Zeros and Subcompositionally Coherenent Estimators

John Bear[1] and Dean Billheimer[2]
[1]Tucson, AZ, USA; *jb@johnbear.net*
[2]University of Arizona, Tucson, AZ, USA

Abstract

Subcompositional coherence demands that we reach the same inferential conclusion about subcompositions under two different operations:

1. take a subcomposition of the data, and then performing inference, or

2. performing inference, and then taking a subcomposition.

   (Aitchison (1986), Aitchison and Egozcue (2005), Egozcue (2009), Egozcue and Pawlowsky-Glahn (2011).)

Some researchers have argued that subcompositional coherence might not be maintainable when modeling compositional data sets containing essential zeros, Butler and Glasbey (2008), Scealy and Welsh (2014). Others have argued that perhaps it is possible to get "close enough" to subcompositional coherence, Greenacre (2011).

We offer formal criteria for subcompositional coherence using functions of location and dispersion estimators.We show that a given statistical model can have both subcompositionally coherent and noncoherent estimators. Note that we are extending the definition of subcompositional coherence from a property of a probability model to one of parameter estimators.

We illustrate with examples, including examples containing essential zeros. The benefit of these explicit criteria is that they make clearer the relationships between zeros, parameter estimators, and subcompositional coherence of models.

References

Aitchison, J. (1986). *The statistical analysis of compositional data*. Chapman & Hall, Ltd.

Aitchison, J. and Egozcue, J. J. (2005). Compositional data analysis:where are we and where should we be heading? *Mathematical Geology*, **37**(7), 829–850.

Butler, A. and Glasbey, C. (2008). A latent gaussian model for compositional data with zeros. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **57**(5), 505–520.

Egozcue, J. (2009). Reply to "on the harker variation diagrams;..." by J.A. Cortes. *Mathematical Geosciences*, **41**(7), 829–834.

Egozcue, J. J. and Pawlowsky-Glahn, V. (2011). *Basic Concepts and Procedures*, pages 12–28. In Pawlowsky-Glahn and Buccianti (2011).

Greenacre, M. (2011). Measuring subcompositional incoherence. *Mathematical Geosciences*, **43**, 681–693.

Pawlowsky-Glahn, V. and Buccianti, A., editors (2011). *Compositional Data Analysis: Theory and Applications*. Wiley.

Scealy, J. and Welsh, A. (2014). Colours and cocktails: Compositional data analysis 2013 lancaster lecture. *Australian & New Zealand Journal of Statistics*, **56**(2), 145–169.