Stat 5100 Handout #14.c – SAS: Logistic Regression with Polytomous Response

Example: Individuals were surveyed regarding how important they viewed AC and power steering in cars. The sex (M or W), age (1=18-23, 2=24-40, 3=40+), and response (1=little importance, 2=important, 3=very important) of each individual was recorded. The count of responses in each sex/age/response combination was summarized. We want to determine whether and how the sex and age of individuals affects their response.

```
/* Define options */
ods html image_dpi=300 style=journal;


/*************** Example 1 *****************************/
       /* Nominal Logistic Regression */
/*******************************************************/

data car; input sex $ age response count @@; cards;
  W 1 1 26     W 1 2 12     W 1 3  7     W 2 1  9
  W 2 2 21     W 2 3 15     W 3 1  5     W 3 2 14
  W 3 3 41     M 1 1 40     M 1 2 17     M 1 3  8
  M 2 1 17     M 2 2 15     M 2 3 12     M 3 1  8
  M 3 2 15     M 3 3 18
;

/* Note that this is equivalent to:
     data car; input sex $ age response; cards;
       W 1 1
       W 1 1
       ... (26 times)
       W 1 1
       W 1 2
       W 1 2
       ... (12 times)
       W 1 2
       ...
       ...  (all the other categorical combinations)
       ...
       M 3 3
       M 3 3
       ... (18 times)
       M 3 3
       ;
 */
```

```
/* Define dummy variables */
data car; set car;
  S = 1;
  if sex = 'W' then S = 0;
  A2 = 0;
  if age = 2 then A2 = 1;
  A3 = 0;
  if age = 3 then A3 = 1;
run;

/* Run nominal logistic regression */
/* I want response=1 to be last so
   it will be the reference category */
proc sort data=car; by descending response;
proc logistic data=car;
  freq count;
  model response(order=data) = S A2 A3 / link=glogit;
    /* glogit is generalized logit function,
       specifies nominal logistic regression */
  title1 'Nominal Logistic Regression';
run;
```

Nominal Logistic Regression

Response Profile		
Ordered Value	response	Total Frequency
1	3	101
2	2	94
3	1	105

Logits modeled use response=1 as the reference category.

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

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	77.8419	6	<.0001
Score	74.9761	6	<.0001
Wald	62.9703	6	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter	**response**	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	**3**	1	-1.0391	0.3305	9.8843	0.0017
Intercept	**2**	1	-0.5908	0.2840	4.3286	0.0375
S	**3**	1	-0.8129	0.3210	6.4122	0.0113
S	**2**	1	-0.3881	0.3005	1.6677	0.1966
A2	**3**	1	1.4780	0.4009	13.5912	0.0002
A2	**2**	1	1.1283	0.3416	10.9059	0.0010
A3	**3**	1	2.9165	0.4229	47.5594	<.0001
A3	**2**	1	1.5876	0.4029	15.5270	<.0001

Odds Ratio Estimates				
Effect	**response**	**Point Estimate**	**95% Wald Confidence Limits**	
S	**3**	0.444	0.236	0.832
S	**2**	0.678	0.376	1.223
A2	**3**	4.384	1.998	9.620
A2	**2**	3.090	1.582	6.037
A3	**3**	18.477	8.066	42.327
A3	**2**	4.892	2.221	10.775

```
/*************** Example 2 ******************************/
       /* Ordinal Logistic Regression */
/********************************************************/

proc logistic data=car;
  freq count; /* defines weights on each row of data */
  model response(order=data) = S A2 A3;
    /* no link specified, defaults to logit,
       which gives proportional odds model for
       ordinal logistic regression */
  title1 'Ordinal Logistic Regression';
run;
```

Ordinal Logistic Regression

Response Profile		
Ordered Value	**response**	**Total Frequency**
1	3	101
2	2	94
3	1	105

Probabilities modeled are cumulated over the lower Ordered Values.

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

Score Test for the Proportional Odds Assumption		
Chi-Square	**DF**	**Pr > ChiSq**
0.7139	3	0.8699

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	77.2485	3	<.0001
Score	70.0452	3	<.0001
Wald	68.0278	3	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	**3**	1	-1.6546	0.2536	42.5742	<.0001
Intercept	**2**	1	-0.0433	0.2303	0.0354	0.8508
S		1	-0.5762	0.2261	6.4936	0.0108
A2		1	1.1468	0.2773	17.1079	<.0001
A3		1	2.2322	0.2904	59.0806	<.0001

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
S	0.562	0.361	0.875
A2	3.148	1.828	5.421
A3	9.320	5.275	16.467