

Adjusting Weights "On the Fly" in SAS¹

Non-random samples usually require weighting. The weight a case has is usually a function of the likelihood of inclusion in the sample. Many datasets from ICPSR, and other sources of data include a variable that should be used as a weight.

In analysis, the weighted number of cases should equal the unweighted number of cases, and the mean of the weights should be 1. If analysis is done by subgroup, the mean of weights for each subgroup should be 1. If it is not, tests of significance will not be correct. Some procedures in SAS adjust weights, and some do not. (PROC LOGISTIC has an undocumented option NORMWT which will adjust the weights. For example:

```
PROC LOGISTIC NORMWT DESCENDING; MODEL Y= X1 X2 X3; WEIGHT WTVAR; BY GENDER;
```

Weights can be adjusted by multiplying the weight by (N OF CASES / SUM OF WEIGHTS). The relative values of the weights are not changed, but they are set so that the mean is 1, and the sum of weights equals the N of cases. The N OF CASES and SUM OF WEIGHTS can be determined by running a PROC UNIVARIATE on the weight variable.

Following is a method of adjusting weights "on the fly" in SAS so that the weights will have a mean of 1.

```
** ADJUST SAS. ADJUST WEIGHTS "ON THE FLY" IN SAS ;
OPTIONS COMPRESS=YES LINESIZE=72 ;
LIBNAME SASB '/FULL/PATH/NAME' ;
DATA ALL; SET SASB.FARM40 ;

* SELECT CASES YOU WANT TO USE IN THE ANALYSIS ;
IF RACE NE . AND SEX NE . AND WEIGHT NE .
AND MARSTAT2 NE . AND AGECAT = 1 ;

* KEEP ONLY THE VARIABLES YOU WILL USE TO SAVE TIME AND SPACE IN THE SORT. ;

* IF YOU ARE NOT DOING SEPARATE ANALYSIS FOR SUBGROUPS, ;
* CREATE A CONSTANT TO USE IN MERGE STATEMENT. ;
CONSTANT=1 ;

KEEP RACE SEX WEIGHT MARSTAT2 CONSTANT ;

* SORT CASES BY SUBGROUPS ;
* (NO NEED TO SORT IF NOT DOING SEPARATE ANALYSIS BY SUBGROUP) ;
PROC SORT; BY RACE SEX ;

* CREATE A NEW DATASET "WTVARS" WHICH WILL HAVE ;
* CASES FOR EACH SUBGROUP. THE VARIABLES WILL BE NWT - THE
* N OF CASES IN THE DATASET, AND SUMWT - THE SUM OF WEIGHTS. ;

PROC UNIVARIATE NOPRINT ;
VAR WEIGHT;
* SUBSTITUTE "BY CONSTANT;" IF NOT DOING SEPARATE ANALYSIS BY SUBGROUPS.;
BY RACE SEX ;
OUTPUT OUT=WTVARS N=NWT SUM=SUMWT ;

* MERGE NEW DATASET WITH ORIGINAL, BY SUBGROUPS, OR BY CONSTANT ;
* IF NOT DOING SEPARATE ANALYSIS BY SUBGROUPS;
DATA FIN; MERGE ALL WTVARS ;
BY RACE SEX ;

* CREATE ADJUSTED WEIGHT ;
ADJWT=(WEIGHT * (NWT / SUMWT) ) ; LABEL ADJWT='WEIGHT ADJ FOR RACE SEX';

* WILL HAVE MEANS OF 1. ;
PROC MEANS; VAR WEIGHT; BY RACE SEX ;
PROC FREQ; TABLES MARSTAT2 ;
BY RACE SEX;
WEIGHT ADJWT ;
```

¹Prepared by Patty Glynn and distributed courtesy of the Center for Social and Demographic Analysis, University at Albany, State University of New York. December 19, 1994