Logistic (RLOGIST) Example #4

SUDAAN Statements and Results Illustrated

- SEs by replicate method
- REPWGT
- EFFECTS
- EXP option
- REFLEVEL

Input Data Set(s): NH3MI1.SAS7bdat - NH3MI5.SAS7bdat

Example

Using the NHANES III Multiply Imputed Datasets, model the probability of being classified as overweight by weighted logistic regression with standard errors obtained by the replicate method.

This example also demonstrates the use of the EXP option (customized odds ratios) in the context of an interaction model.

Solution

NHIS The covariates in this model include the following:

Variable	Definition
HSSEX	Sex (1=Male, 2=Female)
AGEGRP	Age group (1=20-39, 2=40-59, 3=60+)
DMARETHN	Race-Ethnicity (1=non-Hispanic white, 2-non-Hispanic black, 3=Mexican-American, 4=other)
HAB1MI	Self-rating of health status (1=excellent, 2=very good, 3=good, 4=fair, 5=poor)
HAT28MI	Compare own activity level to others (1=more active, 2=less active, 3=about the same)
POVERTY	Poverty status (1=at or below the poverty line, 2=above)

The response variable for this model is:

OVERWT	Overweight status, as determined by BMI (1=overweight, 0=not overweight)

This example uses PROC RLOGIST (SAS-Callable SUDAAN) to model the risk of being classified as overweight as a function of the above covariates, in the context of multiply imputed data.

Computationally, this example highlights the use of the EFFECTS statements in the context of an interaction model, and the estimation of default and user-defined odds ratios and their confidence limits.

The main effects plus interaction model is specified on the MODEL statement (see ***Exhibit 1***). Each of the variables to be modeled as categorical also appear on the CLASS statement. Frequencies for CLASS variables are provided in the printed output. The default Wald-***F*** test is used for all tests of hypotheses.

The REFLEVEL statement defines the reference level for race to be the first level (non-Hispanic White). For all other categorical variables not listed on the REFLEVEL statement, the last level of each of these variables will be used as the reference level.

Finally, the EFFECTS statement forms a contrast comparing activity level 1 vs. 3 (more vs. less active than others your age) when race=1 (non-Hispanic White), 2 (non-Hispanic Black) and 3 (Mexican-American). The EXP option will exponentiate the contrast to provide the user-requested odds ratio for being classified as overweight for those who are more vs. less active within *each* level of race (the default odds ratios compare more vs. less active for race=1 only).

We include two PRINT statements. The first requests specific statistics in the *betas* group, and the second requests the default statistics in the *risk*, *test*, and *expcntrst* groups. Two PRINT statements allow us to set up different default print environments (SETENV statements) for different groups. The PRINT statements are used in this example to request the PRINT groups of interest and to specify a variety of formats for those printed statistics. Without the PRINT statement, the default statistics are produced, with default formats.

The SETENV statements are optional. They set up default formats for printed statistics and manipulate the printout to the needs of the user.

The RFORMAT statements associate the SAS formats with the variables used in the RLOGIST procedure. The RLABEL statement defines variable labels for use in the current procedure only. Without the RLABEL statement, SAS variable labels would be used if already defined.

The SAS-callable SUDAAN code is presented first. It contains the data steps that read in the 5 multiply-imputed SAS datasets (*nh3mi1—nh3mi5*) originally derived from files provided with the NHANES III public use documentation and creates the necessary model variables. The single call to RLOGIST is at the end of the program. The options DATA=MI1 and MI_COUNT=5 specify that temporary SAS datasets MI1—MI5 are the multiply imputed datasets to be used in the SUDAAN analysis.

Exhibit 1. SAS-Callable SUDAAN Code

```
libname in v604 "c:\10winbetatest\examplemanual\logistic";
libname in "c:\903winbetatest\nhanes3";
options nocenter linesize=85 pagesize=68;

proc format;
  value age 1="1=20-39"
            2="2=40-59"
            3="3=60+";
  value overwt 1="1=Overweight"
               0="0=Not Overweight";
  value hab 1="1=Excellent"
            2="2=Very Good"
            3="3=Good"
            4="4=Fair"
            5="5=Poor";
  value active 1="1=More Active"
               2="2=Same"
               3="3=Less Active";
  value sex 1="1=Male"
            2="2=Female";
  value poverty 1="1=At/Below Poverty"
                2="2=Above Poverty";
  value race 1="1=nH White"
             2="2=nH Black"
             3="3=Mex-Amer"
             4="4=Other";

data mi1; set in.nh3mi1;
      BMI = BMPWTMI / (BMPHTMI/100)**2;
      OVERWT = 0;
      IF HSSEX = 1 AND BMI GE 27.8 THEN OVERWT = 1;
      IF HSSEX = 2 AND BMI GE 27.3 THEN OVERWT = 1;

      AGE = HSAGEIR;
      IF HSAGEU = 1 THEN AGE = AGE / 12;
      IF AGE GE 20 AND AGE LE 39 THEN AGEGRP = 1;
      ELSE IF AGE GE 40 AND AGE LE 59 THEN AGEGRP = 2;
      ELSE IF AGE GE 60 THEN AGEGRP = 3;

      IF DMPPIRMI LE 1.0 THEN POVERTY = 1;
      ELSE IF DMPPIRMI GT 1.0 THEN POVERTY = 2;

      if hat28mi=1 then active_mi=1;  /* more */
      else if hat28mi=3 then active_mi=2; /* same */
      else if hat28mi=2 then active_mi=3; /* less */

      keep sdpstra6 sdppsu6 wtpfqx6 WTPQRP1-WTPQRP52 age agegrp dmarethn poverty
           overwt hssex hab1mi hat28mi active_mi;
proc sort data=mi1; by SDPSTRA6 SDPPSU6;
```

Exhibit 1.   SAS-Callable SUDAAN Code – Cont.

```
data mi2; set in.nh3mi2;
      BMI = BMPWTMI / (BMPHTMI/100)**2;
      OVERWT = 0;
      IF HSSEX = 1 AND BMI GE 27.8 THEN OVERWT = 1;
      IF HSSEX = 2 AND BMI GE 27.3 THEN OVERWT = 1;

      AGE = HSAGEIR;
      IF HSAGEU = 1 THEN AGE = AGE / 12;
      IF AGE GE 20 AND AGE LE 39 THEN AGEGRP = 1;
      ELSE IF AGE GE 40 AND AGE LE 59 THEN AGEGRP = 2;
      ELSE IF AGE GE 60 THEN AGEGRP = 3;

      IF DMPPIRMI LE 1.0 THEN POVERTY = 1;
      ELSE IF DMPPIRMI GT 1.0 THEN POVERTY = 2;

      if hat28mi=1 then active_mi=1;  /* more */
      else if hat28mi=3 then active_mi=2; /* same */
      else if hat28mi=2 then active_mi=3; /* less */

      keep sdpstra6 sdppsu6 wtpfqx6 WTPQRP1-WTPQRP52 age agegrp dmarethn poverty
           overwt hssex hab1mi hat28mi active_mi;
proc sort data=mi2; by SDPSTRA6 SDPPSU6;

data mi3; set in.nh3mi3;
      BMI = BMPWTMI / (BMPHTMI/100)**2;
      OVERWT = 0;
      IF HSSEX = 1 AND BMI GE 27.8 THEN OVERWT = 1;
      IF HSSEX = 2 AND BMI GE 27.3 THEN OVERWT = 1;

      AGE = HSAGEIR;
      IF HSAGEU = 1 THEN AGE = AGE / 12;
      IF AGE GE 20 AND AGE LE 39 THEN AGEGRP = 1;
      ELSE IF AGE GE 40 AND AGE LE 59 THEN AGEGRP = 2;
      ELSE IF AGE GE 60 THEN AGEGRP = 3;

      IF DMPPIRMI LE 1.0 THEN POVERTY = 1;
      ELSE IF DMPPIRMI GT 1.0 THEN POVERTY = 2;

      if hat28mi=1 then active_mi=1;  /* more */
      else if hat28mi=3 then active_mi=2; /* same */
      else if hat28mi=2 then active_mi=3; /* less */

      keep sdpstra6 sdppsu6 wtpfqx6 WTPQRP1-WTPQRP52 age agegrp dmarethn poverty
           overwt hssex hab1mi hat28mi active_mi;
proc sort data=mi3; by SDPSTRA6 SDPPSU6;
```

Exhibit 1. SAS-Callable SUDAAN Code – cont.

```
data mi4; set in.nh3mi4;
      BMI = BMPWTMI / (BMPHTMI/100)**2;
      OVERWT = 0;
      IF HSSEX = 1 AND BMI GE 27.8 THEN OVERWT = 1;
      IF HSSEX = 2 AND BMI GE 27.3 THEN OVERWT = 1;

      AGE = HSAGEIR;
      IF HSAGEU = 1 THEN AGE = AGE / 12;
      IF AGE GE 20 AND AGE LE 39 THEN AGEGRP = 1;
      ELSE IF AGE GE 40 AND AGE LE 59 THEN AGEGRP = 2;
      ELSE IF AGE GE 60 THEN AGEGRP = 3;

      IF DMPPIRMI LE 1.0 THEN POVERTY = 1;
      ELSE IF DMPPIRMI GT 1.0 THEN POVERTY = 2;

      if hat28mi=1 then active_mi=1;  /* more */
      else if hat28mi=3 then active_mi=2; /* same */
      else if hat28mi=2 then active_mi=3; /* less */

      keep sdpstra6 sdppsu6 wtpfqx6 WTPQRP1-WTPQRP52 age agegrp dmarethn poverty
           overwt hssex hab1mi hat28mi active_mi;
proc sort data=mi4; by SDPSTRA6 SDPPSU6;

data mi5; set in.nh3mi5;
      BMI = BMPWTMI / (BMPHTMI/100)**2;
      OVERWT = 0;
      IF HSSEX = 1 AND BMI GE 27.8 THEN OVERWT = 1;
      IF HSSEX = 2 AND BMI GE 27.3 THEN OVERWT = 1;

      AGE = HSAGEIR;
      IF HSAGEU = 1 THEN AGE = AGE / 12;
      IF AGE GE 20 AND AGE LE 39 THEN AGEGRP = 1;
      ELSE IF AGE GE 40 AND AGE LE 59 THEN AGEGRP = 2;
      ELSE IF AGE GE 60 THEN AGEGRP = 3;

      IF DMPPIRMI LE 1.0 THEN POVERTY = 1;
      ELSE IF DMPPIRMI GT 1.0 THEN POVERTY = 2;

      if hat28mi=1 then active_mi=1;  /* more */
      else if hat28mi=3 then active_mi=2; /* same */
      else if hat28mi=2 then active_mi=3; /* less */

      keep sdpstra6 sdppsu6 wtpfqx6 WTPQRP1-WTPQRP52 age agegrp dmarethn poverty
           overwt hssex hab1mi hat28mi active_mi;
proc sort data=mi5; by SDPSTRA6 SDPPSU6;
```

Exhibit 1. SAS-Callable SUDAAN Code – cont.

```
PROC RLOGIST DATA=MI1 FILETYPE=SAS MI_COUNT=5 DESIGN=BRR;
REPWGT WTPQRP1-WTPQRP52 / adjfay=2.0408;
WEIGHT WTPFQX6;

SUBPOPX AGE ge 20 / name="Age 20+";
CLASS AGEGRP DMARETHN POVERTY ACTIVE_MI HSSEX HAB1MI;

REFLEVEL dmarethn=1;
MODEL overwt = agegrp dmarethn poverty active_mi hssex hab1mi DMARETHN*ACTIVE_MI;

EFFECTS ACTIVE_MI=(1 0 -1) / DMARETHN=1 EXP name="Active=More vs Less: nH White";
EFFECTS ACTIVE_MI=(1 0 -1) / DMARETHN=2 EXP name="Active=More vs Less: nH Black";
EFFECTS ACTIVE_MI=(1 0 -1) / DMARETHN=3 EXP name="Active=More vs Less: Mex Amer";

setenv labwidth=32 colspce=1;
print beta sebeta t_beta p_beta / betafmt=f7.4 sebetafmt=f8.4 t_betafmt=f6.2
                                  p_betafmt=f7.4;

setenv labwidth=32 colspce=4;
print / risk=default tests=default expcntrst=default
        waldpfmt=f7.4 dffmt=f7.0 loworfmt=f9.2 uporfmt=f9.2;

RLABEL HAB1MI="Self-Rating Health Status";
RLABEL POVERTY="Poverty Status";
RLABEL ACTIVE_MI="Activity level compared with others your age";
RLABEL OVERWT="Overweight-Derived from BMI";
RLABEL DMARETHN="Race/Ethnicity";

RFORMAT agegrp age.;
RFORMAT overwt overwt.;
RFORMAT hab1mi hab.;
RFORMAT active_mi active.;
RFORMAT dmarethn race.;
RFORMAT hssex sex.;
RFORMAT poverty poverty.;
RTITLE "Modelling Prob of Being Classified as Overweight with MI Data (BRR)";
```

Exhibit 2. First Page of SUDAAN Results

```
                                  S U D A A N
            Software for the Statistical Analysis of Correlated Data
           Copyright      Research Triangle Institute    December 2011
                                Release 11.0.0

DESIGN SUMMARY: Variances will be computed using the Balanced Repeated Replication
(BRR) Method
  Sample Weight: WTPFQX6
  Replicate Sample Weights:
        WTPQRP1  WTPQRP2  WTPQRP3  WTPQRP4  WTPQRP5  WTPQRP6  WTPQRP7
        WTPQRP8  WTPQRP9  WTPQRP10 WTPQRP11 WTPQRP12 WTPQRP13 WTPQRP14
        WTPQRP15 WTPQRP16 WTPQRP17 WTPQRP18 WTPQRP19 WTPQRP20 WTPQRP21
        WTPQRP22 WTPQRP23 WTPQRP24 WTPQRP25 WTPQRP26 WTPQRP27 WTPQRP28
        WTPQRP29 WTPQRP30 WTPQRP31 WTPQRP32 WTPQRP33 WTPQRP34 WTPQRP35
        WTPQRP36 WTPQRP37 WTPQRP38 WTPQRP39 WTPQRP40 WTPQRP41 WTPQRP42
        WTPQRP43 WTPQRP44 WTPQRP45 WTPQRP46 WTPQRP47 WTPQRP48 WTPQRP49
        WTPQRP50 WTPQRP51 WTPQRP52
    Multiplier Associated with Replicate Weights: 2.04

Processing data for set 1 of imputed variables:
Processing data for set 2 of imputed variables:
Processing data for set 3 of imputed variables:
Processing data for set 4 of imputed variables:
Processing data for set 5 of imputed variables:


Processing data for set 1 of imputed variables:

Number of zero responses     : 11454
Number of non-zero responses :  7371

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     52

Maximum number of estimable parameters for the model is 20

Sample and Population Counts for Response Variable OVERWT
Based on observations used in the analysis
0:  Sample Count    11454    Population Count 116652970
1:  Sample Count     7371    Population Count  60527700

R-Square for dependent variable OVERWT (Cox & Snell, 1989): 0.062499


Processing data for set 2 of imputed variables:

Number of zero responses     : 11480
Number of non-zero responses :  7345

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     52

Maximum number of estimable parameters for the model is 20
```

Exhibit 2. First Page of SUDAAN Results – Cont.

```
Sample and Population Counts for Response Variable OVERWT
Based on observations used in the analysis
0:  Sample Count    11480    Population Count 116654450
1:  Sample Count     7345    Population Count  60526220

R-Square for dependent variable OVERWT (Cox & Snell, 1989): 0.061448

Processing data for set 3 of imputed variables:

Number of zero responses     : 11435
Number of non-zero responses :  7390

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     52

Maximum number of estimable parameters for the model is 20

Sample and Population Counts for Response Variable OVERWT
Based on observations used in the analysis
0:  Sample Count    11435    Population Count 116618763
1:  Sample Count     7390    Population Count  60561907

R-Square for dependent variable OVERWT (Cox & Snell, 1989): 0.061541


Processing data for set 4 of imputed variables:

Number of zero responses     : 11427
Number of non-zero responses :  7398

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     52

Maximum number of estimable parameters for the model is 20

Sample and Population Counts for Response Variable OVERWT
Based on observations used in the analysis
0:  Sample Count    11427    Population Count 116440809
1:  Sample Count     7398    Population Count  60739861

R-Square for dependent variable OVERWT (Cox & Snell, 1989): 0.059800


Processing data for set 5 of imputed variables:

Number of zero responses     : 11456
Number of non-zero responses :  7369

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     52
```

Exhibit 2. First Page of SUDAAN Results - Cont

```
Maximum number of estimable parameters for the model is 20

Sample and Population Counts for Response Variable OVERWT
Based on observations used in the analysis
0:  Sample Count    11456    Population Count 116566094
1:  Sample Count     7369    Population Count  60614576

R-Square for dependent variable OVERWT (Cox & Snell, 1989): 0.061515

Overall degrees of freedom (Rubin): 30.11
```

Exhibit 3. Frequency Distributions for CLASS Variables (AGEGRP)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: AGEGRP.
------------------------------------
AGEGRP          Frequency      Value
------------------------------------
Ordered
  Position:
  1                  7377    1=20-39
Ordered
  Position:
  2                  4852    2=40-59
Ordered
  Position:
  3                  6596      3=60+
------------------------------------
```

Exhibit 4. Frequency Distributions for CLASS Variables (DMARETHN)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Race/Ethnicity.
---------------------------------------
Race/Ethnic-
  ity           Frequency         Value
---------------------------------------
Ordered
  Position:
  1                  8146    1=nH White
Ordered
  Position:
  2                  5071    2=nH Black
Ordered
  Position:
  3                  4893    3=Mex-Amer
Ordered
  Position:
  4                   715       4=Other
---------------------------------------
```

Exhibit 5. Frequency Distributions for CLASS Variables (POVERTY)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Poverty Status.
---------------------------------------------------
Poverty
  Status             Frequency                Value
---------------------------------------------------
Ordered
  Position:
  1                   4446.800    1=At/Below Poverty
Ordered
  Position:
  2                  14378.200       2=Above Poverty
---------------------------------------------------
```

Exhibit 6. Frequency Distributions for CLASS Variables (ACTIVE_MI)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Activity level compared with others your age.
----------------------------------------------
Activity
  level
  compared
  with
  others
  your age           Frequency           Value
----------------------------------------------
Ordered
  Position:
  1                   5938.200    1=More Active
Ordered
  Position:
  2                   8611.800           2=Same
Ordered
  Position:
  3                   4275.000    3=Less Active
----------------------------------------------
```

Exhibit 7. Frequency Distributions for CLASS Variables (HSSEX)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Sex.
-------------------------------------
Sex             Frequency       Value
-------------------------------------
Ordered
  Position:
  1                  8816      1=Male
Ordered
  Position:
  2                 10009    2=Female
-------------------------------------
```

Exhibit 8. Frequency Distributions for CLASS Variables (HAB1MI)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Self-Rating Health Status.
--------------------------------------------
Self-Rating
  Health
  Status            Frequency          Value
--------------------------------------------
Ordered
  Position:
  1                  2823.600    1=Excellent
Ordered
  Position:
  2                  4388.200    2=Very Good
Ordered
  Position:
  3                  6741.000         3=Good
Ordered
  Position:
  4                  3834.800         4=Fair
Ordered
  Position:
  5                  1037.400         5=Poor
--------------------------------------------
```

Note that HAB1MI (self-rating of general health status), POVERTY (at or below vs. above), and ACTIVE_MI (self-reported more/same/less activity vs. others your age) are all multiply imputed variables. They were defined with other non-MI variables on the CLASS statement in ***Exhibit 1***. The above frequency distributions for the MI variables represent the average frequency of the MI variables on the 5 datasets. Therefore, decimal places are allowed in their frequency distributions.

Exhibit 9. Regression Coefficients (BETAS group)

```
Variance Estimation Method: BRR Using Multiply Imputed Data
Working Correlations: Independent
Link Function: Logit
Response variable OVERWT: Overweight-Derived from BMI
For Subpopulation: Age 20+

Modelling Prob of Being Classified as Overweight with MI Data (BRR)

Results for Summary Over All Imputations
-----------------------------------------------------------------------
Independent Variables and
  Effects                                                     P-value
                                   Beta                T-Test  T-Test
                                   Coeff.     SE Beta  B=0     B=0
-----------------------------------------------------------------------
Intercept                          -0.2096     0.0916   -2.29    0.0267
AGEGRP
  1=20-39                          -0.5980     0.0766   -7.81    0.0000
  2=40-59                           0.0956     0.0687    1.39    0.1706
  3=60+                             0.0000     0.0000     .        .
Race/Ethnicity
  1=nH White                        0.0000     0.0000     .        .
  2=nH Black                        0.2780     0.0850    3.27    0.0020
  3=Mex-Amer                        0.4858     0.0783    6.20    0.0000
  4=Other                          -0.2183     0.3113   -0.70    0.4885
Poverty Status
  1=At/Below Poverty                0.0021     0.0643    0.03    0.9746
  2=Above Poverty                   0.0000     0.0000     .        .
Activity level compared with
  others your age
  1=More Active                    -0.7158     0.0882   -8.12    0.0000
  2=Same                           -0.2764     0.0718   -3.85    0.0003
  3=Less Active                     0.0000     0.0000     .        .
Sex
  1=Male                           -0.0484     0.0510   -0.95    0.3473
  2=Female                          0.0000     0.0000     .        .
Self-Rating Health Status
  1=Excellent                      -0.3833     0.1222   -3.14    0.0030
  2=Very Good                       0.0441     0.0949    0.46    0.6446
  3=Good                            0.3065     0.0831    3.69    0.0006
  4=Fair                            0.3973     0.0889    4.47    0.0000
  5=Poor                            0.0000     0.0000     .        .
Race/Ethnicity, Activity level
  compared with others your age
  1=nH White, 1=More Active         0.0000     0.0000     .        .
  1=nH White, 2=Same                0.0000     0.0000     .        .
  1=nH White, 3=Less Active         0.0000     0.0000     .        .
  2=nH Black, 1=More Active         0.3240     0.1144    2.83    0.0067
  2=nH Black, 2=Same                0.1155     0.0960    1.20    0.2350
  2=nH Black, 3=Less Active         0.0000     0.0000     .        .
  3=Mex-Amer, 1=More Active         0.0947     0.1188    0.80    0.4306
  3=Mex-Amer, 2=Same               -0.2345     0.0965   -2.43    0.0192
  3=Mex-Amer, 3=Less Active         0.0000     0.0000     .        .
  4=Other, 1=More Active           -0.0268     0.4020   -0.07    0.9472
  4=Other, 2=Same                   0.0656     0.3523    0.19    0.8534
  4=Other, 3=Less Active            0.0000     0.0000     .        .
-----------------------------------------------------------------------
```

Exhibit 9 shows the estimated regression coefficients for the main effects plus interaction model. The two significant pairwise interaction *p*-values (p=0.0067 and p=0.0192) indicate that there may be a significant interaction between activity level and race on overweight status.

Exhibit 10. ANOVA Table (TESTS group)

```
Variance Estimation Method: BRR Using Multiply Imputed Data
Working Correlations: Independent
Link Function: Logit
Response variable OVERWT: Overweight-Derived from BMI
For Subpopulation: Age 20+

Modelling Prob of Being Classified as Overweight with MI Data (BRR)

Results for Summary Over All Imputations
by: Contrast.
-----------------------------------------------------------------
Contrast                           Degrees
                                   of                     P-value
                                   Freedom       Wald F   Wald F
-----------------------------------------------------------------
OVERALL MODEL                           20        87.61    0.0000
MODEL MINUS INTERCEPT                   19        53.47    0.0000
INTERCEPT                                .          .        .
AGEGRP                                   2        79.20    0.0000
DMARETHN                                 .          .        .
POVERTY                                  1         0.00    0.9746
ACTIVE_MI                                .          .        .
HSSEX                                    1         0.90    0.3499
HAB1MI                                   4        41.34    0.0000
DMARETHN * ACTIVE_MI                     6         3.33    0.0123
Active=More vs Less: nH White            1        65.92    0.0000
Active=More vs Less: nH Black            1        23.06    0.0000
Active=More vs Less: Mex Amer            1        53.36    0.0000
-----------------------------------------------------------------
```

The ANOVA table above (***Exhibit 10***) indicates that there is a significant (*p*=0.0123) interaction effect between race and activity level on overweight status. The last three lines are produced by the EFFECTS statement. They tell us that the effect of more vs. less activity on overweight status is significant for each racial group. Odds ratios are one way of exploring the interaction effect in more detail. Another way would be to estimate risks (marginals) and risk ratios, but only the odds ratio approach is pursued in this example.

Exhibit 11. Default Odds Ratios (RISK group)

```
Variance Estimation Method: BRR Using Multiply Imputed Data
Working Correlations: Independent
Link Function: Logit
Response variable OVERWT: Overweight-Derived from BMI
For Subpopulation: Age 20+

Modelling Prob of Being Classified as Overweight with MI Data (BRR)

Results for Summary Over All Imputations
by: Independent Variables and Effects.
---------------------------------------------------------------------
Independent Variables and
  Effects                                    Lower 95%   Upper 95%
                                 Odds Ratio   Limit OR    Limit OR
---------------------------------------------------------------------
Intercept                              0.81        0.67        0.98
AGEGRP
  1=20-39                              0.55        0.47        0.64
  2=40-59                              1.10        0.96        1.26
  3=60+                                1.00         .           .
Race/Ethnicity
  1=nH White                           1.00         .           .
  2=nH Black                           1.32        1.11        1.57
  3=Mex-Amer                           1.63        1.39        1.90
  4=Other                              0.80        0.43        1.52
Poverty Status
  1=At/Below Poverty                   1.00        0.88        1.14
  2=Above Poverty                      1.00         .           .
Activity level compared with
  others your age
  1=More Active                        0.49        0.41        0.58
  2=Same                               0.76        0.66        0.88
  3=Less Active                        1.00         .           .
Sex
  1=Male                               0.95        0.86        1.06
  2=Female                             1.00         .           .
Self-Rating Health Status
  1=Excellent                          0.68        0.53        0.87
  2=Very Good                          1.05        0.86        1.27
  3=Good                               1.36        1.15        1.61
  4=Fair                               1.49        1.24        1.78
  5=Poor                               1.00         .           .
Race/Ethnicity, Activity level
  compared with others your age
  1=nH White, 1=More Active            1.00         .           .
  1=nH White, 2=Same                   1.00         .           .
  1=nH White, 3=Less Active            1.00         .           .
  2=nH Black, 1=More Active            1.38        1.10        1.74
  2=nH Black, 2=Same                   1.12        0.93        1.36
  2=nH Black, 3=Less Active            1.00         .           .
  3=Mex-Amer, 1=More Active            1.10        0.86        1.40
  3=Mex-Amer, 2=Same                   0.79        0.65        0.96
  3=Mex-Amer, 3=Less Active            1.00         .           .
  4=Other, 1=More Active               0.97        0.43        2.20
  4=Other, 2=Same                      1.07        0.52        2.19
  4=Other, 3=Less Active               1.00         .           .
---------------------------------------------------------------------
```

Exhibit 11 contains the *default* odds ratios that are obtained by exponentiating each regression coefficient in the model. They can be hard to interpret when there is an interaction effect in the model. For example, this table tells us that the odds of being classified as overweight for people who are more active compared to those less active is 0.49 (*i.e.*, odds are reduced by more than half), but this is specifically for Whites (the user-specified reference cell for the interacting variable DMARETHN). The odds ratios for the other races require exponentiating a linear combination of the regression coefficients. This is where the EXP option (next set of results) comes into play.

Exhibit 12. Customized Odds Ratios (EXPCNTRST group)

```
Variance Estimation Method: BRR Using Multiply Imputed Data
Working Correlations: Independent
Link Function: Logit
Response variable OVERWT: Overweight-Derived from BMI
For Subpopulation: Age 20+

Modelling Prob of Being Classified as Overweight with MI Data (BRR)

Results for Summary Over All Imputations
-------------------------------------------------------------------
Contrast                                            Lower    Upper
                                                    95%      95%
                                    EXP(Contrast)   Limit    Limit
-------------------------------------------------------------------
Active=More vs Less: nH White                0.49     0.41     0.58
Active=More vs Less: nH Black                0.68     0.57     0.80
Active=More vs Less: Mex Amer                0.54     0.45     0.64
-------------------------------------------------------------------
```

This table contains the results produced by specifying the EXP option on the EFFECTS statements in ***Exhibit 1***:

```
EFFECTS ACTIVE_MI=(1 0 -1) / DMARETHN=1 EXP name="Active=More vs Less: nH White";
EFFECTS ACTIVE_MI=(1 0 -1) / DMARETHN=2 EXP name="Active=More vs Less: nH Black";
EFFECTS ACTIVE_MI=(1 0 -1) / DMARETHN=3 EXP name="Active=More vs Less: Mex Amer";
```

The results under the column headed EXP(Contrast) in ***Exhibit 12*** are the user-specified odds ratios obtained by exponentiating the contrast defined on the EFFECTS statement. The results show that the odds of being overweight for those who are more vs. less active is estimated as 0.49 for Whites, 0.68 for Blacks, and 0.54 for Mexican Americans (the default odds ratios produced the 0.49 for Whites, but told us nothing about the other races). Each odds ratio shows a reduction in odds for being overweight when people are more active (as one would hope!). However, the reduction is estimated to be largest among Whites, followed by Mexican Americans and Blacks. Also, each odds ratio is significantly different from the null value of 1.0, since the value 1.0 is not contained in any of the confidence intervals.

The odds ratio for Whites seems to show a significantly greater effect from increased activity when compared to that for Blacks, since the confidence interval for Whites shows very little overlap with the confidence interval for Blacks (***Exhibit 12***). This is apparently why we found a significant interaction effect between activity level and race on overweight status. Both races benefit from increased activity, but Whites appear to benefit more.