Improving the interpretation of binary and count models.

Max A. Halvorson, Connor J. McCabe, Xiaolin Cao, David Huh, David Atkins, Kevin M. King

University of Washington

Non-linear outcome distributions are common in clinical science, because researchers study phenomena that are often discrete by nature. These include binary outcomes (such as treatment relapse, presence/absence of a diagnosis) or outcomes that are counts (such as number of symptoms or problem behaviors endorsed). Such models are estimated using the generalized linear model (GLM). However, due to non-linearity inherent within these models, methods of interpreting findings from these models do not translate directly from those of linear models (such as regression). Additional care and nuance is required to interpret these models with appropriate accuracy and clarity. Whether researchers who apply the GLM to binary and count distributions effectively communicate their results in an intuitive manner remains an open question. We reviewed the clinical science literature and found frequent examples (XX%) of researchers providing inaccurate interpretation of GLM results, and very few cases (XX%)v where findings were communicated in a way that non-experts could understand. The goal of the current manuscript is to provide a tutorial on interpreting model coefficients from binary and count models. These recommendations include computing and interpreting quantities of substantive interest produced by these models (such as predicted probabilities and counts), considering the effects of covariates included within these models that alter the substantive interpretation of these values, and creating visual displays that better communicate the results of these models. We introduce a free and easy to use web application allowing researchers to display model results in an easy-to-understand format.