

Bernd Klaus (bernd.klaus@imise.uni-leipzig.de)
Verena Zuber (verena.zuber@imise.uni-leipzig.de)

<http://uni-leipzig.de/~zuber/teaching/ws11/r-kurs/>

1 Aufgabe: Lineare Einfachregression

Wir betrachten den Datensatz “*Spielzeugautos*”. Er beschreibt die Wegstrecke, die 3 verschiedene Spielzeugautos zurückgelegt haben, nachdem man sie in unterschiedlichen Winkeln eine Rampe herunterfahren ließ.

- *angle*: Winkel der Rampe
 - *distance*: Zurückgelegte Strecke des Spielzeugautos
 - *car*: Autotyp (1, 2 oder 3)
- (a) Lesen Sie den Datensatz “*Spielzeugautos*” in einen dataframe `data` ein und wandeln Sie die Variable “`car`” des Datensatzes in einen Faktor (`as.factor`) um.
- (b) Erstellen Sie drei Boxplots, die die zurückgelegte Strecke getrennt nach dem Faktor “`car`” darstellen.
- (c) Schätzen Sie für **jedes** der 3 Autos **separat** die Parameter des folgenden linearen Modells mit Hilfe der Funktion “`lm()`”

$$\text{distance}_i = \beta_0 + \beta_1 \cdot \text{angle}_i + \varepsilon_i$$

- (d) Überprüfen Sie deskriptiv die Anpassung der drei Modelle, indem Sie die Regressionsgerade in einen Plot von *distance* gegen *angle* einfügen. Deutet das R^2 jeweils auf eine gute Modellanpassung hin?
- (e) Führen Sie weitere deskriptive Diagnosen mit Hilfe der `plot.lm()` Funktion durch. Besteht ein linearer Zusammenhang? Sind die Residuen normalverteilt? Haben die Fehler gleiche Varianz?

2 Aufgabe: Überprüfung einer Daumenregel mittels linearer Einfachregression

Für den Zusammenhang zwischen *Alter* und *Herzfrequenz* gibt es folgende Daumenregel:

$$\text{Herzfrequenz} = 220 - \text{Alter}$$

Im Datensatz *Herz.csv* auf der Homepage finden Sie 15 empirische Herzfrequenzmesswerte aus verschiedenen Altersgruppen.

- (a) Lesen Sie den Datensatz “*Herz*” in einem dataframe `HerData` ein und erstellen Sie einen Scatterplot. Deutet der Plot auf einen linearen Zusammenhang zwischen *Alter* und *Herzfrequenz* hin?
- (b) Überprüfen Sie mittels eines linearen Modells, ob die Daumenregel stimmt.