Modelling censored compositional data in the correct censored space– CoDaWork 2017

T.C.Leung[1] and J. Bacon-Shone[1]

[1]Social Sciences Research Centre, The University of Hong Kong, Hong Kong, *h8894118@hku.hk* and *johnbs@hku.hk*

Abstract

Different versions of the logratio transform (Aitchison, 1982) have been widely accepted as a standard method for analysing compositional data, where zeros can be handled by multiplicative replacement or other methods. (Martin-Fernandez, Barcelo-Vidal and Pawlowsky-Glahn, 2003; Aitchison and Kay, 2003; Bacon-Shone, 2003; Butler and Glasbey, 2008; Palarea-Albaladejo and Martin-Fernandez, 2015).

Bayesian analysis assuming interval censoring has been explored using WinBUGS and applied to specific data sets. (T.C.Leung and J. Bacon-Shone, 2015) This interval censoring approach avoids the need for specific treatment of zeros beforehand and moderate size datasets can be analysed in a simple way.

However, all the above methods all require at least one component without zeros and apply censoring on the logratio space, not on the direct censored space.

In this presentation, modelling of censored compositional data in the correct censored space, i.e. directly on the simplex, is demonstrated for simulated and real datasets.

References

AITCHISON, J. 1982. The statistical analysis of compositional data. Journal of the Royal Statistical Society. Series B (Methodological), 139-177.

AITCHISON, J. & KAY, J. W. 2003. Possible solution of some essential zero problems in compositional data analysis.

BACON SHONE, J. 1992. Ranking methods for compositional data. Applied Statistics, 533-537.

BUTLER, A. & GLASBEY, C. 2008. A latent Gaussian model for compositional data with zeros. Journal of the Royal Statistical Society: Series C (Applied Statistics), 57, 505-520.

FRY, J. M., FRY, T. R. & MCLAREN, K. R. 2000. Compositional data analysis and zeros in micro data. Applied Economics, 32, 953-959.

MARTIN-FERNANDEZ, J. A., BARCELO-VIDAL, C. & PAWLOWSKY-GLAHN, V. 2003. Dealing with zeros and missing values in compositional data sets using nonparametric imputation. Mathematical Geology, 35, 253-278.

PALAREA-ALBALADEJO, J. & MARTIN-FERNANDEZ, J. A. 2015. zCompositions—R package for multivariate imputation of left-censored data under a compositional approach.

T.C.LEUNG and J. BACON-SHONE, Compositional Data Analysis and the zero problem: interval censoring approach. *CoDaWork 2015*. Girona, Spain.