RATIO Example #2

SUDAAN Statements and Results Illustrated

- POLY statement replaces TABLES statement
- Test of trends
- NEST
- WEIGHT
- NEWVAR

Input Data Set(s): NHANES3S3.SAS7bdat

Example

This example determines whether there is a linear, quadratic, or cubic trend in the proportion of DMF teeth resulting from disease over age decades, using adults in NHANES III.

Solution

The data set is adults aged 17+ from NHANES III. Most of the syntax for this example is the same as in the previous example (i.e., NEST, WEIGHT, NEWVAR, SUBPOPX, NUMER, DENOM, and CLASS statements). The TABLES statement is replaced by a POLY (or POLYNOMIAL) statement to request a trend test of the estimated proportions across the levels of the AGEGRP variable (see ***Exhibit 1***).

Since recoded age has seven levels (approximately equally spaced in decades), we can test for an approximate linear, quadratic, and cubic trend, and even up to a 6-degree polynomial. For illustrative purposes, we request linear contrasts through the 6^{th} degree by using AGEGRP=6 on the POLY statement.

The SETENV and PRINT statements are used to manipulate the appearance of the printout. NOMARG is used on the PROC statement to avoid obtaining the same printout two times, since SUDAAN generates a TABLES _ONE_ statement for this program.

This example was run in SAS-Callable SUDAAN, and the programming code is presented in ***Exhibit 1***. 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 age 1="1=17-29"
            2="2=30-39"
            3="3=40-49"
            4="4=50-59"
            5="5=60-69"
            6="6=70-79"
            7="7=80+";

PROC RATIO DATA=in.HANES3S3 FILETYPE=SAS DESIGN=WR NOMARG;
  NEST   SDPSTRA6 SDPPSU6;
  WEIGHT WTPFEX6;

  SUBPOPX DEPEXFLG=1 / NAME="Had Dental Exam in MEC";

  NUMER  DEPDMFT1;     /* # DMF TEETH DUE TO DISEASE */
  DENOM  DEPDMFT2;     /* # DMF TEETH DUE TO ANY CAUSE */

  NEWVAR AGEGRP:     If 17 le HSAGEIR le 29 then AGEGRP=1
                 elseif 30 le HSAGEIR le 39 then AGEGRP=2
                 elseif 40 le HSAGEIR le 49 then AGEGRP=3
                 elseif 50 le HSAGEIR le 59 then AGEGRP=4
                 elseif 60 le HSAGEIR le 69 then AGEGRP=5
                 elseif 70 le HSAGEIR le 79 then AGEGRP=6
                 else if HSAGEIR ge 80 then AGEGRP=7;
  CLASS AGEGRP;
  POLY AGEGRP=6 / NAME="AGE TREND";

  SETENV LABWIDTH=18 COLSPCE=1;
  PRINT NSUM="SamSize" RHAT="Estimate" SERHAT="SE" T_RHAT="T_Stat"
        P_RHAT="P-Value" / NSUMFMT=F8.0 RHATFMT=F8.3 SERHATFMT=F7.4
        T_RHATFMT=F7.2 P_RHATFMT=F9.5 STYLE=NCHS;

  RLABEL depdmft1="DMFT Disease";
  RLABEL depdmft2="DMFT Any Cause";
  RFORMAT agegrp age.;
  RTITLE "TREND, OVER 7 AGE GROUPS, IN PROPORTION OF DMF TEETH DUE TO DISEASE"
         "U.S. ADULTS AGED 17+ YEARS";
  RFOOTNOTE "NHANES-III, 1988-1994, JULY 1997 DATA RELEASE";
```

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: WTPFEX6
    Stratification Variables(s): SDPSTRA6
    Primary Sampling Unit: SDPPSU6

Number of observations read    :  17705    Weighted count :187521389
Number of observations skipped :   2345
(WEIGHT variable nonpositive)
Observations in subpopulation  :  17235    Weighted count :182292666
Denominator degrees of freedom :     49
```

The values in ***Exhibit 2*** are explained in ***Example 2***.

Exhibit 3. Frequencies for CLASS Variable AGEGRP

```
Frequencies and Values for CLASS Variables
-----------------------------------
AGEGRP          Frequency      Value
-----------------------------------
Ordered
  Position:
  1                  4541    1=17-29
Ordered
  Position:
  2                  3244    2=30-39
Ordered
  Position:
  3                  2513    3=40-49
Ordered
  Position:
  4                  1801    4=50-59
Ordered
  Position:
  5                  2236    5=60-69
Ordered
  Position:
  6                  1701    6=70-79
Ordered
  Position:
  7                  1199      7=80+
-----------------------------------
```

Exhibit 4. RATIO Results

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: Had Dental Exam in MEC

TREND, OVER 7 AGE GROUPS, IN PROPORTION OF DMF TEETH DUE TO DISEASE
U.S. ADULTS AGED 17+ YEARS

for: Variable = DMFT Disease/DMFT Any Cause.
-------------------------------------------------------------------------
SUDAAN Reserved
  Variable One
   Contrast            SamSize   Estimate        SE    T_Stat     P-Value
-------------------------------------------------------------------------
1
   AGE TREND:
     Linear              15419      0.245    0.0146     16.81     0.00000
   AGE TREND:
     Quadratic           12278     -0.189    0.0298     -6.35     0.00000
   AGE TREND:
     Cubic               15419      0.131    0.0375      3.49     0.00103
   AGE TREND:
     Quartic             17219     -0.088    0.0586     -1.51     0.13808
   AGE TREND:
     Quintic             15419     -0.096    0.0686     -1.40     0.16885
   AGE TREND:
     6th_Degree          17219      0.013    0.0536      0.25     0.80587
-------------------------------------------------------------------------
NHANES-III, 1988-1994, JULY 1997 DATA RELEASE
```

SUDAAN calculates the value of six linear contrasts (***Exhibit 4***), one contrast for each degree of the specified polynomial equation; these are the orthogonal polynomial linear contrasts displayed in many statistics books. The contrasts are not orthogonal in this weighted survey data example. However, they

do provide tests of linear, quadratic, etc., trend, since the categories of age are approximately equally spaced (in decades). The null hypothesis for each *t*-test is that there is no trend of that degree.

The *p*-values for the linear contrasts for degrees 4, 5, and 6 indicate that these null hypotheses are not rejected (see ***Exhibit 4***). The linear contrasts for degrees 1, 2, and 3 indicate a significant linear and quadratic component, and a less strong cubic component.