UNIVERSITI TEKNOLOGI MARA

BAYESIAN ESTIMATION OF LATENT TRAITS PARAMETERS IN A POLYTOMOUS RASCH RATING SCALE MODEL FOR SUBJECTIVE WELL-BEING RESPONSES SURVEY

NURUL HAFIZAH BINTI AZIZAN

Thesis submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Statistics)

College of Computing Informatics and Mathematics

November 2024

ABSTRACT

Generally, large sample sizes and normality assumption play vital roles in estimating model parameters. Unfortunately, small sample sizes and non-normally distributed data are common issues encountered in the real-world quantitative survey research, including studies on the subjective well-being. The frequentist approach, like the classical maximum likelihood (CML) estimation technique, has often been found impractical under these two circumstances, as demonstrated by previous scholars. While several studies have been carried out to address the issues of poor estimation accuracy resulting from CML under small sample sizes and non-normally distributed datasets, to date, very few studies have been conducted on the Rasch measurement model with polytomous responses survey datasets (i.e., Likert scale), and none have specifically attempted on the rating scale model. Therefore, the present study highlights the challenges related to accuracy in the estimation of the latent traits parameters in the polytomous Rasch rating scale model (pR-RSM). Specifically, this study proposes a more powerful technique known as the Bayesian estimation (BE) to address the negative impact of the small sample sizes with skewed data distributions on the estimation accuracy of latent traits in the pR-RSM obtained through the CML. The BE approach has been successfully tested in the dichotomous Rasch model, and the polytomous Rasch model (i.e., graded response model and the generalized partial credit model) in dealing with unfavourable datasets including small sample sizes and non-normally distributed data. Due to these convincing results, the present study proposed to extend the use of BE approach in the pR-RSM. The analyses in this study were performed using both simulated and actual survey datasets. The simulated survey data were generated according to the pR-RSM, incorporating a combination of shape of distributions (i.e., the standard normal distribution and the standard skew-normal distribution with various skewness values), along with different sample sizes and numbers of items. For the simulated survey data, the performance of the BE approach was evaluated and compared with the CML according to the accuracy and bias measures. The root mean square error (RMSE) and mean absolute error (MAE) were used to examine the accuracy of the parameter estimates. Meanwhile, bias in estimation was assessed through the mean difference of the estimates and true values of the latent traits parameters. To comprehensively evaluate the estimation performance for the simulated survey datasets, Markov Chain Monte Carlo (MCMC) simulation was carried out using 1000 iterations. Furthermore, the performance of BE and CML on the actual survey datasets was assessed and compared using two goodness of fit measures, namely Akaike information criterion (AIC) and Bayesian information criterion (BIC). Overall, the results of this study successfully revealed that small sample sizes and skewed data distributions were negatively impact the performance of the CML in estimating the latent traits parameters of the pR-RSM. The findings show that BE consistently outperformed CML, producing more accurate and less biased estimates, particularly in cases of small sample sizes with skewed data distributions. Consequently, this study suggests that BE emerges as the preferable and most powerful technique in treating the issues highlighted. Ultimately, this study provides a useful guideline for future researchers in selecting the most suitable parameter estimation technique in the context of the pR-RSM, based on the sample sizes and the skewness values.

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and the Most Merciful. First and foremost, I wish to thank Allah for the opportunities given to me to embark on my PhD in University Teknologi MARA (UiTM). All praises to Allah and His blessing for the completion of this long and challenging journey successfully. I would like to express my deepest thanks and gratitude to my supervisor Prof Dr Zamalia Mahmud and Dr Adzhar Rambli for their valuable advice, comments, and consistent guidance throughout these four years of my PhD journey.

I also would like to extend my utmost appreciation to the Ministry of Higher Education (MOHE) Malaysia for giving me scholarship under *Skim Latihan Akademik Bumiputra (SLAB)* and funding this research under Fundamental of Research Grant Scheme (FRGS/1/2019/STG06/UiTM/02/8). I also wish to thank all staff in *Bahagian Kemajuan Kerjaya (BKK), Jabatan Pembangunan Sumber Manusia (JPbSM)* for their co-operation in managing my study leave and scholarship. I would like to express my special thanks to Head and Coordinator of Postgraduate Studies for their continuous support and help me in finishing my PhD journey. Sincere thanks to all the academic staffs and technicians in Faculty of Computer and Mathematical Sciences for their co-operations and providing me with all the facilities that were required to fulfil my study. I would also like to acknowledge the *Institut Pengajian Siswazah (IPSis)* for providing useful training courses which helped me a lot to increase my knowledge.

I also appreciate all the support and advice I received from my colleagues and friends in going through this challenging journey.

Finally, I would like to express my strong gratitude to Ummi for her continuous encouragement, support and dua all through my studies. Her full supports and encouragement in process to complete this thesis is very valuable for me. My appreciation also goes out to my beloved parent and family members. This journey would not have been possible without the dedicated support of my family**.** This thesis is dedicated to the very supportive family for the vision and determination to educate me. Alhamdulillah. Thank You.

TABLE OF CONTENTS

CHAPTER ONE
INTRODUCTION

1.0 Introduction

This chapter encompasses the background of the study, problem statement, research questions, research objectives, scope and limitation, and significance of the study. In the background of the study, an overview of the subject matter and the rational of the study are provided. The following section highlights the specific issues, challenges and gaps of the research that necessitates its exploration. The formulated research questions and research objectives that outline the overall goal of this study are listed in the next two sections. The scope and limitations encountered in this study are provided in the following section, which defines the focus and all the constraints that may impact the research findings. The last section emphasizes the value and the potential contribution of the research outcomes.

1.1 Background of the Study

This section is organized into the several sub-sections. This includes, subjective well-being, sample size in survey research, normality in survey research, previous studies on subjective well-being, Rasch measurement model and its parameter estimation and development of new parameter estimates.

1.1.1 Subjective Well-Being

Subjective well-being (SWB) is the scientific term commonly used to describe happiness (Diener et al., 2009; Pavot & Diener, 2013; Diener, 2016), hence commonly, these two terms are used interchangeably in earlier studies. Often, SWB is measured through self-reported survey of how people evaluate their life as a whole or based on specific domains of life. Therefore, a survey of SWB mostly conducted to assess multidimensional constructs such as personal life, family, career, health condition, and more. The study of SWB has attracted a lot of attention from academicians and