



Distribution and Symmetric Distribution Regression Model for Histogram-Valued Variables

Sónia Dias, Paula Brito

(Submitted on 25 Mar 2013)

Histogram-valued variables are a particular kind of variables studied in Symbolic Data Analysis where to each entity under analysis corresponds a distribution that may be represented by a histogram or by a quantile function. Linear regression models for this type of data are necessarily more complex than a simple generalization of the classical model: the parameters cannot be negative still the linear relationship between the variables must be allowed to be either direct or inverse. In this work we propose a new linear regression model for histogram-valued variables that solves this problem, named Distribution and Symmetric Distribution Regression Model. To determine the parameters of this model it is necessary to solve a quadratic optimization problem, subject to non-negativity constraints on the unknowns; the error measure between the predicted and observed distributions uses the Mallows distance. As in classical analysis, the model is associated with a goodness-of-fit measure whose values range between 0 and 1. Using the proposed model, applications with real and simulated data are presented.

Comments: 49 pages 17 figures

Subjects: **Methodology (stat.ME)**

Cite as: **arXiv:1303.6199 [stat.ME]**

(or **arXiv:1303.6199v1 [stat.ME]** for this version)

Submission history

From: Sonia Dias Mrs [[view email](#)]

[v1] Mon, 25 Mar 2013 16:27:57 GMT (703kb,D)

[Which authors of this paper are endorsers?](#)

Download:

- [PDF](#)
- [Other formats](#)

Current browse context:

stat.ME

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1303](#)

Change to browse by:

[stat](#)

References & Citations

- [NASA ADS](#)

Bookmark([what is this?](#))

