

---

Subject: Post-stratification for DHS data

Posted by [sd535](#) on Thu, 17 Nov 2016 03:36:00 GMT

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

---

Hi,

I wish to post-stratify the Bangladesh DHS data at district-level according to the total number of children per district extracted from the recent census.

I use the following code for post-stratification. However, I didn't get any change after post-stratification. Am I doing any wrong?

Preliminary Design:

```
DHSdesign <- svydesign(id = child.data.HAZ$V001, strata=child.data.HAZ$V023, weights =
child.data.HAZ$V005/1000000, data=child.data.HAZ)
```

```
ps.weights <-
data.frame(
  CODIST = c( PostStrata.Wt.Cal$dist.id ) , # District ID (relates to CODIST )
  Freq = c( PostStrata.Wt.Cal$children.census ) # Number of children under 5
)
```

```
DHSdesign.Post.Strata <-
postStratify(
  DHSdesign ,
  strata = ~CODIST ,
  population = ps.weights
)
```

```
DE.MEAN.HAZ.District<-svyby(~HW70, ~CODIST, DHSdesign, svymean)
PS.DE.MEAN.HAZ.District<-svyby(~HW70, ~CODIST, DHSdesign.Post.Strata, svymean)
SE.DE.MEAN.HAZ.District<-svyby(~HW70, ~CODIST, DHSdesign, svymean)[,3]*100
PS.SE.DE.MEAN.HAZ.District<-svyby(~HW70, ~CODIST, DHSdesign.Post.Strata,
svymean)[,3]*100
```

```
> summary(DE.MEAN.HAZ.District)
  Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
-241.6 -185.3 -162.2 -167.6 -146.8 -112.0
```

```
> summary(PS.DE.MEAN.HAZ.District)
  Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
-241.6 -185.3 -162.2 -167.6 -146.8 -112.0
```

```
> summary(SE.DE.MEAN.HAZ.District)
  Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
  0.0  935.5 1326.0 1298.0 1612.0 2536.0
```

```
> summary(PS.SE.DE.MEAN.HAZ.District)
  Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
  0.0  935.5 1326.0 1298.0 1612.0 2536.0
```

There is no difference. And the most important problem is that the SE for one area is zero.

Can you please inform me where I am doing the wrong.

---

Subject: Re: Post-stratification for DHS data  
Posted by [Bridgette-DHS](#) on Thu, 17 Nov 2016 17:19:24 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Quote:I'm sorry, but I don't use SPSS, and I can't figure out what you did.

I would calculate the total number of weighted cases that you want to have in each district. I think of these as "target" totals. They are obtained by multiplying the total number of cases (unweighted) in your file by the proportions that are in each district in the census. You then multiply the weight (v005) in each district by the ratio of the target total to the current weighted number of cases in the district.

I will illustrate how you would do this in Stata to obtain a uniform distribution of weighted cases across the seven region (v024):

```
set more off
use e:\DHS\DHS_data\KR_files\BDKR70FL.dta, clear
sort v024
save e:\DHS\DHS_data\scratch\BDtemp.dta, replace
```

```
keep v024 v005
```

```
gen wtd_n=v005/1000000
gen unwtd_n=1
collapse (sum) *wtd_n, by(v024)
list, table clean
```

\* adjust v005 so that the total weight will be the same in each region

```
summarize unwtd_n
scalar stotal=r(sum)
scalar list stotal
gen target=stotal/7
gen v005_factor=target/wtd_n
list, table clean
keep v005_factor v024
```

```
sort v024
merge v024 using e:\DHS\DHS_data\scratch\BDtemp.dta
```

tab \_merge

```
gen v005_rewtd=round(v005*v005_factor)
```

\*check that the new distribution matches the targets

```
gen wtd_n=v005/1000000
```

```
gen rewtd_n=v005_rewtd/1000000
```

```
gen unwtd_n=1
```

```
collapse (sum) *wtd_n, by(v024)
```

```
list, table clean
```

---

Subject: Re: Post-stratification for DHS data  
Posted by [sd535](#) on Sun, 11 Dec 2016 13:35:59 GMT

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

Dear Tom Pullum,

We followed your instruction and get results exactly same with/without post-stratification. For your convenience, we are sending you the data with stata code and a theoretical statement in a pdf file. Can you please instruct us whether we are in right track to provide some statistics (with SE) at district-level?

Regards,  
Sumon

#### File Attachments

- 1) [Post Stratification for BDHS.rar](#), downloaded 427 times
- 2) [Post-stratification Weight Calculation.pdf](#), downloaded 464 times

---

Subject: Re: Post-stratification for DHS data  
Posted by [Bridgette-DHS](#) on Tue, 13 Dec 2016 20:36:40 GMT

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

Following is a response from DHS Senior Research Associate, Shireen Assaf:

Dear Sumon,

Following Tom's instructions, you forgot to perform one step which is important to have the total weight same in each district (i.e. divide by the number of districts which is 64). I performed this step in the attached do file, and when I use the svy with the different weights I obtain different standard errors.

Please change the paths in the do file back to your paths.

Thank you.

Best regards,

Shireen Assaf

## File Attachments

---

1) [DistrictSpecificWeight\\_Syntax.do](#), downloaded 614 times

---