Stewart, Connie

An approach to measure distance between compositional diet estimates containing essential zeros. (English) Zbl 07282089

J. Appl. Stat. 44, No. 7, 1137-1152 (2017).

Summary: For many applications involving compositional data, it is necessary to establish a valid measure of distance, yet when essential zeros are present traditional distance measures are problematic. In quantitative fatty acid signature analysis (QFASA), compositional diet estimates are produced that often contain many zeros. In order to test for a difference in diet between two populations of predators using the QFASA diet estimates, a legitimate measure of distance for use in the test statistic is necessary. Since ecologists using QFASA must first select the potential species of prey in the predator's diet, the chosen measure of distance should be such that the distance between samples does not decrease as the number of species considered increases, a property known in general as subcompositional coherence. In this paper we compare three measures of distance for compositional data capable of handling zeros, but not satisfying some of the well-accepted principles of compositional data analysis. For compositional diet estimates, the most relevant of these is the property of subcompositionally coherence and we show that this property may be approximately satisfied. Based on the results of a simulation study and an application to real-life QFASA diet estimates of grey seals, we recommend the chi-square measure of distance.

MSC:

62 Statistics

Keywords:

chi-square distance; compositional data; essential zeros; subcompositional coherence; QFASA

Software:

QFASA

Full Text: DOI

References:

[1] J. Aitchison, \textit{On criteria for measures of compositional difference}, Math. Geol. 24 (1992), pp. 365-379. · Zbl 0970.86531

[2] C.A. Beck, S.J. Iverson, W.D. Bowen, and W. Blanchard, \textit{Sex differences in grey seal diet reflect seasonal variation in foraging behavior and reproductive expenditure: Evidence from quantitative fatty acid signature analysis}, J. Animal Ecol. (2007), pp. 490-502.

[3] S.M. Budge, S.J. Iverson, W.D. Bowen, and R.G. Ackman, \textit{Among- and within-species variation in fatty acid signatures of marine fish and invertebrates on the Scotian Shelf, Georges Bank and southern Gulf of St. Lawrence}, Can. J. Fish. Aquat. Sci. 59 (2002), pp. 886-898.

[4] J. Egozcue and V. Pawlowsky-Glahn, \textit{Basic concepts and procedures}, in \textit{Compositional Data Analysis: Theory and Applications}, V. Pawlowsky-Glahn and A. Buccianti, eds., Wiley, New York, 2011, pp. 12-28.

[5] M. Greenacre, \textit{Log-ratio analysis is a limiting case of correspondence analysis}, Math. Geosci. 42 (2010), pp. 129-134.

[6] M. Greenacre, \textit{Compositional data and correspondence analysis}, in \textit{Compositional Data Analysis: Theory and Applications}, V. Pawlowsky-Glahn and A. Buccianti, eds., Wiley, New York, 2011, pp. 104-113.

[7] M. Greenacre, \textit{Measuring subcompositional incoherence}, Math. Geosci, 43 (2011), pp. 681-693.

[8] S.J. Iverson, C. Field, W.D. Bowen, and W. Blanchard, \textit{Quantitative fatty acid signature analysis: A new method of estimating predator diets}, Ecol. Monographs 72 (2004), pp. 211-235.

[9] D.A. Jackson, \textit{Compositional data in community ecology: The paradigm or peril of proportions?}, Ecology 78 (1997), pp. 929-940.

[10] R. Johnson and D. Wichern, \textit{Applied Multivariate Statistical Analysis}, Pearson Education, Inc., Upper Saddle River, NJ, 2007. · Zbl 1269.62044

[11] P.E. Kirsch, S.J. Iverson, and W.D. Bowen, \textit{Effect of a low-fat diet on body composition and blubber fatty acids of captive juvenile harp seals (Phoca groenlandica)}, Physiol. Biochem. Zool. 73 (2000), pp. 45-59.

[12] J. Martín-Fernández, C. Barceló-Vidal, and V. Pawlowsky-Glahn, Measures of difference for compositional data and hierarchical clustering methods, Proceedings of IAMG, Naples, 1998. · Zbl 1052.62531

[13] J.A. Martín-Fernandez, J. Palarea-Albaladejo, and R.A. Olea, \textit{Dealing with zeros}, in \textit{Compositional Data Analysis: Theory and Applications}, V. Pawlowsky-Glahn and A. Buccianti, eds., Wiley, New York, 2011, pp. 43-58.

[14] B.H. McArdle and M.J. Anderson, \textit{Fitting multivariate models to community data: A comment on distance-based redundancy analysis}, Ecology 82 (2001), pp. 290-297.

[15] J. Palarea-Albaladejo and J.A. Martín-Fernández, \textit{Values below detection limit in compositional chemical data}, J. Anal. Chim. Acta 764 (2013), pp. 32-43.

[16] J. Palarea Albaladejo, J.A. Martín-Fernández, and J. Soto, \textit{Dealing with distances and transformations for fuzzy }\(c\)-means clustering of compositional data, J. Classification 29 (2012), pp. 744-169. · Zbl 1360.62347

[17] D.A.S. Rosen and D.J. Tollit, \textit{Effects of phylogeny and prey type on fatty acid calibration coefficients in three pinniped species: implications for the QFASA dietary quantification technique}, Marine Ecol. Progress Ser. 467 (2012), pp. 263-276.

[18] C. Stewart and C. Field, \textit{Managing the essential zeros in quantitative fatty acid signature analysis}, J. Agric. Biol. Environ. Stat. 16 (2011), pp. 45-69. · Zbl 1306.62237

[19] C. Stewart, S. Iverson, and C. Field, \textit{Testing for a change in diet using fatty acid signatures}, Environ. Ecol. Stat. 21 (2014), pp. 775-792, doi:10.1007/s10651-014-0280-9.

[20] Y. Takane, W. Young, and J. De Leeuw, \textit{Non-metric individual differences multidimensional scaling: Alternating least squares with optimal scaling features}, Psycometrika 42 (1977), pp. 7-67. · Zbl 0354.92048

[21] M. Tsagris, S. Preston, and A. Wood, A data-based power transformation for compositional data, Proceedings of the Fourth International Workshop on Compositional Data Analysis, Girona, Spain, 2011.