**Comprehension Check**

**Conceptual questions**

1. What are the assumptions of multinomial logistic regression?

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.

2. What are the data requirements?

The outcome variable must be categorical with more than 2 levels. The predictor variables can be either categorical or continuous. The relationships between the independent and dependent variables must be linear. The observations/records in the data must be independent of each other, meaning that they should not influence each other. Independent variables should not be highly correlated with each other.

3. In what field can we use multinomial logistic regression?

Multinomial logistic regression can be used in many different fields, such as marketing, healthcare, and social sciences, to predict outcomes like consumer preferences, disease diagnoses, and political affiliations.

**Implementation questions**

4. What packages, if any, are required to run multinomial logistic regression?

Requires library (nnet) for multinom, to estimate a multinomial logistic regression model.

5. What is one way to measure for high multicollinearity?

We can use a variance inflation factor, which measures how much the behavior (variance) of an independent variable is influenced, or inflated, by its interaction/correlation with the other independent variables.

6. What metrics can you use to judge your model’s performance?

Precision and recall are very important metrics to judge your classification model for an imbalanced dataset.