Subject: incorrect stunting rates
Posted by margovg on Wed, 07 Mar 2018 13:45:31 GMT
View Forum Message <> Reply to Message

Hello,

Im trying to calculate the prevalence of stunting using the Burundi 2010 DHS dataset.
I have used the PR file and - presumably- the correct survey settings, the overall percentage that I have calculated is correct (57,7% stunted), however, when I try to calculate the sex specific prevalence I'm of by a bit.

I am using R:

```
# making a variable for stunting
data2$stunt <- ifelse(data2$hc70 < -200,'stunted','not_stunted')

#survey settings

int.design <- function (){
  data.w <<-
    svydesign(
      id = ~ hv001 ,
      data = data ,
      weight = ~ hv005 ,
      strata = ~ hv022)
}

int.design()

> svyby(~stunt, ~hv104, data.w, svymean)
       hv104 stuntnot_stunted stuntstunted se.stuntnot_stunted se.stuntstunted
male     male       0.3750139    0.6249861         0.01372043      0.01372043
female female       0.4729965    0.5270035         0.01436399      0.01436399
```

According to the DHS report 62,1% of boys is stunted and 53,1% of the girls. As you can see from the output above, my percentages come down to 52,7% for girls and 62,5% for boys.

I have deleted from my dataset:
- Children with missing data for variable HC70
- Children with HC70>9000
- Children who did not sleep in the hh the night before the survey(hv103==1)

Can anybody tell me what I doing wrong, and how I can get the correct percentages?

Kind regards,

Margo

Subject: Re: incorrect stunting rates
Posted by Trevor-DHS on Wed, 07 Mar 2018 20:09:16 GMT
View Forum Message <> Reply to Message

I'm matching the numbers from the report with the following code:

```
library(foreign)
library(survey)
# switch to directory where my data exists
setwd("C:/Data/DHS_stata/")

# read the PR dataset
data <- read.dta("BUPR61FL.dta", convert.factors = TRUE)

# subset for children measured with valid measures who are de facto
data2 <- data[!is.na(data$hc70) & data$hc70<9000 & data$hv103==1,]

# making a variable for stunting
data2$stunt <- ifelse(data2$hc70 < -200,'stunted','not_stunted')

# create weight variable
data2$wt <- data2$hv005/1000000

# survey settings
data.w <- svydesign(id = ~ hv001, data = data2, weight = ~ wt, strata = ~ hv022)

# tabulate stunting by sex
svyby(~stunt, ~hv104, data.w, svymean)
      hv104 stuntnot_stunted stuntstunted se.stuntnot_stunted se.stuntstunted
male   male       0.3790181    0.6209819         0.01352742      0.01352742
female female     0.4687265    0.5312735         0.01374611      0.01374611
```

Subject: Re: incorrect stunting rates
Posted by Nicholus Tint Zaw on Sun, 11 Mar 2018 15:38:08 GMT
View Forum Message <> Reply to Message

H, I am Nicholus from Myanmar and I am also experiencing same issued as margovg did before.

I am now working on child nutrition status data analysis and used KR at first and use following stat

code to identify the mean HAZ score and stunting rate of children. But my result was quite different in term of denominator with the number from DHS Myanmar report. (4,213 in my analysis and 4,089 in DHS report)

[global
global kr_child D:\SPA SCI\Technical Knowledge\GOVERMENT INFO\DHS 2015-16\dataset\MM_2015-16_DHS_05022017_255_106315\mmkr71dt\ MMKR71FL.dta

global pr_person D:\SPA SCI\Technical Knowledge\GOVERMENT INFO\DHS 2015-16\dataset\MM_2015-16_DHS_05022017_255_106315\mmpr71dt\ MMPR71FL.dta

*---------- ----------

use "${kr_child}", clear

** Calculate Stunting Rate **
** Use DHS HAZ Variable **

tab hw70, m

gen haz_dhs_revised = hw70
replace haz_dhs_revised = .n if hw70 >= 9996
replace haz_dhs_revised = hw70/100 if !mi(haz_dhs_revised)
tab haz_dhs_revised, m

gen stunt = (haz_dhs_revised < -2)
replace stunt = .n if mi(haz_dhs_revised)
tab stunt, m

** Construct WEIGHT var using Women Insidividual Sample Weight **
gen wgt = v005/1000000
tab wgt, m

lookfor sampling // search for samping unit var

** Generate Stunting Rate & HAZ Mean Score by Weighted Data **
svyset, clear
svyset [pw = wgt], psu(v021) strata (v022)
svy: mean haz_dhs_revised stunt
svyset, clear][/code]

Then, when I found this post in user forum, I revised my code and dataset usage with PR file (as follow). But, still got the different figure compare with DHS report and this time the denominator discrepancy became more larger as my analysis figure become 4,640.

[code][/clear

```
use "${pr_person}", clear

** Check with DHS Calculated HAZ Variable **

tab hc70, m

gen haz_dhs_revised = hc70
replace haz_dhs_revised = .n if hc70 >= 9996
replace haz_dhs_revised = hc70/100 if !mi(haz_dhs_revised)
replace haz_dhs_revised = .n if hv103 != 1
tab haz_dhs_revised, m

tab hv103, m // should be 1

** Construct WEIGHT var using Women Insidividual Sample Weight **
gen wgt = hv005/1000000
tab wgt, m

lookfor sampling // search for samping unit var

gen stunt = (haz_dhs_revised < -2 & hv103 == 1)
replace stunt = .n if hv103 != 1 | mi(haz_dhs_revised)
tab stunt, m

svyset, clear
svyset [pw = wgt], psu(hv021) strata (hv022)
svy: mean haz_dhs_revised stunt
svyset, clear

]
```

So, I was wondering there is anyone who can review my code and provide guidance to fix this issue. Thanks in advance.

Best regards,
Nicholus

Subject: Re: incorrect stunting rates
Posted by Trevor-DHS on Thu, 15 Mar 2018 15:16:30 GMT
View Forum Message <> Reply to Message

In the output you get, the number of observations is 4640, but this is the unweighted number of cases. If you look at the Population Size, this is the weighted number of cases and should match the 4089 in the report.

Subject: Re: incorrect stunting rates
Posted by Nicholus Tint Zaw on Thu, 15 Mar 2018 15:25:04 GMT
View Forum Message <> Reply to Message

Yes, the weighted population size was identical with DHS publication report.
So, for further analysis, I should check with that weight population size and DHS report instead of # of observation. right?

Thanks a lot for your reply.

best regards,
Nicholus

Subject: Re: incorrect stunting rates
Posted by Trevor-DHS on Thu, 15 Mar 2018 15:26:48 GMT
View Forum Message <> Reply to Message

Yes, the DHS report presents the weighted number of cases which are given as the weighted population size in the Stata output, and not the unweighted number of observations

Subject: Re: incorrect stunting rates
Posted by Nicholus Tint Zaw on Fri, 16 Mar 2018 08:43:49 GMT
View Forum Message <> Reply to Message

Thanks a lot.

By the way, one last question remained. Why we used PR data file instead of KR for this child stunting indicator.
To my knowledge, KR is the children dataset file. right?

best regards,
Nicholus

Subject: Re: incorrect stunting rates
Posted by Trevor-DHS on Fri, 16 Mar 2018 13:54:19 GMT
View Forum Message <> Reply to Message

KR is the children's recode file for children of interviewed women. The PR file includes all people who are usually resident in the household (de jure) or stayed in the household the prior night (de facto). By analyzing the children based on the PR file we include children who do not live with their mother or whose mother has died, which would be excluded if you used the KR file.

Subject: Re: incorrect stunting rates
Posted by Nicholus Tint Zaw on Fri, 16 Mar 2018 14:49:48 GMT
View Forum Message <> Reply to Message

Thanks for your kind explanation. Now, it is very clear to me.

Subject: Re: incorrect stunting rates
Posted by anneclaireclaire on Wed, 08 May 2019 15:31:44 GMT
View Forum Message <> Reply to Message

Dear Trevor, I am trying to match the stunting numbers for the RW (Rwanda) DHS 2014-2015 but it does not work . Can you please help me . I am using the following code
Thanks in advance

```
library(haven)
library(data.table)
RWPR70FL <- as.data.table(read_dta("~/Data/RW/2015_DHS/RWPR70FL.DTA", convert.factors = TRUE)))

RWPR70_2 <- RWPR70FL[!is.na(hc70) & hc70<9000 & hv103==1,]


RWPR70_2$stunt2 <- cut(RWPR70_2$hc70,
                       breaks=c(-Inf, -200, Inf),
                       labels=c("stunted","not sunted"))

RWPR70_2[,wt:= hv005/1000000]
RWPR70_2[, district:=as.factor(shdistrict)]

RWPR70.w2 <- svydesign(id = ~ hv001, data = RWPR70_2, strata = ~ hv022, weights=~wt)
svyby(~factor(stunt2), RWPR70_2$district, RWPR70.w, svymean, verbose=TRUE)
```