**Multinomial Logistic Regression**

Multinomial logistic regression (MLR) is a statistical method used to predict the probability of categorical outcomes with more than two possible values. It is an extension of logistic regression, which is used for binary outcomes (only two possible values). The independent variables for MLR can be either continuous or categorical. Some examples of when to use MLR would be, predicting the type of disease a patient has based on their symptoms or a bank using it to classify a customer as low, medium, or high risk for loans.

MLR uses a set of coefficients to weight the predictor variables and then applies a transformation to convert these weighted sums into probabilities for each category. The coefficients are estimated using maximum likelihood estimation, which involves finding the values of the coefficients that maximize the likelihood of the observed data.

MLR can be more complex to interpret than binary logistic regression, as there are multiple coefficients for each independent variable. One of the independent variable categories becomes a baseline variable. You then compare your baseline category with the remaining independent variable categories. The output represents the log-odds when comparing your baseline variable to the corresponding category that is being analyzed. The log-odds increases if the number is positive and decreases if the number is negative.

Multinomial Logistic Regression assumes that the data is linear, that all observations are independent, that outcome categories are mutually exclusive, and that there is no multicollinearity between independent variables. We can assess the assumption of linearity by simply looking at a residual vs. fitted plot of the data or performing a Box Tidewell Test The assumption of independence can be assessed by understanding the data and how it was sampled. We can plot a correlation matrix of the data to visually assess multicollinearity or by using variance inflation factor.

Team members: Jackie Voge, Ignacio Luque, Leo Walker, Sarah Millard

References:

Outside Reference 1:

https://stats.oarc.ucla.edu/r/dae/multinomial-logistic-regression/

Outside Reference 2:

https://www.mygreatlearning.com/blog/multinomial-logistic-regression/

Outside Reference 3:

https://www.statstest.com/multinomial-logistic-regression/

Outside Reference 4:

https://bookdown.org/chua/ber642\_advanced\_regression/multinomial-logistic-regression.html

Outside Reference 5:

https://towardsdatascience.com/multinomial-logistic-regression-in-r-428d9bb7dc70

Outside Reference 6:

https://en.wikipedia.org/wiki/Multinomial\_logistic\_regression

Outside Reference 7:

https://resources.nu.edu/statsresources/Multinomiallogistic

Outside Reference 8 (R reference):

https://www.r-bloggers.com/2020/05/multinomial-logistic-regression-with-r/

Outside Reference 9 (R reference):

https://bookdown.org/chua/ber642\_advanced\_regression/multinomial-logistic-regression.html

Outside Reference 10 (R reference):

https://datasciencebeginners.com/2018/12/20/multinomial-logistic-regression-using-r/