Centering categorical predictors in multilevel models

Haley E. Yaremych[1], Kristopher J. Preacher[1], & Donald Hedeker[2]

[1]Quantitative Methods, Department of Psychology & Human Development, Vanderbilt University; [2]Department of Public Health Sciences, University of Chicago

VANDERBILT UNIVERSITY

Background

- Categorical predictors are used ubiquitously in multilevel models across the fields of psychology, education, medicine, and organizational research.
- The topic of centering has been discussed at great length in the methodological literature concerning multilevel modeling.[1,2,3]
  - ➢ The effects of using uncentered, grand-mean-centered (CGM), and cluster-mean-centered (CWC) predictors are well-known, regarding parameter estimates and their corresponding interpretations.[4]
  - ➢ Problematically, this work has focused almost exclusively on continuous predictors.

Current Aims

1. **Centering:** to clarify how and why categorical predictors should be centered in multilevel models, and to demonstrate that, when categorical predictors are centered incorrectly, models will yield biased/inaccurate estimates.
2. **Parameter interpretation**: to explain how MLM coefficients resulting from centered categorical predictors should be interpreted. Namely slopes, as these have not been previously addressed, and with particular attention paid to multi-categorical predictors.

Literature Review

- We began by conducting a literature review to clarify whether and how categorical predictors are centered by applied researchers employing multilevel models.
- We observed inconsistency, a lack of rationale for centering decisions, and most commonly, no mention at all of centering categorical predictors.
- This suggests a need for a comprehensive resource that explicates how categorical predictors should be treated.

Aim 1: Centering

Logic and Algebra

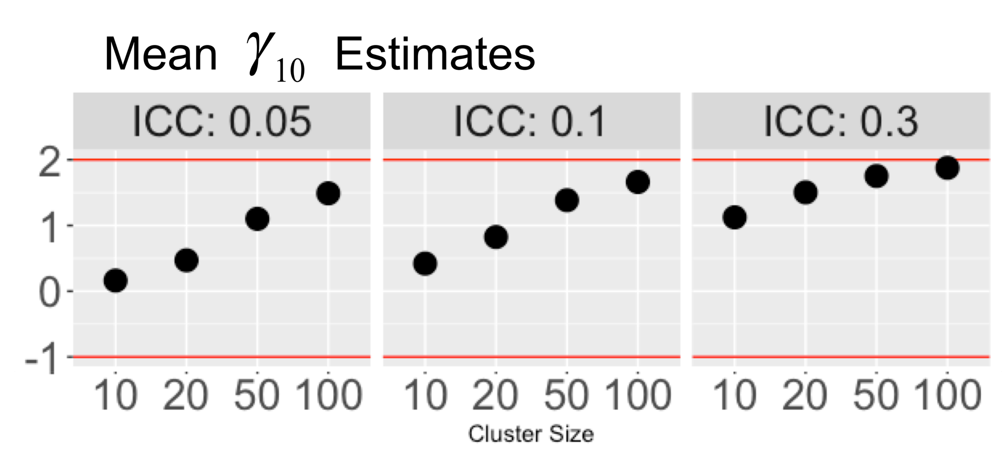
- A three-group categorical predictor represented by two dummy codes can be partitioned into its 'pure' within-cluster part and its 'pure' between-cluster part by the typical method of subtracting, and then reintroducing, each cluster mean:[1] $x_{1ij} = (x_{1ij} - \bar{x}_{1j}) + \bar{x}_{1j}$ and $x_{2ij} = (x_{2ij} - \bar{x}_{2j}) + \bar{x}_{2j}$
- Of note, cluster means now represent *proportions* of people in cluster *j* that belong to group 1 and group 2 of the multi-group predictor, and both the original dummy code and the CWC dummy code can take on just two possible values.
- Despite these differences, given that the within- and between-cluster components of a multi-categorical predictor are preserved, it follows that the consequences of various centering decisions should remain.

Simulation Study

- We generated multilevel datasets with a three-group categorical predictor and continuous outcome, fit two models, and recorded point estimates.
- Factors varied included cluster size and ICC(Y).
- Results for coefficients associated with the first dummy code are shown below; population-level within-cluster effect = 2 and between-cluster effect = -1.

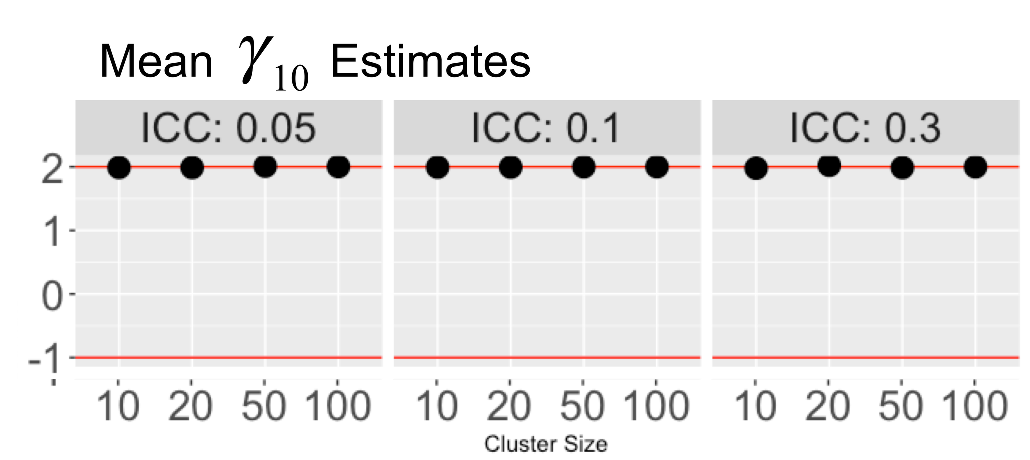
Model 1:

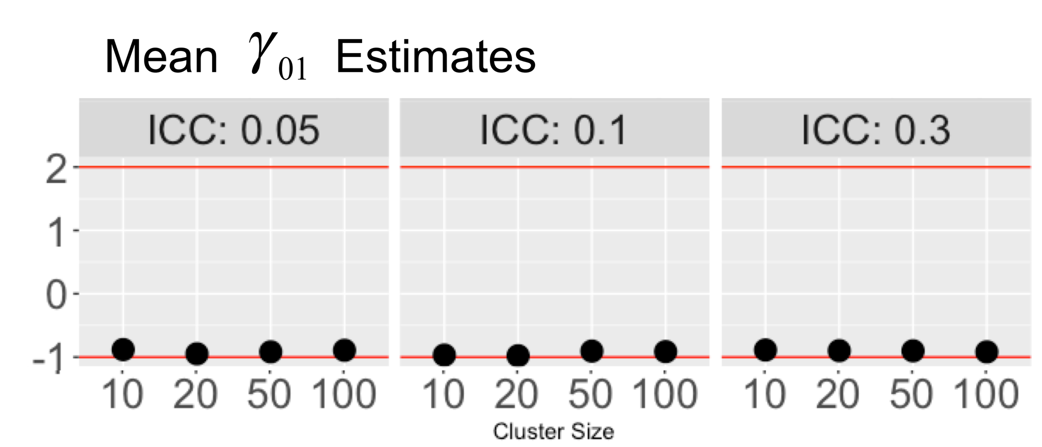
$$y_{ij} = \gamma_{00} + \gamma_{10}x_{1ij} + \gamma_{20}x_{2ij} + u_{0j} + e_{ij}$$



Model 2:

$$y_{ij} = \gamma_{00} + \gamma_{10}(x_{1ij} - \bar{x}_{1j}) + \gamma_{20}(x_{2ij} - \bar{x}_{2j}) + \gamma_{01}\bar{x}_{1j} + \gamma_{02}\bar{x}_{2j} + u_{0j} + e_{ij}$$





Results and Conclusions

- Model 1, which contained only uncentered dummy codes representing a three-group categorical predictor, yielded conflated slope estimates that were a meaningless blend of true within- and between-cluster effects.
- Model 2, which contained CWC dummy codes alongside cluster means as predictors, recovered unbiased estimates of the true within- and between-cluster effects.
- It appears that centering guidelines in place for continuous predictors should be implemented analogously for categorical predictors.
- Of note, the simulation study showed that the degree of conflation observed in Model 1 was dependent upon extraneous characteristics of the data, including cluster size, ICC(Y), and ICC of the categorical predictor.

Aim 2: Parameter Interpretation

Logic and Algebra

- We conducted expected-value derivations, a method drawn from the single-level regression literature. We began with Model 2 because, as shown, it elucidates 'pure' within- and between-cluster effects.
- Our goal was to derive the expected value of Y, separately for each group of the categorical predictor, in order to elucidate the relationship between slope coefficients and group mean differences on Y.
- Using a three-group categorical predictor represented by two dummy codes as our example, derivations yielded the following:
  - ➢ When there is *no* between-cluster variability with respect to the categorical predictor (i.e., ICC of all dummy codes = 0), group mean differences are equal to the within-cluster effects of the predictor:

$$E(y_{ij})|_{g1} - E(y_{ij})|_{g0} = \gamma_{10} \qquad E(y_{ij})|_{g2} - E(y_{ij})|_{g0} = \gamma_{20}$$

  - ➢ When there is *no* within-cluster variability with respect to the categorical predictor (i.e., ICC of all dummy codes = 1), group mean differences are equal to the between-cluster effects of the predictor:

$$E(y_{ij})|_{g1} - E(y_{ij})|_{g0} = \gamma_{01} \qquad E(y_{ij})|_{g2} - E(y_{ij})|_{g0} = \gamma_{02}$$

Conclusions

- Direct correspondences between coefficients and group mean differences under certain conditions serves to illuminate what the coefficients represent, and therefore how they should be correctly interpreted.
- The within-cluster effect of a dummy-coded categorical predictor should be interpreted as the mean difference on Y between members of group *k* and members of the reference group, within clusters, on average.
- The between-cluster effect should be interpreted as the mean difference on Y upon moving from a cluster composed entirely of reference group members to a cluster composed entirely of group *k* members.
- In practice, researchers often utilize uncentered dummy codes but interpret their effects as a "pure" within-cluster effect.
- In reality, such coefficients are meaningless blends of within- and between-cluster effects and therefore cannot be interpreted; the categorical predictor and its corresponding effects must be appropriately partitioned into level-specific parts before such interpretation is warranted.

References

1. Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods, 2nd edition* (Vol. 1). Sage.
2. Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, *12*(2), 121.
3. Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, *24*(5), 623-641.
4. Kreft, I. G., De Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate behavioral research*, *30*(1), 1-21.