ITEM RESPONSE THEORY (IRT) FOR ASSESSING STUDENT PROFICIENCY

Sunil Kalekar

Assist. Professor, Adhyapak Mahavidyalaya, Aranyeshwar, Pune 9

Abstract

Classical Test theory provides easy approach for assessing students' achievement but it takes in to account the overall performance of the student in a given test. It ignores the ability of student to attempt individual test items. We know that all test items are not of the same difficulty level. Item Response Theory (ITR) helps us to assess students' performance in more methodical way. One can analyze an individual students performance using ITR as well as one can compare performances of individual students belonging to a group. Based on parameters ITR models are proposed. The present paper tries to explain some basic concepts related to ITR. Much more explanation and understanding of this concept is needed before it can be implemented in present education system. ITR may not be the ultimate solution for making assessment full proof and accurate but it can definitely provide sound basis for more scientific and systematic approach for assessing students' proficiency.

Key terms: Classical Test Theory (CTT), Item response Theory, models of IRT – 1PL, 2Pl, 3 PL, Features of ITR

Introduction

Suppose you have given a test of ten questions to your students. Out of these ten questions, two questions are very difficult, two are trivial and other six are equally difficult. Now think that two of your students have answered nine questions correctly. They would get 90 % scores each. But student "A" has attempted a simple question incorrectly and student "B" has attempted a very difficult one incorrectly. Based on this situation how can one identify the student with higher ability? This situation points out a major problem with traditional method of testing. Classical Testing Theory (CTT): It is the easiest and most widely used form of

analysis. The analyses are performed on the test as a whole rather than on the item. CTT is based on true score model:

$X = T + E$ where T is the true score and E is the error.

In CTT we assume that the error:

1.   Is normally distributed.

2.   Uncorrected with true score

3.   Has a mean of zero.

CTT statistics analyze the test difficulty and discrimination with respect to item level statistics whereas the reliability of the test is analyzed with respect to test level statistics. Classical analysis uses a test as a basis and not the test item for analysis. Although the statistics generated are often generalized to similar students taking a similar test whereas they only apply to those students taking that test.

What is Item Response Theory?

Item Response Theory (IRT) was first proposed in the field of psychometric for the purpose of ability assessment. It is widely used in education to evaluate items in tests, questionnaires and other instruments and to score subjects on their abilities, attitudes or other traits. Today all major educational tests such as scholastic Aptitude Test(SAT) and Graduate Record Examination (GRE) are developed by using item response theory, because the methodology can significantly improve measurement accuracy and reliability while providing potentially significant reductions in assessment time and effort. IRT attempts to model student ability using item level performance instead of aggregate test level performance. Instead of assuming that all items contribute equally to our understanding of students' abilities, IRT provides a multi- layered view on the information that each item provides about the student. IRT is a family of models that describes the interaction between examinees and test items. The performance of the student can be predicted in terms of the underlying trait. IRT provides a means for estimating scores for people and characteristics of items. The three basic components of IRT are:

a.   Item response Function: Mathematical function that relates the latent trait to the probability of supporting an item.

b.   Item Information Function: an indication of item quality: an item's ability to differentiate among respondents.

c.   Invariance: position on the latent trait can be estimated by any items with known IRFs and item characteristics are population independent within a linear transformation.

Models of IRT are classified based on the number of parameters considered. They are:

One parameter logistics (1 PL), two parameter logistics (2 PL), three parameter logistics (3 PL).

One parameter logistics (1 PL):

It allows each test item to have its own independent difficulty variable. It models the probability of a correct answer using the logistic function: where 'Q' represents the question of interest, theta is the students' ability and beta is item 'Q's difficulty. This function is also known as the item response function.
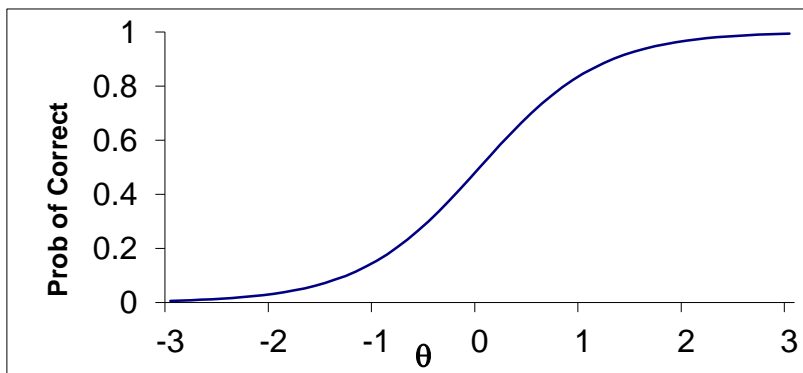

It points out two things:

1.   The probability of a correct answer increases as item difficulty decreases.

2.   For a given item difficulty level, the probability of a correct answer increases as student ability increases.

Two Parameter Logistic (2PL):

Suppose an item in a test is attempted by all the students correctly, then we can say that the particular item has a low discrimination value since it cannot discriminate between students of high or low ability. Ideally a good question will maximally separate students into two groups; those with the ability to answer correctly and those without. 2PL model incorporates this idea by attempting to model each item's level of discrimination between high and low ability students.
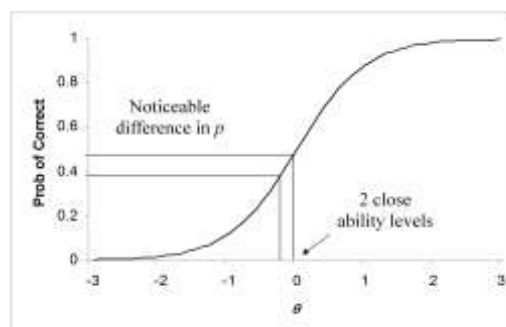
Three parametric Logistic (3PL):

Sometimes students will identify the correct answer by simply guessing. Students can increase their odds of guessing a question correctly by ignoring answers that are obviously wrong. The three parameter logistic (3PL) model uses 'guess ability' in the model.('Chi' represents the item's pseudo guess value. chi is not considered a pure guessing value as students can use some strategy or knowledge to eliminate bad guesses.

Main features of IRT procedures:

The IRT procedure enables you to estimate various IRT models. The following list summarizes some of the main features of IRT:

1.   Fits the following classes of models: Rasch model; one-, two-, three- parameter models; and graded response models

2.   Supports logistic and probity links

3.   Calibrates items that can have different response models

4.   Performs multidimensional exploratory and confirmatory analysis

5.   Performs multiple-group analysis, with fixed values and equality constraints within and between groups

6.   Estimates factor scores by using the maximum likelihood (ML), maximum a posteriori (MAP), or expected a posteriori (EAP) method.

IRT and Teacher Education:

Although IRT is very commonly used theory for assessment in competitive exams like GRE and SAT, it remains a new concept in teacher education. Teacher education has incorporated CCE as a component in educational evaluation and also used CCE for evaluation in teacher education. Many a times it happens that a new approach is used in secondary schools and later incorporated in teacher education. This should not happen with new trends in educational evaluation. Training of teachers should be completed before the implementation of any new approach in secondary schools so that the teachers can successfully implement those approaches correctly. IRT basics must be taught to student teachers so that they can use it in their students' evaluation.

Conclusion:

CTT is well rooted and well understood theory. In present scenario, we are using the CTT in all school based testing procedures but there are certain disadvantages of this theory. When we talk about validity of a test, with present type of test analysis and conclusions that we draw the validity of the test is under question. IRT provides sound basis for test analysis and

test inferences using appropriate mathematical tool. IRT helps to provide real information about students' abilities from test analysis.

Website

http://en.wikipedia.org/wiki/Item_response_theory

http://www.creative-wisdom.com/computer/sas/IRT.pdf

http://info.worldbank.org/etools/docs/library/117765/Item%20Response%20Theory%20-%20F%20Baker.pdf

www.csun.edu/~ata20315/psy427/Topic08_IntroIRT.ppt

https://www.metheval.uni-jena.de/irt/VisualIRT.pdf

http://harvey.psyc.vt.edu/Documents/TCP_IRT98.pdf