What is Multinomial Logit Regression and how can it be applied in analyzing data?

Authored by
stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *What is Multinomial Logit Regression and how can it be applied in analyzing data?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=160096

Multinomial Logit Regression is a statistical method used to analyze categorical data with more than two categories. It is based on the logistic regression model and is used to predict the probability of an event occurring in one category compared to the other categories. This technique is commonly used in fields such as marketing, economics, and social sciences to understand the relationship between a set of independent variables and a categorical outcome. It can be applied to analyze data by identifying the important factors that influence the outcome and determining the relative impact of each factor on the different categories. Multinomial Logit Regression is a powerful tool for understanding and predicting behavior in complex systems with multiple outcomes.

Multinomial Logit Regression | Mplus Annotated Output

This page shows an example of multinomial 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 3.6) 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.

Stata Example

Here is a multinomial 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/ex3.6.dat, clear

mlogit u1 x1 x3

Iteration 0: log likelihood = -539.2303
Iteration 1: log likelihood = -446.49742
Iteration 2: log likelihood = -434.20483
Iteration 3: log likelihood = -433.4331
Iteration 4: log likelihood = -433.42628
Iteration 5: log likelihood = -433.42628

Multinomial logistic regression Number of obs = 500

LR chi2(4) = 211.61

Prob > chi2 = 0.0000

Log likelihood = -433.42628 Pseudo R2 = 0.1962

--

u1 | Coef. Std. Err. z P>|z|

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

0 |

x1 | .7686261C .1567749 4.90 0.000 .461353 1.075899

x3 | 2.259422C .2144306 10.54 0.000 1.839146 2.679699

_cons | -.7488877E .1702198 -4.40 0.000 -1.082512 -.4152631

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

1 |

x1 | .2798667D .1131474 2.47 0.013 .0581018 .5016316

x3 | .885101D .1402897 6.31 0.000 .6101382 1.160064

_cons | .2622508E .1198104 2.19 0.029 .0274268 .4970748

--

(u1==2 is the base outcome)

estat ic

--

Model | Obs ll(null) ll(model) df AIC BIC

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

. | 500 -539.2303 -433.4263A 6 878.8526B 904.1402B

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 x3, are
significantly related to predicting the comparison of level 0 to level 2 of the
outcome variable, u1. Likewise, x1 and x3, are
significantly related to predicting the comparison of level 1 to level 2 of the
outcome variable, u1. 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

Here is the same example illustrated in Mplus based on the ex3.6 data file.

TITLE:

this is an example of a multinomial logistic regression for an unordered categorical (nominal) dependent variable with two covariates

DATA:

FILE IS https://stats.idre.ucla.edu/wp-content/uploads/2016/02/ex3.6.dat;

VARIABLE:

NAMES ARE u1 x1 x3;

NOMINAL IS u1;

MODEL:

u1#1 u1#2 ON x1 x3;

Number of observations 500

Estimator MLR

THE MODEL ESTIMATION TERMINATED NORMALLY

TESTS OF MODEL FIT

Loglikelihood

H0 Value -433.426A

Information Criteria

Number of Free Parameters 6

Akaike (AIC) 878.853B

Bayesian (BIC) 904.140B

Sample-Size Adjusted BIC 885.096

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

MODEL RESULTS

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

U1#1 ON

X1 0.769C 0.165 4.670

X3 2.259C 0.203 11.148

U1#2 ON

X1 0.280D 0.114 2.444

X3 0.885D 0.143 6.200

Intercepts

U1#1 -0.749E 0.158 -4.728

U1#2 0.262E 0.120 2.192