SUDAAN Analysis Example Replication C6

```
* Sudaan Analysis Examples Replication for ASDA 2nd Edition
* Berglund April 2017
* Chapter 6 ;
libname d "P:\ASDA 2\Data sets\nhanes 2011_2012\" ;
ods graphics off ;
options nodate nonumber ls=135 ps=68 ;
title ;
ods listing ;
data c6_nhanes ;
  set d.nhanes1112_sub_8aug2016 ;
run ;
* sort data before use of Sudaan ;
proc sort ;
 by sdmvstra sdmvpsu ;
run ;

title "Example 6.1: Estimating the Proportion of the U.S. Adult Population with an Irregular Heart Beat. " ;
* CI based on symmetric interval ;
proc descript data=c6_nhanes filetype=sas deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class irregular age18p / nofreq;
table age18p ;
var irregular ;
setenv decwidth=5 colwidth=14 ;
run ;
* CI based on logit ;
proc crosstab data=c6_nhanes filetype=sas deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class age18p irregular / nofreq;
subpopn age18p=1 ;
tables irregular ;
setenv decwidth=5 colwidth=14 ;
run ;

title " Example 6.2: Estimating the Proportion of U.S. Adults by Race and Ethnicity using NHANES data. " ;
options ls=120 ps=64 ;
proc crosstab data=c6_nhanes 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 ;

title " Example 6.3: Estimating the Proportions of U.S. Adults by Blood Pressure Category using the 2011-2012 NHANES
Data." ;
options ls=120 ps=64 ;
proc crosstab data=c6_nhanes deft1 ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class bp cat/ nofreq ;
subpopn age18p=1 ;
tables bp_cat ;
setenv decwidth=3 ;
print nsum rowper serow lowrow uprow deffrow;
run ;

libname russia "P:\ASDA 2\Data sets\ESS6 Russia" ;
data c6_russia ;
  set russia.ess6_russia_20aug2016 ;
run ;
* Note: Sudaan requires numeric design variables, these are in data set and ready to use ;
proc sort ;
 by nstrat psu ;
run ;
title "Example 6.4: A Goodness of Fit Test for Proportions of Russians age 15+ by Marital Status." ;
proc crosstab data=c6_russia deft1 ;
nest nstrat psu ;
weight pspwght ;
class marcat / nofreq ;
```

```
tables marcat ;
gofit marcat = (.5 .25 .25) / waldchi adjwaldf ;
setenv decwidth=3 ;
print nsum rowper serow lowrow uprow deffrow / gof = default ;
run ;
data _null_ ;
file print ;
put "Graphics Not Available in Sudaan: Example 6.5: Pie Charts and Vertical Bar Charts of the Estimated Proportions of
Russians age 15+ by Marital Status." ;
run ;

libname ncsr "P:\ASDA 2\Data sets\ncsr\" ;
data c6_ncsr ;
 set ncsr.ncsr_sub_5apr2017 ;
run ;
proc sort ;
 by sestrat seclustr ;
run ;

title " Example 6.6: Estimation of Total and Row Proportions for the Crosstabulation of Gender and Lifetime Major
Depression Status (Source: NCS-R)." ;
options ls=120 ps=64 ;
proc crosstab data=c6_ncsr filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtsh ;
class sex mde / nofreq ;
tables sex*mde ;
print / style=nchs ;
test chisq llchisq ;
run ;

title "Example 6.7: Comparing the Proportions of U.S. Adult Men and Women with Lifetime Major Depression." ;
proc descript data=c6_ncsr filetype=sas ;
nest sestrat seclustr ;
weight ncsrwtsh ;
class sex / nofreq;
var mde ;
contrast sex =(1 -1) / name="Sex Contrast for MDE" ;
setenv decwidth=3 ;
print / style=nchs ;
run ;

title " Example 6.8: Testing the Independence of MDE and Gender in U.S. Adults Using the NCS-R data." ;
proc crosstab data=c6_ncsr filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtsh ;
class sex mde / nofreq ;
tables sex*mde ;
print / style=nchs ;
test chisq llchisq ;
run ;

data c6_ncsr1 ;
 set c6_ncsr ;
 * create indicator for subpopulation of interest ;
 age18_28=0 ;
 if 18<=age<=28 then age18_28=1 ;
 * age categories for Example 6.11 : create agecat in different age groups than in ag4cat ;
 if age<=29 then agecat=1 ;
 else if age <=39 then agecat=2 ;
 else if age <=49 then agecat=3 ;
 else agecat=4 ;
 run ;
proc sort  ;
by sestrat seclustr ;
run ;

title "Example 6.9: Testing the Independence of Alcohol Dependence and Education Level in Young Adults (Ages 18-28) using
the NCS-R data. " ;
options ls=120 ps=64 ;
proc crosstab data=c6_ncsr1 filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtlg ;
class ed4cat ald / nofreq ;
subpopn age18_28=1 ;
tables ed4cat * ald ;
```

```
print rowper serow lowrow uprow deffrow / stest=all style=nchs ;
test chisq llchisq ;
setenv decwidth=3 ;
run ;

title "Example 6.10: Simple Logistic Regression to Estimate the NCS-R Male/Female Odds Ratio for Lifetime Major
Depressive Episode. " ;
proc rlogist data=c6_ncsr filetype=sas deft1 ;
nest sestrat seclustr ;
weight ncsrwtsh ;
model mde = sexm ;
setenv decwidth=3 ;
run ;

title "Example 6.11: Using the NCS-R Data to Estimate and Test the Association between Gender and Depression in the U.S.
Adult Population when controlling for Age." ;
title2 "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." ;
options ls=120 ps=64 ;
proc crosstab data=c6_ncsr1 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 /stest=all adjrisk=all atest=all ;
run ;
title ;

data _null_ ;
file print ;
put "Not Available in Sudaan: Example 6.12: A Simple Log-linear Model to Test the Association between Lifetime Major
Depression Episode and Sex. " ;
run ;
ods listing close ;
```

Output SUDAAN Analysis Example Replication C6

Example 6.1: Estimating the Proportion of the U.S. Adult Population with an Irregular Heart Beat.

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute June 2013
Release 11.0.1

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 : 9338 Weighted count :306590681
Number of observations skipped : 418
(WEIGHT variable nonpositive)
Denominator degrees of freedom : 17

Date: 05-15-2017 SUDAAN
Page: 1
Time: 13:09:34
Table: 1

Variance Estimation Method: Taylor Series (WR)
by: Variable, Age >=18: 1=Yes 0=No.

Variable		Age >=18: 1=Yes 0=No		
		Total	0	1
1=yes 0=no	Sample Size	9027.00000	3653.00000	5374.00000
	Weighted Size	**********	73011719.98572	**********
	Total	3971887.40801	305579.82178	3666307.58623
	Lower 95% Limit Total	2868493.21259	60374.24113	2695880.65165
	Upper 95% Limit Total	5075281.60343	550785.40242	4636734.52081
	Mean	0.01340	0.00419	0.01642
	SE Mean	0.00146	0.00154	0.00168
	Lower 95% Limit Mean	0.01032	0.00094	0.01288
	Upper 95% Limit Mean	0.01648	0.00743	0.01996
	DEFF Mean #1	1.45431	1.26100	1.18545
	DEFF Total #1	2.12465	1.34750	1.77861

Example 6.1: Estimating the Proportion of the U.S. Adult Population with an Irregular Heart Beat.

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute June 2013
Release 11.0.1

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 : 9338 Weighted count :306590681
Number of observations skipped : 418
(WEIGHT variable nonpositive)
Observations in subpopulation : 5615 Weighted count :232002539
Denominator degrees of freedom : 17

Date: 05-15-2017 SUDAAN
Page: 1
Time: 13:09:34
Table: 1

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

		1=yes 0=no		
		Total	0	1
	Sample Size	5374.00000	5264.00000	110.00000
	Weighted Size	**********	**********	3666307.58623
	SE Weighted	13647988.38992	13382260.55750	459958.56877
	DEFF Weighted	93.50807	87.54942	1.77861
	Row Percent	100.00000	98.35810	1.64190
	SE Row Percent	0.00000	0.16779	0.16779
	Lower 95% Limit			
	ROWPER	.	97.96390	1.32298
	Upper 95% Limit			
	ROWPER	.	98.67702	2.03610
	DEFF Row Percent			
	#1	.	1.18545	1.18545
	Col Percent	100.00000	98.35810	1.64190
	SE Col Percent	0.00000	0.16779	0.16779
	Lower 95% Limit			
	COLPER	.	97.96390	1.32298
	Upper 95% Limit			
	COLPER	.	98.67702	2.03610
	DEFF Col Percent			
	#1	.	1.18545	1.18545
	Tot Percent	100.00000	98.35810	1.64190
	SE Tot Percent	0.00000	0.16779	0.16779
	Lower 95% Limit			
	TOTPER	.	97.96390	1.32298
	Upper 95% Limit			
	TOTPER	.	98.67702	2.03610
	DEFF Tot Percent			
	#1	.	1.18545	1.18545

Example 6.2: Estimating the Proportion of U.S. Adults by Race and Ethnicity using NHANES data.

```
                          S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright       Research Triangle Institute       June 2013
                                Release 11.0.1


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    :   9338    Weighted count :306590681
Number of observations skipped :    418
(WEIGHT variable nonpositive)
Observations in subpopulation  :   5615    Weighted count :232002539
Denominator degrees of freedom :     17
```


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	1	2	3	4	5
	Sample Size	5615.0000	569.0000	577.0000	2014.0000	1505.0000	950.0000
	Row Percent	100.0000	7.9168	6.6224	65.9386	11.7185	7.8037
	SE Row Percent	0.0000	1.7251	1.5193	3.8892	2.3370	1.0917
	Lower 95% Limit ROWPER	.	4.9597	4.0523	57.3274	7.6141	5.7896
	Upper 95% Limit ROWPER	.	12.4068	10.6420	73.6120	17.6135	10.4407
	DEFF Row Percent #1	.	28.8416	26.3723	47.5827	37.3007	11.7043

Example 6.3: Estimating the Proportions of U.S. Adults by Blood Pressure Category using the 2011-2012 NHANES Data.

```
                              S U D A A N
                Software for the Statistical Analysis of Correlated Data
              Copyright       Research Triangle Institute       June 2013
                                    Release 11.0.1


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    :   9338    Weighted count :306590681
Number of observations skipped :    418
(WEIGHT variable nonpositive)
Observations in subpopulation  :   5615    Weighted count :232002539
Denominator degrees of freedom :     17
```


Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18P = 1
by: Blood Pressure: 1=Normal 2=Pre-Hypertension 3=Hypertension Stage 1 4=Hypertension Stage 2.

		Blood Pressure: 1=Normal 2=Pre-Hypertension 3=Hypertension Stage 1 4=Hypertension Stage 2				
		Total	1	2	3	4
	Sample Size	5356.000	2438.000	2284.000	489.000	145.000
	Row Percent	100.000	47.222	42.799	7.978	2.001
	SE Row Percent	0.000	1.552	1.204	0.582	0.438
	Lower 95% Limit ROWPER	.	43.964	40.280	6.834	1.258
	Upper 95% Limit ROWPER	.	50.504	45.355	9.294	3.169
	DEFF Row Percent #1	.	6.557	4.014	3.125	6.650

Example 6.4: A Goodness of Fit Test for Proportions of Russians age 15+ by Marital Status.

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute June 2013
Release 11.0.1

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

Number of observations read : 2484 Weighted count : 2484
Denominator degrees of freedom : 176

Date: 06-06-2017 SUDAAN
Page: 1
Time: 13:06:37
Table: 1

Variance Estimation Method: Taylor Series (WR)
by: Marital Status: 1=Currently Married 2=Previously Married 3=Never Married.

		Marital Status: 1=Currently Married 2=Previously Married 3=Never Married			
		Total	1	2	3
	Sample Size	2444.000	1066.000	791.000	587.000
	Row Percent	100.000	50.386	23.007	26.607
	SE Row Percent	0.000	1.288	1.154	1.340
	Lower 95% Limit ROWPER	.	47.846	20.809	24.048
	Upper 95% Limit ROWPER	.	52.924	25.362	29.334
	DEFF Row Percent #1	.	1.627	1.842	2.255

Date: 06-06-2017 SUDAAN
Page: 2
Time: 13:06:37
Table: 1

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Goodness-of-Fit Hypotheses
For: MARCAT = (0.500 0.250 0.250)
by: Test Statistic.

Test Statistic	DF	Test Value	P-Value
Wald chi-square	2.00	3.171	0.205
Adj Wald F	2.00	1.576	0.210

Graphics Not Available in Sudaan: Example 6.5: Pie Charts and Vertical Bar Charts of the Estimated Proportions of Russia ns age 15+ by Marital Status.

Example 6.6: Estimation of Total and Row Proportions for the Crosstabulation of Gender and Lifetime Major Depression

```
                              S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright       Research Triangle Institute       June 2013
                                Release 11.0.1


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: Sex 1=Male 2=Female, Major Depressive Episode 1=Yes 0=No.

Sex 1=Male 2=Female Major Depressive Episode 1=Yes 0=No	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	.	.
0	7453	7502.54	349.58	84.95	80.83	0.49	79.83	81.79
1	1829	1779.46	113.96	9.03	19.17	0.49	18.21	20.17
1								
Total	4139	4444.71	215.70	20.08	100.00	0.00	.	.
0	3522	3774.47	169.19	12.78	84.92	0.77	83.29	86.42
1	617	670.23	57.70	5.35	15.08	0.77	13.58	16.71
2								
Total	5143	4837.29	248.29	26.61	100.00	0.00	.	.
0	3931	3728.06	195.08	17.06	77.07	0.56	75.91	78.19
1	1212	1109.23	61.50	3.87	22.93	0.56	21.81	24.09

Variance Estimation Method: Taylor Series (WR)
by: Sex 1=Male 2=Female, Major Depressive Episode 1=Yes 0=No.

Sex 1=Male 2=Female Major Depressive Episode 1=Yes 0=No	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	.
0	1.42	100.00	0.00	.	.	.	80.83	0.49	79.83
1	1.42	100.00	0.00	.	.	.	19.17	0.49	18.21
1									
Total	.	47.89	0.53	46.81	48.96	1.05	47.89	0.53	46.81
0	2.08	50.31	0.71	48.88	51.74	1.50	40.66	0.70	39.26
1	2.08	37.66	1.19	35.29	40.10	1.08	7.22	0.34	6.56
2									
Total	.	52.11	0.53	51.04	53.19	1.05	52.11	0.53	51.04
0	0.87	49.69	0.71	48.26	51.12	1.50	40.16	0.54	39.09
1	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 1=Male 2=Female, Major Depressive Episode 1=Yes 0=No.

Sex 1=Male 2=Female / Major Depressive Episode 1=Yes 0=No	Upper 95% Limit TOTPER	DEFF Tot Percent #1
Total		
Total	.	.
0	81.79	1.42
1	20.17	1.42
1		
Total	48.96	1.05
0	42.08	1.87
1	7.95	1.64
2		
Total	53.19	1.05
0	41.25	1.11
1	12.58	0.81

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Stratum-Specific Hypotheses
Variable SEX by Variable MDE
by: Hypothesis Test, Test Statistic.

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

Example 6.7: Comparing the Proportions of U.S. Adult Men and Women with Lifetime Major Depression.

```
                        S U D A A N
          Software for the Statistical Analysis of Correlated Data
        Copyright       Research Triangle Institute       June 2013
                              Release 11.0.1


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: Variable, SUDAAN Reserved Variable One, Contrast.

for: Variable = Major Depressive Episode 1=Yes 0=No.

SUDAAN Reserved Variable One Contrast	Sample Size	Weighted Size	Cntrst Total	Lower 95% Limit Cntrst Total	Upper 95% Limit Cntrst Total	Cntrst Mean	SE Cntrst 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

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

for: Variable = Major Depressive Episode 1=Yes 0=No.

SUDAAN Reserved Variable One Contrast	Lower 95% Limit Cntrst Mean	Upper 95% Limit Cntrst Mean	T-Test Cont.Mean-=0	P-value T-Test Cont. 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

Example 6.8: Testing the Independence of MDE and Gender in U.S. Adults Using the NCS-R data.

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute June 2013
Release 11.0.1

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: Sex 1=Male 2=Female, Major Depressive Episode 1=Yes 0=No.

Sex 1=Male 2=Female Major Depressive Episode 1=Yes 0=No	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	.	.
0	7453	7502.54	349.58	84.95	80.83	0.49	79.83	81.79
1	1829	1779.46	113.96	9.03	19.17	0.49	18.21	20.17
1								
Total	4139	4444.71	215.70	20.08	100.00	0.00	.	.
0	3522	3774.47	169.19	12.78	84.92	0.77	83.29	86.42
1	617	670.23	57.70	5.35	15.08	0.77	13.58	16.71
2								
Total	5143	4837.29	248.29	26.61	100.00	0.00	.	.
0	3931	3728.06	195.08	17.06	77.07	0.56	75.91	78.19
1	1212	1109.23	61.50	3.87	22.93	0.56	21.81	24.09

Variance Estimation Method: Taylor Series (WR)
by: Sex 1=Male 2=Female, Major Depressive Episode 1=Yes 0=No.

Sex 1=Male 2=Female Major Depressive Episode 1=Yes 0=No	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	.
0	1.42	100.00	0.00	.	.	.	80.83	0.49	79.83
1	1.42	100.00	0.00	.	.	.	19.17	0.49	18.21
1									
Total	.	47.89	0.53	46.81	48.96	1.05	47.89	0.53	46.81
0	2.08	50.31	0.71	48.88	51.74	1.50	40.66	0.70	39.26
1	2.08	37.66	1.19	35.29	40.10	1.08	7.22	0.34	6.56
2									
Total	.	52.11	0.53	51.04	53.19	1.05	52.11	0.53	51.04
0	0.87	49.69	0.71	48.26	51.12	1.50	40.16	0.54	39.09
1	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 1=Male 2=Female, Major Depressive Episode 1=Yes 0=No.

Sex 1=Male 2=Female Major Depressive Episode 1=Yes 0=No	Upper 95% Limit TOTPER	DEFF Tot Percent #1
Total		
Total	.	.
0	81.79	1.42
1	20.17	1.42
1		
Total	48.96	1.05
0	42.08	1.87
1	7.95	1.64
2		
Total	53.19	1.05
0	41.25	1.11
1	12.58	0.81

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Stratum-Specific Hypotheses
Variable SEX by Variable MDE
by: Hypothesis Test, Test Statistic.

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

Example 6.9: Testing the Independence of Alcohol Dependence and Education Level in Young Adults (Ages 18-28) using the NCSR data

```
                              S U D A A N
          Software for the Statistical Analysis of Correlated Data
        Copyright       Research Triangle Institute       June 2013
                              Release 11.0.1


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
```


Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18_28 = 1
by: Education 1=0-11 2=12 3=13-15 4=16+ Yrs, Alcohol Dependence 1=Yes 0=No.

Education 1=0-11 2=12 3=13-15 4=16+ Yrs	Alcohol Dependence 1=Yes 0=No	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
1						
	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
2						
	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
3						
	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
4						
	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

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: AGE18_28 = 1
Test Statistics for Stratum-Specific Hypotheses
Variable ED4CAT by Variable ALD
by: Hypothesis Test, Test Statistic.

Hypothesis Test Test Statistic	DF	Test Value	P-Value
CHISQ (Obs - Exp) Wald-F	3	1.0742	0.3703
LLCHISQ (Log-Lin Model) Wald-F	3	1.3837	0.2609

Example 6.10: Simple Logistic Regression to Estimate the NCS-R Male/Female Odds Ratio for Lifetime Major Depressive Episode

```
                              S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright       Research Triangle Institute       June 2013
                                Release 11.0.1


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 C6_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: Major Depressive Episode 1=Yes 0=No
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
Male 1=Yes 0=No	-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: Major Depressive Episode 1=Yes 0=No
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: Major Depressive Episode 1=Yes 0=No
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
Male 1=Yes 0=No	0.597	0.520	0.685

Example 6.11: Using the NCS-R Data to Estimate and Test the Association between Gender and Depression in the U.S. Adult
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.

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute June 2013
Release 11.0.1

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, Male 1=Yes 0=No, Major Depressive Episode 1=Yes 0=No.

for: AGECAT = Total.

Male 1=Yes 0=No		Major Depressive Episode 1=Yes 0=No		
		Total	0	1
Total	Sample Size	9282	7453	1829
	Weighted Size	9282.00	7502.54	1779.46
	Row Percent	100.00	80.83	19.17
	SE Row Percent	0.00	0.49	0.49
	Col Percent	100.00	100.00	100.00
	SE Col Percent	0.00	0.00	0.00
0	Sample Size	5143	3931	1212
	Weighted Size	4837.29	3728.06	1109.23
	Row Percent	100.00	77.07	22.93
	SE Row Percent	0.00	0.56	0.56
	Col Percent	52.11	49.69	62.34
	SE Col Percent	0.53	0.71	1.19
1	Sample Size	4139	3522	617
	Weighted Size	4444.71	3774.47	670.23
	Row Percent	100.00	84.92	15.08
	SE Row Percent	0.00	0.77	0.77
	Col Percent	47.89	50.31	37.66
	SE Col Percent	0.53	0.71	1.19

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, Male 1=Yes 0=No, Major Depressive Episode 1=Yes 0=No.

for: AGECAT = 1.

Male 1=Yes 0=No		Major Depressive Episode 1=Yes 0=No		
		Total	0	1
Total	Sample Size	2104	1698	406
	Weighted Size	2167.86	1765.87	401.99
	Row Percent	100.00	81.46	18.54
	SE Row Percent	0.00	0.83	0.83
	Col Percent	100.00	100.00	100.00
	SE Col Percent	0.00	0.00	0.00
0	Sample Size	1138	870	268
	Weighted Size	1066.19	824.18	242.00
	Row Percent	100.00	77.30	22.70
	SE Row Percent	0.00	1.16	1.16
	Col Percent	49.18	46.67	60.20
	SE Col Percent	1.10	1.33	2.35
1	Sample Size	966	828	138
	Weighted Size	1101.67	941.69	159.99
	Row Percent	100.00	85.48	14.52
	SE Row Percent	0.00	1.21	1.21
	Col Percent	50.82	53.33	39.80
	SE Col Percent	1.10	1.33	2.35

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, Male 1=Yes 0=No, Major Depressive Episode 1=Yes 0=No.

for: AGECAT = 2.

Male 1=Yes 0=No		Major Depressive Episode 1=Yes 0=No		
		Total	0	1
Total	Sample Size	1904	1483	421
	Weighted Size	1696.12	1317.20	378.92
	Row Percent	100.00	77.66	22.34
	SE Row Percent	0.00	1.32	1.32
	Col Percent	100.00	100.00	100.00
	SE Col Percent	0.00	0.00	0.00
0	Sample Size	1040	758	282
	Weighted Size	884.20	640.97	243.23
	Row Percent	100.00	72.49	27.51
	SE Row Percent	0.00	1.54	1.54
	Col Percent	52.13	48.66	64.19
	SE Col Percent	1.20	1.52	2.67
1	Sample Size	864	725	139
	Weighted Size	811.92	676.23	135.70
	Row Percent	100.00	83.29	16.71
	SE Row Percent	0.00	1.88	1.88
	Col Percent	47.87	51.34	35.81
	SE Col Percent	1.20	1.52	2.67

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, Male 1=Yes 0=No, Major Depressive Episode 1=Yes 0=No.

for: AGECAT = 3.

Male 1=Yes 0=No		Major Depressive Episode 1=Yes 0=No		
		Total	0	1
Total	Sample Size	1891	1447	444
	Weighted Size	1978.32	1528.92	449.40
	Row Percent	100.00	77.28	22.72
	SE Row Percent	0.00	0.87	0.87
	Col Percent	100.00	100.00	100.00
	SE Col Percent	0.00	0.00	0.00
0	Sample Size	1003	732	271
	Weighted Size	999.63	741.34	258.29
	Row Percent	100.00	74.16	25.84
	SE Row Percent	0.00	1.08	1.08
	Col Percent	50.53	48.49	57.47
	SE Col Percent	1.15	1.41	2.47
1	Sample Size	888	715	173
	Weighted Size	978.68	787.58	191.11
	Row Percent	100.00	80.47	19.53
	SE Row Percent	0.00	1.57	1.57
	Col Percent	49.47	51.51	42.53
	SE Col Percent	1.15	1.41	2.47

Variance Estimation Method: Taylor Series (WR)
by: AGECAT, Male 1=Yes 0=No, Major Depressive Episode 1=Yes 0=No.

for: AGECAT = 4.

Male 1=Yes 0=No		Major Depressive Episode 1=Yes 0=No		
		Total	0	1
Total	Sample Size	3383	2825	558
	Weighted Size	3439.70	2890.55	549.15
	Row Percent	100.00	84.03	15.97
	SE Row Percent	0.00	0.79	0.79
	Col Percent	100.00	100.00	100.00
	SE Col Percent	0.00	0.00	0.00
0	Sample Size	1962	1571	391
	Weighted Size	1887.27	1521.56	365.71
	Row Percent	100.00	80.62	19.38
	SE Row Percent	0.00	0.90	0.90
	Col Percent	54.87	52.64	66.60
	SE Col Percent	0.97	1.09	2.25
1	Sample Size	1421	1254	167
	Weighted Size	1552.42	1368.99	183.44
	Row Percent	100.00	88.18	11.82
	SE Row Percent	0.00	1.14	1.14
	Col Percent	45.13	47.36	33.40
	SE Col Percent	0.97	1.09	2.25

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)
Test Statistics for Stratum-Specific Hypotheses
Variable SEXM by Variable MDE
by: AGECAT, Hypothesis Test, Test Statistic.

for: AGECAT = Total.

Hypothesis Test / Test Statistic	DF	Test Value	P-Value
CHISQ (Obs - Exp)			
Wald-F	1	87.8030	0.0000

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Stratum-Specific Hypotheses
Variable SEXM by Variable MDE
by: AGECAT, Hypothesis Test, Test Statistic.

for: AGECAT = 1.

Hypothesis Test / Test Statistic	DF	Test Value	P-Value
CHISQ (Obs - Exp)			
Wald-F	1	23.4661	0.0000

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Stratum-Specific Hypotheses
Variable SEXM by Variable MDE
by: AGECAT, Hypothesis Test, Test Statistic.

for: AGECAT = 2.

Hypothesis Test / Test Statistic	DF	Test Value	P-Value
CHISQ (Obs - Exp)			

Wald-F	1	23.2237	0.0000

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Stratum-Specific Hypotheses
Variable SEXM by Variable MDE
by: AGECAT, Hypothesis Test, Test Statistic.

for: AGECAT = 3.

Hypothesis Test / Test Statistic	DF	Test Value	P-Value
CHISQ (Obs - Exp)			
Wald-F	1	9.1165	0.0043

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Stratum-Specific Hypotheses
Variable SEXM by Variable MDE
by: AGECAT, Hypothesis Test, Test Statistic.

for: AGECAT = 4.

Hypothesis Test / Test Statistic	DF	Test Value	P-Value
CHISQ (Obs - Exp)			
Wald-F	1	38.2751	0.0000

Variance Estimation Method: Taylor Series (WR)
Test Statistics for Stratum-Adjusted Hypotheses
Variable SEXM by Variable MDE
Controlling for: Variable AGECAT
by: Hypothesis Test, Test Statistic.

Hypothesis Test / Test Statistic	DF	Test Value	P-Value
CMH General Association			
Wald-F	1	92.4582	0.0000

Not Available in Sudaan: Example 6.12: A Simple Log-linear Model to Test the Association between Lifetime Major Depressi
on Episode and Sex.