

DESCRIPT Example #7

SUDAAN Statements and Results Illustrated

- VAR
- CATLEVEL
- CONTRAST
- SETENV
- RFORMAT

Input Data Set(s): NHANES3S3.SAS7bdat

Example

Compare male and female adults on arthritis prevalence, within age group, using NHANES III.

Solution

See *Example 6* for the estimates of arthritis prevalence and the general setup of the SUDAAN program below. The VAR and CATLEVEL statements (*Exhibit 1*) identify the dependent variable (arthritis). The CONTRAST statement with the variable HSSEX (1=male, 2=female) is used to compare males and females on arthritis prevalence by subtracting the male prevalence from the female prevalence. The TABLES statement requests that the contrast statement be implemented for each level of the variable AGEGRP4 (*i.e.*, for each of four age groups).

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 yesno 1="1=Yes";
  value sex 1="Male"
           2="Female";
  value age 1="1=17-34"
           2="2=35-49"
           3="3=50-64"
           4="4=65-90+";

PROC DESCRIPT DATA=in.HANES3S3 FILETYPE=SAS DESIGN=WR;
  NEST SDPSTRA6 SDPPSU6;
  WEIGHT WTPFQX6;

  VAR HAC1A;
  CATLEVEL 1;

  SUBGROUP AGEGRP4 HSSEX;
  LEVELS 4 2;
  TABLES AGEGRP4;
  CONTRAST HSSEX=(-1 1) / NAME="FEMALE-MALE";

  SETENV COLWIDTH=12;
  PRINT NSUM="Sample Size" PERCENT SEPERCENT LOWPCT="Lower 95% Limit"
        UPPCT="Upper 95% Limit" T_PCT="T-Test" P_PCT="P-Value";
  RFORMAT agegrp4 age.;
  RFORMAT hac1a yesno.;
  RTITLE "COMPARISON OF MALES & FEMALES ON ARTHRITIS PREVALENCE, BY AGE"
        "U.S. ADULTS (17+), NHANES-3, 1988-1994";
  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: WTPFQX6
Stratification Variables(s): SDPSTRA6
Primary Sampling Unit: SDPPSU6

Number of observations read      : 20050      Weighted count :187647206
Denominator degrees of freedom :      49
```

See *Example 5* for a discussion of the above printout.

The output in *Exhibit 3* shows that females have a significantly higher prevalence of arthritis than do males, both overall and for each age group except the youngest (17-34 years). The estimated difference of gender-specific prevalences is given by the “Cntrst Pct” line (*e.g.*, 8.15% over all ages, the difference between the female prevalence of 21.31% and the male prevalence of 13.15%, as given in *Example 6*).

In addition to the estimated difference in percentages, SUDAAN also provides the SE for the contrast, the lower and upper 95% confidence limits for the contrast, the *t*-statistic for testing H_0 : contrast=0, and the associated 2-sided *p*-value. Note that the estimated contrast is not statistically significant for the age group in which the confidence limits contain the null value of 0.

Exhibit 3. DESCRIPT Results: Sex Differences Within Age

Variance Estimation Method: Taylor Series (WR)

COMPARISON OF MALES & FEMALES ON ARTHRITIS PREVALENCE, BY AGE
 U.S. ADULTS (17+), NHANES-3, 1988-1994

by: Variable, AGEGRP4, Contrast.

for: Variable = Doctor ever told you had: arthritis: 1=Yes.

		Contrast
AGEGRP4		-----
		FEMALE-MALE

Total	Sample Size	20046
	Cntrst Pct	8.15
	SE Cntrst Pct	0.69
	Lower 95% Limit	6.76
	Upper 95% Limit	9.54
	T-Test	11.79
	P-Value	0.0000

1=17-34	Sample Size	6900
	Cntrst Pct	1.26
	SE Cntrst Pct	0.82
	Lower 95% Limit	-0.38
	Upper 95% Limit	2.90
	T-Test	1.54
	P-Value	0.1294

2=35-49	Sample Size	4496
	Cntrst Pct	3.78
	SE Cntrst Pct	1.39
	Lower 95% Limit	0.97
	Upper 95% Limit	6.58
	T-Test	2.71
	P-Value	0.0093

3=50-64	Sample Size	3401
	Cntrst Pct	15.89
	SE Cntrst Pct	2.31
	Lower 95% Limit	11.24
	Upper 95% Limit	20.53
	T-Test	6.88
	P-Value	0.0000

4=65-90+	Sample Size	5249
	Cntrst Pct	14.82
	SE Cntrst Pct	1.94
	Lower 95% Limit	10.92
	Upper 95% Limit	18.72
	T-Test	7.64
	P-Value	0.0000

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