

Adjusting Weights "On the Fly" in SPSSX¹

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. SPSSX does not adjust weights automatically.

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 "sum of weights" can be determined by running a `FREQ` on the weight variable, and asking for `SUM` in the statistics statement. "`FREQ VAR = WEIGHT / STATISTICS = SUM`"

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

```
** ADJUST SPSSX. ADJUST WEIGHTS "ON THE FLY" IN SPSSX
* KEEP ONLY THE VARIABLES YOU WILL USE TO SAVE TIME AND SPACE IN THE SORT.
GET FILE=FARM40 / KEEP WEIGHT RACE SEX MARSTAT2
* TURN WEIGHTS OFF
WEIGHT OFF

* SELECT CASES YOU WANT TO USE IN THE ANALYSIS.
SELECT IF NOT(MISSING(RACE)) AND NOT(MISSING(SEX)) AND
      NOT(MISSING(MARSTAT2)) AND AGECAT=1

* IF YOU ARE NOT DOING SEPARATE ANALYSIS FOR SUBGROUPS,
  CREATE A CONSTANT TO USE IN "AGGREGATE" AND "MATCH FILES".
COMPUTE CONSTANT=1

* NO NEED TO SORT CASES IF NOT DOING SEPARATE ANALYSIS FOR SUBGROUPS.
SORT CASES BY RACE SEX
SAVE OUT= "$SPSSTMPDIR/orig.sav"

* 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.

* SUBSTITUTE "CONSTANT" FOR "RACE SEX" IF NOT DOING SEPARATE ANALYSIS
  BY SUBGROUP.
AGGREGATE OUTFILE= "$SPSSTMPDIR/WTVARS.sav"
  / BREAK = RACE SEX
  / NWT = N(WEIGHT)
  / SUMWT = SUM(WEIGHT)

* MERGE NEW DATASET WITH ORIGINAL, BY SUBGROUPS
* SUBSTITUTE "CONSTANT" FOR "RACE SEX" IF NOT DOING SEPARATE ANALYSIS BY SUBGROUP.
MATCH FILES TABLES = "$SPSSTMPDIR/WTVARS.sav"
  / FILE = "$SPSSTMPDIR/orig.sav"
  / BY RACE SEX

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

* IF DOING SEPARATE ANALYSIS BY SUBGROUP -
SPLIT FILE BY RACE SEX

* UNWEIGHTED MEANS WILL BE 1
DESC VAR = ADJWT
WEIGHT BY ADJWT
FREQ VAR=MARSTAT2
```

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, updated August 29, 2000.