

ABSTRAK

Nama : Sha'fiah Az Zahra
NIM : 1187010072
Judul : **Estimasi Parameter Laten Menggunakan Metode *Item Weighted Likelihood Estimation* dan Uji Kecocokan Model pada Teori Respon Butir Dikotomus**

Evaluasi instrumen pengukuran merupakan hal krusial dalam bidang pendidikan karena hal tersebut memengaruhi peningkatan mutu pendidikan. Salah satu bentuk instrumen pengukuran adalah tes/ujian, yang mana hal itu bertujuan untuk mengukur suatu aspek tertentu. Oleh sebab itu, diperlukan analisis butir untuk melihat kualitas instrumen pengukuran menggunakan pendekatan teori respon butir. Jenis pengukuran yang dilakukan pada ujian adalah pengukuran kemampuan terhadap suatu bidang tertentu. Kemampuan seseorang termasuk dalam hal yang bersifat laten. Maka dari itu, penelitian ini bertujuan untuk menentukan nilai estimasi parameter laten menggunakan metode *item weighted likelihood estimation* (IWLE) dan menentukan model terbaik bagi teori respon butir dikotomus. Nilai estimasi parameter laten diperoleh menggunakan metode IWLE dengan cara memaksimalkan fungsi likelihood kemudian dilakukan penyelesaian secara numerik oleh metode Fisher Scoring, sehingga diperoleh nilai estimasi untuk model teori respon butir dikotomus yang terdiri dari model Rasch, model 2PL, dan model 3PL. Cara pemilihan model terbaik diantara ketiga model tersebut menggunakan statistik uji *Chi-square*. Studi kasus pada penelitian ini menggunakan data jawaban penilaian akhir tahun (PAT) bahasa Inggris tahun pembelajaran 2021/2022 dari 150 siswa kelas XI SMAN Tanjungsari. Berdasarkan hasil perhitungan diperoleh nilai estimasi ketiga model yang kemudian nilai-nilai tersebut dikategorikan berdasarkan tingkat kemampuan. Hasilnya menunjukkan bahwa siswa-siswa kelas XI dikatakan memiliki pemahaman bahasa Inggris yang baik karena sebagian besar siswa mempunyai tingkat kemampuan berbahasa Inggris yang tinggi dan tidak ada yang kemampuannya rendah sekali. Kemudian dilakukan uji *Chi-square* untuk menentukan model terbaik antara model Rasch dan model 2PL, karena model 3PL menghasilkan nilai yang divergen maka tidak dilakukan uji kecocokan model. Menurut hasil perhitungan, model Rasch dikatakan model terbaik untuk data jawaban PAT karena memuat lebih banyak butir yang cocok dibandingkan dengan model 2PL.

Kata Kunci : teori respon butir, dikotomus, *item weighted likelihood estimation*, Fisher Scoring, *Chi-square*.

ABSTRACT

Name : Sha'fiah Az Zahra

NIM : 1187010072

Title : *Latent Parameter Estimation Using Items-Weighted Likelihood Estimation and Goodness-of-fit Test of Dichotomous Item Response Theory Models*

Measurement instrument evaluation is crucial in the field of education because it affects the quality improvement of education. One form of the measurement instrument is a test/exam, which aims to measure a specific aspect. Therefore, item analysis is needed to assess the quality of the measurement instrument using the item response theory approach. The type of measurement conducted in the exam measures the ability in a specific field. A person's ability is considered latent. Therefore, this study aims to determine the latent parameter estimation values using the item-weighted likelihood estimation (IWLE) method and determine the best model for the dichotomous item response theory. The latent parameter estimation values are obtained using the IWLE method by maximizing the likelihood function and then numerically solving it by the Fisher Scoring method, resulting in estimation values for the dichotomous item response theory model consisting of the Rasch model, 2PL model, and 3PL model. The best model selection method among these three models uses the Chi-square test statistic. This research uses the final assessment (PAT) English language answer data for the 2021/2022 academic year from 150 grade 11 students at SMAN Tanggulangsi. Based on the calculation results, the estimation values for the three models got obtained, then these values were categorized based on ability levels. The results showed that grade 11 students had a good understanding of the English language because most students had a high level of English proficiency, and none had meager abilities. The 3PL model produced diverging values, so a Chi-square test got conducted to determine the best model between the Rasch models and the 2PL model. According to the calculation results, the Rasch model is better on the PAT answer data than the 2PL model because it has more suitable items compared to the 2PL model.

Keywords : *item response theory, dichotomous, item weighted likelihood estimation, Fisher Scoring, Chi-square .*