J. R. Statist. Soc. A (2008)
171, *Part* 1, *pp.* 137–157

The analysis of change, Newton's law of gravity and association models

Mark de Rooij

Leiden University, The Netherlands

[Received September 2004. Final revision April 2007]

Summary. Newton's law of gravity states that the force between two objects in the universe is equal to the product of the masses of the two objects divided by the square of the distance between the two objects. In the first part of the paper it is shown that a model with a 'law-of-gravity' interpretation applies well to the analysis of longitudinal categorical data where the number of people changing their behaviour or choice from one category to another is a measure of force and the goal is to obtain estimates of mass for the two categories and an estimate of the distance between them. To provide a better description of the data dynamic masses and dynamic positions are introduced. It is shown that this generalized law of gravity is equivalent to Goodman's RC(*M*) association model. In the second part of the paper the model is generalized to two kinds of three-way data. The first case is when there are multiple two-way tables and in the second case we have change over three points of time. Conditional and partial association models are related to three-way distance models, like the INDSCAL model, and triadic distance models respectively.

Keywords: Categorical data; Euclidean distance; Gravity model; Longitudinal data; Square tables; Triadic distance

1. Introduction

This paper will be concerned with longitudinal categorical data, i.e. repeated measurements on a number of observational units with the same instrument. The main interest in studying longitudinal data is whether change occurred and, if so, what the nature of the change is. We shall confine ourselves to the case of categorical data. Our questions concern qualitative change, i.e. changes in attitude, opinion, behaviour or any other categorical variable. This is typically different from continuous data where it might be possible to describe change in terms of better or worse; for categorical data descriptions are in terms of 'different' or 'the same'.

Once longitudinal categorical data have been collected they can be represented in transition frequency tables, which are contingency tables where each way corresponds to the categories of a variable measured at a specific time point (we adopt the way mode terminology for the tables of Carroll and Arabie (1980)). The number of time points defines the number of ways of the transition frequency table. Having measured a group of people twice on a categorical variable, a square transition frequency table arises. If measurements are obtained at three time points the data can be gathered in a three-way contingency table, and so forth.

An example of such data is obtained from Upton and Särlvik (1981) who discussed changes in political voting in Sweden. The data are shown in Table 1. There are five political parties: the *Communists* COM; the *Social Democrats* SD; the *Centre Party* C; the *People's Party* P; the

Address for correspondence: Mark de Rooij, Methodology and Statistics Unit, Institute for Psychological Research, Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands.
E-mail: rooijm@fsw.leidenuniv.nl

 0964–1998/08/171137