Subject: Computing Newborn Mortality Proportions
Posted by Hanu on Fri, 08 Jul 2016 07:58:49 GMT
View Forum Message <> Reply to Message

Hello,

I am using stata to analyze some of Ethiopia's 2011 DHS data. I am using the Birth Recode file and I am interested in creating a dichotomous variable that shows the new born mortality proportion. The numerator of this variable would show the proportions of all children who died at day 30 (or month 1) or below and the denominator would be all children (living or dead) who would be over 30 days had they survived. I used this code to generate the proportion:

gen age = v008-b3
gen nbmort = b6<=201 if age>30

nonetheless the proportion of new born mortality I got was only 6% whereas in the DHS report this number is around 65% for 0-4 years preceding the survey (Table C.6). I would like to replicate some of the information contained in this table and appreciate any help you may provide with this regard.

Thank you

Subject: Re: Computing Newborn Mortality Proportions
Posted by Bridgette-DHS on Fri, 08 Jul 2016 18:06:07 GMT
View Forum Message <> Reply to Message

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

First, you are mixing days and months. v008-b3 is months since the birth. It's not a good idea to refer to this as age, because it can only be interpreted as age for children who have survived. When you restrict to this being >30, you are selecting births that were more than 30 months before the survey, not more than 30 days.

Neonatal deaths can occur in either the month of birth or the next month. To take account of censoring, which is very minor, I would restrict to births with v008-b3>1.

Second, there are two b variables for age at death, b6 and b7. b6 is more detailed and must be used for deaths in the first day or the first week but for deaths in the first month we normally use b7, which is a recode of b6, and specifically b7=0, which identifies deaths at less than one month. Someone could make a fuss over the use of "one month" rather than "28 days" but the error due to this difference is negligible when compared with other measurement errors for these estimates. If you do, say, "tab b6 if b7==0" you will see the correspondence between b6 and b7.

It is also relevant that b6 and b7 are not applicable, or ".", if the child survived. If, say you look at the proportion of values of b7 that are 0, you are getting the proportion of deaths that were neonatal, not the proportion of births that were neonatal deaths.

Table C.6 in the Ethiopia 2011 main report is included mainly for assessing data quality assessment. That's why, for example, it does not use b7, but instead uses b6, and it breaks out the response "one year" for children whose age at death is just stated to be "one year", without specification of months.

In the first column of that table, 65.4% is the percentage of infant deaths that are neonatal (among births in the past five years). It is not the neonatal death rate or the proportion of children who have a neonatal death.

In the Guide to DHS Statistics, there is a description of how the under-five mortality rates are calculated. What you are calculating, as a proportion, is not the same thing as a neonatal mortality rate.

You may not need this, but here is Stata code for the construction of table C.6.

```
* how to get table C.6 in the Ethiopia 2011 DHS main report

* open ETBR61FL.dta

set more off

* prepare the years ago variable
gen months_ago=v008-b3
gen years_ago5=1+int(months_ago/60)
drop if years_ago5>4

* construct the age in months variable used in Table C.6
describe b6
label list b6
* tab b6

gen age_tableC6=.
replace age_tableC6=0      if b6<=130
replace age_tableC6=1      if b6>130 & b6<200
replace age_tableC6=b6-200 if b6>200 & b6<300
replace age_tableC6=24     if b6>=224 & b6<301
replace age_tableC6=99     if b6==301
replace age_tableC6=.      if b6==199 | b6==299 | b6>=399
tab age_tableC6 years_ago5 [iweight=v005/1000000]

gen nn_to_infant=.
replace nn_to_infant=0 if age_tableC6<=11
replace nn_to_infant=1 if age_tableC6==0
replace nn_to_infant=.  if b6>=399
summarize nn_to_infant if years_ago5==1 [iweight=v005/1000000]
summarize nn_to_infant if years_ago5==2 [iweight=v005/1000000]
summarize nn_to_infant if years_ago5==3 [iweight=v005/1000000]
summarize nn_to_infant if years_ago5==4 [iweight=v005/1000000]
```

Subject: Re: Computing Newborn Mortality Proportions
Posted by Mark on Tue, 07 Jan 2020 06:47:26 GMT
View Forum Message <> Reply to Message

I have found the thread posted by one of our experts, Mr. Trevor Croft (https://userforum.dhsprogram.com/index.php?t=msg&goto=15 138&S=Google). It has helped me a lot. Thank you so much for that. However, the provided code remove other variables, which I want to consider them in the analysis model. I mean, the Stata code provides mortality rate of each age group but it removes other variables that could be used as an explanatory variable. Would you please provide me your usual hand to separately compute neonatal mortality, infant mortality and under-5 mortality without dropping other variables? I, basically, want to use each mortality as an outcome of the study using the 2016 EDHS. Many thanks for your help.
Regards,

Subject: Re: Computing Newborn Mortality Proportions
Posted by Bridgette-DHS on Fri, 10 Jan 2020 14:48:32 GMT
View Forum Message <> Reply to Message

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

A rate is inherently a calculation for a group, not for an individual case. If you want to use individual-level characteristics (those of the child, such as sex, or of the mother, such as her level of education) as predictors of child mortality, you can use the individual-level data for children in the KR or BR files, with a binary outcome. Neonatal mortality is easiest. A child died in the first month if b7 (imputed age in months at death) is 0. (You do not need to take b5 into account, because b7 is coded NA if the child survived.) Construct a binary outcome: died=1 if b7=0 and otherwise died=0. Then to a logit regression of died on the predictors of interest. For an infant death, define died=1 if b7<12, and otherwise died=0. For anything past month 0, you need to take account of potential censoring, because a child born in the past year, say, has not had full exposure to the risk of an infant death.

Alternatively you can calculate the rates for small subpopulations, identified by the predictors, and try to make sense of the differences between the rates for boys and girls, for example, but this is statistically inefficient and clumsy. I strongly recommend the use of individual-level logit regression to analyze child mortality.

Subject: Re: Computing Newborn Mortality Proportions
Posted by Mark on Sat, 11 Jan 2020 02:34:22 GMT
View Forum Message <> Reply to Message

I Can't Thank You Enough. Your explanation clarifies those unclear points and helps me a lot.
Cheers,