

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 LINEAR REGRESSION 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, however output data sets can be saved and used in other software packages. We demonstrate this technique at the end of this document and show how to save output statistics from Sudaan and then graph using SAS PROC GPLOT (residuals*linear prediction) and PROC UNIVARIATE (histogram of residuals).

SUDAAN PROC REGRESS can perform all of the analyses presented in Chapter 7 of ASDA with the exception of the graphical displays. 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, RFORMAT and REFLEVEL for use with formatted variables and optional reference level changes, and a TEST statement for hypothesis tests and many other options for analysis/output. Please see the Sudaan 10.0.1 Language and Examples Guides for additional detail.

```

title "Analysis Example 7.5: Bivariate Testing of Predictors: NHANES" ;
proc regress data=nhanes0506 filetype=sas ;
nest sdmvstra sdmvpsu ;
weight wtmecl2yr ;
class ridreth1 / nofreq;
reflevel ridreth1=1 ;
subpopn age18p=1 ;
model bpxdi1_1 = ridreth1 ;
setenv decwidth=3 ;
print / betas=all tests=all ;
run ;

```

Analysis Example 7.5: Bivariate Testing of Predictors: NHANES

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
Observations used in the analysis	: 4581	Weighted count:190012694
Denominator degrees of freedom	: 15	

Maximum number of estimable parameters for the model is 5

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 248 records
 Minimum cluster size is 98 records

Weighted mean response is 70.611417

Multiple R-Square for the dependent variable BPXDI1_1: 0.004513

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	68.300	0.412	0.412	67.420	69.179	165.587	0.000
1=mex 2=oth hisp 3=white 4=black 5=other							
Mexican	0.000	.	0.000	0.000	0.000	.	.
Other Hispanic	1.592	0.883	1.109	-0.771	3.956	1.436	0.172
White	2.428	0.670	0.554	1.246	3.609	4.380	0.001
Black	3.728	0.789	0.753	2.122	5.333	4.949	0.000
Other	1.785	1.014	1.030	-0.410	3.980	1.733	0.104

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Var Beta
Intercept	0.170
1=mex 2=oth hisp 3=white 4=black 5=other	
Mexican	0.000
Other Hispanic	1.229
White	0.307
Black	0.567
Other	1.061

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite		P-value		P-value		Wald F	P-value Wald F
		Adj DF	Adj F	S_waite Adj F	S_waite Adj ChiSq	S_waite ChiSq	ChiSq		
OVERALL MODEL	5.000	3.154	18500.058	0.000	58342.280	0.000	21588.580	0.000	
MODEL MINUS									
INTERCEPT	4.000	2.657	4.156	0.028	11.041	0.008	7.787	0.001	
INTERCEPT	
RIDRETH1	4.000	2.657	4.156	0.028	11.041	0.008	7.787	0.001	

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	P-value		P-value	
	Adj Wald F	Adj Wald F	Wald ChiSq	Wald ChiSq
OVERALL MODEL	15831.625	0.000	107942.898	0.000
MODEL MINUS				
INTERCEPT	6.229	0.006	31.147	0.000
INTERCEPT
RIDRETH1	6.229	0.006	31.147	0.000

```
proc regress data=nhanes0506 filetype=sas ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
class marcat / nofreq ;
reflevel marcat=1 ;
subpopn age18p=1 ;
model bpxdi1_1 = marcat ;
setenv decwidth=3 ;
print / betas=all tests=all ;
run ;
```

Analysis Example 7.5: Bivariate Testing of Predictors: NHANES

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
Observations used in the analysis	:	4578	Weighted count:189848122
Denominator degrees of freedom	:	15	

Maximum number of estimable parameters for the model is 3

File NHANES0506 contains 30 Clusters

30 clusters were used to fit the model

Maximum cluster size is 247 records

Minimum cluster size is 98 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.018187

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	71.392	4.352	0.468	70.395	72.388	152.696	0.000
1=married 2=prev married 3=never married							
Married	0.000	.	0.000	0.000	0.000	.	.
Previously Married	-0.073	1.976	0.681	-1.525	1.378	-0.108	0.916
Never Married	-4.386	1.390	0.573	-5.608	-3.165	-7.654	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Var Beta
Intercept	0.219
1=married 2=prev married 3=never married	
Married	0.000
Previously Married	0.464
Never Married	0.328

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite Adj DF	S_waite Adj F	P-value S_waite Adj F	S_waite Adj ChiSq	P-value S_waite ChiSq	Wald F	P-value Wald F
OVERALL MODEL	3.000	2.102	21758.207	0.000	45735.355	0.000	15679.632	0.000
MODEL MINUS INTERCEPT	2.000	1.728	27.039	0.000	46.719	0.000	40.157	0.000
INTERCEPT
MARCAT	2.000	1.728	27.039	0.000	46.719	0.000	40.157	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Adj Wald F	P-value Adj Wald F	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	13589.014	0.000	47038.896	0.000
MODEL MINUS INTERCEPT	37.480	0.000	80.314	0.000
INTERCEPT
MARCAT	37.480	0.000	80.314	0.000

```

proc regress data=nhanes0506 filetype=sas ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
subpopn age18p=1 ;
model bpxdi1_1 = female ;
print / betas=all tests=all ;
setenv decwidth=3 ;
run ;
Analysis Example 7.5: Bivariate Testing of Predictors: NHANES

```

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
Observations used in the analysis : 4581   Weighted count:190012694
Denominator degrees of freedom  : 15

```

Maximum number of estimable parameters for the model is 2

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 248 records
 Minimum cluster size is 98 records

Weighted mean response is 70.611417

Multiple R-Square for the dependent variable BPXDI1_1: 0.013316

Date: 03-16-2010
Time: 14:45:55

SUDAAN

Page: 1
Table: 1

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	72.07	2.64	0.42	71.17	72.97	171.37	0.0000
FEMALE	-2.84	1.10	0.38	-3.65	-2.04	-7.51	0.0000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Var Beta
Intercept	0.18
FEMALE	0.14

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite		P-value		P-value		Wald F	P-value
		Adj DF	Adj F	Adj F	Adj ChiSq	Adj ChiSq	Wald F		
OVERALL MODEL	2	1.54	34756.78	0.0000	53609.72	0.0000	24260.14	0.0000	
MODEL MINUS									
INTERCEPT	1	1.00	56.43	0.0000	56.43	0.0000	56.43	0.0000	
INTERCEPT	1	1.00	29368.09	0.0000	29368.09	0.0000	29368.09	0.0000	
FEMALE	1	1.00	56.43	0.0000	56.43	0.0000	56.43	0.0000	

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	P-value		P-value	
	Adj Wald F	Adj Wald F	Wald ChiSq	Wald ChiSq
OVERALL MODEL	22642.80	0.0000	48520.28	0.0000
MODEL MINUS				
INTERCEPT	56.43	0.0000	56.43	0.0000
INTERCEPT	29368.09	0.0000	29368.09	0.0000
FEMALE	56.43	0.0000	56.43	0.0000

```

proc regress data=nhanes0506 filetype=sas ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
subpopn age18p=1 ;
model bpxdi1_1 = agec ;
print / betas=all tests=all ;
setenv decwidth=3 ;
run ;
Analysis Example 7.5: Bivariate Testing of Predictors: NHANES

```

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
Observations used in the analysis : 4581      Weighted count:190012694
Denominator degrees of freedom  : 15

```

Maximum number of estimable parameters for the model is 2

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 248 records
 Minimum cluster size is 98 records

Weighted mean response is 70.611417

Multiple R-Square for the dependent variable BPXDI1_1: 0.006472

Date: 03-16-2010
Time: 14:47:46

SUDAAN

Page: 1
Table: 1

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	70.62	3.71	0.35	69.87	71.36	201.94	0.0000
AGEC	0.06	3.88	0.02	0.01	0.10	2.77	0.0142

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Identity
Response variable BPXDI1_1: BPXDI1_1
For Subpopulation: AGE18P = 1
by: Independent Variables and Effects.

```

-----
Independent
  Variables and
  Effects              Var Beta
-----
Intercept              0.12
AGEC                   0.00
-----

```

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Identity
Response variable BPXDI1_1: BPXDI1_1
For Subpopulation: AGE18P = 1
by: Contrast.

```

-----
Contrast              Degrees of Freedom    S_waite Adj DF    S_waite Adj F    P-value S_waite Adj F    S_waite Adj ChiSq    P-value S_waite Adj ChiSq    Wald F    P-value Wald F
-----
OVERALL MODEL              2          1.85    19977.19    0.0000    36956.32    0.0000    22061.12    0.0000
MODEL MINUS
  INTERCEPT              1          1.00         7.70    0.0142         7.70    0.0055         7.70    0.0142
INTERCEPT              1          1.00    40780.47    0.0000    40780.47    0.0000    40780.47    0.0000
AGEC                      1          1.00         7.70    0.0142         7.70    0.0055         7.70    0.0142
-----

```

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Identity
Response variable BPXDI1_1: BPXDI1_1
For Subpopulation: AGE18P = 1
by: Contrast.

```

-----
Contrast              Adj Wald F    P-value Adj Wald F    Wald ChiSq    P-value Wald ChiSq
-----
OVERALL MODEL              20590.38    0.0000    44122.24    0.0000
MODEL MINUS
  INTERCEPT              7.70    0.0142         7.70    0.0055
INTERCEPT              40780.47    0.0000    40780.47    0.0000
AGEC                      7.70    0.0142         7.70    0.0055
-----

```

```

proc regress data=nhanes0506 design=srs ;
rtitle "Analysis Example 7.5: Unweighted and Without Design Correction: NHANES" ;
class ridreth1 marcat / nofreq ;
reflevel ridreth1=1 marcat=1 ;
subpopn age18p=1 ;
model bpxdi1_1 = ridreth1 marcat female agec ;
setenv decwidth=3 ;
print / betas=all tests=all style=nchs ;
run ;
The SAS System

```

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 Simple Random Sample (SRS) Design

```

Number of observations read      : 10348   Weighted count: 10348
Observations in subpopulation   : 5563    Weighted count: 5563
Observations used in the analysis : 4578   Weighted count: 4578
Denominator degrees of freedom  : 10347

```

Maximum number of estimable parameters for the model is 9

```

File NHANES0506 contains 10348 Clusters
4578 clusters were used to fit the model
Maximum cluster size is 1 records
Minimum cluster size is 1 records

```

Weighted mean response is 68.991263

Multiple R-Square for the dependent variable BPXDI1_1: 0.059885

```

Date: 04-01-2010          SUDAAN          Page: 1
Time: 09:01:06           Table: 1

```

```

Variance Estimation Method: Taylor Series (SRS)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Identity
Response variable BPXDI1_1: BPXDI1_1
For Subpopulation: AGE18P = 1
Analysis Example 7.5: Unweighted and Without Design Correction: NHANES
by: Independent Variables and Effects.

```

Independent Variables and Effects	Beta Coeff.	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta
Intercept	69.672	0.441	68.807	70.537
1=mex 2=oth hisp 3=white 4=black 5=other				
1	0.000	0.000	0.000	0.000
2	1.898	1.063	-0.185	3.982
3	1.672	0.465	0.761	2.582
4	4.508	0.555	3.421	5.596
5	2.312	0.980	0.390	4.233
1=married 2=prev married 3=never married				
1	0.000	0.000	0.000	0.000
2	0.327	0.554	-0.759	1.413
3	-4.216	0.510	-5.216	-3.216
FEMALE	-3.402	0.375	-4.137	-2.667
AGEC	0.039	0.013	0.014	0.064

Variance Estimation Method: Taylor Series (SRS)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 Analysis Example 7.5: Unweighted and Without Design Correction: NHANES
 by: Independent Variables and Effects.

Independent Variables and Effects	P-value		Var Beta
	T-Test B=0	T-Test B=0	
Intercept	157.884	0.000	0.195
1=mex 2=oth hisp 3=white 4=black 5=other			
1	.	.	0.000
2	1.786	0.074	1.130
3	3.599	0.000	0.216
4	8.125	0.000	0.308
5	2.359	0.018	0.961
1=married 2=prev married 3=never married			
1	.	.	0.000
2	0.590	0.555	0.307
3	-8.265	0.000	0.260
FEMALE	-9.069	0.000	0.141
AGEC	3.087	0.002	0.000

Variance Estimation Method: Taylor Series (SRS)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 Analysis Example 7.5: Unweighted and Without Design Correction: NHANES
 by: Contrast.

Contrast	Degrees of Freedom	S_waite		P-value	
		Adj DF	Adj F	Adj F	Adj ChiSq
OVERALL MODEL	9.000	8.782	15284.563	0.000	134235.451
MODEL MINUS					
INTERCEPT	8.000	7.857	35.655	0.000	280.122
INTERCEPT
RIDRETH1	4.000	3.974	17.447	0.000	69.342
MARCAT	2.000	1.991	33.062	0.000	65.840
FEMALE	1.000	1.000	82.250	0.000	82.250
AGEC	1.000	1.000	9.532	0.002	9.532

Variance Estimation Method: Taylor Series (SRS)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 Analysis Example 7.5: Unweighted and Without Design Correction: NHANES
 by: Contrast.

Contrast	P-value	Wald F	P-value	Adj Wald F	P-value
	S_waite ChiSq		Wald F		Adj Wald F
OVERALL MODEL	0.000	15953.272	0.000	15940.938	0.000
MODEL MINUS					
INTERCEPT	0.000	38.132	0.000	38.106	0.000
INTERCEPT
RIDRETH1	0.000	16.792	0.000	16.787	0.000
MARCAT	0.000	34.877	0.000	34.874	0.000
FEMALE	0.000	82.250	0.000	82.250	0.000
AGEC	0.002	9.532	0.002	9.532	0.002

Variance Estimation Method: Taylor Series (SRS)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 Analysis Example 7.5: Unweighted and Without Design Correction: NHANES
 by: Contrast.

Contrast	P-value	
	Wald ChiSq	Wald ChiSq
OVERALL MODEL	143579.452	0.000
MODEL MINUS		
INTERCEPT	305.053	0.000
INTERCEPT	.	.
RIDRETH1	67.168	0.000
MARCAT	69.754	0.000
FEMALE	82.250	0.000
AGEC	9.532	0.002

```
title "Analysis Example 7.5: Weighted and Without Design Correction: NHANES" ;
```

```
proc regress data=nhanes0506 ;  
nest strata_one ;  
class ridreth1 marcat / nofreq ;  
weight wtmec2yr ;  
reflevel ridreth1=1 marcat=1 ;  
subpopn age18p=1 ;  
model bpxdi1_1 = ridreth1 marcat female agec ;  
setenv decwidth=3 ;  
print / betas=all tests=all ;  
run ;
```

Analysis Example 7.5: Weighted and Without Design Correction: NHANES

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): STRATA_ONE
Primary Sampling Unit: Observation Number

Number of observations read	:	9950	Weighted count:291616892
Number of observations skipped	:	398	
(WEIGHT variable nonpositive)			
Observations in subpopulation	:	5334	Weighted count:217700471
Observations used in the analysis	:	4578	Weighted count:189848122
Denominator degrees of freedom	:	9949	

Maximum number of estimable parameters for the model is 9

File NHANES0506 contains 9950 Clusters
4578 clusters were used to fit the model
Maximum cluster size is 1 records
Minimum cluster size is 1 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.039028

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	70.678	0.537	0.488	69.721	71.636	144.709	0.000
1=mex 2=oth hisp 3=white 4=black 5=other							
1	0.000	.	0.000	0.000	0.000	.	.
2	1.787	1.269	1.307	-0.775	4.348	1.367	0.172
3	2.192	0.592	0.518	1.176	3.208	4.229	0.000
4	4.409	0.528	0.611	3.211	5.606	7.216	0.000
5	1.958	1.065	1.039	-0.077	3.994	1.886	0.059
1=married 2=prev married 3=never married							
1	0.000	.	0.000	0.000	0.000	.	.
2	0.017	1.729	0.662	-1.280	1.315	0.026	0.979
3	-4.356	1.468	0.635	-5.601	-3.112	-6.862	0.000
FEMALE	-2.997	1.489	0.440	-3.860	-2.135	-6.811	0.000
AGEC	0.017	1.548	0.015	-0.012	0.046	1.141	0.254

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Var Beta
Intercept	0.239
1=mex 2=oth hisp 3=white 4=black 5=other	
1	0.000
2	1.707
3	0.269
4	0.373
5	1.079
1=married 2=prev married 3=never married	
1	0.000
2	0.438
3	0.403
FEMALE	0.194
AGEC	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite Adj DF	S_waite Adj F	P-value S_waite Adj F	S_waite Adj ChiSq	P-value S_waite ChiSq	Wald F	P-value Wald F
OVERALL MODEL	9.000	7.939	12987.093	0.000	103110.630	0.000	14412.186	0.000
MODEL MINUS								
INTERCEPT	8.000	7.157	17.618	0.000	126.086	0.000	21.633	0.000
INTERCEPT
RIDRETH1	4.000	3.387	6.369	0.000	21.570	0.000	13.031	0.000
MARCAT	2.000	1.985	21.776	0.000	43.235	0.000	23.670	0.000
FEMALE	1.000	1.000	46.388	0.000	46.388	0.000	46.388	0.000
AGEC	1.000	1.000	1.302	0.254	1.302	0.254	1.302	0.254

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Adj Wald F	P-value Adj Wald F	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	14400.597	0.000	129709.675	0.000
MODEL MINUS				
INTERCEPT	21.617	0.000	173.061	0.000
INTERCEPT
RIDRETH1	13.027	0.000	52.124	0.000
MARCAT	23.668	0.000	47.341	0.000
FEMALE	46.388	0.000	46.388	0.000
AGEC	1.302	0.254	1.302	0.254

```

title "Analysis Example 7.5: Weighted and With Design Correction: NHANES" ;
proc regress data=nhanes0506 deft4;
nest sdmvstra sdmvpsu ;
class ridreth1 marcat / nofreq ;
weight wtmec2yr ;
reflevel ridreth1=1 marcat=1 ;
subpopn age18p=1 ;
model bpxdi1_1 = ridreth1 marcat female agec ;
setenv decwidth=3 ;
print / betas=all tests=all ;
run ;

```

Analysis Example 7.5: Weighted and With Design Correction: NHANES

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
Observations used in the analysis	: 4578	Weighted count:189848122
Denominator degrees of freedom	: 15	

Maximum number of estimable parameters for the model is 9

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 247 records
 Minimum cluster size is 98 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.039028

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	70.678	0.564	0.501	69.611	71.745	141.141	0.000
1=mex 2=oth hisp 3=white 4=black 5=other							
1	0.000	.	0.000	0.000	0.000	.	.
2	1.787	0.969	1.142	-0.648	4.221	1.564	0.139
3	2.192	0.806	0.605	0.903	3.481	3.624	0.002
4	4.409	0.820	0.761	2.786	6.031	5.792	0.000
5	1.958	0.964	0.988	-0.148	4.064	1.982	0.066
1=married 2=prev married 3=never married							
1	0.000	.	0.000	0.000	0.000	.	.
2	0.017	2.034	0.718	-1.513	1.547	0.024	0.981
3	-4.356	1.162	0.565	-5.560	-3.152	-7.710	0.000
FEMALE	-2.997	0.843	0.331	-3.703	-2.292	-9.052	0.000
AGEC	0.017	3.321	0.022	-0.030	0.064	0.779	0.448

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Var Beta
Intercept	0.251
1=mex 2=oth hisp 3=white 4=black 5=other	
1	0.000
2	1.305
3	0.366
4	0.579
5	0.976
1=married 2=prev married 3=never married	
1	0.000
2	0.515
3	0.319
FEMALE	0.110
AGEC	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite		P-value		Wald F	P-value	
		Adj DF	Adj F	Adj F	Adj ChiSq			ChiSq
OVERALL MODEL	9.000	4.739	9402.661	0.000	44556.050	0.000	12543.800	0.000
MODEL MINUS INTERCEPT	8.000	4.381	14.779	0.000	64.742	0.000	23.743	0.000
INTERCEPT
RIDRETH1	4.000	2.843	5.863	0.008	16.665	0.001	9.446	0.001
MARCAT	2.000	1.789	22.248	0.000	39.807	0.000	32.736	0.000
FEMALE	1.000	1.000	81.941	0.000	81.941	0.000	81.941	0.000
AGEC	1.000	1.000	0.607	0.448	0.607	0.436	0.607	0.448

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	P-value		P-value	
	Adj Wald F	Adj Wald F	Wald ChiSq	Wald ChiSq
OVERALL MODEL	5853.773	0.000	112894.197	0.000
MODEL MINUS INTERCEPT	12.663	0.001	189.944	0.000
INTERCEPT
RIDRETH1	7.557	0.003	37.786	0.000
MARCAT	30.554	0.000	65.473	0.000
FEMALE	81.941	0.000	81.941	0.000
AGEC	0.607	0.448	0.607	0.436

```

* run weighted and design corrected regression;
title "Analysis Example 7.5: Weighted and With Design Correction Final Model with Age Squared: NHANES" ;
proc regress data=nhanes0506 deft4;
nest sdmvstra sdmvpsu ;
class ridreth1 marcat / nofreq ;
weight wtmec2yr ;
reflevel ridreth1=1 marcat=1 ;
subpopn age18p=1 ;
model bpxdi1_1 = ridreth1 marcat female agec agecsq ;
setenv decwidth=3 ;
print / betas=all tests=all ;
run ;

```

Analysis Example 7.5: Weighted and With Design Correction Final Model with Age Squared: NHANES

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
Observations used in the analysis	: 4578	Weighted count:189848122
Denominator degrees of freedom	: 15	

Maximum number of estimable parameters for the model is 10

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 247 records
 Minimum cluster size is 98 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.133543

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	73.859	0.491	0.455	72.889	74.829	162.369	0.000
1=mex 2=oth hisp 3=white 4=black 5=other							
1	0.000	.	0.000	0.000	0.000	.	.
2	1.189	0.972	1.087	-1.127	3.505	1.094	0.291
3	1.781	0.971	0.631	0.436	3.125	2.823	0.013
4	3.465	0.950	0.779	1.804	5.126	4.447	0.000
5	1.189	0.954	0.934	-0.803	3.180	1.272	0.223
1=married 2=prev married 3=never married							
1	0.000	.	0.000	0.000	0.000	.	.
2	1.040	1.677	0.622	-0.285	2.366	1.673	0.115
3	-0.343	1.209	0.582	-1.583	0.897	-0.590	0.564
FEMALE	-2.721	0.970	0.338	-3.441	-2.002	-8.061	0.000
AGEC	0.125	1.432	0.015	0.094	0.157	8.454	0.000
AGECSQ	-0.012	1.867	0.001	-0.014	-0.011	-16.336	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Var Beta
Intercept	0.207
1=mex 2=oth hisp 3=white 4=black 5=other	
1	0.000
2	1.181
3	0.398
4	0.607
5	0.873
1=married 2=prev married 3=never married	
1	0.000
2	0.387
3	0.339
FEMALE	0.114
AGEC	0.000
AGECSQ	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite		P-value		P-value		Wald F	P-value
		Adj DF	Adj F	Adj F	Adj ChiSq	ChiSq	Wald F		
OVERALL MODEL	10.000	5.176	11119.301	0.000	57553.870	0.000	21893.478	0.000	
MODEL MINUS									
INTERCEPT	9.000	5.454	59.273	0.000	323.291	0.000	186.676	0.000	
INTERCEPT	
RIDRETH1	4.000	2.849	4.098	0.027	11.675	0.007	6.398	0.003	
MARCAT	2.000	1.936	1.899	0.185	3.677	0.151	1.613	0.232	
FEMALE	1.000	1.000	64.985	0.000	64.985	0.000	64.985	0.000	
AGEC	1.000	1.000	71.462	0.000	71.462	0.000	71.462	0.000	
AGECSQ	1.000	1.000	266.864	0.000	266.864	0.000	266.864	0.000	

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	P-value		P-value	
	Adj Wald F	Adj Wald F	Wald ChiSq	Wald ChiSq
OVERALL MODEL	8757.391	0.000	218934.781	0.000
MODEL MINUS				
INTERCEPT	87.115	0.000	1680.084	0.000
INTERCEPT
RIDRETH1	5.118	0.012	25.590	0.000
MARCAT	1.506	0.256	3.227	0.199
FEMALE	64.985	0.000	64.985	0.000
AGEC	71.462	0.000	71.462	0.000
AGECSQ	266.864	0.000	266.864	0.000

```

title "Analysis Example 7.5 : Final Model with Main Effects and Interactions: NHANES " ;
proc regress data=nhanes0506 filetype=sas ;
nest sdmvstra sdmvpsu ;
weight wtmecl2yr ;
subpopn age18p=1 ;
class ridreth1 marcat other black white othhis nevmar prevmar female / nofreq;
reflevel other=0 black=0 white=0 othhis=0 nevmar=0 prevmar=0 female=0 ;
model bpxdi1_1 = other black white othhis nevmar prevmar female agec agecsq
      other*agec black*agec white*agec othhis*agec
      other*agecsq black*agecsq white*agecsq othhis*agecsq ;
effects other*agec black*agec white*agec othhis*agec other*agecsq black*agecsq white*agecsq othhis*agecsq
      / name="Race and Age Interactions" ;
print / tests=all ;
run ;

```

Analysis Example 7.5 : Final Model with Main Effects and Interactions: NHANES

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
Observations used in the analysis	: 4578	Weighted count:189848122
Denominator degrees of freedom	: 15	

WARNING: DDF (15) < maximum number of independent parameters in the model (18)
 Tests of hypothesis may be different for different choices of reference level

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 247 records
 Minimum cluster size is 98 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.135070

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXD11_1: BPXD11_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S _{waite} Adj DF	S _{waite} Adj F	P-value S _{waite} Adj F	S _{waite} Adj ChiSq	P-value S _{waite} ChiSq	Wald F	P-value Wald F
OVERALL MODEL	15	6.74	7652.38	0.0000	51594.77	0.0000	56789.68	0.0000
MODEL MINUS								
INTERCEPT	15	7.22	37.60	0.0000	271.37	0.0000	8502.08	0.0000
INTERCEPT
OTHER	1	1.00	0.36	0.5581	0.36	0.5492	0.36	0.5581
BLACK	1	1.00	10.86	0.0049	10.86	0.0010	10.86	0.0049
WHITE	1	1.00	6.32	0.0239	6.32	0.0120	6.32	0.0239
OTHHIS	1	1.00	0.24	0.6340	0.24	0.6270	0.24	0.6340
NEVMAR	1	1.00	0.33	0.5752	0.33	0.5667	0.33	0.5752
PREVMAR	1	1.00	2.51	0.1337	2.51	0.1129	2.51	0.1337
FEMALE	1	1.00	63.29	0.0000	63.29	0.0000	63.29	0.0000
AGEC
AGECSQ
AGEC * OTHER	1	1.00	2.92	0.1080	2.92	0.0874	2.92	0.1080
AGECSQ * OTHER	1	1.00	0.00	0.9453	0.00	0.9443	0.00	0.9453
AGEC * BLACK	1	1.00	1.26	0.2785	1.26	0.2609	1.26	0.2785
AGECSQ * BLACK	1	1.00	1.37	0.2594	1.37	0.2411	1.37	0.2594
AGEC * WHITE	1	1.00	0.11	0.7425	0.11	0.7379	0.11	0.7425
AGECSQ * WHITE	1	1.00	0.94	0.3484	0.94	0.3331	0.94	0.3484
AGEC * OTHHIS	1	1.00	0.75	0.3998	0.75	0.3862	0.75	0.3998
AGECSQ * OTHHIS	1	1.00	1.36	0.2625	1.36	0.2443	1.36	0.2625
Race and Age Interactions	8	4.24	1.13	0.3796	4.82	0.3386	1.84	0.1461

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	P-value		P-value	
	Adj Wald F	Adj Wald F	Wald ChiSq	Wald ChiSq
OVERALL MODEL	3785.98	0.0128	851845.14	0.0000
MODEL MINUS				
INTERCEPT	566.81	0.0329	127531.27	0.0000
INTERCEPT
OTHER	0.36	0.5581	0.36	0.5492
BLACK	10.86	0.0049	10.86	0.0010
WHITE	6.32	0.0239	6.32	0.0120
OTHHIS	0.24	0.6340	0.24	0.6270
NEVMAR	0.33	0.5752	0.33	0.5667
PREVMAR	2.51	0.1337	2.51	0.1129
FEMALE	63.29	0.0000	63.29	0.0000
AGEC
AGECSQ
AGEC * OTHER	2.92	0.1080	2.92	0.0874
AGECSQ * OTHER	0.00	0.9453	0.00	0.9443
AGEC * BLACK	1.26	0.2785	1.26	0.2609
AGECSQ * BLACK	1.37	0.2594	1.37	0.2411
AGEC * WHITE	0.11	0.7425	0.11	0.7379
AGECSQ * WHITE	0.94	0.3484	0.94	0.3331
AGEC * OTHHIS	0.75	0.3998	0.75	0.3862
AGECSQ * OTHHIS	1.36	0.2625	1.36	0.2443
Race and Age				
Interactions	0.98	0.5091	14.75	0.0641

```

title "Analysis Example 7.5 : Final Model with Main Effects and Interactions: NHANES " ;
proc regress data=nhanes0506 filetype=sas ;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
subpopn age18p=1 ;
class ridreth1 marcat other black white othhis nevmar prevmar female / nofreq;
reflevel other=0 black=0 white=0 othhis=0 nevmar=0 prevmar=0 female=0 ridreth1=1 marcat=1 ;
model bpxdi1_1 = ridreth1 marcat female agec agecsq female*agec female*agecsq ;
effects female*agec female*agecsq / name="Gender and Age Interactions" ;
print / tests=all ;
run ;

```

Analysis Example 7.5 : Final Model with Main Effects and Interactions: NHANES

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
Observations used in the analysis	: 4578	Weighted count:189848122
Denominator degrees of freedom	: 15	

Maximum number of estimable parameters for the model is 12

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 247 records
 Minimum cluster size is 98 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.134374

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite		P-value		P-value		Wald F	P-value Wald F
		Adj DF	Adj F	Adj F	Adj ChiSq	ChiSq	ChiSq		
OVERALL MODEL	12	5.95	9052.08	0.0000	53824.50	0.0000	22964.91	0.0000	
MODEL MINUS INTERCEPT	11	5.94	45.84	0.0000	272.23	0.0000	287.90	0.0000	
INTERCEPT	
RIDRETH1	4	2.87	4.15	0.0262	11.90	0.0068	6.45	0.0032	
MARCAT	2	1.92	1.41	0.2751	2.69	0.2455	1.19	0.3311	
FEMALE	1	1.00	20.59	0.0004	20.59	0.0000	20.59	0.0004	
AGEC	
AGECSQ	
AGEC * FEMALE	1	1.00	0.25	0.6210	0.25	0.6137	0.25	0.6210	
AGECSQ * FEMALE	1	1.00	1.16	0.2983	1.16	0.2813	1.16	0.2983	
Gender and Age Interactions	2	1.64	1.17	0.3260	1.93	0.3023	1.86	0.1905	

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	P-value		P-value	
	Adj Wald F	Adj Wald F	Wald ChiSq	Wald ChiSq
OVERALL MODEL	6123.98	0.0000	275578.91	0.0000
MODEL MINUS INTERCEPT	95.97	0.0000	3166.88	0.0000
INTERCEPT
RIDRETH1	5.16	0.0118	25.80	0.0000
MARCAT	1.11	0.3564	2.38	0.3039
FEMALE	20.59	0.0004	20.59	0.0000
AGEC
AGECSQ
AGEC * FEMALE	0.25	0.6210	0.25	0.6137
AGECSQ * FEMALE	1.16	0.2983	1.16	0.2813
Gender and Age Interactions	1.73	0.2127	3.71	0.1563

```

title "Analysis Example 7.5: Final Model of Main Effects: NHANES " ;
* note use of default deft4 for design effects, see Sudaan documentation for details ;
proc regress data=nhanes0506 filetype=sas deft4;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
subpopn age18p=1 ;
class ridreth1 marcat / nofreq ;
reflevel ridreth1=1 marcat=1 ;
model bpxdi1_1 = ridreth1 marcat female agec agecsq ;
setenv decwidth=3 ;
print / betas=all tests=all ;
run ;

```

Analysis Example 7.5: Final Model of Main Effects: NHANES

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
Observations used in the analysis	: 4578	Weighted count:189848122
Denominator degrees of freedom	: 15	

Maximum number of estimable parameters for the model is 10

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 247 records
 Minimum cluster size is 98 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.133543

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta
Intercept	73.859	0.491	0.455	72.889
1=mex 2=oth hisp 3=white 4=black 5=other				
1	0.000	.	0.000	0.000
2	1.189	0.972	1.087	-1.127
3	1.781	0.971	0.631	0.436
4	3.465	0.950	0.779	1.804
5	1.189	0.954	0.934	-0.803
1=married 2=prev married 3=never married				
1	0.000	.	0.000	0.000
2	1.040	1.677	0.622	-0.285
3	-0.343	1.209	0.582	-1.583
FEMALE	-2.721	0.970	0.338	-3.441
AGEC	0.125	1.432	0.015	0.094
AGECSQ	-0.012	1.867	0.001	-0.014

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0	Var Beta
Intercept	74.829	162.369	0.000	0.207
1=mex 2=oth hisp 3=white 4=black 5=other				
1	0.000	.	.	0.000
2	3.505	1.094	0.291	1.181
3	3.125	2.823	0.013	0.398
4	5.126	4.447	0.000	0.607
5	3.180	1.272	0.223	0.873
1=married 2=prev married 3=never married				
1	0.000	.	.	0.000
2	2.366	1.673	0.115	0.387
3	0.897	-0.590	0.564	0.339
FEMALE	-2.002	-8.061	0.000	0.114
AGEC	0.157	8.454	0.000	0.000
AGECSQ	-0.011	-16.336	0.000	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	Degrees of Freedom	S_waite		P-value	
		Adj DF	Adj F	Adj F	Adj ChiSq
OVERALL MODEL	10.000	5.176	11119.301	0.000	57553.870
MODEL MINUS INTERCEPT	9.000	5.454	59.273	0.000	323.291
INTERCEPT
RIDRETH1	4.000	2.849	4.098	0.027	11.675
MARCAT	2.000	1.936	1.899	0.185	3.677
FEMALE	1.000	1.000	64.985	0.000	64.985
AGEC	1.000	1.000	71.462	0.000	71.462
AGECSQ	1.000	1.000	266.864	0.000	266.864

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

Contrast	P-value		P-value		P-value
	S_waite ChiSq	Wald F	Wald F	Adj Wald F	Adj Wald F
OVERALL MODEL	0.000	21893.478	0.000	8757.391	0.000
MODEL MINUS INTERCEPT	0.000	186.676	0.000	87.115	0.000
INTERCEPT
RIDRETH1	0.007	6.398	0.003	5.118	0.012
MARCAT	0.151	1.613	0.232	1.506	0.256
FEMALE	0.000	64.985	0.000	64.985	0.000
AGEC	0.000	71.462	0.000	71.462	0.000
AGECSQ	0.000	266.864	0.000	266.864	0.000

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Identity
Response variable BPXDI1_1: BPXDI1_1
For Subpopulation: AGE18P = 1
by: Contrast.

Contrast	P-value	
	Wald ChiSq	Wald ChiSq
OVERALL MODEL	218934.781	0.000
MODEL MINUS INTERCEPT	1680.084	0.000
INTERCEPT	.	.
RIDRETH1	25.590	0.000
MARCAT	3.227	0.199
FEMALE	64.985	0.000
AGEC	71.462	0.000
AGECSQ	266.864	0.000

```

* example of saving output to a SAS file for further use ;
* residual analysis uses graphics not available in Sudaan ;

title "Analysis Example 7.5: Final Model of Main Effects: NHANES " ;

* note use of default deft4 for design effects, see Sudaan documentation for details ;

proc regress data=nhanes0506 filetype=sas deft4;
nest sdmvstra sdmvpsu ;
weight wtmec2yr ;
subpopn age18p=1 ;
class ridreth1 marcat / nofreq ;
reflevel ridreth1=1 marcat=1 ;
model bpxdi1_1 = ridreth1 marcat female agec agecsq ;
setenv deewidth=3 ;
*print / betas=all tests=all ;
output residual expected / filetype=sas filename=work.regressresiduals replace ;
run ;

proc means data=work.regressresiduals ;
run ;

```

Analysis Example 7.5: Final Model of Main Effects: NHANES

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
Observations used in the analysis	: 4578	Weighted count:189848122
Denominator degrees of freedom	: 15	

Maximum number of estimable parameters for the model is 10

File NHANES0506 contains 30 Clusters
 30 clusters were used to fit the model
 Maximum cluster size is 247 records
 Minimum cluster size is 98 records

Weighted mean response is 70.610633

Multiple R-Square for the dependent variable BPXDI1_1: 0.133543

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Beta Coeff.	DEFF Beta #4	SE Beta	Lower 95% Limit Beta
Intercept	73.859	0.491	0.455	72.889
1=mex 2=oth hisp 3=white 4=black 5=other				
1	0.000	.	0.000	0.000
2	1.189	0.972	1.087	-1.127
3	1.781	0.971	0.631	0.436
4	3.465	0.950	0.779	1.804
5	1.189	0.954	0.934	-0.803
1=married 2=prev married 3=never married				
1	0.000	.	0.000	0.000
2	1.040	1.677	0.622	-0.285
3	-0.343	1.209	0.582	-1.583
FEMALE	-2.721	0.970	0.338	-3.441
AGEC	0.125	1.432	0.015	0.094
AGECSQ	-0.012	1.867	0.001	-0.014

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Independent Variables and Effects.

Independent Variables and Effects	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	74.829	162.369	0.000
1=mex 2=oth hisp 3=white 4=black 5=other			
1	0.000	.	.
2	3.505	1.094	0.291
3	3.125	2.823	0.013
4	5.126	4.447	0.000
5	3.180	1.272	0.223
1=married 2=prev married 3=never married			
1	0.000	.	.
2	2.366	1.673	0.115
3	0.897	-0.590	0.564
FEMALE	-2.002	-8.061	0.000
AGEC	0.157	8.454	0.000
AGECSQ	-0.011	-16.336	0.000

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Identity
 Response variable BPXDI1_1: BPXDI1_1
 For Subpopulation: AGE18P = 1
 by: Contrast.

```
-----
```

Contrast	Degrees of Freedom	Wald F	P-value Wald F
OVERALL MODEL	10.000	21893.478	0.000
MODEL MINUS			
INTERCEPT	9.000	186.676	0.000
INTERCEPT	.	.	.
RIDRETH1	4.000	6.398	0.003
MARCAT	2.000	1.613	0.232
FEMALE	1.000	64.985	0.000
AGEC	1.000	71.462	0.000
AGECSQ	1.000	266.864	0.000

```
-----
```

Analysis Example 7.5: Final Model of Main Effects: NHANES

The MEANS Procedure

```
-----
```

Variable	Label	N	Mean	Std Dev	Minimum
PROCNUM	Procedure Number	9950	4.0000000	0	4.0000000
RESIDUAL	Residual	5526	0.0233139	11.7070323	-60.0204935
EXPECTED	Expected	6423	67.4341033	6.1853206	54.3787294

```
-----
```

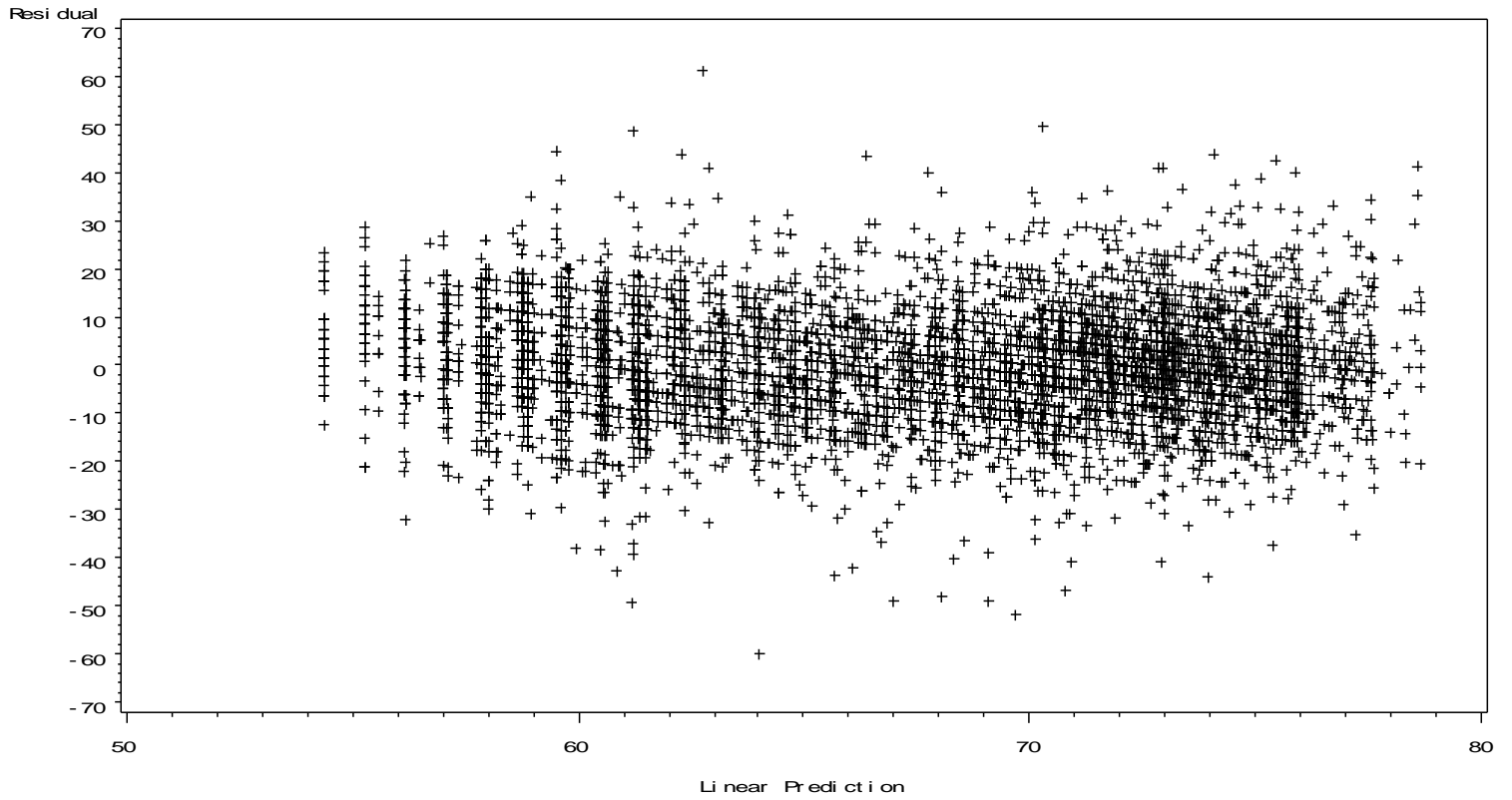
```
-----
```

Variable	Label	Maximum
PROCNUM	Procedure Number	4.0000000
RESIDUAL	Residual	61.2439364
EXPECTED	Expected	78.6772230

```
-----
```

```
proc gplot data=work.regressresiduals ;  
plot residual*expected ;  
title "NHANES: Final Model Diagnostics: Residual by Expected" ;  
label expected='Linear Prediction' ;  
run ;
```

NHANES: Final Model Diagnostics: Residual by Expected



```
symbol1 color=red ;  
proc univariate data=work.regressresiduals ;  
  histogram residual / normal;  
  title "NHANES: Final Model Diagnostics: Histogram of Residuals" ;  
run ;
```

NHANES: Final Model Diagnostics: Histogram of Residuals

