

# MSc. Research Methods – Statistikteil

## Beispielhafte Lösung einer Übungs-/Klausuraufgabe

**Abzugeben sind immer drei Elemente:**

- (1) Lauffähiges R-Script
- (2) Begründung von gewählten Methoden/Schritten aufgrund von Zwischenergebnissen (z. B. Abbildung von EDA, Residualplots,...)
- (3) Ein zugehöriger, ausformulierter Methoden- und Ergebnistext wie in einer wissenschaftlichen Arbeit

(1) und (2) können ggf. in einem Dokument sein; (3) muss immer ein separates Word- oder pdf-Dokument sein.

### **Beispielaufgabe**

#### **Regressionsanalyse mit decay.csv**

Der Datensatz (decay.csv, siehe online verfügbare Materialien zum Lehrbuch von Crawley 2015) beschreibt in einem physikalischen Experiment die Zahl der radioaktiven Zerfälle pro Minute in Abhängigkeit vom Zeitpunkt (min nach Start des Experimentes).

- Ladet den Datensatz in R und macht eine explorative Datenanalyse.
- Wählt unter den schon gelernten Methoden der Regressionsanalyse ein adäquates Vorgehen zur Analyse dieser Daten und führt diese dann durch.
- Prüft anhand der Residuen, ob die Modellvoraussetzungen erfüllt waren
- Stellt die erhaltenen Ergebnisse angemessen dar (Text, Abbildung und/oder Tabelle).
- Kennt ihr ggf. noch eine andere geeignete Herangehensweise?

### **Beispiel für Element (3)**

#### **Methoden**

Es wurden die radioaktiven Zerfallsvorgänge pro Zeiteinheit in 31 gleichen Zeitintervallen nach dem Start des Experimentes ermittelt.

Mittels linear Regression wurde ein geeignetes Modell ermittelt und die Erfüllung der Modellvoraussetzungen anschliessend in den Residualplots visuell geprüft. Da die abhängige Variable (Zerfälle pro Zeiteinheit) stark rechtsschief verteilt war, wurde sie vor der Berechnung des Modells  $\log_e$ -transformiert.

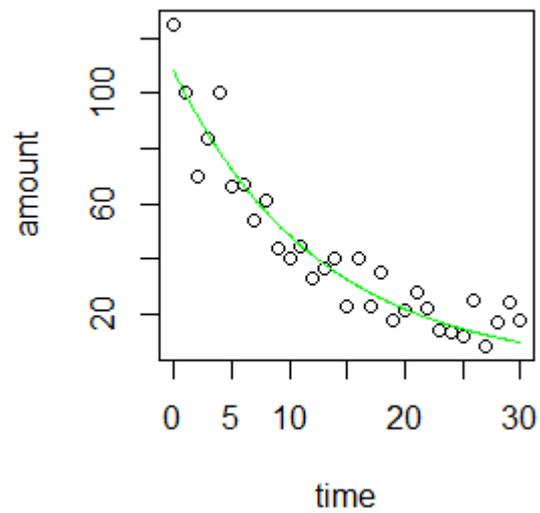
#### **Ergebnisse**

Das für die  $\log_e$ -transformierte Zerfallsrate ermittelte höchstsignifikante ( $p < 0.001$ ) Regressionsmodell lautet

$$\log_e(\text{Zerfallsrate}) = 4.55 - 0.07 * \text{time}$$

und erklärt dabei 83% der Variabilität der Daten im  $\log_e$ -transformierten Raum. In Originaldaten lautet das Modell wie folgt (Abb. 1):

$$\text{Zerfallsrate} = e^{4.55} * \exp(-0.07 * \text{time}) = 94.6 * \exp(-0.07 * \text{time})$$



**Abb. 1:** Radioaktiver Zerfall mit beobachteten Werten und modellierter Funktion.