

Konzeption Lehrveranstaltung

Ingo Rohlf, 22. März 2014

Titel: Statistische Datenanalyse mit R

Art: Seminar

Umfang: 4 Doppelstunden Präsenz mit praktischen Übungen

Zielgruppe: Doktoranden, Wissenschaftler und fortgeschrittene Studenten

Voraussetzungen: Erfahrung im Umgang mit Daten, z.B. Messwerten

Lernziel: Mit dem Werkzeug „R“ reale Daten strukturieren und mit wichtigen Methoden der modernen Datenanalyse untersuchen können

Motivation: Die stark wachsende Menge an Primärdaten verlangt nach immer besseren Methoden der Verarbeitung. „R“ ist eine freie Datenverarbeitungssoftware mit einem starken Fokus auf statistische Analyse, Visualisierung und auf grosse Datenmengen („Big Data“). Auch wegen der Fülle von Erweiterungen entwickelt es sich zunehmend zum Marktführer, besonders im wissenschaftlichen Umfeld. Aufgrund seiner Mächtigkeit und seines spezialisierten Konzeptes ist der Einstieg mit deutlichem Lernaufwand verbunden.

Strukturierung der Unterrichtseinheiten

Tag 1: Überblick über die Methoden der statistischen Datenanalyse – Einsatzgebiete und Aufbau des Systems „R“ – Datentypen in R, Kontrollstrukturen – vektororientierte Problemlösung – Erweiterungen und Module

Tag 2: Datenerfassung von hoch- und mittelstrukturierten Daten – Datenmodelle und -strukturierung – Statistische Kennzahlen und Verteilungsfunktionen – Graphische Datenanalyse

Tag 3: Umgang mit realen Daten – Robuste Statistik, Ausreisser und fehlende Werte – nichtparametrische Verfahren – Analyse grosser Datenmengen – fortgeschrittene graphische Datenanalyse

Tag 4: Auswertung von Messreihen physikalischer Grössen – Multivariate lineare Regression – Zeitreihenanalyse mit Prognosen