What is Logit Regression and how does it appear in the Mplus Annotated Output?

Authored by

stats writer

June 29, 2024

Logit Regression is a statistical method used to model the relationship between a categorical dependent variable and one or more independent variables. It is commonly used in social science research to analyze binary or ordinal outcome data.

In the Mplus Annotated Output, Logit Regression appears as part of the Model Results section. It provides information on the estimated coefficients, standard errors, odds ratios, and p-values for each independent variable included in the model. This allows researchers to determine the strength and significance of the relationship between the dependent variable and the independent variables. Additionally, Mplus also provides model fit statistics such as the chi-square test and the Akaike Information Criterion (AIC) to evaluate the overall fit of the Logit Regression model. Overall, the Logit Regression results in the Mplus Annotated Output provide valuable insights into the relationship between variables and aid in drawing meaningful conclusions from the data.

Logit Regression | Mplus Annotated Output

**This page shows an example of logit regression with footnotes
explaining the output. First an example is shown using Stata, and then an
example is shown using Mplus, to help you relate the output you are likely to be
familiar with (Stata) to output that may be new to you (Mplus). We suggest that
you view this page using two web browsers so you can show the page side by side
showing the Stata output in one browser and the corresponding Mplus output in
the other browser.**

This example is from the Mplus User's Guide (example 3.5) and we suggest that
you see the Mplus User's Guide for more details about this example. We thank the
kind people at Muthén & Muthén for permission to use examples from their manual.

Example Using Stata

Here is a logit regression example using Stata with two continuous predictors
x1 and x2 used to predict a binary outcome variable, u1.

```
infile u1 x1 x3 using https://stats.idre.ucla.edu/wp-content/uploads/2016/02/ex3.5.dat, clear

tabulate u1

u1 | Freq. Percent Cum.
------------+-----------------------------------
0 | 327 65.40A 65.40
1 | 173 34.60A 100.00
------------+-----------------------------------
```

Total | 500 100.00

A. These are the percent of cases with 0 and 1 on the variable u1

logit u1 x1 x3

Iteration 0: log likelihood = -322.46763
Iteration 1: log likelihood = -216.57883
Iteration 2: log likelihood = -203.79479
Iteration 3: log likelihood = -202.63515
Iteration 4: log likelihood = -202.61995
Iteration 5: log likelihood = -202.61995

Logistic regression Number of obs = 500
LR chi2(2) = 239.70
Prob > chi2 = 0.0000
Log likelihood = -202.61995 Pseudo R2 = 0.3717

--
u1 | Coef. Std. Err. z P>|z|
-------------+--
x1 | 1.071767E .1428573 7.50 0.000 .791772 1.351762
x3 | 1.838588E .1794923 10.24 0.000 1.486789 2.190386
_cons | -1.025842D .1369173 -7.49 0.000 -1.294195 -

.7574886

---

logit , or

Logistic regression Number of obs = 500
LR chi2(2) = 239.70
Prob > chi2 = 0.0000
Log likelihood = -202.61995 Pseudo R2 = 0.3717

---

u1 | Odds Ratio Std. Err. z P>|z|

----------+----------

x1 | 2.920536F .4172198 7.50 0.000 2.207304 3.864229
x3 | 6.287652F 1.128585 10.24 0.000 4.422872 8.938663

---

estat ic

Model | Obs ll(null) ll(model)B df AICC BICC

----------+----------

. | 500 -322.4676 -202.6199 3 411.2399 423.8837

The output is labeled with superscripts to help you relate the later Mplus output to this Stata output. To summarize the output, both predictors in this model, x1 and x2, are significantly related to the outcome variable, u1. The coefficients from the logit output can be exponentiated to obtain odds ratios, as shown in the output from the logit, or command. For a one unit increase in x1, the odds of u1 equaling 1 (as compared to u1 equaling 0) increases by a factor of 2.92. The estat ic command produces fit indices for the model including the log likelihood for the empty (null) model, the log likelihood for the model, as well as the AIC and BIC fit indices.

Mplus Example #1

Here is the same example illustrated in Mplus based on

the https://stats.idre.ucla.edu/wp-content/uploads/2016/02/ex3.5.dat data file.

```
TITLE:
this is an example of a logistic
regression for a categorical observed
dependent variable with two covariates
DATA:
FILE = https://stats.idre.ucla.edu/wp-content/uploads/2016/02/ex3.5.dat;
VARIABLE:
NAMES = u1 x1 x3;
CATEGORICAL = u1;
ANALYSIS:
ESTIMATOR = ML;
! need to use estimator = ml to make this a logistic model;
MODEL:
u1 ON x1 x3;
```

SUMMARY OF ANALYSIS

Number of observations 500

Estimator MLR

<some output was omitted to save space>

SUMMARY OF CATEGORICAL DATA PROPORTIONS

U1

Category 1 0.654A

Category 2 0.346A

TESTS OF MODEL FIT

Loglikelihood

H0 Value -202.620B

Information Criteria

Number of Free Parameters 3

Akaike (AIC) 411.240C

Bayesian (BIC) 423.884C

Sample-Size Adjusted BIC 414.362

(n* = (n + 2) / 24)

MODEL RESULTS

Estimates S.E. Est./S.E.

U1 ON
X1 1.072D 0.143 7.502
X3 1.839D 0.179 10.243

Thresholds
U1$1 1.026E 0.137 7.492

LOGISTIC REGRESSION ODDS RATIO RESULTS

U1 ON
X1 2.921F
X3 6.288F

Mplus Example #2

Here is another version of this example in Mplus. Note that by using
estimator=ml; (maximum likelihood) the results are shown in a logit metric.
Had we specified something like estimator=wls; (weighted least squares)
then the results would be shown in a probit scale. Because this analysis does

not use the type=logistic option (unlike example #1), the format of the
output is somewhat different (notably omitting odds ratios from the output).

```
TITLE:
this is an example of a logistic
regression for a categorical observed
dependent variable with two covariates.
DATA:
FILE =
https://stats.idre.ucla.edu/wp-content/uploads/2016/02/ex3.5.dat;
VARIABLE:
NAMES = u1 x1 x3;
CATEGORICAL = u1;
! note using Maximum Likelihood produces results in Logit scale
! using GLS produces results in Probit scale
analysis:
estimator=ml;
MODEL:
u1 ON x1 x3;
```

SUMMARY OF ANALYSIS

Number of observations 500

Estimator ML

<some output omitted to save space>

SUMMARY OF CATEGORICAL DATA PROPORTIONS

U1

Category 1 0.654A

Category 2 0.346A

THE MODEL ESTIMATION TERMINATED NORMALLY

TESTS OF MODEL FIT

Loglikelihood

H0 Value -202.620B

Information Criteria

Number of Free Parameters 3

Akaike (AIC) 411.240C

Bayesian (BIC) 423.884C

Sample-Size Adjusted BIC 414.362

(n* = (n + 2) / 24)

MODEL RESULTS

Estimates S.E. Est./S.E.

U1 ON

X1 1.072E 0.143 7.503

X3 1.839E 0.179 10.245

Thresholds

U1$1 1.026D 0.137 7.493