

## Projekt 5: Regressionsanalyse anhand einer Studie zum Eiskonsum

Der Datensatz ICECREAM.DAT enthält den Eiskonsum  $Y$  pro Kopf in Pints (1 Pint = 0,473 l) von 30 vierwöchigen Perioden im Zeitraum vom 18. März 1951 bis zum 11. Juli 1953. Zusätzlich wurde der durchschnittliche Preis  $X_1$  pro Pint Eis in Dollar, das durchschnittliche wöchentliche Familieneinkommen  $X_2$  in Dollar so wie die durchschnittliche Temperatur  $X_3$  in °F erfasst. Es soll untersucht werden, welchen Einfluss der Eispreis, das Familieneinkommen und die Temperatur auf den Eiskonsum haben. Dabei sind die Daten im Datensatz ICECREAM.DAT von folgender Form:

	$Y$	$X_1$	$X_2$	$X_3$
1	0.386	0.270	78	41
2	0.374	0.282	79	56
3	0.393	0.277	81	63
4	0.425	0.280	80	68
⋮	⋮	⋮	⋮	⋮
30	0.548	0.260	90	71

### Aufgaben

1. Welche Einflussvariablen haben einen signifikanten Einfluss auf den Eiskonsum? Untersuchen Sie dazu, ob eine Kollinearität der Regressoren vorliegt. Betrachten Sie einmal ein Modell mit allen möglichen Wechselwirkungen und ein Modell nur mit Haupteffekten. Beachten Sie dabei, dass der Versuchsplan nicht balanciert ist.
2. Bestimmen Sie unter Verwendung von Variablenselektion- sowie Modelldiagnoseverfahren ein möglichst „gutes“ Modell, welches den Konsum anhand der Einflussfaktoren erklärt.
3. Sagen Sie mit dem in 2. bestimmten Modell den Eiskonsum bei einer Temperatur von 50°F, einem Preis von 0.27\$ und einem wöchentlichen Einkommen von 80\$ voraus.

## Literatur

Elpelt, B., Hartung, J. (2004). *Grundkurs Statistik. Lehr- und Übungsbuch der angewandten Statistik.*, Oldenbourg, München.

Sachs, L. und Hedderich, J. (2009). *Angewandte Statistik: Methodensammlung mit R.* Springer, Berlin.

Faraway, J.J. (2005). *Linear models with R.* CRC Press, Boca Raton.

Faraway, J.J. (2007). *Practical Regression and ANOVA using R.*

<http://csyue.nccu.edu.tw/Practical%20Regression%20and%20Anova%20using%20R.pdf>,  
12.3.2013

Groß, J. (2003). *Linear Regression.* Springer, Berlin.

[http://en.wikipedia.org/wiki/Akaike\\_information\\_criterion](http://en.wikipedia.org/wiki/Akaike_information_criterion), 12.2.2010

## Literatur zu den Daten

Koteswara Rao Kadiyala (1970). Testing for the independence of regression disturbances. *Econometrica* **38**, 97-117.

aus

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. und Ostrowski, E. (1994). *A Handbook of Small Data Sets.* Chapman & Hall, London, Seite 214.

## Abgabetermin

Abgabe bis spätestens **Donnerstag, den 13. Juni 2019**, in der Veranstaltung.