Paper 445-2013

Models for Ordinal Response Data

Robin High, University of Nebraska Medical Center, Omaha, NE

Abstract

Ordinal responses are commonly collected for data analysis, yet the computations of results possible with response data having ordered categories with SAS® procedures are not well known. Although the cumulative logit model is readily available with several statistical procedures, other types of ordinal response models can also be evaluated through statements entered into the NLMIXED procedure. This paper introduces several models for ordinal response data including the partial proportional odds, adjacent logit, continuation ratio, and stereotype models along with the statements to estimate them with PROC NLMIXED. The process requires no restructuring of the input data set, as required with SAS procedures that can produce a subset of these models. Interpretation of ordinal logistic regression models depends on the coding of both the response and explanatory data and whether formats are applied. Computation of odds ratios are illustrated with programming statements and the goodness of fit of these models is tested. Implementation of these models assumes a background with generalized linear models and categorical data analysis including maximum likelihood equations and computing odds ratios with binary data.

Introduction

This paper introduces four types of analyses with ordinal data as the response variable and example SAS code to estimate them. The concepts of a analyzing a binary response variable as a function of a categorical explanatory variable with logistic regression are extended to cover situations where the dependent variable has three or more discrete levels that have an inherent order. The proportional odds assumption for ordinal data is relaxed by demonstrating through programming techniques how to compute non-proportional odds ratios.

Ordinal Data

Ordinal data are defined as variables with a relatively small number of discrete values (at least three) having an increasing or decreasing order. Data collected as numbers having a continuous distribution should in general be analyzed with an appropriate statistical model and should not recoded into ordered categories; working with the actual numerical values as the response is virtually always preferable to recoded data in that it reduces unnecessary measurement error and thus potentially increases the power to detect a significant effect. However, situations may arise where response variables cannot realistically or accurately be measured with a number on an interval or ratio scale. Ordinal data consist of two types of ordered responses:

1. Measurements where all subjects start out at the same initial level and may transition from one level to the next ordered level. For example, in biomedical studies, a subject progresses through levels of an ordered response that is difficult or impossible to quantify, such as durations (categorical time to an event), disease severity levels (stages of cancer), or a patient's response following treatment of an illness. Other examples include level of education (highest grade completed or degree received) or level of proficiency achieved in a skill. One trait of the response assumed for this paper is the ordered progression is one-directional without skipping levels.

2. Cross sectional responses of subjects opinions measured on a Likert scale; for example, select an ordered response measuring level of agreement with a statement (disagree to agree), the frequency a subject engages in some specified activity (Never, Sometimes, Often, Always), or preference for a consumer to purchase a product (not at all to extremely likely), among many others.

Although ordinal responses are recommended here to be coded as integers (e.g., sequentially increasing values of 1,2,3,..), the actual numerical values assigned do not imply that a difference between any adjacent pair of responses has a meaningful quantitative interpretation relative to another adjacent pair. Statistical methods designed for the analysis of continuous data based on a normal distribution of the residuals are much too often invoked for the analysis of ordinal data, essentially not considering implications due non-normality or non-constant variance as well as floor and ceiling effects. One goal of this paper is to introduce other methods for modeling ordinal data and SAS code to implement them.

Ordinal data analysis in this paper is defined as statistical methods to evaluate responses that have a well-defined and meaningful order with at least J=3 distinct categories; the estimation process can easily be adapted for more levels and will be illustrated with J=4 response categories for most of the examples. However, working with six or more levels of an ordinal response variable is less common and may increase the difficulty of interpretation,

especially the ability of the chosen model to distinguish between levels. Although the direction of the ordering is arbitrary, methods described here assume an increasing order for the response with the larger values of greater interest. Results are summarized with odds ratios with a logit link which will be the basis for all methods presented (though other links with their respective interpretations are also possible). The purpose of this paper is to explore methods to analyze ordinal response data, in particular how the NLMIXED procedure from SAS/STAT software may be applied for the models described in Chapters 3 and 4 of "Analysis of Ordinal Data, " 2nd Ed. by Agresti (2010).

Ordinal Data Models

Among the ordinal data models illustrated are the proportional odds model, adjacent category logit, and continuation ratio models. They evaluate an ordinal response variable with J levels (J ≥ 3) coded numerically in the positive direction (i.e., the specified ordering proceeds from smallest to largest). Proceeding in the opposite direction (largest to smallest) requires adjustments to the methods shown here. For each respective model, odds are found with the two probabilities, p_A and p_B computed from the J-1 cut-points of the response, j = 1,2,3....J-1:

```
CL: Cumulative Logit  compares  p_A = ∑ p_i  i > j      with p_B = ∑ p_i, i ≤ j
AC: Adjacent Category compares  p_A = p_{j+1}           with p_B = p_j
CR: Continuation Ratio compares p_A = p_j               with p_B = ∑ p_i, i > j
GL: Generalized Logit compares  p_A = p_i  i > j        with p_B = p_j
```

The pattern of these response categories for the corresponding dichotomous conditional probabilities of each model, p_A and p_B, are computed according to the cut-points defined by level j = 1 to J-1. All models compute the odds of response for level A vs level B and compares them across two specified levels of an explanatory variable (continuous or discrete) to form an odds ratio. The CL model includes the entire range of the responses for all J-1 comparisons produces cumulative odds ratios; the AC model only compares adjacent levels and produces local odds ratios; the CR model specifies a level j and makes comparisons with responses having higher levels. The CL and AC models have similar analytical features and often produce similar results (Agresti, 2010, p. 89). The ordering of the ratio of probabilities for the response categories, p_A / p_B as defined above, is assumed in the subsequent examples. For the CL, AC, and GL models the ordering indicates comparisons are to be made as one would apply the descending option for a binary response model produced with PROCs LOGISTIC, GENMOD, or GLIMMIX. The ordering of the response levels for the CR model is different than the other three and will be explained in that section.

The proportional odds models for CL, AC, & CR are all based on the logit link for p_A vs p_B which is:

$$\text{Log odds} = \text{LOG}(p_A / p_B) = \alpha_j + \text{ß}*x \text{ for } j = 1 .. J-1$$

The estimated coefficient for ß leads to an odds ratio for comparisons of the odds across the first level of the categorical explanatory variable with the odds from the reference category. The ordering of the response comparisons associates an estimated parameter ß such that a positive value implies the effect of increasing the explanatory variable is associated with comparing larger vs. smaller values of the response (which also determines how a categorical explanatory variable is interpreted with dummy variables). The intercepts, α_j, are dependent upon the number of responses in the jth response category and are needed to compute predicted probabilities, but usually are not of substantive interest by themselves. It is also important to note that with one estimated coefficient ß, the proportional odds assumptions implies the odds ratio has the same value across all J-1 comparisons. As will be shown, the CL, AL, and CR models can be modified when the assumption of proportional odds across cut-points is not satisfied; the examples demonstrate how either constrained or unconstrained procedures for estimation of ß can be incorporated into the ordinal response models which may improve the fit of the model and result in improved interpretations.

The fourth type of ordinal statistical model is the stereotype (ST) model has a close relationship to the generalized logit (GL) model listed above. Like the GL model, it is not inherently ordinal (although with some data analyses it can be). Perhaps its most important application to ordinal data analysis is to examine response data levels for distinguishability as will be described in the concluding section.

Example Data

As a basic introduction to estimating ordinal response models, the data set chosen for most of the examples has one response variable with four levels and one categorical explanatory variable with two levels. These data are taken from a paper by Ananth and Kleinbaum (1997) concerning patients' responses from a clinical trial for two single-dose, post-operative analgesic drugs.

```
-------------------------------------------------------------------------
|             |             Rating            |      |      |      |      |      |      |
|             |-------------------------------|      |      |      |      |      |      |
|             |      |      |      | VGd/ |      |      |      |      | VGd/ |
|             | Poor | Fair | Good | Excl | All  | Poor | Fair | Good | Excl |
|             |------+------+------+------+------+------+------+------+------|
|             |  N   |  N   |  N   |  N   |  N   |Row % |Row % |Row % |Row % |
|-------------+------+------+------+------+------+------+------+------+------|
|Drug         |      |      |      |      |      |      |      |      |      |
|C15 & C60    |   17 |   18 |   20 |    5 |   60 | 28.3 | 30.0 | 33.3 |  8.3 |
|Z100 & EC4   |   10 |    4 |   13 |   34 |   61 | 16.4 |  6.6 | 21.3 | 55.7 |
-------------------------------------------------------------------------
```

The SAS data set consists of three variables. The independent variable is coded as drug=1 for "C15 & C60" and drug=2 for "Z100 & EC4." The four ordinal ratings are coded numerically in an increasing order of the response as y=1 (Poor), y=2 (Fair), y=3 (Good), and y=4 (VGd/Excl) (the fourth and fifth response levels were combined due to sparse cell counts). The combination of these two variables results in a data set with eight records. With discrete explanatory data, the third variable in the data set contains the counts, positive integers printed in the left half of the table, that indicate the frequency for each combination of drug and response.

With responses coded as single digit integers, the actual text value can be attached with a format. Never code ordinal responses as text, since the default alphabetical listing will likely not be the same as the desired ordering (e.g., the four responses would be sorted as "Fair / Good / Poor / Vgd/Excl" losing their inherent ascending order). When working with SAS procedures and numerical values with formats apply the order=internal or rorder=internal option (or =data if the coded numbers are entered or sorted with an increasing order) on the PROC statement so that the comparative ordering of the responses remains intact for proper interpretation.

SAS Procedures for Analyzing Ordinal Data

Ordinal response data with one or more categorical or numerical explanatory variables may be analyzed with the CATMOD, LOGISTIC, GENMOD, and GLIMMIX procedures (among others) from SAS/STAT software. With a logit link these procedures estimate odds ratios which also have a natural interpretation with ordinal data (similar to working with binary data) and are the output that underlies most of the examples.

With PROCs LOGISTIC, GENMOD, and GLIMMIX two types of models are available by specifying link=clogit for the cumulative logit or link=glogit for the generalized logit on the MODEL statement (both are also available with PROC CATMOD). The cumulative logit (CL) model represents the most general situation, also known as the proportional odds model (POM). At the other extreme, the most specific model is the generalized logit (GL) for nominal data and will usually not retain the ordinal nature of the data. Methods to analyze ordinal response models to compute odds ratios that lie between these two extremes are not well known or readily available; some of these models are available with these procedures, but the data set first needs to be restructured.

The PROC LOGISTIC code to analyze the example data with a cumulative logit (CL) model is:

```
PROC LOGISTIC order=internal descending
CLASS drug / PARAM=ref;
MODEL rsp = drug / link=clogit expb;
FREQ count;
FORMAT drug xx. rsp yy. ;
RUN;
```

The output includes the Maximum Likelihood Estimates:

```
                              Standard      Wald
Parameter             DF     Estimate      Error   Chi-Square  Pr > ChiSq  Exp(Est)

Intercept VGd/Excl    1       0.070       0.2499      0.0786      0.779
Intercept Good        1       1.469       0.2915     25.3998      <.001
Intercept Fair        1       2.420       0.3378     51.3196      <.001
drug      C15 & C60   1      -1.771       0.3625     23.8683      <.001      0.170
```

The odds ratio comparing the two drugs across the three cut-points is the exponent of the estimated coefficient for drug, EXP(-1.771)=0.170 which implies that over all response levels, subjects with drug=1 (C15 & C60) showed less improvement than subjects with drug=2 (Z100 & EC4). The authors used these data to demonstrate various types of ordinal response models with an ascending ordering of the two response probabilities p_B / p_A as defined earlier; these models will also be described here. For the CL model, PROC LOGISTIC also provides a score test for the

3

proportional odds assumption (Chi-square = 14.2 with 2 d.f., P<0.001) which indicates the proportional odds assumption is not appropriate with these data; the reason will be examined later.

PROC NLMIXED STATEMENTS

SAS has several procedures which offer programming capabilities to analyze ordinal data including the NLIN, MODEL, and NLMIXED procedures. Based on the complexity of the models and the number of statements required, the NLMIXED procedure efficiently estimates these ordinal data models with maximum likelihood, and random effects, if present can also be included if the data are clustered. The required PROC NLMIXED statements for the examples which assume a four level response variable include the following:

```
PROC NLMIXED DATA=indat;
PARMS < enter initial estimates for the parameters >;

* Statements for the linear predictors (eta1, eta2, eta3);
* Statements to compute response probabilities (p1, p2, p3, p4);

* Enter one of the Multinomial Likelihood Equations to Maximize;
lk = (p1**(y=4)) * (p2**(y=3)) * (p3**(y=2)) * (p4**(y=1)); * descending;
lk = (p1**(y=1)) * (p2**(y=2)) * (p3**(y=3)) * (p4**(y=4)); * ascending;

* Compute loglikelihood;
IF (lk > 1e-8) then lglk = LOG(lk); else lglk=-1e100; * to avoid log(0);

* Estimate the model;
MODEL y ~ general(lglk);

REPLICATE count; * for categorical explanatory data entered as counts;
ESTIMATE statements for odds ratios or statistical tests on parameter estimates;
ID p1-p4;
PREDICT statement to save predicted probabilities (p1, p2, p3, p4);
RUN;
```

ODS statements can also be entered to print or save results for subsequent computations, data summaries, and graphical displays. Edits required for ordinal response models are made to a subset of these statements including the PARMS statement (to set initial values), the linear predictors eta1, eta2, & eta3 (three values of eta with J=4 response categories, since eta4 is constrained to be 0), and to compute the four probabilities p1, p2, p3, p4, which must sum to 1. ESTIMATE statements written with the parameter estimate names will compute odds ratios. The two types of log likelihood equations illustrate how to specify either the descending or ascending order to the responses. All remaining statements for the various ordinal data models with one two-level categorical explanatory variable (except as noted for the continuation ratio model) will be the same as the proportional odds model below. Note that there are other ways one may encounter NLMIXED code written to maximize log-likelihood equations and compute probabilities; the layout of these statements work well together.

Cumulative Logit Model

The proportional odds (PO) or cumulative logit (CL) model is based on the logit of the dichotomization of J-1 cumulative probabilities across the J response levels:

```
LOG(p_A/p_B)  = LOG[Prob(Y > p_j)/Prob(Y ≤ p_j)]
              = LOG[(π_{j+1} + … + π_J) / (π_1 + … + π_j)]        for j = 1 .. J-1
```

With four levels of an ordinal response variable, the cumulative logit model is fitted with the following complete set of PROC NLMIXED statements to estimate the odds ratio as found with the PROC LOGISTIC code given above. Replicating the results from PROC LOGISTIC provides a convenient starting point to understand how to further adapt the programming statements for the other types of ordinal response models described the following sections.

```
PROC NLMIXED DATA=cltr ;
PARMS _Intrcpt1 -1 _Intrcpt2 0 _Intrcpt3 1 drg .1 ;
*linear predictors;
eta1 = _Intrcpt1 + drg*(drug=1); * drug=2(bottom row) as reference;
eta2 = _Intrcpt2 + drg*(drug=1);
eta3 = _Intrcpt3 + drg*(drug=1);
```

```
* ordered logit model estimates cumulative probabilities;
cp1 = 1 / (1 + exp(-eta1));
cp2 = 1 / (1 + exp(-eta2));
cp3 = 1 / (1 + exp(-eta3));

* compute individual response probabilities by subtraction;
p1 = cp1;   p2 = cp2-cp1;   p3 = cp3-cp2;  p4 = 1-cp3;

* Maximize likelihood equation;
lk = (p1**(y=4)) * (p2**(y=3)) * (p3**(y=2)) * (p4**(y=1)); * descending;
IF (lk > 1e-8) then lglk = log(lk); else lglk=-1e100;

MODEL y ~ general(lglk);
REPLICATE count;
ESTIMATE 'Odds Ratio: C15 & C60 vs Z100 & EC4' EXP(drg);
RUN;
```

The linear predictors, eta1, eta2, and eta3, are written with reference category coding (drug=2 is the reference level) and then associating them with the cumulative probabilities where cp1 < cp2 < cp3; for this reason, initial values on a log scale for the 3 intercepts should be entered with an increasing order from negative to positive in the PARMS statement. This process then computes valid probabilities – p1, p2, p3, p4 – where all values lie between 0 and 1 so that the taking the log of 0 will not produce an error and cause NLMIXED to stop. The descending order of the response variable levels are listed in the likelihood statement, so that p1 is associated with y=4, p2 with y=3, etc.

The cumulative odds ratio from this model with NLMIXED is 0.170 (from the ESTIMATE statement, the same result from PROC LOGISTIC). To observe why the score test from PROC LOGISTIC rejects the proportional odds assumption, compare the individual estimates and the respective odds ratios for the three 2x2 tables formed by collapsing adjacent response values, estimated with PROC LOGISTIC:

```
  Comparison     Estimate    Odds Ratio
1  4 vs 123     -2.628       0.072
2  34 vs 12     -1.548       0.213
3  234 vs 1     -0.701       0.496
```

A proportional odds model implies the values under the columns for both the estimates and odds ratios should be of a similar magnitude. In this example the estimates are nearly linear when plotted against the three comparisons coded as 1, 2, & 3. With at least three comparisons (i.e., four or more response levels) the plotted logits can be approximated by a straight line (constrained). Thus computing different estimates based on the specific comparison is an important component of the modeling process that leads to the partial proportional odds model.

Linear Predictors with Partial Proportional Odds

The NLMIXED procedure is a flexible tool for estimating partial proportional odds models. To do so one learns to write statements for the linear predictors; the process is much like specifying linear regression equations. First, the number of independent linear predictors, eta, is one less than the number of ordinal response levels. All statements generally begin with an intercept. Numerical explanatory data are entered directly into the equations with the variable name and a different yet identifiable coefficient name. Since PROC NLMIXED does not have a CLASS statement, the linear predictors with a categorical explanatory variable may be written with reference coding following the same guidelines applied when categorical data are dummy coded for linear models in PROCs GLM, MIXED, or GLIMMIX. For any type of analysis with NLMIXED, categorical explanatory data must be dummy coded within a previous DATA step, or with an output data set from the GLMMOD procedure, or with programming statements; for ease of illustration the latter approach is illustrated here for the simplest case of one explanatory variable with two levels. Within NLMIXED, classification variables with no missing data may be dummy coded directly with the discrete levels of categorical predictors with a binary test, e.g., (drug=1) = 1 when the record is for drug 1, =0 otherwise. This result is multiplied by an arbitrary coefficient name for drug. For example, as shown above, with one categorical explanatory variable drug with two levels coded 1 or 2 assuming proportional odds:

```
eta1 = _Intrcpt1 + drg*(drug=1); * (drug=1) implies drug=2 is the reference;
eta2 = _Intrcpt2 + drg*(drug=1);
eta3 = _Intrcpt3 + drg*(drug=1);
```

For ordinal response models, partial proportional odds coefficients may be included in two ways by modifying the drug coefficient of these linear predictors. With at least 4 response levels, the constrained version adds a linearly increasing component to the estimated coefficient for the explanatory variable:

```
eta1 = _Intrcpt1 + (drg + 0*d1)*(drug=1); * add a linear trend to the coefficient;
eta2 = _Intrcpt2 + (drg + 1*d1)*(drug=1);
eta3 = _Intrcpt3 + (drg + 2*d1)*(drug=1);
```

The unconstrained version of partial proportional odds model adds J-2 dummy coded variables to the explanatory variable coefficient:

```
eta1 = _Intrcpt1 + (drg + 0*d2 + 0*d3)*(drug=1); * add dummy codes to d2, d3;
eta2 = _Intrcpt2 + (drg + 1*d2 + 0*d3)*(drug=1); * to the linear predictor;
eta3 = _Intrcpt3 + (drg + 0*d2 + 1*d3)*(drug=1);
```

The extension of the partial proportional odds model with one categorical factor having three or more levels is analogous to these statements; dummy-coded variables are entered for each level except the reference. One may also include the constrained (linear) or unconstrained (dummy codes) additions to the linear predictors to one level and leave the other level unaffected. For drug with 3 levels (drug=3 is the reference), comparisons of level 1 may be kept proportional (one coefficient drg1 for drug=1) and for comparisons of drug=2 with drug=3 add to its coefficient drg2 a constrained linear non-proportional fit. With J=4 ordinal response levels, the revised linear predictors are:

```
eta1= _Intrcpt1 + drg1*(drug=1) + (drg2 + 0*dg)*(drug=2); * drug=3 as reference;
eta2= _Intrcpt2 + drg1*(drug=1) + (drg2 + 1*dg)*(drug=2); * drug 2 has constrained;
eta3= _Intrcpt3 + drg1*(drug=1) + (drg2 + 2*dg)*(drug=2); *        linear term ;
```

The process works the same way with multiple categorical explanatory variables. In the next set of linear predictors, both classification variables drug and gender have two levels. For each of the two variables the non-proportionality assumption can be tested separately. For example, one two level variable gender (coded F/M) may be kept proportional and the second variable, drug, given a constrained linear fit:

```
eta1 = _Intrcpt1 + gnd*(gender='F') + (drg + 0*dg)*(drug=1); * Gender=M is reference;
eta2 = _Intrcpt2 + gnd*(gender='F') + (drg + 1*dg)*(drug=1); *  Drug=2 is reference;
eta3 = _Intrcpt3 + gnd*(gender='F') + (drg + 2*dg)*(drug=1);
```

Odds ratios for a variable from any of these linear predictors are found by taking the exponent of the respective parameter estimate and its linear or unconstrained adjustment. For example, given the third equation for eta3 above the modified odds ratio for drug is:

```
ESTIMATE 'A vs B for Drug 1 vs 2' EXP(drg + 2*dg);
```

Model Fit

The constrained (linear) and unconstrained partial proportional odds models computed with these modified linear predictors are nested within the proportional odds model. To determine if the fit is improved by the addition of a linear or categorical modification, the difference in -2 log likelihoods (deviances) provides a statistical test of their equality (provided in the Fitstatistics table from NLMIXED). This difference is distributed as a chi-square random variable under the null hypothesis of no difference with degrees of freedom equal to the difference in the number of parameters between the two models. AIC and BIC fitstatistics can also be compared.

Predicted probabilities of the response for the values of the covariates are also an important component that assists with measuring the goodness of fit of these models. With categorical explanatory data, comparisons of the predicted counts with the actual values can be obtained from an output data set produced with an ID and PREDICT statements.

```
ID p1-p4;
PREDICT p1 out=prdprb(KEEP= rsp drug count p1-p4);
```

For each record the four predicted probabilities p1, p2, p3, p4 listed on the ID statement are saved into the data set prdprb along with identifying variables rsp, drug, and count. The corresponding predicted value for each row is identified by the value of the response variable, rsp, e.g., for rsp=1, its predicted value is under column p4 (with the descending option applied). Predicted counts are then obtained by extracting the predicted probabilities from this data set and multiplying by the respective row total for each drug.

Partial Proportional Odds Model

As mentioned above, the unequal values of the individual logits for aggregated 2x2 tables indicate a partial proportional odds model may be an alternative for these data. Three or more comparisons (i.e., four or more response levels) are required to fit the constrained linear version. The NLMIXED statements required for the Partial

Proportional Odds Model (PPOM) are the same as listed above for the CL model except a linearly increasing component, $(i-1)*d1$, for drug is added to the estimate for drg in the three linear predictors:

```
eta1= _Intrcpt1 + (drg + 0*d1)*(drug=1); * (drug=1) implies drug=2 is the reference;
eta2= _Intrcpt2 + (drg + 1*d1)*(drug=1); * add an ascending linear trend;
eta3= _Intrcpt3 + (drg + 2*d1)*(drug=1); * to the estimate with d1;
```

Odds ratios are found with the coefficient for drug and the linear adjustments entered into ESTIMATE statements:

```
ESTIMATE '1: A 4 v B 123' EXP(drg + 0*d1);
ESTIMATE '2: A 34 v B 12' EXP(drg + 1*d1);
ESTIMATE '3: A 234 v B 1' EXP(drg + 2*d1);
```

Parameter	Estimate	Standard Error	Probt
_Intrcpt1	0.221	0.257	0.414
_Intrcpt2	1.243	0.296	0.003
_Intrcpt3	1.636	0.344	0.0014
drg	-2.533	0.455	<.001
d1	0.922	0.266	0.008

The estimate for drug (drg=-2.533) differs from results by Anathe and Kleinbaum since their computations assume an ascending order of comparisons of the response. The modified logits (Estimate) and odds ratios can be compared to the proportional odds model above.

	Comparison	Estimate	Odds Ratio
1	A 4 v B 123	-2.533	0.079
2	A 34 v B 12	-1.611	0.200
3	A 234 v B 1	-0.690	0.502

In Section 15.13 of Stokes, et. al. (2000) a method with PROC GENMOD is presented that computes the unconstrained Partial Proportional Odds Model by first restructuring the data set to compute a dichotomous variable for the two respective levels of the J-1 comparisons and then including a REPEATED statement (to work with the multiple observations from each subject the restructuring produces; see Appendix for SAS code). With NLMIXED and the original data set, the two unconstrained adjustments to the three linear predictors are added directly to the coefficients for drug rather than the linearly increasing adjustment for a constrained fit:

```
eta1 = _Intrcpt1 + (drg + 0*d2 + 0*d3)*(drug=1); * 0/1 multipliers for d2, d3;
eta2 = _Intrcpt2 + (drg + 1*d2 + 0*d3)*(drug=1); * in each linear predictor;
eta3 = _Intrcpt3 + (drg + 0*d2 + 1*d3)*(drug=1);
```

Compute the odds ratios for each comparison by exponentiating drg and adding the adjustments d2 or d3:

	Comparison	Estimate	Odds Ratio
1	A 4 vs B 123	-2.628	0.072
2	A 34 vs B 12	-1.548	0.213
3	A 234 vs B 1	-0.701	0.496

With an unconstrained model and one two-level categorical explanatory variable, the odds ratios are the same as produced by the analysis of individual 2x2 tables by grouping adjacent columns. A more complex example of running the unconstrained partial proportional odds model with PROC NLMIXED is described in SAS Usage Note 22954 and also is now featured with PROC LOGISTIC (Version 12.1). Chapter 9.2.4 of Stokes, et. al. (2012) describes how to run the unconstrained model with the UNEQUALSLOPES option in the MODEL statement. Details concerning writing NLMIXED statements for the constrained PPO model with four ordered responses and two explanatory variables are analogous to those presented above for one variable.

Adjacent Logit Model

Adjacent logit (AL) models provide comparisons of the J=4 responses for all pairs of adjacent categories, that is, it compares the ratio of two probabilities (Π_{i+1}/Π_i), i=1,..,J-1 across the levels of the explanatory variable (rather than the underlying continuum of all responses). The adjacent-categories logits are defined with conditional probabilities at each of the J-1 cut-points:

```
LOG(p_A/p_B) = Logit[P( Y=i+1 | Y=i or Y=i+1 )]
            = LOG( p_{i+1} / p_i )                      for i = 1 to J-1
```

For two adjacent columns the linear predictors compare the ratios of adjacent cell probabilities from right-to-left (response codes 4,3,2,1) for drugs coded as 1 and 2 (rows).

$$\text{Log}(p_4/p_3) = \alpha_1 + \text{ß}*x$$
$$\text{Log}(p_3/p_2) = \alpha_2 + \text{ß}*x$$
$$\text{Log}(p_2/p_1) = \alpha_3 + \text{ß}*x$$

The NLMIXED statements to estimate the adjacent logit model are the same as above for the proportional odds model, including the PARMS, linear predictors, and ESTIMATE statements, with modifications only needed to compute the individual probabilities of the response categories, p1, p2, p3, p4. The derivation of these probability statements written within NLMIXED is given in the Appendix. To maintain the order for comparisons of the response categories, the statements to compute the four probabilities are entered in a descending order in NLMIXED:

```
p4 = 1 / total;
p3 = EXP(eta3)*p4;
p2 = EXP(eta2)*p3;
p1 = EXP(eta1)*p2;
```

where the denominator, total for p4, necessary for the four probabilities to sum to 1 is:

```
total = 1 + exp(eta3) + exp(eta2+eta3) + exp(eta1+eta2+eta3);
```

The odds ratio estimate from this model is:

```
                    Standard
Parameter     Estimate     Error      Pr > |t|     EXP(Est)

drg            -0.809      0.188        0.003        0.445
```

The estimated coefficient for drug (-0.809) is the log of an odds ratio which is 0.455, again indicating less improvement for the responses of drug=1. This odds ratio applies across the three adjacent 2x2 tables of collapsed predicted values. If the drug coefficient is not assumed to be equal across these three adjacent 2x2 tables, that is, if they are allowed to vary freely (consider how an unconstrained version of the three linear predictors, eta, would be written), the predicted counts equal the actual, the same results produced by the generalized logit model.

Continuation Ratio Model

The continuation ratio (CR) model follows a comparative process different than the previous models discussed. For a specified cut-point, it sequentially estimates the logit of the probability of each level of the response, given the number of responses at that level and higher.

```
LOG(p_A/p_B)  = LOGIT[PROB(Y = i) | j ≥ i ]
            = LOG[ π_i / ( π_{i+1} + … + π_J) ]      for i = 1 .. J-1
```

Continuation logits are relevant when a sequential mechanism determines the value of a categorized outcome, such as survival time (< 6 months, 6 months to 1 year, 1 to 5 years, > 5 years), tumor size, or skill or proficiency level achieved. An important feature of this model is that all subjects begin at the lowest coded level of the response and have the potential to move sequentially forward through each level, not reversing the trend. The odds ratio compares two groups for the occurrence of an event at level j, given that it did not already experience the event. The form of the dichotomous comparison at each stage is analogous to the discrete time hazard ratio $H(t) = P(Y=t) / P(Y \geq t)$ in survival analysis. With J=4 ordered response categories, the model assumes the following sequential comparisons of the response: 1 vs. 2,3,4 / 2 vs. 3,4 / 3 vs. 4.

The mathematical derivation of the CR model as a product of binomials is outlined in the Appendix. Its equivalence to the stage-wise estimation process illustrated in the table below indicates how NLMIXED statements are written to work with the original data set (no data restructuring is necessary) by entering these statements for the four probabilities:

```
p1 =  1 / (1 + exp(-eta1));
p2 = (1 / (1 + exp(-eta2)))*(1 - p1);
p3 = (1 / (1 + exp(-eta3)))*(1 - p1 - p2);
p4 =  1 - (p1 + p2 + p3);
```

To interpret results the same forward directional manner as the hazard ratio from the PHREG procedure with continuous time to event data, the estimation of probabilities with this model proceeds from smallest response to largest, so the ascending order is required in the multinomial likelihood statement:

```
lk = (p1**(y=1)) * (p2**(y=2)) * (p3**(y=3)) * (p4**(y=4)); * ascending;
```

For the example data with proportional odds, NLMIXED gives the coefficient of drg=1.597 with an odds ratio of 4.94.

To compute the CR model with PROC LOGISTIC, binary data analysis procedures can be utilized, but the data set first needs to be restructured with a DATA step as demonstrated in the Appendix (see Allison, 2012, Chapter 6). The result from this new data set is shown in the table below which illustrates well how comparisons proceed with stagewise estimation and the binary ordering assumptions of the response variable. With a four level ordinal response variable, the CR model examines the four ordered responses categorized into two levels at three stages. The restructuring process groups the ordinal data into J-1=3 sets of 2x2 tables where the binary response variable y2=(1/0) at stage j implies testing for the response equal to the jth category compared to a response greater than j.

```
----------------------------------------------
|Rating        |Comparison |     |     |     |
|Comparisons   |-----------|     |     |     |
|at each stage |  1 |  0  | All |  1  |  0  |
|              |----+-----+-----+-----+-----|
|              |  N |  N  |  N  |Row %|Row %|
|------------- +----+-----+-----+-----+-----|
|a 1 vs 234    |    |     |     |     |     |
|  C15 & C60   | 17 |  43 |  60 | 28.3| 71.7|
|  Z100 & EC4  | 10 |  51 |  61 | 16.4| 83.6|
|b 2 vs 34     |    |     |     |     |     |
|  C15 & C60   | 18 |  25 |  43 | 41.9| 58.1|
|  Z100 & EC4  |  4 |  47 |  51 |  7.8| 92.2|
|c 3 vs 4      |    |     |     |     |     |
|  C15 & C60   | 20 |   5 |  25 | 80.0| 20.0|
|  Z100 & EC4  | 13 |  34 |  47 | 27.7| 72.3|
----------------------------------------------
```

Note that in the table from the restructured data given above, the increasingly wider differences in proportions of y2=1 between drugs observed at the three stages makes the CR model suspect. Constrained or unconstrained linear adjustments to the estimate for drug in the linear predictors can also be included in NLMIXED with the same coding of the classification factors for eta as shown for the partial proportional odds (PPO) models. Cole and Anath (2001) utilize unconstrained linear predictors to compute the partial proportional odds for a five level ordinal response.

With the CR model, if one analyzes the same data with the comparisons of the response levels made in the reverse order [i.e., 4 vs 1,2,3 / 3 vs 12 / 2 vs 1] the resulting continuation odds ratio of 0.341 which does not have the equivalent inverse relationship which conveniently exists with the descending option for the CL or AL proportional odds models (i.e., 0.341 is not equal to 1/4.94=0.202). The methodology for implementation of the CR model (with a restructured data set) with PROC LOGISTIC and how to interpret and diagnose it is further described in Chapter 6, Section 7 of Paul Allison's second edition of "Logistic Regression Using SAS" (2012).

Stereotype Model

The stereotype model is inherently related to the generalized logit model. First consider the linear predictors for the generalized logit model with four response levels (J=4) having descending comparisons and three dummy coded explanatory variables (p=3) with two levels each. The three linear predictors have separate $ß_{ik}$ coefficients for each row and each values of x_k, k = 1 to p:

$$Log(\pi_4(\mathbf{x}) / \pi_1(\mathbf{x})) = \alpha_1 + ß_{41}{}^*x_1 + ß_{42}{}^*x_2 + ß_{43}{}^*x_3$$
$$Log(\pi_3(\mathbf{x}) / \pi_1(\mathbf{x})) = \alpha_2 + ß_{31}{}^*x_1 + ß_{32}{}^*x_2 + ß_{33}{}^*x_3$$
$$Log(\pi_2(\mathbf{x}) / \pi_1(\mathbf{x})) = \alpha_3 + ß_{21}{}^*x_1 + ß_{22}{}^*x_2 + ß_{23}{}^*x_3$$

Estimation of the 12 parameters for this model produces a perfect fit (with categorical predictors). Consider how the equations for the 9 $ß_{ik}$ coefficients for the 3 dummy coded explanatory variables, x_1, x_2, and x_3 could be modified to estimate one $ß_k$ coefficient for each x_k across the three linear predictors:

$Log(\pi_4(x) / \pi_1(x)) = \alpha_1 + \Phi_1{}^*(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)$
$Log(\pi_3(x) / \pi_1(x)) = \alpha_2 + \Phi_2{}^*(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)$
$Log(\pi_2(x) / \pi_1(x)) = \alpha_3 + \Phi_3{}^*(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)$

That is, the coefficient β_{ik} of the generalized logit model is replaced by the product of β_k for column k and a row coefficient, Φ_i, as these modified coefficients demonstrate:

$$\Phi_1{}^*\beta_1 \text{ vs } \beta_{41} \quad \Phi_1{}^*\beta_2 \text{ vs } \beta_{42} \quad \Phi_1{}^*\beta_3 \text{ vs } \beta_{43}$$
$$\Phi_2{}^*\beta_1 \text{ vs } \beta_{31} \quad \Phi_2{}^*\beta_2 \text{ vs } \beta_{32} \quad \Phi_2{}^*\beta_3 \text{ vs } \beta_{33}$$
$$\Phi_3{}^*\beta_1 \text{ vs } \beta_{21} \quad \Phi_3{}^*\beta_2 \text{ vs } \beta_{22} \quad \Phi_3{}^*\beta_3 \text{ vs } \beta_{23}$$

This adjusted model requires fewer parameters to estimate. For identifiability, arbitrary boundary values are set for the outer Φ coefficients, e.g., $\Phi_1 =1$ and $\Phi_4 =0$ are the usual choices ($\alpha_4 = 0$ is assumed from a fourth equation for response level 4). In this example, the stereotype model requires the estimation of 8 parameters, 4 fewer than the 12 required for the generalized logit. The model is interpreted by means of a latent common covariate effect which becomes stronger moving sequentially from lowest to highest through the response categories. For the cumulative logit, stereotype, and generalized baseline models, the number of parameters to estimate depends on the number of response categories, J, and the number of explanatory variables p:

```
   Cumulative Logit (CL): (J-1) + p
           Stereotype (ST):  2*J + p - 3
   Generalized Logit (GL): (J-1)*(1+p)
```

This table shows the efficiency of a stereotype model (ST) compared with a cumulative logit (POM) or generalized baseline (GL) for up to 6 response categories and 4 explanatory variables, each with two levels:

```
------------------------------------------------------------
|         |   Number of Explanatory Vars, 2 levels each    |
|         |------------------------------------------------|
|         |     1     |     2     |     3     |     4      |
|         |-----------+-----------+-----------+-----------|
|         |CL |ST |GL |CL |ST |GL |CL |ST |GL |CL |ST |GL |
|--------+---+---+---+---+---+---+---+---+---+---+---+---|
|Response|   |   |   |   |   |   |   |   |   |   |   |   |
|3       |  3|  4|  4|  4|  5|  6|  5|  6|  8|  6|  7| 10|
|4       |  4|  6|  6|  5|  7|  9|  6|  8| 12|  7|  9| 15|
|5       |  5|  8|  8|  6|  9| 12|  7| 10| 16|  8| 11| 20|
|6       |  6| 10| 10|  7| 11| 15|  8| 12| 20|  9| 13| 25|
------------------------------------------------------------
```

The example data which was evaluated through the three types of ordinal response models discussed earlier does not gain any efficiency with the stereotype model because, as the table indicates, with one categorical explanatory variable, the stereotype model and generalized logit models require the same number of parameters, thus making them equivalent (which is also true for one numerical explanatory variable). For an ordinal data analysis to apply with the stereotype model, the estimated values of the inner Φ coefficients need to exhibit a specific ordering, bounded between 1 and 0:

$$\Phi_1 =1 > \Phi_2 > \Phi_3 > \Phi_4 =0$$

With the stereotype model estimation of the individual Φ coefficients should not determine order of response levels; how would a model be interpreted if, say Φ_3 is significantly greater than Φ_2? The following examples demonstrate an important application of the stereotype model with ordinal response data is how it works as a diagnostic tool.

Distinguishability

The values of Φ_j, j=1 to J fixed or estimated from the stereotype model provide a measure of the distinguishability (Anderson, 1984) of the individual response categories. If two or more of the adjacent Φ coefficients estimated in the model are statistically similar to each other, the evidence suggests that the categories are indistinguishable. Equality of adjacent Φ coefficients to the boundary values of 0 and 1 can be tested with ESTIMATE statements in NLMIXED which may lead to a decision to collapse two or more response categories into one. An example of this process is shown with data from the following table comparing an ordinal response with six levels across two levels of a categorical explanatory variable:

```
---------------------------------------------
|Counts  |                y            |     |
|        |-----------------------------|     |
|        | 1 | 2  | 3  | 4  | 5  | 6  | All|
|--------+----+----+----+----+----+----+----|
|Level 1 |  28|  32|  16|  29|  21|  23| 149|
|Level 2 |   6|   8|   6|  25|  13|  18|  76|
---------------------------------------------
```

PROC NLMIXED statements to estimate the stereotype model are very similar to those already presented; the layout of the statements specific to this example are:

```
PARMS theta1 .1 theta2 -.1 theta3 -.1 theta4 .1 theta5 .1
      phi2 .8 phi3 .6 phi4 .4 phi5 .2  drg .1 ;

eta1 = theta1 +    1 * ( drg*(drug=1));
eta2 = theta2 + phi2 * ( drg*(drug=1));
eta3 = theta3 + phi3 * ( drg*(drug=1));
eta4 = theta4 + phi4 * ( drg*(drug=1));
eta5 = theta5 + phi5 * ( drg*(drug=1));
eta6 =      0 +    0 * ( drg*(drug=1));

total = exp(eta1) + exp(eta2) + exp(eta3) + exp(eta4) + exp(eta5) + exp(eta6);

p1 = exp(eta1) / total;
p2 = exp(eta2) / total;
p3 = exp(eta3) / total;
p4 = exp(eta4) / total;
p5 = exp(eta5) / total;
p6 = exp(eta6) / total;
```

The estimates of the coefficients are computed with the descending order of the response:

```
lk = (p1**(y=6))*(p2**(y=5))*(p3**(y=4))*(p4**(y=3))*(p5**(y=2))*(p6**(y=1));
```

```
                         Standard
Parameter    Estimate     Error      Probt
 phi2          0.819      0.343      0.034 *
 phi3          1.075      0.336      0.008 *
 phi4          0.432      0.432      0.34
 phi5          0.119      0.433      0.79
```

Is there evidence the six response categories are indistinguishable? To test whether the estimates of Φ_2 and Φ_3 in the above table are equal to 1, include these two ESTIMATE statements:

```
ESTIMATE 'Phi2=1?' 1-phi2;
ESTIMATE 'Phi3=1?' 1-phi3;
```

The output indicates the estimates of Φ_2 and Φ_3 are close to 1:

```
                  Standard
Label     Estimate    Error    Pr > |t|

Phi2=1?   0.181      0.343      0.60
Phi3=1?  -0.075      0.337      0.82
```

The pvalues for the other two estimated values of Φ_4 and Φ_5 from the parameter estimates table indicate they are both not different than 0. These tests of the individual values of Φ to the boundary conditions of 0 and 1 suggest that the six ordinal response categories are not distinguishable. This implies the adjacent ordered responses may be placed into fewer distinct categories. An alternative analysis would be to consider the six response values grouped into two levels: that is, responses 4,5,6 are evaluated as level 1 and responses 1,2,3 are evaluated as level 0. No recoding is necessary with NLMIXED, as the revised model can be estimated directly with the stereotype code by entering 1 for the Φ coefficients for eta1, eta2, and eta3 and 0 for the Φ coefficients for eta4, eta5, and eta6:

11

```
eta1 = theta1 + 1*(drg*(drug=1));
eta2 = theta2 + 1*(drg*(drug=1));
eta3 = theta3 + 1*(drg*(drug=1));
eta4 = theta4 + 0*(drg*(drug=1));
eta5 = theta5 + 0*(drg*(drug=1));
eta6 =      0 + 0*(drg*(drug=1));
```

The result is the same as binary logistic regression with an estimated coefficient of -1.070 and an odds ratio of 0.343, which is also the result obtained from PROC LOGISTIC when entering a recoded outcome y2=1 for responses 4,5,6 and y2=0 for responses 1,2,3. The estimated coefficient and odds ratio retaining all six response levels in the stereotype model are -1.295 and 0.274 respectively, though the estimated value of phi3=1.075 is greater than 1 and also greater than the estimated phi2=0.819, another compelling reason to simplify the model.

One may also want to consider recoding in order to distinguish between proportional odds, adjacent category, or generalized baseline models. For example, the stereotype model can replicate results for the adjacent categories (AC) model by entering descending integer values for the Φ coefficients in the linear predictors. From the example demonstrated earlier for the adjacent logits model with four response levels they are coded:

```
eta1 = theta1 + 3*drg*(drug=1);
eta2 = theta2 + 2*drg*(drug=1);
eta3 = theta3 + 1*drg*(drug=1);
eta4 =      0 + 0*drg*(drug=1);
```

The output provides the same estimated coefficient and odds ratio when comparing the response from better to worst conditions in the example above with adjacent logits (drg= -0.809, exp(drg)=0.445).

As an illustration of the stereotype model itself as an ordinal data analysis method, the following example data have J=4 responses and p=2 explanatory variables with two levels (coded 1/2) with a total 400 observations.

```
---------------------------------------
|          |            y             |
|          |--------------------------|
|          |  1  |  2  |  3  |  4  |
|----------+-----+-----+-----+-----|
|x1    x2  |     |     |     |     |
|1     1   |  16 | 12  | 24  | 46  |
|      2   |  20 |  6  | 46  | 91  |
|2     1   |  17 |  3  | 15  | 23  |
|      2   |  15 |  4  | 29  | 33  |
---------------------------------------
```

The NLMIXED statements for PARMS and the linear predictors to estimate the stereotype model:

```
PARMS theta1 .1 theta2 .1 theta3 .1  phi2 .3 phi3 .7  b1 .1 b2 .1 ;

eta1 = theta1 +    1*(b1*(x1=1) + b2*(x2=1));
eta2 = theta2 + phi2*(b1*(x1=1) + b2*(x2=1));
eta3 = theta3 + phi3*(b1*(x1=1) + b2*(x2=1));
eta4 =      0 +    0*(b1*(x1=1) + b2*(x2=1));
```

Coefficients for the stereotype model are compared with the proportional odds and generalized logit models:

Parameter		Cumulative Logit	Stereo-type	Generalized Logit
Intercept 1		-0.246	0.953	0.781
Intercept 2		1.043	0.468	0.571
Intercept 3		1.442	-1.003	-1.800
b1	1	0.484	0.589	0.758
	2		0.437	0.325
	3		0.109	0.845
b2	1	-0.378	-0.657	-0.500
	2		-0.488	-0.583
	3		-0.122	0.494
			$\Phi_2 = 0.74$	
			$\Phi_3 = 0.19$	

Notice how the two estimated coefficients b1 and b2 from the CL model lie within the smallest and largest values of the stereotype parameter estimates for each of the two variables. The generalized logit coefficients do not follow an increasing or decreasing pattern, which may be a reason to reject an ordinal data interpretation.

In summary, as a method to analyze ordinal data, the stereotype model is of limited use, though it is interesting to both see how it can function as a diagnostic tool indicating logistic regression or the adjacent category logit models may be relevant choices, given the flexibility of entering specified values for the phi coefficients.

Conclusion

PROC NLMIXED has already proven that it is a flexible procedure to compute results for many types of non-standard statistical models. These examples show how it also offers convenient and perhaps improved ways to estimate several types of ordinal response data models not readily available in other SAS procedures or statistical software.

References

Ananth, Cande V and David G Kleinbaum, (1997) *Regression Models for Ordinal Responses: A Review of Methods and Applications,* International Journal of Epidemiology, Volume 26, No. 6, pp. 1323-1333.

Anderson, J. A. (1984) *Regression and ordered categorical variables*. Journal of the Royal Statistical Society, Series B., 46(1):1-30.

Agresti, Alan. (1984) Analysis of Ordinal Categorical Data, John Wiley & Sons, New York, NY.

Agresti, Alan. (2010) Analysis of Ordinal Categorical Data, Second Ed., John Wiley & Sons, New York, NY.

Allison, Paul D. (2010) *Survival Analysis Using SAS@: A Practical Guide, Second Edition*. Cary, NC: SAS Institute Inc.

Allison, Paul D. (2012) *Logistic Regression Using SAS@: Theory and Application, Second Edition*. Cary, NC: SAS Institute Inc.

Cole, Stephen R and Anath, Cande V. (2001) "Regression models for unconstrained, partially or fully constrained continuation odds ratios" International Journal of Epidemiology; 30: 1379-1383.

Kuss, Oliver, (2006) *On the estimation of the stereotype regression model*, Computational Statistics & Data Analysis 50:1877 – 1890

SAS Usage Note 22954. "The PROC LOGISTIC proportional odds test and how to fit a partial proportional odds model." March 12, 2013. Available at http://support.sas.com/kb/22/954.html

Stokes, Maura E., Charles S. Davis, and Gary G. Koch. (2000) *Categorical Data Analysis Using the SAS@ System, Second Edition*, Cary, NC: SAS Institute Inc.

Stokes, Maura E., Charles S. Davis, and Gary G. Koch. (2012) *Categorical Data Analysis Using the SAS@ System, Third Edition*, Cary, NC: SAS Institute Inc.

Your comments and questions are valued and encouraged. Contact the author at:

Robin High
Statistical Coordinator
College of Public Health
Department of Biostatistics
University of Nebraska Medical Center
984375 Nebraska Medical Center
Omaha, NE 68198-4375
Phone: (402) 559-2929
email: rhigh@unmc.edu

APPENDIX

Partial Proportional Odds Model

The restructured data set for computations of the partial proportional odds model with PROCs GENMOD, LOGISTIC, or GLIMMIX is produced as follows:

```
* Data set with individual records for each response identified by id;

DATA cltrP2;
DROP rsp count i;
INPUT drug rsp count @@;
RETAIN idd 0;
split=1; y2=(rsp IN (4));
        id=idd; DO i = 1 to count; id+1; OUTPUT; END;
split=2; y2=(rsp IN (3,4));
        id=idd; DO i = 1 to count; id+1; OUTPUT; END;
split=3; y2=(rsp IN (2,3,4));
        id=idd; DO i = 1 to count; id+1; OUTPUT; END;
idd+count;
CARDS;
1 1 17  1 2 18  1 3 20  1 4  5
2 1 10  2 2  4  2 3 13  2 4 34
;

PROC SORT DATA=cltrP2;
BY id descending split y2 drug;
RUN;

ODS SELECT GEEEmpPEst;
ODS OUTPUT diffs=dfsGMD;
PROC GENMOD DATA=cltrp2 descending order=data;
CLASS split drug id;
MODEL y2 = split | drug / noint d=binomial link=logit ;
LSMEANS split*drug / diff oddsratio;
REPEATED subject=id / within=split type=un;
RUN;

PROC SORT DATA =dfsGMD; BY split;
PROC PRINT DATA=dfsGMD;
WHERE split = _split;
RUN;
```

Adjacent Logits

To compute the four probabilities p1, p2, p3, p4 with a descending comparison of the response variable y:

```
p1=P(y=4), p2=P(y=3), p3=P(y=2), p4=P(y=1)
```

Working with a descending order of the coded values of the ordinal response, 4,3,2,1, the linear predictors are the logs of the comparisons

```
LOG( P(y=4)/P(y=3)) = LOG(p1/p2) = eta1
LOG( P(y=3)/P(y=2)) = LOG(p2/p3) = eta2
LOG( P(y=2)/P(y=1)) = LOG(p3/p4) = eta3
```

Solve for p1, p2, p3 and in reverse order and substitute:

```
p3 =              p4*EXP(eta3);
p2 = p3*EXP(eta2) = p4*EXP(eta3)*EXP(eta2)
p1 = p2*EXP(eta1) = p4*EXP(eta3)*EXP(eta2)*EXP(eta1)
```

Since p4 = 1 – (p1 + p2 + p3)

```
p4 = 1 - p4[ EXP(eta1 + eta2 + eta3) + EXP(eta2 + eta3) + EXP(eta3) ]
```

14

which reduces to:

```
p4 = 1 / [1 + EXP(eta3) + EXP(eta2+eta3) + EXP(eta1+eta2+eta3)]
   = 1 / total
```

The individual probabilities are computed with these equations in the NLMIXED code.

```
p4 = 1 / total;
p3 = EXP(eta3)*p4;
p2 = EXP(eta2)*p3;
p1 = EXP(eta1)*p2;
```

In addition to working with these individual probabilities, it can be demonstrated that by writing out the conditional probabilities for two adjacent columns such that the two values add to 1 (as the AL model was defined), i.e., p12=p1/(p1+p2), p23=p2/(p2+p3), and p34=p3/(p3+p4) produces the common odds ratio. Enter the respective equations for p1, p2, p3, & p4 above and algebraically solve the ratio of p12/(1-p12), p23/(1-p23), p34/(1-p34) for the two drugs.

Continuation Ratio Model

The continuation ratio model is derived from the probabilities from a multinomial distribution for J ordinal response categories with individual sample sizes defined by $\mathbf{N} = (n_1, n_2, \ldots, n_J)$ probabilities $\mathbf{P} = \{p_1, p_2, \ldots, p_J\}$. For J= 4, the total sample size is $N = n_1 + n_2 + n_3 + n_4$. The probability density function for the multinomial distribution:

$$Pr(Y_1=y_1,\ Y_2=y_2,\ Y_3=y_3,\ Y_4=y_4)\ |\ \mathbf{N},\mathbf{P}) = \frac{N!}{n_1!\ n_2!\ n_3!\ n_4!}\ (p_1^{n_1}) * (p_2^{n_2}) * (p_3^{n_3}) * (p_4^{n_4})$$

can be factored into a product of J-1=3 binomial distributions:

$$Bin(N, n_1, \pi_1) * Bin(N-n_1, n_2, \pi_2) * Bin(N-n_1-n_2, n_3, \pi_3)$$

where $\pi_1 = p_1$
$\pi_2 = p_2/(1-p_1)$
$\pi_3 = p_3/(1-p_1-p_2)$

See Problem 3 in Chapter 6 of Agresti (1984) for more details. The individual response probabilities can be found by solving for the p_j, j=1 to 4, in this order:

$$p_1 = \pi_1$$
$$p_2 = \pi_2 * (1 - p_1)$$
$$p_3 = \pi_3 * (1 - p_1 - p_2)$$
$$p_4 = 1 - (p_1 + p_2 + p_3)$$

For a proportional odds model defined by the statements for eta, a common value of π is to be estimated at each stage. These equations define how to write the statements to compute the individual response category probabilities.

```
p1 =  1 / (1 + exp(-eta1));
p2 = (1 / (1 + exp(-eta2))) * (1 - p1);
p3 = (1 / (1 + exp(-eta3))) * (1 - p1 - p2);
p4 = 1 - (p1 + p2 + p3);
```

Restructuring of the Data Set

When running the continuation ratio model with PROC LOGISTIC, various sources (e.g., Allison, "Logistic Regression' 2012) demonstrate how to restructure the original dataset (modified here with one DATA step).

```
DATA cltrCR;
KEEP drug count stage y y2 ;
INPUT drug y count @@;
stage=1; y2 = (y EQ 1); OUTPUT; IF y EQ 1 then DELETE;
stage=2; y2 = (y EQ 2); OUTPUT; IF y EQ 2 then DELETE;
stage=3; y2 = (y EQ 3); OUTPUT;
CARDS;
1 1 17  1 2 18  1 3 20  1 4  5
2 1 10  2 2  4  2 3 13  2 4 34
;
```

```
PROC SORT DATA=cltrCR;
BY stage drug y;
RUN;
```

With the restructured dataset and the binary response variable (y2=1/0) PROC LOGISTIC produces the same results as NLMIXED (notice the param=GLM coding in the CLASS statement and NOint option on the MODEL statement):

```
ODS SELECT parameterestimates;

PROC LOGISTIC DATA=cltrCR order=internal descending;
CLASS stage drug / param=GLM;
FREQ count;
MODEL y2 = stage drug / NOINT aggregate scale=none expb;
RUN;


Parameter                 DF     Estimate   Odds Ratio

stage   1 vs 2,3,4     1     -2.216
stage   2 vs 3,4       1     -2.084
stage   3 vs 4         1     -0.723
drug                   1      1.597        4.94
```

The parameter estimates for the intercepts for the stages are the same as the NLMIXED continuation ratio model and the odds ratio, assumed to be heterogeneous across the 3 stages, is also 4.94.

An important identity occurs with the CR model when the cloglog link is chosen rather than the clogit. This is the preferred choice when the ordered response categories represent categorical times to an event since it has a connection with the hazard's ratio from the proportional hazards model. In this situation, the beta coefficients from PROC LOGISTIC are comparable to those estimated with the proportional hazards model using PROC PHREG (Allison, "Analysis of Survival Data," Chapter 7, 2010). For a time to event coded as sequentially increasing integers y=1, 2,3,4 or binary coded as demonstrated above with the restructured file (y2=0,1), the estimated coefficient for drug is the same when running PROC LOGISTIC with either the original or the restructured data sets.

```
PROC LOGISTIC DATA=<original data set>;
CLASS drug / param=glm;
FREQ count;
MODEL y = drug / link = cloglog; * y is the original ordinal responses;
RUN;

PROC LOGISTIC DATA=<restructured data set> descending;
CLASS stage drug / param=GLM;
FREQ count;
MODEL y2 = stage drug / NOINT link = cloglog; * y2 is binary;
RUN;
```

Although the example data here do not represent categorical times to event, for illustration both approaches compute a coefficient of 1.345 (the intercepts are necessarily different). Choosing the cloglog link avoids the need to restructure the data file for the CR model, though as shown in this paper, the logit link can be applied with the NLMIXED code and the original data set. [With GENMOD, specify `cumcll` as the link function.] For comparison to survival analysis results, the corresponding code with PHREG with ties=exact is:

```
PROC PHREG DATA =<original data set>;
CLASS drug;
FREQ count;
MODEL rsp = drug / ties=exact;
HAZARDRATIO drug;
run;
```

with parameter estimate=1.335 with hazard ratio of EXP(1.335) = 3.80.