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 discretely measured. These include binary outcomes (such as treatment relapse, presence/absence of a diagnosis) and outcomes that are counts (such as number of symptoms or problem behaviors endorsed). The general linear model and its extensions have long provided a flexible means of quantifying these non-linear distributions. However, due to inherent non-linearity in these models, the interpretation of their parameters is less straightforward than interpretation of linear models (such as regression). We assert that researchers who apply the GLM to binary and count distributions should strive to communicate their results in a manner that is both careful and accessible to a broad audience. We reviewed the clinical science literature and found that common practice (XX% of papers) is for researchers to provide odds ratios and risk ratios. Very few papers (XX%), however, communicated findings 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. We advocate for extracting quantities of direct substantive interest such as predicted probabilities and counts, and displaying these quantities in simple visual displays. We introduce a free and easy-to-use web application allowing researchers to display model results in an easy-to-understand format.