A FLEXIBLE LATENT TRAIT MODEL FOR RESPONSE TIMES IN TESTS

JOCHEN RANGER
UNIVERSITY OF GIESSEN

JORG-TOBIAS KUHN
UNIVERSITY OF MUNSTER

Latent trait models for response times in tests have become popular recently. One challenge for response time modeling is the fact that the distribution of response times can differ considerably even in similar tests. In order to reduce the need for tailor-made models, a model is proposed that unifies two popular approaches to response time modeling: Proportional hazard models and the accelerated failure time model with log–normally distributed response times. This is accomplished by resorting to discrete time. The categorization of response time allows the formulation of a response time model within the framework of generalized linear models by using a flexible link function. Item parameters of the proposed model can be estimated with marginal maximum likelihood estimation. Applicability of the proposed approach is demonstrated with a simulation study and an empirical application. Additionally, means for the evaluation of model fit are suggested.

Key words: response time, proportional hazard model, accelerated failure time model.

1. A Flexible Latent Trait Model for Response Times in Tests

In psychological testing, the consideration of response times is becoming more and more popular. Response times in test items can be used in order to detect random guessing, to detect collusion between test takers or as collateral information on parameters of item response models—for an overview over possible applications see van der Linden (2009) and Schnipke and Scrams (2002). However, in order to use the full diagnostic potential of response times, latent trait models are needed that relate the observed response times to characteristics of the test takers. Such a model was first sketched in the 1950s by Furneaux (1952). Since then, further models have been proposed (Scheiblechner, 1979; Maris, 1993; van Breukelen, 1997; Douglas, Kosorok, & Chewing, 1999; van der Linden, 2006). However, there does not seem to exist a standard model that can routinely be used for response times in tests. One of the reasons for this could be the fact that response time distributions from different items can be diverse such that a specific model might not be applicable to all sorts of tests. For example, although models based on the log–normal distribution in general and the approach of van der Linden (2006) in particular often yield good model fit (Schnipke & Scrams, 2002), the log–normal distribution sometimes fails (Klein Entink, van der Linden, & Fox, 2009).

The following example is meant to illustrate this finding. Figure 1 displays features of the response time distribution in two items from different personality scales. The plots in the first row illustrate the response time distribution in a neuroticism item from the German version of the NEO-FFI (Borkenau & Ostendorf, 1993). The plots in the second row depict the response time distribution in an anxious-calm item of the Eysenck Personality Profiler (Eysenck, Wilson, & Jackson, 1998). The first column contains Nelson–Aalen estimators of the cumulative hazard.
FIGURE 1.
First column: Cumulative hazard rate of the neuroticism item (top) and the anxious-calm item (bottom). Second column: Q–Q plot of the neuroticism item (top) and the anxious-calm item (bottom). Dashed line indicates the cumulative hazard rate of the best fitting exponential distribution.

rate.1 The second column contains Q–Q plots for different distributions: In the case of the neuroticism item, the Q–Q plot is based on a shifted exponential distribution, whereas in case of the anxious-calm item, the Q–Q plot is based on the log–normal distribution.

The cumulative hazard rate of the neuroticism item (first row) resembles a straight line, if one ignores the deviation at the last five observations that probably represent outliers. This particular shape of the cumulative hazard rate suggests an exponential distribution. However, in order to achieve good fit to the exponential distribution, the response times have to be shifted toward the origin, as there seems to exist a minimum time limit of about 3.0 s no subject can fall below. This is obviously due to a minimal time requirement for reading this item. Having subtracted the response time minimum from all response times, the exponential distribution fits the data excellently. This can be seen in the Q–Q plot where the observed quantiles of the shifted response times are plotted against the theoretical quantiles of the corresponding exponential distribution.

The cumulative hazard rate of the anxious-calm item has no characteristic and unambiguously interpretable shape. It clearly deviates from the dotted line, indicating that the response times are not distributed according to the exponential distribution. Nevertheless, the log–normal distribution fits very well as can be seen in the Q–Q plot in the second row where the observed quantiles are plotted against the theoretical quantiles of the log–normal distribution. In this case, the Q–Q plot suggests the presence of four outliers.

1The cumulative hazard rate is the negative log survival probability \(-\log S(t)\). Response time distributions can be characterized by their cumulative hazard rate. The cumulative hazard rate of the exponential distribution, for example, is a straight line. The cumulative hazard rate can be estimated by the nonparametric Nelson–Aalen estimator.