CHAPTER 10 ASDA ANALYSIS EXAMPLES REPLICATION-MPLUS 5.21

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 IN MPLUS 5.21

The analysis replication examples were all run using Mplus 5.21. Mplus is an advanced modeling tool and offers the ability to correctly account for complex sample survey data for all analytic techniques.

Mplus can perform most of the modeling tasks presented in Chapter 10 of ASDA including Cox Proportional Hazards models and discrete-time logistic regression (with the appropriate person year data set) but not Kaplan Meier survival curves or the Clog-log logistic model for comparison to the logit model.

Some of the fine points of this tool are use of a unique cluster variable with a different value for each person in the data set, use of a SUBPOPULATION statement for subpopulation analyses, use of TYPE=COMPLEX and ESTIMATOR=MLR on the ANALYSIS command, and a MODELTEST statement for linear contrasts providing a Wald ChiSq test for selected parameter estimates. Due to the difficulty of running multi-parameter tests in Mplus (requires a separate run with the use of MODEL TEST: for each group of variables tested), this technique is not demonstrated for each group of parameters in the models presented. See Chapter 8 logistic examples for how to do this and also see the Mplus User's Guide for additional detail.

ANALYSIS EXAMPLE 10.1 COX PH MODEL NCS-R DATA

Mplus VERSION 5.21
MUTHEN & MUTHEN
04/07/2010   1:20 PM

INPUT INSTRUCTIONS

```
  TITLE: ANALYSIS EXAMPLE 10.1 NCSR DATA SURVIVAL ANALYSIS COX MODEL
  DATA:
    FILE IS "F:\applied_analysis_book\Mplus\ncsrsurv.txt";

  VARIABLE:
    NAMES ARE numsecu ageonsetmde mde age sexm mar3cat ed4cat ed011 ed12 ed1315
      ed16 racecat other
      hispanic black white caseid sestrat seclustr ncsrwtlg ncsrwtsh prevmar nevmar  ;

    USEVARIABLES ARE ageonsetmde mde age sexm prevmar nevmar ed12 ed1315
      ed16 hispanic black white ;

    missing are . ;

    WEIGHT IS NCSRWTsh  ;
    stratification is sestrat ;
    cluster is numsecu ;
    survival=ageonsetmde (all) ;
    timecensored=mde (O=right 1=not) ;
  ANALYSIS:
    type is complex;

  Model:
    ageonsetmde on age sexm prevmar nevmar ed12 ed1315 ed16 hispanic black white ;
  output:
      cint basehazard ;
```

ANALYSIS EXAMPLE 10.1 NCSR DATA SURVIVAL ANALYSIS COX MODEL

SUMMARY OF ANALYSIS

Number of groups   1
Number of observations   9282

Number of dependent variables   1
Number of independent variables   10
Number of continuous latent variables   0

Observed dependent variables

  Time-to-event (survival)
   AGEONSET

Observed independent variables
   AGE   SEXM   PREVMAR   NEVMAR   ED12   ED1315
   ED16   HISPANIC   BLACK   WHITE

Variables with special functions

  Stratification   SESTRAT
  Cluster variable   NUMSECU
  Weight variable   NCSRWTSH
  Time-censoring variables
   MDE

Estimator   MLR
Information matrix   OBSERVED
Optimization Specifications for the Quasi-Newton Algorithm for
Continuous Outcomes
  Maximum number of iterations   100
  Convergence criterion   0.100D-05
Optimization Specifications for the EM Algorithm
  Maximum number of iterations   500
  Convergence criteria
    Loglikelihood change   0.100D-02
    Relative loglikelihood change   0.100D-05
    Derivative   0.100D-02
Optimization Specifications for the M step of the EM Algorithm for
Categorical Latent variables
  Number of M step iterations   1
  M step convergence criterion   0.100D-02
  Basis for M step termination   ITERATION
Optimization Specifications for the M step of the EM Algorithm for
Censored, Binary or Ordered Categorical (Ordinal), Unordered
Categorical (Nominal) and Count Outcomes
  Number of M step iterations   1
  M step convergence criterion   0.100D-02
  Basis for M step termination   ITERATION
  Maximum value for logit thresholds   15
  Minimum value for logit thresholds   -15
  Minimum expected cell size for chi-square   0.100D-01
Maximum number of iterations for H1   2000
Convergence criterion for H1   0.100D-03
Optimization algorithm   EMA
Integration Specifications
  Type   STANDARD
  Number of integration points   15
  Dimensions of numerical integration   0
  Adaptive quadrature   ON
Base Hazard   OFF
Cholesky   OFF

```
Input data file(s)
  F:\applied_analysis_book\Mplus\ncsrsurv.txt
Input data format  FREE


SUMMARY OF DATA

     Number of missing data patterns             0
     Number of strata            42
     Number of clusters          84


COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value    0.100



THE MODEL ESTIMATION TERMINATED NORMALLY



TESTS OF MODEL FIT

Loglikelihood

        HO Value                     -10578.163
        HO Scaling Correction Factor     1.150
          for MLR

Information Criteria

        Number of Free Parameters         10
        Akaike (AIC)                21176.327
        Bayesian (BIC)              21247.685
        Sample-Size Adjusted BIC    21215.906
          (n* = (n + 2) / 24)



MODEL RESULTS

                                              Two-Tailed
                   Estimate      S.E.  Est./S.E.    P-Value

 AGEONSETMD ON
    AGE              -0.050     0.002   -20.802      0.000
    SEXM             -0.454     0.062    -7.280      0.000
    PREVMAR           0.502     0.060     8.371      0.000
    NEVMAR            0.081     0.089     0.908      0.364
    ED12             -0.057     0.067    -0.848      0.396
    ED1315            0.046     0.058     0.787      0.431
    ED16             -0.090     0.064    -1.417      0.157
    HISPANIC         -0.250     0.134    -1.860      0.063
    BLACK            -0.479     0.149    -3.216      0.001
    WHITE             0.078     0.117     0.662      0.508


QUALITY OF NUMERICAL RESULTS

     Condition Number for the Information Matrix          0.870E-04
        (ratio of smallest to largest eigenvalue)
```

CONFIDENCE INTERVALS OF MODEL RESULTS

	Lower .5%	Lower 2.5%	Estimate	Upper 2.5%	Upper .5%
AGEONSET ON					
AGE	-0.056	-0.054	-0.050	-0.045	-0.043
SEXM	-0.614	-0.576	-0.454	-0.331	-0.293
PREVMAR	0.348	0.385	0.502	0.620	0.657
NEVMAR	-0.148	-0.093	0.081	0.255	0.309
ED12	-0.230	-0.188	-0.057	0.075	0.116
ED1315	-0.104	-0.068	0.046	0.159	0.195
ED16	-0.254	-0.215	-0.090	0.035	0.074
HISPANIC	-0.596	-0.513	-0.250	0.013	0.096
BLACK	-0.863	-0.771	-0.479	-0.187	-0.095
WHITE	-0.225	-0.152	0.078	0.308	0.380

	Lower .5%	Lower 2.5%	Estimate	Upper 2.5%	Upper .5%

ANALYSIS EXAMPLE 10.2 DISCRETE TIME LOGISTIC REGRESSION NCS-R DATA

Mplus VERSION 5.21
MUTHEN & MUTHEN
04/07/2010   3:31 PM

INPUT INSTRUCTIONS

  TITLE: ANALYSIS EXAMPLE 10.2 NCSR DATA DISCRETE TIME LOGISTIC REGRESSION
  DATA:
    FILE IS "F:\applied_analysis_book\Mplus\ncsrdt.txt";

  VARIABLE:
    NAMES ARE numsecu mdetv pyr age sexm ed011 ed12 ed1315 ed16 other hispanic black white
          caseid sestrat seclustr ncsrwtlg ncsrwtsh prevmar nevmar ageonsetmde ;

    USEVARIABLES ARE numsecu mdetv pyr age sexm ed12 ed1315 ed16 hispanic black white
          prevmar nevmar ageonsetmde ;

    missing are . ;

    WEIGHT IS NCSRWTsh ;
    stratification is sestrat ;
    cluster is numsecu ;
    SUBPOPULATION =(PYR <= AGEONSETMDE) ;
    categorical = mdetv ;

  ANALYSIS:
    type is complex;
    estimator=mlr ;
  Model:
    mdetv on pyr age sexm ed12 ed1315 ed16 hispanic black white prevmar nevmar ;

ANALYSIS EXAMPLE 10.2 NCSR DATA DISCRETE TIME LOGISTIC REGRESSION

SUMMARY OF ANALYSIS

Number of groups   1
Number of observations   385696

Number of dependent variables   2
Number of independent variables   11
Number of continuous latent variables   0

Observed dependent variables

  Continuous
   AGEONSETMD

  Binary and ordered categorical (ordinal)
   MDETV

Observed independent variables
   PYR   AGE   SEXM   ED12   ED1315   ED16
   HISPANIC   BLACK   WHITE   PREVMAR   NEVMAR

Variables with special functions

  Stratification   SESTRAT
  Cluster variable   NUMSECU
  Weight variable   NCSRWTSH

Estimator   MLR
Information matrix   OBSERVED
Optimization Specifications for the Quasi-Newton Algorithm for
Continuous Outcomes
  Maximum number of iterations   100
  Convergence criterion   0.100D-05
Optimization Specifications for the EM Algorithm
  Maximum number of iterations   500
  Convergence criteria
    Loglikelihood change   0.100D-02
    Relative loglikelihood change   0.100D-05
    Derivative   0.100D-02
Optimization Specifications for the M step of the EM Algorithm for
Categorical Latent variables
  Number of M step iterations   1
  M step convergence criterion   0.100D-02
  Basis for M step termination   ITERATION
Optimization Specifications for the M step of the EM Algorithm for
Censored, Binary or Ordered Categorical (Ordinal), Unordered
Categorical (Nominal) and Count Outcomes
  Number of M step iterations   1
  M step convergence criterion   0.100D-02
  Basis for M step termination   ITERATION
  Maximum value for logit thresholds   15
  Minimum value for logit thresholds   -15
  Minimum expected cell size for chi-square   0.100D-01
Maximum number of iterations for H1   2000
Convergence criterion for H1   0.100D-03
Optimization algorithm   EMA
Integration Specifications
  Type   STANDARD
  Number of integration points   15
  Dimensions of numerical integration   0
  Adaptive quadrature   ON
Link   LOGIT
Cholesky   OFF

```
Input data file(s)
  F:\applied_analysis_book\Mplus\ncsrdt.txt
Input data format  FREE




SUMMARY OF CATEGORICAL DATA PROPORTIONS

     MDETV
       Category 1    0.995
       Category 2    0.005

THE MODEL ESTIMATION TERMINATED NORMALLY

TESTS OF MODEL FIT

Loglikelihood

          H0 Value                   -1690195.042
          H0 Scaling Correction Factor     18.100
            for MLR

Information Criteria

          Number of Free Parameters        14
          Akaike (AIC)               3380418.084
          Bayesian (BIC)             3380570.163
          Sample-Size Adjusted BIC   3380525.671
            (n* = (n + 2) / 24)



MODEL RESULTS

                                                  Two-Tailed
                     Estimate     S.E.  Est./S.E.   P-Value


 MDETV      ON
    PYR              0.033       0.002    15.816     0.000
    AGE             -0.058       0.002   -23.823     0.000
    SEXM            -0.445       0.062    -7.142     0.000
    ED12            -0.020       0.066    -0.305     0.761
    ED1315           0.093       0.057     1.618     0.106
    ED16            -0.019       0.063    -0.307     0.759
    HISPANIC        -0.248       0.135    -1.843     0.065
    BLACK           -0.457       0.150    -3.049     0.002
    WHITE            0.074       0.118     0.626     0.531
    PREVMAR          0.494       0.061     8.101     0.000
    NEVMAR          -0.035       0.088    -0.402     0.688

 Means
    AGEONSETMD      50.356       0.370   136.029     0.000


 Thresholds
    MDETV$1          3.436       0.162    21.209     0.000


 Variances
    AGEONSETMD     354.439       7.689    46.095     0.000
```

```
LOGISTIC REGRESSION ODDS RATIO RESULTS

 MDETV       ON
    PYR             1.033
    AGE             0.943
    SEXM            0.641
    ED12            0.980
    ED1315          1.097
    ED16            0.981
    HISPANIC        0.780
    BLACK           0.633
    WHITE           1.077
    PREVMAR         1.639
    NEVMAR          0.965


QUALITY OF NUMERICAL RESULTS

    Condition Number for the Information Matrix        0.412E-05
       (ratio of smallest to largest eigenvalue)


    Beginning Time:  15:31:31
      Ending Time:  15:34:48
     Elapsed Time:  00:03:17
```