CHAPTER 6 ASDA ANALYSIS EXAMPLES REPLICATION-SUDAAN 10.0.1

GENERAL NOTES ABOUT ANALYSIS EXAMPLES REPLICATION

These examples are intended to provide guidance on how to use the commands/procedures for analysis of complex sample survey data and assume all data management and other preliminary work is done. The relevant syntax for the procedure of interest is shown first along with the associated output for that procedure(s). In some examples, there may be more than one block of syntax and in this case all syntax is first presented followed by the output produced.

In some software packages certain procedures or options are not available but we have made every attempt to demonstrate how to match the output produced by Stata 10+ in the textbook. Check the ASDA website for updates to the various software tools we cover.

NOTES ABOUT CATEGORICAL DATA ANALYSIS IN SUDAAN 10.0.1

The analysis replication examples were all run using SAS-callable SUDAAN version 10.0.1. There are very few differences between SAS-callable and stand-alone SUDAAN with the exception of the names of the procedures are sometimes slightly different as to avoid confusion with SAS procedures.

SUDAAN does not offer the ability to perform graphical analyses within the program therefore are not included in this output however output data sets can be saved and used in other software packages. SUDAAN PROC CROSSTAB/DESCRIPT can perform nearly all of the analyses presented in Chapter 6 of ASDA with the exception of the graphics. Some of the fine points of these procedures are the use of a SUBPOPN statement for subpopulation analyses, a CLASS statement for declaration of categorical variables, a TEST statement for hypothesis tests, a new GOF test for PROC CROSSTAB, and many other options for analysis/output. Please see the Sudaan 10.0.1 Language and Examples Guides for additional detail.

```
title "Analysis Example 6.1: Proportions: NHANES Data" ;
proc descript data=nhanes0506 filetype=sas deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class irregular age18p / nofreq;
table age18p ;
var irregular ;
setenv decwidth=3 colwidth=14 ;
run ;
```

Analysis Example 6.1: Proportions: NHANES Data

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute January 2009
Release 10.0.1

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
Sample Weight: WTMEC2YR
Stratification Variables(s): SDMVSTRA
Primary Sampling Unit: SDMVPSU

Number of observations read : 9950 Weighted count :291616892
Number of observations skipped : 398
(WEIGHT variable nonpositive)
Denominator degrees of freedom : 15

Variance Estimation Method: Taylor Series (WR)
by: Variable, AGE18P.

Variable		AGE18P		
		Total	0	1
1=yes 0=no	Sample Size	9438.000	4317.000	5121.000
	Weighted Size	280288865.407	69973494.921	210315370.487
	Total	6791921.890	581091.836	6210830.054
	Lower 95% Limit			
	Total	3591566.789	159099.538	3362457.622
	Upper 95% Limit			
	Total	9992276.992	1003084.134	9059202.487
	Mean	0.024	0.008	0.030
	SE Mean	0.006	0.003	0.007
	Lower 95% Limit			
	Mean	0.012	0.002	0.015
	Upper 95% Limit			
	Mean	0.036	0.014	0.044
	DEFF Mean #1	12.614	2.362	11.087
	DEFF Total #1	11.595	2.306	10.023

```
proc crosstab data=nhanes0506 filetype=sas deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class age18p irregular / nofreq;
subpopn age18p=1 ;
tables age18p*irregular ;
setenv decwidth=3 colwidth=14 ;
run ;
```

Analysis Example 6.1: Proportions: NHANES Data

```
                                S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      January 2009
                                 Release 10.0.1


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
    Sample Weight: WTMEC2YR
    Stratification Variables(s): SDMVSTRA
    Primary Sampling Unit: SDMVPSU


Number of observations read    :   9950    Weighted count :291616892
Number of observations skipped :    398
(WEIGHT variable nonpositive)
Observations in subpopulation  :   5334    Weighted count :217700471
Denominator degrees of freedom :     15

Date: 03-16-2010                                        SUDAAN                                                    Page:   1
Time: 12:13:04                                                                                                    Table: 1

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
by: AGE18P, 1=yes 0=no.
```

AGE18P		1=yes 0=no Total	0	1
Total	Sample Size	5121.000	4944.000	177.000
	Weighted Size	210315370.487	204104540.432	6210830.054
	SE Weighted	12813550.938	12922102.008	1336354.646
	DEFF Weighted	95.532	93.009	10.023
	Row Percent	100.000	97.047	2.953
	SE Row Percent	0.000	0.665	0.665
	Lower 95% Limit ROWPER	.	95.246	1.821
	Upper 95% Limit ROWPER	.	98.179	4.754
	DEFF Row Percent #1	.	11.087	11.087
	Col Percent	100.000	100.000	100.000
	SE Col Percent	0.000	0.000	0.000
	Lower 95% Limit COLPER	.	.	.
	Upper 95% Limit COLPER	.	.	.
	DEFF Col Percent #1	.	.	.
	Tot Percent	100.000	97.047	2.953
	SE Tot Percent	0.000	0.665	0.665
	Lower 95% Limit TOTPER	.	95.246	1.821
	Upper 95% Limit TOTPER	.	98.179	4.754
	DEFF Tot Percent #1	.	11.087	11.087

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
by: AGE18P, 1=yes 0=no.

AGE18P		1=yes 0=no		
		Total	0	1
1	Sample Size	5121.000	4944.000	177.000
	Weighted Size	210315370.487	204104540.432	6210830.054
	SE Weighted	12813550.938	12922102.008	1336354.646
	DEFF Weighted	95.532	93.009	10.023
	Row Percent	100.000	97.047	2.953
	SE Row Percent	0.000	0.665	0.665
	Lower 95% Limit ROWPER	.	95.246	1.821
	Upper 95% Limit ROWPER	.	98.179	4.754
	DEFF Row Percent #1	.	11.087	11.087
	Col Percent	100.000	100.000	100.000
	SE Col Percent	0.000	0.000	0.000
	Lower 95% Limit COLPER	.	.	.
	Upper 95% Limit COLPER	.	.	.
	DEFF Col Percent #1	.	.	.
	Tot Percent	100.000	97.047	2.953
	SE Tot Percent	0.000	0.665	0.665
	Lower 95% Limit TOTPER	.	95.246	1.821
	Upper 95% Limit TOTPER	.	98.179	4.754
	DEFF Tot Percent #1	.	11.087	11.087

```
proc descript data=nhanes0506 filetype=sas deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
subpopn age18p=1 ;
var irregular ;
setenv decwidth=3 colwidth=14 ;
run ;
```

Analysis Example 6.1: Proportions: NHANES Data

```
                                S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      January 2009
                                 Release 10.0.1


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
    Sample Weight: WTMEC2YR
    Stratification Variables(s): SDMVSTRA
    Primary Sampling Unit: SDMVPSU


Number of observations read    :   9950    Weighted count :291616892
Number of observations skipped :    398
(WEIGHT variable nonpositive)
Observations in subpopulation  :   5334    Weighted count :217700471
Denominator degrees of freedom :     15

Date: 03-16-2010                                          SUDAAN                                                Page:   1
Time: 12:15:42                                                                                                  Table: 1

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
by: Variable, SUDAAN Reserved Variable One.
```

Variable		SUDAAN Reserved Variable One	
		Total	1
1=yes 0=no	Sample Size	5121.000	5121.000
	Weighted Size	210315370.487	210315370.487
	Total	6210830.054	6210830.054
	Lower 95% Limit		
	Total	3362457.622	3362457.622
	Upper 95% Limit		
	Total	9059202.487	9059202.487
	Mean	0.030	0.030
	SE Mean	0.007	0.007
	Lower 95% Limit		
	Mean	0.015	0.015
	Upper 95% Limit		
	Mean	0.044	0.044
	DEFF Mean #1	11.087	11.087
	DEFF Total #1	10.023	10.023

```
title  "Analysis Example 6.2: Proportions by Subgroups: NHANES Data " ;
options ls=120 ps=64 ;
proc crosstab data=nhanes0506 deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class ridreth1 / nofreq ;
subpopn age18p=1 ;
tables ridreth1 ;
setenv decwidth=4 colwidth=10;
print nsum rowper serow lowrow uprow deffrow;
run ;
```

Analysis Example 6.2: Proportions by Subgroups: NHANES Data

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute January 2009
Release 10.0.1

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
Sample Weight: WTMEC2YR
Stratification Variables(s): SDMVSTRA
Primary Sampling Unit: SDMVPSU

Number of observations read : 9950 Weighted count :291616892
Number of observations skipped : 398
(WEIGHT variable nonpositive)
Observations in subpopulation : 5334 Weighted count :217700471
Denominator degrees of freedom : 15

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
by: 1=mex 2=oth hisp 3=white 4=black 5=other.

		1=mex 2=oth hisp 3=white 4=black 5=other					
		Total	Mexican	Other Hispanic	White	Black	Other
Sample Size		5334.0000	1133.0000	164.0000	2516.0000	1300.0000	221.0000
Row Percent		100.0000	8.0783	3.3785	71.4143	11.7263	5.4026
SE Row Percent		0.0000	1.0053	0.7421	2.7698	1.9849	0.5825
Lower 95% Limit	ROWPER	.	6.1785	2.1084	65.1672	8.1112	4.2873
Upper 95% Limit	ROWPER	.	10.4970	5.3718	76.9375	16.6604	6.7874
DEFF Row Percent	#1	.	10.1091	12.5313	27.9121	28.2693	4.9311

```
title  "Analysis Example 6.3: Proportions by Subgroups: NHANES Data " ;
options ls=120 ps=64 ;
proc crosstab data=nhanes0506 deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class bp_cat/ nofreq ;
subpopn age18p=1 ;
rformat bp_cat bpf. ;
tables bp_cat ;
setenv decwidth=3 ;
print nsum rowper serow lowrow uprow deffrow;
run ;
```

```
Analysis Example 6.3: Proportions by Subgroups: NHANES Data
                                S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      January 2009
                                 Release 10.0.1


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
    Sample Weight: WTMEC2YR
    Stratification Variables(s): SDMVSTRA
    Primary Sampling Unit: SDMVPSU


Number of observations read    :   9950    Weighted count :291616892
Number of observations skipped :    398
(WEIGHT variable nonpositive)
Observations in subpopulation  :   5334    Weighted count :217700471
Denominator degrees of freedom :     15

Date: 03-16-2010                                          SUDAAN                                                    Page:   1
Time: 12:34:43                                                                                                      Table: 1

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
by: BP_CAT.
```

		BP_CAT				
		Total	Normal	Pre-Hypertension	Stage 1 Hypertension	Stage 2 Hypertension
	Sample Size	5057.000	2441.000	1988.000	470.000	158.000
	Row Percent	100.000	47.108	41.854	8.641	2.397
	SE Row Percent	0.000	1.109	1.179	0.621	0.240
	Lower 95% Limit ROWPER	.	44.753	39.365	7.406	1.934
	Upper 95% Limit ROWPER	.	49.476	44.386	10.059	2.967
	DEFF Row Percent #1	.	3.500	4.051	3.464	1.754

```
* goodness of fit statistic is new in Sudaan 10.0.1 : used here to test specific proportions ;
title  "Analysis Example 6.4: Proportions by Subgroups with GOF Test: NHANES Data " ;
options ls=120 ps=64 ;
proc crosstab data=nhanes0506 deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class bp_cat/ nofreq ;
subpopn age18p=1 ;
rformat bp_cat bpf. ;
tables bp_cat ;
gofit bp_cat = (.5 .3 .15 .05) / waldchisq adjwaldf ;
setenv decwidth=3 ;
print nsum rowper serow lowrow uprow deffrow / gof = default ;
run ;
```

Analysis Example 6.4: Proportions by Subgroups with GOF Test: NHANES Data

```
                                S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      January 2009
                                 Release 10.0.1
```

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
    Sample Weight: WTMEC2YR
    Stratification Variables(s): SDMVSTRA
    Primary Sampling Unit: SDMVPSU

Number of observations read   :   9950   Weighted count :291616892
Number of observations skipped :   398
(WEIGHT variable nonpositive)
Observations in subpopulation :   5334   Weighted count :217700471
Denominator degrees of freedom :   15

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
by: BP_CAT.

		BP_CAT				
		Total	Normal	Pre-Hypertension	Stage 1 Hypertension	Stage 2 Hypertension
	Sample Size	5057.000	2441.000	1988.000	470.000	158.000
	Row Percent	100.000	47.108	41.854	8.641	2.397
	SE Row Percent	0.000	1.109	1.179	0.621	0.240
	Lower 95% Limit					
	ROWPER	.	44.753	39.365	7.406	1.934
	Upper 95% Limit					
	ROWPER	.	49.476	44.386	10.059	2.967
	DEFF Row Percent					
	#1	.	3.500	4.051	3.464	1.754

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
  Test Statistics for Goodness-of-Fit Hypotheses
  For: BP_CAT = (0.500 0.300 0.150 0.050)
by: Test Statistic.

Test Statistic	DF	Test Value	P-Value
Wald chi-square	3.00	317.867	0.000
Adj Wald F	3.00	91.828	0.000

```
*NOTE: analysis example 6.5 includes weighted pie and bar charts, not available directly in Sudaan therefore not included

title  "Analysis Example 6.6 Proportions by Gender and MDE:  NCSR" ;
options ls=120 ps=64 ;
proc crosstab data=ncsr filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtsh ;
rformat sex sf. ; rformat mde mdef. ;
class sex mde / nofreq ;
tables sex*mde ;
print / style=nchs ;
test chisq llchisq ;
run ;
```

Analysis Example 6.6 Proportions by Gender and MDE: NCSR

```
                              S U D A A N
           Software for the Statistical Analysis of Correlated Data
         Copyright      Research Triangle Institute      January 2009
                               Release 10.0.1


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


Number of observations read    :   9282    Weighted count :     9282
Denominator degrees of freedom :     42


Date: 03-16-2010                                        SUDAAN                                                  Page:  1
Time: 12:47:34                                                                                                  Table: 1

Variance Estimation Method: Taylor Series (WR)
by: Sex, MDE.
```

Sex MDE	Sample Size	Weighted Size	SE Weighted	DEFF Weighted	Row Percent	SE Row Percent	Lower 95% Limit ROWPER	Upper 95% Limit ROWPER
Total								
Total	9282	9282.00	453.55	.	100.00	0.00	.	.
No	7453	7502.54	349.58	84.95	80.83	0.49	79.83	81.79
Yes	1829	1779.46	113.96	9.03	19.17	0.49	18.21	20.17
Male								
Total	4139	4444.71	215.70	20.08	100.00	0.00	.	.
No	3522	3774.47	169.19	12.78	84.92	0.77	83.29	86.42
Yes	617	670.23	57.70	5.35	15.08	0.77	13.58	16.71
Female								
Total	5143	4837.29	248.29	26.61	100.00	0.00	.	.
No	3931	3728.06	195.08	17.06	77.07	0.56	75.91	78.19
Yes	1212	1109.23	61.50	3.87	22.93	0.56	21.81	24.09

Variance Estimation Method: Taylor Series (WR)
by: Sex, MDE.

Sex MDE	DEFF Row Percent #1	Col Percent	SE Col Percent	Lower 95% Limit COLPER	Upper 95% Limit COLPER	DEFF Col Percent #1	Tot Percent	SE Tot Percent	Lower 95% Limit TOTPER
Total									
Total	.	100.00	0.00	.	.	.	100.00	0.00	.
No	1.42	100.00	0.00	.	.	.	80.83	0.49	79.83
Yes	1.42	100.00	0.00	.	.	.	19.17	0.49	18.21
Male									
Total	.	47.89	0.53	46.81	48.96	1.05	47.89	0.53	46.81
No	2.08	50.31	0.71	48.88	51.74	1.50	40.66	0.70	39.26
Yes	2.08	37.66	1.19	35.29	40.10	1.08	7.22	0.34	6.56
Female									
Total	.	52.11	0.53	51.04	53.19	1.05	52.11	0.53	51.04
No	0.87	49.69	0.71	48.26	51.12	1.50	40.16	0.54	39.09
Yes	0.87	62.34	1.19	59.90	64.71	1.08	11.95	0.30	11.35

Variance Estimation Method: Taylor Series (WR)
by: Sex, MDE.

Sex MDE	Upper 95% Limit TOTPER	DEFF Tot Percent #1
Total		
Total	.	.
No	81.79	1.42
Yes	20.17	1.42
Male		
Total	48.96	1.05
No	42.08	1.87
Yes	7.95	1.64
Female		
Total	53.19	1.05
No	41.25	1.11
Yes	12.58	0.81

Hypothesis Test Test Statistic	DF	Test Value	P-Value
CHISQ (Obs - Exp)			
Wald-F	1	87.8030	0.0000
LLCHISQ (Log-Lin Model)			
Wald-F	1	57.2767	0.0000

```
title  "Analysis Example 6.7 Comparing Proportions of MDE by Gender:  NCSR" ;
proc crosstab data=ncsr filetype=sas ;
nest sestrat seclustr ;
weight ncsrwtsh ;
class sex mde / nofreq;
rformat sex sf. ; rformat mde mdef. ;
tables mde*sex  ;
setenv decwidth=3 colwidth=8 ;
print colper secol lowcol upcol / style=nchs ;
run ;
```

```
Analysis Example 6.7 Comparing Proportions of MDE by Gender:  NCSR
                              S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      January 2009
                                 Release 10.0.1
DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
    Sample Weight: NCSRWTSH
    Stratification Variables(s): SESTRAT
    Primary Sampling Unit: SECLUSTR


Number of observations read    :   9282    Weighted count :     9282
Denominator degrees of freedom :     42


Date: 03-16-2010                                          SUDAAN                                                Page:  1
Time: 12:59:48                                                                                                  Table: 1

Variance Estimation Method: Taylor Series (WR)
by: MDE, Sex.
```

MDE Sex	Col Percent	SE Col Percent	Lower 95% Limit COLPER	Upper 95% Limit COLPER
Total				
Total	100.000	0.000	.	.
Male	100.000	0.000	.	.
Female	100.000	0.000	.	.
No				
Total	80.829	0.488	79.825	81.794
Male	84.921	0.775	83.290	86.419
Female	77.069	0.565	75.910	78.189
Yes				
Total	19.171	0.488	18.206	20.175
Male	15.079	0.775	13.581	16.710
Female	22.931	0.565	21.811	24.090

```
proc descript data=ncsr filetype=sas ;
nest sestrat seclustr ;
weight ncsrwtsh ;
class sex / nofreq;
rformat sex sf. ; rformat mde mdef. ;
var mde ;
contrast sex =(1 -1) / name="Sex Contrast for MDE" ;
setenv decwidth=3 ;
print / style=nchs ;
run ;
```

Analysis Example 6.7 Comparing Proportions of MDE by Gender: NCSR

```
                                S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      January 2009
                                 Release 10.0.1


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


Number of observations read    :   9282    Weighted count :     9282
Denominator degrees of freedom :     42

Date: 03-16-2010                                          SUDAAN                                                  Page:  1
Time: 13:01:26                                                                                                    Table: 1

Variance Estimation Method: Taylor Series (WR)
by: Variable, SUDAAN Reserved Variable One, Contrast.

for: Variable = MDE.

-------------------------------------------------------------------------------------------------------------------------
SUDAAN Reserved
  Variable One                                                 Lower 95%      Upper 95%
   Contrast            Sample     Weighted                   Limit Cntrst   Limit Cntrst   Cntrst       SE Cntrst
                       Size       Size         Cntrst Total  Total          Total          Mean         Mean
-------------------------------------------------------------------------------------------------------------------------
Total
   Sex Contrast for
     MDE               9282.000   9282.000       -439.000      -509.994       -368.005       -0.079        0.010
1
   Sex Contrast for
     MDE               9282.000   9282.000       -439.000      -509.994       -368.005       -0.079        0.010
-------------------------------------------------------------------------------------------------------------------------

Date: 03-16-2010                                          SUDAAN                                                  Page:  2
Time: 13:01:26                                                                                                    Table: 1

Variance Estimation Method: Taylor Series (WR)
by: Variable, SUDAAN Reserved Variable One, Contrast.

for: Variable = MDE.

-----------------------------------------------------------------------
SUDAAN Reserved        Lower 95%    Upper 95%                P-value
  Variable One         Limit        Limit        T-Test      T-Test
   Contrast            Cntrst       Cntrst       Cont.Mean-  Cont.
                       Mean         Mean         =0          Mean=0
-----------------------------------------------------------------------
Total
   Sex Contrast for
     MDE                 -0.098       -0.059       -8.220      0.000
1
   Sex Contrast for
     MDE                 -0.098       -0.059       -8.220      0.000
-----------------------------------------------------------------------
```

```
title  "Analysis Example 6.8 Independence of Alcohol Dependence and Education Levels in Young Adults: NCSR" ;
options ls=120 ps=64 ;
proc crosstab data=ncsr filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtlg ;
class ed4cat ald / nofreq ;
subpopn age < 29 ;
rformat ed4cat edf. ;
tables ed4cat * ald ;
test chisq llchisq ;
print rowper serow lowrow uprow deffrow / test=all style=nchs ;
setenv decwidth=3 ;
run ;
```

Analysis Example 6.8 Independence of Alcohol Dependence and Education Levels in Young Adults: NCSR

```
                              S U D A A N
           Software for the Statistical Analysis of Correlated Data
         Copyright      Research Triangle Institute      January 2009
                               Release 10.0.1


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With Replacement (WR)
Design
    Sample Weight: NCSRWTLG
    Stratification Variables(s): SESTRAT
    Primary Sampling Unit: SECLUSTR
Number of observations read    :   5692    Weighted count :     5692
Number of observations skipped :   3590
(WEIGHT variable nonpositive)
Observations in subpopulation  :   1275    Weighted count :     1267
Denominator degrees of freedom :     42
Date: 03-16-2010                                        SUDAAN                                                    Page:  1
Time: 13:29:02                                                                                                    Table: 1

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE < 29
by: Years of education-4 categories, ALD.
```

Years of education-4 categories ALD	Row Percent	SE Row Percent	Lower 95% Limit ROWPER	Upper 95% Limit ROWPER	DEFF Row Percent #1
Total					
Total	100.00	0.00	.	.	.
0	94.05	0.88	92.00	95.60	1.76
1	5.95	0.88	4.40	8.00	1.76
0-11 Yrs					
Total	100.00	0.00	.	.	.
0	90.87	2.94	82.97	95.31	2.30
1	9.13	2.94	4.69	17.03	2.30
12 Yrs					
Total	100.00	0.00	.	.	.
0	95.14	1.35	91.59	97.24	1.59
1	4.86	1.35	2.76	8.41	1.59
11-13 Yrs					
Total	100.00	0.00	.	.	.
0	95.10	1.00	92.63	96.78	0.94
1	4.90	1.00	3.22	7.37	0.94
16+ Yrs					
Total	100.00	0.00	.	.	.
0	93.10	1.36	89.78	95.39	0.60
1	6.90	1.36	4.61	10.22	0.60

For Subpopulation: AGE < 29
  Chi Square Test of Independence for Years of education-4 categories and ALD

ChiSq	P-value ChiSq	Degrees of Freedom Chi-Sq	LLChiSq	P-value LLChiSq	Degrees of Freedom LLC-hiSq
1.07	0.3703	3	1.38	0.2609	3

```
title  "Analysis Example 6.9 Simple Logistic Regression to Estimate the NCS-R Male:Female Odds Ratio for MDE" ;
options ls=120 ps=64 ;
proc rlogist data=ncsr filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtsh  ;
model mde = sexm ;
setenv decwidth=3 ;
run ;
```

```
Analysis Example 6.9 Simple Logistic Regression to Estimate the NCS-R Male:Female Odds Ratio for MDE

                                S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      January 2009
                                 Release 10.0.1


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


Number of zero responses     :  7453
Number of non-zero responses :  1829

Independence parameters have converged in 5 iterations

Number of observations read       :   9282    Weighted count:     9282
Observations used in the analysis :   9282    Weighted count:     9282
Denominator degrees of freedom    :     42


Maximum number of estimable parameters for the model is  2

File NCSR contains   84 Clusters
  84 clusters were used to fit the model
Maximum cluster size is 234 records
Minimum cluster size is  29 records


Sample and Population Counts for Response Variable MDE
Based on observations used in the analysis
0:  Sample Count     7453    Population Count      7503
1:  Sample Count     1829    Population Count      1779

R-Square for dependent variable MDE (Cox & Snell, 1989): 0.009981

-2 * Normalized Log-Likelihood with Intercepts Only :  9072.13
-2 * Normalized Log-Likelihood Full Model          :  8979.02
Approximate Chi-Square (-2 * Log-L Ratio)          :    93.11
Degrees of Freedom                                 :        1

Note: The approximate Chi-Square is not adjusted for clustering.
      Refer to hypothesis test table for adjusted test.
```

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Logit
Response variable MDE: MDE
by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #1	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	-1.212	0.873	0.032	-1.277	-1.148	-37.935	0.000
SEXM	-0.516	1.589	0.068	-0.654	-0.379	-7.568	0.000

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Logit
Response variable MDE: MDE
by: Contrast.

Contrast	Degrees of Freedom	Wald F	P-value Wald F
OVERALL MODEL	2.000	1119.190	0.000
MODEL MINUS INTERCEPT	1.000	57.277	0.000
INTERCEPT	1.000	1439.096	0.000
SEXM	1.000	57.277	0.000

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Logit
Response variable MDE: MDE
by: Independent Variables and Effects.

Independent Variables and Effects	Odds Ratio	Lower 95% Limit OR	Upper 95% Limit OR
Intercept	0.298	0.279	0.317
SEXM	0.597	0.520	0.685

```
title  "Analysis Example 6.10 Test of Association between Gender and MDE When Controlling for Age: NCSR " ;
options ls=120 ps=64 ;
proc crosstab data=ncsr filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtsh  ;
class agecat mde sexm / nofreq ;
tables agecat*sexm*mde ;
risk mhor mhrr1 lor lrr1 ;
test cmh chisq ;
print nsum wsum rowper serow colper secol /tests=all adjrisk=all cmhtest=all ;
run ;
```

NOTE: CODES FOR AGECAT 1=18-29 2=30-39 3=40-49 4=50+ YEARS OF AGE, MDE 0=NO 1=YES, SEXM 0=FEMALE 1=MALE.

Analysis Example 6.11 Test of Association between Gender and MDE When Controlling for Age: NCSR

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute January 2009
Release 10.0.1

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

Number of observations read : 9282 Weighted count : 9282
Denominator degrees of freedom : 42

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, SEXM, MDE.

for: AGECAT = Total.

SEXM / MDE	Sample Size	Weighted Size	Row Percent	SE Row Percent	Col Percent	SE Col Percent
Total						
Total	9282	9282.00	100.00	0.00	100.00	0.00
0	7453	7502.54	80.83	0.49	100.00	0.00
1	1829	1779.46	19.17	0.49	100.00	0.00
0						
Total	5143	4837.29	100.00	0.00	52.11	0.53
0	3931	3728.06	77.07	0.56	49.69	0.71
1	1212	1109.23	22.93	0.56	62.34	1.19
1						
Total	4139	4444.71	100.00	0.00	47.89	0.53
0	3522	3774.47	84.92	0.77	50.31	0.71
1	617	670.23	15.08	0.77	37.66	1.19

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, SEXM, MDE.

for: AGECAT = 1.

SEXM / MDE	Sample Size	Weighted Size	Row Percent	SE Row Percent	Col Percent	SE Col Percent
Total						
Total	2104	2167.86	100.00	0.00	100.00	0.00
0	1698	1765.87	81.46	0.83	100.00	0.00
1	406	401.99	18.54	0.83	100.00	0.00
0						
Total	1138	1066.19	100.00	0.00	49.18	1.10
0	870	824.18	77.30	1.16	46.67	1.33
1	268	242.00	22.70	1.16	60.20	2.35
1						
Total	966	1101.67	100.00	0.00	50.82	1.10
0	828	941.69	85.48	1.21	53.33	1.33
1	138	159.99	14.52	1.21	39.80	2.35

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, SEXM, MDE.

for: AGECAT = 2.

SEXM / MDE	Sample Size	Weighted Size	Row Percent	SE Row Percent	Col Percent	SE Col Percent
Total						
Total	1904	1696.12	100.00	0.00	100.00	0.00
0	1483	1317.20	77.66	1.32	100.00	0.00
1	421	378.92	22.34	1.32	100.00	0.00
0						
Total	1040	884.20	100.00	0.00	52.13	1.20
0	758	640.97	72.49	1.54	48.66	1.52
1	282	243.23	27.51	1.54	64.19	2.67
1						
Total	864	811.92	100.00	0.00	47.87	1.20
0	725	676.23	83.29	1.88	51.34	1.52
1	139	135.70	16.71	1.88	35.81	2.67

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, SEXM, MDE.

for: AGECAT = 3.

SEXM MDE	Sample Size	Weighted Size	Row Percent	SE Row Percent	Col Percent	SE Col Percent
Total						
Total	1891	1978.32	100.00	0.00	100.00	0.00
0	1447	1528.92	77.28	0.87	100.00	0.00
1	444	449.40	22.72	0.87	100.00	0.00
0						
Total	1003	999.63	100.00	0.00	50.53	1.15
0	732	741.34	74.16	1.08	48.49	1.41
1	271	258.29	25.84	1.08	57.47	2.47
1						
Total	888	978.68	100.00	0.00	49.47	1.15
0	715	787.58	80.47	1.57	51.51	1.41
1	173	191.11	19.53	1.57	42.53	2.47

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, SEXM, MDE.

for: AGECAT = 4.

SEXM MDE	Sample Size	Weighted Size	Row Percent	SE Row Percent	Col Percent	SE Col Percent
Total						
Total	3383	3439.70	100.00	0.00	100.00	0.00
0	2825	2890.55	84.03	0.79	100.00	0.00
1	558	549.15	15.97	0.79	100.00	0.00
0						
Total	1962	1887.27	100.00	0.00	54.87	0.97
0	1571	1521.56	80.62	0.90	52.64	1.09
1	391	365.71	19.38	0.90	66.60	2.25
1						
Total	1421	1552.42	100.00	0.00	45.13	0.97
0	1254	1368.99	88.18	1.14	47.36	1.09
1	167	183.44	11.82	1.14	33.40	2.25

Variance Estimation Method: Taylor Series (WR)
Cochran-Mantel-Haenszel Test of Association for SEXM and MDE
Stratified by: AGECAT

	Cochran-Mantel-Haenszel Chi-Square	Degrees of Freedom CMH	P-value CMH Test
1	92.4582	1	0.0000

Variance Estimation Method: Taylor Series (WR)
Summary Statistics for Variable SEXM by Variable MDE
Controlling for: Variable AGECAT
by: Adjusted Risk.

Adjusted Risk	Value	Lower 95% Limit	Upper 95% Limit	Log(Value)	SE Log(Value)	VAR Log(Value)
MH Common OR	0.59	0.51	0.67	-0.53	0.07	0.00
MH Common RR Col 1	0.91	0.88	0.93	-0.10	0.01	0.00
Logit Common OR	0.60	0.52	0.68	-0.51	0.07	0.00
Logit Common RR Col 1	0.91	0.89	0.93	-0.10	0.01	0.00

Variance Estimation Method: Taylor Series (WR)
Chi Square Test of Independence for SEXM and MDE
by: AGECAT.

AGECAT	ChiSq	P-value ChiSq	Degrees of Freedom Chi-Sq	LLChiSq	P-value LLChiSq	Degrees of Freedom LLChiSq
Total	87.80	0.0000	1	57.28	0.0000	1
1	23.47	0.0000	1	20.43	0.0000	1
2	23.22	0.0000	1	19.49	0.0001	1
3	9.12	0.0043	1	8.56	0.0055	1
4	38.28	0.0000	1	24.75	0.0000	1