Subject: Re: Multilevel model with region random effects
Posted by Anonymous on Wed, 29 Mar 2023 23:59:23 GMT
View Forum Message <> Reply to Message

Hi Tom,

I am facing a similar situation as Sophia and I found your answer to be quite logical. But I have seen some contradicting information in a previous DHS analytical study (here: https://dhsprogram.com/pubs/pdf/AS26/AS26.pdf) and I would like to get a clarification.

I am conducting an analysis where I plan to study associations between health system capacity (extracted from SPA survey) and length of stay in health facilities after birth (from the DHS). The SPA survey for this country is not a census, hence only a sample domain linkage is possible (DHS Spatial analysis report 10). Thus I plan to link the two databases using the administrative level variable (i.e. Region), and consequently conduct a multi-level modeling approach since the outcome is at the individual level, but some predictors are at the regional level. To quote Wang et. al : "Using the standard statistical approach, questions arise about the appropriate unit of analysis. Analyzing data at the individual level ignores the nesting of people within regions, which results in underestimating the standard errors and increasing the chance of incorrectly rejecting null hypotheses (Raudenbush and Bryk 2002). Alternatively, if the unit of analysis is the region, it will be difficult to include individual-level variables in the analysis. We can address these problems with multilevel modeling and allow for simultaneous investigation of the effects of group-level and individual-level predictors on individual-level outcomes."

In their report, Wang and colleagues do not go into the details of adjusting for sampling weights, and you might be familiar with the fact that running an melogit in Stata fails when the level2 variable is Region because there are no weights for this level.

I failed to find solutions online up to now. Do you have any suggestions on how to conduct this analysis without ignoring the hierarchical structure of the data, and still correctly adjusting for sampling weights? Would including "region" in the multivariable model be sufficient in this case?

Many thanks,
Aline