Abstract

The macros presented in this paper use output from PROC MLOGIT to generate (currently only for the polytomous logit model): 1) all pairwise combinations of the effects of explanatory variables on the relative probabilities of all outcomes, and 2) the predicted probabilities of each of the dependent variable outcomes, evaluated at the means of the explanatory variables, or any user-specified values.

Introduction

The MLOGIT procedure calculates parameter estimates of the effects of each of the k explanatory variables on the logarithms of p-1 probabilities of outcomes relative to the probability of the pth outcome. The coefficients on all pairwise combinations of relative probabilities may be calculated directly as the differences between the coefficients on the p-1 relative probabilities. Similarly, the standard errors for each of these coefficients may be calculated from the variance-covariance matrix of the MLOGIT output as:

$\text{SE}(\beta_i - \beta_j) = \sqrt{\text{SE}(\beta_i)^2 + \text{SE}(\beta_j)^2 - 2 \cdot \text{cov}(\beta_i, \beta_j)}$

These direct calculations, however, can get cumbersome with a large number of outcomes or explanatory variables. The PERMUTE Macro performs these calculations automatically for up to approximately 179 coefficients.

It is also frequently desirable to calculate the predicted probabilities of each outcome, evaluated at a user-specified set of values for the explanatory variables. The macros PCARDS and PROB permit the user to calculate these probabilities evaluated at default values of the means of each variable or any user-specified values based on the following equations:

$$\Pr(Y_i) = \frac{\exp[X \beta_i]}{\sum_{j=1}^{p-1} \exp[X \beta_j]} \quad i = 1, \ldots, p-1$$

$$\Pr(Y_p) = \frac{1}{\sum_{j=1}^{p-1} \exp[X \beta_j]}$$

where $\Pr(Y_i)$ is the predicted probability of outcome $Y_i$, $X$ is the user-specified $1 \times k$ vector of explanatory variable values and $\beta_i$ is the $i$th $k \times 1$ vector of calculated coefficients from the MLOGIT output.

Outline of the Macros

The PERMUTE macro (see Figure 4) reads the raw output from PROC MLOGIT line by line as character variables of length 133. (See Stone, 1985, for a detailed example of the method.) It scans through these lines, searching for combinations of keywords which indicate necessary data and saving these as SAS variables. Some effort has gone into ensuring that only the pertinent data are read in; occasional error messages in the PROC MLOGIT output generally will not interrupt this scan. However, if stray characters are read in as pertinent data, the number of variables read in will be thrown off and the program will stop, indicating this with an error message. In this case, it may be necessary to manually edit the MLOGIT output, removing all stray lines.

After the raw output is read, the SAS variables are fetched into a PROC MATRIX routine in which the full set of coefficients and standard errors are calculated and finally printed out.

If requested, the PCARDS and PROB macros use the parameter estimates to calculate a set of predicted probabilities for each value of the dependent variable based on the means of the explanatory variables as well as an additional set of probabilities for every set of user-specified values.

Requirements

The macros must be included in your program before they are invoked. This is done via a %INCLUDE statement shown below. The MLOGIT raw output must also exist as an OS dataset before the macros are invoked. One can use PROC PRINTTO, for example, to direct the output to an OS file. Use of PROC PRINTTO permits the user to invoke the PERMUTE macro in the same job that runs PROC MLOGIT. Figure 1 illustrates this.

Alternatively, assuming that the MLOGIT output (which must include the variance-covariance matrix) already exists, the macros may be invoked using the following required [and optional] statements. Note that the 'pgmfile' DD statement names the program library containing the macros.
//jobname JOB account.number.name
// EXEC SAS, OPTIONS='DQUOTE'
/pgmfile DD DSN=mlogit.proglib,DISP=SHR
/in DD DSN=mlogit.raw.output,DISP=SHR
/SYSIN DD *
%INCLUDE pgmfile(PERMUTE);
%PERMUTE(P=p,K=k,INDD=in,STAT=TRAT)
%PCARDS;
[1 vlname=vlvaluel v2name=v2value1 ... /
  vnname=vnvaluel]
[2 vlname=vlvaluel v2name=v2value2 ... /
  vnname=vnvaluel]
[ q vlname=vlvaluel ... ];

%PROB;

where

pgmfile = DD name for the OS partitioned dataset containing the PERMUTE macro,

p = number of choices for the dependent variable specified in PROC MLOGIT

k = number of explanatory variables specified in PROC MLOGIT

in = DD name for the OS file containing the raw MLOGIT output (defaults to 'IN').

STAT = TRAT [optional] invoked if T-ratios are desired, rather than standard errors.

Limitations

The probability calculations generated by %PROB may use incorrect mean values for the explanatory variables when P, the number of outcomes, exceeds 5. Consequently, when P exceeds 5, it may be necessary to explicitly define the mean values of all explanatory variables. To determine if explicit definition is necessary, users may check the MLOGIT output to determine whether the column labelled 'MEANS FOR ALL' are the true means, rather than one of the sets of conditional means.

Explanation of Figures

Figure 1 shows a job in which the macros PERMUTE, PCARDS, and PROB are invoked following PROC MLOGIT. The output from PROC MLOGIT is written to a temporary data set (DD name FT20FP001). In the call to PERMUTE, the number of categories of the dependent variable is specified as 3 and the number of independent variables as 8. The output from the macros, shown in Figure 3, (1) repeats the summary statistics of the MLOGIT output (K, P, N, ...); (2) repeats the (P-1)*K set of coefficients and standard errors and reports the remaining coefficients and standard errors; and (3) reports the predicted probabilities of the alternative outcomes PROB (y=1), PROB (y=2), PROB (y=3)) while the explanatory variables are evaluated at the means (MEANS) and at the values specified in MODEL 1.

In this example age (AGEYRS), geographic location (CURNORTH, CURMID, CURWEST) and educational achievement (EDOT8, ED911, ED1315, ED16) are used to estimate the relative probabilities of outcomes equal to 1, 2 or 3. In MODEL 1 the predicted outcome probabilities are calculated for an observation with age=30, location=north and education=0 to 8.

Figure 2 shows the output from PROC MLOGIT which is the input to the PERMUTE macro. Figure 3 is the output from the PERMUTE and PROB macros and Figure 4 is the source listing of the macros.

Acknowledgments

The authors wish to acknowledge the support given by the Carolina Population Center for this programming project. It was done in part as the programming support on projects funded by grants NICHD-1-R23-RD-15521-01 and NICHD-1-R01-RD18101 from the National Institute of Child Health and Human Development with Amy O. Taut and Ronald R. Rindfuss, respectively, as principal investigators.
Figure 2. OUTPUT FROM PROC MLOGIT (INPUT TO THE PERMUTE MACRO)

(Variance-covariance matrix is not shown.)

ESTIMATION OF THE POLYTOMOUS LOGIT MODEL

THE NUMBER OF X VARIABLES IS 9 AND THE NUMBER OF S VARIABLES IS 9

IMPLEMENTING A TOTAL NUMBER OF PARAMETERS TO BE ESTIMATED = 18

THE METHOD IS NEWTON-RAPHSON WITH THE FOLLOWING ITERATION PARAMETERS:

CONVERGENCE ON EACH PARAMETER = 0.000100

CONVERGENCE ON THE SUMS OF SQUARES = 0.000010

MAXIMUM NUMBER OF ITERATIONS = 20

NUMBER OF CASES IN EACH CHOICE CLASS:

1 1453
2 907
3 605

2965 CASES CONSISTING OF 8895 CHOICES WERE READ.

INDEPENDENT VARIABLE MEANS

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MEAN FOR D = 1</th>
<th>MEAN FOR D = 2</th>
<th>MEAN FOR D = 3</th>
<th>MEAN FOR ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>1.0000000</td>
<td>1.0000000</td>
<td>1.0000000</td>
<td>1.0000000</td>
</tr>
<tr>
<td>AGED</td>
<td>27.523056</td>
<td>31.910695</td>
<td>32.328926</td>
<td>31.529935</td>
</tr>
<tr>
<td>CURNORTH</td>
<td>0.15141087</td>
<td>0.26240353</td>
<td>0.2132314</td>
<td>0.23124954</td>
</tr>
<tr>
<td>CURWEST</td>
<td>0.16173434</td>
<td>0.12679152</td>
<td>0.12046161</td>
<td>0.12612644</td>
</tr>
<tr>
<td>EDOT</td>
<td>0.399174120-01</td>
<td>0.529217200-01</td>
<td>0.11074380</td>
<td>0.39577860-01</td>
</tr>
<tr>
<td>ED911</td>
<td>0.16517550</td>
<td>0.10474090</td>
<td>0.17685950</td>
<td>0.15982068</td>
</tr>
<tr>
<td>ED1315</td>
<td>0.16104611</td>
<td>0.18652816</td>
<td>0.15196669</td>
<td>0.15853295</td>
</tr>
</tbody>
</table>

CONVERGENCE ACHIEVED.

NUMBER OF ITERATIONS = 5

LOG LIKELIHOOD AT CONVERGENCE IS -2875.1273

CHI-SQUARED = 394.28341 WITH 16 DEGREES OF FREEDOM.

<table>
<thead>
<tr>
<th>CHOICE</th>
<th>VARIABLE</th>
<th>PARAMETER</th>
<th>STANDARD ERROR</th>
<th>T-STATISTIC</th>
<th>LAST CHANGE</th>
<th>DERIVATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CONSTANT</td>
<td>.58233492</td>
<td>.25838508</td>
<td>-2.2537483</td>
<td>-130820360</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>1</td>
<td>AGED</td>
<td>.7651631</td>
<td>.2080515</td>
<td>-15.684286</td>
<td>.2387957D-11</td>
<td>.794923690-01</td>
</tr>
<tr>
<td>2</td>
<td>EDOT</td>
<td>-.9080839</td>
<td>.20521828</td>
<td>.94249658</td>
<td>.583473860-01</td>
<td>.14907251</td>
</tr>
<tr>
<td>3</td>
<td>ED911</td>
<td>-.1325057</td>
<td>.14271942</td>
<td>-3.2914738</td>
<td>-.105370970-11</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>4</td>
<td>ED1315</td>
<td>.31222461</td>
<td>.15732821</td>
<td>1.9845430</td>
<td>.12221405D-10</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>5</td>
<td>AGEYRS</td>
<td>-.5171810</td>
<td>.74031894</td>
<td>.69859217</td>
<td>.59343814D-09</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>6</td>
<td>CURNORTH</td>
<td>.16550745</td>
<td>.14446642</td>
<td>1.1475644</td>
<td>.33855993D-01</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>7</td>
<td>CURWEST</td>
<td>-.11355316</td>
<td>.13339419</td>
<td>-1.0452036</td>
<td>-.10212999D-01</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>8</td>
<td>EDOT</td>
<td>.31222461</td>
<td>.15732821</td>
<td>1.9845430</td>
<td>.12221405D-10</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>9</td>
<td>ED16</td>
<td>.41941934</td>
<td>.15840158</td>
<td>2.6478230</td>
<td>.6846908D-09</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>10</td>
<td>EDOT</td>
<td>.35079753</td>
<td>.10837347</td>
<td>-3.2914738</td>
<td>-.105370970-11</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>11</td>
<td>AGED</td>
<td>.7651631</td>
<td>.2080515</td>
<td>-15.684286</td>
<td>.2387957D-11</td>
<td>.794923690-01</td>
</tr>
<tr>
<td>12</td>
<td>EDOT</td>
<td>-.9080839</td>
<td>.20521828</td>
<td>.94249658</td>
<td>.583473860-01</td>
<td>.14907251</td>
</tr>
<tr>
<td>13</td>
<td>ED911</td>
<td>-.1325057</td>
<td>.14271942</td>
<td>-3.2914738</td>
<td>-.105370970-11</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>14</td>
<td>ED1315</td>
<td>.31222461</td>
<td>.15732821</td>
<td>1.9845430</td>
<td>.12221405D-10</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>15</td>
<td>AGEYRS</td>
<td>-.5171810</td>
<td>.74031894</td>
<td>.69859217</td>
<td>.59343814D-09</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>16</td>
<td>CURNORTH</td>
<td>.16550745</td>
<td>.14446642</td>
<td>1.1475644</td>
<td>.33855993D-01</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>17</td>
<td>CURWEST</td>
<td>-.11355316</td>
<td>.13339419</td>
<td>-1.0452036</td>
<td>-.10212999D-01</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>18</td>
<td>ED16</td>
<td>.41941934</td>
<td>.15840158</td>
<td>2.6478230</td>
<td>.6846908D-09</td>
<td>.67951382D-09</td>
</tr>
<tr>
<td>OBJS</td>
<td>X</td>
<td>P</td>
<td>N</td>
<td>NIT</td>
<td>LOC_LIK</td>
<td>CHI_SQR</td>
</tr>
<tr>
<td>------</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>-----</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>3</td>
<td>2965</td>
<td>5</td>
<td>-2875.1</td>
<td>394.283</td>
</tr>
</tbody>
</table>

**Figure 3.** OUTPUT FROM THE PERMUTE AND PROB MACROS

- **OBS**: Observation count
- **X**: Predictor variable
- **P**: Predictor variable
- **N**: Number of observations
- **NIT**: Number of iterations
- **LOC_LIK**: Log likelihood
- **CHI_SQR**: Chi-square statistic
- **DF**: Degrees of freedom
- **L(P1/P2) STD_ERR**: Standard error of the log odds ratio for model 1
- **L(P2/P3) STD_ERR**: Standard error of the log odds ratio for model 2
- **L(P1/P2) STD_ERR**: Standard error of the log odds ratio for model 1

**PROBABILITY MEANS**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MEANS</th>
<th>MODEL 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>3.16116310</td>
<td>0.23980515</td>
</tr>
<tr>
<td>AGEYRS</td>
<td>-0.09385114</td>
<td>0.00119388</td>
</tr>
<tr>
<td>CURNORTI</td>
<td>-0.32995403</td>
<td>0.14442161</td>
</tr>
<tr>
<td>CURMID</td>
<td>-0.09495428</td>
<td>0.12410172</td>
</tr>
<tr>
<td>CURWEST</td>
<td>0.11110228</td>
<td>0.16353820</td>
</tr>
<tr>
<td>EDOT8</td>
<td>-0.90808389</td>
<td>0.20521828</td>
</tr>
<tr>
<td>ED911</td>
<td>-0.13250576</td>
<td>0.14271942</td>
</tr>
<tr>
<td>ED1315</td>
<td>0.09119632</td>
<td>0.15290250</td>
</tr>
<tr>
<td>ED16</td>
<td>0.41941934</td>
<td>0.15840158</td>
</tr>
</tbody>
</table>

**Figure 4.** SOURCE LISTING OF THE MACROS

---

**Problems:**
- Reject bad values
- Error handling
- Example code snippets

**Data:**
- Example data input
- Example data processing

**Program:**
- Example program execution
- Error handling

---

686
**PARTITION THE VARIANCE-COVARIANCE MATRIX INTO THE MAIN AND THE REMAINING VARIANCES ARE COMPUTED USING THE DIAGONALS OF THE SUB-MATRICES OF SO; **COMPUTE THE FULL SET OF COEFFICIENTS; STANDARD ERRORS; **THE FIRST KV-P ARMS ARE REPORTED DIRECTLY IN SO. THE REMAINING KV-P ARMS ARE COMPUTED FROM THE DIFFERENCES BETWEEN **EACH OF THE ESTIMATED ARMS. THEIR VARIANCES ARE EQUAL TO THE **SUM OF THE CORRESPONDING VARIANCES, PLUS THE COVARIANCE** **BETWEEN THEM; NDO 1=1 NDO 6P; B[4,4]=0(NN[4,4]); NDO 1=1 NDO 6P; B[4,4]=0(NN[4,4]); NDO 1=1 NDO 6P; B[4,4]=0(NN[4,4]); NDO 1=1 NDO 6P; B[4,4]=0(NN[4,4]); Wert**
CALCULATE THE PROBABILITIES

(SEE SCHMIDT AND STRAUSS (1975)


** \[ \exp \left( \beta \right) \]

** \[ Y(t) = \exp(\beta) \cdot A \]

** \[ Y^T = \exp(\beta) \cdot M \]

DO I=1 TO M;
  \[ Y_{,I} = Y_{,I} \cdot \frac{1}{A_{,I}} \]
  \[ Y = Y^T \]
** END;

** 1. MAKE COLUMNS NEXT TO THE TRANSPOSED MATRICES FOR LABELS;
** PROBILITY = \[ Y_{,I}/A_{,I} \]
** 2. ASSIGN VALUES TO EACH COLUMN TO DENOTE THE VALUE OF Y;
** DO I=1 TO M; PROBILITY(I)=I+1; END;
** 3. MAKE MEANS LABELS AND MODEL LABELS FOR A RAINY DAY;
** MEANCOLS = "MEANS" "MODEL #1"
** MEANCOLS = "MEANS"
** END;
** PRINT PROBILITY FORMAT=MHz.. MEANNAME=MEANCOLS;
** PRINT X_MATRIX FORMAT=MHz.. MEANNAME=MEANCOLS;
** MEAN;

References


Footnotes

1 Proc MLOGIT, an unsupported SAS* procedure (Salford Systems, 1985), computes maximum likelihood parameters for both the polytomous and conditional logit models.

2 This limit \((k*(p-1)<179)\) ensures that the entire variance-covariance matrix as well as all coefficient vectors fit within the PROC MATRIX limit of 32,767 elements.

3 This problem is due to an apparent formatting error in the version of MLOGIT used for this program. The last column of means labelled 'MEANS FOR ALL' actually reports the \(p\) th set of conditional means when \(P\) is between 8 and 10, inclusive. For \(P\) greater than 10 the effect of this error is uncertain.

*SAS is the registered trademark of SAS Institute Inc., Cary, NC, USA.