Logistic Regression for EpiInfo

Gerard E Dallal

Brixton Books

This edition published and distributed by Brixton Books,
Unit K, Station Building, Llanidloes, Powys SY18 6EB

ISBN 1 873937 06 7

LOGISTIC

Logistic Regression

Version 3.11Ef

Notice

Version 3.11E analyses data in EpiInfo REC files and uses EpiInfo style commands (e.g., ROUTE, RUN, READ).

Please acknowledge LOGISTIC in any manuscript that uses its calculations. A suitable reference is: Dallal GE (1988), "LOGISTIC: A Logistic Regression Program for the IBM PC," The American Statistician, 42, 272.

Disclaimer

LOGISTIC is provided 'as is' without warranty of any kind. The entire risk as to the quality, performance, and fitness for intended purpose is with you. You assume responsibility for the selection of the program and for the use of results obtained from that program.

Program Description

LOGISTIC fits the multiple logistic regression model:

$$\log(P(Y=1 \mid x_1,...,x_p)/P(Y=0 \mid x_1,...,x_p)) = b_0 + b_1*x_1 + ... + b_p*x_p$$

or, equivalently:

$$P(Y=1 \mid x_1,...,x_p) = 1 / [1 + \exp(- b_0 - b_1*x_1 - ... - b_p*x_p)]$$

where Y is a binary response variable and $X_1,...,X_p$ are explanatory variables. The program runs interactively or in batch mode.

LOGISTIC (LOGISTIC.EXE) can run as a stand alone program or from within the EpiInfo ANALYSIS module as an add-on.

LOGISTIC is written in Microsoft FORTRAN version 5.1 and was compiled with optimisation disabled. Double precision arithmetic is used throughout.

Command Language

LOGISTIC commands are:

MODEL	Specify the model to be fitted
READ	Select a file for analysis
NAMES	Display variable names
ADD	Add a variable to the current model
RUN	Execute a file of commands
HELP	Display a help screen
DELETE	Remove a variable from the current model
ROUTE	Send output to a disk file
QUIT	Exit LOGISTIC
ESTIMATE	Fit the specified model to the data
WEIGHT	Specify a weighting variable for summary records
LR	Display the LR statistics for each variable
TABLE	Construct a classification table
CODES	Change legal values of the binary response
LEVEL	Specify the width of the confidence interval
TOLERANCE	Specify the convergence criterion
SWITCH	Controls display of the regression coefficients

Command Reference

Only the first two letters of a LOGISTIC command are significant. RE and REA, for example, are equivalent to READ. Items in square brackets are optional. Items in angle brackets should be replaced by the appropriate number, filename, or variable list. A vertical bar separates option choices. The brackets themselves are never entered as part of a command. Commands may span many lines (up to a total of 800 characters) provided all but the last line end in a back slash (\). Only the first 80 characters of any line are recognised; LOGISTIC ignores characters that wrap around the screen.

Command Reference

MODEL <response> = CONSTANT + [<independent variables>]
MODEL <successes>/<total> = CONSTANT + [<independent variables>]

Specifies the model to be fitted. Terms on the right hand side are separated by plus signs (+); spaces may be used to improve readability. Data records can be individual cases containing a binary response or they can be summary records. Summary records may contain either a binary response and a weighting variable or they may include variables containing the number of successes and the number of cases. CONSTANT must be spelled out in full and must be the first word to the right of the equals sign.

Interactions such as AGE*WEIGHT can be included in the MODEL command. The right hand side of a MODEL command can contain up to 42 variable names, including duplicates, in up to 20 terms, excluding CONSTANT.

LOGISTIC cannot analyse character data, nor can it transform a variable taking on discrete values into a set of indicator variables. However, it automatically treats YES/NO variables as 1/0 variables when fitting models.

ADD <variable name>

Adds a variable to the current model. ADD cannot be used to enter an interaction.

DELETE <variable name>

Removes a variable from the current model. DELETE cannot be used if the model contains any interactions.

ESTIMATE

Fits the specified model to the data. Output includes the likelihood ratio statistic for overall significance, parameter estimates, exponentiated parameter estimates (which are the odds ratios corresponding to a unit change in the independent variables), Wald statistics for assessing the effects of independent variables, and confidence intervals for the regression parameters.

Command Reference

LR

Gives the likelihood ratio statistics for the significance of each variable. They can be used as a check on the corresponding Wald statistics, which Hauck and Donner (1977) have shown to be misleading sometimes. LR was implemented as a separate command rather than made part of ESTIMATE because of the work it generates: a different model must be fitted iteratively to assess the effect of each variable. LR is not available when the MODEL command contains interactions.

WEIGHT <variable name>

Specifies a weighting variable for summary records. The balance of the record can be thought of as appearing as many times as specified by the weight. This command is useful for analysing data files creating using the OUTPUT TABLES command in EpiInfo ANALYSIS.

TABLE

Constructs a classification table:

observed response vs. P(observed response = 1)

with probabilities, grouped in tenths, determined from the fitted regression model to aid in assessing the adequacy of the fitted model. If 3 of more rows have positive totals, a Hosmer-Lemeshow goodness-of-fit statistic (Hosmer and Lemeshow, 1989, sec. 5.2.2) is computed along with its P-value.

CODES

Changes legal values of the binary response from the default (0/1). It may also be used to reverse the values of the binary response.

LEVEL <#>

Specifies the width of the confidence intervals for the coefficients. The default value is .95.

TOLERANCE <#>

Specifies the convergence criterion. Iterations cease when the largest relative change in any coefficient between successive iterations is smaller than the specified tolerance. The default value is 0.0001. The maximum number of iterations is set at 50; it cannot be altered.

Command Reference

SWITCH ON|OFF

Permits the display of the regression coefficients after each iteration of the fitting procedure. The default is OFF.

READ <filename>

Selects a file for analysis. The filename is assumed to have the qualifier REC, which is **not** typed as part of the command.

RUN <filename>

Allows an ASCII file of commands to be executed in batch mode. The filename is assumed to have the qualifier PGM, which is **not** typed as part of the command. When the CODES command is included in a command file, it should be followed immediately by two lines: the first line contains the value to be treated as a success (1), the second contains the value to be treated as a failure (0).

ROUTE <filename>

Sends output to a disk file and screen simultaneously. The filename may include a qualifier; no default qualifier is appended. ROUTE PRINTER sends output to the printer and screen simultaneously. ROUTE SCREEN sends output to the screen; its main use is to stop output from going to a file or the printer.

NAMES

Stamps the output file with the file name and displays the variable names.

HELP

Displays a help screen.

Method

LOGISTIC carries out an iterative Newton-Raphson procedure to determine the maximum likelihood estimates of the regression coefficients. (See Walker and Duncan, 1967, sec. 4.)

Caveat

LOGISTIC recognises collinearity in the independent variables and many completely separable data patterns (Albert and Anderson, 1984; Santner and Duffy, 1977) in which there is a linear combination of independent variables for which every observation with a response of 1 has a value greater than (or all less than) that of any observation with a response of 0. The program does not recognise quasi-separation, however. The user is responsible for insuring that a problem is well posed.

Missing Values

Only cases containing all variables specified in the model are analysed. When the LR command is invoked, only those cases which are complete for the full model are used.

Problem Size & Temporary Files

LOGISTIC is limited to 20 terms on the right hand side of the model, excluding CONSTANT. These terms can contain a total of up to 42 variable names, including duplicates. The only other limitations are available disk space and a restriction to 200 variables in the data file.

LOGISTIC analyses data contained in EpiInfo REC files. LOGISTIC can make use of numerical variables and YES/NO variables, which it treats as 1 (for YES) and 0 (for NO). The original file is not modified.

LOGISTIC creates at least one and possibly two temporary disk files. First, it creates a temporary file of numerical and recoded YES/NO variables. The amount of storage in bytes that must be available is slightly more than 8 times (the number of subjects) times (the number of variables). Then, all data are placed in memory if (the number of complete cases) times (the number of terms plus 2) is no greater than 6000. Otherwise, LOGISTIC creates a temporary file to hold the data values required for the analysis.

Running Logistic From Within Analysis

LOGISTIC can be run as though it were part of ANALYSIS itself. You must copy LOGISTIC.EXE into the sub-directory that contains EpiInfo and the add the following line to the CONFIG.EPI file:

```
CMD command_name TABLES OUTPUT FILE=LOGISTIC.EXE
```

where command_name is the name to be used within EpiInfo to invoke LOGISTIC. For example, to choose the command name LOGIT, add the line:

```
CMD LOGIT TABLES OUTPUT FILE=LOGISTIC.EXE
```

to CONFIG.EPI. If you do not have a CONFIG.EPI file then, at the DOS prompt, type the two lines:

```
COPY CON CONFIG.EPI
CMD LOGIT TABLES OUTPUT FILE=LOGISTIC.EXE
```

and then press Ctrl-Z (that is, hold down the Ctrl key and simultaneously press Z) followed by the Enter key.) Now, from within ANALYSIS, typing:

```
LOGIT <list of variable names>
```

will pass the named variables to LOGISTIC for further analysis. Because ANALYSIS passes the data to LOGISTIC in the form of a table, you will probably need to use the variable COUNT in a WEIGHT command.

There is a minor side effect. Immediately upon exiting LOGISTIC, most of the ANALYSIS screen will be missing. To restore the screen, type:

```
SET SPLIT = ON
```

at the EPI> prompt. The fault is EpiInfo's.

Within LOGISTIC, the arrow keys will not recover commands nor can the F10 key be used to quit. (The ADD and DELETE commands can be used to avoid retyping many models, however.) PgUp and PgDn cannot be used to inspect LOGISTIC's output. To save results for later review, enter a ROUTE command from within LOGISTIC.

How this works

When LOGISTIC is invoked from within ANALYSIS, ANALYSIS creates the file ANAOUT.REC, which contains the variables in the specified list, and passes control to LOGISTIC. It is as though the commands:

```
ROUTE ANAOUT
OUTPUT TABLES <variable list>
DOS LOGISTIC
```

had been typed at the EPI> prompt. LOGISTIC looks to see if the file ANAOUT.REC exists. If ANAOUT.REC is found, LOGISTIC automatically loads it, as though the command READ ANAOUT had been typed at the LOGIT> prompt.

Some minor consequences

ANALYSIS erases ANAOUT.REC only when a new version is created by another external command. Therefore, the latest copy will remain in the current sub-directory when ANALYSIS is terminated. The DOS DEL command can be used to erase ANAOUT.REC, if you wish.

If LOGISTIC is used as a stand-alone procedure, it will automatically load ANAOUT.REC if there is one in the current sub-directory. A different file can be analysed simply by READing it.

Using Logistic with the EpiInfo Menu

If you are using EpiInfo v6.xx you can edit the EpiInfo menu definition file (EPI6.MNU) to add Logistic to the Programs menu. If you have not edited the EpiInfo menu file to suit your own needs then you need only copy the files EPI6.MNU and EPI6.SCR on the distribution disk to your EpiInfo directory. The next time you start EpiInfo there will be an option for 'LOGISTIC Regression' on the Programs menu.

If you have already edited the EPI6.MNU file to suit your needs then copying the EPI6.MNU file from the distribution disk will overwrite your existing menu system. If you want to keep your existing menu system you will have to edit the EPI6.MNU file yourself. You can do this with any text editor (such as EPED or the EPI6 file editor). Open the file EPI6.MNU. Change the first section to read:

```
BEGIN
  POPUP "&Programs"
  BEGIN
    MENUITEM "&EPED word processor^#14", DoEped
    MENUITEM "E&NTER data^#14", DoEnter
    MENUITEM "&ANALYSIS of data^#14", DoAnalysis
    MENUITEM "&LOGISTIC Regression", DoLogReg
    MENUITEM SEPARATOR
```

After the section of the file labelled 'DoAnalysis' add a new section:

```
DoLogReg
BEGIN
 LOGISTIC
END
```

In the section of the file labelled 'STRINGTABLE' add a description line:

```
STRINGTABLE
BEGIN
  DoEPED     "« Create questionnaires ... "
  DoENTER    "« Make .REC files ... "
  DoANALYSIS "« Do LISTS/FREQuencies/ ..."
  DoLogReg   "« Logistic Regression from Brixton Books"
```

Save the file. Now copy the file EPI6.SCR on the distribution disk to the EpiInfo directory. The next time you start EpiInfo there will be an option for 'LOGISTIC Regression' on the Programs menu.

You can use a similar procedure to add cLogistic to the EpiInfo menu.

Algorithms

Freeman PR (1982), "Remark AS R44. A remark on AS6 and AS7. Triangular Decomposition of a Symmetric Matrix and Inversion of a Positive Semi-definite Symmetric Matrix," Applied Statistics, 31, 336-339.

Hill ID (1973), "Algorithm AS 66. The normal integral," Applied Statistics, 22, 424-427.

Odeh RE and Evans JO (1974), "Algorithm AS 70. The percentage points of the normal distribution," Applied Statistics, 23, 96-97.

Pike MC and Hill ID (1966), "Algorithm 291. Logarithm of the gamma function," Communications of the Association of Computing Machinery, 9, 684.

References

Albert A and Anderson JA (1984), "On the existence of maximum likelihood estimates in logistic regression models," Biometrika, 71, 1-10.

Hauck WW, Jr. and Donner A (1977), "Wald's Test as Applied to Hypotheses in Logit Models," Journal of the American Statistical Association, 72, 851-853.

Hosmer DW, Jr. and Lemeshow S (1989), Applied Logistic Regression, New York: John Wiley & Sons, Inc.

Santner TJ and Duffy DE (1986), "A note on A. Albert and J. A. Anderson's conditions for the existence of maximum likelihood estimates in logistic regression models," Biometrika, 73, 755-758.

Walker SH and Duncan DB (1967), "Estimation of the Probability of an Event as a Function of Several Independent Variables," Biometrika, 42, 167-179.

cLOGISTIC

Conditional Logistic Regression

Version 1.00E

Notice

Version 1.00E analyses data in EpiInfo REC files and uses EpiInfo style commands (e.g., ROUTE, RUN, READ).

Please acknowledge cLOGISTIC in any manuscript that uses its calculations. A suitable reference is Dallal GE, (1989), "cLOGISTIC: A Conditional Logistic Regression Program for the IBM-PC," The American Statistician, 43, 125.

Disclaimer

cLOGISTIC is provided 'as is' without warranty of any kind. The entire risk as to the quality, performance, and fitness for intended purpose is with you. You assume responsibility for the selection of the program and for the use of results obtained from that program.

Program Description

cLOGISTIC fits multiple logistic models to stratified case-control data. cLOGISTIC models the log odds of being a case as a linear function of the specified predictor variables, that is, the probability of being a case is:

$$P(\text{case} \mid x_1,...,x_p,\text{k-th stratum}) = 1 / [1 + \exp(- a_k - b_1*x_1 - ... - b_p*x_p)]$$

where $X_1,...,X_p$ are explanatory variables, $\{a_k:k=1,...,\#\text{ of strata}\}$ are stratum specific constants, and $b_1,...,b_p$ may be interpreted as the logarithms of the odds ratios associated with unit changes of each of the explanatory variables. The program runs interactively or in batch mode.

cLOGISTIC (CLOGISTI.EXE) can run as a stand alone program or from within the EpiInfo ANALYSIS module as an add-on.

cLOGISTIC is written in Microsoft FORTRAN version 5.1 and was compiled with optimisation disabled. Double precision arithmetic is used throughout.

Limitations

cLOGISTIC can handle up to twenty explanatory variables, some of which may be interactions, provided the total number of named original variables, including duplicates, is no greater than 42. The number of strata may be no greater than 500 with no more than 10 cases plus controls per stratum. Further, the product of the number of subjects and the number of variables can be no greater than 6000.

Important - Data Organisation

The data **must** be sorted by stratum. Within each stratum, cases and controls may **not** be interspersed - cases must precede controls or controls must precede cases. The same order (cases/controls or controls/cases) must be true for all strata. The STRATUM, CASES, and CONTROLS commands are used to describe the structure of the data.

For example, if your EpiInfo data file contained a variable called MATCHNUM that identified the case stratum and a variable called CACO that identified the case or control status of each observation you would issue the commands:

```
SORT MATCHNUM CACO
ROUTE WORKFILE.REC
WRITE RECFILE
```

In EpiInfo ANALYSIS to prepare a file called WORKFILE.REC for analysis using cLOGISTIC. The cLOGISTIC command to read this file and specify the stratum and case / control order would be:

```
READ WORKFILE
STRATUM MATCHNUM
CASES
```

Command Language

cLOGISTIC commands are:

MODEL	Specify the model to be fitted
READ	Select a file for analysis
NAMES	Display variable names
FILENAME	Display the name of the file being analysed
ADD	Add a variable to the current model
RUN	Execute a file of commands
HELP	Display a help screen
DELETE	Remove a variable from the current model
ROUTE	Send output to a disk file
QUIT	Exit LOGISTIC
ESTIMATE	Fit the specified model to the data
LR	Display the LR statistics for each variable
LEVEL	Specify the width of the confidence interval
TOLERANCE	Specify the convergence criterion
SWITCH	Controls display of the regression coefficients
MESSAGE	Displays author's message
STRATUM	Specifies the variable that identifies a case's stratum
CASES	Within each stratum cases precede controls
CONTROLS	Within each stratum controls precede cases

Command Reference

Only the first two letters of a command are significant. ME and MESS, for example, are equivalent to MESSAGE. Items in square brackets are optional. Items in angle brackets are to be replaced by the appropriate number, filename, or variable list. A vertical bar separates option choices. The brackets themselves are never entered as part of a command. Commands may span many lines (up to a total of 800 characters) provided all but the last line end in a back slash (\). Only the first 80 characters of any line are recognised; cLOGISTIC ignores characters that wrap around the screen.

cLOGISTIC models the log odds of being a case as a linear function of the specified predictor variables.

Command Reference

MODEL <response> = CONSTANT + [<independent variables>]

Specifies the model to be fitted. Terms on the right hand side are separated by plus signs (+); spaces may be used to improve readability. Any variable that takes on only two distinct values and distinguishes between cases and controls is a satisfactory response variable. The particular values of the response variable are immaterial; they do not affect the fitting process. CONSTANT must be spelled out in full and must be the first word to the right of the equals sign.

Interactions such as AGE*WEIGHT can be included in the MODEL command. The right hand side of a MODEL command can contain up to 42 variable names, including duplicates, in up to 20 terms, excluding CONSTANT.

cLOGISTIC cannot analyse character data, nor can it transform a variable taking on discrete values into a set of indicator variables. However, it automatically treats YES/NO variables as 1/0 variables when fitting models.

STRATUM <variable name>

Specifies the variable that identifies an observation's stratum.

ESTIMATE

Fits the specified model to the data. Output includes the likelihood ratio and score statistics for overall significance, parameter estimates, exponentiated parameter estimates (which are the odds ratios corresponding to a unit change in the independent variables), Wald statistics for assessing the effects of independent variables, and confidence intervals for the regression parameters.

LR

Gives the likelihood ratio statistics for the significance of each variable. They can be used as a check on the corresponding Wald statistics, which Hauck and Donner (1977) have shown to be misleading sometimes. LR was implemented as a separate command rather than made part of ESTIMATE because of the work it generates: the model must be refitted iteratively to assess the effect of each variable. The LR command can be issued only after a model had been fitted by using the ESTIMATE command.

Command Reference

CASES

Indicates that within each stratum cases precede controls. This is the default. It is reset with every READ command.

CONTROLS

Indicates that within each stratum controls precede cases.

LEVEL <#>

Specifies the width of the confidence intervals for the coefficients. The default value is .95.

TOLERANCE <#>

Specifies the convergence criterion. Iterations cease when the largest relative change in any coefficient between successive iterations is smaller than the specified tolerance. The default value is 0.0001. The maximum number of iterations is set at 50; it cannot be altered.

READ <filename>

Selects a file for analysis. The filename is assumed to have the qualifier REC, which is not typed as part of the command.

RUN <filename>

Allows an ASCII file of commands to be executed in batch mode. The filename is assumed to have the qualifier PGM, which is **not** typed as part of the command.

ROUTE <filename>

Sends output to a disk file and screen simultaneously. The filename may include a qualifier; no default qualifier is appended. ROUTE PRINTER sends output to the printer and screen simultaneously. ROUTE SCREEN sends output to the screen; its main use is to stop output from going to a file or the printer.

NAMES

Stamps the output file with the name of the variables in the file being analysed.

Command Reference

FILENAME

Stamps the output file with the name of the file being analysed.

HELP

Displays a help screen.

Method

cLOGISTIC carries out an iterative Newton-Raphson procedure to determine the maximum likelihood estimates of the regression coefficients conditionally on the sets of values of the explanatory variables within each stratum (Smith et al 1981, as modified by Krailo 1984). If a stratum contains 'n' cases and 'm' controls, the conditional likelihood of the stratum is a fraction whose denominator is the product of the m+n probabilities of being a case, conditional on $x_1,...,x_p$, and whose numerator is the sum of C(m+n,n) terms, each the product of the probability of being a case for 'n' of the m+n individuals. The conditional likelihood that is maximised is the product of the conditional likelihoods for each stratum.

Caveat

The user is responsible for insuring that a problem is well-posed. The only ill-condition cLOGISTIC can recognise is a degenerate data set.

Missing Values

Only cases containing all variables specified in the model are analysed. When the LR command is invoked, only those cases which are complete for the full model are used.

Data Files & Temporary Files

cLOGISTIC analyses data contained in EpiInfo REC files. cLOGISTIC can make use of numerical variables and YES/NO variables, which it treats as 1 (for YES) and 0 (for NO). The original file is not modified.

cLOGISTIC creates a temporary file of numerical and recoded YES/NO variables. The amount of storage in bytes that must be available is slightly more than 8 times (the number of subjects) times (the number of variables).

Running cLOGISTIC From Within Analysis

cLOGISTIC can be run as though it were part of ANALYSIS itself. You must copy CLOGISTI.EXE into the sub-directory that contains EpiInfo and the add the following line to the CONFIG.EPI file:

```
CMD command_name TABLES OUTPUT FILE=CLOGISTI.EXE
```

where command_name is the name to be used within EpiInfo to invoke cLOGISTIC. For example, to choose the command name cLOGIT, add the line:

```
CMD CLOGIT TABLES OUTPUT FILE=CLOGISTI.EXE
```

to CONFIG.EPI. (If you do not have a CONFIG.EPI file then, at the DOS prompt, type the two lines:

```
COPY CON CONFIG.EPI
CMD CLOGIT TABLES OUTPUT FILE=CLOGISTI.EXE
```

and then press Ctrl-Z (that is, hold down the Ctrl key and simultaneously press Z) followed by the Enter key.) Now, from within ANALYSIS, typing:

```
CLOGIT <list of variable names>
```

will pass the named variables to CLOGISTI.EXE for further analysis.

There is a minor side effect. Immediately upon exiting cLOGISTIC, most of the ANALYSIS screen will be missing. To restore the screen, type:

```
SET SPLIT = ON
```

at the EPI> prompt.

Within cLOGISTIC, the arrow keys will not recover commands nor can the F10 key be used to quit. (The ADD and DELETE commands can be used to avoid retyping many models, however.) PgUp and PgDn cannot be used to inspect cLOGISTIC's output. To save results for later review, enter a ROUTE command from within cLOGISTIC.

How this works

When cLOGISTIC is invoked from within ANALYSIS, ANALYSIS creates the file ANAOUT.REC, which contains the variables in the specified list, and passes control to LOGISTIC. It is as though the commands:

```
ROUTE ANAOUT
OUTPUT TABLES <variable list>
DOS LOGISTIC
```

had been typed at the EPI> prompt. cLOGISTIC looks to see if the file ANAOUT.REC exists. If ANAOUT.REC is found, cLOGISTIC automatically loads it, as though the command READ ANAOUT had been typed at the LOGIT> prompt.

Some minor consequences

ANALYSIS erases ANAOUT.REC only when a new version is created by another external command. Therefore, the latest copy will remain in the current sub-directory when ANALYSIS is terminated. The DOS DEL command can be used to erase ANAOUT.REC, if you wish.

If LOGISTIC is used as a stand-alone procedure, it will automatically load ANAOUT.REC if there is one in the current sub-directory. A different file can be analysed simply by READing it.

Example

From Smith et al. (1981):

Matchnum	Caco	Exposure1	Exposure2
1	1	0	1
1	1	1	2
1	0	0	1
1	0	1	3
2	1	0	1
2	0	1	0
2	0	0	2

If these data are stored in the file TEST.REC, the data are then analysed by cLOGISTIC through the set of commands:

```
READ TEST
STRATUM MATCHNUM
CASES
MODEL CACO = CONSTANT + EXPOSURE1 + EXPOSURE2
ESTIMATE
```

To obtain the parameter estimates:

b_1 = -.5223
b_2 = -.2674

cLOGISTIC and LOGISTIC

LOGISTIC is a program for unconditional logistic regression. LOGISTIC and cLOGISTIC have substantially the same capabilities and command structure with the following exceptions:

LOGISTIC does not require that the observations be in any particular order.

LOGISTIC's TABLE command, not implemented in cLOGISTIC, classifies each observation according to its fitted probability.

LOGISTIC does not permit the use of the LR command when the model contains interactions.

Algorithms

cLOGISTIC makes use of Freeman (1982), Hill (1973), Krailo (1984), Odeh and Evans (1974), Smith et al. (1981), and a FORTRAN translation of Pike and Hill (1966).

References

Freeman, P.R. (1982), "Remark AS R44. A remark on AS6 and AS7. Triangular Decomposition of a Symmetric Matrix and Inversion of a Positive Semi-definite Symmetric Matrix," Applied Statistics, 31, 336-339.

Hauck, Walter W., Jr. and Allan Donner (1977), "Wald's Test as Applied to Hypotheses in Logit Models," Journal of the American Statistical Association, 72, 851-853.

Hill, I.D. (1973), "Algorithm AS 66. The normal integral," Applied Statistics, 22, 424-427.

Krailo, Mark D. (1984), "A Remark on Algorithm AS 162. Multivariate Conditional Logistic Analysis of Stratum-matched Case-control Studies," Applied Statistics, 33, 123.

Odeh, R.E. and J.O. Evans (1974), "Algorithm AS 70. The percentage points of the normal distribution," Applied Statistics, 23, 96-97.

Pike, M.C. and I.D. Hill (1966), "Algorithm 291. Logarithm of the gamma function," Communications of the Association of Computing Machinery, 9, 684.

Smith PG, Pike MC, Hill AP, Breslow NE, and NE Day (1981), "Algorithm AS 162. Multivariate Conditional Logistic Analysis of Stratum-matched Case-control Studies," Applied Statistics, 30, 190-197.