
Subject: Re: Help-Seeking Behaviour of Domestic Violence_Myanmar DHS (2015-16)

Posted by [Janet-DHS](#) on Wed, 31 Jul 2024 18:33:34 GMT

[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member, Tom Pullum:

You are using the weights correctly. v101 is a duplicate of v024--both are "region"--and region x urban/rural defines the strata in the sample design. In the design of the sample, the small strata were over-sampled and the large strata were under-sampled, relatively speaking. If you look at the unweighted distribution of v101, the number of cases is nearly the same in each region, varying between about 751 and 1039. However, the weighted distribution, which is proportional to the actual size in the population, ranges from about 65 to 1649. By sampling approximately the same number of cases in each region, the sample is made more efficient. The standard errors of estimates are relatively equal across regions or strata. I repeat that what you observe is ok. (For most other variables, weighted and unweighted frequencies are typically within 10% to 20% of each other.)

We recommend using weights for all estimates. Indeed, we recommend a full adjustment for the sampling design, taking into account clustering, stratification, and weights, for everything except some initial exploratory analysis. If you don't use weights, it becomes almost impossible to understand changes and differences between surveys and sub-populations.

The argument against weights comes from some econometricians who believe that their models are correctly specified and whose main interest is in test statistics and p-values. I believe that statisticians and demographers uniformly advocate the use of weights to compensate for variation in sampling fractions in complex surveys. With cross-sectional DHS data, our models cannot be correctly specified and we are usually more interested in estimates (including confidence intervals) than in tests.
