DIDACTICS

OF

MATHEMATICS

10(14)

TABLE OF CONTENTS

THE ITEM INFORMATION FUNCTION
IN ONE AND TWO-PARAMETER LOGISTIC MODELS –
A COMPARISON AND USE IN THE ANALYSIS
OF THE RESULTS OF SCHOOL TESTS

Anna Zięba

Abstract. This paper presents a very important problem in the theory of IRT, namely the Item Information Function. It presents the similarities and differences in the approach to this problem for the one-parameter Rasch model and the two-parameter Birnbaum model. All kinds of relationships are presented in graphs. Additionally, special attention is paid to the interpretation of these graphs and the results.

1. Introduction

A test is a didactic tool frequently used in evaluating scholastic achievements. A test provides the possibility to check the student's knowledge and skills, whose measurement is the primary objective of educational research. However the problem is that such characteristic features as mathematical abilities, verbal skills, resistance to stress, intelligence, dissatisfaction, various opinions on a specific subject, etc. are not directly observable. These features are called latent traits and they can be measured only indirectly, e.g. with the use of specially prepared questionnaires where the responses are closely connected with the specific trait which is being studied. The measurement of the student skills and knowledge usually has two objectives. Very often, tests are used to evaluate the level of knowledge of each of the students. The other objective is to build a bank of questions −

Anna Zięba
 Department of Economic Forecasts and Analyses, Wrocław University of Economics, Komandorska Street 118/120, 53-345 Wrocław, Poland.
 E-mail: anna.zieba@ue.wroc.pl

a set of items used in taking the measurement as diligently and precisely as possible. Consequently, the so-called latent trait models have been developed which are used to estimate the values of parameters connected with the human personality. These models provide a different kind of information which in turn, when appropriately used, can improve the development of tests, items or exams. It seems obvious that most information comes from the analysis of the given responses, but it should be considered if a question can provide information. If it can, what does that information regard?

The paper presents an attempt to compare two theoretical approaches regarding statistical tools used in developing tests, items, exams, as well as their use and interpretation of the collected results – the item information function in the one-parameter Rasch model and the two-parameter Birnbaum model (Andersen 1983). Attention is drawn to how to measure the information provided for example. through specific items in the exam sheet, as well as how to use that knowledge to develop better exam sheets. Furthermore, the discussion will be limited to the case where one latent trait is measured and the responses are evaluated with the use of the binary system (zeros and ones).

2. Item information function

2.1. The item information function in the one-parameter logistic model

The notion of information plays an important role in the Item Response Theory (IRT) (Andrich 1978) as it is possible to use it to evaluate precisely how individual items included in the test measure the level of a given latent trait (value of parameter θ_i). That latent trait can include for example the level of student knowledge, intelligence, ability, satisfaction, stress, etc. For example, in educational tests, item parameters represent the difficulty of items while person parameters represent the ability level of the people who are being assessed. The higher the student's ability relative to the difficulty of an item (parameter α_j describes the degree of item difficulty, the level of influence of the item on the respondent), the higher the probability of the correct response to that item. When a person's location on the latent trait is equal to the difficulty of the item, there is by definition a 0.5 probability of a correct response in the Rasch model. The information about a specific value θ_i depends on the number and the qualities of the questions used to evaluate that parameter (Hambleton, Swaminathan, Rogers 1991). In order to be able to use any information about the level of the measured trait, first

the data should be collected to evaluate parameter α_j and θ_i. In the case of IRT models, each item has its specific information function which provides information about how precisely, and what level of the analyzed latent trait, a given item measures. The information function of question j is defined by the following formula (Hambleton, van der Linden 1997):

$$I_j\left(\theta_i\right) = \frac{\left[P_j^{'}\left(\theta_i\right)\right]^2}{P_j\left(\theta_i\right) \cdot \left(1 - P_j\left(\theta_i\right)\right)} \tag{1}$$

where $P_j\left(\theta_i\right)$ – probability of the correct response to question j.

In the case of the one-parameter Rasch model, the following is the information function for question j:

$$I_j\left(\theta\right) = \frac{e^{\theta - \alpha_j}}{(1 + e^{\theta - \alpha_j})^2} \, . \tag{2}$$

The scope in which a given question provides information about parameter θ_i does not depend on the other questions (Brzeziński 2005). As a result of that property it is possible to evaluate how each of the questions contributes to the evaluation of the studied latent trait. It is possible to estimate the level of information provided by a given question in the whole scope of the studied latent trait $\left(-\infty, \infty\right)$.

Function $I_j(\theta)$ is presented graphically as the relationship between the information value of the question and the level of the latent trait. In order to present that relationship, five selected questions with assumed values of parameters α_j were analyzed. These values are as follows: for question 1 it was assumed that its difficulty was estimated at the level of $\alpha_1 = 0.57$, for question 2: $\alpha_2 = 0.31$, for question 3: $\alpha_3 = 1.28$,   for question 4: $\alpha_4 = 0.54$, and for question 5: $\alpha_5 = -0.15$, (the higher value of the alpha parameters – the more difficult the question). It is important to write here that there is a special software for calculations and estimation of the alpha and beta parameters in practice (based on the Rash and Birnbaum models), for example: WinSteps, the ltm package in R, RUMM, Param, Rasch. Figure 1 presents the graphs of the information functions for these questions.
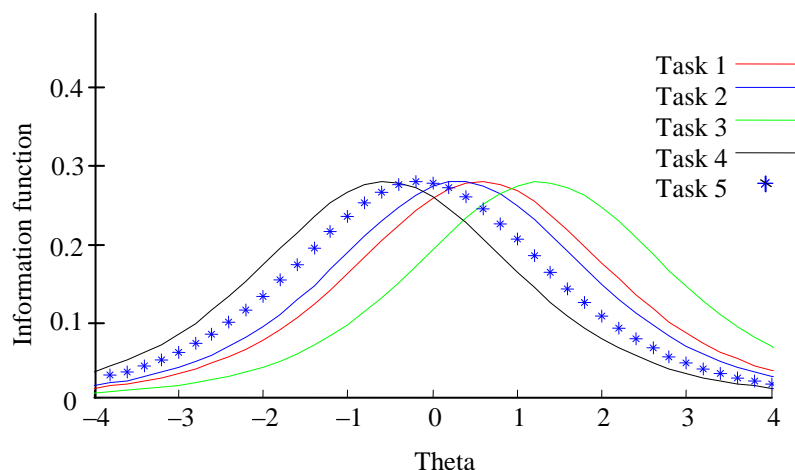
Fig. 1. Graphs of the information functions for five selected questions

Source: own study (assumed data).

The maximum of the information function of question j in the one-parameter Rasch model corresponds to the value of parameter α_j and its value is 0.25. Furthermore, the graph of the information function is symmetrical to parameter α_j. The information functions show what level of the latent trait a given question measures most reliably (their maximum values are reached in point θ_i for which the question provides most information.)

While an information function can be obtained for each item in the test, this is rarely done. The amount of information yielded by each item is rather small (for example: the student's ability – parameter θ – is not estimated with a single item). It is highly significant that due to the local independence of the questions the information functions demonstrate the features of additiveness. Consequently, specific pieces of information about the test items add up and provide the general information about the whole test. The test information is obtained by summing up the item information at a given ability level. The information function regarding the whole test is as follows (compare (Hambleton, van der Linden 1997)):

$$I(\theta) = \sum_{j=1}^{k} I_j(\theta). \tag{3}$$

Below is a graph of the information function for the test built on the basis of the five selected questions analyzed earlier.
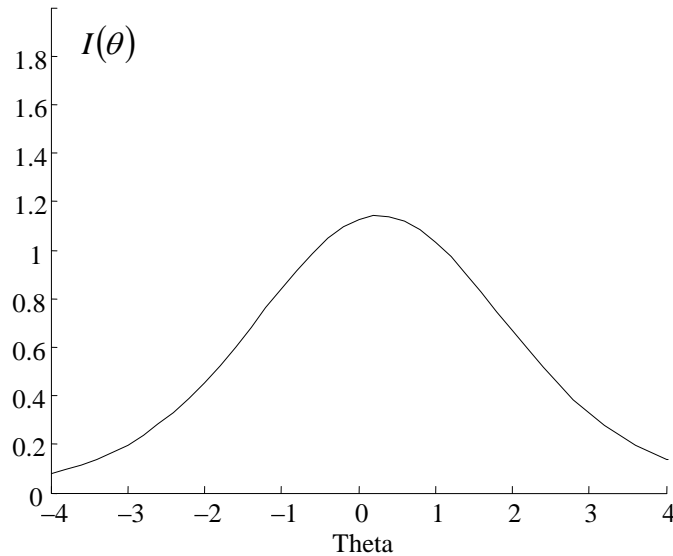
Fig. 2. Graph of the test information function

Source: own study.

That function reaches its maximum at the point close to $\theta_i = 0.25$ and then decreases on both sides so the test built on the basis of five selected questions would measure the latent trait most reliably at the level of $\theta_i = 0.25$ (most information would be provided about average students as the degree of concentration of the latent trait for them in point $\theta_i = 0$ on the scale). Such a kind of measurement tool is useful when we want to divide the respondents into two groups.

2.2. The item information function in the two-parameter logistic model

The situation looks different in the two-parameter logistic model (Andersen 1983). The assumption of the previous model was that the questions differ only in respect of their difficulty. In fact, however, this is not necessarily true. In the two-parameter logistic model it is assumed that two parameters are connected with the question: parameter α_j describing the difficulty of the item (question) and additional parameter β_j – item discrimination. Parameter β (slope of the curve) demonstrates the degree to which the question helps to differentiate between the respondents with a higher level of the analyzed trait and those where that level is lower. That parame-

ter demonstrates how closely the item is connected with the whole test. The correctly selected items should be closely connected with the whole test. The higher the value of that parameter, the greater the item discrimination (the easier it is to select the respondents with a high and those with a low concentration of a given trait.) It is not the rule either that the most difficult item demonstrates the highest item discrimination level.

The information function is defined as follows:

$$I_j(\theta) = \frac{\beta_j \cdot e^{(\theta - \alpha_j)\beta_j}}{\left(1 + e^{(\theta - \alpha_j)\beta_j}\right)^2} \ .$$

(4)

The figure shows the graphs of the information functions for five selected questions in the two-parameter model. Table 1 presents the assumed values of parameter α_j and β_j.

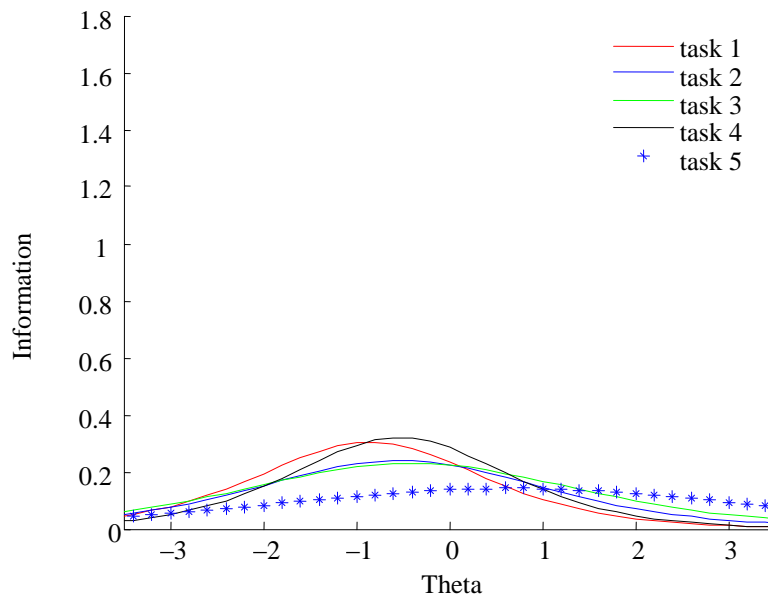

Fig. 3. The information functions in the two-parameter model

Source: own study (assumed data).

Table 1. Examples of estimated parameter α_j and β_j

Question	$\hat{\alpha}_j$	$\hat{\beta}_j$
1	−0.86	1.21
2	−0.53	0.96
3	−0.41	0.82
4	−0.54	1.28
5	0.66	0.57

Source: assumed data.

When analyzing the graph, it is visible that question 4 provides the greatest amount of information about the values of the latent trait $\theta_i \in (-1.0)$. On the other hand, question 5 provides the smallest amount of information for the same respondents; it is an average item characterized by poor discrimination. It can be concluded that question 4 has a large impact on the measurement of the studied latent trait. Furthermore, in this model, different questions provide different amount of information about how precisely and what level of the analyzed latent trait a specific question measures. In the one-parameter model the information function assumed the same value of 0.25 for all the questions included in the test. This results from the fact that the same item discrimination (the same values of parameters $\beta_j = 1$) is assumed for all the questions in the one-parameter model. The narrower the shape of the graphic presentation of the information function, the narrower the scope of the latent trait which can be precisely measured on the basis of a given question. Consequently, the higher the value of the discrimination parameter, the more precise the information about the skills of the student for a specific range.

The following graph presents the information functions of the five questions analyzed above, however it was assumed that they are equally difficult (the same values of parameters $\alpha_j = 1$). The values of parameters β_j remained unchanged and they are as follows: $\beta_1 = 1.21$; $\beta_2 = 0.96$; $\beta_3 = 0.82$; $\beta_4 = 1.28$; $\beta_5 = 0.57$. The objective of such assumptions was to draw attention to the influence of the parameters on the graphs of the functions.
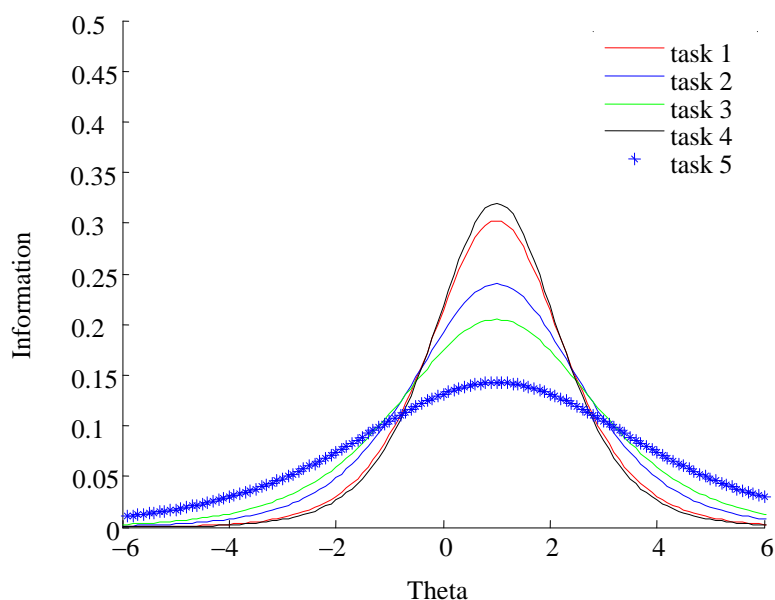
Fig. 4. The information functions in the two-parameter model
for five equally difficult questions

Source: own study (assumed data).

The following graph presents another relationship between the parameters – a fixed value of $\beta_j = 0.57$ was assumed. The values of parameters α_j remained unchanged (Table 1).

Graph 5 is similar to the graphs of the information function for the one-parameter model. The difference is in the amount of information provided by the questions. In the case of the one-parameter logistic model the maximum of the function was 0.25. The change of the value of parameter β_j also changes the maximum value. When $\beta_j = 0.57$ (Fig. 5) the functions reach the maximum which is $0.25 \cdot 0.57 = 0.14$. The maximum value reached by the information functions then increases or decreases depending on whether parameter β_j is respectively higher or lower than 1.
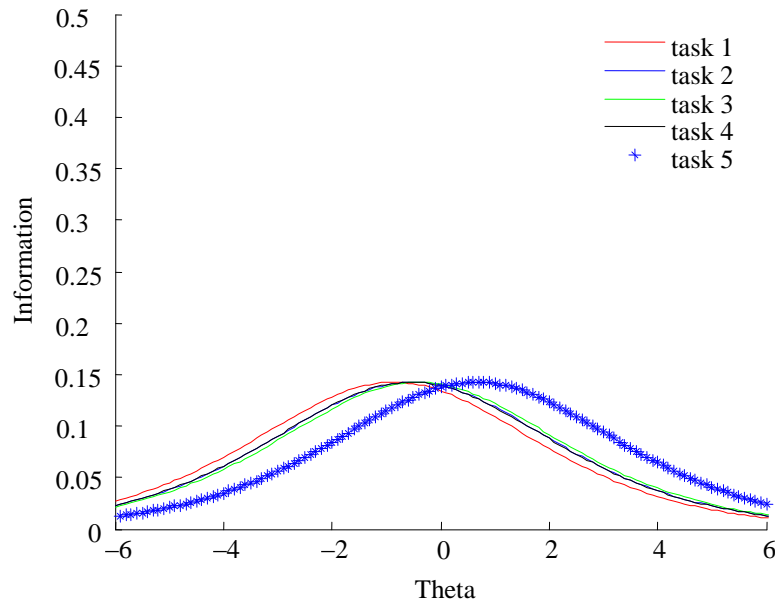
Fig. 5. The information functions in the two-parameter model
for five questions with the same value of parameters $\beta_j = 0.57$

Source: own study (assumed data).

Also in the case of the two-parameter model it is possible to define the information function regarding the whole test, which is the sum of the information functions of all questions included in the examination sheet (compare Andersen (1983), Andrich (1978)):

$$I(\theta) = \sum_{j=1}^{k} I_j(\theta). \tag{5}$$

3. Conclusion

The discussion above can be summed up in the following way: the precision of estimation of the measured latent trait θ with a given question (item) depends on the amount of information which it provides. Furthermore, in the one-parameter logistic model, each question provides most information for the value of the latent trait which precisely corresponds with a given item's difficulty level. The analysis of the graphs leads to the conclusion that the narrower the shape of the graphic presentation of the infor-

mation function, the narrower the scope of the analyzed latent trait which can be precisely measured on the basis of that question. Additionally, in the two-parameter Birnbaum model (assuming that the questions are equally difficult) the greater the item discrimination, the greater the maximum value of the information function.

References

Andersen E.B. (1983). *Latent trait models*. Journal of Econometrics 22. University of Copenhagen. North – Holland Publishing Company. Pp. 215-227 (for Birnbaum and Rasch models).

Andrich D. (1978). *A rating formulation for ordered response categories.* Psychometrika 43. Pp. 561-573.

Brzeziński J. (red.) (2005). *Wprowadzenie do teorii odpowiedzi na pozycje testu.* W: "Trafność i rzetelność testów psychologicznych. Wybór tekstów". Gdańskie Wydawnictwo Psychologiczne. Pp. 213-267.

Hambleton R.K., Swaminathan H., Rogers H.J. (1991). *Fundamentals of Item Response Theory.* Sage Publications. USA.

Hambleton R.K., van der Linden W.J. (1997). *Handbook of Modern Item Response Theory.* Springer-Verlag. New York. Pp. 85-123.