Subject: How to handle continuous variable in constructing the index
Posted by gyamada on Sun, 29 Nov 2015 14:56:46 GMT

Hi,

I am reviewing the document by Dr. Rutstein and SPSS syntax files, and I have questions about how to handle continuous variables in constructing wealth index.

My first question is about how to handle missing values of continuous variables. May I know why we use mean instead of median? I feel that land area and number of household per sleeping room are usually skewed and thought that median might be another option.

My next question is about number of household per sleeping room. May I know why the number is obtained by rounding down to the nearest integer instead of leaving the decimals?

Faithfully,

Goro

Subject: Re: How to handle continuous variable in constructing the index
Posted by Liz-DHS on Thu, 03 Dec 2015 19:45:37 GMT

Dear User,
Here is a response from Dr. Shea Rutstein:
Quote:1. Use of mean substitution instead of median:
a. For land size, most households have no land so the median may end up as zero. However, that variable is the result of two questions, whether the household owns agricultural land and if so, what the size of the land is. The missing and unknown responses for land size are for those with land from the first question and so the median should not be zero.
b. For number of household members per sleeping room, the distribution is not very skewed so the mean and median should produce similar results. In any case, the missing data would come from number of sleeping rooms, not number of household members, and almost all households have less than 10 sleeping rooms.
2. The use of integers for number of members per sleeping room is for convenience when tabulating the frequency distribution and means of wealth by number of members per sleeping room to check on the results. Alternatively, a continuous variable can be and has been used in a few surveys, but hardly any difference in results has been found.
3. One should remember that the wealth index is not an absolute measure of net worth, and is derived from a sample of households, and as such is subject to sampling variance and non-sampling error. Additionally, the wealth index also depends to a small extent on the assets, dwelling materials, amenities and other indicators chosen for the survey. Small differences in household wealth score due to handling of continuous variables will not invalidate the relative position of households. However, some non-linearity is taken into account of these variables for number of farm animals by type, and will be tried for land size as well.

Subject: Re: How to handle continuous variable in constructing the index
Posted by gyamada on Sat, 12 Dec 2015 08:46:37 GMT

Thank you, Dr. Rutstein, for your information!

Following up one of your comments on skewed data (number of farm animals and land area), I would like to know what type of consideration could be more commonly thought. Could log transformation or truncation at certain value/percentile could be an option?

Sincerely,

Goro

Subject: Re: How to handle continuous variable in constructing the index
Posted by Hassen on Mon, 07 May 2018 15:09:02 GMT

Thanks all!!