be formulated as follows:

$$g_x(\mathbf{p}_j(\theta)) = \log \frac{p(X_j \geq x \mid \theta)}{p(X_j < x \mid \theta)} = \alpha_j(\theta - \vartheta_{jx}),\ j=1,\ldots,m,\ x=1,\ldots,l_j-1. \tag{8}$$

By combining the local logit link and the most restrictive parameterization for the item parameters the RSM [Andersen 1977; Andrich 1978] model[6] is obtained:

$$p_{jx}(\theta) = \frac{e^{\sum_{h=1}^{x}\theta-(\vartheta_j+\tau_h)}}{\sum_{r=0}^{l-1} e^{\sum_{h=1}^{r}\theta-(\vartheta_j+\tau_h)}} = \frac{e^{x\theta-\sum_{h=1}^{x}(\vartheta_j+\tau_h)}}{1+\sum_{r=1}^{l-1} e^{r\theta-\sum_{h=1}^{r}(\vartheta_j+\tau_h)'}}. \tag{9}$$

The RSM model can be generally formulated as:

$$g_x(\mathbf{p}_j(\theta)) = \log \frac{p(X_j = x \mid \theta)}{p(X_j = x-1 \mid \theta)} = \theta - (\vartheta_j + \tau_x),$$

$$j=1,\ldots,m,\ x=0,\ldots,l_j-1. \tag{10}$$

The PCM model [Masters 1982] for items with two or more ordered response categories is defined as:

---

5 The Two-Parameter Sequential Model [Hemker et al. 2001; Van der Ark 2001] obtained as a special case of the acceleration model of Semejima [1995].

6 The RSM model is suitable when $l_j = l$.

$$p_{jx}(\theta)=\frac{e^{\sum_{h=1}^{x}(\theta-\vartheta_{jh})}}{\sum_{r=0}^{l_j-1}e^{\sum_{h=1}^{r}(\theta-\vartheta_{jh})}}=\frac{e^{x\theta-\sum_{h=1}^{x}\vartheta_{jh}}}{1+\sum_{r=1}^{l_j-1}e^{r\theta-\sum_{h=1}^{r}\vartheta_{jh}}}. \tag{11}$$

According to the generalization given by equation (2), this model can be formulated as follows:

$$g_x(\mathbf{p}_j(\theta))=\log\frac{p(X_j=x\,|\,\theta)}{p(X_j=x-1\,|\,\theta)}=\theta-\vartheta_{jx},\; j=1,\ldots,m,\; x=0,\ldots,l_j-1. \tag{12}$$

3. LC-IRT models for ordinal polytomous responses

A crucial assumption characterizing the multidimensional LC- IRT models concerns the discreteness of the random variable Θ, with support points $\xi_1,\ldots,\xi_u$ and weights $\pi_1,\ldots,\pi_u$. Each weight π_s ($s=1,\ldots,u$) represents the probability that a subject belongs to class *s*:

$$\pi_s = p(\Theta=\xi_s), \tag{13}$$

assuming that $\sum_{s=1}^{u}\pi_s=1$, $\pi_s\geq 0$, $s=1,\ldots,u.$

For each subject *i*, the manifest distribution of the response vector $\mathbf{X}=X_1,\ldots,X_m$ is given by

$$p(\mathbf{x})=p(\mathbf{X}=\mathbf{x})=\sum_{s=1}^{u}p(\mathbf{x}\,|\,\xi_s)\pi_s. \tag{14}$$

According to the assumption of local independence [Hambleton and Swaminathan 1985]:

$$p(\mathbf{x}\,|\,\xi_s)=\prod_{j=1}^{m}p_j(\xi_s)^{x_{ij}}[1-p_j(\xi_s)]^{1-x_{ij}}. \tag{15}$$

The conditional probabilities depend on the nature of the response variable and each of the item parameterizations presented in Table 1 may be obtained in the latent class IRT approach. The most parsimonious model is obtained by constraining all the discriminating parameters to equal one other, that is $\alpha_j=1$ for all $j=1,\ldots,m.$ In this way a Rasch type model [Rasch 1960] is specified.

As usual in the LC model, individuals do not differ within the latent classes, as the same ability level $\boldsymbol{\xi}_s$ is assumed for all individuals in class *s*. Moreover, the item parameters are supposed to be constant across classes.

4. Estimation methods

The most popular estimation methods for IRT models are: joint (unconditional) maximum likelihood (JML), conditional maximum likelihood (CML), and the marginal maximum likelihood (MML) method. In the empirical part of this article we use the `ltm` and `MultiLCIRT` packages of R based on marginal maximum likelihood estimation (MML).

In the MML method the marginal likelihood corresponding to the manifest probability $p(\mathbf{x})$ is maximized. The distribution that is assumed on the latent trait may be continuous or discrete. Most often it is assumed that Θ is normally distributed, and then the marginal likelihood is obtained by integrating the probability of the item response patterns over the ability distribution as in:

$$p(\mathbf{x}) = \int p(\mathbf{x} \mid \theta) f(\theta) d\theta, \tag{16}$$

(see [Bock and Lieberman 1970; Bock and Aitkin 1981]). Then the random ability parameters are removed and the item parameters estimates are obtained by maximizing the resulting marginal log-likelihood.

The MML method based on the assumption that the latent trait has a discrete distribution with support points and weights which need to be estimated together with the item parameters, is denoted as MML-LC. In such an approach, the marginal likelihood is obtained by summing up the probability of the observed item response patterns over the ability distribution, as in (14).

The estimates of the item parameters obtained from the MML-LC method are equal to those obtained from the CML method and are then consistent (see [Formann 1992; 2007]). Moreover, the maximization of the log-likelihood based on (4) is easier to perform than the maximization of that based on (2) as the problem of solving the integral involved in the MML approach is skipped. Typically, the MML-LC estimates are obtained through the EM algorithm [Bartolucci 2007; Dempster et al. 1977].

In the MML-LC method, a crucial point concerns the choice of the number of support points of the latent distribution or, equivalently, latent classes (s). This choice is very important in applications where s cannot be a priori defined. The most relevant are the Bayesian Information Criterion (BIC) [Schwarz 1978] and the Akaike Information Criterion (AIC) [Akaike 1974]. In these criteria, a term to the likelihood penalizing the complexity of the model is added, so that it may be minimized for more parsimonious parameterizations and smaller numbers of groups than the log-likelihood. Accordingly, the smaller value of the information criteria, the stronger evidence of the model. The BIC can be used to compare models with differing parameterizations, differing number of components, or both.

5. Empirical analysis

In this section we provide a comparison of the estimation results for polytomous IRT and LC-IRT models (assuming the continuousness and discreteness of the latent trait).

The analyses presented below are based on $n = 756$ currently working[7] interviewers[8] who participated in the International Social Survey Programme (ISSP) in 2015. The data collected on these interviewees included responses to questionnaire items about job satisfaction (the public data set, available at *www.diagnoza.com*, see also [*Diagnoza społeczna…* 2015]).

The International Social Survey Programme is a continuous annual programme of cross-national collaboration on surveys covering topics important for social science research nationally representative. We have analyzed the section about the job characteristics and the social dimension of the survey conducted by the Institute for Social Studies (ISS), University of Warsaw.

All computations and graphics in this paper have been done in `ltm` [Rizopolous 2015] and `MultiLCIRT` [Bartoulucci et al. 2016a] packages of R[9].

We have considered the items concerning the different aspects of job satisfaction. The original questions concern the assessment of the primary job:

X_1 (*HSW12_1*) – My job is secure,

X_2 (*HSW12_2*) – My income is high,

X_3 (*HSW12_3*) – My opportunities for advancement are high,

X_4 (*HSW12_4*) – My job is interesting,

X_5 (*HSW12_5*) – I can work independently,

X_6 (*HSW12_6*) – In my job I can help other people,

X_7 (*HSW12_7*) – My job is useful to society,

X_8 (*HSW12_8*) – In my job I have personal contact with other people.

In all of those questions respondents could choose one of the following response options: strongly agree (1), agree(2), neither agree nor disagree (3), disagree (4), strongly disagree (5). However, in order to have a clearer interpretation of the results, the response categories were arranged in increasing order (from absolutely unsatisfied to absolutely satisfied).

The first part of our analysis concerns the models based on the GRM and PCM formulations under the assumption of the normality of the latent trait. The second

[7] Currently working refers to both self-employed, employees, and include persons on leave if they are in an employment relationship.

[8] We dropped records with at least one missing responses and “can’t choose” category.

[9] The first package is suitable to estimate IRT models under the assumption of normality of the latent trait, whereas the second package is used where this trait is represented by a latent trait with a discrete distribution.

part presents the results of the different polytomous LC-IRT models (under the discreteness of the latent trait).

Initially we fitted the Graded Response Model (GRM) in the standard formulation with free discrimination parameters. Then we constrained the discriminant indices to be equal across items (1P-GRM). In the next step the partial credit model (PCM) and the generalization of PCM, i.e. GPCM, with different discrimination parameters were compared.

A final point in the classical approach (continuous latent trait) concerned the comparison between the graded response formulation and the partial credit formulation for the data at issue. We compared the models on the basis on BIC and AIC values as well as the likelihood ratio (LR) test to compare nested models. The results for information criteria are given in Table2 (see models 1-4).

Table 2. Log-likelihood, AIC and BIC results for different polytomous IRT models in discrete and continuous approach

Lp	Model	LL	AIC	BIC
1	GRM	–7497.186	15 074.37	15 259.49
2	1P-GRM	–7539.735	15 145.47	15 298.2
3	GPCM	–7539.727	15 159.45	15 344.58
4	PCM	–7584.5	15 235	15 387.73
5	LC-GRM	–7486.97	15 063.94	15 272.2
6	LC-RS-GRM	–7582.82	15 213.64	15 324.71
7	LC-1P-GRM	–7527.25	15 130.49	15 306.36
8	LC-1P-RS-GRM	–7612.77	15 259.54	15 338.22
9	LC-GPCM	–7526.36	15 142.71	15 350.97
10	LC-RS-GPCM	–7639.02	15 326.04	15 437.11
11	LC-PCM	–7570.83	15 217.65	15 393.52
12	LC-RSM	–7654.55	15 343.1	15 421.78

Source: own calculations in R.

In conclusion, we could state that GRM is the most suitable among the models considered for the data at issue, at least when the latent trait is assumed to be normally distributed. The estimates of the discriminant index (α_j), for each item, together with the estimates of the four threshold parameters (τ_x) under the selected model (in continuous approach) are given in the first part of Table 3.

On the basis of the results presented in Table 3 we conclude that X_4 (`interesting job`) and X_8 (`personal contact with other people`) have the highest and the lowest discriminating power, respectively.

Table 3. The item parameter estimates for GRM and LC-GRM models

	GRM model					LC-GRM model				
Item	τ_1	τ_2	τ_3	τ_4	α_j	τ_1	τ_2	τ_3	τ_4	α_j
1	–3.159	–0.819	0.417	2.950	1.076	–3.368	–0.856	0.438	3.125	0.991
2	**–1.285**	**0.651**	1.925	3.173	1.570	**–1.294**	**0.675**	1.972	3.376	1.468
3	–1.410	**0.650**	**1.995**	**3.428**	1.421	–1.454	**0.677**	**2.072**	**3.634**	1.303
4	–1.923	–0.932	0.204	*1.989*	**2.133**	–1.914	–0.912	0.209	*1.953*	**2.042**
5	–2.307	–0.256	0.593	2.620	1.297	–2.363	–0.270	0.594	2.698	1.239
6	–2.276	–0.667	0.278	2.569	1.404	–2.258	–0.654	0.277	2.566	1.409
7	–3.750	–1.965	–0.430	2.525	1.024	–3.863	–2.018	–0.441	2.604	0.982
8	*–5.374*	*–3.062*	*–1.886*	2.123	*0.760*	*–5.410*	*–3.086*	*–1.902*	2.124	*0.755*

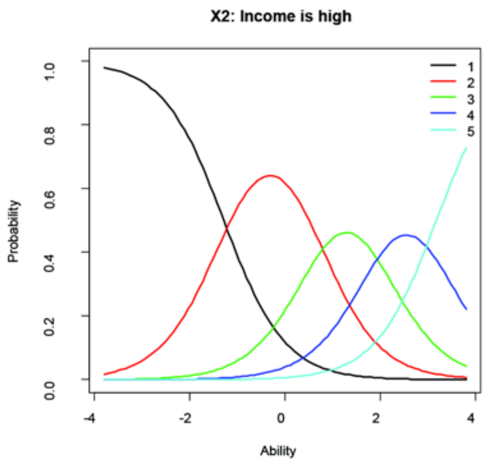
Source: own calculations in R.

Regarding the threshold parameters, we noted that for almost all response categories, the estimate for the second and third items are the highest. This means that items X_2 (*income*) and X_3 (`opportunities for advancement`) are considered to have the highest difficulty level. In turn, the eighth item (`personal contact with other people`) is considered to have the lowest difficulty level (meaning that that this is the aspect of the highest satisfaction of the interviewees).

This conclusion can be confirmed by the probability plot presented in Figure 1. In particular, we can compare the plot which refers to the eighth item with the one which refers to the second one. We observe that the probability of the fourth and fifth response categories (with the highest satisfaction level) are much higher for the eighth than for the second item (i.e. for respondents with the ability level of 1).

As far as the discrete approach is concerned (analysis under discrete distribution for the ability), we followed the consecutive ordered steps [Bacci et al. 2014; Bartolucci et al. 2014; Genge 2016]. The optimal number of clusters was chosen using information criteria for the basic model. On the basis of the BIC and AIC values, we decided to choose four latent classes as the most suitable number for this data. As regards the second step and the choice of the best logit link function, we compared the graded response type model and a partial credit type model[10], assuming $s = 4$, free item discriminating and difficulties parameters and a general multidimensional structure for the data see [Bartolucci et al. 2014].

[10] The continuation ratio logit link function is not suitable in the context of the empirical study because the item response process does not consist of the sequence of successive steps [Bacci et al. 2014].



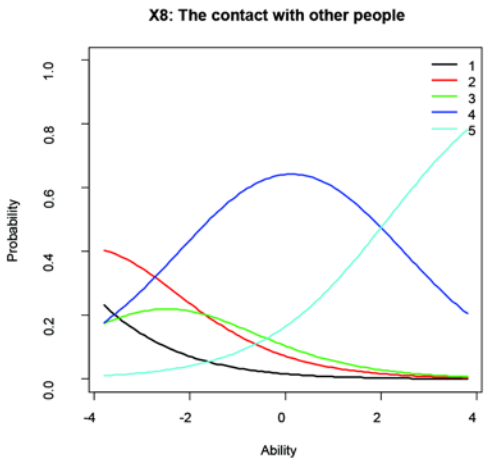


Fig. 1. Item response characteristic curves for the second and eighth items estimated under GRM.

Source: own calculations in R.

Table 4. Graded response and partial credit type models with $s = 4$

Logit	LL	npar	BIC	AIC
Global	–7486.968	45	15 272.198	15 063.936
Local	–7526.356	45	15 350.974	15 142.712

Source: own calculations in R.

Because the global link has to be preferred to a local logit link function (the smaller BIC and AIC values, see Table 4) we fitted different types of the LC graded response type models with a free and constraint discriminating index as well as free and constrained threshold difficulty parameters, for each item. This implies a comparison among four models in accordance with the classification adopted in Table 1 (see LC-GRM, LC-RS-GRM, LC-1P-GRM, LC-1P-RS-GRM models in Table 2). For the sake of completeness the results for the local logit link function are also included in Table 2 (models 9-12).

The results presented in Table 2 show that the LC-GRM model has to be preferred among all of the LC-IRT models considered, that is the LC graded response model with free discriminant and difficulty parameters. Because the compared models are nested, the parametrization is selected on the basis of an LR test and BIC, AIC criteria as well (see models 5-12, Table 2).

In the next step of the analysis it was in our interest to analyze the distribution of the latent variable based on four ordered latent classes (Table 5) as well as the estimates of the item parameters (the second part of Table 3). The estimated support points and probabilities, under the selected model are shown in Table 5.

Table 5. The estimated support points and prior probabilities for the LC-GRM model

Parameter	Cluster 1	Cluster 2	Cluster 3	Cluster 4
ξ_s	–2.711	–0.598	0.737	3.167
π_s	0.040	0.502	0.428	0.029

Source: own calculations in R.

We observe that most of the subjects belong to class 2 (50.2%) and 3 (42.8%), which is characterized by an intermediate level of satisfaction; 4% of subjects are in class 1 and they tend to have the lowest level of satisfaction, whereas the remaining 2.9% of individuals belong to class 4, which is characterized by the highest level of job satisfaction.

Moreover for all of the items, the probabilities of answering with a high response category (denoting a high level of satisfaction) increase from class 1 to 4, whereas the probabilities of answering with a low response category (denoting a low level of satisfaction) decrease from class 1 to class 4 (see Figure 2).

The estimates of the item difficulty levels and discrimination parameters for the LC-GRM model presented in the second part of Table3 confirmed that the most difficult is the second and third item, which concern income and opportunities for advancement, the easiest item is the eighth one (`personal contact with other people`). As far as the discrimination parameters are concerned (`interesting job`) and (`personal contact with other people`) were considered to have the highest and the lowest discriminating power, respectively.

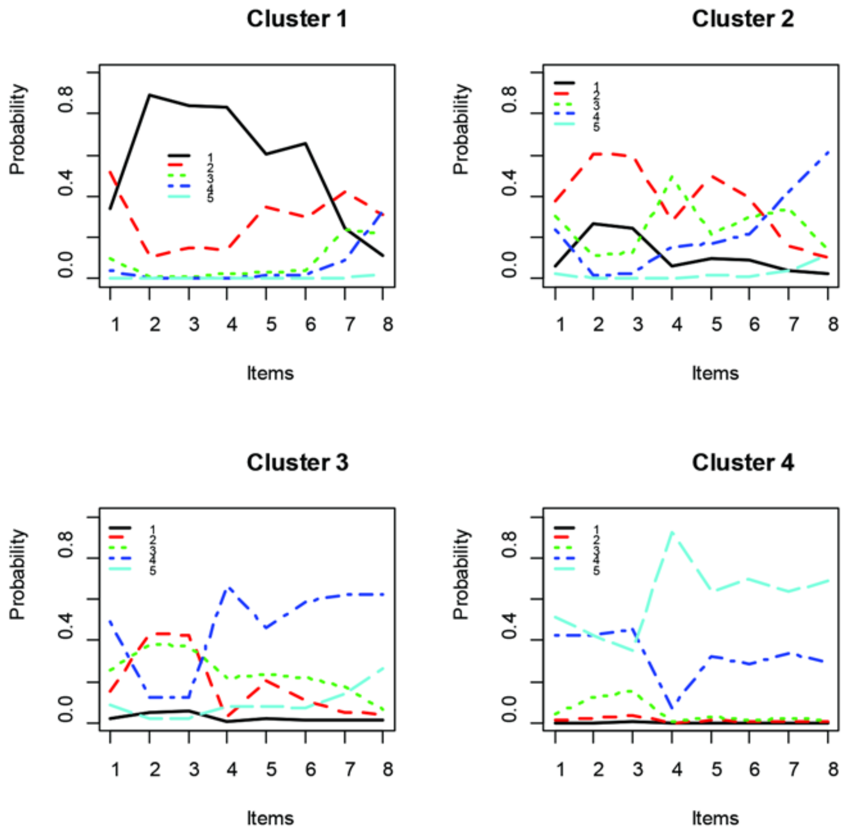


Fig. 2. The conditional probabilities for the fourth-class LC-GRM model

Source: own calculations in R.

6. Conclusion

We presented the analysis of multivariate polytomous data using latent variable models under the Item Response Theory assuming continuous and discreteness of the latent trait.

We fitted and compared the results of the 12 different IRT and LC-IRT models e.g. the GRM, 1P-GRM, PCM, GPCM, LC-GRM, LC-PCM models assuming discrete and continuous distribution of the latent trait. The analyzed item responses (concerning job satisfaction in Poland) can be explained by a graded response type model with items having the same discriminating power and different distances between consecutive response categories as well as in the continuous and discrete approach.

The comparison showed a certain agreement between the two sets of parameter estimates obtained under the different assumptions about the latent distributions. We observed that the estimated parameters are rather smaller under the MML method based on the normal distribution for the ability. However, the comparison among the items in terms of difficulty and discriminating power generally led to the same results under both approaches.

Finally, it should be noted that there is no dominating approach to the latent trait (as far as the IRT models for ordinal polytomous item responses are concerned). Some are good only in some specific simulations or examples. However it is worth noting that this extension of the traditional IRT (i.e. LC-IRT) models, by introducing assumptions of discreteness and also multidimensionality of latent trait (see [Bartolucci et al. 2014; Genge 2016]) may be especially useful in socio-economic data analyses where the normality and unidimensional assumptions of the latent trait (explicitly introduced) are very often restrictive to fulfill.

Bibliography

Agresti A., 2002, *Categorical Data Analysis*, John Wiley&Sons, Hoboken, NJ.

Akaike H., 1974, *A new look at statistical model identification*, IEEE Transactions on Automatic Control, 19, pp. 716-723.

Andersen E.G., 1977, *Sufficient statistics and latent trait models*, Psychometrika, 42, pp. 69-81.

Andrich D., 1978, *A rating formulation for ordered response categories*, Psychometrika, 43, pp. 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, pp. 787-800.

Bartolucci F., 2007, *A class of multidimensional IRT models for testing unidimensionality and clustering items*, Psychometrika, 72, pp. 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, pp. 971-985

Bartolucci F., Bacci S., Gnaldi M., 2016a, *MultiLCIRT: Multidimensional latent class Item Response Theory models. R package version 2.10*, http://CRAN.R-project.org/package=MultiLCIRT.

Bartolucci F., Bacci S., Gnaldi M., 2016b, *Statistical Analysis of Questionnaires. A Unified Approach Based on R and Stata*, A Chapman&Hall/ CRC.

Birnbaum A., 1968, *Some Latent Trait Models and Their Use in Inferring an Examinee's Ability*, [in:] F.M. Lord, M.R. Novick, *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, MA, pp. 395-479.

Bock R.D., Aitkin M., 1981, *Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm*, Psychometrika, 46, pp. 443-459.

Bock R.D., Lieberman M., 1970, *Fitting a response model for n dichotomously scored items*, Psychometrika, 35, pp. 179-197.

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, pp. 1-38.

Diagnoza społeczna 2015. Warunki i jakość życia Polaków (raport), J. Czapiński, T. Panek (eds.), Warszawa, Rada Monitoringu Społecznego (5.01.2017).

Formann A.K.,1992, *Linear logistic latent class analysis for polytomous data*, Journal of the American Statistical Association, 87:476-486. 22

Formann A.K., 2007, *(Almost) Equivalence between Conditional and Mixture Maximum Likelihood Estimates for Some Models of the Rasch Type*, [in:] M. von Davier, C. Carstensen, *Multivariate and Mixture Distribution Rasch Models*, Springer-Verlag: New York, pp. 177-189.

Genge E., 2016, *Teoria reakcji na pozycję w podejściu modelowym w taksonomii* [*Item response theory in model-based clustering*], Ekonometria, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, 1 (51), pp. 9-19.

Hambleton R.K., Swaminathan H., 1985, *Item Response Theory: Principles and Applications*, Kluwer Nijhoff, Boston.

Hemker B.T., van der Ark L.A., Sijtsma K., 2001, *On measurement properties of continuation ration models*, Psychometrika, 66, 4 pp. 87-506.

Masters G., 1982, *A Rasch model for partial credit scoring*, Psychometrika, 47, pp. 149-174.

Muraki E., 1992, *A generalized partial credit model: application of an EM algorithm*, Applied Psychological Measurement, 16, pp. 159-176.

Muraki E., 1997, *A Generalized Partial Credit Model*, [in:] W. Van der Linden, R.K. Hambleton, *Handbook of Modern Item Response Theory*, Springer, pp. 153-164.

Nering M.L., Ostini R., 2010, *Handbook of Polytomous Item Response Theory Models*, Taylor and Francis, New York.

Rasch G., 1960, *Probabilistic Models for Some Intelligence and Attainment Tests*, Danish Institute for Educational Research, Copenhagen.

Rizopoulos D., 2015, *Latent trait models under IRT*, https://cran.r-project.org/web/packages/ltm/ltm.pdf.

Samejima F., 1969, *Estimation of ability using a response pattern of graded scores*, Psychometrika Monograph, 17.

Semejima F., 1995, *Acceleration model in the heterogeneous case of the general graded response model*, Psychometrika, 60, pp. 549-572.

Schwarz G., 1978, *Estimating the dimension of a model*, Annals of Statistics, 6, pp. 461-464.

Thissen D, Wainer H., 2001, *Test Scoring*, Lawrence Erlbaum Associates, Mahwah.

Tutz G., 1990, *Sequential item response models with an ordered response*, British Journal of Mathematical and Statistical Psychology, 43, pp. 39-55.

Van der Ark L.A., 2001, *Relationships and properties of polytomous Item Response Theory models*, Applied Psychological Measurement, 25, pp. 273-282.

Van der Linden W., Hambleton R.K., 1997, *Handbook of modern item response theory*, Springer.

Verhelst N.D., Glas C.A.W., Vries de H.H., 2005, *IRT Models for Rating Scale Data*, [in:] B. Everitt, D. Howell, *Encyclopedia of Statistics in Behavioral Science*, John Wiley&Sons, West Sussex, UK, pp. 995-1005.

Wright B., Masters G., 1982, *Rating Scale Analysis*, Mesa Press, Boston.