

Generalized Linear Models (Logistic Regression), a Comparison of Available Options and Output from the Catmod, Probit, Logistic, and Genmod Procedures.

Margaret A. Connolly, ShellSTAT Inc., Marion Station, MD

ABSTRACT

A baseline overview for a clinical depression study required many variables to be evaluated relative to a three factor model. Most efficacy items were continuous and were fit by an analysis of variance to compare two patient classifications, two treatment groups, and multiple clinical study sites. Pair-wise interactions were also evaluated in this model. Our problem was to do corresponding evaluations of binary, ordered categorical and qualitative categorical items. Each of more than 300 items in the psychological profiles was to be summarized by one or two lines in a table. Would a single procedure be suitable for the three types of categorical outcome? Would the procedure(s) produce stable maximum likelihood solutions? How would sparse cells affect the estimates and how could over-specified models be detected in the extract table? Only Genmod results output to tables. Results tables of other procedures must be parsed from the output listing. What diagnostic statistics should be collected, to be viewed during processing or included in the summary tables? Eventually we chose the Catmod procedure for the three types of categorical evaluation. This paper compares the application of Catmod, Probit, Logistic, and Genmod Procedures for the generalized linear model and enumerates some advantages of using the Probit, Logistic, or Genmod procedure for some of the models.

INTRODUCTION

The following section describes generalizations of the linear model for categorical data. We used these models to produce 'analysis of variance' tables for each baseline measure (Basevalu). Summary tables for continuous variables in the study report included only the treatment group means, the analysis of variance mean squared error and degrees of freedom, and statistical significance tests of each factor in the model. The emphasis here is on significance tests, although we recognize that descriptive statistics and data displays tell the real story of a data sample.

The CATMOD procedure is appropriate for fitting generalized linear models for binary, ordered categorical, or qualitative multinomial outcomes. The CATMOD procedure includes these within the framework of loglinear models – for generally measuring association within a multivariate categorical context. Workers in clinical trials, used to describing independent and outcome variables, may not be comfortable with the symmetric associations implied by the loglinear context. The CATMOD procedure has many options and the syntax is quite different from that of the GLM or REG procedures. For these reasons, a user who is familiar with linear models may wish to begin by duplicating the results of more intuitive procedures – such as the LOGISTIC or PROBIT – before going on to use the CATMOD procedure to solve more general problems.

THE GENERALIZED LINEAR MODEL

The linear model for continuous outcomes specifies group means as summed deviance effects. The linear expectation of Y_{ij} is $[\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}]$. Statistical significance tests for factors in the model compare the deviations (α , β , or $(\alpha\beta)$) to the residual variation. The logistic regression model tests conditional associations. As the Chi-square test for independence in the two-by-two table compares the observed cell frequency (p_{ij}) with the product of marginal frequencies ($p_{i.} \times p_{.j}$), the logistic model for the k-by-two table allows tests of several factors which cause expected cell frequencies to be unequal to the product of marginals.

Section 2.2 of McCullagh and Nelder and Section 4.2 of Agresti are references for the logistic linear model. The linear expectation of Y_{ij} is p_{ij} and $\log[p_{ij} / (1 - p_{ij})] = [\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}]$.

Models for multinomial outcomes must further generalize the logistic model. For an example three category outcome, the logit probabilities of the first and second categories may be modeled independently and fit in a single procedure: $\log[p^{k_{ij}} / (1 - p^{k_{ij}})] = [\mu^k + \alpha^k_i + \beta^k_j + (\alpha\beta)^k_{ij}]$ for $k=1$ and 2 . There may be justification to impose a parallel structure on the model. For qualitative multinomial outcomes one such model is for logits of adjacent categories: $\log[p^{k_{ij}} / (p^{(k-1)}_{ij})]$. For ordered categories the cumulative logistic compares higher scored categories to lower scored: $\log[p^{(2+3)}_{ij} / (p^{(1)}_{ij})]$ and $\log[p^{(3)}_{ij} / (p^{(1+2)}_{ij})]$. Chapter 9 in Agresti has examples of multinomial response models and Chapter 5 of McCullagh and Nelder has example applications of the cumulative logistic model.

FACTORS IN THE MODEL

The independent factors in the model were treatment (A or B), disease classification (simple or secondary), and study site. Patients at each of the eight study sites were classified and then randomized to treatment group. The model, written as a GLM procedure, is

```
proc glm noprint outstat=probset;
  class trt dclass site;
  model basevalu=
    trt dclass site
    trt*dclass trt*site dclass*site
    / ss3;
  lsmeans trt /out=outlsm;
run;
```

Differences among study sites could be expected and some differences among the disease classification groups were also predicted.

THE SAMPLE DATA

Binary and ordinal trinomial scores were derived by applying threshold classifications to a random normal sample.

Table 1. Classification frequencies of 526 patients

Dclass	Treatment A		Treatment B	
	Simple	2°	Simple	2°
Site 1	16	19	17	19
2	22	21	22	21
3	21	15	20	15
4	12	22	12	23
5	18	13	18	12
6	18	19	18	20
7	11	9	11	8
8	13	14	13	14

The sample data set was contrived to represent some significant differences between the disease classification groups, some differences among study sites with respect to the proportions of patients in each classification and also in the characteristics (definitions) of the classification groups. The randomization was successful, so that no significant differences would be found between treatment groups or in the treatment interactions by site or disease classification.

Table 2. PROC GLM Type III significance tests for normal Basevalu

Model Factor	dof	significance	Main Effects
Treatment	1	0.885	0.988
Disease Class	1	0.133	0.045
Site	7	<0.001	<0.001
Trt * Site	1	0.446	
Trt * Dclass	7	0.260	
Dclass * Site	7	0.083	

BINARY / ORDINAL / QUALITATIVE RESPONSE

Three ordinal response categories were created by applying threshold criteria to normal scores represented Tables 1 and 2. The sample frequencies were 189 (36%) in level 1, 232 (44%) in level 2 and 105 (20%) in level 3. The binary classifications distinguished level 1 from the other two. The qualitative response categories (38%, 40%, 22%) were created by randomly setting 15% of the ordinal scores to 1.

The binary scores might represent the numbers of male and female patients or frequencies of a baseline diagnostic variable. The ordinal scores represented a clinical scoring criterion (none, mild, moderate) and the qualitative categories represented status such as marital or personality type.

Although no Basevalu class was rare, each contrived sample included missing cells (among 64) in the four-way cross-tabulation. As shown in Table 1, there were no missing cells in the study design.

MAIN EFFECTS MODELS

Cochran-Mantel-Haenszel

The CMH option of the FREQ procedure provides a simple test of conditional association between the Basevalu and one other factor. The test for treatment association, adjusted for Dclass and Site differences would measure the common association in 16 2-by-2

tables. The table rows would be the Trt within Site by Dclass and the columns Basevalu. The 32 row layout is that underlying the CATMOD analysis. The CATMOD model allows various simultaneous contrasts among the 32 rows, but the Cochran-Mantel-Haenszel statistic tests only the last levels of the cross classification.

Table 3. Three different stratified CMH tests, each applied to binary, ordinal, or qualitative Basevalu scores.

Independent tests	Binary (Row Mean Scores)		Ordinal (Row mean scores)		Qualitative (Gen. Assoc)	
	df	p=	df	p=	df	p=
Treatment	1	0.801	1	0.545	2	0.116
Dis. Class	1	0.041	1	0.084	2	0.241
Site	7	0.001	7	0.001	14	0.001

A reference for some clinical applications of the CMH tests may be found in Chapter 6 of Fleiss.

The PROBIT procedure

The CMH tests are not a simultaneous or hierarchical. They include no contrasts of effects within a general test – of pair-wise differences among the eight Sites, for example. For these options, the PROC PROBIT may be applied to PROC FREQ output:

```
proc freq;
  tables trt * dclass * site * basevalu
    / out=outfrq sparse ;
run;

proc probit data=outfrq
  order=internal;
  weight count;
  class trt dclass site basevalu;
  model basevalu=
    trt dclass site
    / d=logistic lackfit;
run;
```

The LACKFIT option requests goodness of fit tests - the Pearson or L.R. Chi-Square have some of the same utility for comparing models as the mean squared error and the degrees of freedom for error of the PROC GLM

Table 4a. PROBIT procedure global tests for binary Basevalu

Log Likelihood for LOGISTIC -287.04

Goodness-of-fit Tests

Statistic	Value	DF	Prob>Chi-Sq
Pearson Chi-Square	22.458	22	0.433
L.R. Chi-Square	24.309	22	0.331

Response Levels: 2 Number of Covariate Values: 32

Table 4b. PROBIT procedure global tests for ordinal Basevalu

Log Likelihood for LOGISTIC -483.56

Goodness-of-fit Tests

Statistic	Value	DF	Prob>Chi-Sq
Pearson Chi-Square	48.91	53	0.634
L.R. Chi-Square	55.24	53	0.390

Response Levels: 3 Number of Covariate Values: 32

The final page of the PROBIT procedure listing gives global tests of classification effects as well as contrasts among the effects.

Table 5a. PROC PROBIT coefficient estimates and significance tests for the binary Basevalu.

Model Factor	df	significance	Coeff	pr (Coeff)
Treatment	1	0.784	0.056	0.784
Disease Class	1	0.038	0.425	0.038
Site	7	<0.001	1: 2.53	<0.001
			2: 0.64	0.070
			3: -0.67	0.070
			4: -0.10	0.774
			5: 2.55	<0.001
			6: -0.10	0.779
			7: 1.42	0.004
			8: =0	-

Table 5b. PROC PROBIT significance tests for ordinal Basevalu.

Model Factor	df	significance
Treatment	1	0.526
Disease Class	1	0.106
Site	7	<0.001

PROC PROBIT model effects may be classification factors or continuous. The Basevalu dependent variable must also be specified as a classification factor. If Basevalu take two levels then PROC PROBIT finds maximum likelihood solutions for the logistic regression (on dummy variable for the class factors) and if Basevalu take more than two then PROC PROBIT finds iterative re-weighted least squares solutions for the cumulative logistic.

The dummy coefficients for the classification factors assign the highest numbered class to =0, so the significance levels for the other coefficients are significance tests for contrasts to the last classification. This can give a convenient breakdown for doing treatment comparisons to a control.

The PROBIT procedure model definition does not accept interaction terms – Trt*Dclass, Trt*Site, Dclass*Site. Because of the way main effects are coded, it does not seem to be possible to create interaction coefficients, in a data step, which are linearly independent of the main effects.

A bad feature of the PROBIT procedure is that it is sensitive to the input data sort order. The procedure as written would give nonsense results if the OUTFRQ data set had been created by a line

```
tables basevalu *
      trt * dclass * site
      / out=outfrq sparse ;
```

The clue that the model had failed would be found in the line giving Response Levels and Number of Covariate Values.

If only one test were of interest, of Trt, Dclass, or Site, then the CMH result for binary Basevalu (Table 3, column 1) or the ordinal basevalu (Table 3, column 2) would be very similar to the PROC PROBIT results: Table 5a for the binary or Table 5b for the ordinal.

GENMOD PROCEDURE

The most adaptable procedure for the binary Basevalu is PROC GENMOD. In contrast to the PROBIT, CATMOD, and LOGISTIC procedures, GENMOD produces output tables of the statistical tests.

PROC GENMOD requires that Basevalu be expressed as a proportion of successful trials:

```
proc freq;
  tables trt * dclass * site * basevalu
    / out=outfrq sparse ;
run;
data togen; set outfrq;
  by trt dclass site basevalu;
  retain ntot;
  if first.site then ntot=count;
  else do;
    ntot=ntot+count;
  output;
end;
run;
proc genmod data=togen;
  make 'modfit' out=fitstats;
  make 'type3w' out=sigtests;
  class trt dclass site;
  model count/ntot=
    trt dclass site
    / d=b link=logit type3;
run;
```

The MODFIT data set lists:

Table 6. PROC GENMOD 'modfit' table for the binary Basevalu.

CRITERIO	DF	VALUE	VALUEDF
Deviance	22	24.309	1.105
Scaled Deviance	22	24.309	1.105
Pearson Chi-Square	22	22.458	1.021
Scaled Pearson X2	22	22.458	1.021
Log Likelihood	.	-287.04	.

A comparison of Table 6 with Table 4a confirms that the PROBIT and GENMOD procedures maximized the same likelihood equations.

The GENMOD procedure will calculate likelihood ratio tests of model coefficients and factors as well as the WALD tests:

Table 7. PROC GENMOD 'Type3W' (Wald tests) table and the 'Type3LR' (Likelihood ratio) table for binary Basevalu.

SOURCE	DF	Type3W		Type3LR	
		CHISQ	PVALC	CHISQ	PVALC
Trt	1	0.075	0.784	0.075	0.784
Dclass	1	4.30	0.038	4.33	0.037
Site	7	73.7	<0.001	108.0	<0.001

The Wald tests are based on the ratio of the estimate to its standard error. These tests are asymptotically equivalent but less optimal than the likelihood ratio tests. Tests of coefficient contrasts in PROC PROBIT, PROC LOGISTIC, and PROC CATMOD are all Wald tests.

BINARY OUTCOME

INTERACTION EFFECTS AND ZERO CELLS

The four empty cells in the 32 row by two column table were:

Sites 1, 5, and 7, Trt B, secondary Dclass, Basvalu=1;
Site 7, Trt A, secondary Dclass, Basvalu=1.

CATMOD PROCEDURE

Since Site 7, secondary Dclass, Trt A or B, was entirely basevalu=0, the Catmod Procedure

```
proc catmod data=outfrq;
  weight count;
  model binscore =
    trt site dclass
    trt*site trt*dclass site*dclass
    / ml noiter noprofile;
run;
```

indicated that tests for Site and Site*Dclass were confounded:

Table 8. PROC CATMOD tests for over-specified model

MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE					
Source	DF	Chi-Square	Prob ¹	Prob ²	Prob ³
Intercept	1	129.55	0.000	0.316	0.000
TRT	1	0.81	0.368	0.384	0.674
DCLASS	1	15.72	<0.001	0.775	0.832
SITE	6*	.	.	<0.001	<0.001
TRT*DCLASS	1	0.94	0.332	0.332	0.469
TRT*SITE	7	8.18	0.317	0.373	0.555
DCLASS*SITE	6*	.	.	0.699	0.443
LIKELIHOOD-RATIO				(df 7, 9.06)	(df 7, 6.83)
	9	9.09	0.429	0.248	0.447
1. with WEIGHT COUNT					
2. With WEIGHT XCOUNT (replace zero cells with 0.001)					
3. The weighted least squares solution, addcell=0.5					

The 'fix' of replacing the zero cells with a 0.001 results in the Prob² column of statistical significance results. If the replacement for zero is too small then the model remains over-specified. A more standard fix (Prob³) would be to add 0.5 to all cells and use the weighted least squares approximation to the maximum likelihood solution. This last method is known to be conservative.

LIKELIHOOD RATIO OR WALD SIGNIFICANCE TESTS

The GENMOD procedure applied to the table with zero cells gives no error messages. The Wald type results and likelihood ratio significance tests are:

Table 9. PROC GENMOD tests for over-specified model

Source	Wald type tests			Likelihood Ratio		
	DF	Chi-Sq	Prob	DF	Chi-Sq	Prob
Intercept	1		0.597			
TRT	1		0.384	1		0.378
DCLASS	1		<0.001	1		0.290
SITE	7		<0.001	7		<0.001
TRT*DCLASS	1		0.332	1		0.331
TRT*SITE	7		0.373	7		0.304
DCLASS*SITE	6*		0.606	7		0.048
LIKELIHOOD-RATIO	7	9.09		7	9.09	

The likelihood ratio tests, based on iterative re-fitting, do not indicate special problems resulting from the zero cells. It may be coincidental that this last evaluation of the binary Basevalu reflects the true model for the underlying normal data – a Site by Dclass interaction was imposed but no average Dclass difference.

LOGISTIC PROCEDURE

Emulating the CATMOD procedure model using PROC LOGISTIC requires the creation of dummy variables, because PROC LOGISTIC is a regression procedure. The default CATMOD main effects design matrix is

Factor Levels			Design Matrix							
Trt	Dclass	Site	Trt	Dclass	S1	S2	S3	S4	S6	S7
1	1	1	1	1	1	0	0	0	0	0
1	2	2	1	-1	0	1	0	0	0	0
2	1	8	-1	1	-1	-1	-1	-1	-1	-1

Columns in the design matrix for interaction effects are the products of main effects columns.

Coding this design matrix to apply PROC LOGISTIC to the binary Basevalu:

```
data tologist; set outfrq;
  array sitedum{3,7} s1-s7 txs1-txs7
    dcxs1-dcxs7;

  if trt='A' then trtdum=1;
  else trtdum=-1;
  if dclass=1 then dcdum=1;
  else dcdum=-1;
  txdc=trtdum*dcdum;
  do ii=1 to 7;
    sitedum{1,ii}=0;
  end;
  if site < 8 then sitedum{1,site}=1;
  else do;
    do ii=1 to 7;
      sitedum{1,ii}=-1;
    end;
  end;
  do ii=1 to 7;
    sitedum{2,ii}=
      trtdum*sitedum{1,ii};
    sitedum{3,ii}=
      dcdum*sitedum{1,ii};
  end;
run;
proc logistic descending
  data=tologist;

  freq count;
  model binscore =
    trtdum dcdum s1-s7
    txdc txs1-txs7 dcxs1-dcxs7
    / link=logit aggregate scale=n;
  TRT: test trtdum=0;
  DCLASS: test dcdum=0;
  SITE: test s1=s2=s3=s4=s5=s6=s7=0;
  TRT_SITE: test
    txs1=txs2=txs3=txs4=txs5=txs6=txs7=0;
  DC_SITE: test dcxs1=dcxs2=dcxs3=dcxs4=
    dcxs5=dcxs6=dcxs7=0;
run;
```

This code gives a WARNING, that four cells have zero frequencies and the maximum likelihood may not exist. Alternately the statement WEIGHT XCOUNT; (replace zero cells with 0.001) may be used in place of FREQ COUNT; The Wald tests in the analysis of variance are listed in Table 10.

Table 10. PROC LOGISTIC tests for binary Basevalu with zero cells

Source	with 4 zeros			with 4 cells of 0.001		
	DF	Chi-Sq	Prob	DF	Chi-Sq	Prob
Intercept	1		0.953			0.318
TRT	1		0.759	1		0.384
DCLASS	1		0.979	1		0.774
SITE	7		<0.001	7		<0.001
TRT*DCLASS	1		0.332	1		0.332
TRT*SITE	7		0.373	7		0.373
DCLASS*SITE	7		0.717	7		0.700
LIKELIHOOD-RATIO	7	9.09	0.246	7	9.06	0.248

The LOGISTIC procedure, with small positive weights attached to the zero cells, duplicates the corresponding CATMOD result. Uncertain results in this model applied to this data configuration arose because all of the patients of the Dclass secondary at Site 7 had a common Basvalu. If no three-way classification has zero cells then the adjustment for zero cells is not needed (except for the WLS option of PROC CATMOD) and results are unambiguous.

ORDERED CATEGORICAL

The data configuration had zero cells for:

Trt B, Dclass secondary, Basevalu=3, for Sites 3, 6, and 8;

Trt B, Dclass secondary, Basevalu=1 for Sites 1 and 5; and also for

Trt A, Dclass secondary, Basevalu=3 at Site 3.

PROC CATMOD - WEIGHTED LEAST SQUARES SOLUTION

The CATMOD procedure finds weighted least squares solutions for the cumulative logistic model. The ADDCELL option must be used since the table has zero cells. Alternately 0.001 could be substituted for the zero cells:

```
proc catmod data=outfrq;
  weight count;          (or weight
  xcount;)               xcount;)
  response clogits;
  model ordscore = _response_
    trt site dclass
    trt*site trt*dclass site*dclass
  / wls
    addcell=0.5          ( for weight
    count)
    nodesign noprofile;
run;
```

The XCOUNT version (zero to 0.001) appears to be more sensitive, as may be seen by comparing factor significance tests in Table 11.

For the multinomial Basevalu, the goodness of fit (likelihood ratio) test is a check on factors in the model and also on the adequacy of the cumulative logistic assumption.

Table 11. PROC CATMOD cumulative logistic analysis-of-variance

Source	ADDCELL=0.5			WEIGHT XCOUNT (0 to 0.001)		
	DF	Chi-Sq	Prob	DF	Chi-Sq	Prob
Intercept	1		<0.001	1		<0.001
response	1		<0.001	1		<0.001
TRT	1		0.627	1		0.611
DCLASS	1		0.326	1		0.383
SITE	7		<0.001	7		<0.001
TRT*DCLASS	1		0.742	1		0.774
TRT*SITE	7		0.904	7		0.881
DCLASS*SITE	7		0.107	7		0.084
Residual (Pearson Chi-Sq)	38	24.27	0.959	38	21.54	0.985

PROC LOGISTIC - MAXIMUM LIKELIHOOD SOLUTION

By default PROC LOGISTIC applies the cumulative logistic model to dependent variables with more than two classification levels. The form of the cumulative distribution may be controlled by using the DESCENDING and ORDER= options in the PROC LOGISTIC STATEMENT. Zero cells of the ordinal Basevalu cause a NOTE: and the WEIGHT XCOUNT; statement may be used in place of the FREQ COUNT.

Table 12. PROC LOGISTIC cumulative logistic analysis-of-variance

Source	with zero cells			WEIGHT XCOUNT (0 to 0.001)		
	DF	Chi-Sq	Prob	DF	Chi-Sq	Prob
Intercept1	1		<0.001			<0.001
Intercept2	1		<0.001			<0.001
TRT	1		0.621	1		0.621
DCLASS	1		0.287	1		0.287
SITE	7		<0.001	7		<0.001
TRT*DCLASS	1		0.704	1		0.704
TRT*SITE	7		0.775	7		0.775
DCLASS*SITE	7		0.032	7		0.032
Pearson Chi-Sq	38	30.16	0.814	38	30.14	0.814
LIKELIHOOD-RATIO	38	35.02	0.608	38	34.93	0.612

The maximum likelihood solution of PROC LOGISTIC is more precise than the weighted least squares solution of PROC CATMOD and PROC LOGISTIC requires no adjustment for zero cells.

QUALITATIVE OUTCOME

Zero cells of the multinomial sample were:

Trt B, Dclass secondary, Basevalu=3, for Sites 3 and 8;

Trt B, Dclass secondary, Basevalu=1 for Site 5; and

Trt B, Dclass simple, Basevalu=3 at Site 8.

MAIN EFFECTS MODEL

The general main effects model for the qualitative Basevalu, may be compared with the CMH test of general association in Table 3. The effect of excluding or including _RESPONSE_ in the CATMOD model is represented in Table 13.

```
proc catmod data=outfrq;
  weight count;
  model mulscore = (_response_)
    trt site dclass
    / ml noiter noprofile;
run;
```

Table 13. PROC CATMOD main effects models

Source	General Structure			Imposed parallel (=_RESPONSE_)		
	DF	Chi-Sq	Prob	DF	Chi-Sq	Prob
Intercept	2		<0.001	1		<0.001
response				1		0.588
TRT	2		0.108	1		0.036
DCLASS	2		0.231	1		0.535
SITE	14		<0.001	7		<0.001
LIKELIHOOD-RATIO	44	46.69	0.257	53	100.3	<0.001

Results for the general model are similar to the CMH tests for general association. The parallel ratios constraint shows a poor fit in the main effects model, as evaluated by the likelihood ratio goodness-of-fit tests.

INTERACTION FACTORS AND ZERO CELLS

The general multinomial model with interaction effects is over-specified due to zero cells. Alternate models substitute 0.001 for 0 or impose a parallel constraint.

Table 14. PROC CATMOD for multinomial with interactions

Source	General Structure		General, zero to 0.001		Imposed parallel	
	DF	Prob	DF	Prob	DF	Prob
Intercept	2	<0.001	2	0.077	1	<0.001
response					1	<0.001
TRT	2	0.030	2	0.406	1	0.621
DCLASS	2	0.418	2	0.418	1	0.287
SITE	13*	*	14	<0.001	7	<0.001
TRT*DCLASS	2	0.670	2	0.670	1	0.704
TRT*SITE	14	0.078	14	0.219	7	0.775
DCLASS*SITE	14	0.558	14	0.558	7	0.032
LIKELIHOOD-RATIO	15	0.510	14	0.436	38	<0.001

The imposed parallel model does not appear to fit the contrived data sample. The parallel constraint greatly reduces the number of model parameters but is frequently found not to fit the data. The parallel logistic constraint on the outcome probability ratios becomes more restrictive as more terms are added to the model specification.

CONCLUSION

The method of choice for binomial (grouped binary) responses is the GENMOD procedure. This is the only procedure among the four which calculates likelihood ratio tests for the significance of model coefficients and contrasts. It is also the only one which produces table output of these tests. The advantage of the LOGISTIC procedure for the ordered categorical model is that it calculates maximum likelihood coefficients. The data set for the LOGISTIC procedure must include dummy coefficients for classification factors. If the model includes only main effects, then the PROBIT procedure will give equivalent tests of

Model contrasts as the TEST contrasts of PROC LOGISTIC. The PROBIT procedure can give incorrect results if the data are sorted by response category.

PROC CATMOD is the only procedure appropriate to model probabilities of qualitative response categories. The k-1 independent probabilities of the k-multinomial may be fit independently or a restriction on the pair-wise probabilities may be imposed. PROC CATMOD treats zero cells as structural zeros and gives a message of model over-specification. The goodness-of-fit test for the restricted model tests model adequacy with respect to the model factors and the restrictions on response categories.

ACKNOWLEDGMENTS

This work was part of the statistical analysis of a study by Pfizer, Inc. to compare the effectiveness of Zoloft® with a tricyclic antidepressant for treatment of patients with chronic depression or depression and dysthymia.

REFERENCES

- Agresti, Alan (1990), *Categorical Data Analysis*, New York: John Wiley & Sons, Inc.
- Fleiss, Joseph L. (1986), *The Design and Analysis of Clinical Experiments*, New York: John Wiley & Sons, Inc.
- McCullagh, P. and Nelder, J. A. (1983) *Generalized Linear Models*, New York: Chapman and Hall.

Margaret A. Connolly, ShellSTAT Inc.
3431 Wms. Pt. Rd., 21838
(410) 957-1011
Shellstt@shore.intercom.net