

SUDAAN Analysis Example Replication C5

* SUDAAN Analysis Examples Replication for ASDA 2nd Edition, SAS v9.4 TS1M3 ;
* Berglund April 2017
* Chapter 5 ;

```
libname d "P:\ASDA 2\Data sets\nhanes 2011_2012\" ;
ods graphics off ;
options nodate nonumber ;
title ;
* use list output since Sudaan does not work well with other ODS destinations ;
ods listing ;
data _null_ ;
file print ;
put "Examples 5.1 and 5.2 are not available in Sudaan software since it does not offer any graphing abilities"
;
run ;
libname ncsr "P:\ASDA 2\Data sets\ncsr\" ;
data c5_ncsr ;
  set ncsr.ncsr_sub_13nov2015 ;
  * create variables needed for NCSR examples ;
  ncsrwtsh_pop = ncsrwtsh * (209128094 / 9282) ;
run ;
proc sort ;
  by sestrat seclustr ;
run ;
title "ASDA2 Example 5.3 Population Totals of MDE, NCSR Data" ;
options ls=120 ps=64 ;
proc descript data=c5_ncsr filetype=sas deft1 ;
  nest sestrat seclustr ;
  weight ncsrwtsh_pop ;
  var mde ;
  print total settotal lowtotal uptotal defftotal ;
  setenv decwidth=3 ;
run ;
title "ASDA2 Example 5.3 Population Totals of MDE by Marital Status, NCSR Data" ;
options ls=120 ps=64 ;
proc descript data=c5_ncsr filetype=sas deft1 ;
  nest sestrat seclustr ;
  weight ncsrwtsh_pop ;
  class mar3cat / nofreq ;
  var mde ;
  print total settotal lowtotal uptotal defftotal ;
  setenv decwidth=3 ;
run ;
title "Example 5.4 : Mean of Total HH Wealth using HRS 2012 data " ;
libname hrs "P:\ASDA 2\Data sets\HRS 2012\" ;
data c5_hrs ;
  set hrs.hrs sub 28sep2016 ;
  if nfinr=1 then finr=1 ; else if nfinr=0 then finr=0 ; else finr=. ;
  if gender=2 then female=1 ; else female=0 ;
  if nage >=70 then age70=1 ; else age70=0 ;
run ;
proc sort ;
  by stratum secu ;
run ;
proc descript data=c5_hrs filetype=sas deft1 ;
  nest stratum secu ;
  weight nwgthh ;
  subpopn finr=1 ;
  var hllatota ;
  setenv decwidth=1 colwidth=18 ;
  print total settotal lowtotal uptotal defftotal ;
run ;
title "Example 5.5: Estimating the Mean Value of Household Income using the 2012 HRS Data." ;
proc descript data=c5_hrs filetype=sas deft1 ;
  nest stratum secu ;
  weight nwgthh ;
  subpopn finr=1 ;
  var hllitot ;
  setenv decwidth=1 colwidth=18 ;
  print mean semean lowmean upmean deffmean ;
run ;
libname d "P:\ASDA 2\Data sets\nhanes 2011_2012\" ;
data c5_nhanes ;
```

```

    set d.nhanes1112_sub_8aug2016 ;
    int wtme2yr = int(wtme2yr) ;
    female=0 ;
    if riagendr=2 then female = 1 ;
    if age > 45 then age45=1 ; else age45=0 ;
run ;
proc sort ;
  by sdmvstra sdmvpsu ;
run ;
title "Example 5.6: Estimating Mean Systolic Blood Pressure using the NHANES Data." ;
proc descript data=c5_nhanes filetype=sas deft1 ;
  nest sdmvstra sdmvpsu ;
  weight wtme2yr ;
  subpopn age18p=1 ;
  var bpxs1 ;
  setenv decwidth=3 colwidth=18 ;
  print mean semean lowmean upmean deffmean ;
run ;
title "Example 5.7: Estimating the Mean Value of Total Household Wealth using the HRS Data." ;
proc descript data=c5_hrs filetype=sas deft1 ;
  nest stratum secu ;
  weight nwgthh ;
  subpopn finr=1 ;
  var hllatota ;
  setenv decwidth=1 colwidth=18 ;
  print mean semean lowmean upmean deffmean ;
run ;
data _null_ ;
  file print ;
  put "Sudaan software does not easily produce a SD therefore, Example 5.8: Estimation of the Population Standard
  Deviations of NHANES 2011-2012
  Measures of High-density and Total Cholesterol Level, is not shown." ;
run ;
title "Example 5.9: Estimating Population Quantiles for Total Household Wealth Using the HRS Data." ;
proc descript filetype=sas data=c5_hrs design=jackknife ;
  nest stratum secu ;
  weight nwgthh ;
  subpopn nfinr=1 ;
  var hllatota ;
  percentiles 25 75 / median ;
  setenv decwidth=3 colwidth=16 ;
run ;
title "Example 5.10: Estimating the Lorenz Curve and Gini Coefficient for the 2012 HRS Population Distribution
of Total Household Wealth. Not available in Sudaan" ;
* Use SAS for data preparation and pre-analysis ;
title "Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the
2011-2012 NHANES." ;
proc means data=c5_nhanes mean ;
  where age18p=1 ;
  var lbdhdd lbxtc ;
  weight wtme2yr ;
run ;
proc corr data=c5_nhanes ;
  where age18p=1 ;
  var lbdhdd lbxtc ;
  weight wtme2yr ;
run ;
data c5_nhanes_1 ;
  set c5_nhanes ;
  stdlbxtc=(lbxtc - 194.4355)/41.05184 ;
  stdlbdhdd = (lbdhdd - 52.83826) / 14.93157 ;
run ;
proc regress data=c5_nhanes_1 filetype=sas ;
  nest sdmvstra sdmvpsu ;
  weight wtme2yr ;
  subpopn age18p=1 ;
  model stdlbdhdd = stdlbxtc ;
  setenv decwidth=3 ;
run ;
title "Example 5.12: Estimating the Population Ratio of High Density to Total Cholesterol for U.S. Adults. " ;
proc ratio data=c5_nhanes filetype=sas deft1 ;
  nest sdmvstra sdmvpsu ; weight wtme2yr ;
  subpopn age18p=1 ;
  numer lbdhdd ;
  denom lbxtc ;
  setenv decwidth=4 ;

```

```

print nsum rhat serhat lowrhat uprhat deffrhat ;
run ;
title "Example 5.13: Estimating the Proportions of Males and Females Age >= 70 with Diabetes Using the HRS
Data." ;
proc descript filetype=sas deft1 data=c5_hrs ;
nest stratum secu ;
weight nwgtr ;
subpopn age70=1 ;
class gender / nofreq ;
var diabetes ;
setenv decwidth=4 ;
print mean semean lowmean upmean deffmean ;
run ;
title "Example 5.14: Estimating Mean Systolic Blood Pressure for Males and Females Age > 45 using the 2011-2012
NHANES data." ;
proc descript filetype=sas deft1 data=c5_nhanes ;
nest sdmvstra sdmvpsu ;
weight wtmecc2yr ;
subpopn age45=1 ;
class riagendr / nofreq ;
var bpxs1 ;
setenv decwidth=4 ;
print mean semean lowmean upmean deffmean ;
run ;
title "Example 5.15: Estimating Differences in Mean Total Household Wealth Between HRS Subpopulations Defined by
Educational Attainment Level." ;
proc descript filetype=sas data=c5_hrs ;
nest stratum secu ;
weight nwgthh ;
subpopn nfinr=1 ;
class edcat / nofreq ;
*rformat edcat edcatf. ;
var h1latota ;
setenv decwidth=2 colwidth=19 ;
run ;
proc descript filetype=sas data=c5_hrs ;
nest stratum secu ;
weight nwgthh ;
subpopn nfinr=1 ;
class edcat / nofreq ;
*rformat edcat edcatf. ;
var h1latota ;
contrast edcat = (1 0 0 -1) ;
setenv decwidth=3 colwidth=15 ;
run ;
title "Example 5.16: Estimating Differences in Mean Total Household Wealth from 2010 to 2012 using Data from the
HRS study. " ;
libname hrs10_12 "P:\ASDA 2\Data sets\HRS 2012\hrs 2010" ;
data hrs_2010_2012_c5 ;
set hrs10_12.hrs_2010_2012_both ;
* prepare data for analysis ;
hhweight = mwgthh ; if year=2012 then hhweight = nwgthh ;
totwealth=h10atota ; if year=2012 then totwealth=h1latota ;
finr2010 = 0 ; if (year = 2010 & mfinr = 1) then finr2010=1 ;
finr2012 = 0 ; if (year = 2012 & nfinr = 1) then finr2012=1 ;
finr2010_2012 = 0 ; if finr2010 = 1 | finr2012 = 1 then finr2010_2012=1 ;
run ;
proc sort ;
by stratum secu ;
run ;
proc descript filetype=sas data=hrs_2010_2012_c5 ;
nest stratum secu ;
weight hhweight ;
subpopn finr2010_2012=1 ;
class year ;
*rformat edcat edcatf. ;
var totwealth ;
contrast year = (1 -1) ;
setenv decwidth=3 colwidth=15 ;
run ;
ods listing close ;

```

Output SUDAAN Analysis Example Replication C5

Examples 5.1 and 5.2 are not available in Sudaan software since it does not offer any graphing abilities

Example 5.3 Population Totals of MDE by Marital Status, 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: NCSRWTSH_POP
 Stratification Variables(s): SESTRAT
 Primary Sampling Unit: SECLUSTR

Number of observations read : 9282 Weighted count :209128097
 Denominator degrees of freedom : 42

Date: 05-15-2017 SUDAAN
 Page: 1
 Time: 12:10:46
 Table: 1

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

Variable		Marital Status 1=Married 2=Previously Married 3=Never Married			
		Total	1	2	3
Major	Total	40092206.52	20304190.50	10360670.65	9427345.37
Depressive	SE Total	2567487.98	1584108.64	702621.51	773137.58
Episode 1=Yes	Lower 95% Limit				
0=No	Total	34910806.08	17107329.88	8942723.06	7867090.58
	Upper 95% Limit				
	Total	45273606.96	23501051.12	11778618.23	10987600.16
	DEFF Total #1	9.03	6.07	2.22	2.95

Sudaan software does not easily produce a SD therefore, Example 5.8: Estimation of the Population Standard Deviations of NHANES 2011-2012 Measures of High-density and Total Cholesterol Level, is not shown.


```

50.00          8432.365
75.00          16963.764
-----

```

Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the 2011-2012 NHANES

The CORR Procedure

```

2          Variables:  LBDHDD  LBXTC
Weight Variable:      WTMEC2YR

```

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
LBDHDD (mg/dL)	5187	52.83826	3066	1.15527E10	14.00000	175.00000	Direct HDL-Cholesterol
LBXTC (mg/dL)	5187	194.43547	8428	4.25118E10	59.00000	523.00000	Total Cholesterol(

Pearson Correlation Coefficients, N = 5187
 Prob > |r| under H0: Rho=0

	LBDHDD	LBXTC
LBDHDD Direct HDL-Cholesterol (mg/dL)	1.00000	0.24144 <.0001
LBXTC Total Cholesterol(mg/dL)	0.24144 <.0001	1.00000

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
 Observations used in the analysis : 5187 Weighted count:218642036
 Denominator degrees of freedom : 17

Maximum number of estimable parameters for the model is 2

File C5_NHANES_1 contains 31 Clusters
 31 clusters were used to fit the model
 Maximum cluster size is 231 records
 Minimum cluster size is 72 records

Weighted mean response is 0.000000
 Multiple R-Square for the dependent variable STDLBHDD: 0.058295

Date: 05-15-2017 SUDAAN
 Page: 1
 Time: 12:10:48
 Table: 1

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

Independent Variables and Effects	Beta Coeff.	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	0.000	0.034	-0.072	0.072	0.000	1.000
STDLBXTC	0.241	0.014	0.211	0.271	16.988	0.000

Date: 05-15-2017 SUDAAN
 Page: 2
 Time: 12:10:48
 Table: 1

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

Contrast	Degrees of Freedom	Wald F	P-value Wald F
OVERALL MODEL	2.000	163.856	0.000
MODEL MINUS INTERCEPT	1.000	288.582	0.000
INTERCEPT	1.000	0.000	1.000
STDLBXTC	1.000	288.582	0.000

Example 5.15: Estimating Differences in Mean Total Household Wealth Between HRS Subpopulations Defined by Educational Attainment Level.

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: NWGTHH
 Stratification Variables(s): STRATUM
 Primary Sampling Unit: SECU

Number of observations read : 19990 Weighted count : 89174512
 Number of observations skipped : 564
 (WEIGHT variable nonpositive)
 Observations in subpopulation : 13657 Weighted count : 58969863
 Denominator degrees of freedom : 56

Date: 05-15-2017 SUDAAN
 Page: 1
 Time: 12:10:48
 Table: 1

Variance Estimation Method: Taylor Series (WR)
 For Subpopulation: NFINR = 1
 by: Variable, Education 1=0-11 Yrs 2=12 Yrs 3=13-15 Yrs 4=16+ Yrs.

Variable	Education 1=0-11 Yrs 2=12 Yrs 3=13-15 Yrs 4=16+ Yrs			
	Total	1	2	
H11ATOTA:W11	Sample Size	13589.00	2870.00	4222.00
Total of all	Weighted Size	58685309.00	9008461.00	17514052.00
Assets--Cross-wave	Total	25203275952445.47	1099830753953.12	4536615163179.62
	Lower 95% Limit			
	Total	22484820495771.75	943670841569.93	4162111284145.26
	Upper 95% Limit			
	Total	27921731409119.18	1255990666336.30	4911119042213.99
	Mean	429464.82	122088.64	259027.16
	SE Mean	17452.97	10595.60	9802.47
	Lower 95% Limit			
	Mean	394502.33	100863.10	239390.45
	Upper 95% Limit			
	Mean	464427.32	143314.18	278663.87

Date: 05-15-2017
Page: 2
Time: 12:10:48
Table: 1

SUDAAN

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: NFINR = 1
by: Variable, Education 1=0-11 Yrs 2=12 Yrs 3=13-15 Yrs 4=16+ Yrs.

Variable		Education 1=0-11 Yrs 2=12 Yrs 3=13-15 Yrs 4=16+ Yrs	
		3	4
H11ATOTA:W11	Sample Size	3265.00	3232.00
Total of all	Weighted Size	14586188.00	17576608.00
Assets--Cross-	Total	4905460778023.10	14661369257289.63
wave	Lower 95% Limit		
	Total	4320066093107.27	12310841677607.39
	Upper 95% Limit		
	Total	5490855462938.93	17011896836971.87
	Mean	336308.62	834140.99
	SE Mean	17201.79	46477.79
	Lower 95% Limit		
	Mean	301849.30	741034.79
	Upper 95% Limit		
	Mean	370767.94	927247.19

Example 5.16: Estimating Differences in Mean Total Household Wealth from 2010 to 2012 using Data from the HRS study.

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: HHWEIGHT
Stratification Variables(s): STRATUM
Primary Sampling Unit: SECU

Number of observations read : 37291 Weighted count :164299555
Number of observations skipped : 659
(WEIGHT variable nonpositive)
Observations in subpopulation : 25182 Weighted count :107277623
Denominator degrees of freedom : 56

Date: 05-15-2017

SUDAAN

Page: 1

Time: 12:10:49

Table: 1

Frequencies and Values for CLASS Variables
by: YEAR.

```
-----  
YEAR            Frequency    Value  
-----  
Ordered  
  Position:  
  1            12676      2010  
Ordered  
  Position:  
  2            12506      2012  
-----
```

Variance Estimation Method: Taylor Series (WR)
 For Subpopulation: FINR2010_2012 = 1
 by: Variable, SUDAAN Reserved Variable One, Contrast.

for: Variable = TOTWEALTH.

		Contrast
SUDAAN Reserved Variable One		CONTRAST_1
Total	Sample Size	25182.000
	Weighted Size	107277623.000
	Cntrst Total	*****
	Lower 95% Limit	
	Cntrst Total	*****
	Upper 95% Limit	
	Cntrst Total	*****
	Cntrst Mean	-4978.066
	SE Cntrst Mean	7936.797
	Lower 95% Limit	
	Cntrst Mean	-20877.382
	Upper 95% Limit	
	Cntrst Mean	10921.249
	T-Test	
Cont.Mean=0	-0.627	
P-value T-Test		
Cont. Mean=0	0.533	
1	Sample Size	25182.000
	Weighted Size	107277623.000
	Cntrst Total	*****
	Lower 95% Limit	
	Cntrst Total	*****
	Upper 95% Limit	
	Cntrst Total	*****
	Cntrst Mean	-4978.066
	SE Cntrst Mean	7936.797
	Lower 95% Limit	
	Cntrst Mean	-20877.382
	Upper 95% Limit	
	Cntrst Mean	10921.249
	T-Test	
Cont.Mean=0	-0.627	
P-value T-Test		
Cont. Mean=0	0.533	