
Appendix: Some typical compositional problems

Contents

- A.1 Chemical compositions of Roman-British pottery
- A.2 Arctic lake sediments at different depths
- A.3 Household budget patterns
- A.4 Milk composition study
- A.5 A statistician's time budget
- A.6 The MN blood system
- A.7 Mammal's milk
- A.8 Calc-alkaline and tholeiitic volcanic rocks
- A.9 Concentration of minor elements in carbon ashes
- A.10 Paleocological compositions
- A.11 Pollen composition in fossils
- A.12 Food consumption in European countries
- A.13 Household expenditures
- A.14 Serum proteins
- A.15 Physical activity and body mass index
- A.16 Hotel posts in social media
- A.17 The waste composition in Catalonia
- A.18 Employment distribution in EUROSTAT countries

We present the reader with a series of challenging problems in compositional data analysis, with typical data sets and the questions they pose. These come from a number of different disciplines and will be used to elicit the concepts and principles of compositional data analysis. For further detailed information see [Ait86, Sections 1.1-1.14, p. 1-20].

A.1. Chemical compositions of Romano-British pottery

In archaeology, the compositional analysis of raw materials (clays used to make pottery, lithic materials used to make stone tools, etc.) has become a key tool for examining trade and exchange in ancient economies. Different sources for such materials often have distinct chemical 'signatures' that can be identified in places far from their point of origin. One interpretative challenge is to take an often large, complex array of chemical assays and identify patterns that can be exploited in higher level interpretations.

The `pottery` data set consists of data pertaining to the chemical composition of 45 specimens of Romano-British pottery. The method used to generate these data is atomic absorption spectrophotometry, and readings for nine oxides (Al_2O_3 , Fe_2O_3 , MgO , CaO , Na_2O , K_2O , TiO_2 , MnO , and BaO) are provided. These samples come from five different kiln sites, and one of the issues we want to consider is the degree to which compositional data help distinguish pottery from these various kilns.

 [Back to Index](#)

A.2. Arctic lake sediments at different depths

In sedimentology, specimens of sediments are traditionally separated into three mutually exclusive and exhaustive constituents —sand, silt and clay— and the proportions of these parts by weight are quoted as [sand, silt, clay] compositions. The `arctic.lake` data set records the [sand, silt, clay] compositions of 39 sediment samples at different water depths in an Arctic lake. Again we recognise substantial variability between compositions. Questions of obvious interest here are the following. Is sediment composition dependent on water depth? If so, how can we quantify the extent of the dependence? If we regard sedimentation as a process, do these data provide any information on the nature of the process? Even at this stage of investigation we can see that this may be a question of compositional regression.

 [Back to Index](#)

A.3. Household budget patterns

An important aspect in the study of consumer demand is the analysis of household budget surveys, in which attention often focuses on the expenditures of a sample of households on a number of mutually exclusive and exhaustive commodity groups and their relation to total expenditure, income, type of housing, household composition and so on. In the investigation of such data the pattern or composition of expenditures, the proportions of total expenditure allocated to the commodity groups, can be shown to play a central role in a form of budget share approach to the analysis. Assurances of confidentiality and limitations of space preclude