

## Finnish Society of Biostatistics: One-day CoDa-course

|  |
|--|
| <b>COURSE TITLE</b>  |
| Compositional Data Analysis ( <i>CoDaCourse</i> )  |
| <b>COURSE DURATION</b>   |
| 1 day (6-8 hours)  |
| <b>INSTRUCTOR</b>  |
| <b>TITLE:</b> PhD  |
| <b>NAME:</b> Pepus Daunis–i-Estadella  |
| <b>EMAIL:</b> <a href="mailto:pepus.daunis@udg.edu">pepus.daunis@udg.edu</a>   |
| <b>AFFILIATION:</b> Dept. of Computer Science, Applied Mathematics and Statistics (Univ. Girona)   |
| <b>COUNTRY OF RESIDENCE:</b> Spain   |
| <b>BIOGRAPHICAL SKETCH</b>   |
| <p>Pepus Daunis–i-Estadella has a PhD from the Polytechnic University of Catalonia working on “Inertias on Factorial Analysis” and a degree in Mathematics. Currently, he is Associate Professor at the Department of Computer Science, Applied Mathematics and Statistics of the University of Girona, Spain. His interests lie primarily in the statistical analysis of compositional data and applications to the Medicine. More information at <a href="http://imae.udg.edu/~pepus/">http://imae.udg.edu/~pepus/</a>.</p>  |
| <b>RESERVE INSTRUCTOR</b>  |
| <b>TITLE:</b> PhD  |
| <b>NAME:</b> Josep-Antoni Martín-Fernández   |
| <b>EMAIL:</b> <a href="mailto:josepantoni.martin@udg.edu">josepantoni.martin@udg.edu</a>   |
| <b>AFFILIATION:</b> Department of Computer Science, Applied Mathematics and Statistics (Univ. Girona)  |
| <b>COUNTRY OF RESIDENCE:</b> Spain   |
| <b>BIOGRAPHICAL SKETCH</b>   |
| <p>Josep-Antoni Martín-Fernández has a degree in Mathematics. He got his PhD from the Polytechnic University of Catalonia working on “Measurements of difference and non-parametric classification of CoDa”. Currently, he is Full Professor at the Department of Computer Science, Applied Mathematics and Statistics of the University of Girona, Spain. His interests lie primarily in the statistical analysis of compositional data. He focuses his research on the topics “Cluster Analysis of Compositional Data” and “Rounded Zeros and Missing Data”. More information at <a href="http://imae.udg.edu/~jamf/">http://imae.udg.edu/~jamf/</a></p>   |
| <b>COURSE DESCRIPTION</b>  |
| <p>The course will provide an introduction to the theoretical and practical aspects of the statistical analysis of compositional data (CoDa). It will provide mathematical background and an informal discussion forum on more advanced modelling methods and a help to practitioners to avoid common pitfalls in the analysis of high-dimensional CoDa sets). CoDa are vectors whose components show the relative importance of some parts of a whole. Typical examples are data presented in percentages, ppm, ppb, or the like. This type of data appears in most applications, and the interest and importance of consistent statistical methods cannot be underestimated. The log-ratio approach to CoDa was introduced back in the eighties. Since then, steady progress has been done in understanding the geometry peculiar to their sample space, the D-part simplex. The course will consist of lectures and exercises. Exercises are done with the freeware CoDaPack (<a href="http://imae.udg.edu/codapack/">http://imae.udg.edu/codapack/</a>). Some datasets and their particular problems will be presented, analysed and discussed interactively. Visit <a href="http://www.compositionaldata.com">http://www.compositionaldata.com</a> for further information.</p> |

## Finnish Society of Biostatistics: One-day CoDa-course

---

### OUTLINE

The proposed course will introduce the current state of the art in CoDa analysis and will cover the following topics:

1. Hypothesis underlying statistical data analysis (simplex sample space and scale).
2. The Aitchison geometry. Log-ratio coordinate representation. Sequential Binary Partition.
3. Exploratory analysis: centering, variation array, biplot and balances-dendrogram.
4. Working on coordinates: multivariate analysis of CoDa.
5. Pre-processing zero values.
6. Introduction to available software: CoDaPack and R-packages.

### EQUIPMENT REQUIREMENTS

Participants are encouraged to bring their own laptop with installed software for CoDa (CoDaPack <http://imae.udg.edu/codapack/>) for doing computer exercises.