Editorial
Volume 12:6, 2021

Journal of Biometrics & Biostatistics

ISSN: 2155-6180 Open Access

Explanatory variables

Gholamreza Hessian*

Department of Statistics, Payame Noor University, Iran

Introduction

Binomial regression is employed to assess the relationship between a binary response variable and other explanatory variables. Popular instances of binomial regression include examination of the etiology of adverse health states employing a case-control study and development of prediction algorithms for assessing the danger of adverse health outcomes (e.g., risk of a heart attack). In R, a binomial regression model can be fit using the glm() function. In this chapter, we demonstrate the following aspects of binomial regression, with R code, using real data examples: •To highlight the main components of a binomial model fitting using the glm() function•How to evaluate the modeling assumptions in binomial regression?•How to relax the assumptions once they are violated?

•How to fit binomial models for non-independent data?•How to develop and evaluate prediction models for binary response? The chapter is meant to be a quick, practical guide to binomial regression using R. We particularly envision the accompanying task view to be a useful resource on all topics closely related to binomial regression. Negative binomial regression – Negative binomial regression are often used for over-dispersed count data, that's when the conditional variance exceeds the conditional mean. It are often considered as a generalization of Poisson regression since it's an equivalent mean structure as Poisson regression and it's an additional parameter to model the over-dispersion. If the conditional distribution of the result variable is over-dispersed, the arrogance intervals for the Negative binomial regression are likely to be narrower as compared to those from a Poisson regression model. Poisson regression – Poisson regression is usually used for modeling count data. Poisson regression has a number of extensions useful for count models. Zero-inflated regression model – Zero-inflated models plan to account for excess zeros. In other words, two sorts of zeros are thought to exist within the data, "true zeros" and "excess zeros". Zero-inflated models estimate two equations simultaneously, one for the count model and one for the excess zeros. OLS regression – Count outcome variables are sometimes log-transformed and analyzed using OLS regression. Many issues arise with this approach, including loss of knowledge thanks to undefined values generated by taking the log of zero (which is undefined), also because the lack of capacity to model the dispersion. Linear and Logistic regressions are usually the primary algorithms people learn in data science. Due to their popularity, a lot of analysts even end up thinking that they are the only form of regressions. The ones who are slightly more involved think that they're the foremost important among all sorts of multivariate analysis. The truth is that there are innumerable sorts of regressions, which may be performed. Each form has its own importance and a selected condition where they're best suited to use. In this article, I even have explained the foremost commonly used 7 sorts of regression in data science during a simple manner. Through this article, I also hope that people develop an idea of the breadth of regressions, instead of just applying linear/logistic regression to every machine learning problem they come across and hoping that they would just fit!

How to cite this article: Yamamoto, Kouji. "Explanatory variables ." *jbmbs*12 (2021) : 6

***Corresponding author:** *Gholamreza Hessian, Department of Statistics, Payame Noor University, Iran, E- mail: gh.hesamian@pnu.ac.ir*

Received Date: June 01, 2021; **Accepted Date:** June 15, 2021; **Published Date:** June 22, 2021