Birth Weight Part Four: Miscellaneous Examples

Table of contents


```
/* bweight4.sas Try to get a clean LR test */
%include 'bweightread.sas';
title2 'Correct Likelihood Ratio Test in the Presence of Missing Data';

proc logistic;
     title3 'Full model, all the data';
     model low (event=last) = age lwt smoke ptl ht ui ftv r2 r3;

proc logistic;
     title3 'Reduced model, all the data';
     model low (event=last) = age lwt smoke ptl ht ui ftv;

/* G2 = 208.753  - 201.285 = 7.468, df=2 */

proc logistic;
     title3 'Forced stepwise, all the data';
     model low (event=last) = age lwt smoke ptl ht ui ftv r2 r3
           / selection=forward start=7 slentry=1.00  ;

/* Same G2 as above */

data chopped;
     set bigbaby;
     if _n_ < 30 then r2=.;

proc freq; tables r2;

proc logistic;
     title3 'Full model, chopped data';
     model low (event=last) = age lwt smoke ptl ht ui ftv r2 r3;

proc logistic;
     title3 'Reduced model, chopped data (should be same as reduced above)';
     model low (event=last) = age lwt smoke ptl ht ui ftv;

/* G2 = 208.753  - 174.107 = 34.646, df = 2, a disaster  */

proc logistic;
     title3 'Forced stepwise, chopped data';
     model low (event=last) = age lwt smoke ptl ht ui ftv r2 r3
           / selection=forward start=7 slentry=1.00  ;

/* G2 = 181.014 - 174.107 = 6.907, df = 2. That's better! */
```

bweight4.lst Just output from the last proc logistic

Correct Likelihood Ratio Test in the Presence of Missing Data
Forced stepwise, chopped data

The LOGISTIC Procedure

Model Information

Data Set	WORK.CHOPPED	
Response Variable	low	Low Birth Weight
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	189
Number of Observations Used	160

Response Profile

Ordered Value	low	Total Frequency
1	2500 g +	101
2	Under 2500 g	59

Probability modeled is low='Under 2500 g'.

NOTE: 29 observations were deleted due to missing values for the response or explanatory variables.

Forward Selection Procedure

Step 0. The following effects were entered:

Intercept age lwt smoke ptl ht ui ftv

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	212.652	197.014
SC	215.727	221.616
-2 Log L	210.652	181.014

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	29.6375	7	0.0001
Score	27.7530	7	0.0002
Wald	22.5177	7	0.0021

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
6.8536	2	0.0325

Step 1. Effect r2 entered:

Model Convergence Status

Skipping ...

Step 2. Effect r3 entered:

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	212.652	194.107
SC	215.727	224.859
-2 Log L	210.652	174.107

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	36.5448	9	<.0001
Score	33.2599	9	0.0001
Wald	26.6550	9	0.0016

NOTE: All effects have been entered into the model.

Summary of Forward Selection

Step	Effect Entered	DF	Number In	Score Chi-Square	Pr > ChiSq
1	r2	1	8	3.1974	0.0738
2	r3	1	9	3.8088	0.0510

Skipping ...

```
/* bweight5.sas:  Explore ODS */
%include 'bweightread.sas';
title2 'Likelihood ratio test with less output';

/* Starting with SAS Version 9, every table produced by SAS has a name,
   and the Output Delivery System (ODS) allows one to specify what
   is printed. The most convenient way to find out the names in a
   unix/linux environment is with ods trace. */

data chopped;
     set bigbaby;
     if _n_ < 30 then r2=.;

ods trace on / listing;  /* listing option writes trace on list file,
                            rather than log file (default) */
proc logistic;
     title3 'Forced stepwise, chopped data';
     model low (event=last) = age lwt smoke ptl ht ui ftv r2 r3
           / selection=forward start=7 slentry=1.00 covb;
run; /* Need run statement with ods trace */
ods trace off;

ods select Logistic.NObs
           Logistic.Step0.GlobalTests
           Logistic.Step2.GlobalTests;

proc logistic;
     title3 'Reduced Output: Still must use a calculator';
     model low (event=last) = age lwt smoke ptl ht ui ftv r2 r3
           / selection=forward start=7 slentry=1.00;


/* Comments:

      1. It is safer to use Logistic.Step0.GlobalTests etc. rather than
         Logistic.Step0.FitStatistics etc., because the global tests show
         the df. This lets you verify that you're doing the right test.

      2. The names are very systematic, so by counting the number of steps
         required to go from the full to the reduced model, you can use
         ods select with confidence. There is no need to do ods trace
         every time.

      3. It is a great relief to be able to look at only PART of the SAS
         output.

      4. This example barely scratches the surface of what you can do with
         the ods system. For example, you can get SAS to write the list
         file in html format suitable for posting as a Web page, or it
         can write selected tables to SAS data sets for further processing
         within the current SAS job, particularly with proc iml.
```

bweight5.lst

```
------------------------------------------------------------------------------

                          Low Birth Weight Data                              1
                    Likelihood ratio test with less output
                         Forced stepwise, chopped data

                            The LOGISTIC Procedure

Output Added:
-------------
Name:       ModelInfo
Label:      Model Information
Template:   Stat.Logistic.ModelInfo
Path:       Logistic.ModelInfo
-------------
                                Model Information

      Data Set                        WORK.CHOPPED
      Response Variable               low                  Low Birth Weight
      Number of Response Levels       2
      Model                           binary logit
      Optimization Technique          Fisher's scoring

Output Added:
-------------
Name:       NObs
Label:      Observations Summary
Template:   Stat.Logistic.NObs
Path:       Logistic.NObs
-------------


                     Number of Observations Read         189
                     Number of Observations Used         160
```

Skipping ...

```
                             Forward Selection Procedure

Step  0. The following effects were entered:

Intercept  age  lwt  smoke  ptl  ht  ui  ftv
```

Skipping ...

```
Output Added:
-------------
Name:       FitStatistics
Label:      Fit Statistics
Template:   Stat.Logistic.FitStatistics
Path:       Logistic.Step0.FitStatistics
-------------

                               Model Fit Statistics

                                                 Intercept
                                Intercept            and
                     Criterion       Only     Covariates

                     AIC           212.652        197.014
                     SC            215.727        221.616
                     -2 Log L      210.652        181.014
```

```
Output Added:
-------------
Name:       GlobalTests
Label:      Global Tests
Template:   Stat.Logistic.GlobalTests
Path:       Logistic.Step0.GlobalTests
-------------
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	29.6375	7	0.0001
Score	27.7530	7	0.0002
Wald	22.5177	7	0.0021

Skipping ...

Step 1. Effect r2 entered:

Skipping ...

Step 2. Effect r3 entered:

Skipping ...

```
Output Added:
-------------
Name:       GlobalTests
Label:      Global Tests
Template:   Stat.Logistic.GlobalTests
Path:       Logistic.Step2.GlobalTests
-------------
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	36.5448	9	<.0001
Score	33.2599	9	0.0001
Wald	26.6550	9	0.0016

NOTE: All effects have been entered into the model.

Skipping ...

The LOGISTIC Procedure

Number of Observations Read	189
Number of Observations Used	160

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	29.6375	7	0.0001
Score	27.7530	7	0.0002
Wald	22.5177	7	0.0021

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	36.5448	9	<.0001
Score	33.2599	9	0.0001
Wald	26.6550	9	0.0016

(With R)

```
> 36.5448 - 29.6375
[1] 6.9073
```

```
/* bweight6.sas */
%include 'bweightread.sas';
title2 'Do a Wald test "Manually" with proc iml';

ods trace on / listing;  /* listing option writes trace on list file,
                            rather than log file (default) */
proc logistic;
     title3 'The easy way';
     class race /  param = ref;
     model low (event='Under 2500 g') = lwt race / covb;
     race: test raceBlack = raceOther = 0;
run;
ods trace off;

ods listing close; /* No output is sent to the list file */
proc logistic;
     title3 'Make SAS data sets with beta-hat and V-hat';
     ods output ParameterEstimates=MLEdata covb=Vdata;
     class race /  param = ref;
     model low (event='Under 2500 g')= lwt race / covb;
run;

ods listing; /* Turn output to list file back on now. */

/* Take a look at the SAS data sets that have been written
   by proc logistic */
proc print data=MLEdata;
title3 'Parameter Estimates';
proc print data=Vdata;
title3 'Asymptotic Covariance matrix';

proc iml;
     title3 'Proc IML output';
     use MLEdata;
     read all var {estimate} into betahat;
     print betahat;
     use Vdata;
     read all var {Intercept lwt raceBlack raceOther} into vhat;
     print vhat;
     /* Testing H0: C beta = 0 */
     C = {0 0 1 0,
          0 0 0 1};
     W = (C*betahat)` * inv(C*vhat*C`) * (C*betahat);
     pval = 1-probchi(W,2);
     print "Wald Test for race" W pval;
     print "  ";
```

```
% cat bweight6.lst
```

```
                          Low Birth Weight Data                                 1
                  Do a Wald test "Manually" with proc iml
                              The easy way

                         The LOGISTIC Procedure
```

Output Added:

Skipping …

```
                     Linear Hypotheses Testing Results

                                Wald
                Label    Chi-Square      DF    Pr > ChiSq

                race         5.4024       2        0.0671
```

```
                          Low Birth Weight Data                                 2
                  Do a Wald test "Manually" with proc iml
                            Parameter Estimates

                   Class                                                    Prob
Obs    Variable    Val0     DF    Estimate      StdErr     WaldChiSq      ChiSq

 1     Intercept              1      0.8057      0.8452         0.9088     0.3404
 2     lwt                    1     -0.0152     0.00644         5.5886     0.0181
 3     race        Black      1      1.0811      0.4881         4.9065     0.0268
 4     race        Other      1      0.4806      0.3567         1.8156     0.1778
```

```
                          Low Birth Weight Data                                 3
                  Do a Wald test "Manually" with proc iml
                       Asymptotic Covariance matrix

                                                          race        race
      Obs    Parameter    Intercept         lwt         Black       Other

       1     Intercept      0.7143     -0.00521     0.022602     -0.1035
       2     lwt           -0.00521     0.000041    -0.00065     0.000356
       3     raceBlack     0.022602     -0.00065     0.238194       0.0532
       4     raceOther      -0.1035     0.000356       0.0532     0.127216
```

Do a Wald test "Manually" with proc iml
Proc IML output

betahat

0.8057102
-0.015223
1.0810614
0.4806029

vhat

0.7142996	-0.005214	0.0226022	-0.103497
-0.005214	0.0000415	-0.000647	0.0003559
0.0226022	-0.000647	0.238194	0.0532
-0.103497	0.0003559	0.0532	0.1272158

	W	pval
Wald Test for race	5.4023699	0.0671259

Compare earlier output

Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
race	5.4024	2	0.0671

```
/* bweight7.sas */
%include 'bweightread.sas';
title2 'Probit example';

/* Just look at parameter estimates and Z-tests */

ods select ParameterEstimates(persist); /* Persist means keep doing it. */
proc logistic;
     title3 'Logistic Regression, for comparison';
     class race /  param = ref;
     model low (event='Under 2500 g') = lwt race;
     race: test raceBlack = raceOther = 0;

proc logistic;
     title3 'Probit Regression: p = Phi(x-prime beta)';
     class race /  param = ref;
     model low (event='Under 2500 g') = lwt race / link=probit;
     race: test raceBlack = raceOther = 0;
```

Low Birth Weight Data 1
Probit example
Logistic Regression, for comparison

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	0.8057	0.8452	0.9088	0.3404
lwt		1	-0.0152	0.00644	5.5886	0.0181
race	Black	1	1.0811	0.4881	4.9065	0.0268
race	Other	1	0.4806	0.3567	1.8156	0.1778

Low Birth Weight Data 2
Probit example
Probit Regression: p = Phi(x-prime beta)

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	0.4620	0.4938	0.8754	0.3495
lwt		1	-0.00904	0.00370	5.9622	0.0146
race	Black	1	0.6564	0.2959	4.9216	0.0265
race	Other	1	0.2872	0.2151	1.7826	0.1818