MATH MODEL RESEARCH GROUP
DEPARTMENT OF MATHEMATICS
FEDERAL UNIVERSITY OF TECHNOLOGY (FUT), MINNA, NIGERIA

BOOK OF PROCEEDINGS

for the

International Conference on MATHEMATICAL MODELLING, OPTIMIZATION AND ANALYSIS OF DISEASE DYNAMICS (ICMMOADD 2024)

HYBRID (VIRTUAL & PHYSICAL)

Theme:

CONTEMPORARY ISSUES ON THE CONTROLS OF DISEASES EPIDEMICS AND PANDEMIC

DATE: Thursday 22nd February, 2024 **TIME:** 10:00am Prompt

VENUE: Department of Mathematics, Federal University of Technology Minna, Nigeria

PROFESSOR N. I. AKINWANDE FNMS
Convener

ISBN: 978-978-789-930-4

On Unobserved Traits and its Measurements: A Note on Modern Item Response Modelling Approach

O. M. Adetutu

Department of Statistics, Federal University of Technology Minna, Nigeria
*adetutuolayiwola@gmail.com

Abstract

Often, scientists are interested in studying traits and other unobservable attributes of individuals which can only be measured with instruments such as questionnaire, test let, construct, and so on. These instruments are useful in extracting information on these hidden traits. Traditionally, Karl Pearson's classical approach were in used in modelling these traits with its attendant shortcoming ranges from variant estimates to its challenges of validity and reliability. This study present a better and modern item response modelling that remedy shortcomings of classical approach yielding reliable, consistent, and valid estimate through item response functions.

Keywords: unobservable, attributes, traits, reliability, validity

1 Introduction

Unobserved traits are mental and emotional activities that cannot be directly observed or measured, examples are: $(i.)$ psychological constructs such as intelligence, self-esteem, depression, academic proficiency, aptitude achievement, attitude or belief, $(ii.)$ psychological traits and behaviour such as thinking, memory, perception, problem-solving, and decision-making are all hidden traits which can only be estimated through questionnaire, test lets, constructs, schedules, and so on (Embretson *et al.*, 2000).

Two approaches toward measurement of unobserved traits are: classical test theory (CTT), and item response theory (IRT). The former is a body of related psychometric theory that predicts outcomes of psychological testing based on the fact that person's observed or obtained score on a test, construct, and questionnaire is sum of a true score and error score (NCME, 2017). This approach is a sample (test, questionnaire, or schedule) based rather than item base. The model for this approach is positioned in equation (1)

$$X = T_x + E_x: \tag{1}$$

Such that: $E(E_x) = 0$

$$Cov(T_x, E_x) = 0$$

$$Cor((T_x, E_x) = 0$$

Where: X is the obtained score

T_x is a true score

E_x is an error score

The focus of CTT is to understand and improve the reliability and validity of unobserved traits using the entire sample. Furtherance to this, Hambleton *et al.*, (1991) identified various shortcomings of CTT which are: inseparability of respondent's and construct's characteristics, assumptions of equal standard errors for all respondents, construct oriented rather than item oriented, and impossibility of making predictions on how well on individual respondents or even group of respondents may do on a given construct item.

The latter is a modern technique of modelling unobserved traits called latent trait theory, strong true score theory or modern mental test theory. Adetutu and Lawal, (2020) described item response theory modelling as a paradigm for the design, analysis, and scoring of tests, constructs, questionnaire, schedule and other instruments used in measuring hidden traits. This approach makes use of item response functions which make it item base rather than sample base.

Assumptions

Potency of IRT and its assumptions that places IRT over CTT as enumerated by Adetutu and Lawal (2023) are: *(i.)* the number of traits that underlies respondents' behaviours or performances. This could be unidimensional or multi-dimensional; for unidimensional, only one trait is being measured, otherwise, multidimensional. The differences in observing response between respondents would be due to their traits when this assumption hold. *(ii.)* Local independence assumption implies that items in the construct are uncorrelated and hence, response to separate item is mutually independent (Lim, 2020). *(iii.)* A non-linear regression curve relates the probability of response on the item to trait being measured, this is a distinction between different latent models, sometime called trace lines or item response functions, and *(v.)* invariant properties of IRT which guarantee unchanging parameters' estimates even when the characteristics of the respondents change. These assumptions have made it possible to estimate individual, and item properties rather than sample based of CTT.

Properties

Discrimination parameter of modern latent traits theory determines the rate at which the probability of correct response option to an item or construct changes as a function of ability or trait continuums. This is often called slope of item response curves whereby negative values are discarded, hence, it is a monotonic increasing because as ability or trait increases, probability of endorsing correct response category to a construct or test also increases and vice-versa (Baker and Kim, 2004). In addition, *difficulty* parameter describes how easy or hard an item of a construct is being perceived by the respondent in binary IRT, while this is tagged 'propensity of endorsing a response option' in a polytomous IRT (Linden, 2018). However, the third property is *pseudo-guessing* parameter which only apply to educational setting, and this is the impact of chance on observed response. This means a lower trait respondent may have non-zero probability of endorsing an item in a test correctly; even an item that seems to be difficult to endorse.

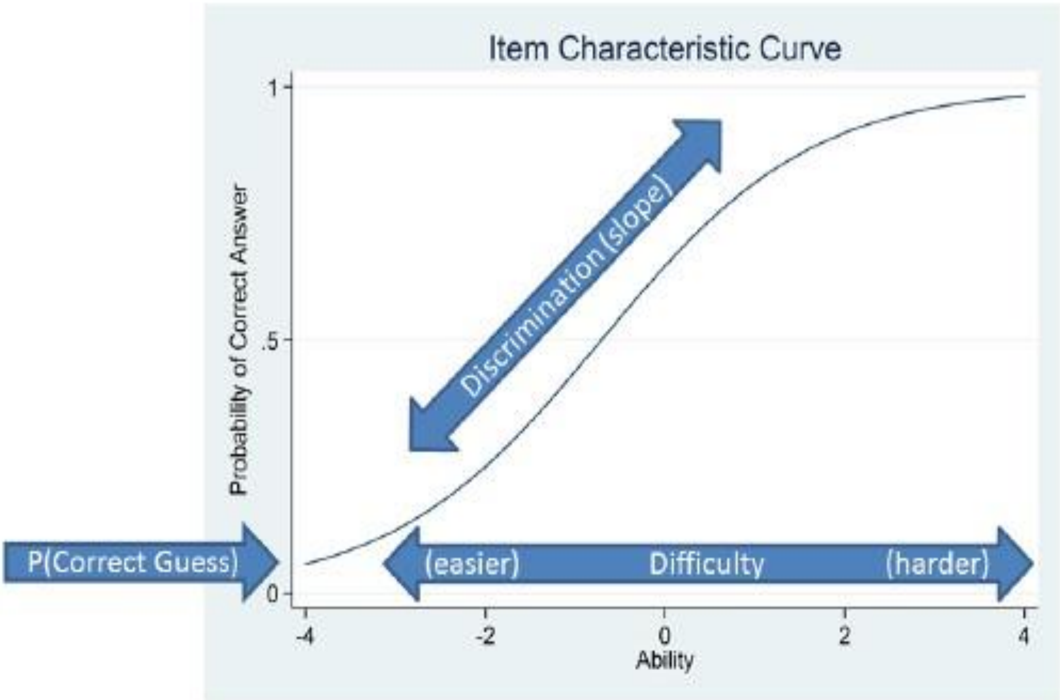

Figure 1: Properties of Unobserved Ability or Traits Model

Vertical axis of the Figure 1 displays probability of correct responses to an item while horizontal axis represent trait or ability continuum; this means, any point on the response curve is a function of respondents' traits with a specific probability of response. The implication of this is that items and individual respondent characteristics can be estimated at any point on the item characteristic curve which is a plus for IRT modelling.

The Aim and Objectives

The study was motivated by the need to efficiently measures and models latent, hidden, not yet manifested, personality characteristic, or trait of respondent devoid of drawbacks associated with traditional approach of CTT. The specific objectives are to describe how to estimate the:

i. Value of an assumed latent trait (discrimination, and difficulty/location parameter)

ii. Attributes of the individual (trait/ability)

iii. Describe variation of the latent parameters among individuals

2. Methodology

2.1 Study Designs and Measures

A well structure construct, questionnaire, multiple choice question, likert-scales or instruments that are scored dichotomous and polytomous would be suitable in extracting information from respondents using suitable sampling techniques on area of interest is applicable here.

2.2 Methods

Responses from the respondents are coded polytomously to form $n \times p$ matrix depending on number of respondents and items on the construct, after which graded response model (GRM) proposed by Samejima (1969) which allowed categorical or ordered item responses function is used.

2.3 Graded Response Model

The likelihood of a respondent with trait θ responds in response category k of item i is:

$$\Pr\left(Y_{ij} \geq k \middle| a_i, b_i, \theta_j\right) = \frac{\exp\{a_i(\theta_j - b_{ik})\}}{1+\exp\{a_i(\theta_j - b_{ik})\}} : \qquad \theta_j \sim N(0,1) \tag{2}$$

Where:

a_i is the discrimination parameter index for item i, and

b_i is the difficulty (location) parameter of category (threshold) for item i

$i(i = 1,2, \dots, I)$

The principle of GRM is based on operating characteristic functions which indicates the likelihood of responding to a particular category given trait or ability level θ are computed by subtracting adjacent $P^*{}_{ik}(\theta)$ as defined in equation (3).

$$P_{ik} = P^*{}_{ik} - P^*{}_{ik+1} \tag{3}$$

Where:

$$P^*{}_{ik} = \frac{e^{\{a_i(\theta_j - b_{ik})\}}}{1+e^{\{a_i(\theta_j - b_{ik})\}}} \tag{4}$$

$$P^*{}_{ik+1} = \frac{e^{\{a_i(\theta_j - b_{ik+1})\}}}{1+e^{\{a_i(\theta_j - b_{ik+1})\}}} \tag{5}$$

2.4 Parameters Estimation in Graded Response Model

The parameterization of equation (2) in term of slope-intercept is presented as

$$\Pr\left(Y_{ij} \geq k \middle| \gamma_i, \omega_i, \theta_j\right) = \frac{\exp\{\gamma_i(\theta_j - \omega_{ik})\}}{1+\exp\{\gamma_i(\theta_j - \omega_{ik})\}} \tag{6}$$

$$\left.\begin{aligned} a_i &= \gamma_i \\ b_{ik} &= \frac{\omega_i}{\gamma_i} \end{aligned}\right\} \tag{7}$$

When y_{ij} is the observed response for Y_{ij} and

$$p_{ij} = \Pr\,(Y_{ij} = y_{ij} | \gamma_i \omega_i\,) \tag{8}$$

Conditional density for respondent j is

$$f\left(y_j \middle| \boldsymbol{G}, \theta_j\right) = \prod_{i=1}^{I} p_{ij} \tag{9}$$

Where:

$y_j = (y_{1j}, y_{2j}, \dots, y_{Ij})$

$\boldsymbol{G} = (\gamma_1, \gamma_2, \dots, \gamma_I, \omega_1, \omega_2, \dots, \omega_I\,)$ and I is the number of items.

The likelihood of the respondent j is

$$L_j(\boldsymbol{G}) = \int_{-\infty}^{\infty} f\left(y_j \middle| \boldsymbol{G}, \theta_j\right) \emptyset(\theta_j) d(\theta_j) \qquad (10)$$

Where: $\emptyset(\theta_j)$ is the density function for standard normal distribution.

For N respondents

$$1ogL(\boldsymbol{B}) = \sum_{j=1}^{N} 1ogL_j(\boldsymbol{G}) \qquad (11)$$

Equation (11) is intractable, numerical method such adaptive quadrature would be implemented in Stata or R software.

4.0 Item Category Characteristic Curve and Discussions

A typical item category characteristic curve is a non-linear regression curves that displays likelihood of respondent responding in each of the mutually exclusive response categories.

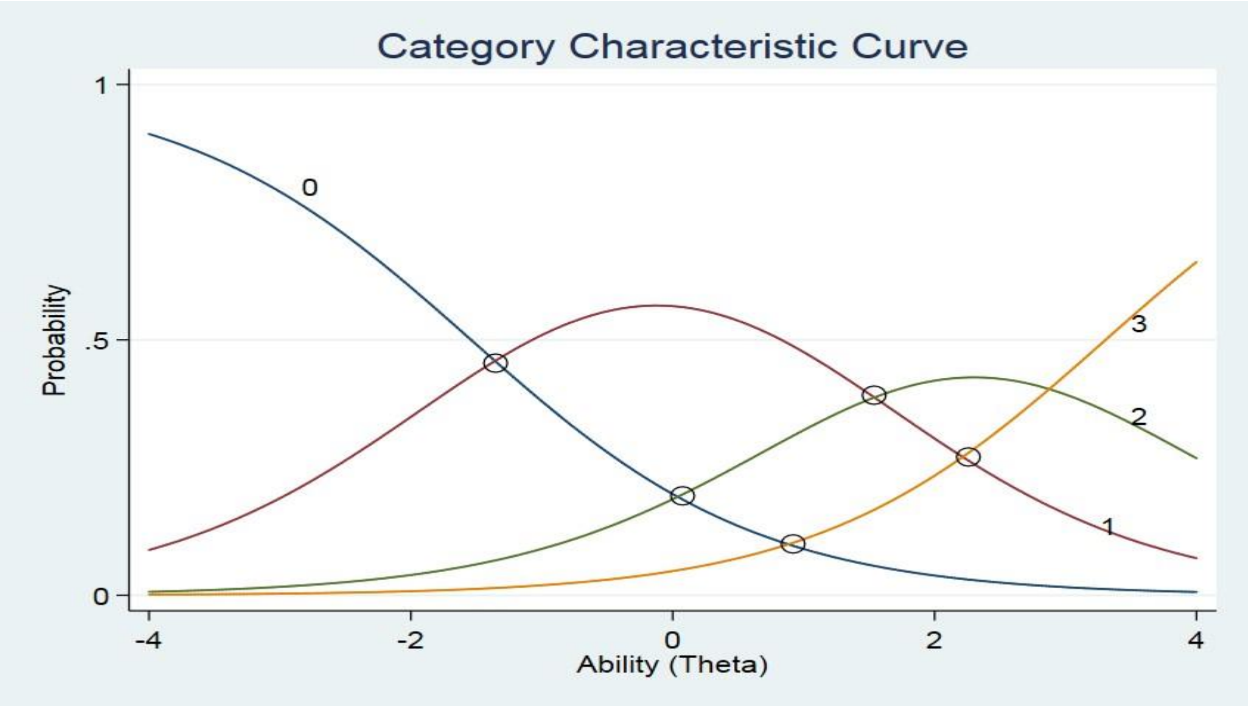

Figure 2: A Typical Category Characteristic Curve of an Item

In Figure 2, there are four responses to this item, curve 1, is taken as a base. The lower trait respondents are most likely to respond to this particular item by endorsing curve 1 instead of a base curve 0 with the corresponding likelihood or probability of endorsing the response option, this is because the first curve to cross base curve from the left is curve 1. This have negative index on ability axis. Moreover, an average individual most likely to endorse curve 2 instead of the base response, this is because where it crosses base curve on the ability axis is close to 0 while those respondents endorse curve 3 have higher trait, this is because curve 3 is the last to cross the base curve towards right.

In addition, response option tagged curve 3 would mostly discriminate respondents. This is because of all the response options, it has the largest slope and curve 2 would have the least, hence least in classifying respondents in term of their abilities.

5. Conclusions

The results from item response theory modelling is item base rather than sample, therefore attributes of individual item such as discriminating and location indices, and respondents' traits are estimable at any point of item response curve, this places IRT over CTT. In addition, parameters estimates remain unchanged under IRT even when attributes of respondents and environment surrounding estimation changes; this is a plus for IRT over CTT. The major focus and challenge of CTT is improving validity and reliability of estimates which are already be taken care of by IRT when its assumptions hold.

6. References

Adetutu, O. M. and Lawal, H. B. (2023). Applications of Item Response Theory Models to Assess Item Properties and Students' Abilities in Dichotomous Responses Items. Open Journal of Educational Development (ISSN: 2734-2050), 3(1), 01-19. Available from https://doi.org/10.52417/ojed.v3i/.304

Adetutu, O. M., and Lawal, H. B., (2020). On comparisons of Frequentist to Bayesian Estimation for item response theory models in the presence of dichotomous response of Science, Technology, Mathematics, and Education, 16(4), 128-137. https//:www.jostmed.futminna.edu.ng

Baker, F. B., and Kim, S. H. (2004). Item Response Theory: Parameter Estimation Techniques (2nd ed.). New York: Taylor and Francis.

Embretson, Susan E.; Reise, Steven P. (2000). Item Response Theory for Psychologists. Psychology Press. ISBN 9780805828191.

Hambleton, R., Swaminathan, H., Rogers, H. (1991). Fundamentals of Item Response Theory. Newbury Park, California: Sage Publications, Inc.

Lim, S. (2022). Review: A Course in Item Response Theory and Modelling with Stata, and Using R for Item Response Theory Model Applications. Structural Equation Modelling: A Multidisciplinary Journal, 27(4), 657-659, doi: 10.1080/10705511.2020.1740886, Available from https://doi.org/10.1080/10705511.2020.1740886

Linden, A. (2018). Review of Tenko Raykov and George Marcoulides's A Course in Item Response Theory and Modeling with Stata. The Stata Journal: Promoting Communications on Statistics and Stata, 18(2), 485–488. https://doi.org/10.1177/1536867X1801800213

National Council on Measurement in Education (2017) http://www.ncme.org/ncme/NCME/Resource_Center/Glossary/NCME/Resource_Center/Glossary1.aspx?hkey=4bb87415-44dc-4088-9ed9-e8515326a061#anchorC Archived 2017-07-22 at the Wayback Machine

Samejima, F. (1969). Estimation of Latent Ability Using a Response Pattern of Graded Scores. Psychometrika Monograph Supplement, No. 7