DESCRIPT Example #3

SUDAAN Statements and Results Illustrated

- SUBPOPN
- SETENV
- Design effect (DEFT2) option
- STYLE option
- NEST

Input Data Set(s): NHANES3S3.SAS7bdat

Example

Estimate mean BMI for male and female adults, within race/ethnicity groups, using NHANES III.

Solution

The target population is the civilian, non-institutionalized population of the United States aged two months and older. In this example, we are interested in adults only, aged 17 years or older. Note that the NHANES III data set can be subsetted to adults only without losing aspects of the sampling design (*i.e.*, without losing any "pseudo-PSUs").

The analysis (dependent) variable on the VAR statement below (***Exhibit 1***) is body mass index. The TABLES statement requests the mean BMI cross-classified by race/ethnicity and gender. The SUBPOPN statement restricts the analysis to those aged 20+ years. The SETENV statement is used to get all columns of the printout table on one screen (or piece of paper). The requested design effect is DEFT2 (not the default). STYLE=NCHS is requested on the PRINT statement, as opposed to the default STYLE=BOX.

This example was run in SAS-Callable SUDAAN, and the programming code is presented below. Note that the basic SUDAAN code is the same for both Standalone and SAS-Callable versions.

Exhibit 1. SAS-Callable SUDAAN Code

```
libname in "\\rtints29\sudaan\data\nhanes3";
options linesize=95 pagesize=60 nocenter;

proc format;
  value sex 1="1=male"
            2="2=female";
  value age 1="1=17-34"
            2="2=35-49"
            3="3=50-64"
            4="4=65-90+";
  value race 1="1=nH_white"
             2="2=nH_black"
             3="3=Mex_Amer"
             4="4=other";

PROC DESCRIPT DATA=in.HANES3S3 FILETYPE=SAS DESIGN=WR DEFT2;
  NEST SDPSTRA6 SDPPSU6;
  WEIGHT WTPFHX6;

  SUBPOPN HSAGEIR > 19 / NAME = "ADULTS AGED 20+";

  CLASS DMARETHN HSSEX;
  VAR BMPBMI_R;
  TABLES DMARETHN*HSSEX;

  SETENV COLWIDTH=10 LABWIDTH=15 COLSPCE=1;
  PRINT  NSUM="SAMSIZE" MEAN SEMEAN LOWMEAN UPMEAN DEFFMEAN="DEFF2" /
         nsumfmt=f8.0 meanfmt=f5.2 deffmeanfmt=f5.2 semeanfmt=f7.2 STYLE=NCHS;
  rformat dmarethn race.;
  rformat hssex sex.;
  RTITLE "MEAN BMI BY SEX, RACE/ETHNICITY GROUPS, U.S. CIVILIAN"
         "NON-INSTITUTIONALIZED POPN. AGED 20 YEARS OR OLDER";
  RFOOTNOTE "NHANES-III, 1988-1994, JULY 1997 DATA RELEASE, BMI UPDATED";
```

Exhibit 2. First page of SUDAAN output (SAS *.lst file)

```
                                S U D A A N
           Software for the Statistical Analysis of Correlated Data
         Copyright     Research Triangle Institute      December 2011
                              Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a
With Replacement (WR) Design
    Sample Weight: WTPFHX6
    Stratification Variables(s): SDPSTRA6
    Primary Sampling Unit: SDPPSU6


Number of observations read    :  18162    Weighted count :187513911
Number of observations skipped :   1888
(WEIGHT variable nonpositive)
Observations in subpopulation  :  17030    Weighted count :177180671
Denominator degrees of freedom :     49
```

Exhibit 3. Frequencies for CLASS Variables (DMARETHN)

```
Frequencies and Values for CLASS Variables
----------------------------------------
Race-
  ethnicity     Frequency         Value
----------------------------------------
Ordered
  Position:
  1                  7121    1=nH_white
Ordered
  Position:
  2                  4709    2=nH_black
Ordered
  Position:
  3                  4528    3=Mex_Amer
Ordered
  Position:
  4                   672       4=other
----------------------------------------
```

Exhibit 3. Frequencies for CLASS Variables (HSSEX)

```
Frequencies and Values for CLASS Variables
------------------------------------
Sex             Frequency       Value
------------------------------------
Ordered
  Position:
  1                  7953      1=male
Ordered
  Position:
  2                  9077    2=female
------------------------------------
```

Exhibit 4. DESCRIPT Results: Race*Sex

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: ADULTS AGED 20+

MEAN BMI BY SEX, RACE/ETHNICITY GROUPS, U.S. CIVILIAN
NON-INSTITUTIONALIZED POPN. AGED 20 YEARS OR OLDER

by: Variable, Race-ethnicity, Sex.
for: Variable = Body Mass Index.

Race-ethnicity / Sex	SAMSIZE	Mean	SE Mean	Lower 95% Limit Mean	Upper 95% Limit Mean	DEFF2
Total						
Total	16969	26.51	0.11	26.29	26.73	6.27
1=male	7933	26.58	0.11	26.37	26.80	3.98
2=female	9036	26.44	0.16	26.13	26.76	5.51
1=nH_white						
Total	7110	26.34	0.13	26.07	26.61	4.12
1=male	3285	26.68	0.13	26.41	26.95	2.55
2=female	3825	26.03	0.19	25.64	26.41	3.72
2=nH_black						
Total	4692	27.75	0.15	27.45	28.04	2.30
1=male	2112	26.52	0.13	26.27	26.78	1.15
2=female	2580	28.71	0.21	28.29	29.13	2.15
3=Mex_Amer						
Total	4499	27.41	0.15	27.11	27.71	3.52
1=male	2250	26.84	0.16	26.52	27.17	2.89
2=female	2249	28.03	0.18	27.67	28.40	1.97
4=other						
Total	668	25.87	0.30	25.28	26.47	2.14
1=male	286	25.53	0.34	24.86	26.21	1.78
2=female	382	26.18	0.45	25.27	27.09	2.20

NHANES-III, 1988-1994, JULY 1997 DATA RELEASE, BMI UPDATED

BMI: Among the entire adult population aged 20 years and older, the estimated mean BMI for males and females appears to be similar (26.58 vs. 26.44, respectively, from ***Exhibit 4***), with overlapping confidence intervals. However, this is not the case when one compares males and females within race/ethnicity groups. Among non-Hispanic whites, for example, males seem to have a higher estimated mean BMI than do females (26.68 vs. 26.03, respectively). Within non-Hispanic blacks and Mexican-Americans, on the other hand, females seem to have a higher estimated mean BMI than males (28.71 vs. 26.52, respectively, for non-Hispanic blacks, and 28.03 vs. 26.84, respectively, for Mexican-Americans). For the "other" subpopulation, females seem to have a higher estimated mean BMI than males (26.18 vs. 25.53), although the estimated standard errors are larger for this subpopulation compared to the other three.

Design Effects: The design effect in this example is DEFT2. The comparison sampling plan is simple random sampling within the subpopulation, with the same sample size of elements as in the complex sample. For example, for the estimated mean BMI of 27.41 for Mexican-Americans, the comparison sampling plan for calculation of DEFT2 is a simple random sample of 4,499 elements from the population of all U.S. Mexican-Americans aged 20 years and older. Note that all design effects are larger than 1.0, since the effect of oversampling on the design effect has been removed. The design effects are larger than 1.0 due to clustering and to variability in the values of the sampling weights (beyond the effects of oversampling specific subpopulations). Note also that design effects are lower within a race/ethnicity group than for the entire population.