

Introduction to data simulations in R

FDZ Herbstakademie / Autumn Academy 2025

- ♦ Dozierende*r / Dr. Nora Wickelmaier (Leibniz-Institut für Wissensmedien (IWM), Tübingen)
Lecturer:
- ♦ Termin / Donnerstag / Thursday, 18.09.2025, 9:00 – 17:00 h
Date and Time:

Abstract

Simulating data can be considered as the key idea behind hypothesis testing in the frequentist framework, where we assume that we can repeat the data collection under identical conditions. Therefore, data simulations are a powerful tool to help understand data generating processes assumed by statistical models and at the same time to better our understanding of any inferences we want to draw from our results. In this workshop, we will focus on applied settings from psychology and educational science and see that data simulations can help applied researchers to get a much better understanding of their data, the underlying assumptions of the data generating processes, and what are valid conclusions that can be drawn from their data.

Inhalte / Contents

- ♦ Introduction to basic concepts of simulation
- ♦ Generating random numbers in R
- ♦ Creating design matrices for applied research designs
- ♦ Better understanding of data generating processes

Voraussetzungen / Previous knowledge required

- ♦ Basic knowledge of R
- ♦ Basic knowledge of statistics, including concepts like random variables, statistical distributions like the normal distribution, t distribution, Binomial distribution, etc., and hypothesis testing

Literatur / Literature

Strobl, C., Henninger, M., Rothacher, Y., & Debelak, R. (2024). *Simulationsstudien in R*. Springer Berlin.

Software / Software requirements

Participants will need to have installed:

- ♦ a current R version (<https://cran.r-project.org/>)
- ♦ an IDE for R (like RStudio or VSCode) or a text editor with syntax highlighting (like Vim or Notepad++)