Abstract:

Although there has been a growing literature in education research on practical guidance for calculating statistical power under education RCTs for continuous outcomes-and, in particular, for student achievement test scores-there is no comparable literature in the education area that systematically examines theoretical and empirical issues related to the calculation of statistical power for binary (0/1) outcomes. This paper will help fill this gap by developing a practical approach for calculating statistical power for binary outcomes under clustered designs where schools are the unit of random assignment. The paper will draw on the large statistical literature in this area in the public health and medical fields that often focus on binary outcomes as their primary outcomes. The paper will focus on the estimation of impacts on proportions using logistic regression methods and a generalized estimation equation (GEE) approach to adjust standard errors for the clustering of students within schools. Power formulas will be provided for models that both include and exclude baseline covariates. The empirical analysis will examine intraclass correlations for a range of binary outcomes using data from recently published IES-funded randomized control trials and other pertinent data sources, and will use these ICCs and the theoretical formulas to examine appropriate school sample sizes to achieve precise binary outcomes.