

MULTILOG Example #4

SUDAAN Statements and Results Illustrated

- EFFECTS
- CUMLOGIT option
- SUBGROUP
- LEVELS
- SETENV

Input Data Set(s): IRONSUD.SSD

Example

Using data from the NHANES I and its Longitudinal Follow-up Study, model the effects of body iron stores, age group, and smoking status on follow-up cancer status. Use the EFFECTS statement in MULTILOG to test the combined effect of age group and smoke, as well as compare current smokers (level 1) to former smokers (level 2).

Solution

Using the NHANES I Study and its Longitudinal Follow-up, we evaluate the effects of body iron stores at initial exam (B_TIBC, continuous); age group at initial exam (AGEGROUP, 1=20-49, 2=50+); and smoking status (SMOKE, 1=current, 2=former, 3=never, 4=unknown) on follow-up cancer status (CANCER12, 1=yes, 2=no) in a logistic regression model.

The EFFECTS statement is used to:

1. Test the combined effect of Age Group and Smoke:

```
EFFECTS AGEGROUP SMOKE / NAME = "Combined Age, Smoke";
```

2. Compare Smoke Level 1 (current smoker) to Level 2 (former smoker) (the default reference level for Smoke is Level 4):

```
EFFECTS SMOKE = (-1 1 0 0) / NAME="Smoke 1 vs 2";
```

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

Exhibit 1. SAS-Callable MULTILOG Code

```
LIBNAME IN V604 "C:\Program Files\SUDAAN\SUDAAN10\data";

PROC MULTILOG DATA=in.IRONSUD FILETYPE=SAS DESIGN=WR DEFT2;
NEST Q_STRATA PSU1;
WEIGHT B_WTIRON;

SUBGROUP CANCER12 AGEGROUP SMOKE;
LEVELS 2 2 4;
MODEL CANCER12 = B_TIBC AGEGROUP SMOKE / CUMLOGIT;
EFFECTS AGEGROUP SMOKE / NAME = "Combined Age, Smoke";
EFFECTS SMOKE=(-1 1 0 0) / NAME = "Smoke 1 vs 2";

SETENV COLSPCE=1 LABWIDTH=25 COLWIDTH=8 DECWIDTH=4;
PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
      P_BETA="P-VALUE" DF WALDCHI WALDCHP /
      T_BETAfmt=F8.2 DEFTfmt=F6.2 DFFMT=F8.0 WALDCHIFMT=F8.2;
RTITLE "EFFECTS Statement Examples";
```

When the outcome has two levels, a logistic regression model is fit to the data. The CUMLOGIT and GENLOGIT links are equivalent in that situation.

Exhibit 2. First Page of MULTILOG Output

```
              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: B_WTIRON
Stratification Variables(s): Q_STRATA
Primary Sampling Unit: PSU1

Independence parameters have converged in 6 iterations

Number of observations read      :    3290      Weighted count: 40570323
Observations used in the analysis :    3290      Weighted count: 40570323
Denominator degrees of freedom   :         35

Maximum number of estimable parameters for the model is 6

File in.IRONSUD contains 67 Clusters
67 clusters were used to fit the model
Maximum cluster size is 111 records
Minimum cluster size is 15 records

Sample and Population Counts for Response Variable CANCER12
Based on observations used in the analysis
 1: Sample Count    232      Population Count  1745695
 2: Sample Count   3058      Population Count  38824628

-2 * Normalized Log-Likelihood with Intercepts Only : 1167.64
-2 * Normalized Log-Likelihood Full Model           : 1026.95
Approximate Chi-Square (-2 * Log-L Ratio)           : 140.70
Degrees of Freedom                                  :         5

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

Exhibit 3. Estimated Regression Coefficients

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Cumulative Logit
 Response variable CANCER12: Cancer Status (1/2)

EFFECTS Statement Examples

CANCER12 (cum-logit), Independent Variables and Effects	BETA	S.E.	DESIGN EFFECT	T:BETA=0	P-VALUE
CANCER12 (cum-logit)					
Intercept 1	-0.8618	0.6605	0.94	-1.30	0.2004
Age Cohort					
1	-2.2525	0.3343	1.89	-6.74	0.0000
2	0.0000	0.0000	.	.	.
Smoking Status					
1	-0.5858	0.2771	0.77	-2.11	0.0417
2	-0.9418	0.2922	0.84	-3.22	0.0027
3	-0.4998	0.2743	0.85	-1.82	0.0770
4	0.0000	0.0000	.	.	.
TOTAL IRON BINDING CAPACITY	-0.0024	0.0018	1.10	-1.29	0.2052

Exhibit 4. ANOVA Table

Variance Estimation Method: Taylor Series (WR)
 SE Method: Robust (Binder, 1983)
 Working Correlations: Independent
 Link Function: Cumulative Logit
 Response variable CANCER12: Cancer Status (1/2)

EFFECTS Statement Examples

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	6	708.28	0.0000
MODEL MINUS INTERCEPT	5	64.47	0.0000
AGEGROUP	1	45.39	0.0000
SMOKE	3	10.60	0.0141
B_TIBC	1	1.67	0.1967
Combined Age, Smoke	4	53.16	0.0000
Smoke 1 vs 2	1	1.72	0.1899

Exhibit 4 indicates that the individual and combined effect of Age Group and Smoking Status on follow-up cancer status are all statistically significant ($p=0.0000$). However, current smokers (SMOKE=1) were not significantly different ($p=0.1899$) from former smokers (SMOKE=2) in their follow-up cancer status.