

Übungen zur Vorlesung QM II Lineare Regression

Aufgabe 7.1

Ein Unternehmen möchte den Einfluss seiner Werbemaßnahmen auf den erzielten Umsatz quantifizieren. Hierfür werden die jährlichen Ausgaben für Werbung und die jährlich erzielten Umsätze über einen Zeitraum von acht Jahren erfasst. Die Daten sind der folgenden Tabelle zu entnehmen:

Jahr	Werbung in 1000 €	Umsatz in 100 000 €
2006	6	9
2007	9	10
2008	10	13
2009	11	12
2010	12	10
2011	12	16
2012	13	15
2013	15	25

- a) Berechnen und interpretieren Sie eine geeignete statistische Maßzahl zur Messung des linearen Zusammenhangs zwischen den jährlichen Werbeausgaben und dem Umsatz.
- b) Angenommen das Unternehmen plant, 14 000 € für Werbung auszugeben. Mit welchem Umsatz kann es dann rechnen?
- c) Angenommen das Unternehmen plant, einen Umsatz von 2 600 000 € zu erzielen. Wie viel Euro sind dann für Werbung auszugeben?
- d) für wie verlässlich halten Sie Ihre Ergebnisse aus b) und c)? (Begründung!)

Aufgabe 7.2

Die deutsch-amerikanische Umweltschutz-Organisation International Council on Clean Transportation (kurz ICCT) stellte für europäische Neuwagen folgende durchschnittlichen Abweichungen in Prozent zwischen dem vom Hersteller angegeben Spritverbrauch und dem tatsächlichen Spritverbrauch im Straßenverkehr fest:

Jahr	durchsch. Abweichung
2001	+8%
2013	+25%
2015	+40%

Quelle: Süddeutsche vom 07.11.2015

- a) Prognostizieren Sie die Abweichung in Prozent für das Jahr 2016 mittels einer geeigneten statistischen Methode, falls vom Gesetzgeber bis dahin keine neuen Hersteller-Regeln für die Feststellung des Spritverbrauchs erfolgen.
- b) Nehmen Sie kritisch zur Aussagekraft des Prognosewertes Stellung.

- c) Beurteilen Sie die Güte Ihres Prognosemodells mit Hilfe der Berechnung einer geeigneten Maßzahl.

Aufgabe 7.3

Hinweis: Diese Aufgabe bringt wirklich Spaß!

Bearbeiten Sie das Applet „Guessing Correlations“. Sie gelangen zu dem Applet wie folgt:

<http://istics.net/Correlations/>

Es erscheinen vier Streudiagramme und vier Korrelationswerte. Ordnen Sie die Korrelationswerte den jeweiligen Streudiagrammen zu. Starten Sie bitte das Applet mehrmals neu und ordnen den Korrelationen die richtigen Streudiagramme zu.

Aufgabe 7.4

Bearbeiten Sie das Applet „Regression“. Sie gelangen zu dem Applet wie folgt:

<http://www.shodor.org/interactive/activities/Regression/>

Zeichnen Sie per Maus-Klick einige Punkte in das Streudiagramm ein und beobachten Sie dabei den Wert des Korrelationskoeffizienten. Und tragen Sie die Regressionsgerade ins Streudiagramm ein (Display line of best fit). Verändern Sie Punkte (Move points) in dem Streudiagramm, so dass sich das Vorzeichen der Steigung verändert. Gelingt Ihnen das?

Lösung zu Aufgabe 7.1:

X =Werbeausgaben (in 1 000 Euro pro Jahr)

Y =Jahres-Umsatz (in 100 000 Euro)

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
6	9			
9	10			
10	13			
11	12			
12	10			
12	16			
13	15			
15	25			
88	110	1 288	1 020	1 700

a) 1. Lösungsweg:

$$b_1 = \frac{8 \cdot 1\,288 - 88 \cdot 110}{8 \cdot 1\,020 - 88 \cdot 88} = \frac{624}{416} = 1,5$$

$$b_2 = \frac{624}{8 \cdot 1\,700 - 110 \cdot 110} = \frac{624}{1\,500} = 0,416$$

$$r = +\sqrt{1,5 \cdot 0,416} = 0,789\,936\,7$$

d.h. es liegt positive mittelstarke Korrelation vor.

2. Lösungsweg:

$$\bar{x} = \frac{88}{8} = 11$$

$$s_x^2 = \frac{1\,020}{8} - 11^2 = 6,5 \text{ und } s_x = \sqrt{6,5} = 2,54951$$

$$\bar{y} = \frac{110}{8} = 13,75$$

$$s_y^2 = \frac{1\,700}{8} - 13,75^2 = 23,4375 \text{ und } s_y = \sqrt{23,4375} = 4,841229$$

$$s_{xy} = \frac{1\,288}{8} - 11 \cdot 13,75 = 9,75$$

$$r_{xy} = \frac{9,75}{2,54951 \cdot 4,841229} = 0,789\,936\,7$$

$$b_1 = \frac{9,75}{6,5} = 1,5$$

$$b_2 = \frac{9,75}{23,4375} = 0,416$$

b) Gesucht: $a_1 + b_1 \cdot 14 = ?$

$$a_1 = \frac{110}{8} - 1,5 \cdot \frac{88}{8} = \frac{110 - 1,5 \cdot 88}{8} = -2,75$$

$$\text{Prognosewert} = -2,75 + 1,5 \cdot 14 = 18,25$$

d.h. es ist mit einem Umsatz von 1 825 000 Euro zu rechnen.

c) Gesucht: $a_2 + b_2 \cdot 26 = ?$

$$a_2 = \frac{88}{8} - 0,416 \cdot \frac{110}{8} = \frac{88 - 0,416 \cdot 110}{8} = 5,28$$

Prognosewert: $5,28 + 0,416 \cdot 26 = 16,096$

d.h. es ist mit Ausgaben für Werbung in Höhe von 16 000 Euro zu rechnen.

- d) Da es sich bei der Berechnung unter b) um eine Interpolation handelt und da der lineare Zusammenhang beinahe stark ist, hat der Prognosewert Verlass. Hingegen ist der Prognosewert unter c) ein extrapoliertes Wert, auf den kein Verlass ist.

Lösung zu Aufgabe 7.2:

X = Zeitpunkte

Y = Abweichungen (in %)

Regressionsgerade $f(x) = a_1 + b_1x$

Jahr	x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2001	1	8	8	1	64
2013	13	25	325	169	625
2015	15	40	600	225	1 600
Σ	29	73	933	395	2 289

1. Lösungsweg:

$$b_1 = \frac{3 \cdot 933 - 29 \cdot 73}{3 \cdot 395 - 29 \cdot 29} = \frac{682}{344} = 1,98$$

$$a_1 = \frac{73 - 1,98 \cdot 29}{3} = \frac{15,58}{3} = 5,19$$

2. Lösungsweg:

$$\bar{x} = \frac{29}{3} = 9,6\bar{6}$$

$$s_x^2 = \frac{395}{3} - 9,6\bar{6}^2 = 38,2\bar{2} \text{ und } s_x = \sqrt{38,2\bar{2}} = 6,182412$$

$$\bar{y} = \frac{73}{3} = 24,3\bar{3}$$

$$s_y^2 = \frac{2\,289}{3} - 24,3\bar{3}^2 = 170,8\bar{8} \text{ und } s_y = \sqrt{170,8\bar{8}} = 13,07245$$

$$s_{xy} = \frac{933}{3} - 9,6\bar{6} \cdot 24,3\bar{3} = 75,7\bar{7}$$

$$r_{xy} = \frac{75,7\bar{7}}{6,182412 \cdot 13,07245} = 0,9376202$$

$$b_1 = \frac{75,7\bar{7}}{38,2\bar{2}} = 1,98$$

$$a_1 = 24,3\bar{3} - 1,98 \cdot 9,6\bar{6} = 5,19$$

a) Jahr 2016 $\hat{=}$ Zeitpunkt 16

$$f(16) = 5,19 + 1,98 \cdot 16 = 36,87 \approx 37$$

d.h. für das Jahr 2016 ergibt sich ein Prognosewert von etwa 37%.

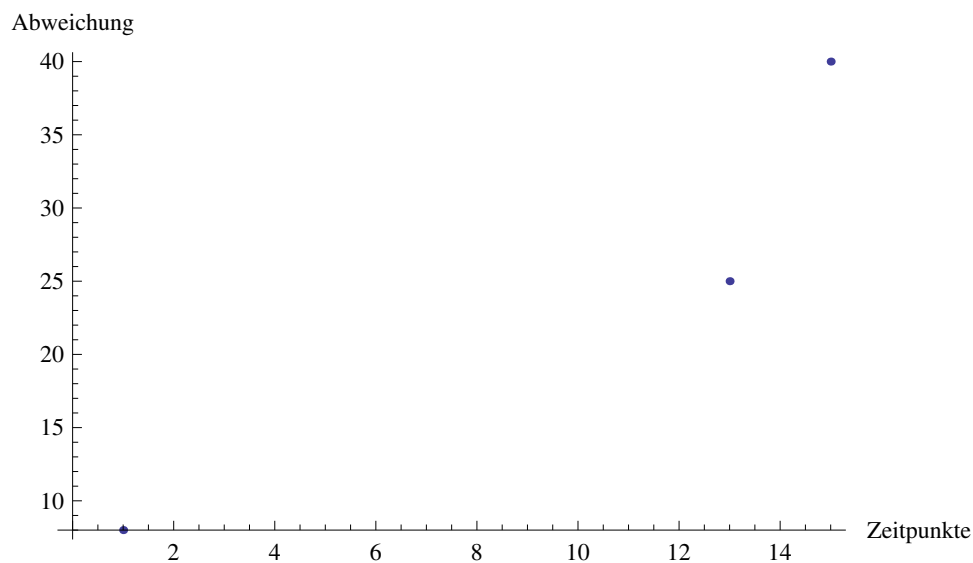
b) Der Zeitpunkt $x = 16$ liegt außerhalb des Beobachtungszeitraums [1;15], deshalb

ist der vorhergesagte Wert ein extrapoliertes Wert. Aus statistischer Sicht besteht kein Verlass auf extrapolierte Werte.

$$c) b_2 = \frac{682}{3 \cdot 2289 - 73^2} = \frac{682}{1538} = 0,44$$

$$r_{xy} = +\sqrt{1,98 \cdot 0,44} = 0,93$$

d.h. die Korrelation ist positiv und stark. Bei einem Blick auf das Streudiagramm:



lässt sich erkennen, dass der Zusammenhang zwischen x und y nicht linear, sondern quadratisch zu sein scheint.

Technische Hochschule Köln
Fakultät für Wirtschafts- und Rechtswissenschaften
 Prof. Dr. Arrenberg
 Raum 221, Tel. 39 14
 jutta.arrenberg@th-koeln.de

Übungen zur Vorlesung QM II
 Arbeitsblatt: Lineare Regression

Beispiel 1

X =Anteil der Führerscheininhaber unter 21 Jahren an allen Führerscheinelizenzen einer Stadt

Y =Anteil der tödlichen Unfälle pro 1 000 Führerscheinelizenzen in dieser Stadt

$n= 12$ städte in den USA

i	x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	13	3,0	39,0	169	9,00
2	12	0,7	8,4	144	0,49
3	8	0,9			
4	11	2,1			
5	18	3,8			
6	9	1,0			
7	16	2,8			
8	12	1,4			
9	17	4,1			
10	14	2,9			
11	10	1,0			
12	9	0,9	8,1	81	0,81
\sum	149	24,6	345,1	1 969	66,98

Wie hoch ist die Korrelation?

$$b_1 = \frac{12 \cdot 345,1 - 149 \cdot 24,6}{12 \cdot 1969 - 149^2} = \frac{475,8}{1\,427} = 0,3334268$$

$$b_2 = \frac{12 \cdot 66,98 - 149 \cdot 24,6}{12 \cdot 66,98 - 24,6^2} = \frac{475,8}{198,6} = 2,39577$$

$$r = +\sqrt{0,3334268 \cdot 2,39577} = 0,8937639$$

Beispiel 2

Gibt es bei den Bundestagswahlen einen Zusammenhang zwischen dem Stimmanteil für die CDU/CSU und der insgesamten Wahlbeteiligung?

Aus den Jahren 1949 bis 2017 liegen folgende Daten vor:

Jahr	Wahlbeteiligung (in %)	CDU/CSU (in %)	SPD (in %)
1949	78,5	31,0	29,2
1953	85,5	45,2	28,8
1957	87,8	50,2	31,8
1961	87,7	45,3	36,2
1965	86,8	47,6	39,3
1969	86,7	46,1	42,7
1972	91,1	44,9	45,8
1976	90,7	48,6	42,6
1980	88,6	44,5	42,9
1983	89,1	48,8	38,2
1987	84,4	44,3	37,0
1990	77,8	43,8	33,5
1994	79,0	41,5	36,4
1998	82,2	35,1	40,9
2002	79,1	38,5	38,5
2005	77,7	35,2	34,2
2009	70,8	33,8	23,0
2013	71,5	41,5	25,7
2017	76,2	33,0	20,5

wahlen.sav

Stellen Sie ein Modell auf, das den Stimmenanteil der CDU/CSU erklärt durch die Höhe der Wahlbeteiligung.

Lösung:

X = Wahlbeteiligung (in Prozent)

Y = Stimmenanteil der CDU/CSU

Modell: $\boxed{\text{CDU/CSU-Stimmenanteil}} \approx a_1 + b_1 \cdot \boxed{\text{Wahlbeteiligung}}$

$b_1 = 0,702$ d.h. steigt die Wahlbeteiligung um einen Prozentpunkt, so steigt der Stimmenanteil für die CDU/CSU um etwa 0,7 Prozentpunkte.

$a_1 = -16,018$

$r = 0,747$ d.h. positive fast starke Korrelation; d.h. je Höher die Wahlbeteiligung, desto Höher ist der Stimmenanteil für die CDU/CSU.

Technische Hochschule Köln
Fakultät für Wirtschafts- und Rechtswissenschaften
Prof. Dr. Arrenberg
Raum 221, Tel. 39 14
jutta.arrenberg@th-koeln.de

Vorlesung QM II
Arbeitsblatt: Lineare Regression

Beispiel 1

X =Ausgaben für Werbung in 1 000 Euro

Y =Jahresumsatz in 100 000 Euro

i	x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	4	4	16	16	16
2	4	5	20	16	25
3	5	6			
4	6	6			
5	8	8			
6	8	10			
7	10	12	120	100	144
8	11	13	143	121	169
\sum	56	64	509	442	590

$$b_1 = 1,22 \text{ und } b_2 = 0,7820513$$

$$a_1 = -0,54 \text{ und } a_2 = 0,7435897$$

$$r = +\sqrt{1,22 \cdot 0,7820513} = \sqrt{0,9541026} = 0,9767817 \approx 0,98$$

- Führen Sie den Jahresumsatz zurück auf die Ausgaben für Werbung und ermitteln Sie die lineare Regressionsgerade.
- Mit welchem Jahresumsatz ist zu rechnen, wenn 7 000 Euro in die Werbung gesteckt werden?
- Wie viel Geld ist für Werbung auszugeben, wenn ein Jahresumsatz von 900 000 Euro erzielt werden soll?

Lösung:

a) $\boxed{\text{Jahresumsatz}} \approx a_1 + b_1 \cdot \boxed{\text{Ausgaben für Werbung}}$

$$\text{Regressionsgerade} = f(x) = a_1 + b_1 \cdot x = -0,54 + 1,22 \cdot x$$

b) Gesucht $a_1 + b_1 \cdot 7 = ?$

$$a_1 + b_1 \cdot 7 = -0,54 + 1,22 \cdot 7 = 8$$

d.h. wenn 7 000 Euro für Werbung ausgegeben werden, so ist mit einem Jahresumsatz von 800 000 Euro zu rechnen.

c) Gesucht $a_2 + b_2 \cdot 9 = ?$

$$a_2 + b_2 \cdot 9 = 0,7435897 + 0,7820513 \cdot 9 = 7,782051 \approx 8$$

d.h. um einen Jahresumsatz von 900 000 Euro zu erzielen, sind 8 000 Euro in die Werbung zu stecken.