

Beta Regression: Shaken, Stirred, Mixed, and Partitioned

Achim Zeileis¹, Bettina Grün^{2,3*}, Francisco Cribari-Neto⁴

1. Department of Statistics, Universität Innsbruck
2. Institut für Statistik, Ludwig-Maximilians-Universität München
3. Institut für Angewandte Statistik, Johannes Kepler Universität Linz
4. Departamento de Estatística, Universidade Federal de Pernambuco

*Contact author: Bettina.Gruen@stat.uni-muenchen.de

Keywords: beta regression, finite mixture, proportion, rate, model-based recursive partitioning

The class of beta regression models is commonly used by practitioners to model variables that assume values in the open standard unit interval $(0, 1)$. It is based on the assumption that the dependent variable is beta-distributed with its mean and precision parameters depending on some regressors. We explore various flavors of parametrizations for the dependence of the response on the regressors. To shake and stir, either a single or double index model is employed for mean and/or precision, i.e., the parameters are related to potentially different sets of regressors through a linear predictor plus link function(s). This approach naturally incorporates features such as heteroskedasticity or skewness which are commonly observed in data taking values in the standard unit interval, such as rates or proportions. Furthermore, additional heterogeneity in the data can be captured by mixing or partitioning: If covariates are available that explain the heterogeneity, a model-based recursive partitioning approach can be employed. If not, latent class regression is an alternative. All flavors of models are implemented in the *R* package **betareg**, leveraging the **flexmix** and **party** packages.

References

- Cribari-Neto, F. and A. Zeileis (2010). Beta regression in *R*. *Journal of Statistical Software* 34(2), 1–24.
- Grün, B. and F. Leisch (2008). Flexmix version 2: Finite mixtures with concomitant variables and varying and constant parameters. *Journal of Statistical Software* 28(4), 1–35.
- Zeileis, A., T. Hothorn, and K. Hornik (2008). Model-based recursive partitioning. *Journal of Computational and Graphical Statistics* 17(2), 492–514.