SAS FAQ

How do I do a conditional logit model analysis in SAS 9.1?

PROC LOGISITC has been improved in SAS 9.1 It does a lot more than just logistic regression on binary outcome variables. On this page, we show two examples on using proc logistic for conditional logit models. For conditional logit model, proc logistic is very easy to use and it handles all kinds of matching, 1-1, 1-M matching, and in fact M-N matching.

Example 1: 1-1 Matching

This example is adapted from Chapter 7 of Applied Logistic Regression by Hosmer & Lemeshow (2000). You can download the SAS data file lbwt11.sas7bdat here.

The first 20 observations are listed below. Notice that variable **pairid** indicates that the observations are paired.

pairid	lbwt	age	lastwt	race	smoke	ptd	ht	ui	race1
1	0	14	135	1	0	0	0	0	1
1	1	14	101	3	1	1	0	0	0
2	0	15	98	2	0	0	0	0	0
2	1	15	115	3	0	0	0	1	0
3	0	16	95	3	0	0	0	0	0
3	1	16	130	3	0	0	0	0	0
4	0	17	103	3	0	0	0	0	0
4	1	17	130	3	1	1	0	1	0
5	0	17	122	1	1	0	0	0	1
5	1	17	110	1	1	0	0	0	1
6	0	17	113	2	0	0	0	0	0
6	1	17	120	1	1	0	0	0	1
7	0	17	113	2	0	0	0	0	0
7	1	17	120	2	0	0	0	0	0
8	0	17	119	3	0	0	0	0	0
8	1	17	142	2	0	0	1	0	0
9	0	18	100	1	1	0	0	0	1
9	1	18	148	3	0	0	0	0	0
10	0	18	90	1	1	0	0	1	1
10	1	18	110	2	1	1	0	0	0

```
proc logistic data = lbwt11 descending;
  model lbwt = lastwt smoke race2 race3 ptd ht ui ;
  strata pairid;
run;
```

```
The LOGISTIC Procedure

Conditional Analysis

              Model Information

Data Set                          ATS.LBWT11
Response Variable                 lbwt
Number of Response Levels         2
Number of Strata                  56
Model                             binary logit
Optimization Technique            Newton-Raphson ridge

           Model Information

low brth wt < 2500g

Number of Observations Read          112
Number of Observations Used          112

          Response Profile

 Ordered                      Total
   Value          lbwt     Frequency
```

```
       1            1            56
       2            0            56

Probability modeled is lbwt=1.

                Strata Summary

             lbwt
Response     ------     Number of
 Pattern     1    0        Strata     Frequency

       1     1    1            56           112

Newton-Raphson Ridge Optimization

Without Parameter Scaling

         Convergence criterion (GCONV=1E-8) satisfied.

         Model Fit Statistics

                   Without          With
Criterion       Covariates     Covariates

AIC                77.632         65.589
SC                 77.632         84.618
-2 Log L           77.632         51.589

        Testing Global Null Hypothesis: BETA=0

Test                 Chi-Square       DF     Pr > ChiSq

Likelihood Ratio        26.0439        7         0.0005
Score                   20.2669        7         0.0050
Wald                    12.7208        7         0.0792

            Analysis of Maximum Likelihood Estimates

                                    Standard          Wald
Parameter    DF     Estimate         Error     Chi-Square    Pr > ChiSq

lastwt        1      -0.0184        0.0101         3.3229        0.0683
smoke         1       1.4007        0.6278         4.9770        0.0257
race2         1       0.5714        0.6896         0.6864        0.4074
race3         1      -0.0253        0.6992         0.0013        0.9711
ptd           1       1.8080        0.7887         5.2557        0.0219
ht            1       2.3612        1.0861         4.7259        0.0297
ui            1       1.4019        0.6962         4.0554        0.0440

           Odds Ratio Estimates

             Point          95% Wald
Effect    Estimate      Confidence Limits

lastwt       0.982        0.963       1.001
smoke        4.058        1.185      13.890
race2        1.771        0.458       6.842
race3        0.975        0.248       3.839
ptd          6.098        1.300      28.609
ht          10.603        1.262      89.115
ui           4.063        1.038      15.901
```

Example 2: 1-M matching

This example is adapted from Chapter 7 of Applied Logistic Regression by Hosmer & Lemeshow (2000). You can download the SAS data file bbdm13.sas7bdat here.

The first 20 observations are listed below. Notice that variable **str** indicates that there are four choices for each subject.

str	obs	fndx	chk	agmn	wt	mod	wid	nvmr
1	1	1	1	13	118	55	0	0
1	2	0	2	11	175	1	0	0
1	3	0	2	12	135	1	0	0
1	4	0	1	11	125	55	0	0
2	1	1	1	14	118	55	0	0
2	2	0	2	15	183	55	0	0
2	3	0	2	11	218	55	0	0
2	4	0	1	13	192	55	0	0
3	1	1	1	15	125	55	0	0
3	2	0	2	14	123	55	0	0
3	3	0	1	13	140	55	0	0
3	4	0	1	13	160	55	0	0
4	1	1	1	14	150	55	0	1
4	2	0	1	13	130	1	0	0
4	3	0	2	14	140	55	0	0
4	4	0	1	16	130	55	0	0
5	1	1	1	17	150	1	0	0
5	2	0	2	12	148	55	0	0
5	3	0	1	13	134	55	0	0
5	4	0	1	14	138	55	1	0

```
proc logistic data = bbdm13 descending;
  model fndx = chk agmn wt mod wid nvmr ;
  strata str;
run;
```

```
        The LOGISTIC Procedure

Conditional Analysis

               Model Information

Data Set                          ATS.BBDM13
Response Variable                 fndx
Number of Response Levels         2
Number of Strata                  50
Model                             binary logit
Optimization Technique            Newton-Raphson ridge

           Model Information

final diagnosis

Number of Observations Read         200
Number of Observations Used         200

          Response Profile

 Ordered                      Total
   Value         fndx     Frequency

       1            1            50
       2            0           150

Probability modeled is fndx=1.

               Strata Summary

             fndx
Response     ------     Number of
 Pattern    1    0        Strata     Frequency

       1    1    3            50           200

Newton-Raphson Ridge Optimization

Without Parameter Scaling

         Convergence criterion (GCONV=1E-8) satisfied.

         Model Fit Statistics
```

Criterion	Without Covariates	With Covariates
AIC	138.629	102.430
SC	138.629	122.220
-2 Log L	138.629	90.430

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	48.1998	6	<.0001
Score	39.9247	6	<.0001
Wald	25.2218	6	0.0003

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
chk	1	-1.1218	0.4474	6.2862	0.0122
agmn	1	0.3561	0.1292	7.6013	0.0058
wt	1	-0.0284	0.00998	8.0771	0.0045
mod	1	0.00376	0.0120	0.0984	0.7538
wid	1	-0.4916	0.8173	0.3618	0.5475
nvmr	1	1.4722	0.7582	3.7701	0.0522

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
chk	0.326	0.135	0.783
agmn	1.428	1.108	1.839
wt	0.972	0.953	0.991
mod	1.004	0.980	1.028
wid	0.612	0.123	3.035
nvmr	4.359	0.986	19.264