ABSTRACT

An analysis of achievement test using logistic models

This study had three main objectives. First, to demonstrate the procedures of item analysis, using logistic models, including the interpretation of the results. Second, to compare the results of item analyses using logistic models and the classical model. Third, to examine the agreement and the contradiction of the results from the logistic models and the classical model.

The data used in this study were responses of 1502 Prathom Suksa VI students to the 60-item mathematics achievement test. The testing, a regular school examination, was taken place at the end of the second semester of the academic year 1986. These students were from schools in Bangkapi area in Bangkok.

The preliminary analyses indicated that the test scores were unidimensional.

Under the three parameter logistic model -- 3PL model, the parameter estimation results were as follows: difficulty estimates ranged from -3.7490 to 3.9110, discrimination estimates ranged from .1471 to 2.000, and the lower asymptote or guessing estimates ranged from .0418 to .3966. Item no. 14 caused the incompleted run, therefore it was excluded from the

estimation. Four items had discrimiation estimates less than .3000, these were item nos. 15, 30, 41, and 60.

Under the Rasch model -- 1PL model, the difficulty parameter estimates of 59 items were between -2.68 and 1.90. The fit analysis showed that four items did not fit the medel. They were item nos. 30, 34, 45, and 60.

All sixty items were included in the item analysis under the classical model. The easyness parameter estimates -- proportion correct, ranged from .1853 to .9401. Three items had easyness indices above .9000, these were item nos. 1, 3, and 17. The discrimination estimates --point biserial correlation between item i and total test excluding item i, ranged from -.0891 to .4170. Item no. 14 had the lowest discrimination, and it was negative. Five items had discrimination indices lower than .10, these were item nos. 14, 30, 34, 58, and 60.

Further analyses were done to examine whether or not the choices chosen by the examinees, among 4 alternatives for each item, were in accord with the examiner 's expectations. The results showed that from 18 items out of 60, only one alternative was found inappropriate for each item, and 2 alternatives from another 3 items were of the same problem.

Ranks of item difficulties estimated from the three different models were in high agreement, especially between the 1PL model and the classical model.

The 3PL model rejected less number of items than the 1PL model did. Eventhough the classical model gave results that dependent upon the group ability, it gave useful information for the alternative inprovement. Therefore, at the stage of test development, the classical model should be used in couple with the logistic model. For the final test forms, the item parameters should be estimated under the logistic model.