

Testing

Prediction tests

Michel Bierlaire

Introduction to choice models

Outlier analysis

Outlier analysis

Procedure

- ▶ Apply the model on the sample
- ▶ Examine observations where the predicted probability is the smallest for the observed choice
- ▶ Test model sensitivity to outliers, as a small probability has a significant impact on the log likelihood
- ▶ Potential causes of low probability:
 - ▶ Coding or measurement error in the data
 - ▶ Model misspecification
 - ▶ Inexplicable variation in choice behavior

Coding or measurement error in the data

Look for signs of data errors

- ▶ Travel time is negative
- ▶ Number is coded as a string
- ▶ etc.

Correct or remove the observation

- ▶ Go back to the original survey
- ▶ Correct only if you are certain

Model misspecification

Improve the specification

- ▶ Seek clues of missing variables from the observation.
- ▶ Why is the model associating such a low probability for this choice?
- ▶ Did we forget to account for age, income, or any other variable ?
- ▶ Should a nonlinear specification be investigated?
- ▶ Use a behavioral intuition.

Inexplicable variation in choice behavior

Keep the observation

- ▶ If no acceptable explanation is found, keep the observation.
- ▶ Avoid overfitting of the model to the data.
- ▶ The model should reflect how people behave, not how they should behave.