Title:
Evaluating the performance of a Bayesian-multiplicative treatment of zeros in compositional count data sets.

Authors:
J.A. Martín-Fernández[a], K. Hron[b], M. Templ[c,d], P. Filzmoser[c], J. Palarea-Albaladejo[e]
[a] Department of Computer Science and Applied Mathematics, University of Girona, Campus Montilivi, P4, E-17071 Girona, Spain
[b] Department of Mathematical Analysis and Applications of Mathematics, Faculty of Science, Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic
[c] Department of Statistics and Probability Theory, Vienna University of Technology, Wiedner Hauptstraße 8-10, 1040 Vienna, Austria
[d] Department of Methodology, Statistics Austria, Guglgasse 13, 1110 Vienna, Austria
[e] Biomathematics & Statistics Scotland, JCMB, The King's Buildings, Edinburgh, EH9 3JZ, UK

Session:
CS04 Multivariate Data Analysis

Abstract:
Counts data are discrete vectors representing the numbers of outcomes falling into each of several mutually exclusive categories. Compositional techniques based on log-ratio methodology are appropriate in those cases where the total sum of the vector is not of interest. Such compositional count data sets usually contain zero values resulting from insufficiently large samples. That is, they refer to unobserved positive values that may have been observed with a larger number of trials. Because the log-ratio transformations require data with positive values, any statistical analysis of count compositions must be preceded by a proper replacement of the zeros. In recent years, a Bayesian-multiplicative approach, coherent with the vector space structure of the simplex, has been proposed for addressing the "zero problem". This treatment involves the Dirichlet prior distribution as the conjugate distribution of the multinomial and a multiplicative modification of the non-zero values in the vector of counts. Different parameterizations of the prior distribution provide different zero replacement results. Their performance will be evaluated from both theoretic and computational points of view.