Test Construction from Item Banks

Ellen Timminga\textsuperscript{1} and Jos J. Adema\textsuperscript{2}

ABSTRACT This chapter discusses the topic of computerized test construction from item banks if a dichotomous Rasch model holds for all items of the bank. It is shown how the task of selecting items optimally can be formulated as a 0-1 linear programming problem. Next, integer linear programming is introduced as an easier way to near-optimal item selection. Then, the techniques of 0-1 and integer linear programming are discussed. Finally, an overview of the literature in this field and a discussion of various approaches to the construction of parallel tests is given.

7.1 Introduction

Computerized test construction assumes the existence of a large pool of pre-calibrated items, a so-called item bank, which stores psychometric and other characteristics of items. Here we assume that all items in an item bank are scored dichotomously and perfectly fit a Rasch model (RM). Moreover, the difficulty parameters of the items are treated as known true values. In practice, however, there will exist no banks of items perfectly fitting a RM, and difficulty parameters will always be estimates rather than true values. Thus, the importance of selecting the test with the best psychometric qualities should not be exaggerated. Satisfying the practical constraints is the major goal.

Two approaches to computerized test construction can be distinguished:

1. the construction of complete tests, and
2. adaptive testing.

The construction of complete tests, which is the topic of this chapter, means to select, on each occasion, a test from the item bank meeting certain requirements; this test is then given to one or more examinees. In adaptive testing, examinees are presented with individualized tests. One item is selected at a time on the basis of the currently available estimate of the examinee's ability (e.g., Lord, 1980; Weiss, 1982). The ability estimate is updated after the administration of each item, and the next optimal item is selected. A criterion is needed for deciding when to stop the process; such criteria often focus on accuracy of ability.
estimation. This chapter, however, does not deal with the problem of adaptive testing.

Good tests meet requirements that are specified by the test constructor. Meeting these requirements, however, is problematic for at least two reasons: one is the impossibility of examining all possible tests in order to find the one that best fits the test constructor’s requirements. Even for a relatively small item bank of 200 items, say, in theory $2^{200} - 1$ tests have to be compared. The other is related to the practical requirements that tests have to meet. The problem is that several requirements have to be taken into account simultaneously, which often is almost impossible.

The main concern in this chapter is to optimize the psychometric qualities of the test to be selected. It will also be indicated how to consider other kinds of requirements. The psychometric characteristics used are item difficulty parameters, and item and test information functions.

It is well-known that the maximum likelihood estimator of $\theta$ asymptotically has the smallest variance attainable. This minimum variance is $[I(\theta)]^{-1}$, where $I(\theta)$ is called the test information. Thus, an information function indicates the maximum accuracy with which an examinee’s ability parameter can be estimated over the total ability range. The general formula of the test information function for the RM is (see also (4.6)):

$$I(\theta) = \sum_{i=1}^{k} P_i(\theta)[1 - P_i(\theta)],$$

(7.1)

where $k$ is the number of test items.

Observe that (7.1) consists of independent and additive contributions of items. This is a result of the local independence assumption that is made for maximum likelihood estimation (see Chapters 1 and 4). Thus, the information provided by a test is simply the sum of the information functions of the items in the test:

$$I(\theta) = \sum_{i=1}^{k} I_i(\theta),$$

(7.2)

where $I_i(\theta)$ denotes the information function of item $I_i$, which can also be formulated as follows:

$$I_i(\theta) = P_i(\theta)[1 - P_i(\theta)] = [1 + \exp(-(\theta - \beta_i))]^{-1}.$$

(7.3)

For all items fitting the same RM, the information functions are identically bell-shaped functions that only differ in their location. Expression (7.3) clearly shows that the maximum of the information function is obtained at ability level $\theta = \beta_i$, and that this maximum amount is equal to 0.25. A thorough discussion of information functions can, for instance, be found in Birnbaum (1968), Lord (1980), and Hambleton and Swaminathan (1985).