CHAPTER 10 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 SURVIVAL ANALYSIS USING 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 survival analyses presented in Chapter 10 of ASDA. Cox PH regression and discrete time logistic regression are both available in IVEware while Kaplan Meier survival curves are not. Please note that the Cox model example using both %regress and %sasmod is not included as of Spring 2010 as the software code is under construction. When this command is operational this example will be updated and re-posted to the ASDA web site.

The %regress 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 Cox PH the PHREG 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 10 are omitted here.

*NOTE %regress and %sasmod for Cox PH models:
THIS PROCEDURE IS NOT CURRENTLY WORKING CORRECTLY IN IVEWARE, IS UNDER CONSTRUCTION AS OF SPRING 2010

DISCRETE TIME LOGISTIC MODEL: DATA CONSTRUCTION FIRST

```
data finalc10pyr ;
set ncsrpersonyr ;
if pyr <= ageonsetmde ;

%regress (name=10_2 , setup=new, dir=. ) ;
title "Analysis Example for Table 10.3: Discrete Time Logistic Regression : NCS-R" ;
datain  finalc10pyr ;
stratum sestrat;
cluster seclustr ;
weight  ncsrwtsh ;
link logistic  ;
dependent mdetv ;
predictor pyr intwage male ed12 ed1315 ed16 hispanic black white prevmar nevmar ;
run ;
```

```
IVEware Setup Checker, Thu Mar 11 15:39:14 2010                                  1

Setup listing:

title "Analysis Example for Table 10.3: Discrete Time Logistic Regression :
 NCS-R" ;
datain  finalc10pyr ;
stratum sestrat;
cluster seclustr ;
weight  ncsrwtsh ;
link logistic  ;
dependent mdetv ;
predictor pyr intwage male ed12 ed1315 ed16 hispanic black white prevmar nevmar ;
run ;
```

"Analysis Example for Table 10.3: Discrete Time Logistic Regression: NCS-R"

```
Regression type:        Logistic
Dependent variable:     mdetv
Predictors:             pyr
                        Intwage
                        male
                        ed12
                        ed1315
                        ed16
                        hispanic
                        black
                        white
                        prevmar
                        nevmar
Cat. var. ref. codes:   mdetv  2
Stratum variable:       SESTRAT  SAMPLING ERROR STRATUM
Cluster variable:       SECLUSTR  SAMPLING ERROR CLUSTER
Weight variable:        NCSRWTSH  NCSR sample part 1 weight

Valid cases             385696
Sum weights        386866.0469
Replicates                 42

Degr freedom               42

-2 LogLike         21657.91151
```

Variable	Estimate	Std Error	Wald test	Prob > Chi
Intercept	-3.4355246	0.1630052	444.20497	0.00000
pyr	0.0327978	0.0020831	247.90405	0.00000
intwage	-0.0583337	0.0024527	565.64520	0.00000
male	-0.4448692	0.0633800	49.26746	0.00000
ed12	-0.0201363	0.0670542	0.09018	0.76395
ed1315	0.0929188	0.0579647	2.56968	0.10893
ed16	-0.0194510	0.0634037	0.09411	0.75901
hispanic	-0.2484217	0.1349829	3.38705	0.06571
black	-0.4569678	0.1505834	9.20910	0.00241
white	0.0739955	0.1178912	0.39396	0.53023
prevmar	0.4942501	0.0617942	63.97312	0.00000
nevmar	-0.0353462	0.0880183	0.16126	0.68799

Variable	Odds Ratio	95% Confidence Interval Lower	Upper
Intercept			
pyr	1.0333416	1.0290067	1.0376946
intwage	0.9433351	0.9386773	0.9480160
male	0.6409081	0.5639583	0.7283574
ed12	0.9800651	0.8560240	1.1220803
ed1315	1.0973726	0.9762285	1.2335500
ed16	0.9807370	0.8629448	1.1146079
hispanic	0.7800309	0.5940286	1.0242743
black	0.6332007	0.4672659	0.8580621
white	1.0768020	0.8488115	1.3660306
prevmar	1.6392685	1.4470753	1.8569879
nevmar	0.9652712	0.8081771	1.1529014

Variable	Design Effect	SRS Estimate	% Diff SRS v Est
Intercept	1.34068	-3.4209702	-0.42364
pyr	1.18835	0.0329478	0.45733
intwage	1.25033	-0.0577725	-0.96205
male	1.59559	-0.4961003	11.51599
ed12	0.72923	0.0066541	-133.04516
ed1315	0.54601	0.0782755	-15.75920
ed16	0.60420	-0.0161028	-17.21346
hispanic	1.13643	-0.2190248	-11.83347
black	1.46987	-0.5001870	9.45782
white	1.31095	0.0150557	-79.65318
prevmar	1.09388	0.5206784	5.34715
nevmar	1.81968	0.0656526	-285.74163