Subject: Weigthing and pooling multicountry datasets
Posted by Denise_kpebo on Thu, 02 Jun 2022 11:32:21 GMT
View Forum Message <> Reply to Message

Hi,

I am working on a multilevel logistic regression for multi-country (n=15) study using the DHS guide (https://www.dhsprogram.com/pubs/pdf/MR27/MR27.pdf) for the calculation of level weight.

Firstly, I have conducted the analysis within each country separately to determine which level of alpha should be used (e.g., allocation of variation in weights to the level-1 and level-2 units).
Then based on alpha for each country, I've weighted each dataset separately before merging the 15 countries datasets

Is this sufficient? May I start my analysis at this point with the pooled database using the svyset command?
Or do I still need to apply additional weights considering the differences in population size, since the population are not the same? And if so, how can I do that?

Many thanks in anticipation of your help

Subject: Re: Weigthing and pooling multicountry datasets
Posted by Janet-DHS on Fri, 03 Jun 2022 12:45:15 GMT
View Forum Message <> Reply to Message

Following is response from DHS Research & Data Analysis Director, Tom Pullum:

There have many postings on how to weight pooled surveys. There are basically two options. The first is to re-scale the weight for each country so that the total weight for country X is proportional to the population of country X at the time of the survey. You can get the estimated population size from the UN Population Division website, World Population Prospects 2019. The second option is to rescale so that the total weight is the same for each country (or survey). That is, if you pool 10 countries, you re-scale so that each survey has 1/10 of the total weight. Specific steps for both options have been given on the forum.

The first option has the problem that typically one large country, such as India or Nigeria, will completely dominate the results.

At DHS we often leave the weights alone and pool surveys into a single file just to simplify the data processing. We give results separately for each survey, but do not give results for all the surveys combined. Pooled surveys, from different countries and different years, do not describe a well-defined population. It's very hard to interpret a mean or percentage or coefficient from a mix of different surveys.

Subject: Re: Weigthing and pooling multicountry datasets
Posted by Denise_kpebo on Wed, 08 Jun 2022 02:45:13 GMT
View Forum Message <> Reply to Message

Thanks so much for your help on this

I'm going to use the first option and I have the population size for each country but this is my first time to weight data based on country population size and I actually have no idea of the state command to be used, could you please help me with that one last time?

Best

Subject: Re: Weigthing and pooling multicountry datasets
Posted by Janet-DHS on Thu, 09 Jun 2022 12:38:37 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

If you search through previous forum responses you will find examples of how to do this.

Subject: Re: Weigthing and pooling multicountry datasets
Posted by jamwiggly on Sat, 13 May 2023 16:01:02 GMT
View Forum Message <> Reply to Message

I was searching but maybe those posts have been deleted. :(
coreball