9.2 Ordinal Response: Proportional Odds Model

9.2.1 Methodology

Consider the arthritis pain data in Table 9.1. Male and female subjects received an active or placebo treatment for their arthritis pain, and the subsequent extent of improvement was recorded as marked, some, or none (Koch and Edwards 1988).

Table 9.1 Arthritis Data

		Improvement			
Sex	Treatment	Marked	Some	None	Total
Female	Active	16	5	6	27
Female	Placebo	6	7	19	32
Male	Active	5	2	7	14
Male	Placebo	1	0	10	11

One possible analysis strategy is to create a dichotomous response variable by combining two of the response categories, basing a model on either Pr{marked improvement} versus Pr{some or no improvement} or Pr{marked or some improvement} versus Pr{no improvement}. However, since there is a natural ordering to these response levels, it makes sense to consider a strategy that takes advantage of this ordering.

Consider the quantities

$$\theta_{hi1} = \pi_{hi1}, \quad \theta_{hi2} = \pi_{hi1} + \pi_{hi2}$$

where π_{hi1} denotes the probability of marked improvement, π_{hi2} denotes the probability of some improvement, and π_{hi3} denotes the probability of no improvement. The $\{\theta_{hij}\}$ represent cumulative probabilities: θ_{hi1} is the probability of marked improvement, and θ_{hi2} is the probability of marked or some improvement (h = 1 for females, h = 2 for males; i = 1 for active treatment, i = 2 for placebo).

For a dichotomous response, you compute a logit function for each subpopulation. For a multilevel response, you create more than one logit function for each subpopulation. With ordinal data, you can compute *cumulative logits*, which are based on the cumulative probabilities. For three response levels, you compute two cumulative logits:

$$\text{logit}(\theta_{hi1}) = \log\left[\frac{\pi_{hi1}}{\pi_{hi2} + \pi_{hi3}}\right], \quad \text{logit}(\theta_{hi2}) = \log\left[\frac{\pi_{hi1} + \pi_{hi2}}{\pi_{hi3}}\right]$$

These cumulative logits are the log odds of marked improvement to none or some improvement and the log odds of marked or some improvement to no improvement, respectively. Both log odds focus on more favorable to less favorable response. The proportional odds model takes both of these odds into account.

Assuming that the data arise from a stratified simple random sample or are at least conceptually representative of a stratified population, they have the following likelihood:

$$\Pr\{n_{hij}\} = \prod_{h=1}^{2}\prod_{i=1}^{2} n_{hi+}! \prod_{j=1}^{3} \frac{{\pi_{hij}}^{n_{hij}}}{n_{hij}!}$$

where

$$\sum_{j=1}^{3} \pi_{hij} = 1$$

You could write a model that applies to both logits simultaneously for each combination of gender and treatment:

$$\text{logit}(\theta_{hik}) = \alpha_k + \mathbf{x}'_{hi}\boldsymbol{\beta}_k$$

where k indexes the two logits. This says that there are separate intercept parameters (α_k) and different sets of regression parameters ($\boldsymbol{\beta}_k$) for each logit.

If you take the difference in logits between two subpopulations for this model, you get

$$\text{logit}(\theta_{hik}) - \text{logit}(\theta_{hi'k}) = (\mathbf{x}_{hi} - \mathbf{x}_{hi'})'\boldsymbol{\beta}_k \text{ for } k = 1, 2$$

Thus, you would need to look at two differences in logits simultaneously to compare the response between two subpopulations. This is the same number of comparisons you would need to compare two subpopulations for a three-level nominal response, for example, in a test for association in a contingency table (that is, $r - 1$, where r is the number of response outcomes). Therefore, this general model doesn't take advantage of the ordinality of the data.

The proportional odds assumption is that $\boldsymbol{\beta}_k = \boldsymbol{\beta}$ for all k, simplifying the model to

$$\text{logit}(\theta_{hik}) = \alpha_k + \mathbf{x}'_{hi}\boldsymbol{\beta}$$

If you take the difference in logits for this model, you obtain the equations

$$\text{logit}(\theta_{hi1}) - \text{logit}(\theta_{hi'1}) = \log\left[\frac{\pi_{hi1}/(\pi_{hi2} + \pi_{hi3})}{\pi_{hi'1}/(\pi_{hi'2} + \pi_{hi'3})}\right] = (\mathbf{x}_{hi} - \mathbf{x}_{hi'})'\boldsymbol{\beta}$$

$$\text{logit}(\theta_{hi2}) - \text{logit}(\theta_{hi'2}) = \log\left[\frac{(\pi_{hi1} + \pi_{hi2})/\pi_{hi3}}{(\pi_{hi'1} + \pi_{hi'2})/\pi_{hi'3}}\right] = (\mathbf{x}_{hi} - \mathbf{x}_{hi'})'\boldsymbol{\beta}$$

This says that the log cumulative odds are proportional to the distance between the explanatory variable values and that the influence of the explanatory variables is independent of the cutpoint for the cumulative logit. In this case, there is a “cut” at marked improvement to form logit(θ_{hi1}) and a cut at some improvement to form logit(θ_{hi2}). This proportionality is what gives the proportional odds model its name. For a single continuous explanatory variable, the regression lines would be parallel to each other, their relative position determined by the values of the intercept parameter.

This model can also be stated as

$$\theta_{hik} = \frac{\exp(\alpha_k + \mathbf{x}'_{hi}\boldsymbol{\beta})}{1 + \exp(\alpha_k + \mathbf{x}'_{hi}\boldsymbol{\beta})}$$

and is written in summation notation as

$$\theta_{hik} = \frac{\exp\{\alpha_k + \sum_{g=1}^{t} \beta_g x_{hig}\}}{1 + \exp\{\alpha_k + \sum_{g=1}^{t} \beta_g x_{hig}\}}$$

where $g = (1, 2, \ldots, t)$ references the explanatory variables. This model is similar to the previous logistic regression models and is also fit with maximum likelihood methods. You can determine the values for π_{hij} from this model by performing the appropriate subtractions of the θ_{hik}.

$$\pi_{hi1} = \theta_{hi1}$$
$$\pi_{hi2} = \theta_{hi2} - \theta_{hi1}$$
$$\pi_{hi3} = 1 - \theta_{hi2}$$

The main effects model is an appropriate starting point for the analysis of the arthritis data. You can write this model in matrix notation as

$$\begin{bmatrix} \text{logit}(\theta_{111}) \\ \text{logit}(\theta_{112}) \\ \text{logit}(\theta_{121}) \\ \text{logit}(\theta_{122}) \\ \text{logit}(\theta_{211}) \\ \text{logit}(\theta_{212}) \\ \text{logit}(\theta_{221}) \\ \text{logit}(\theta_{222}) \end{bmatrix} = \begin{bmatrix} \alpha_1 & & +\beta_1 & +\beta_2 \\ & \alpha_2 & +\beta_1 & +\beta_2 \\ \alpha_1 & & +\beta_1 & \\ & \alpha_2 & +\beta_1 & \\ \alpha_1 & & & +\beta_2 \\ & \alpha_2 & & +\beta_2 \\ \alpha_1 & & & \\ & \alpha_2 & & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

This is very similar to the models described in Chapter 8 except that there are two intercept parameters corresponding to the two cumulative logit functions being modeled for each group. The parameter α_1 is the intercept for the first cumulative logit, α_2 is the intercept for the second cumulative logit, β_1 is an incremental effect for females, and β_2 is an incremental effect for active treatment. Males on placebo constitute the reference cell.

Table 9.2 contains the cell probabilities for marked improvement and no improvement based on this model. Table 9.3 contains the odds. The cell probabilities for marked improvement are based on the model for the first logit function, and the probabilities for no improvement are based on the model for the second logit function (these probabilities are computed from $1 - \theta_{hi2}$). Since the probabilities for all three levels sum to 1, you can determine the cell probabilities for some improvement through subtraction.

The odds ratio for females versus males is e^{β_1}, and the odds ratio for active treatment versus placebo is e^{β_2}. The odds ratios are computed in the same manner as for the logistic regression analysis for a dichotomous response—you form the ratio of the appropriate odds.

Table 9.2 Formulas for Cell Probabilities

		Improvement	
Sex	Treatment	Marked	None
Female	Active	$e^{\alpha_1+\beta_1+\beta_2}/(1+e^{\alpha_1+\beta_1+\beta_2})$	$1/(1+e^{\alpha_2+\beta_1+\beta_2})$
Female	Placebo	$e^{\alpha_1+\beta_1}/(1+e^{\alpha_1+\beta_1})$	$1/(1+e^{\alpha_2+\beta_1})$
Male	Active	$e^{\alpha_1+\beta_2}/(1+e^{\alpha_1+\beta_2})$	$1/(1+e^{\alpha_2+\beta_2})$
Male	Placebo	$e^{\alpha_1}/(1+e^{\alpha_1})$	$1/(1+e^{\alpha_2})$

Table 9.3 Formulas for Model Odds

Sex	Treatment	Improvement: Marked Versus Some or None	Improvement: Marked or Some Versus None
Female	Active	$e^{\alpha_1+\beta_1+\beta_2}$	$e^{\alpha_2+\beta_1+\beta_2}$
Female	Placebo	$e^{\alpha_1+\beta_1}$	$e^{\alpha_2+\beta_1}$
Male	Active	$e^{\alpha_1+\beta_2}$	$e^{\alpha_2+\beta_2}$
Male	Placebo	e^{α_1}	e^{α_2}

For example, when you compare the odds of marked improvement versus some or no improvement for active females versus active males, you obtain

$$\frac{e^{\alpha_1+\beta_1+\beta_2}}{e^{\alpha_1+\beta_2}} = e^{\beta_1}$$

As constrained by the proportional odds model, this is also the odds ratio for marked or some improvement versus no improvement.

9.2.2 Fitting the Proportional Odds Model with PROC LOGISTIC

PROC LOGISTIC fits the proportional odds model by default when the response variable has more than two levels. Thus, you need to ensure that you have an ordinal response variable because PROC LOGISTIC assumes that you do. The GENMOD, GLIMMIX, and PROBIT procedures also fit the proportional odds model with maximum likelihood estimation.

The following SAS statements create the data set ARTHRITIS. Note that these data are in the form of counts, so a variable named COUNT is created to contain the frequencies for each table cell. The variable IMPROVE is a character variable that takes the values marked, some, or none to indicate the subject's extent of improvement of arthritic pain. The variable SEX takes the values male and female, and the variable TREATMENT takes the values active and placebo.

```
data arthritis;
   input sex $ treatment $ improve $ count @@;
   datalines;
female active  marked 16 female active  some 5 female active  none  6
female placebo marked  6 female placebo some 7 female placebo none 19
male   active  marked  5 male   active  some 2 male   active  none  7
male   placebo marked  1 male   placebo some 0 male   placebo none 10
;
```

The use of PROC LOGISTIC is identical to previous invocations for dichotomous response logistic regression. The response variable is listed on the left-hand side of the equal sign and the explanatory variables are listed on the right-hand side. Since the ORDER=DATA option is specified in the PROC statement, the values for IMPROVE are ordered in the sequence in which PROC LOGISTIC encounters them in the data, which is marked, some, and none. (Another legitimate ordering would be none, some, and marked.) It is crucial to ensure that the ordering is correct when you are using ordinal data strategies. The procedure still performs an analysis if the response values are ordered incorrectly, but the results will be erroneous. The burden is on the user to specify the correct order and then to check the results.

The following statements requests that PROC LOGISTIC fit a proportional odds model.

```
proc logistic order=data;
   freq count;
   class treatment sex / param=reference;
   model improve = sex treatment / scale=none aggregate;
run;
```

The "Response Profile" table displayed in Output 9.1 shows that the response variable values are ordered correctly in terms of decreasing improvement. Thus, the cumulative logits modeled are based on more to less improvement. The procedure also prints out a note that a zero count observation has been encountered. For these data, this is not a problem since the total row counts are acceptably large. Computationally, zero counts are discarded. The model still produces predicted values for the cell that corresponds to the zero cell, males on placebo who showed some improvement.

Output 9.1 Response Profiles

Response Profile		
Ordered Value	improve	**Total Frequency**
1	marked	28
2	some	14
3	none	42

Probabilities modeled are cumulated over the lower Ordered Values.

The procedure next prints the "Class Level Information" table, which shows that the parameterization takes the form of incremental effects for active treatment and females.

Output 9.2 Class Levels

Class Level Information			
Class	**Value**	**Design Variables**	
treatment	**active**		1
	placebo		0
sex	**female**		1
	male	0	

Next, PROC LOGISTIC prints out a test for the appropriateness of the proportional odds assumption. The test performed is a score test that determines whether, if you fit a different set of explanatory variable parameters $\boldsymbol{\beta}_k$ for each logit function, those sets of parameters are equivalent. Thus, the model considered is

$$\text{logit}(\theta_{hik}) = \alpha_k + \mathbf{x}'_{hi}\boldsymbol{\beta}_k$$

The hypothesis tested is that there is a common parameter vector $\boldsymbol{\beta}$ instead of distinct $\boldsymbol{\beta}_k$. The hypothesis can be stated as $\boldsymbol{\beta}_k = \boldsymbol{\beta}$ for all k. Thus, if you reject the null hypothesis, you reject the assumption of proportional odds and you need to consider a different approach. If the null hypothesis is not rejected, then the test supports the assumption of proportional odds. Since the test is comparing t parameters for the t explanatory variables across $(r - 1)$ logits, where r is the number of response levels, it has $t * (r - 2)$ degrees of freedom.

The sample size requirements for this test are moderately demanding; you need approximately five observations at each outcome at each level of each main effect, or roughly the same sample size as if you were fitting a generalized logit model. Small samples may artificially make the statistic large, meaning that any resulting significance needs to be interpreted cautiously. However, nonsignificant results are always informative.

The partial proportional odds model is an alternative model that can be fit when the proportionality assumption does not hold for all explanatory variables, but there is proportionality for some (Peterson and Harrell, 1990). Section 9.2.4 describes this approach See Koch, Amara, and Singer (1985) for another discussion of this model. When there appears to be no proportionality, you can fit the model with different parameters for each of the cumulative logits.

Output 9.3 displays the score test for the proportional odds assumption.

Output 9.3 Proportional Odds Test

Score Test for the Proportional Odds Assumption		
Chi-Square	**DF**	**Pr > ChiSq**
1.8833	2	0.3900

Q_{RS} takes the value 1.883 with 2 df. This is clearly nonsignificant, and so the assumption of proportional odds is a reasonable one for these data.

Output 9.4 contains the goodness-of-fit statistics. With values of 2.7121 and 1.9099, respectively, and 4 df, Q_L and Q_P support the adequacy of the model. The 4 df come from (3−1)(4−1)−2 = 4.

Output 9.4 Goodness-of-Fit Statistics

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	2.7121	4	0.6780	0.6071
Pearson	1.9099	4	0.4775	0.7523

Number of unique profiles: 4

The tests for assessing model fit through explanatory capability are also supportive of the model; the likelihood ratio test has a value of 19.8865 with 2 df and the score test has a value of 17.8677 with 2 df, as displayed in Output 9.5.

Output 9.5 Global Tests

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	19.8865	2	<.0001
Score	17.8677	2	0.0001
Wald	16.7745	2	0.0002

You can also investigate goodness of fit by performing the score test for a set of additional terms not in the model. In this case, this effect would simply be the treatment × sex interaction. The following code requests that PROC LOGISTIC fit a main effects model and then perform a score test for the other effect listed in the MODEL statement, which is the interaction.

```
proc logistic order=data;
   freq count;
   class sex treatment / param=reference;
   model improve = sex treatment sex*treatment /
                   selection=forward start=2;
run;
```

The score test of interest is labeled “Residual Chi-Square” and is printed after the “Testing Global Null Hypothesis: BETA=0” table; it is displayed in Output 9.6. The value of the test statistic is 0.2801 (1 df since you are testing the addition of one term to the model) with $p = 0.5967$. This indicates that the main effects model is adequate.

Output 9.6 Score Statistic to Evaluate Goodness of Fit

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
0.2801	1	0.5967

An alternative goodness-of-fit test is the difference in the likelihood ratios for the main effects model and the saturated model. Although the output is not displayed here, the difference in these statistics is (150.029 − 149.721) = 0.308. This is also clearly nonsignificant, compared to a chi-square distribution with 1 df. (Again, note that whenever you form a test statistic based on the difference in likelihoods, then the corresponding degrees of freedom are equal to the difference in the number of parameters for the two models.)

Output 9.7 contains the “Type 3 Analysis of Effects” table. Both sex and treatment are influential effects. Since these effects have 1 df each, the tests are the same as printed for the parameter estimates listed in Output 9.8.

Output 9.7 Type 3 Analysis of Effects

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
sex	1	6.2096	0.0127
treatment	1	14.4493	0.0001

Output 9.8 Parameter Estimates

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	marked	1	-2.6671	0.5997	19.7800	<.0001
Intercept	some	1	-1.8127	0.5566	10.6064	0.0011
sex	female	1	1.3187	0.5292	6.2096	0.0127
treatment	active	1	1.7973	0.4728	14.4493	0.0001

Table 9.4 displays the parameter interpretations.

Table 9.4 Parameter Estimates

Parameter	Estimate(SE)	Interpretation
α_1	−2.667(0.600)	log odds of marked improvement versus some or no improvement for males receiving placebo
α_2	−1.813(0.557)	log odds of marked or some improvement versus no improvement for males receiving placebo
β_1	1.319(0.529)	increment for both types of log odds due to female sex
β_2	1.797(0.473)	increment for both types of log odds due to active drug

Females have $e^{1.319} = 3.7$ times higher odds of showing improvement as males, both for marked improvement versus some or no improvement and for marked or some improvement versus no improvement. Those subjects receiving the active drug have $e^{1.8} = 6$ times higher odds of showing improvement as those on placebo, both for marked improvement versus some or no improvement and for some or marked improvement versus no improvement. These odds ratio estimates are displayed in Output 9.9.

Output 9.9 Odds Ratio Estimates

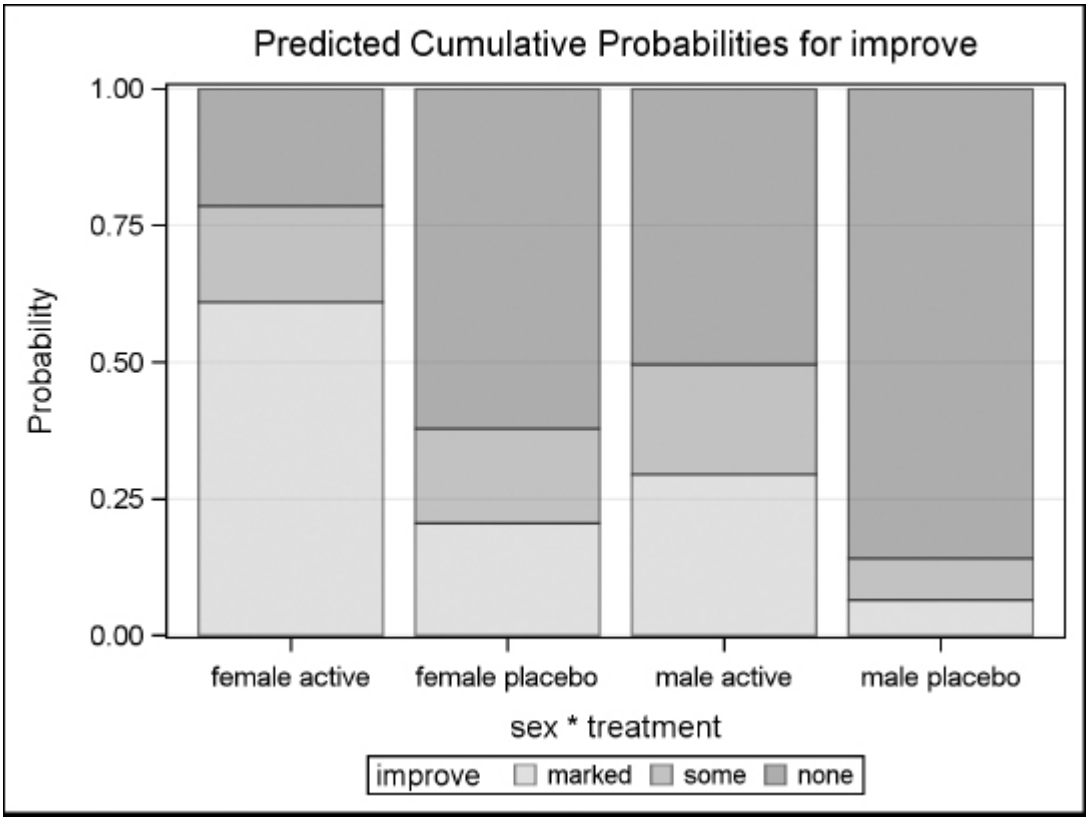
Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
sex female vs male	3.739	1.325	10.547
treatment active vs placebo	6.033	2.388	15.241

A graph of the predicted cumulative probabilities also provides a useful interpretation of the results of this analysis. The PLOTS=EFFECT option, with the POLYBAR and X=TREATMENT*SEX suboptions, specifies this plot from PROC LOGISTIC. The X= suboption specifies the cross-classication of the main effects for which you want to see predicted probabilities.

```
proc logistic order=data plots=effect(polybar x=treatment*sex);
   freq count;
   class sex treatment / param=reference;
   model improve = sex treatment sex*treatment /
                   selection=forward start=2;
run;
```

It's clear from the graph in Output 9.10 that the predicted probabilities of marked improvement and marked or some improvement, are highest for females and active treatment.

Output 9.10 Predicted Cumulative Probabilities



9.2.3 Multiple Qualitative Explanatory Variables

The inclusion of multiple explanatory variables in a proportional odds model produces no additional problems. The data in Table 9.5 are from an epidemiological study of chronic respiratory disease analyzed in Semenya and Koch (1980). Researchers collected information on subjects' exposure to general air pollution, exposure to pollution in their jobs, and whether they smoked. The response measured was chronic respiratory disease status. Subjects were assigned to one of four possible categories.

- Level I: no symptoms
- Level II: cough or phlegm less than three months a year
- Level III: cough or phlegm more than three months a year
- Level IV: cough and phlegm plus shortness of breath more than three months a year

Table 9.5 Chronic Respiratory Disease Data

Air Pollution	Job Exposure	Smoking Status	Response Level I	II	III	IV	Total
Low	No	Non	158	9	5	0	172
Low	No	Ex	167	19	5	3	194
Low	No	Current	307	102	83	68	560
Low	Yes	Non	26	5	5	1	37
Low	Yes	Ex	38	12	4	4	58
Low	Yes	Current	94	48	46	60	248
High	No	Non	94	7	5	1	107
High	No	Ex	67	8	4	3	82
High	No	Current	184	65	33	36	318
High	Yes	Non	32	3	6	1	42
High	Yes	Ex	39	11	4	2	56
High	Yes	Current	77	48	39	51	215

The outcome is clearly ordinal, although there is no obvious distance between adjacent levels. You could combine response categories and fit the set of models that compared Level I versus Level II, III, and IV; Levels I and II versus Levels III and IV; and Levels I, II, and III versus Level IV. Note that if you did this, you would be computing models for the individual cumulative logits. The proportional odds model addresses these cumulative logits simultaneously by assuming that the slope parameters for the explanatory variables are the same regardless of the cumulative logit cutpoints.

From these data, you form three cumulative logits:

$$\text{logit}(\theta_{i1}) = \log\left[\frac{\pi_{i1}}{\pi_{i2} + \pi_{i3} + \pi_{i4}}\right]$$

$$\text{logit}(\theta_{i2}) = \log\left[\frac{\pi_{i1} + \pi_{i2}}{\pi_{i3} + \pi_{i4}}\right]$$

$$\text{logit}(\theta_{i3}) = \log\left[\frac{\pi_{i1} + \pi_{i2} + \pi_{i3}}{\pi_{i4}}\right]$$

where i = 1, 2,..., 12 references the 12 populations determined by the levels of air pollution, job exposure, and smoking status, as ordered in Table 9.5. These cumulative logits are the log odds of a Level I response to a Level II, III, or IV response; the log odds of a Level I or II response to a Level III or IV response; and the log odds of a Level I, II, or III response to a Level IV response, respectively.

However, if you are more interested in the odds of more severe responses to less severe responses, you may want to order the cumulative logits in the opposite direction:

$$\text{logit}(\theta_{i1}) = \log\left[\frac{\pi_{i4}}{\pi_{i3} + \pi_{i2} + \pi_{i1}}\right]$$

$$\text{logit}(\theta_{i2}) = \log\left[\frac{\pi_{i4} + \pi_{i3}}{\pi_{i2} + \pi_{i1}}\right]$$

$$\text{logit}(\theta_{i3}) = \log\left[\frac{\pi_{i4} + \pi_{i3} + \pi_{i2}}{\pi_{i1}}\right]$$

You can generate this ordering in PROC LOGISTIC by using the DESCENDING option in the PROC statement, as shown in the following analysis.

The primary model of interest for these data is a main effects model. Besides three intercept terms α_1, α_2, and α_3 for the three cumulative logits, the main effects model includes the parameters β_1, β_2, β_3 and β_4 for incremental effects for air pollution exposure, job pollution exposure, ex-smoker status, and current smoking status, respectively.

The following SAS statements create the data set RESPIRE.

```
data respire;
   input air $ exposure $ smoking $ level count @@;
   datalines;
low no non   1 158 low  no  non 2   9
low no ex    1 167 low  no  ex  2  19
low no cur   1 307 low  no  cur 2 102
low yes non  1  26 low  yes non 2   5
low yes ex   1  38 low  yes ex  2  12
low yes cur  1  94 low  yes cur 2  48
high no non  1  94 high no  non  2  7
high no ex   1  67 high no  ex   2  8
high no cur  1 184 high no  cur  2 65
high yes non 1  32 high yes non 2   3
high yes ex  1  39 high yes ex  2  11
high yes cur 1  77 high yes cur 2  48
low  no  non 3  5  low  no non  4  0
low  no  ex  3  5  low  no ex   4  3
low  no  cur 3 83  low  no cur  4 68
low  yes non 3  5  low  yes non 4  1
low  yes ex  3  4  low  yes ex  4  4
low  yes cur 3 46  low  yes cur 4 60
high no  non 3  5  high no  non 4  1
high no  ex  3  4  high no  ex  4  3
high no  cur 3 33  high no  cur 4 36
high yes non 3  6  high yes non 4  1
high yes ex  3  4  high yes ex  4  2
high yes cur 3 39  high yes cur 4 51
;
```

The following PROC LOGISTIC code requests a main effects proportional odds model. The MODEL statement generates a score statistic for the goodness of fit of the expanded model containing all pairwise interaction terms. The SCALE=NONE and AGGREGATE=(AIR EXPOSURE SMOKING) options request the goodness-of-fit tests based on the 12 subpopulations. The REF='no' option specified for the EXPOSURE variable in the CLASS statement causes no exposure to be the reference level.

```
proc logistic descending;
   freq count;
   class air exposure(ref='no') smoking / param=reference;
   model level = air exposure smoking
         air*exposure air*smoking exposure*smoking /
         selection=forward include=3 scale=none
         aggregate=(air exposure smoking);
run;
```

Output 9.11 shows the internal ordered values that PROC LOGISTIC uses. Since the response variable LEVEL has numeric values, the DESCENDING option causes PROC LOGISTIC to sort the values numerically, then reverses them to form the ordered values.

Output 9.11 Response Profile

Response Profile		
Ordered Value	**level**	**Total Frequency**
1	**4**	230

23	239
32	337
41	1283

Probabilities modeled are cumulated over the lower Ordered Values.

The score test for the proportional odds assumption takes the value Q_{RS} = 12.0745 (p = 0.1479) with 8 df (4(4 − 2)), as shown in Output 9.12. Thus, the proportional odds assumption is not contradicted.

Output 9.12 Test for Proportionality

Score Test for the Proportional Odds Assumption		
Chi-Square	**DF**	**Pr > ChiSq**
12.0745	8	0.1479

The three intercepts and four indicator variables representing the main effects are first entered into the model. The residual chi-square has a value of 2.7220 with 5 df and p = 0.7428, so this measure of goodness of fit suggests that the model-predicted cell proportions are acceptably close to the observed proportions.

Output 9.13 Assessment of Fit

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
2.7220	5	0.7428

Output 9.14 displays the goodness-of-fit statistics. Q_L = 29.9969 and Q_P = 28.0796, both with $(r - 1)(s - 1) - t = 29$ df (r = 4, s = 12, t = 4). Model adequacy is again supported.

Output 9.14 Goodness-of-Fit Statistics

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	29.9969	29	1.0344	0.4142
Pearson	28.0796	29	0.9683	0.5137

Number of unique profiles: 12

The “Type III Analysis of Effects” table displayed in Output 9.15 suggests a strong effect for job pollution exposure but no significant effect for outside air pollution (p = 0.675). The smoking effect is also highly significant.

Output 9.15 Type III Analysis of Effects

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
air	1	0.1758	0.6750
exposure	1	82.0603	<.0001
smoking	2	209.8507	<.0001

The parameter estimates are displayed in Output 9.16.

Output 9.16 Parameter Estimates

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	4	1	-3.8938	0.1779	479.2836	<.0001
Intercept	3	1	-2.9696	0.1693	307.7931	<.0001
Intercept	2	1	-2.0884	0.1633	163.5861	<.0001
air	high	1	-0.0393	0.0937	0.1758	0.6750
exposure	yes	1	0.8648	0.0955	82.0603	<.0001
smoking	cur	1	1.8527	0.1650	126.0383	<.0001
smoking	ex	1	0.4000	0.2019	3.9267	0.0475

The predicted odds ratios illustrate this model's conclusions. Persons with job exposure have $e^{0.8648} = 2.374$ times higher odds of having serious problems to less serious problems compared to persons not exposed on the job. Current smokers have $e^{1.8527} = 6.377$ times higher odds of having serious problems to less serious problems compared to nonsmokers. Both of these odds ratios have been adjusted for the other variables in the model.

Output 9.17 Odds Ratios

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
air high vs low	0.961	0.800	1.155
exposure yes vs no	2.374	1.969	2.863
smoking cur vs non	6.377	4.615	8.812
smoking ex vs non	1.492	1.004	2.216

Note that if you fit the same model without reversing the order of the cumulative logits with the DESCENDING option, you fit an equivalent model. The intercepts will be in the opposite order and have opposite signs; that is, INTERCEP3 will have the value of this model's INTERCEP1 with the opposite sign. The parameters for the effects will have opposite signs, and the odds ratios will be inverted since they would represent the odds of less serious response to more serious response.

9.2.4 Partial Proportional Odds Model

Table 9.6 contains data from a study by a bicycling clothing manufacturer who wanted to assess their test glove. It was designed to combat the hand problems experienced by cyclists dealing with carpal tunnel syndrome. Cyclists wore either the company's standard gel glove or the new test glove for a week's worth of their standard rides. Then they reported whether they experienced major, moderate, or no relief from their usual symptoms (numbness and nerve pain) on the bike.

Table 9.6 Cycling Glove Data

Glove Type	Gender	Relief From Symptoms			Total
		Major	Moderate	None	
Test	Female	12	8	5	25
Test	Male	8	14	15	37
Gel	Female	5	5	9	19
Gel	Male	8	4	20	32

The data layout is similar to that for the arthritis data in Table 9.1: the response is ordinal in nature, and both explanatory variables have two levels. The proportional odds model is a reasonable one to consider.

The following DATA step inputs the cycling glove data.

```
data wrist;
  input glove $ gender $ relief $ count @@;
  datalines;
test  female   major   12 test female    moderate 8  test female   none  5
test  male major    8 test male  moderate 14 test male none 15
gel  female   major    5 gel female    moderate 5  gel female   none  9
gel  male major    8 gel male  moderate 4  gel male none 20
;
```

First, a saturated model is fit (not shown here). The interaction term was nonsignificant, so the following main effects model is fit.

```
proc logistic order=data;
   freq count;
   class glove gender / param=reference order=data;
   model relief= glove gender / scale=none aggregate;
run;
```

Output 9.18 displays the response profiles, listed from major to none, so the two cumulative logits modeled compare major symptom relief to moderate or no relief and major or moderate relief compared to no relief.

Output 9.18 Response Profile

Response Profile	
Ordered Valuerelief	**Total Frequency**
1major	33
2moderate	31
3none	49

Probabilities modeled are cumulated over the lower Ordered Values.

Output 9.19 shows the score test for the proportional odds assumption; it has the value 4.1734, 2 df, for $p = 0.1241$. Strictly speaking, that p-value would not lead you to reject the hypothesis of proportionality at the $\alpha = 0.05$ level, but it can be considered a marginal result. The partial proportional odds model may be appropriate in this situation if the parameters corresponding to one of the explanatory variables are consistent with proportional odds assumption but the parameters corresponding to the other explanatory variable are not consistent with the assumption.

Output 9.19 Proportional Odds Test

Score Test for the Proportional Odds Assumption		
Chi-Square	**DF**	**Pr > ChiSq**
4.1734	2	0.1241

If neither set of parameters is consistent with proportional odds, then you could fit a model for the cumulative logits with different parameters for each cumulative logit as previously discussed:

$$\text{logit}(\theta_{hik}) = \alpha_k + \mathbf{x}'_{hi}\beta_k$$

where k indexes the two logits. This says that there are separate intercept parameters (α_k) and different sets of regression parameters ($\boldsymbol{\beta}_k$) for each cumulative logit.

But if proportionality holds for one set of coefficients, you could write the model as

$$\text{logit}(\theta_{hik}) = \alpha_k + \mathbf{x}'_{hi1}\beta_k + \mathbf{x}'_{hi2}\gamma$$

where $\mathbf{x}_{hi1}$ represents the explanatory variable with unequal slopes, $\mathbf{x}_{hi2}$ represents the explanatory variable with equal slopes, $\boldsymbol{\beta}_k$ represents the regression parameters for the $\mathbf{x}_{hi1}$, and $\boldsymbol{\gamma}$ represents the parameters for the $\mathbf{x}_{hi2}$.

The LOGISTIC procedure fits the general cumulative logits model as well as the partial proportional odds model. One way to determine whether you have partial proportional odds is to fit the general model and then perform contrast tests to see whether an effect's parameters are the same.

The following statements perform this analysis for the cycling glove data. The UNEQUALSLOPES option in the MODEL statement specifies the general model. Since two cumulative logits are being modeled, the model includes two intercept parameters, two parameters for the gender effect, and two parameters for the glove effect—each parameter corresponds to one of the two cumulative logits. The first TEST statement produces a contrast test to assess the equality of the two parameters for gender, and the second TEST statement produces a contrast test to assess the equality of the glove parameters. (You can determine the internal SAS names of the parameter effects by creating an OUTEST= SAS data set in the PROC LOGISTIC statement and printing it—not shown here).

```
proc logistic order=data;
   freq count;
   class glove gender / param=reference order=data;
   model relief = glove gender / link=clogit
                  scale=none aggregate unequalslopes;
   pogender: test genderfemale_major=genderfemale_moderate;
   poglove:  test glovetest_major=glovetest_moderate;
run;
```

Output 9.20 displays the goodness-of-fit statistics for this model, which are adequate.

Output 9.20 Goodness of Fit

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	1.8487	2	0.9243	0.3968
Pearson	1.8365	2	0.9183	0.3992

Number of unique profiles: 4

Output 9.21 shows the results of the Type 3 tests. Both gender and glove type appear to be important factors.

Output 9.21 Type 3 Tests

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
glove	2	8.1812	0.0167
gender	2	5.1230	0.0772

The parameter estimates for the general model are displayed in Output 9.22. The parameter estimates for the gender effect appear to be similar for both the cumulative logit comparing major relief to moderate or no relief and the cumulative logit comparing any relief (major and moderate relief) compared to no relief.

Output 9.22 Parameter Estimates

Analysis of Maximum Likelihood Estimates							
Parameter		relief	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		major	1	-1.3704	0.3799	13.0128	0.0003
Intercept		moderate	1	-0.6480	0.3362	3.7155	0.0539
glove	test	major	1	0.3209	0.4261	0.5672	0.4514
glove	test	moderate	1	1.0855	0.4022	7.2838	0.0070
gender	female	major	1	0.7189	0.4209	2.9177	0.0876
gender	female	moderate	1	0.9067	0.4143	4.7905	0.0286

Output 9.23 contains the results of formal tests of equal slopes for gender and glove type, considered separately. The *p*-value of 0.6307 for the gender effect indicates that the equal slopes assumption is viable, and the *p*-value of 0.0474 for glove type indicates that the equal slopes assumption is not viable.

Output 9.23 Proportional Odds Test

Linear Hypotheses Testing Results				
	Label	Wald Chi-Square	DF	Pr > ChiSq
	pogender	0.2311	1	0.6307
	poglove	3.9326	1	0.0474

The following PROC LOGISTIC statements request a partial proportional odds model, where gender is handled with a single slope for both cumulative logits modeled and glove type is handled with a different parameter for each cumulative logit. This model is specified with the UNEQUALSLOPES=GLOVE option in the MODEL statement.

```
proc logistic order=data;
   freq count;
   class glove gender / param=reference order=data;
   model relief= glove gender / scale=none aggregate unequalslopes=glove;
run;
```

Output 9.24 provides the goodness of fit for the partial proportional odds model. The *p*-values of 0.3531 and 0.3500 for Q_L and Q_P, respectively, indicate a reasonable fit. The degrees of freedom are $(r - 1)(s - 1) - t$.

Output 9.24 Goodness of Fit

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	2.0819	3	0.6940	0.5556
Pearson	2.0994	3	0.6998	0.5520

Number of unique profiles: 4

Output 9.25 displays the Type 3 Analysis of Effects results. Both gender and type of glove remain influential effects; the effect for gender has 1 df since that effect assumes the proportional odds assumption, and the effect for glove has 2 df since proportionality is not assumed (unequal slopes).

Output 9.25 Type 3 Tests

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
glove	2	7.9611	0.0187
gender	1	4.9395	0.0263

Output 9.26 shows the parameter estimates for the partial proportional odds model.

Output 9.26 Parameter Estimates

Analysis of Maximum Likelihood Estimates							
Parameter		relief	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		major	1	-1.4269	0.3669	15.1215	0.0001
Intercept		moderate	1	-0.6123	0.3284	3.4773	0.0622
glove	test	major	1	0.3278	0.4269	0.5895	0.4426
glove	test	moderate	1	1.0672	0.3997	7.1292	0.0076
gender	female		1	0.8180	0.3680	4.9395	0.0263

Output 9.27 displays the odds ratios.

Output 9.27 Odds Ratios

Odds Ratio Estimates				
Effect	relief	Point Estimate	95% Wald Confidence Limits	
glove test vs gel	major	1.388	0.601	3.205
glove test vs gel	moderate	2.907	1.328	6.364
gender female vs male		2.266	1.101	4.661

There is only one odds ratio listed for gender since that effect was handled with just one parameter. Females reported 2.266 times higher odds of relief compared to males, for both major relief compared to moderate or no relief and for any relief compared to no relief. However, note that the Wald confidence interval barely excludes the value 1.

Two odds ratios are reported for glove type. When major relief is compared to moderate or no relief, people with the test glove had 1.33 times higher odds of reporting relief than people with the gel glove; however, this is not a significant result when you consider its 95% Wald confidence interval (0.601, 3.205). When any relief is compared to no relief, people with the test glove had 2.907 times higher odds of reporting relief than people with the gel glove.

To summarize this partial proportional odds model, gender was an influential effect—females reported relief more often than males—which needed to be accounted for in the model and this behavior held up for both cumulative logits. Test glove performed better than the gel glove, but that performance depended on the cumulative logit that was considered.

Peterson and Harrell (1990) described an unconstrained model and a constrained model, in which various constraints, such as linearity, are imposed on the parameters for an effect. The analysis described in this section is what is called an unconstrained model, where proportional odds can be imposed for some effects but not others, and that is the analysis provided by the UNEQUALSLOPES option. (The constrained model may be available in future releases of the LOGISTIC procedure.) See Koch, Amara, and Singer (1985) for other examples of data suitable for partial proportional odds analysis.