DESCRIPT Example #4

SUDAAN Statements and Results Illustrated

- CONTRAST
- PAIRWISE
- DIFFVAR
- SUBPOPN
- SETENV

Input Data Set(s): NHANES3S3.SAS7bdat

Example

Determine whether male and female adults differ on mean BMI, for the entire population and also within race/ethnicity, 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 subset 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 is body mass index (***Exhibit 1***). Males and females are compared on mean BMI three different ways (via CONTRAST, PAIRWISE, and DIFFVAR statements) to illustrate that equivalent results are obtained when there are only two groups being compared. All three contrast statements use the variable HSSEX (1=male, 2=female). The user-defined CONTRAST statement specifies a linear contrast of males vs. females on mean BMI (mean BMI for females minus mean BMI for males). The NAME option on the CONTRAST statement will be printed out to remind the user of the specified direction of the gender difference (F-M). The PAIRWISE statement instructs SUDAAN to compare all levels of HSSEX, two at a time, and to construct the linear contrasts for these comparisons. Since HSSEX is only at two levels, there is only one contrast; SUDAAN specifies the gender difference as male BMI minus female BMI. The DIFFVAR statement also instructs SUDAAN to construct a linear contrast comparing males and females, with the difference as male BMI minus female BMI.

The SUBPOPN statement restricts the analysis to those aged 20 years and older. The SETENV statement is used to get all columns of the printout table on one screen (or piece of paper).

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="M"
            2="F";
  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;
  NEST SDPSTRA6 SDPPSU6;
  WEIGHT WTPFHX6;

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

  SUBGROUP DMARETHN HSSEX;
  LEVELS   4        2;
  VAR BMPBMI_R;
  TABLES DMARETHN;

  CONTRAST HSSEX = (-1  1) / NAME="CONTRAST F-M";
  PAIRWISE HSSEX / NAME="PAIRWISE";
  DIFFVAR  HSSEX = (1 2) / NAME="DIFFVAR M-F";

  SETENV labwidth=20 COLWIDTH=6 COLSPCE=1;
  PRINT MEAN="Difference" SEMEAN="SE" LOWMEAN="Lower 95% Limit"
        UPMEAN="Upper 95% Limit" T_MEAN="T-Stat" P_MEAN="P-Value" /
        style=nchs meanfmt=f10.2 p_meanfmt=f7.4;
  RFORMAT dmarethn race.;
  RFORMAT hssex sex.;
  RTITLE "TEST NULL HYPOTHESIS:MALES/FEMALES HAVE SAME MEAN BMI, WITHIN"
         "RACE/ETHNICITY GROUP, U.S.CIVILIAN, NONINST. POPN. AGED 20+";
  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. DESCRIPT Results: Sex Differences Within Race

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

TEST NULL HYPOTHESIS:MALES/FEMALES HAVE SAME MEAN BMI, WITHIN
RACE/ETHNICITY GROUP, U.S.CIVILIAN, NONINST. POPN. AGED 20+

by: Variable, Race-ethnicity, CONTRAST.
for: Variable = Body Mass Index.
------------------------------------------------------------------------------
Race-ethnicity
   CONTRAST                                 Lower    Upper
                                            95%      95%
                     Difference       SE   Limit    Limit    T-Stat   P-Value
------------------------------------------------------------------------------
Total
   CONTRAST F-M           -0.14     0.16    -0.47     0.19    -0.87    0.3902
   PAIRWISE: (M,F)         0.14     0.16    -0.19     0.47     0.87    0.3902
   DIFFVAR M-F             0.14     0.16    -0.19     0.47     0.87    0.3902
1 = nH_white
   CONTRAST F-M           -0.65     0.20    -1.05    -0.25    -3.27    0.0019
   PAIRWISE: (M,F)         0.65     0.20     0.25     1.05     3.27    0.0019
   DIFFVAR M-F             0.65     0.20     0.25     1.05     3.27    0.0019
2 = nH_black
   CONTRAST F-M            2.19     0.22     1.75     2.63     9.93    0.0000
   PAIRWISE: (M,F)        -2.19     0.22    -2.63    -1.75    -9.93    0.0000
   DIFFVAR M-F            -2.19     0.22    -2.63    -1.75    -9.93    0.0000
3 = Mex_Amer
   CONTRAST F-M            1.19     0.17     0.86     1.52     7.17    0.0000
   PAIRWISE: (M,F)        -1.19     0.17    -1.52    -0.86    -7.17    0.0000
   DIFFVAR M-F            -1.19     0.17    -1.52    -0.86    -7.17    0.0000
4 = Other
   CONTRAST F-M            0.65     0.55    -0.46     1.76     1.17    0.2468
   PAIRWISE: (M,F)        -0.65     0.55    -1.76     0.46    -1.17    0.2468
   DIFFVAR M-F            -0.65     0.55    -1.76     0.46    -1.17    0.2468
------------------------------------------------------------------------------
NHANES-III, 1988-1994, JULY 1997 DATA RELEASE, BMI UPDATED
```

Exhibit 3 provides the results of the CONTRAST, PAIRWISE, and DIFFVAR statements, respectively. Each of these three rows within a given race/ethnicity level results in equivalent calculations for testing the null hypothesis of no gender differences on mean BMI among adults aged 20 years and older.

The first three rows of the table (Total) in ***Exhibit 3*** indicates that, in the adult population aged 20 years and older, the estimated difference in mean BMI for males and females is 0.14 or -0.14, depending upon whether the difference is calculated as (M-F) or as (F-M). ***Example 3*** shows the means that are subtracted (26.58 for males, 26.44 for females). The estimated standard error for the estimated difference is .16. The *t*-statistic (ratio of estimated difference to its estimated standard error) is 0.87 or -0.87. The two-sided *p*-value from the *t*-distribution with 49 df is .3902. And the 95% confidence limits for the difference contain the null value of 0. Hence, the null hypothesis is not rejected. Thus, males and females do not differ significantly on mean BMI in the adult population aged 20 and older. See ***Example 1*** for more discussion of the denominator degrees of freedom in NHANES III.

The gender differences within each race/ethnicity group indicate that males and females do differ significantly for three of the four race/ethnicity levels. For non-Hispanic whites, <u>males</u> have a significantly larger mean BMI (by .65 units). For non-Hispanic blacks and for Mexican-Americans, however, <u>females</u> have a significantly larger mean BMI (by 2.19 and 1.19 units, respectively). Accordingly, confidence limits within each of these 3 race/ethnicity groups do not contain the null value of 0 difference. The change in direction of the gender effect across race/ethnicity groups is accounting for the overall lack of gender's significance when averaged over the whole subpopulation of adults.