Characterizing Level 2 Automation in a Naturalistic Driving Fleet

Data exercises

Background

- The introduction of automation features into the vehicle fleet is disrupting the way vehicles operate
  - Potentially affects what drivers do with these features and how they expect them to perform
- Data from 55 vehicles from the Virginia Tech Transportation Institute Level 2 (L2) Naturalistic Driving Study were used to support analyses of driver behavior with L2 automation features
- The data are available at https://doi.org/10.15787/VTT1/42MUF1
- This module provides exercises related to analysis approaches applicable to this data

Approach 1: Regression

- Regression can support the development of models that capture the relationship between different variables

- https://en.wikipedia.org/wiki/Regression_analysis

- Exercise:
  - Plot "Activation duration [ms]" vs. "Total trip duration [ms]"
  - Is a linear regression model appropriate?
  - Build a linear regression model that predicts the total trip duration based on the duration of activations and summarize the resulting regression equation and R-square

Approach 2: ANOVA

- Analysis of Variance (ANOVA) allows to determine whether different groups are statistically different

- https://en.wikipedia.org/wiki/Analysis_of_variance

- Exercise:
  - Develop boxplots of "Hands-off-wheel time [sec]" for the different "Lateral warning provided [5 sec]" groups
  - Run an ANOVA test on the "Hands-off-wheel time [sec]" variable with the "Lateral warning provided [5 sec]" as a factor
    - Is there a significant effect?
    - Perform a suitable *post hoc* test to determine differences between the different levels of the "Lateral warning provided [5 sec]" factor

Approach 3: Frequency analysis

- Frequency analysis allows determining whether the frequency of occurrence of some circumstances are related to specific factors

- https://en.wikipedia.org/wiki/Chi-squared_test

- Exercise:
    - Determine whether the frequency of different types of "Lateral warning provided [5 sec]" is associated with "Curve presence"

Closing

- These are just a small subset of the statistical approaches that can be used on transportation datasets