

Models for dependent paired comparison data

Manuela Cattelan¹ and Cristiano Varin²

¹*Department of Statistical Sciences, University of Padova*

²*D.A.I.S., Ca' Foscari University - Venice*

Abstract

Paired comparison data are binary data that originate from the comparison of objects in couples. This type of data arises in many applications including animal behaviour experiments, sport tournaments and psychometric experiments. The aim of the analysis is to determine whether some covariates can predict the results of comparisons or to estimate a “worth” parameter for each of the elements compared which can be used for ranking them.

The Bradley-Terry and the Thurstone models are usually employed to analyse paired comparison data. The Bradley-Terry model is a logistic model in which the linear predictor is described as the difference of the worth parameters of the two elements compared. The Thurstone model is similarly specified, but it employs a probit link function. However, usually both models are estimated assuming that all comparisons are independent. This assumption is often unrealistic, for example in animal behaviour experiments it seems inappropriate to assume that the results of two different fights involving a common animal are independent. Extensions of traditional models that overcome this limit are proposed. Unfortunately, the inclusion of dependence generates some difficulties in the estimation of the model since the dependence among comparisons involving the same element produces an intricate scheme of cross-correlations among observations and sometimes, as in animal behaviour experiments, it is not possible to distinguish groups of independent data. The problem of estimating the extension of the Bradley-Terry model for dependent data is overcome by resorting to optimal estimating equations. The dependence structure is taken into account in the covariance matrix of the observations which includes parameters that cannot be estimated independently of the regression parameters. This problem is solved by means of a hybrid pairwise likelihood estimation which cycles between optimal estimating equations for estimation of the regression parameters given the covariance parameters, and maximum pairwise likelihood for estimation of the covariance parameters given the regression parameters. On the contrary, the extension of the Thurstone model can be seen as a multivariate probit model, so many different estimation techniques can be employed. The properties of different estimation methods in models for dependent paired comparison data will be discussed and illustrations to real data applications will be presented.

Keywords: Optimal estimating equations, Paired comparisons, Pairwise likelihood.

AMS subject classifications: 62F10, 62J12.