

Pinheiro, Hildete P.; Maia, Rafael P.; Lima Neto, Eufrásio A.; Rodrigues-Motta, Mariana
Zero-one augmented beta and zero-inflated discrete models with heterogeneous dispersion for the analysis of student academic performance. (English) [Zbl 1437.62148](#)
Stat. Methods Appl. 28, No. 4, 749-767 (2019).

Summary: The purpose of this work is to present suitable statistical methods to study the performance of undergraduate students based on the incidence/proportion of failed courses/subjects. Three approaches are considered: first, the proportion of failed subjects is modeled considering a zero-one augmented beta distribution; second, discrete models are used to model the probability of failing subjects with logit link; third, incidence is modeled using regression for count data with log link and the logarithm of the total number of subjects as an offset. Zero-inflated versions are used to account for the excess of zeros in the data when appropriate and we also considered the heterogeneous dispersion parameter, when applicable. Overall, the zero-inflated negative binomial and zero inflated beta-binomial models, with regression on the mean and the dispersion parameters, present good measures of goodness of fit to the data. The database consists of records of Engineering major students who entered the State University of Campinas, Brazil, from 2000 to 2005. Entrance exam scores and demographic variables as well as socio-economic status are considered as covariates in the models.

MSC:

- 62G08 Nonparametric regression and quantile regression
- 62J12 Generalized linear models (logistic models)
- 62P25 Applications of statistics to social sciences

Keywords:

academic performance; heteroscedasticity; quantile residuals; residual analysis; overdispersion; zero-inflated discrete models; zero-one augmented beta models

Software:

GAMLSS; R

Full Text: [DOI](#)

References:

- [1] Adelfio G, Boscaino G (2016) Degree course change and student performance: a mixed-effect model approach. *J Appl Stat* 43(1):3-15
- [2] Adelfio G, Boscaino G, Capursi V (2014) A new indicator for higher education student performance. *High Educ* 68:653-668
- [3] Attanasio M, Boscaino G, Capursi V, Plaia A (2013) Can students' career be helpful in predicting an increase in universities income? In: Giudici P, Ingrassia A, Vichi M (eds) *Statistical models for data analysis. Studies in classification, data analysis and knowledge organization*. Springer, Switzerland, pp 9-16
- [4] Bayer F, Cribari-Neto F (2017) Model selection criteria in beta regression with varying dispersion. *Commun Stat Simul Comput* 46(1):729-746 · [Zbl 1364.62191](#)
- [5] Bianconcini S, Cagnone S (2012) Multivariate latent growth models for mixed data with covariate effects. *Commun Stat Theory Methods* 41:3079-3093 · [Zbl 1296.62119](#)
- [6] Birch ER, Miller P (2006) Student outcome at university in australia: a quantile regression approach. *Aust Econ Pap* 45:1-17
- [7] Chambers JM, Cleveland WS, Kleiner B, Tukey PA (1983) Tutorial on methods for interval-censored data and their implementation in R. In: *Graphical methods for data Analysis*. Wadsworth, Belmont, CA
- [8] Dunn PK, Smyth GK (1996) Randomized quantile residuals. *J Comput Graph Stat* 5:236-244
- [9] Ferrari SLP, Cribari-Neto F (2004) Beta regression for modeling rates and proportions. *J Appl Stat* 7:799-815 · [Zbl 1121.62367](#)
- [10] Grilli L, Rampichini C, Varriale A (2015) Binomial mixture modelling of university credits. *Commun Stat Theory Methods* 44:4866-4879 · [Zbl 1365.62476](#)
- [11] Grilli L, Rampichini C, Varriale A (2016) Statistical modelling of gained university credits to evaluate the role of pre-enrollment assessment tests: an approach based on quantile regression for counts. *Stat Model* 16:47-66

- [12] Maia RP, Pinheiro HP, Pinheiro A (2016) Academic performance of students from entrance to graduation via quasi u-statistics: a study at a brazilian research university. *J Appl Stat* 43(1):72-86
- [13] Masserini L, Bini M, Pratesi M (2017) Effectiveness of non-selective evaluation test scores for predicting first-year performance in university career: a zero-inflated beta regression approach. *Qual Quant Int J Methodol* 51(2):693-708
- [14] McCullagh P, Nelder JA (1992) *Generalized linear models*. Chapman & Hall, London · [Zbl 0588.62104](#)
- [15] Murray-Harvey R (1993) Identifying characteristics of successful tertiary students using path analysis. *Aust Educ Res* 20:63-81
- [16] Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *J Classif* 31:274-295 · [Zbl 1360.62344](#)
- [17] Ospina R, Ferrari SLP (2010) Inflated beta distributions. *Stat Pap* 51:111-126 · [Zbl 1247.62043](#)
- [18] Ospina R, Ferrari SLP (2012) A general class of zero-or-one inflated beta regression models. *Comput Stat Data Anal* 56:1609-1623 · [Zbl 1243.62099](#)
- [19] Pedrosa RHL, Dachs JNW, Maia RP, Andrade CY, Carvalho BS (2007) Academic performance, student's background and affirmative action at a Brazilian research university. *High Educ Manag Policy* 19(3):1-20
- [20] Pinheiro A, Sen PK, Pinheiro HP (2009) Decomposability of high-dimensional diversity measures: quasi u-statistics, martingales and nonstandard asymptotics. *J Multivar Anal* 100(8):1645-1656 · [Zbl 1190.62130](#)
- [21] Pinheiro A, Sen PK, Pinheiro HP (2011) A class of asymptotically normal degenerate quasi u-statistics. *Ann Inst Math Stat* 63:1165-1182 · [Zbl 1230.62069](#)
- [22] Pinheiro HP, Sen PK, Pinheiro A, Kiihl S (2018) A nonparametric approach to asses undergraduate performance. arXiv:181000678 [statME] (submitted)
- [23] R Core Team (2013) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- [24] Rigby RA, Stasinopoulos DM (2005) Generalized additive models for location, scale and shape. *J R Stat Soc Ser C (Appl Stat)* 54(3):507-554 · [Zbl 05188697](#)
- [25] Rigby RA, Stasinopoulos DM, Heller G, Voudouris V (2014) The distribution toolbox of GAMLSS. <http://www.gamlss.org/>
- [26] Salehi M, Roudbari M (2015) Zero inflated poisson and negative binomial regression models: application in education. *Med J Islam Repub Iran* 29:297-297

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.