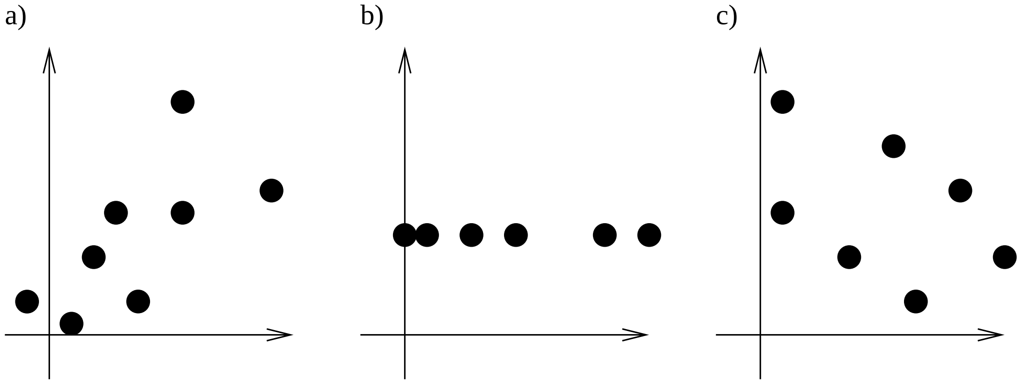
Fachbereich Mathematik
Prof. Dr. Michael Kohler
Dipl.-Math. Andreas Fromkorth
Dipl.-Inf. Jens Mehnert

TECHNISCHE
UNIVERSITÄT
DARMSTADT

WS 08/09
19.11.2008

5. Übungsblatt zur „Statistik I für Human- und Sozialwissenschaft“

Aufgabe 15 (3 Punkte)



Welche Aussage können Sie über die Größe der Korrelation der Datenmengen machen (z.B. $r_{x,y} = -1, -1 < r_{x,y} < 0, r_{x,y} = 0, 0 < r_{x,y} < 1$ oder $r_{x,y} = 1$)? Begründen Sie Ihre Aussage!

Lösung:

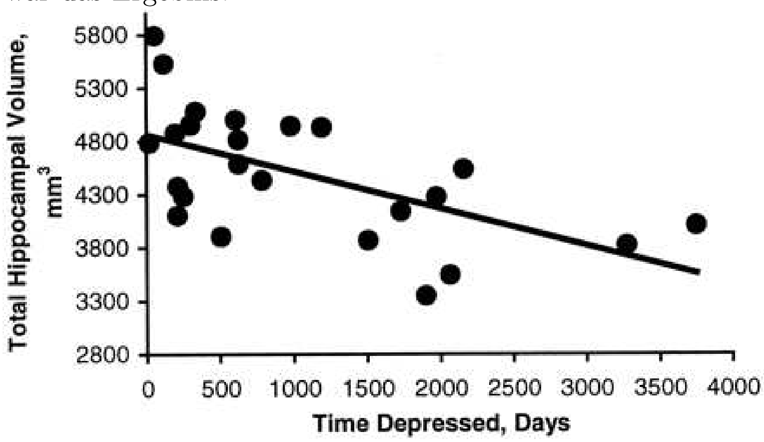
(a) Durch die erste Datenmenge kann man eine steigende Regressionsgerade legen. Daher gilt für die Korrelation $0 < r_{x,y} < 1$.

(b) Durch Datenmenge b) lässt sich eine Gerade legen die alle Datenpunkte enthält. Daher gilt $r_{x,y} = 0$.

(c) Durch Datenmenge c) läßt sich eine fallenden Gerade legen. Daher gilt $-1 < r_{x,y} < 0$.

Aufgabe 16 (3 Punkte)

Eine Gruppe von Versuchspersonen wurde nach der Anzahl von Tagen gefragt, an welchen diese depressiv waren. Zudem wurde ihr Gehirnvolumen gemessen. Folgender Graph mit eingezeichneter Regressionsgerade war das Ergebnis:



(a) Welches Vorzeichen hat die empirische Kovarianz?

(b) Liegt hier die empirische Korelation im Intervall $(0, 1]$ oder im Intervall $[-1, 0)$? Warum kann die empirische Korrelation hier nicht gleich 0 sein?

(c) Wenn wir die Punkte weglassen, welche einen Wert kleiner als 1400 auf der "Time Depressed, Days"-Achse haben, wie sähe dann die Regressionsgerade bzw. das Vorzeichen der empirische Korrelation aus?

Lösung:

(a) Negativ, da die Regressionsgerade fallend ist.

(b) Im Intervall $[-1, 0)$, da die Regressionsgerade fallend ist. Die Regressionsgerade ist fallend, womit die Korrelation nicht 0 sein kann.

(c) Die Regressionsgerade wäre nahezu waagerecht und die Korrelation nahezu 0.

Aufgabe 17 (3 Punkte)

Ein Lehrer stuft die Aufsätze seiner 6 Schüler auf einer Skala von 1 bis 5 ein, ob das Thema eher kreativ oder weniger kreativ bearbeitet wurde. Ferner bringt er die Schüler nach ihren allgemeinen Leistungen im Deutschunterricht in eine Rangreihe.

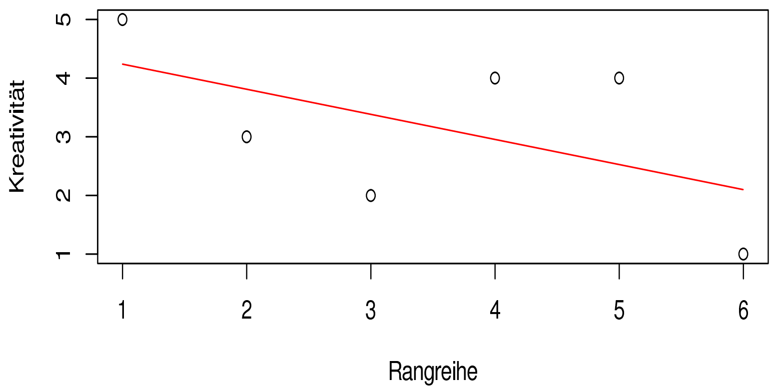
Schüler Nr.	Kreativität (x)	Deutschleistung (y)
1	2	3
2	4	4
3	5	1
4	4	5
5	1	6
6	3	2

(a) Zeichnen Sie einen Scatterplot nach den letzten beiden Spalten der Tabelle.

(b) Berechnen Sie die empirische Kovarianz sowie die empirische Korrelation.

(c) Berechnen Sie die Regressionsgerade mit Hilfe der Formel aus der Vorlesung. Zeichnen Sie die Regressionsgerade in den Scatterplot.

Lösung:



(a)

(b) Die Krorrelation $r_{x,y}$ und die kovarianz $s_{x,y}$ ist

$$r_{x,y} = \frac{s_{x,y}}{s_x \cdot s_y} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})}\cdot\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})}} = \frac{-1{,}5}{2{,}75} \approx -0{,}55.$$

(c) Es gilt $\bar{x} = \frac{1}{n}\sum x_i = \frac{19}{6} \approx 3,17$, $\bar{y} = \frac{7}{2} = 3,5$ und $\hat{a} = \frac{s_{x,y}}{s_x^2} = \frac{-7{,}5}{10{,}83} = -0,69$, womit $y = -0.69 \cdot (x - 3,17) + 3,5$ ist.

Aufgabe 18 (3 Punkte)

Rechnen Sie folgende Summen aus:

a) $\sum_{i=1}^{4} i$ b) $\sum_{0\leq k\leq 3} (k+\frac{1}{2})$ c) $\sum_{x=2}^{5} x^2$

d) $\sum_{k=0}^{2} 2k$ e) $\sum_{k=1}^{2} \frac{1}{k}$ f) $\sum_{k=1}^{2} \frac{1}{2^k}$

Lösung:

- a) $\sum_{i=1}^{4} i = 1+2+3+4 = 10$
- b) $\sum_{0\leq k\leq 3} (k+\frac{1}{2}) = \sum_{k=0}^{3} (k+\frac{1}{2}) = (0+\frac{1}{2}) + (1+\frac{1}{2}) + (2+\frac{1}{2}) + (3+\frac{1}{2}) = 6 + 4\cdot\frac{1}{2} = 8$
- c) $\sum_{x=2}^{5} x^2 = 2^2+3^2+4^2+5^2 = 4+9+16+25 = 54$
- d) $\sum_{k=0}^{2} f(k) = 0+2+4 = 6$
- e) $\sum_{k=1}^{2} \frac{1}{k} = \frac{1}{1}+\frac{1}{2} = 1,5$
- f) $\sum_{k=1}^{2} \frac{1}{2^k} = \frac{1}{2}+\frac{1}{4} = \frac{3}{4} = 0,75$

Abgabe der Übung: Eine Woche nachdem das Übungsblatt zu Ihrem Übungstermin bearbeitet wurde, zu Beginn der nächsten Übung bei Ihrer Übungsgruppenleiterin oder bei Ihrem Übungsgruppenleiter.