

Hron, Karel; Kubáček, Lubomír

Statistical properties of the total variation estimator for compositional data. (English)

Zbl 1434.62093

Metrika 74, No. 2, 221-230 (2011).

Summary: The sample space of compositional data, a simplex, induces a different kind of geometry, known as Aitchison geometry, with the Euclidean space property. For this reason, the standard statistical analysis is not meaningful here, and this is also true for measures of location and covariance. The measure of location, called centre, is the best linear unbiased estimator of the central tendency of the distribution of a random composition with respect to the geometry on the simplex [*V. Pawlowsky-Glahn* and *J. J. Egozcue*, *Stoch. Environ. Res. Risk Assess.* 15, No. 5, 384–398 (2001; [Zbl 0987.62001](#)); *Math. Geol.* 34, No. 3, 259–274 (2002; [Zbl 1031.86007](#))]. Its covariance structure is described through a variation matrix, which induces the so called total variation as a measure of dispersion. The aim of the paper is to show that its sample counterpart has theoretical properties, corresponding to the standard multivariate case, like unbiasedness and convergence in probability. Moreover, its distribution in the case of normality on the simplex is developed.

MSC:

[62H11](#) Directional data; spatial statistics
[62E20](#) Asymptotic distribution theory in statistics
[62H10](#) Multivariate distribution of statistics

Cited in **2** Documents

Keywords:

Aitchison geometry; center of compositional data; unbiased and consistent estimator; normal distribution on the simplex; asymptotic distribution

Full Text: [DOI](#)

References:

- [1] Aitchison J (1986) The statistical analysis of compositional data. Monographs on statistics and applied probability. Chapman and Hall, London · [Zbl 0688.62004](#)
- [2] Aitchison J, Shen SM (1980) Logistic-normal distributions. Some properties and uses. *Biometrika* 67: 261–272 · [Zbl 0433.62012](#) · [doi:10.2307/2335470](#)
- [3] Anděl J (1978) Mathematical statistics (in Czech). SNTL/Alfa, Prague
- [4] Buccianti A, Pawlowsky-Glahn V (2005) New perspectives on water chemistry and compositional data analysis. *Math Geol* 37: 703–727 · [Zbl 1103.62111](#) · [doi:10.1007/s11004-005-7376-6](#)
- [5] Buccianti A, Mateu-Figueras G, Pawlowsky-Glahn V (eds) (2006) Compositional data analysis in the geosciences: from theory to practice. Geological Society, London (Special Publications), p 264 · [Zbl 1155.86002](#)
- [6] Daunis-i-Estadella J, Barceló-Vidal C, Buccianti A (2006) Exploratory compositional data analysis. In: Buccianti A, Mateu-Figueras G, Pawlowsky-Glahn V (eds) Compositional data analysis in the geosciences: from theory to practice, vol 264. Geological Society, London (Special Publications), pp 161–174 · [Zbl 1158.86333](#)
- [7] Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003) Isometric logratio transformations for compositional data analysis. *Math Geol* 35: 279–300 · [Zbl 1302.86024](#) · [doi:10.1023/A:1023818214614](#)
- [8] Egozcue JJ, Pawlowsky-Glahn V (2005) Groups of parts and their balances in compositional data analysis. *Math Geol* 37: 795–828 · [Zbl 1177.86018](#) · [doi:10.1007/s11004-005-7381-9](#)
- [9] Egozcue JJ, Pawlowsky-Glahn V (2006) Simplicial geometry for compositional data. In: Buccianti A, Mateu-Figueras G, Pawlowsky-Glahn V (eds) Compositional data analysis in the geosciences: from theory to practice, vol 264. Geological Society, London (Special Publications), pp 145–160 · [Zbl 1156.86307](#)
- [10] Filzmoser P, Hron K (2008) Outlier detection for compositional data using robust methods. *Math Geosci* 40: 233–248 · [Zbl 1135.62040](#) · [doi:10.1007/s11004-007-9141-5](#)
- [11] Filzmoser P, Hron K, Reimann C (2009) Principal component analysis for compositional data with outliers. *Environmetrics* 20: 621–632 · [doi:10.1002/env.966](#)
- [12] Glueck DH, Muller KE (1998) On the trace of a Wishart. *Commun Stat Theory Methods* 21: 2137–2141 · [Zbl 0907.62065](#)

- [13] Halmos PR (1974) Measure theory. Springer, New York · [Zbl 0283.28001](#)
- [14] Kshirsagar AM (1972) Multivariate analysis. Dekker, New York · [Zbl 0246.62064](#)
- [15] Mateu-Figueras G, Pawlowsky-Glahn V (2008) A critical approach to probability laws in geochemistry. *Math Geosci* 40: 489–502 · [Zbl 1153.86338](#) · [doi:10.1007/s11004-008-9169-1](#)
- [16] Mathai AM (1980) Moments of the trace of a noncentral Wishart matrix. *Commun Stat Theory Methods* 9: 795–801 · [Zbl 0462.62038](#) · [doi:10.1080/03610928008827921](#)
- [17] Mathai AM, Pillai KCS (1982) Further results on the trace of a noncentral Wishart matrix. *Commun Stat Theory Methods* 11: 1077–1086 · [Zbl 0521.62042](#) · [doi:10.1080/03610928208828294](#)
- [18] Pawlowsky-Glahn V, Egozcue JJ (2001) Geometric approach to statistical analysis on the simplex. *Stoch Envir Res Risk Ass* 15: 384–398 · [Zbl 0987.62001](#) · [doi:10.1007/s004770100077](#)
- [19] Pawlowsky-Glahn V, Egozcue JJ (2002) BLU estimators and compositional data. *Math Geol* 34: 259–274 · [Zbl 1031.86007](#) · [doi:10.1023/A:1014890722372](#)
- [20] Pawlowsky-Glahn V, Egozcue JJ, Tolosana-Delgado J (2008) Lecture notes on compositional data analysis. <http://hdl.handle.net/10256/297> . Accessed 17 Feb 2009
- [21] Reimann C, Filzmoser P, Garrett RG, Dutter R (2008) Statistical data analysis explained. Applied environmental statistics with R. Wiley, Chichester
- [22] Tolosana-Delgado R, Otero N, Pawlowsky-Glahn V, Soler A (2005) Latent compositional factors in the Llobragat river basin (Spain) hydrochemistry. *Math Geol* 37: 681–702 · [doi:10.1007/s11004-005-7375-7](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.