LOGLINK Example #2

SUDAAN Statements and Results Illustrated

- Log-linear regression modeling
- SEMETHOD
- REFLEVEL
- EFFECTS
- PREDMARG

Input Data Set(s): PERSONSX.SAS7BDAT

Example

Using the 2006 National Health Interview Survey (NHIS), Predict Self-Reported Doctor's Visits During the Past 2 Weeks.

This example highlights new features in SUDAAN 11, namely 1) the UNITS option for obtaining IDRs associated with a multiple-unit change in a continuous covariate, and 2) confidence intervals for predicted marginals.

Solution

In this example, we analyze the 2006 NHIS, a survey conducted annually by the National Center for Health Statistics (NCHS).The sampling design statements will include DESIGN=WR on the PROC statement, with design variables *stratum_p* and *psu_p* and a sampling weight variable (*wtfa*), as follows:

```
PROC LOGLINK ... DESIGN=WR;
NEST stratum_p psu_p;
WEIGHT wtfa;
```

In the following example (see ***Exhibit 1***), we model the number of doctor visits in the past 2 weeks using the LOGLINK procedure. Since this is an analysis of complex survey data, we use the default robust variance estimator (SEMETHOD=BINDER) in all examples.

Using PROC CROSSTAB, we request frequency distributions and covariate means for the variables in our model. We use the SUBPOPX statement to subset the analysis to the subpopulation of interest.

This example was run in SAS-Callable SUDAAN, and the SAS program and *.LST files are provided.

Exhibit 1. SAS-Callable SUDAAN Code for Data Manipulation and CROSSTAB Procedure

```
libname in "c:\10winbetatest\amjepid";

proc format;
 value m  1="Married"
          2="Unmarried";
 value i_ 1="Less than 20K"
          2="20K+"
          3="Unknown";
 value  E_ 1="Less Than HS"
           2="HS"
           3="Greater Than HS"
           4="Unknown";
 value  H_ 1="Excellent/Very Good"
           2="Good"
           3="Fair/Poor"
           4="Unknown";
 value RE_ 1="White"
           2="Black"
           3="Hispanic"
           4="Other";
 value s  1="Males"
          2="Females";

data nhis06; set in.personsx;
  if cdcmstat=3 then married=1;
  else married=2;

  if ernyr_p <=04 then INC20K=1;
  else if ernyr_p <=11 then inc20k=2;
  else inc20k=3;

  if educ1 <=12 then educ=1;
  else if educ1 <=15 then educ=2;
  else if educ1 <=21 then educ=3;
  else if educ1=96 then educ=.;
  else educ=4;

  if phstat in ( 1 2) then health=1;
  else if phstat=3 then health=2;
  else if phstat in ( 4 5) then health=3;
  else health=4;

  if ORIGIN_I=1 then RACETHN = 3;
  else if (origin i=2 and racreci3=1) then racethn=1;
  else if  (origin_i=2 and racreci3=2) then racethn=2;
  else racethn=4;

  if phcdv2w=1 then do;
  if phcdvn2w < 50 then DOCVIS2W=phcdvn2w;
  else DOCVIS2W=.;
  end;
  else if phcdv2w=2 then DOCVIS2W=0;

  label married="Married";
  label health="Health Status";
  label racethn="Race/Ethnicity";
  label inc20k="Income";
  label educ="Educ Responsible Adult";
  label docvis2w="Doctor visits, 2 Weeks";
  format married m_. inc20k i_. educ e_. health h_. racethn RE_. sex s_.;

proc sort data=nhis06;
  by STRAT_P PSU_P;


PROC CROSSTAB DATA=nhis06 FILETYPE=SAS DESIGN=WR;
  NEST STRAT_P PSU_P;
  WEIGHT wtfa;
```

```
SUBPOPX MARRIED=1 AND 18 le AGE_P le 64 / NAME="Married: Ages 18-64 years";
CLASS SEX HEALTH RACETHN INC20K EDUC;

SETENV colwidth=7 decwidth=2;
PRINT NSUM WSUM COLPER SECOL / STYLE=NCHS NSUMFMT=F6.0 WSUMFMT=F9.0;
RTITLE "NHIS Data (2006)"  "Covariate Means and Percentage Distributions";
```

Exhibit 2. First Page of CROSSTAB Output (*.lst file)

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


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: WTFA
    Stratification Variables(s): STRAT_P
    Primary Sampling Unit: PSU_P


Number of observations read    :  75456    Weighted count :293755796
Number of observations skipped :    260
(WEIGHT variable nonpositive)
Observations in subpopulation  :  26420    Weighted count :106108785
Denominator degrees of freedom :    300
```

The frequencies for each CLASS variable are shown below in ***Exhibit 3*** through ***Exhibit 7***.

Exhibit 3. Frequencies and Values for CLASS Variable SEX

```
Frequencies and Values for CLASS Variables
-------------------------------------
Sex             Frequency      Value
-------------------------------------
Ordered
  Position:
  1                 12888      Males
Ordered
  Position:
  2                 13532    Females
-------------------------------------
```

Exhibit 4. Frequencies and Values for CLASS Variable HEALTH STATUS

```
Frequencies and Values for CLASS Variables
-------------------------------------------------
Health
  Status        Frequency                  Value
-------------------------------------------------
Ordered
  Position:
  1                 17044    Excellent/Very Good
Ordered
  Position:
  2                  6885                   Good
Ordered
  Position:
  3                  2411              Fair/Poor
Ordered
  Position:
  4                    80                Unknown
-------------------------------------------------
```

Exhibit 5. Frequencies and Values for CLASS Variable Race/Ethnicity

```
Frequencies and Values for CLASS Variables
-----------------------------------
Race/Ethnic-
  ity           Frequency       Value
-----------------------------------
Ordered
  Position:
  1                15760       White
Ordered
  Position:
  2                 2634       Black
Ordered
  Position:
  3                 5831    Hispanic
Ordered
  Position:
  4                 2195       Other
-----------------------------------
```

Exhibit 6. Frequencies and Values for CLASS Variable INCOME

```
Frequencies and Values for CLASS Variables
-----------------------------------------
Income          Frequency            Value
-----------------------------------------
Ordered
  Position:
  1                 9105    Less than 20K
Ordered
  Position:
  2                10431            20K+
Ordered
  Position:
  3                 6884          Unknown
-----------------------------------------
```

Exhibit 7. Frequencies and Values for CLASS Variable EDUC

```
Frequencies and Values for CLASS Variables
-------------------------------------------
Educ
  Responsible
  Adult         Frequency              Value
-------------------------------------------
Ordered
  Position:
  1                 4198       Less Than HS
Ordered
  Position:
  2                11064                 HS
Ordered
  Position:
  3                10165    Greater Than HS
Ordered
  Position:
  4                  993            Unknown
-------------------------------------------
```

The requested CROSSTAB output appears next in ***Exhibit 8*** through ***Exhibit 12***. Note that the percentages are population estimates.

Exhibit 8. CROSSTAB Results for SEX

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: Married: Ages 18-64 years

NHIS Data (2006)
Covariate Means and Percentage Distributions

------------------------------------------------------------
Sex
                      Sample   Weighted    Col       SE Col
                      Size     Size        Percent   Percent
------------------------------------------------------------
Total                  26420   106108785    100.00      0.00
Males                  12888    52725327     49.69      0.11
Females                13532    53383458     50.31      0.11
------------------------------------------------------------
```

Exhibit 9. CROSSTAB Results for Health Status

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: Married: Ages 18-64 years

NHIS Data (2006)
Covariate Means and Percentage Distributions

------------------------------------------------------------
Health Status
                      Sample   Weighted    Col       SE Col
                      Size     Size        Percent   Percent
------------------------------------------------------------
Total                  26420   106108785    100.00      0.00
Excellent/Very Good    17044    70366220     66.32      0.47
Good                    6885    26339048     24.82      0.40
Fair/Poor               2411     9058757      8.54      0.24
Unknown                   80      344760      0.32      0.05
------------------------------------------------------------
```

Exhibit 10. CROSSTAB Results for Race/Ethnicity

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: Married: Ages 18-64 years

NHIS Data (2006)
Covariate Means and Percentage Distributions

------------------------------------------------------------
Race/Ethnicity
                      Sample   Weighted    Col       SE Col
                      Size     Size        Percent   Percent
------------------------------------------------------------
Total                  26420   106108785    100.00      0.00
White                  15760    76178846     71.79      0.54
Black                   2634     8503698      8.01      0.29
Hispanic                5831    14925177     14.07      0.41
Other                   2195     6501064      6.13      0.26
------------------------------------------------------------
```

Exhibit 11. CROSSTAB Results for INCOME

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: Married: Ages 18-64 years

NHIS Data (2006)
Covariate Means and Percentage Distributions

------------------------------------------------------------
Income
                      Sample   Weighted    Col       SE Col
                      Size     Size        Percent   Percent
------------------------------------------------------------
Total                  26420   106108785    100.00      0.00
Less than 20K           9105    34888178     32.88      0.39
20K+                   10431    44443561     41.88      0.54
Unknown                 6884    26777046     25.24      0.54
------------------------------------------------------------
```

Exhibit 12. CROSSTAB Results for EDUC

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: Married: Ages 18-64 years

NHIS Data (2006)
Covariate Means and Percentage Distributions

------------------------------------------------------------
Educ Responsible
  Adult               Sample   Weighted    Col       SE Col
                      Size     Size        Percent   Percent
------------------------------------------------------------
Total                  26420   106108785    100.00      0.00
Less Than HS            4198    13454859     12.68      0.36
HS                     11064    45498584     42.88      0.50
Greater Than HS        10165    43434165     40.93      0.57
Unknown                  993     3721177      3.51      0.23
------------------------------------------------------------
```

Exhibit 13 contains the DESCRIPT code for basic descriptive statistics for the response variable (DOCVIS2W) and a continuous covariate (AGE_P). Note that the means are population estimates.

Exhibit 13. SAS-callable SUDAAN Code for DESCRIPT Procedure

```
PROC DESCRIPT DATA=nhis06 FILETYPE=SAS DESIGN=WR NOMARG;
NEST STRAT_P PSU_P;
WEIGHT wtfa;

SUBPOPX MARRIED=1 AND 18 le AGE_P le 64 / NAME="Married: Ages 18-64 years";
VAR AGE_P DOCVIS2W;

setenv colwidth=8 decwidth=2;
PRINT NSUM WSUM MEAN SEMEAN / STYLE=NCHS NSUMFMT=F6.0 WSUMFMT=F9.0;
RTITLE "NHIS Data (2006)"  "Covariate Means and Percentage Distributions";
```

Exhibit 14. First Page of Output for DESCRIPT Procedure

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


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: WTFA
    Stratification Variables(s): STRAT_P
    Primary Sampling Unit: PSU_P


Number of observations read    :  75456    Weighted count :293755796
Number of observations skipped :    260
(WEIGHT variable nonpositive)
Observations in subpopulation  :  26420    Weighted count :106108785
Denominator degrees of freedom :    300
```

Exhibit 15. DESCRIPT Procedure Results for AGE and Doctor Visits

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: Married: Ages 18-64 years

NHIS Data (2006)
Covariate Means and Percentage Distributions
-------------------------------------------------------------------
Variable
   SUDAAN Reserved       Sample   Weighted
     Variable One        Size     Size            Mean    SE Mean
-------------------------------------------------------------------
Age
   1                     26420   106108785      43.82       0.11
Doctor visits, 2 weeks
   1                     26196   105141005       0.20       0.01
-------------------------------------------------------------------
```

Next, we perform the LOGLINK regression (***Exhibit 16***).

Exhibit 16. SAS-callable SUDAAN Code for LOGLINK Procedure

```
PROC LOGLINK DATA=nhis06 FILETYPE=SAS DESIGN=WR;
  NEST STRAT_P PSU_P;
  WEIGHT wtfa;

  SUBPOPX MARRIED=1 AND 18 le AGE_P le 64 / NAME="Married: Ages 18-64 years";

  REFLEVEL INC20K=2 HEALTH=1 RACETHN=2 EDUC=1;
  CLASS SEX HEALTH RACETHN INC20K EDUC;
  MODEL DOCVIS2W = SEX INC20K EDUC HEALTH RACETHN AGE_P;

  EFFECTS AGE_P / exp units=10 name="AGE";
  PREDMARG SEX;
  PREDMARG HEALTH;
  PRED_EFF SEX=(1 -1) / name="Males-Females";
  PRED EFF HEALTH=(-1 1 0 0) / name="Excellent vs Good Health";
  PRED_EFF HEALTH=(-1 0 1 0) / name="Excellent vs Fair/Poor Health";

  SETENV COLSPCE=1 TOPMGN=0 COLWIDTH=7 DECWIDTH=4 LABWIDTH=25;
  PRINT / betas=default tests=default t_betafmt=f6.2 waldffmt=f6.2 dffmt=f7.0;

  SETENV labwidth=25 colwidth=6 decwidth=3;
  PRINT / idratio=default idrfmt=f9.3;

  SETENV COLSPCE=1 TOPMGN=0 COLWIDTH=5 DECWIDTH=3 LABWIDTH=20;
  PRINT / expcntrst=default exp_cntrstfmt=f13.3 unitsfmt=f5.0;

  SETENV COLSPCE=1 TOPMGN=0 COLWIDTH=7 DECWIDTH=4 LABWIDTH=25;
  PRINT / pred_mrg=default prmgcons=default lowpmfmt=f5.3 uppmfmt=f5.3
          predmrgfmt=f9.4 t_prdmrgfmt=f8.2 prmgconfmt=f8.4 t_pmconfmt=f7.2;

  rlabel docvis2w="Doctor visits, 2 Weeks";
  rformat educ e_.;
  rformat health h_.;
  rformat racethn re_.;
  rformat inc20k i_.;
  rformat sex s_.;
  RTITLE "Modelling Number of Doctor Visits, Past 2 Weeks"  " ";
  RFOOTNOTE "Source: National Center for Health Statistics, 2006";
```

We model the number of doctor visits in the past 2 weeks as a function of sex, income, education, health status, race, and age. Age is modeled as a continuous covariate, and all other independent variables are modeled as categorical (all variables appearing on the CLASS statements are modeled as categorical). The REFLEVEL statement defines the reference cell for each categorical covariate (the default reference cell is the last level of each categorical covariate).

We request the estimated predicted marginals and contrasts among predicted marginals for the variables sex and health status. ***New in SUDAAN 11:*** Confidence intervals for the marginals are included by default, in addition to the standard error and *t*-test. Predicted marginals in nonlinear models are analogous to adjusted means in linear models.

Finally, we request the IDR associated with a 10-year increase in AGE ***(also new in SUDAAN 11).*** IDRs associated with a 1-unit increase in each covariate are included by default.

Exhibit 17. First Page of LOGLINK Output (*.LST file)

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


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: WTFA
    Stratification Variables(s): STRAT_P
    Primary Sampling Unit: PSU_P

Independence parameters have converged in 8 iterations

Number of observations read       :  75456    Weighted count:293755796
Number of observations skipped    :    260
(WEIGHT variable nonpositive)
Observations in subpopulation     :  26420    Weighted count:106108785
Observations used in the analysis :  26196    Weighted count:105141005
Denominator degrees of freedom    :    300


Maximum number of estimable parameters for the model is 14

File NHIS06 contains  600 Clusters
 600 clusters were used to fit the model
Maximum cluster size is 122 records
Minimum cluster size is   7 records

Weighted mean response is 0.204257

-2 * Normalized Log-Likelihood with Intercepts Only : -27699.31
-2 * Normalized Log-Likelihood Full Model           : -25610.11
Approximate Chi-Square (-2 * Log-L Ratio)           :  2089.20
Degrees of Freedom                                  :       13

Note: The approximate Chi-Square is not adjusted for clustering.
      Refer to hypothesis test table for adjusted test.
```

The SUBPOPX statement (***Exhibit 16***) restricts the regression analysis to those who are married, aged 18-64 years (26,420 observations).

Exhibit 18 contains the estimated regression coefficients for the fitted model. Men had significantly fewer doctor visits in the two weeks compared to women, and people reporting excellent health visited the doctor significantly less frequently than those reporting fair/poor health.

Exhibit 18. LOGLINK Regression Coefficients

Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks

Independent Variables and Effects	Beta Coeff.	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
Intercept	-2.4041	0.1386	-2.6769	-2.1312	-17.34	0.0000
Sex						
Males	-0.3856	0.0493	-0.4827	-0.2885	-7.82	0.0000
Females	0.0000	0.0000	0.0000	0.0000	.	.
Health Status						
Excellent/Very Good	0.0000	0.0000	0.0000	0.0000	.	.
Good	0.4145	0.0511	0.3139	0.5152	8.11	0.0000
Fair/Poor	1.3701	0.0651	1.2420	1.4982	21.04	0.0000
Unknown	1.4109	0.4895	0.4477	2.3741	2.88	0.0042
Race/Ethnicity						
White	0.0831	0.0849	-0.0840	0.2502	0.98	0.3286
Black	0.0000	0.0000	0.0000	0.0000	.	.
Hispanic	-0.3842	0.1051	-0.5911	-0.1773	-3.65	0.0003
Other	-0.1707	0.1199	-0.4067	0.0652	-1.42	0.1554
Income						
Less than 20K	0.0166	0.0512	-0.0840	0.1173	0.33	0.7452
20K+	0.0000	0.0000	0.0000	0.0000	.	.
Unknown	-0.2483	0.0758	-0.3975	-0.0990	-3.27	0.0012
Educ Responsible Adult						
Less Than HS	0.0000	0.0000	0.0000	0.0000	.	.
HS	0.2263	0.0732	0.0823	0.3702	3.09	0.0022
Greater Than HS	0.4281	0.0726	0.2853	0.5709	5.90	0.0000
Unknown	-1.1472	0.3178	-1.7726	-0.5219	-3.61	0.0004
Age	0.0097	0.0020	0.0058	0.0136	4.91	0.0000

Source: National Center for Health Statistics, 2006

Exhibit 19. Incidence Density Ratios (IDRs)

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks
```

Independent Variables and Effects	Incidence Density Ratio	Lower 95% Limit IDR	Upper 95% Limit IDR
Intercept	0.090	0.069	0.119
Sex			
Males	0.680	0.617	0.749
Females	1.000	1.000	1.000
Health Status			
Excellent/Very Good	1.000	1.000	1.000
Good	1.514	1.369	1.674
Fair/Poor	3.936	3.462	4.474
Unknown	4.100	1.565	10.742
Race/Ethnicity			
White	1.087	0.919	1.284
Black	1.000	1.000	1.000
Hispanic	0.681	0.554	0.838
Other	0.843	0.666	1.067
Income			
Less than 20K	1.017	0.919	1.124
20K+	1.000	1.000	1.000
Unknown	0.780	0.672	0.906
Educ Responsible Adult			
Less Than HS	1.000	1.000	1.000
HS	1.254	1.086	1.448
Greater Than HS	1.534	1.330	1.770
Unknown	0.318	0.170	0.593
Age	1.010	1.006	1.014

```
Source: National Center for Health Statistics, 2006
```

Exponentiating the estimated regression coefficients yields the incidence density ratios (IDR), which can be interpreted as the ratios of event rates for each one-unit increase in the corresponding covariate. We see from ***Exhibit 19*** that the estimated IDR for males vs. females is 0.68, indicating a reduction in the rate of doctor visits by about one-third among males. Also, the IDR for people reporting fair/poor vs. excellent health is 3.94, indicating more than a tripling in the rate of doctor visits among those in fair/poor health.

Exhibit 20. User-Requested IDR for AGE (10-year increase)

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks
-------------------------------------------------------------
Contrast                                        Lower   Upper
                                                95%     95%
                       Units   EXP(Contrast)    Limit   Limit
-------------------------------------------------------------
AGE                       10           1.102   1.060   1.146
-------------------------------------------------------------
Source: National Center for Health Statistics, 2006
```

New The following EFFECTS statement produced the results in ***Exhibit 20:***

```
EFFECTS age_p / exp units=10 name="AGE";
```

SUDAAN estimated the IDR associated with a 10-year increase in AGE, which can be interpreted as the ratio of event rates for each 10-year increase in AGE. The IDR of 1.102 indicates a 10.2% increase in the rate of doctor visits for each 10-year increase in age. The 95% confidence limits indicate that this IDR is significantly different from the null value of 1.0.

Exhibit 21. ANOVA Table

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks

-------------------------------------------------------
Contrast                    Degrees
                            of                  P-value
                            Freedom   Wald F    Wald F
-------------------------------------------------------
OVERALL MODEL                    14   391.57     0.0000
MODEL MINUS INTERCEPT            13    71.36     0.0000
INTERCEPT                         .        .          .
SEX                               1    61.08     0.0000
HEALTH                            3   147.75     0.0000
RACETHN                           3    14.65     0.0000
INC20K                            2     7.35     0.0008
EDUC                              3    21.09     0.0000
AGE_P                             1    24.15     0.0000
-------------------------------------------------------
Source: National Center for Health Statistics, 2006
```

Exhibit 21 contains the test for each model term. Adjusted for all other covariates, sex, health status, race/ethnicity, income, education, and age are all statistically significant.

Exhibit 22. Predicted Marginals for SEX

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks

OFFSET=Unit
-------------------------------------------------------------------------------
Predicted Marginal #1                              Lower   Upper
                          Predicted                95%     95%
                          Marginal         SE      Limit   Limit   T:Marg=0   P-value
-------------------------------------------------------------------------------
Sex
  Males                      0.1647    0.0069   0.152   0.179      24.03    0.0000
  Females                    0.2422    0.0069   0.229   0.256      35.20    0.0000
-------------------------------------------------------------------------------
Source: National Center for Health Statistics, 2006
```

The estimated average marginal prediction (marginalizing over the covariate distribution in the population) for the number of doctor visits is 0.24 for females, and 0.16 for males (***Exhibit 22***). The confidence limits for males and females (***new in Release 11***) are non-overlapping, indicating an obvious significant difference.

New

Exhibit 23. Predicted Marginals for Health Status

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks

OFFSET=Unit
-------------------------------------------------------------------------------
Predicted Marginal #2                              Lower   Upper
                          Predicted                95%     95%
                          Marginal         SE      Limit   Limit   T:Marg=0   P-value
-------------------------------------------------------------------------------
Health Status
  Excellent/Very Good        0.1484    0.0049   0.139   0.158      30.20    0.0000
  Good                       0.2247    0.0094   0.207   0.244      24.03    0.0000
  Fair/Poor                  0.5841    0.0320   0.524   0.651      18.25    0.0000
  Unknown                    0.6085    0.2966   0.233   1.588       2.05    0.0411
-------------------------------------------------------------------------------
Source: National Center for Health Statistics, 2006
```

The estimated average marginal prediction for the number of doctor visits in the past 2 weeks is 0.15 for people reporting excellent/very good health, 0.22 for people reporting good health, and 0.58 for people

reporting fair/poor health (***Exhibit 23***). The confidence limits for these 3 groups are all non-overlapping, indicating obvious significant differences between the groups.

Exhibit 24. Differences in Predicted Marginals: males - females

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks

OFFSET=Unit
------------------------------------------------------------------
Contrasted Predicted
  Marginal #1               PREDMARG
                            Contrast        SE    T-Stat   P-value
------------------------------------------------------------------
Males-Females                -0.0775    0.0095     -8.16    0.0000
------------------------------------------------------------------
Source: National Center for Health Statistics, 2006
```

The estimated difference in the number of doctor visits in the past 2 weeks between males and females is -0.075 (p=0.0000, females significantly higher than males), as indicated in ***Exhibit 24***.

Exhibit 25 indicates that the estimated difference in the rate of doctor visits between people reporting to be in excellent vs. good health is 0.076 (p=0.0000, excellent health fewer visits than good health).

Exhibit 25. Differences in Predicted Marginals: Excellent vs. Good Health

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks

OFFSET=Unit
------------------------------------------------------------------
Contrasted Predicted
  Marginal #2               PREDMARG
                            Contrast        SE    T-Stat   P-value
------------------------------------------------------------------
Excellent vs Good Health      0.0762    0.0102      7.46    0.0000
------------------------------------------------------------------
Source: National Center for Health Statistics, 2006
```

Exhibit 26. Differences in Predicted Marginals: Excellent vs. Fair/Poor Health

```
Variance Estimation Method: Taylor Series (WR)
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Log
Response variable DOCVIS2W: Doctor visits, 2 Weeks
Offset variable : None
For Subpopulation: Married: Ages 18-64 years

Modelling Number of Doctor Visits, Past 2 Weeks

OFFSET=Unit
-----------------------------------------------------------------
Contrasted Predicted
  Marginal #3               PREDMARG
                            Contrast        SE    T-Stat   P-value
-----------------------------------------------------------------
Excellent vs Fair/Poor
  Health                      0.4357    0.0326     13.38    0.0000
-----------------------------------------------------------------
Source: National Center for Health Statistics, 2006
```

Exhibit 26 indicates that the estimated difference in number of doctor visits in the past 2 weeks between people reporting to be in excellent vs. fair/poor health is 0.4357 (p=0.0000, excellent health fewer visits than fair/poor health).