MAP ∞ PDMA

Seminar 2021/2022

Compositional data: some challenges in the world of Multivariate Statistics

Adelaide Freitas

CIDMA, Department of Mathematics, University of Aveiro

adelaide@ua.pt

January 14, 2022, 16h15, Sala 11.3.21

In the relative scale, 5% is a half of 10% and 45% forms a fraction of 0.9 of 50%. Obviously! However, in absolute scale, both comparisons produce the same difference. Whenever multivariate observations in a data set represent quantitative descriptions of the parts of some whole, conveying only relative information between parts, statistical techniques adequate to analyze compositional data should be used. Since compositional data are positive multivariate data with constant sum constraint, classic statistical methods (dealing with differences) can be not appropriate to be considered on them. Compositional data has emerged over the last years in numerous scientific fields. We illustrate some examples and list some difficulties to work with compositional data, namely in the area of exploratory multivariate statistics. We review some transformations proposed to overcumber the constraint imposed by the definition and discuss some challenges in the analysis of compositions of compositional data.