CHAPTER 9 ASDA ANALYSIS EXAMPLES REPLICATION-IVEware

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 MULTINOMIAL, ORDINAL, POISSON AND NEGATIVE BINOMIAL REGRESSION ANALYSES IN IVEware
Special note: IVEware MUST BE RUN IN THE REGULAR PROGRAM EDITOR IF RUNNING UNDER SAS!! THE ENHANCED EDITOR DOES NOT WORK WITH SAS BASED IVEware!!

IVEware %regress can perform some of the regression analyses presented in Chapter 9 of ASDA. Multinomial and Poisson regression are available in IVEware while ordinal, negative binomial and zero-inflated Poisson and negative binomial models are not available in IVEware via %regress. There is the possibility of using the %sasmod module of IVEware for complex sample corrected SE's for some selected procedures not covered by the %regress module but use of the %sasmod macro can be complex due to the repeated replication of complex models which often encounter convergence problems. See the IVEware documentation for details and examples.

This module uses the JRR method for variance estimation while the %describe command uses the Taylor Series Linearization method for variance estimation. This is reflected in the log file when running IVEware. Some of the fine points of this procedure are the use of a BY statement for subpopulation analyses, use of a CATEGORICAL statement for class variables in the regression, use of the highest category as the reference group for the outcome variable (which requires a reverse coding scheme for variables used as categorical in the CATEGORICAL statement), and various links depending on the type of regression desired. For logistic regression, use of the LOGISTIC link is required while for Poisson the LOG link is used.

The %regress module does not include an easy way to tests of parameters as a group (i.e. testing whether race is significantly different from zero) so the adjusted F or Wald tests for groups of predictors demonstrated in the ASDA text in Chapter 8 are omitted here.

```
*example 9.2 multinomial logistic regression;

%regress (name=9_2, setup=new, dir=. ) ;
title "Analysis Example 9.2: Multinomial Logistic Regression : NCSR" ;
datain  ncsr ;
stratum sestrat ;
cluster seclustr ;
weight  ncsrwtlg ;
categorical workstat3c_rev ;
dependent workstat3c_rev ;
predictor sexm ald mde ed12 ed1315 ed16 age44 age59 age60 prevmar nevmar ;
link logistic ;
run ;


NOTE: CODES FOR WORKSTAT3C_REV 1=NLF 2=UNEMPLOYED 3=EMPLOYED, THIS IS THE REVERSE OF THE ORIGINAL VARIABLE WKSTAT3C
WHERE 1=EMPLOYED 2=UNEMPLOYED 3=NLF.


IVEware Setup Checker, Wed Mar 10 14:43:54 2010                                    1

Setup listing:

title "Analysis Example 9.2: Multinomial Logistic Regression : NCSR" ;
datain  ncsr ;
stratum sestrat ;
cluster seclustr ;
weight  ncsrwtlg ;
categorical workstat3c_rev ;
dependent workstat3c_rev ;
predictor sexm ald mde ed12 ed1315 ed16 age44 age59 age60 prevmar nevmar ;
link logistic ;
run ;
```

"Analysis Example 9.2: Multinomial Logistic Regression : NCSR"

Regression type: Polytomous
Dependent variable: workstat3c_rev
Predictors: sexm
ald
mde
ed12
ed1315
ed16
age44
age59
age60
prevmar
nevmar
Cat. var. ref. codes: workstat3c_rev 3
Stratum variable: SESTRAT SAMPLING ERROR STRATUM
Cluster variable: SECLUSTR SAMPLING ERROR CLUSTER
Weight variable: NCSRWTLG NCSR sample part 2 weight

Valid cases 5679
Sum weights 5667.184998
Replicates 42

Degr freedom 42

-2 LogLike 7351.903365

Variable	Estimate	Std Error	Wald test	Prob > Chi
workstat3c_rev.1workstat3c_rev.2				
Intercept	-0.3794741	0.1721162	4.86095	0.02747
sexm	-0.6402555	0.1103050	33.69118	0.00000
ald	0.3332477	0.1305425	6.51674	0.01069
mde	0.0985220	0.0875571	1.26615	0.26049
ed12	-0.6514012	0.1403631	21.53732	0.00000
ed1315	-0.9169420	0.1471796	38.81399	0.00000
ed16	-1.2295007	0.1606919	58.54230	0.00000
age44	-0.3164450	0.1288982	6.02702	0.01409
age59	0.0649931	0.1732136	0.14079	0.70750
age60	2.3806067	0.1733245	188.64935	0.00000
prevmar	-0.0522634	0.1049048	0.24820	0.61834
nevmar	0.5527891	0.1325248	17.39901	0.00003
workstat3c_rev.2				
Intercept	-0.6438012	0.2965853	4.71199	0.02995
sexm	-1.3931969	0.1960786	50.48528	0.00000
ald	-0.1637813	0.3517969	0.21674	0.64153
mde	-0.1397560	0.1572275	0.79010	0.37407
ed12	-0.8470398	0.2364028	12.83814	0.00034
ed1315	-1.3653017	0.2593296	27.71745	0.00000
ed16	-1.7309570	0.3083845	31.50559	0.00000
age44	-0.8523907	0.2968191	8.24696	0.00408
age59	-0.8377006	0.2553588	10.76157	0.00104
age60	1.8283949	0.2860155	40.86587	0.00000
prevmar	-0.5899026	0.2237106	6.95325	0.00837
nevmar	-2.7845661	0.3801235	53.66185	0.00000

Variable	Odds Ratio	95% Confidence Interval Lower	Upper
workstat3c_rev.1			
Intercept			
sexm	0.5271577	0.4219537	0.6585918
ald	1.3954930	1.0722966	1.8161025
mde	1.1035387	0.9248024	1.3168193
ed12	0.5213148	0.3927174	0.6920220
ed1315	0.3997396	0.2970181	0.5379864
ed16	0.2924386	0.2114452	0.4044562
age44	0.7287351	0.5618211	0.9452383
age59	1.0671516	0.7523410	1.5136920
age60	10.8114598	7.6203645	15.3388545
prevmar	0.9490789	0.7679969	1.1728571
nevmar	1.7380940	1.3302190	2.2710325
workstat3c_rev.2			
Intercept			
sexm	0.2482803	0.1671441	0.3688022
ald	0.8489276	0.4173896	1.7266316
mde	0.8695704	0.6331465	1.1942775
ed12	0.4286821	0.2660372	0.6907617
ed1315	0.2553036	0.1512760	0.4308677
ed16	0.1771148	0.0950549	0.3300164
age44	0.4263943	0.2342433	0.7761678
age59	0.4327044	0.2584547	0.7244327
age60	6.2238889	3.4945100	11.0850428
prevmar	0.5543813	0.3529715	0.8707180
nevmar	0.0617559	0.0286763	0.1329946

Variable	Design Effect	SRS Estimate	% Diff SRS v Est
workstat3c_rev.1			
Intercept	2.07678	-0.3686795	-2.84461
sexm	2.49941	-0.4438028	-30.68349
ald	1.19542	0.2821847	-15.32285
mde	1.47541	0.0795390	-19.26784
ed12	2.03152	-0.7373261	13.19077
ed1315	2.15313	-1.0199367	11.23241
ed16	2.17003	-1.4038339	14.17919
age44	1.69636	-0.2191842	-30.73545
age59	2.77945	0.1692781	160.45562
age60	2.11489	2.2639728	-4.89934
prevmar	1.47002	0.0577880	-210.57067
nevmar	2.04157	0.4474283	-19.05986
workstat3c_rev.2			
Intercept	1.99630	-0.8881030	37.94679
sexm	1.37287	-1.5207141	9.15285

"Analysis Example 9.2: Multinomial Logistic Regression : NCSR"

Variable	Design Effect	SRS Estimate	% Diff SRS v Est
ald	1.17249	-0.2156997	31.69982
mde	1.15013	-0.1527761	9.31631
ed12	1.68134	-0.7941425	-6.24496
ed1315	1.83966	-1.0604353	-22.32959
ed16	2.23333	-1.2714246	-26.54788
age44	2.20823	-0.5414749	-36.47574
age59	1.36185	-0.6921360	-17.37668
age60	2.02538	1.8149489	-0.73540
prevmar	1.86898	-0.5842458	-0.95894
nevmar	1.24676	-2.0139938	-27.67298

```
* example 9.3 : note: IVEware does not provide option for ordinal logistic regression ;


*9.4 Poisson regression;  * note that only the age65p=1 results are shown here for brevity ;

%regress (name=9_4, setup=new, dir=. ) ;
title "Analysis Example 9.4: Poisson Regression : HRS" ;
datain  hrs  ;
stratum stratum  ;
cluster secu ;
weight  kwgtr  ;
by age65p  ;
offsets numfalls24 (offset2) ;
categorical revage3cat ;
dependent numfalls24 ;
predictor male revage3cat arthritis diabetes bodywgt totheight  ;
link log ;
run ;

CODES FOR AGE3CAT 1=65-74 2=75-84 3=85+ WHILE THESE CODES ARE REVERSED FOR REVAGE3CAT.


IVEware Setup Checker, Thu Mar 11 14:13:00 2010                                    1

Setup listing:

title "Analysis Example 9.4: Poisson Regression : NHANES" ;
datain  hrs  ;
stratum stratum  ;
cluster secu ;
weight  kwgtr  ;
by age65p  ;
offsets numfalls24 (offset2) ;
categorical revage3cat ;
dependent numfalls24 ;
predictor male revage3cat arthritis diabetes bodywgt totheight  ;
link log ;
run ;
```

"Analysis Example 9.4: Poisson Regression : NHANES"

```
Regression type:        Poisson
Dependent variable:     numfalls24
Offset variable:        offset2
Predictors:             male
                        revage3cat
                        arthritis
                        DIABETES
                        bodywgt
                        totheight
Cat. var. ref. codes:   revage3cat  3
By variables:           age65p
Stratum variable:       STRATUM  stratum id
Cluster variable:       SECU  sampling error computation unit
Weight variable:        KWGTR  2006 weight: respondent level

By variable                Code
  age65p                      1

Valid cases               10440
Sum weights            35017430
Replicates                   52

Degr freedom                 52

-2 LogLike          66001018.09
```

Variable	Estimate	Std Error	Wald test	Prob > Chi
Intercept	-0.1993226	0.6364245	0.09809	0.75414
male	0.1831258	0.1069761	2.93039	0.08693
revage3cat.1	0.5838654	0.0915594	40.66487	0.00000
revage3cat.2	0.2383983	0.0534778	19.87278	0.00001
arthritis	0.4867153	0.0823982	34.89110	0.00000
DIABETES	0.2596115	0.0686111	14.31724	0.00015
bodywgt	0.0009237	0.0008907	1.07545	0.29972
totheight	-0.0224337	0.0110437	4.12642	0.04222

Variable		95% Confidence Interval	
	Exp	Lower	Upper
Intercept			
male	1.2009655	0.9689548	1.4885297
revage3cat.1	1.7929555	1.4920307	2.1545732
revage3cat.2	1.2692146	1.1400673	1.4129919
arthritis	1.6269634	1.3790176	1.9194896
DIABETES	1.2964263	1.1296788	1.4877867
bodywgt	1.0009242	0.9991367	1.0027148
totheight	0.9778160	0.9563852	0.9997271

Variable	Design Effect	SRS Estimate	% Diff SRS v Est
Intercept	6.88641	-0.4025539	101.96095
male	11.85974	0.1899902	3.74847
revage3cat.1	8.60835	0.5228645	-10.44776
revage3cat.2	4.84797	0.2286585	-4.08551
arthritis	9.79654	0.4615868	-5.16288
DIABETES	7.53327	0.2493978	-3.93423
bodywgt	6.73659	0.0004927	-46.65810
totheight	7.63996	-0.0181658	-19.02439