Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 10, n. 1, 2023.

Trabalho apresentado no XLII CNMAC, Universidade Federal de Mato Grosso do Sul - Bonito - MS, 2023

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics
Preprint

A brief introduction to the geometry of mixed distributions for the study of fractional data

Alejandra Rada[1], Cristian Coletti[2] and Sandra Zapata[3]
CMCC - UFABC, Santo André, SP

In this work we propose to introduce the differential geometry arising in the study of mixed distributions used in the modeling of fractional data. Many authors focus on rates or proportions which are measured on the interval $(0, 1)$. However, there are many situations in which we should include at least one of the extreme points of $(0, 1)$.

For statistical modeling of continuously distributed data in the interval $(0, 1)$ which include observations at one or both extremes (called fractional data), we consider that the distribution of the data is a mix between a continuous distribution defined on the interval $(0, 1)$ and the Bernoulli distribution . In this situation, the model is part of the class of inflated models defined and explored for instance, in [1, 2].

The beta distribution is useful to model data that are measured continuously on the open interval $(0, 1)$; see [3] and [4]. Following [2], we consider a mixed continuous-discrete distributions to model data that are observed on [0, 1), (0, 1] or [0, 1]. The distributions are called *inflated beta distributions*.

We show that the class of inflated beta distribution is a manifold with coordinate system given by a simple function of its parameter.

We use divergence functions, differential geometry and its relationship with information theory to find properties of the inflated beta distribution that allow us to make inferences about fractional data.

Acknowledgement

This work is financially supported by São Paulo Research Foundation (FAPESP) under grant 2022/08948-2.

Referências

[1] A. Shun-ichi. **Information Geometry and Its Applications**. Springer Publishing Company, Incorporated, 2018. ISBN: 4431567437.

[2] S. Ferrari e R. Ospina. "Inflated beta distributions". Em: **Statistical Papers** 51 (2010), pp. 111–126. DOI: 10.1007/s00362-008-0125-4.

[3] N. Balakrishnan, N. Johnson e S. Kotz. **Continuous univariate distributions**. Vol. 1, 2nd Edition, John Wiley e Sons, 1995. ISBN: 0471584959.

[1] alejandra.rada@ufabc.edu.br
[2] cristian.coletti@ufabc.edu.br
[3] sandra.maria@ufabc.edu.br

[4] F. Cribari-Neto e S. Ferrari. "Beta Regression for Modelling Rates and Proportions". Em: **Journal of Applied Statistics** 31.7 (2004), pp. 799–815. DOI: 10.1080/0266476042000214501.