Subject: Population mean from Sample - NFHS 1 and 2
Posted by Niranjana on Mon, 18 May 2020 09:06:37 GMT
View Forum Message <> Reply to Message

Hi Everyone,

I have been working with the DHS's India Women's Birth Recode data rounds I and II. I was able to create a woman-year unbalanced panel where a woman enters the panel at the birth of the first child and exits at the time of interview. So using this now I am able to get a retrospective birth history of the woman. Now when I combine this data with information from the harmonized district data on NFHS provided by IPUMS, I can get the district of residence of the woman. I want to now generate some population means for the state/district using sampling strategy. For example, I have the number of people in the sample reporting to be sterilized in a particular year(v316), so I can generate the number of people in the survey who report being sterilized in a state in a particular year. Is there anyway, I can generate a population mean of people being sterilized in say year 1985 using the NFHS BR data by combining it with other population covariates (Population growth rates, NFHS sampling strategy, sample mean etc)?

Do let me know if there has been a precedent to calculate this population means from sample mean and if yes, how to go about it. Thanks so much!

Best regards,
Niranjana

Subject: Re: Population mean from Sample - NFHS 1 and 2
Posted by Bridgette-DHS on Wed, 20 May 2020 13:40:35 GMT
View Forum Message <> Reply to Message

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

You are making this question far more complicated than it needs to be. A term such as "unbalanced panel" is completely unnecessary. DHS surveys are cross-sectional and include some retrospective data. That's sufficient. In the NFHS1 (the 1992-93 survey) I find a variable that is year of sterilization, v316. Of the 25,581 (unweighted) women in that survey who had been sterilized, 1819, or 7.11%, were sterilized in 1985. I assume that the subsequent NFHS's include a similar question. Instead of district, I will refer to state; district is not included in all the surveys and only the NFHS4 had sufficient cases for district-level analysis. Here's the sort of thing that you could do for a state, specifically Uttar Pradesh, which had code v024=24 in the NFHS1. In Stata, open the IR file and enter the following three lines:
"gen Y=0 if v024==24" and "replace Y=1 if Y==0 & v316==85" and "tab Y". You get this:

That is, 1.35% of the women age 15-49 living in Uttar Pradesh in 1992-93 were sterilized in 1985. You could take this as an estimate of the percentage of women age 15-49 living in Uttar Pradesh

in 1985 who were sterilized in that year, which I believe is what you are looking for. However, between 1985 and 1992-93 some women died, and all of those who survived got older, and some women who were sterilized in Uttar Pradesh were not living in Uttar Pradesh at the time of the survey, and some women living in Uttar Pradesh in 1992-93 were not living there in 1985, when they were sterilized. There are various ways in which you could refine the estimate, but you would never be able to match something that might come out of a health management information system (HMIS). NFHS4 is the only one for which I would try to produce district-level estimates, and they would have wide confidence intervals.

File Attachments

1) tab.png, downloaded 419 times