

CHAPTER 6 ANALYSIS EXAMPLES REPLICATION WesVar 4.3

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.

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

NOTES ABOUT CATEGORICAL DATA ANALYSIS USING WesVar 4.3

WesVar offers repeated replication variance estimation methods such as Jackknife and Balanced Repeated Replication.

WesVar is a point and click tool with log and output files that echo the options and variables selected for the particular analysis. As a result the output presented for WesVar examples consists of the log file and selected output. The exact syntax is not presented since it is not generated by the program nor is it possible to run WesVar with just user-written syntax but "Workbook" files can be created for a record of the analysis session. The workbook files will be posted on the ASDA web site in the near future and would enhance this output. From the output provided, you can determine the data used, output options, variables analyzed and other details of the analysis.

WesVar software can perform nearly all of the examples in Chapter 6 through use of the Tables and Regression menus.

Some of the fine points of this tool are the use of the subpopulation filter in the regression request statement, creation of variables used in the analyses (means, ratios, differences, etc.), various output options to specify the statistics of interest and a number of Repeated Replication variance estimation methods (JK1, JK2, BRR, etc.). For these examples, the JK2 method was used throughout but other methods are available. See the WesVar User's Guide for details.

ANALYSIS EXAMPLE 6.1 NHANES DATA

WESVAR VERSION NUMBER : 4.3
 TIME THE JOB EXECUTED : 12:36:14 03/26/2010
 INPUT DATASET NAME : C:\Program Files\Westat\WesVar\Data\nhanes0506_july11_2009_JK2.var
 TIME THE INPUT DATASET CREATED : 12:27:25 03/26/2010
 FULL SAMPLE WEIGHT : WTMEC2YR
 REPLICATE WEIGHTS : RPL01...RPL15
 VARIANCE ESTIMATION METHOD : JK2

OPTION COMPLETE : ON
 OPTION FUNCTION LOG : OFF
 OPTION VARIABLE LABEL : OFF
 OPTION VALUE LABEL : OFF
 OPTION OUTPUT REPLICATE ESTIMATES : OFF
 FINITE POPULATION CORRECTION FACTOR : 1.00000
 VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
 DEGREES OF FREEDOM : 15
 t VALUE : 2.131
 SUBSET CRITERIA : age >= 18

ANALYSIS VARIABLES : irregular
 COMPUTED STATISTIC : None Specified.
 TABLE(S) : irregular

FACTOR(S) : 1.00

NUMBER OF REPLICATES : 15
 NUMBER OF OBSERVATIONS READ : 5563
 WEIGHTED NUMBER OF OBSERVATIONS READ : 217700471.284
 EXCLUDED: 229 observations excluded from CELL_n because of zero weights.

irregular	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%	DEFF
0	SUM_WTS	PERCENT	97.05	0.674	95.61	98.48	8.114
1	SUM_WTS	PERCENT	2.95	0.674	1.52	4.39	8.114
MARGINAL	SUM_WTS	PERCENT	100.00	.	.	.	

WESVAR VERSION NUMBER : 4.3
 TIME THE JOB EXECUTED : 12:41:26 03/26/2010
 INPUT DATASET NAME : C:\Program Files\Westat\WesVar\Data\nhanes0506_july11_2009_JK2.var
 TIME THE INPUT DATASET CREATED : 12:27:25 03/26/2010
 FULL SAMPLE WEIGHT : WTMEC2YR
 REPLICATE WEIGHTS : RPL01...RPL15
 VARIANCE ESTIMATION METHOD : JK2

OPTION COMPLETE : ON
 OPTION FUNCTION LOG : OFF
 OPTION VARIABLE LABEL : OFF
 OPTION VALUE LABEL : OFF
 OPTION OUTPUT REPLICATE ESTIMATES : OFF
 FINITE POPULATION CORRECTION FACTOR : 1.00000
 VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
 DEGREES OF FREEDOM : 15
 t VALUE : 2.131
 SUBSET CRITERIA : age >= 18

ANALYSIS VARIABLES : irregular
 COMPUTED STATISTIC : M_irregular = MEAN(irregular)
 TABLE(S) : irregular

FACTOR(S) : 1.00

NUMBER OF REPLICATES : 15
 NUMBER OF OBSERVATIONS READ : 5563
 WEIGHTED NUMBER OF OBSERVATIONS READ : 217700471.284
 EXCLUDED: 229 observations excluded from CELL_n because of zero weights.

STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%	DEFF
SUM_WTS	VALUE	2.10e+08	1.281e+07	1.83e+08	2.38e+08	N/A
irregular	VALUE	6210830.05	1.336e+06	3362457.07	9059203.04	N/A
M_irregular	VALUE	0.03	0.007	0.02	0.04	8.114

NOTE: CODES FOR IRREGULAR: 0=NO 1=YES

ANALYSIS EXAMPLE 6.2 NHANES DATA

WESVAR VERSION NUMBER : 4.3
 TIME THE JOB EXECUTED : 15:01:10 04/05/2010
 INPUT DATASET NAME : C:\Program Files\Westat\WesVar\Data\nhanes0506_july11_2009_JK2.var
 TIME THE INPUT DATASET CREATED : 14:17:14 04/05/2010
 FULL SAMPLE WEIGHT : WTMEC2YR
 REPLICATE WEIGHTS : RPL01...RPL15
 VARIANCE ESTIMATION METHOD : JK2

OPTION COMPLETE : ON
 OPTION FUNCTION LOG : ON
 OPTION VARIABLE LABEL : ON
 OPTION VALUE LABEL : ON
 OPTION OUTPUT REPLICATE ESTIMATES : OFF
 FINITE POPULATION CORRECTION FACTOR : 1.00000
 VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
 DEGREES OF FREEDOM : 15
 t VALUE : 2.131
 SUBSET CRITERIA : age >= 18

ANALYSIS VARIABLES : RIDRETH1
 COMPUTED STATISTIC : None Specified.
 TABLE(S) : RIDRETH1

FACTOR(S) : 1.00

NUMBER OF REPLICATES : 15
 NUMBER OF OBSERVATIONS READ : 5563
 WEIGHTED NUMBER OF OBSERVATIONS READ : 217700471.284
 EXCLUDED: 229 observations excluded from CELL_n because of zero weights.

	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%	DEFF
Mexican	SUM_WTS	PERCENT	8.08	1.006	5.93	10.22	7.265
Other Hispanic	SUM_WTS	PERCENT	3.38	0.745	1.79	4.97	9.059
White	SUM_WTS	PERCENT	71.41	2.772	65.51	77.32	20.083
Black	SUM_WTS	PERCENT	11.73	1.986	7.49	15.96	20.320
Other	SUM_WTS	PERCENT	5.40	0.588	4.15	6.66	3.612
MARGINAL	SUM_WTS	PERCENT	100.00

ANALYSIS EXAMPLE 6.3 NHANES DATA

WESVAR VERSION NUMBER : 4.3
 TIME THE JOB EXECUTED : 15:04:04 04/05/2010
 INPUT DATASET NAME : C:\Program Files\Westat\WesVar\Data\nhanes0506_july11_2009_JK2.var
 TIME THE INPUT DATASET CREATED : 14:17:14 04/05/2010
 FULL SAMPLE WEIGHT : WTMEC2YR
 REPLICATE WEIGHTS : RPL01...RPL15
 VARIANCE ESTIMATION METHOD : JK2

OPTION COMPLETE : ON
 OPTION FUNCTION LOG : OFF
 OPTION VARIABLE LABEL : ON
 OPTION VALUE LABEL : ON
 OPTION OUTPUT REPLICATE ESTIMATES : OFF
 FINITE POPULATION CORRECTION FACTOR : 1.00000
 VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
 DEGREES OF FREEDOM : 15
 t VALUE : 2.131
 SUBSET CRITERIA : AGE >= 18

ANALYSIS VARIABLES : None Specified.
 COMPUTED STATISTIC : None Specified.
 TABLE(S) : bp_cat

FACTOR(S) : 1.00

NUMBER OF REPLICATES : 15
 NUMBER OF OBSERVATIONS READ : 5563
 WEIGHTED NUMBER OF OBSERVATIONS READ : 217700471.284
 EXCLUDED: 229 observations excluded from CELL_n because of zero weights.

bp_cat 1=Normal 2=PreHBP 3=Stage1 4=Stage 2

	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%	DEFF
Normal	SUM_WTS	PERCENT	47.11	1.123	44.71	49.50	2.560
Pre-Hypertensive	SUM_WTS	PERCENT	41.85	1.182	39.33	44.37	2.905
Stage 1 HBP	SUM_WTS	PERCENT	8.64	0.621	7.32	9.96	2.467
Stage 2 HBP	SUM_WTS	PERCENT	2.40	0.241	1.88	2.91	1.260
MARGINAL	SUM_WTS	PERCENT	100.00

ANALYSIS EXAMPLE 6.4: GOF WITH DEFINED POPULATION PROPORTIONS: NOT AVAILABLE IN WESVAR

ANALYSIS EXAMPLE 6.5: GRAPHING NOT AVAILABLE IN WESVAR

ANALYSIS EXAMPLE 6.6 NCS-R DATA

WESVAR VERSION NUMBER : 4.3
TIME THE JOB EXECUTED : 15:15:30 04/05/2010
INPUT DATASET NAME : C:\Program Files\Westat\WesVar\Data\final_ncsr_P1_JK2.var
TIME THE INPUT DATASET CREATED : 15:15:03 04/05/2010
FULL SAMPLE WEIGHT : NCSRWTSH
REPLICATE WEIGHTS : RPL01...RPL42
VARIANCE ESTIMATION METHOD : JK2

OPTION COMPLETE : ON
OPTION FUNCTION LOG : OFF
OPTION VARIABLE LABEL : ON
OPTION VALUE LABEL : ON
OPTION OUTPUT REPLICATE ESTIMATES : OFF
FINITE POPULATION CORRECTION FACTOR : 1.00000
VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
DEGREES OF FREEDOM : 42
t VALUE : 2.018

ANALYSIS VARIABLES : None Specified.
COMPUTED STATISTIC : None Specified.
TABLE(S) : SEX*mde

FACTOR(S) : 1.00

NUMBER OF REPLICATES : 42
NUMBER OF OBSERVATIONS READ : 9282
WEIGHTED NUMBER OF OBSERVATIONS READ : 9282.000

Sex	mde	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%	DEFF
Male	No	SUM_WTS	PERCENT	40.66	0.694	39.26	42.06	1.852
Male	Yes	SUM_WTS	PERCENT	7.22	0.339	6.54	7.91	1.596
Male		MARGINALSUM_WTS	PERCENT	47.89	0.531	46.81	48.96	1.050
Female	No	SUM_WTS	PERCENT	40.16	0.536	39.08	41.25	1.109
Female	Yes	SUM_WTS	PERCENT	11.95	0.303	11.34	12.56	0.808
Female		MARGINALSUM_WTS	PERCENT	52.11	0.531	51.04	53.19	1.050
MARGINAL	No	SUM_WTS	PERCENT	80.83	0.485	79.85	81.81	1.408
MARGINAL	Yes	SUM_WTS	PERCENT	19.17	0.485	18.19	20.15	1.408
MARGINAL		MARGINALSUM_WTS	PERCENT	100.00
Male	No	SUM_WTS	ROWPCT	84.92	0.766	83.38	86.47	1.896
Male	Yes	SUM_WTS	ROWPCT	15.08	0.766	13.53	16.62	1.896
Male		MARGINALSUM_WTS	ROWPCT	100.00
Female	No	SUM_WTS	ROWPCT	77.07	0.564	75.93	78.21	0.927
Female	Yes	SUM_WTS	ROWPCT	22.93	0.564	21.79	24.07	0.927
Female		MARGINALSUM_WTS	ROWPCT	100.00
MARGINAL	No	SUM_WTS	ROWPCT	80.83	0.485	79.85	81.81	1.408
MARGINAL	Yes	SUM_WTS	ROWPCT	19.17	0.485	18.19	20.15	1.408
MARGINAL		MARGINALSUM_WTS	ROWPCT	100.00

ANALYSIS EXAMPLE 6.7 NCS-R DATA

WESVAR VERSION NUMBER : 4.3
 TIME THE JOB EXECUTED : 15:19:49 04/05/2010
 INPUT DATASET NAME : C:\Program Files\Westat\WesVar\Data\final_ncsr_P1_JK2.var
 TIME THE INPUT DATASET CREATED : 15:15:03 04/05/2010
 FULL SAMPLE WEIGHT : NCSRWTSH
 REPLICATE WEIGHTS : RPL01...RPL42
 VARIANCE ESTIMATION METHOD : JK2

OPTION COMPLETE : ON
 OPTION FUNCTION LOG : OFF
 OPTION VARIABLE LABEL : ON
 OPTION VALUE LABEL : ON
 OPTION OUTPUT REPLICATE ESTIMATES : OFF
 FINITE POPULATION CORRECTION FACTOR : 1.00000
 VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
 DEGREES OF FREEDOM : 42
 t VALUE : 2.018

ANALYSIS VARIABLES : None Specified.
 COMPUTED STATISTIC : M_mde = MEAN(mde)
 TABLE(S) : SEX

FACTOR(S) : 1.00

NUMBER OF REPLICATES : 42
 NUMBER OF OBSERVATIONS READ : 9282
 WEIGHTED NUMBER OF OBSERVATIONS READ : 9282.000

Sex	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%	DEFF
Male	M_mde	VALUE	0.15	0.008	0.14	0.17	1.896
Female	M_mde	VALUE	0.23	0.006	0.22	0.24	0.927
MARGINALM	mde	VALUE	0.19	0.005	0.18	0.20	1.408

LABEL	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%
DIFF_M_F	M_mde	VALUE	-0.08	0.009	-0.10	-0.06

ANALYSIS EXAMPLE 6.8 NCS-R DATA

WESVAR VERSION NUMBER : 4.3
 TIME THE JOB EXECUTED : 15:09:22 04/05/2010
 INPUT DATASET NAME : C:\Program Files\Westat\WesVar\Data\final_ncsr_part2weight_JK2.var
 TIME THE INPUT DATASET CREATED : 14:14:12 04/05/2010
 FULL SAMPLE WEIGHT : NCSRWTLG
 REPLICATE WEIGHTS : RPL01...RPL42
 VARIANCE ESTIMATION METHOD : JK2

OPTION COMPLETE : ON
 OPTION FUNCTION LOG : OFF
 OPTION VARIABLE LABEL : ON
 OPTION VALUE LABEL : ON
 OPTION OUTPUT REPLICATE ESTIMATES : OFF
 FINITE POPULATION CORRECTION FACTOR : 1.00000
 VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
 DEGREES OF FREEDOM : 42
 t VALUE : 2.018
 SUBSET CRITERIA : AGE < 29

ANALYSIS VARIABLES : None Specified.
 COMPUTED STATISTIC : None Specified.
 TABLE(S) : ald*ED4CAT

FACTOR(S) : 1.00

NUMBER OF REPLICATES : 42
 NUMBER OF OBSERVATIONS READ : 1275
 WEIGHTED NUMBER OF OBSERVATIONS READ : 1266.557

ald	Educ	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	LOWER 95%	UPPER 95%	DEFF
No	0-11	SUM_WTS	COLPCT	90.87	2.917	84.98	96.76	2.134
No	12	SUM_WTS	COLPCT	95.14	1.339	92.44	97.85	1.481
No	13-15	SUM_WTS	COLPCT	95.10	1.005	93.08	97.13	0.993
No	16+	SUM_WTS	COLPCT	93.10	1.371	90.33	95.86	0.664
No	MARGINAL	SUM_WTS	COLPCT	94.05	0.884	92.27	95.84	1.782
Yes	0-11	SUM_WTS	COLPCT	9.13	2.917	3.24	15.02	2.134
Yes	12	SUM_WTS	COLPCT	4.86	1.339	2.15	7.56	1.481
Yes	13-15	SUM_WTS	COLPCT	4.90	1.005	2.87	6.92	0.993
Yes	16+	SUM_WTS	COLPCT	6.90	1.371	4.14	9.67	0.664
Yes	MARGINAL	SUM_WTS	COLPCT	5.95	0.884	4.16	7.73	1.782
MARGINAL	0-11	SUM_WTS	COLPCT	100.00
MARGINAL	12	SUM_WTS	COLPCT	100.00
MARGINAL	13-15	SUM_WTS	COLPCT	100.00
MARGINAL	16+	SUM_WTS	COLPCT	100.00
MARGINAL		SUM_WTS	COLPCT	100.00

ANALYSIS EXAMPLE 6.9 NCS-R DATA

OPTIONS : Intercept,
 No Standardized Coefficient,
 Degrees of Freedom = 42
 t VALUE : 2.018
 STARTING VALUES : INTERCEPT : 0.0000
 SEX : 0.0000
 BY : None Specified.
 MISSING : 0 (UNWEIGHTED)
 0.000000 (WEIGHTED)
 NONMISSING : 9282 (UNWEIGHTED)
 9282.000152 (WEIGHTED)
 Success = records with dependent value equal to 0 : 7453 (UNWEIGHTED)
 7502.536409 (WEIGHTED)
 Failure = records with dependent value equal to 1 : 1829 (UNWEIGHTED)
 1779.463743 (WEIGHTED)
 ITERATIONS REQUIRED FOR FULL SAMPLE : 5
 MAXIMUM ITERATIONS FOR REPLICATE SAMPLE : 5
 -2 LOG LIKELIHOOD FOR FULL SAMPLE : 8979.02383
 -2 LOG LIKELIHOOD FOR MODEL CONTAINING INTERCEPT ONLY : 9072.13015

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB> T	LOWER 95%	UPPER 95%
INTERCEPT	2.24	0.123	18.266	0.000	1.997	2.493
SEX	-0.52	0.067	-7.680	0.000	-0.652	-0.381

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F	NOTE
OVERALL FIT	58.978		1	42	0.000
SEX	58.978		1	42	0.000

PARAMETER	ESTIMATE	LOWER 95%	UPPER 95%
SEX	0.60	0.521	0.683

ANALYSIS EXAMPLE 6.10 NOT AVAILABLE IN WesVar