Computational Statistics and Data Analysis 78 (2014) 100–118

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

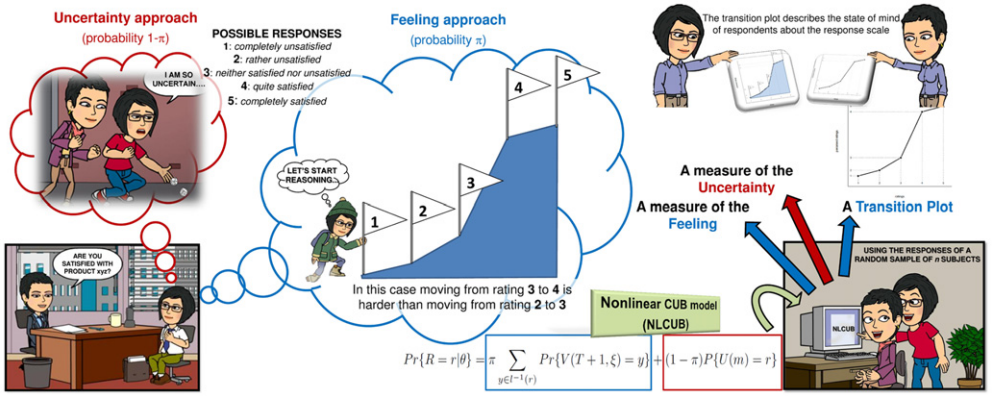
journal homepage: www.elsevier.com/locate/csda

Modeling rating data with Nonlinear CUB models

Marica Manisera, Paola Zuccolotto *

University of Brescia, C.da S. Chiara 50, 25122 Brescia, Italy

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 3 May 2013
Received in revised form 3 April 2014
Accepted 4 April 2014
Available online 13 April 2014

Keywords:
CUB models
Decision process
Latent variables
Ordinal data
Rating data
Transition probability
Eurobarometer

ABSTRACT

A general statistical model for ordinal or rating data, which includes some existing approaches as special cases, is proposed. The focus is on the CUB models and a new class of models, called Nonlinear CUB, which generalize CUB. In the framework of the Nonlinear CUB models, it is possible to express a transition probability, i.e. the probability of increasing one rating point at a given step of the decision process. Transition probabilities and the related transition plots are able to describe the state of mind of the respondents about the response scale used to express judgments. Unlike classical CUB, the Nonlinear CUB models are able to model decision processes with non-constant transition probabilities.

1. Introduction

The latent variable approach is very wide, covers a variety of methods and techniques and can be applied in several fields (see, for example, Borsboom et al., 2003 and the references therein). In this paper, we propose a general statistical model

* Correspondence to: C.da S. Chiara, 50, 25122 Brescia, Italy. Tel.: +39 030 2988634; fax: +39 302400925.
E-mail addresses: marica.manisera@unibs.it (M. Manisera), paola.zuccolotto@unibs.it (P. Zuccolotto).

http://dx.doi.org/10.1016/j.csda.2014.04.001
0167-9473/

for ordinal data that provides insights into the cognitive mechanism connected to the latent variables driving individuals' responses on a rating scale. In our model, the variety of complex interactions determining the final response is summarized by two components that, borrowing the terminology of the CUB models (Piccolo, 2003; D'Elia and Piccolo, 2005), can be called *feeling* and *uncertainty*. The feeling component accounts for reasoning and logical thinking as well as the set of emotions, perceptions, and subjective evaluations that individuals have with regard to the latent trait being evaluated. In our proposal, this is the outcome of a step-by-step decision process. The uncertainty component is concerned with the indecision inherently present in any human choice and does not generally depend on the individuals' position on the latent variable. A very interesting feature of the general model proposed in this paper is given by the so-called *transition probability*, defined as the probability of increasing one rating point at a given step of the decision process. Transition probabilities allow us to understand and describe the individuals' mental stance towards the response scale. The proposed general statistical model includes several approaches existing in the literature as special cases. In particular, we focus on the CUB models that were recently introduced to analyze ordinal data (starting with Piccolo, 2003, who introduced a mixture of two latent components to model ranking or rating data, and D'Elia, 2000, who previously suggested motivating a binomial component for exploring feeling in the paired comparisons model when the preferences are expressed by ranks) and generalize them by introducing a new class of models, called Nonlinear CUB, which, unlike classical CUB, are able to describe decision processes with non-constant transition probabilities.

The paper is organized as follows: in Section 2, the basic idea motivating the paper is intuitively presented through two simple examples. In Section 3, we propose a general framework for the decision process that is able to translate the individual's position on a latent variable into a rating. Section 4 briefly describes the CUB models and shows how they can be viewed as a special case in the proposed general framework. In Section 5, the Nonlinear CUB models are proposed as a generalization of the classical CUB models. Estimation issues are discussed in Section 6, where an illustrative example on artificial data is also shown. Results from simulation studies and real data analyses are presented in Sections 7 and 8, respectively. Section 9 concludes the paper.

2. Basic idea: the unconscious decision process in CUB and Nonlinear CUB models

In this paper, we propose a framework for the Decision Process (DP) which leads an individual to express a rating r about a latent trait on a given ordinal scale. In the followings, we refer to ratings as Likert-scaled responses, using the concept of Likert-type item in its broadest sense of ordered categorical item. In our proposal, r is generated by a latent DP composed of a feeling and an uncertainty component, consistently with the founding idea of the CUB models. The idea is that the feeling component should take into account any reasoned assessment, as well as the set of emotions, sentiments and perceptions logically connected with the object being evaluated, while the uncertainty component accounts for other elements, such as, for example, the unconscious willingness to delight the interviewer or the indecision deriving from the difficulty one can find in evaluating some specific objects about which he/she has not a definite opinion. The two components result from two different approaches that individuals can follow to finally express their judgment. They are not necessarily consecutive but possibly coexist in the DP.

The *feeling approach* determines the feeling component of the DP and consists of a step-by-step process, called *feeling path*. The feeling path proceeds through T steps; at each step t an elementary judgment is given, which leads to a provisional rating r_t, so that the last rating r_T results from the accumulation of T elementary judgments. The uncertainty component of the DP is instead related to the *uncertainty approach*, which, for several reasons, leads an individual to formulate a completely random rating. The rating expressed by the respondent to the interviewer is derived by the feeling or the uncertainty approach, with given probabilities. However, for each answer, we do not know whether the respondent has decided according to the former or the latter approach.

As shown in the following two examples, this DP is suitable for explaining the generation of ratings distributed as postulated both by the CUB models (Piccolo, 2003), and by the Nonlinear CUB models, which are proposed in this paper.

Example 1 (*DP of CUB Models*)**.** Suppose a person is asked to express a judgment about his/her satisfaction with a product by using a Likert scale from 1 to $m = 5$. In the feeling approach, the respondent asks him/herself for $T = m - 1 = 4$ times 'Do I have a positive sensation about this product? Yes or no?' and gives a quick and instinctive response each time. The idea underlying this mechanism is that, while reasoning, positive and negative sensations come to mind randomly and disorderly, according to the individual's experience about the latent trait being evaluated. In the end, 1 plus the total number of 'Yes' responses is the last rating r_T of the feeling path (Table 1). In the uncertainty approach, for a wide variety of reasons (for example the respondent may not be completely confident on his reasoning, or he/she could be reluctant to let the interviewer know his real opinion), the rating is drawn from a discrete Uniform distribution in $(1, 2, \ldots, 5)$. The expressed rating can be formulated by the feeling or the uncertainty approach with probabilities π and $1 - \pi$, respectively. As we will see in Section 4, the distribution of the ratings assumed by the CUB models (Piccolo, 2003) is consistent with this unconscious mechanism.

Example 2 (*DP of Nonlinear CUB Models*)**.** As in Example 1, suppose a person is asked to express a judgment about his/her satisfaction with a product by using a Likert scale from 1 to $m = 5$. In the feeling approach, the person unconsciously asks him/herself for $T > m - 1$ times 'Do I have a positive sensation about this product? Yes or no?' and gives a quick and