



## **Goodness of fit and precision of the Graded Response Model estimates of a psychometric scale at varying number of categories.**

Yuliana Mora Cedeño\*

School of Statistics, University of Costa Rica, San Jose, Costa Rica –  
[yulianamoracedeno22@gmail.com](mailto:yulianamoracedeno22@gmail.com)

This study aims to evaluate the absolute goodness of fit of Graded Response Model (GRM), proposed by Samejima in 1969, when the number of categories of a polytomous scale is changed, and to analyze the precision of GRM estimates of the latent trait. These polytomous scales are frequently used in psychometric measurements of skill levels, such as quantitative skills or reading skills. The analysis performs a simulation study design to assess these objectives. It was simulated a polytomous scale derived from normal variables based on a one-factor model. It was computed 100 replications of 1000 cases and 14 items for each scale of 3, 4, 5, and 6 categories. The GRM is estimated for each replication. The absolute goodness of fit is evaluated using the Likelihood Ratio Test (LRT) that contrasts the model with the saturated model: the  $G^2$  test. I analyze precision computing the Mean Bias and the Root-Mean-Square-Error (RMSE) of the GRM theta parameter. The results show that, based on the  $G^2$  test, the best fit of the GRM is obtained with a scale of 5 categories, and the fit worsens with a 6 categories scale. There were no good predictions of the skill level with any of the 4 scales because all of them had RMSE of 0.97 and Mean Biases of 0.87. The study results suggests that it is better to operationalize skill levels with a 5-categories scale.

**Keywords:** absolute goodness; polytomous scale; One factor model.