

CAN LIKERT SCALES BE COMPOSITIONAL?

B. Ferrer-Rosell¹, V. Pawlowsky-Glahn², J.J. Egozcue³ and G. Coenders⁴

¹Unviersity of Lleida, Lleida, Spain; *berta.ferrer@aegeun.udl.cat*

²Universtiy of Girona, Girona, Spain; *vera.pawlowsky@udg.edu*

³Technical University of Catalonia, Barcelona, Spain; *juan.jose.egozcue@upc.edu*

⁴University of Girona, Girona, Spain; *germa.coenders@udg.edu*

Abstract

Many surveys consist in presenting D items, and for each of them, the individual is required to assign value or score in a scale, for instance, from 1 to 10. This is frequently known as scoring in a (discrete or continuous) Likert scale. The meaning of these scales is object of frequent discussions. For instance, are scores in an absolute, although subjective, scale? Or is the information provided only ordinal? Consequently, the methods to analyse this type of data are also controversial. Here, data in Likert ITEMS are assumed to convey mainly ordinal information, but adding some mild assumptions they can be treated as compositional data.

The proposed way is as follows. The collection of D items can be RANKED ROW-WISE according to their ORIGINAL LIKERT scores. Ties are admissible just admitting fractional RANKS. According to Thurston, the ranks of the items can be viewed as a composition. The reasons for this were that the ranks add to a constant. From an updated compositional point of view, this constant sum is not critical for considering a vector as a composition. The main point is its invariance under scaling by positive constants of the scores, which in fact holds in this case. For instance, with $D=4$, items could be ordered by preferences as (1, 2.5, 2.5, 4) but the ordering is equivalent to that in (10, 25, 25, 40). Once these ranking vectors are placed in the D -part simplex, the Aitchison geometry of the simplex and its derived log-ratio procedures are available. Under the assumption that the Aitchison geometry makes sense, the survey data can be analyzed using log-ratio methods.

Some details may be important. When there is a non-response in the Likert scales, the ranks of items are evaluated in a subcomposition (the constant sum is lost, but the data is still defined in a subcomposition). Association between items is measured by the variation matrix; the answers of an individual, coded as a D -part composition, can be represented in log-ratio coordinates; the relation between items and individuals can be visualized using compositional biplots; total variance can be decomposed into variances of simple log-ratios, of centered log-ratio components, or of ilr-coordinates, giving the opportunity of further traditional statistical analyses. A simple case of a survey on Likert scales is used to illustrate the potential of the methods.