

Übungsaufgaben 3

Abgabetermin: 06.11.2002

Lineare Regression.

Alle Hilfsmittel wie Taschenrechner oder PC-Programme sind diesmal erlaubt; es soll jeweils das angewandte Verfahren kurz beschrieben und eine Zeichnung angefertigt (oder ausgedruckt) werden.

1. Bei der ersten Regressions-Aufgaben soll *EXCEL* (oder ein anderes Tabellenkalkulations-Programm) verwendet werden. Betrachte die gleichen Datenpaare (x_i, y_i) wie in Aufgabe 1 des zweiten Aufgabenzettels:

x_i [cm]	48	49	50	51	52
y_i [Tage]	277, 1	279, 3	281, 4	283, 2	284, 8

Wie in Aufgabe 4 des zweiten Aufgabenzettels bestimmen wir die Regressionsgerade $x = a' + b'y$, dabei ist

$$b' = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum y_i^2 - n \cdot \bar{y}^2},$$

$$a' = \bar{x} - b' \bar{y}.$$

Zusätzlich soll der Korrelations-Koeffizient

$$r_{xy} = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sqrt{\sum x_i^2 - n \cdot \bar{x}^2} \cdot \sqrt{\sum y_i^2 - n \cdot \bar{y}^2}}.$$

bestimmt werden.

2. Bestimme die Regressionsgeraden $y = a + bx$ und $x = a' + b'y$ und den Korrelations-Koeffizienten zu folgender Stichprobe (Winter, p.138): gemessen wurde die Körpergröße x (in cm) und das Körpergewicht y (in kg) von zwölf 18-jährigen Männern

x_i [cm]	170	163	172	160	166	170	174	165	162	167	165	168
y_i [kg]	67, 5	65	80	57, 5	68	74	79	70	56	62, 5	59, 5	66

3. Gegeben seien folgende Datenpaare

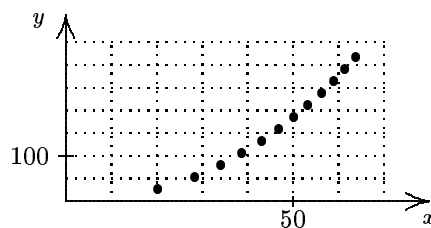
x_i	3	5	6	4	7
y_i	4	5	3	2	6

Durch Hinzufügen zweier Zahlenpaare (x_6, y_6) , (x_7, y_7) soll erreicht werden, daß sich jeweils die Mittelwerte $\bar{x}$ und $\bar{y}$ **nicht** ändern, und daß die Steigung der Regressionsgeraden

- (a) größer wird,
- (b) gleich bleibt,
- (c) kleiner wird.

4. Bestimmung der Beschleunigung einer gleichmäßig beschleunigten Bewegung; zum Beispiel Anfahren eines Schnellzugs (Lambacher-Schweizer 11, p.67). Beim Anfahren eines IC in Köln Hbf wurden von Schülern die Zeiten x_i gestoppt, zu denen die 12 Wagen (Wagenlänge 26,4 m) die Meßstelle (in Höhe des hinteren Puffers der stehenden Lok) passierten. Gegeben sind also für die Zeitpunkte $x_0, x_1, \dots$ in [sek] der zurückgelegte Weg $y_0, y_1, \dots$ in [m], siehe folgende Tabelle und Graphik.

i	0	1	2	3	4	5	6	7	8	9	10	11	12
x_i	0	20,39	28,38	34,14	38,93	43,22	46,94	50,31	53,43	56,35	59,07	61,54	64,01
y_i	0	26,4	52,8	79,2	105,6	132	158,4	184,8	211,2	237,6	264	290,4	316,8



Der vom Zug zurückgelegte Weg scheint quadratisch von der verstrichenen Zeit abzuhängen, also arbeiten wir mit dem Ansatz $y = cx^2$. Um c mit **linearer** Regression zu bestimmen, betrachten wir statt der Funktion cx^2 die Quadratwurzel $z = \sqrt{cx}$. Wir wollen demnach $\sqrt{c}$ (und damit c) durch lineare Regression bestimmen.

- (1) Bilden Sie also eine neue Tabelle mit den Datenpaaren (x_i, z_i) , wobei $z_i = \sqrt{y_i}$ gilt, für $i = 0, 1, \dots, 12$ und berechnen Sie die zugehörige Regressionsgerade $z = a + bx$.
- (2) Welchen Wert erhalten wir demnach für c ? Zeichnen Sie die x - y -Punktwolke zusammen mit der Parabel $y = cx^2$ in ein Koordinatensystem.
- (3) Wann hätte die Lok den Bahnhof Köln-Deutz (Entfernung von Köln Hbf 2 200 m) erreicht, wenn der Zug mit gleicher Beschleunigung weiterfahren könnte?

5. Zusatzaufgabe (freiwillig): Gib jeweils vier Zahlenpaare (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) an, so daß gilt:

- (a) $r_{xy} = 1$.
- (b) $r_{xy} = -1$.
- (c) $r_{xy} = 0$.
- (d) $r_{xy} = \frac{1}{2}$.