Search or Article-id

(Help | Advance

All papers

Statistics > Methodology

A data-based power transformation for compositional data

Michail T. Tsagris, Simon Preston, Andrew T.A. Wood

(Submitted on 7 Jun 2011 (v1), last revised 16 Jun 2011 (this version, v2))

Compositional data analysis is carried out either by neglecting the compositional constraint and applying standard multivariate data analysis, or by transforming the data using the logs of the ratios of the components. In this work we examine a more general transformation which includes both approaches as special cases. It is a power transformation and involves a single parameter, α. The transformation has two equivalent versions. The first is the stay-in-the-simplex version, which is the power transformation as defined by Aitchison in 1986. The second version, which is a linear transformation of the power transformation, is a Box-Cox type transformation. We discuss a parametric way of estimating the value of α, which is maximization of its profile likelihood (assuming multivariate normality of the transformed data) and the equivalence between the two versions is exhibited. Other ways include maximization of the correct classification probability in discriminant analysis and maximization of the pseudo R-squared (as defined by Aitchison in 1986) in linear regression. We examine the relationship between the α-transformation, the raw data approach and the isometric log-ratio transformation. Furthermore, we also define a suitable family of metrics corresponding to the family of α-transformation and consider the corresponding family of Frechet means.

Submission history

Which authors of this paper are endorsers?

Download:

- PDF
- PostScript
- Other formats

Current browse cont

stat.ME
< prev | next >
new | recent | 1106

Change to browse b

stat

References & Citatio

- NASA ADS

Bookmark(what is this?)

ScienceWISE

Link back to: arXiv, form interface, contact.