

Handling Non-significant Linear Terms in Regression Models

Introduction

- In regression models with interactions, the treatment of non-significant linear terms varies between physicists and statisticians.
- This discrepancy arises from different perspectives on modeling and interpretation.
- Let's explore the rationale and implications of each approach.

- **Focus on Simplicity and Relevance:**

- Physicists prioritize parsimonious models with clear physical interpretations.
- Non-significant linear terms are often removed, even if interactions involving these terms are retained.

- **Rationale:**

- If a term does not significantly affect the outcome on its own, it is considered irrelevant.
- Simplifies the model and highlights only significant predictors.

- **Hierarchical Principle:**

- Statisticians follow the principle of hierarchy.
- If an interaction term is included, all lower-order terms (main effects) are also included, regardless of their significance.

- **Rationale:**

- Removing a main effect can distort the interpretation of the interaction term.
- Ensures that interactions are interpreted in the context of main effects.

Implications for Model Interpretation

- **Interpretation:**

- Removing a non-significant linear term can lead to misleading interpretation of interactions.
- Interactions measure how the effect of one variable changes depending on the level of another variable.

- **Model Stability:**

- Removing terms can lead to increased variability in parameter estimates.
- Lower-order terms provide a baseline for understanding interactions.

- **Effect on Coefficient Estimates:**

- Removing terms changes the reference level for interactions.
- Can alter the magnitude and significance of other coefficients.

- **Implications:**

- Standardized coefficients and interpretations may vary.
- Changes the scale and reference level of the variables.

- **Model Purpose:**

- For predictive models, hierarchical principles often apply.
- For simpler, interpretable models, removing non-significant terms may be justified.

- **Trade-off:**

- Balance between model simplicity and accurate representation of relationships.
- Consider the impact on interpretation and model robustness.

Conclusion

- The decision to keep or remove non-significant terms depends on the modeling goals and interpretation needs.
- Statisticians aim for a comprehensive model interpretation, while physicists focus on physical relevance and simplicity.
- Understanding these perspectives helps in making informed modeling choices.