Dominance Analysis

Adon Rosen

Date: 2020-09-17

Question 1 - Purpose of a Dominance Analysis:

The dominance analysis, as described by Azen and Budescu (2003), was proposed to be more intuitive and limit possible misinterpretations when exploring variable importance in multiple regression models. The authors argue this by first presenting commonly used techniques used to analyze variable importance including standardized coefficients, squared partial correlations, and commonality analysis as well as the limitations of these techniques such as susceptibility to the influence of multicollinearity, order dependent effects, and difficulty interpreting when higher order interactions are present. In order to overcome these limitations the authors reintroduce Dominance Analysis (DA). DA is a technique proposed by Budescu (1993) which allows for the relative importance of a variable to be explored through all $2^{(p-2)}$ possible model permutations. This exhaustive procedure allows for the growth of a model to be explored throughout the addition of individual variables. This procedure affords insight into both the predictive and inferential processes of multiple regression. It can be used for the predictive application by finding possible variance explained across all models and attempting to find subsets that satisfy the experimenters desires. This allows for more affordable bias-variance tradeoffs by exploring for possible plateaus in additional variance explained throughout the model building procedure. It benefits inferential approaches by informing an experimenter when a variable can describe portions of variance and also when multicollinearity may be an issue among variables.

The authors also introduced language in order to benefit potential inferential procedures such as concepts of complete dominance, relative importance, conditional dominance, and also general dominance. Complete dominance details variables that across all model folds are more dominant than another variable. If there were three possible variables X_1 would be completely dominant over X_3 if the variance explained in models consisting of X_1 and $X_1 + X_2$ explain more variance than the same permutations exchanging X_1 for X_3. Relative importance details the additional variance explained when a variable is added to a model. Conditional dominance explains how a variable has a greater relative importance when the model is conditioned on other variables. Finally, general dominance details how all possible permutations of a model at a specific level explain greater variance than other possible variable permutations.

Question 2 - Example and explanation of a Dominance Analysis:

	model	level	fit	incom	hs	Urb
1	1	0	0.0000000	0.1881393	0.2180067	0.4588271
incm	incom	1	0.1881393	NA	0.0407708	0.2787241
hs	hs	1	0.2180067	0.0109034	NA	0.2534126
Urb	Urb	1	0.4588271	0.0080364	0.0125922	NA
Avl1	Average level 1	1	NA	0.0094699	0.0266815	0.2660684
inc+	incom+hs	2	0.2289101	NA	NA	0.2438976
in+U	incom+Urb	2	0.4668635	NA	0.0059442	NA
hs+U	hs+Urb	2	0.4714193	0.0013883	NA	NA
Avl2	Average level 2	2	NA	0.0013883	0.0059442	0.2438976
i++U	incom+hs+Urb	3	0.4728077	NA	NA	NA

This table details the variance explained when models were fitted predicting crime rate as functions of: average income for an area; number of individuals with a high school diploma; and finally how urban an area was (population density). Here we see that Urbanicity is completely dominant, High school is conditionally dominant over income at level 2, and is generally dominant over income as well at level 1 and 2.

Bibliography

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. Psychological Methods, 8(2), 129–148. https://doi.org/10.1037/1082-989x.8.2.129

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. Psychological Bulletin, 114(3), 542–551. https://doi.org/10.1037/0033-2909.114.3.542