Statistics in R - Assumptions

Miranda Y Mortlock

2 Oct 2019

Assumptions

- Assumptions for Statistical Inference
- What
- Why
- So what ?
- Practical considerations
- Other considerations

Four key Assumptions

1. Normality: Data have a normal distribution (or at least is symmetric)
2. Homogeneity of variances: Data from multiple groups have the same variance.
3. Linearity: Data have a linear relationship.
4. Independence: Data are independent.

Why it is important

- . . . . . because both assumptions and limitations affect the inferences you can draw from your study.
- One of the more common assumptions made in survey research is the assumption of honesty and truthful responses.

Polls v Surveys

- "No clue" v Simple Random Sample
- Main research Question
- Design and sample size
- Pilots
- Questionnaire design and method (Mail, CATI)

Known Population

- Population
- Target Population
- Actual Population (for sampling)
- Plan to achieve the sample you want

Results and data

- Non Response
- RTS (dead, out of country)
- Understand non response
- Incomplete data
- Missing and modeling
- Present possible solutions if your data fails to meet the required assumptions.

Therefore

- Design well
- Use Pilots
- Previous studies can inform
- Be aware of biases
- Use statistical principles
- Retain your reputation
- Respect Company/Agency reputation too