

Welcome to **instats**

The Session Will Begin Shortly

START

Statistics in R with Tidyverse

Session 11: Inference for Regression

Introduction to Statistical Inference for Regression

- Regression helps us explore relationships between variables
- Inference allows us to make conclusions about these relationships
- Types of regression:
 - **Simple Linear Regression:** One predictor/regressor, one response
 - **Multiple Linear Regression:** More than one predictor/regressor for the response

Simple Linear Regression (SLR)

- Focuses on the relationship between one explanatory variable and one response variable
- Goal: Estimate how much the response variable changes with one unit increase in the predictor

Inference for SLR Coefficients

- Confidence intervals: Estimate the range of the true slope for one predictor
- Hypothesis testing: Check if the predictor has a significant effect on the response
 - Null hypothesis: No effect (slope = 0)
 - Alternative hypothesis: There is an effect (slope $\neq 0$)

Multiple Linear Regression

- Multiple predictors are used to explain the response variable
- Each predictor explains a unique part of the variability in the response
- Coefficients are adjusted for the presence of other predictors
- Identifies the contribution of each predictor while controlling for others
- Inference involves understanding the significance and confidence intervals for each predictor's coefficient

Inference for MLR Coefficients

- Each predictor has its own **coefficient**, representing its unique impact
- Confidence intervals: Estimate the range of the true coefficients for each predictor
- Hypothesis testing: Check if each predictor has a significant effect on the response, controlling for other predictors

Key Differences in Inference

- In SLR: We are assessing the relationship between one predictor and the response
- In MLR: We assess each predictor's effect **while controlling for the others**

Model Fit and Assumptions

- For both SLR and MLR:
 - The relationships should be linear
 - Residuals should appear random and normally distributed
 - LINE acronym
- In MLR: Check for **multicollinearity** (predictors should not be highly correlated)

Hypothesis Tests for Partial Slopes

- Formulating hypothesis tests
 - Null and alternative hypotheses
 - Conducting t-tests for regression coefficients
 - p-values and statistical significance
 - Practical implications and decision-making

Interpreting Results in MLR

- Each coefficient in MLR shows the effect of one predictor **while holding the others constant**
- MLR can handle interaction terms and transformations
- Bootstrap methods can estimate confidence intervals when assumptions are violated
- Permutation tests can check the significance of predictors in MLR

Demo & Exercises

Q & A

STOP