Subject: AIC and BIC model fit tests on weighted DHS data
Posted by nobleman00 on Sun, 29 Mar 2015 14:42:28 GMT
View Forum Message <> Reply to Message

Hello,

I am currently working on DHS household record (HR) datasets for multiple countries using Stata. The specific variables which I use are hv201 and hv205. For example, I am interested in looking at the proportions of the respondents who answered for each category for hv201 out of total respondents. To get the proportions, I first create dummy variables for each category of the variable and sum them up by collapsing. When collapsing, the DHS sample weight (hv005) is used. The proportions are then calculated by dividing the summed frequency of each dummy variable by the total frequency. These proportions are used as independent variables for a count model.
My question is that for count models in general, it is possible to compare models by using AIC or BIC fit tests. However, for weighted or clustered data, these tests are not recommended, and log-likelihoods are not available when running regression with the "svy" command in Stata. In my case, I do not need to use either the "svy" command or the [weight] option when running regression because the final dataset was already weighted when creating independent variables by collapsing. Thus, I can run regression as if I use a non-weighted dataset (although it was weighted in the previous step) in the software, and I do not have any issues on getting the AIC or BIC test results. However, I am wondering if these model fit tests (AIC, BIC) can be still used as valid tools for the model comparison in this case where the dataset was weighted in the previous step.
Thank you.

Jung