

Notes 3-15-01

1

The FREQ Procedure

Table of LOW by SMOKE

LOW(< 2500g)	SMOKE		Total
	0	1	
Frequency			
Percent			
Row Pct			
Col Pct			
0	86	44	130
	45.50	23.28	68.78
	66.15	33.85	
	74.78	59.46	
1	29	30	59
	15.34	15.87	31.22
	49.15	50.85	
	25.22	40.54	
Total	115	74	189
	60.85	39.15	100.00

logistic model low = smoke  
 $\beta_0 =$   
 (smoke)  $\beta_1 = 0.7040$   
 $\exp(0.704) = 2.022$   
 Wald p-value = 0.0276

page 81 Table 3.18 Cross-classification of low birth weight by smoking status stratified by RACE.

```
proc freq data=lowbwt32;
  tables race*low*smoke;
run;
```

The FREQ Procedure

Table 1 of LOW by SMOKE  
 Controlling for RACE=1

LOW(< 2500g)	SMOKE		Total
	0	1	
Frequency			
Percent			
Row Pct			
Col Pct			
0	40	33	73
	41.67	34.38	76.04
	54.79	45.21	
	90.91	63.46	
1	4	19	23
	4.17	19.79	23.96
	17.39	82.61	
	9.09	36.54	
Total	44	52	96
	45.83	54.17	100.00

Race 1  $p(\text{smoke}) = 0.54$

$$\hat{OR} = \frac{(20)(19)}{(33)(4)} \approx 5.75$$

Table 2 of LOW by SMOKE  
 Controlling for RACE=2

LOW(< 2500g) SMOKE

Frequency	Percent	Row Pct	Col Pct	Total
			0   1	
0	11	4		15
	42.31	15.38		57.69
	73.33	26.67		
	68.75	40.00		
1	5	6		11
	19.23	23.08		42.31
	45.45	54.55		
	31.25	60.00		
Total	16	10		26
	61.54	38.46		100.00

The FREQ Procedure

Table 3 of LOW by SMOKE  
Controlling for RACE=3

LOW (< 2500g)	SMOKE		Total
	0	1	
0	35	7	42
	52.24	10.45	62.69
	83.33	16.67	
	63.64	58.33	
1	20	5	25
	29.85	7.46	37.31
	80.00	20.00	
	36.36	41.67	
Total	55	12	67
	82.09	17.91	100.00

```
proc logistic data=lowbwt34 desc;
  model low = smoke / clparm=wald;
run;
```

The LOGISTIC Procedure

Model Information

Data Set	WORK.LOWBWT34
Response Variable	LOW < 2500g
Number of Response Levels	2
Number of Observations	189
Link Function	Logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered	Total
---------	-------

Race 2  
 $p(\text{smoke}) = 0.3846$

$$\hat{OR} = \frac{(11)(6)}{(4)(9)} \approx 1.3$$

$$\hat{OR} = \frac{(35)(5)}{(20)(7)} = 1.25$$

Race 3  
 $p(\text{smoke}) = 0.1791$

Model low = smoke dummy vars race 1 race 2

smoke  $\hat{\beta}_0 = 1.116$   $\exp(1.116) = 3.053$   
 race 1  $\hat{\beta}_1 = 1.0841$   
 race 2  $\hat{\beta}_2 = 1.1086$

Value	LOW	Frequency
1	1	59
2	0	130

3

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

Model Fit Statistics

Criterion	Intercept	Intercept and Covariates
	Only	
AIC	236.672	233.805
SC	239.914	240.288
-2 Log L	234.672	<b>229.805</b>

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.8674	1	0.0274
Score	4.9237	1	0.0265
Wald	4.8516	1	0.0276

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	-1.0870	0.2147	25.6244	<.0001
SMOKE	1	<b>0.7040</b>	0.3196	4.8516	0.0276

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
SMOKE	2.022	1.081	3.783

Association of Predicted Probabilities and Observed Responses

Percent Concordant	33.6	Somers' D	0.170
Percent Discordant	16.6	Gamma	0.338
Percent Tied	49.7	Tau-a	0.073
Pairs	7670	c	0.585

Wald Confidence Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	-1.0870	-1.5078	-0.6661
SMOKE	0.7040	0.0776	1.3305

```
proc logistic data=lowbwt34 desc;
  model low = smoke race2 race3 / clparm=wald;
run;
```

The LOGISTIC Procedure

Model Information

```
Data Set          WORK.LOWBWT34
Response Variable  LOW          < 2500q
```

Number of Response Levels 2  
 Number of Observations 189  
 Link Function Logit  
 Optimization Technique Fisher's scoring

Response Profile

Ordered Value	LOW	Total Frequency
1	1	59
2	0	130

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	236.672	227.975
SC	239.914	240.942
-2 Log L	234.672	219.975

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	14.6973	3	0.0021
Score	14.1265	3	0.0027
Wald	12.8812	3	0.0049

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	-1.8405	0.3529	27.2065	<.0001
SMOKE	1	1.1160	0.3692	9.1357	0.0025
race2	1	1.0841	0.4900	4.8951	0.0269
race3	1	1.1086	0.4003	7.6689	0.0056

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
SMOKE	3.053	1.480	6.294
race2	2.957	1.132	7.725
race3	3.030	1.383	6.640

Association of Predicted Probabilities and Observed Responses

Percent Concordant	54.5	Somers' D	0.299
Percent Discordant	24.6	Gamma	0.378

4  
 $\hat{\beta}_1$  (w/ confounders)  
 $\hat{\beta}_1$  (w/o confounders)  
 | 1.1160 - 0.704 |

0.704

0.412

That is the log odds

changed 41.2% when confounders

were added ... since greater than 15-25% they are important. Also

$H_0: \beta_2 = \beta_3 = 0$

$H_A: \text{Not } H_0$

$LR \chi^2 = 229.805 - 219.975 = 9.83$

$\chi^2_{2df} p = 0.007$

Reject  $H_0$

So these race variables show

strong evidence they

should be in the model

1) They change ln(odds) > 15-25%

2) Also statistically significant.

Percent Tied	20.9	Tau-a	0.129
Pairs	7670	c	0.650

5

Wald Confidence Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	-1.8405	-2.5321	-1.1489
SMOKE	1.1160	0.3923	1.8397
race2	1.0841	0.1237	2.0444
race3	1.1086	0.3240	1.8931

```
proc logistic data=lowbwt34 desc;
  model low = smoke race2 race3 race2sm race3sm / clparm=wald;
run;
quit;
```

The LOGISTIC Procedure

Model Information

Data Set	WORK.LOWBWT34	
Response Variable	LOW	< 2500g
Number of Response Levels	2	
Number of Observations	189	
Link Function	Logit	
Optimization Technique	Fisher's scoring	

Response Profile

Ordered Value	LOW	Total Frequency
1	1	59
2	0	130

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	236.672	228.818
SC	239.914	248.268
-2 Log L	234.672	<b>216.818</b>

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.8542	5	0.0031
Score	15.8649	5	0.0072
Wald	13.1634	5	0.0219

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
>Intercept	1	-2.3026	0.5244	19.2796	<.0001
SMOKE	1	<b>1.7505</b>	0.5983	8.5611	0.0034
race2	1	1.5141	0.7523	4.0511	0.0441
race3	1	1.7430	0.5946	8.5921	0.0034
race2sm	1	-0.5566	1.0322	0.2907	0.5897
race3sm	1	-1.5274	0.8828	2.9933	0.0836

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
SMOKE	5.758	1.782	18.599
race2	4.545	1.041	19.857
race3	5.714	1.782	18.327
race2sm	0.573	0.076	4.334
race3sm	0.217	0.038	1.225

Association of Predicted Probabilities and Observed Responses

Percent Concordant	54.8	Somers' D	0.305
Percent Discordant	24.3	Gamma	0.386
Percent Tied	20.9	Tau-a	0.132
Pairs	7670	c	0.653

Wald Confidence Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	-2.3026	-3.3304	-1.2748
1 SMOKE	1.7505	0.5779	2.9231
2 race2	1.5141	0.0397	2.9885
3 race3	1.7430	0.5775	2.9084
4 race2sm	-0.5566	-2.5797	1.4666
5 race3sm	-1.5274	-3.2577	0.2029

Interaction (Effect Modification)

$$H_0: \beta_4 = \beta_5 = 0$$

$H_a$ : Not  $H_0$

$$LR \chi^2 = 219.975 - 216.975 = 3.00$$

$$p \approx 0.223$$

Do not reject implies not an effect modifier