Mixture Distribution Rasch Models

Jürgen Rost and Matthias von Davier

ABSTRACT This chapter deals with the generalization of the Rasch model to a discrete mixture distribution model. Its basic assumption is that the Rasch model holds within subpopulations of individuals, but with different parameter values in each subgroup. These subpopulations are not defined by manifest indicators, rather they have to be identified by applying the model. Model equations are derived by conditioning out the class-specific ability parameters and introducing class-specific score probabilities as model parameters. The model can be used to test the fit of the ordinary Rasch model. By means of an example it is illustrated that this goodness-of-fit test can be more powerful for detecting model violations than the conditional likelihood ratio test by Andersen.

14.1 Introduction

The chapter deals with a generalization of the Rasch model (RM) to a discrete (finite) mixture distribution model. Stated briefly, the mixed Rasch model (MRM) assumes that the RM does not hold for the entire population, but does so within (latent) subpopulations of individuals which are not known beforehand. Section 14.1 describes the general structure of discrete and continuous mixture models. The RM itself being a special kind of a continuous mixture model, its generalization to the MRM assumes discrete mixtures of continuous mixture distributions. The central properties of the MRM, including parameter estimation, are outlined in Section 14.2. One of the most relevant applications of the MRM is the testing of fit of the ordinary RM. Some related results are presented in Section 14.3. The unrestricted MRM does not require any assumption about the ability distributions within the classes. This entails, however, a large number of model parameters. The latter can be reduced by imposing restrictions on the latent score distributions as proposed in Section 14.4. The mixture generalization has also been developed for ordinal RMs, like the partial credit or the rating scale model. These extensions of the MRM, however, are described in Chapter 20.

14.2 Continuous and Discrete Mixture Distribution Models

The basic assumption of mixture distribution models (MDMs) is that the distribution of a (possibly vector valued) observed random variable is not adequately

---

1Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, Olshausenstraße 62, D-24098 Kiel; e-mail: ROST@IPN.UNI-KIEL.DE
described by a single uniform probability function, but rather by a number of conditional probability functions. The variable on which the probabilities are conditioned is referred to as the mixing variable. The probability function of a discrete\(^2\) variable \(X\) is formally defined as

\[
P (X = x) = \int_{-\infty}^{\infty} P (X = x|\theta) dF (\theta), \tag{14.1}
\]

where \(\theta\) is a continuous mixing variable, and \(F(\theta)\) its probability distribution. The unconditional probability \(P(X = x)\) is obtained by integrating over the density of the mixing variable (Everitt and Hand, 1981). The mixing variable can also be discrete, which reduces the integral structure to a sum,

\[
P (X = x) = \sum_{c=1}^{C} P (c) P (X = x|\theta = c). \tag{14.2}
\]

In this case, the unconditional probability is a weighted sum, where the weights \(P(c)\) are the probabilities that the mixing variable has value \(c\), and the terms are summed over all values of \(c\). Parameter \(\theta\) in such a discrete MDM defines a (usually small) number of components of the mixture or classes of individuals. Therefore, the weights \(P(c)\) define the relative size of the components of the mixture or class sizes. Item response models (Lord, 1980) are usually MD models, either continuous or discrete. The observed random variables described by IRT models are the response vectors of the individuals, \(x = (x_1, \ldots, x_k)\), in the present case with binary components \(x_i \in \{0, 1\}\). The mixing variable \(\theta\) is referred to as the ‘latent variable’. The conditional probabilities are often defined as the logistic function, e.g., for the RM,

\[
P (X = x|\theta) = \prod_{i=1}^{k} \frac{\exp [x_i (\theta + \beta_i)]}{1 + \exp (\theta + \beta_i)}, \tag{14.3}
\]

where \(\beta_i\) is an easiness parameter of item \(I_i\). Written in this way, the RM is a continuous MDM, because \(\theta\) is a real valued latent variable. In the RM, however, it is not necessary to know or to specify the distribution of the mixing variable, \(F(\theta)\), as is done in other IRT models.

Since the sum score of a response vector, \(R = \sum_{i=1}^{k} X_i\), is a sufficient statistic for estimating \(\theta\), the classification of individuals according to their sum scores \(r\) is equivalent to the classification according to their estimated trait values \(\hat{\theta}\) (but not their true trait values \(\theta\)). Hence, the RM can also be written as a discrete MDM,

\[
P (X = x) = P (r_{\bar{x}}) P (X = x|r_{\bar{x}}), \tag{14.4}
\]

\(^2\)Densities of continuous variables and their mixtures are not treated here.