

RATIO Example #4

SUDAAN Statements and Results Illustrated

- Compares PROC CROSSTAB versus PROC RATIO
- SUBPOPX
- LEVELS
- SUBGROUP
- SETENV

Input Data Set(s): SADLTST3.SSD

Example

This example use CROSSTAB instead of RATIO to perform a similar ratio analysis.

Solution

Example 3 can be done in either DESCRIPT or CROSSTAB, as well as in RATIO (shown above). Here we show how *Example 3* can be done in CROSSTAB. This example also illustrates that many calculations in CROSSTAB and DESCRIPT are actually ratio estimators of the type in PROC RATIO.

The SUBPOPX statement below (*Exhibit 1*) restricts the analysis to females who have ever had cancer (including non-melanoma skin cancer). Without a TABLES statement, CROSSTAB will produce a one-way percentage distribution for the SUBGROUP variable CNKIND5 (*i.e.*, ever had breast cancer). The LEVELS statement restricts the analysis to women who answered CNKIND5 with either *yes* (=1) or *no* (=2), effectively excluding from the analysis women who answered *don't know* (7) or *refuse* (9) to CNKIND5 (ever had breast cancer?). The SETENV statement manipulates the appearance of the printout. The PRINT statement changes some of the default labels in the printout and some of the default number of places after the decimal point.

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: CROSSTAB

```
libname in v604 "\\rtints29\sudaan\data\nhis1997";
options linesize=95 pagesize=60 nocenter;

PROC FORMAT;
  value yesno 1="1=Yes"
              2="2=No";

PROC CROSSTAB DATA=in.sadltst3 DESIGN=WR FILETYPE=SAS;
  NEST  STRATUM PSU;
  WEIGHT WTFA_SA;

  SUBPOPX SEX=2 and CANEV=1 / NAME="FEMALES AND CANCER EVER = YES";
  SUBGROUP CNKIND5;
  LEVELS 2;

  SETENV COLWIDTH=10 DECWIDTH=2 ROWWIDTH=10;
  PRINT NSUM="SAMSIZE" WSUM="POPSIZE" SEWGT="SE POPSIZE" ROWPER SEROW
        LOWROW="Lower 95% Limit PCT" UPROW="Upper 95% Limit PCT" / NSUMFMT=F10.0
        WSUMFMT=F10.0 SEWGFMT=F10.0;
  RFORMAT CNKIND5 yesno.;
  RTITLE "PERCENT OF ADULT FEMALE CANCER PATIENTS WHO HAD BREAST CANCER"
        "U.S., 1997";
  RFOOTNOTE "NHIS, 1997, SAMPLE ADULT DATASET";
```

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: WTFA_SA
  Stratification Variables(s): STRATUM
  Primary Sampling Unit: PSU

Number of observations read      : 36116      Weighted count :195276321
Observations in subpopulation  : 1531      Weighted count : 7243176
Denominator degrees of freedom : 339
```

Exhibit 2 indicates that there are 1,531 sample adults in the subpopulation of females with cancer (including non-melanoma skin cancer), and they represent an estimated 7,243,176 women in the adult U.S. population.

Exhibit 3. CROSSTAB Results

Variance Estimation Method: Taylor Series (WR)
 For Subpopulation: FEMALES AND CANCER EVER = YES

PERCENT OF ADULT FEMALE CANCER PATIENTS WHO HAD BREAST CANCER
 U.S., 1997

by: What kind of cancer...breast.

		What kind of cancer...breast		
		Total	1=Yes	2=No
	SAMSIZE	1521	436	1085
	POPSIZE	7202079	1974567	5227512
	SE POPSIZE	200925	110402	175011
	Row Percent	100.00	27.42	72.58
	SE Row Percent	0.00	1.33	1.33
	Lower 95% Limit			
	PCT	.	24.87	69.89
	Upper 95% Limit			
	PCT	.	30.11	75.13

NHIS, 1997, SAMPLE ADULT DATASET

Of the 1,531 women in the subpopulation under analysis, 1,521 gave a *yes* or *no* answer to the breast cancer question; these appear in the table above (*Exhibit 3*). The estimated percentage of female cancer survivors (or patients) who have had breast cancer is 27.42%, with an estimated standard error of 1.33%. This is the same result as in *Example 3* using RATIO (one exception—the 95% confidence limits are slightly different in CROSSTAB vs. RATIO, because CROSSTAB uses the default logit transformation that produces asymmetric confidence limits, while RATIO produces symmetric confidence intervals).

The CROSSTAB program can produce additional information, such as an estimated standard error for the estimated population sizes. Thus, the estimated number of adult female cancer survivors (including non-melanoma skin cancer) in the U.S. is 7,202,079, with an estimated standard error of 200,925. The estimated number of adult female breast cancer survivors in the U.S. is 1,974,567, with an estimated standard error of 110,402.

In order to conduct the above analysis in DESCRIPT, the SUBPOPX statement would be defined as women who ever had breast cancer and answered *yes* or *no* to CNKIND5 (breast cancer). The VAR statement would be CNKIND5, and the CATLEVEL statement would be level 1 (=yes) of CNKIND5. The confidence intervals will match those in CROSSTAB (asymmetric confidence intervals for percentages).