Adjusting, or Normalizing Weights "On the Fly" in SAS[1]

Non-random samples usually require weighting. The weight a case has is usually a function of the likelihood of inclusion in the sample. Many data sets 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 separately by subgroup, the mean of weights for each subgroup should be 1. If the weighted number of cases is very different from the unweighted number of cases, tests of significance will not be valid. Some procedures in SAS adjust weights by default, and some do not. And, some procedures will "normalize" weights if asked to do so. For example, PROC LOGISTIC has an option NORMWT which will adjust the weights. (Example: **PROC LOGISTIC NORMWT DESCENDING; MODEL Y=X1 X2 X3; WEIGHT WTVAR; BY SEX;**)

Weights can be adjusted by dividing the weight variable by the mean of the weight variable. The relative values of the weights are not changed, but they are adjusted so that the mean is 1, and the sum of weights equals the N of cases. The mean of the weights can be running a proc means or 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 ;

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

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

* KEEP ONLY THE VARIABLES YOU WILL USE TO SAVE TIME AND SPACE IN THE SORT.  ;
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 A CASE 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  MEAN = MEANWHT ;

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

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

* THE UNWEIGHTED MEAN OF THE NEW WEIGHT VARIABLE WILL BE 1. ;
PROC MEANS; VAR ADJWT ; BY RACE SEX ;

* TO APPLY THE WEIGHT, USE THE "WEIGHT" COMMAND IN THE PROCEDURE. ;
PROC FREQ; TABLES MARSTAT2 ; BY RACE SEX; WEIGHT ADJWT ;
```

---

[1]Prepared by Patty Glynn, University of Washington. 8/14/01, updated 6/16/02