

Multinomial Logistic Regression - Important Considerations

The Independence of Irrelevant Alternatives (IIA) assumption:

Roughly, the IIA assumption means that adding or deleting alternative outcome categories does not affect the odds among the remaining outcomes. There are alternative modeling methods, such as alternative-specific multinomial probit model, or nested logit model to relax the IIA assumption.

Diagnostics and model fit:

Unlike logistic regression where there are many statistics for performing model diagnostics, it is not as straightforward to do diagnostics with multinomial logistic regression models. For the purpose of detecting outliers or influential data points, one can run separate logit models and use the diagnostics tools on each model.

Sample size:

Multinomial regression uses a maximum likelihood estimation method, it requires a large sample size. It also uses multiple equations. This implies that it requires an even larger sample size than ordinal or binary logistic regression.

Complete or quasi-complete separation:

Complete separation means that the outcome variable separate a predictor variable completely, leading perfect prediction by the predictor variable. Perfect prediction means that only one value of a predictor variable is associated with only one value of the response variable. But you can tell from the output of the regression coefficients that something is wrong. You can then do a two-way tabulation of the outcome variable with the problematic variable to confirm this and then rerun the model without the problematic variable.

Empty cells or small cells:

You should check for empty or small cells by doing a cross-tabulation between categorical predictors and the outcome variable. If a cell has very few cases (a small cell), the model may become unstable or it might not even run at all.