5.3 Zweites Beispiel

Wir verwenden die Daten aus Kapitel 3.3 und wollen nun mit Hilfe der linearen Regressionsanalyse den Einfluss der Körpergröße (v2) und des Alters (v3) auf das Gewicht (v1) untersuchen.

Den folgenden Output erhalten Sie, wenn Sie unter Lineare Regression für x1 die Variable v2, für x2 die Variable v3 und für y die Variable v1 auswählen und danach **→(Lineare) Regression →Lineare Regression** wählen.

Lineare Regression

Hier sind die eingegebenen Daten zu sehen:

Werte der unabhängigen Variablen	Werte der abhängigen Variable (y_i)
170; 20	54
170; 21	67
167; 22	60
177; 23	63
182; 23	75
167; 24	63
164; 25	56
170; 25	65
169; 26	61
176; 27	80

Die Regressionsfunktion: $y = b_2*x_2 + b_1*x_1 + b_0$

Parameter	Schätzer	geschätzte Standardabweichung	Prüfgröße ('t-Value') für Test (H0: Parameter = 0; gegen H1: Parameter <> 0)	p-Wert
b_0	-147.03661736442	57.045775341904	-2.5775198335574	0.0366
b_1	1.0371521684633	0.31696646574044	3.272119547538	0.0136
b_2	1.4354307679452	0.78105528655225	1.8378094261182	0.1087

Stichprobenumfang n	10
geschätzte Fehlervarianz	27.078771383209
Bestimmtheitsmaß r^2	0.67114607966263

ANOVA-Tabelle

	Freiheitsgrade	Quadratsummen	Mittlere Quadratsummen	Prüfgröße ('F-Value')	p-Wert
Modell (Model)	2	386.84860031754	193.42430015877	7.1430234932561	0.0204
Fehler (Error)	7	189.55139968246	27.078771383209		
Gesamt (Total)	9	576.4			

Obige Prüfgröße ist für den Test H0: b_i = 0 mit i >= 1; gegen H1: Es existiert mindestens ein i >= 1 mit b_i <> 0.

Das Bestimmtheitsmaß hat hier einen Wert von 0,6711… . In diesem Modell ist ebenfalls der Parameter β_1 (dieser Parameter erfasst die Steigung der Regressionsgeraden bezüglich der Körpergröße) signifikant von Null verschieden (0,0136 $\leq$ 0,05), während β_2 (dieser Parameter erfasst die Steigung der Regressionsfunktion bezüglich des Alters) nicht signifikant von Null verschieden (0,1087 > 0,05) ist. Der Parameter β_0 (der Achsenabschnitt der Regressionsfunktion) ist signifikant von Null verschieden (0,0366 $\leq$ 0,05). Da β_1 signifikant ungleich Null ist, kann ein Einfluss der Körpergröße (v2) auf das Gewicht v1 nachgewiesen werden, während kein Einfluss des Alters (v3) nachgewiesen werden kann. Bei dem globalen F-Test kommt man zu dem Ergebnis, dass mindestens ein Steigungsparameter (β_i für $i > 0$) signifikant ungleich Null ist ($0{,}0204 \leq 0{,}05$), was sich bereits bestätigt hat.

Auf die gleiche Art wie die einfache lineare Regression wird nun die multiple lineare Regression durchgeführt. Hierzu wird einfach nur die Designmatrix mit der Spalte, in der die Altersangaben stehen, erweitert. Der Rest verläuft analog (mit gleichem Vektor $\bar{y}$ wie im Kapitel 3.4.1):

$$X = \begin{pmatrix} 1 & 170 & 20 \\ 1 & 170 & 21 \\ 1 & 167 & 22 \\ 1 & 177 & 23 \\ 1 & 182 & 23 \\ 1 & 167 & 24 \\ 1 & 164 & 25 \\ 1 & 170 & 25 \\ 1 & 169 & 26 \\ 1 & 176 & 27 \end{pmatrix}$$

$$\hat{\bar{\beta}}\left(X^t X\right)^{-1} X^T \bar{y} \approx \begin{pmatrix} -147{,}037 \\ 1{,}03715 \\ 1{,}43543 \end{pmatrix}$$

Die Komponenten dieses Vektors findet man im Output wieder unter „Schätzer“.

Umsetzung mit SAS:

```
data dat1;
input x z y;
datalines;
170    20    54
170    21    67
167    22    60
177    23    63
182    23    75
167    24    63
164    25    56
170    25    65
169    26    61
176    27    80
run;

proc reg data=dat1;
model y=x z /dw covb;
run;
```

SAS-Output zur Prozedur REG:

Die Prozedur REG
Model: MODEL1
Dependent Variable: y

Number of Observations Read	10
Number of Observations Used	10

Varianzanalyse

Quelle	DF	Summe der Quadrate	Mittleres Quadrat	F-Statistik	Pr > F
Model	2	386.84860	193.42430	7.14	0.0204
Error	7	189.55140	27.07877		
Corrected Total	9	576.40000			

Root MSE	5.20373	**R-Square**	0.6711
Dependent Mean	64.40000	**Adj R-Sq**	0.5772
Coeff Var	8.08032		

Parameterschätzer

Variable	DF	Parameter-schätzer	Standard-fehler	t-Wert	Pr > |t|
Intercept	1	-147.03662	57.04578	-2.58	0.0366
x	1	1.03715	0.31697	3.27	0.0136
z	1	1.43543	0.78106	1.84	0.1087

Kovarianz der Schätzer

Variable	Intercept	x	z
Intercept	3254.2204844	-17.10395397	-13.69981729
x	-17.10395397	0.1004677404	-0.004073017
z	-13.69981729	-0.004073017	0.6100473607