

## Übungen zur Regression und Korrelation

### Aufgabe 1

Der Zusammenhang zwischen den beiden Merkmalen X und Y wird untersucht. Zu diesem Zweck werden 6 Wertepaare  $(x_i, y_i)$  betrachtet, die der nachfolgenden Tabelle zu entnehmen sind.

$x_i$	2	4	5	5	6	8
$y_i$	10	5	4	8	7	2

- Bestimmen Sie die arithmetischen Mittelwerte  $\bar{X}$ ,  $\bar{Y}$ ,  $\overline{X^2}$ ,  $\overline{Y^2}$  und  $\overline{XY}$ .
- Bestimmen Sie die Varianzen  $\sigma_x^2$  und  $\sigma_y^2$  sowie die Standardabweichungen  $\sigma_x$  und  $\sigma_y$ .
- Bestimmen Sie die Kovarianz und den Korrelationskoeffizienten.
- Bestimmen Sie die Gleichung der linearen Regressionsgeraden.

### Aufgabe 2

Der Zusammenhang zwischen den beiden Merkmalen X und Y wird untersucht. Zu diesem Zweck werden 20 Wertepaare  $(x_i, y_i)$  betrachtet, die der nachfolgenden Tabelle zu entnehmen sind.

$x_i$	1	2	3	3	3	4	5
$y_i$	5	3	1	3	4	2	1
$N_i$	2	3	2	4	2	4	3

- Bestimmen Sie die arithmetischen Mittelwerte  $\bar{X}$ ,  $\bar{Y}$ ,  $\overline{X^2}$ ,  $\overline{Y^2}$  und  $\overline{XY}$ .
- Bestimmen Sie die Varianzen  $\sigma_x^2$  und  $\sigma_y^2$  sowie die Standardabweichungen  $\sigma_x$  und  $\sigma_y$ .
- Bestimmen Sie die Kovarianz und den Korrelationskoeffizienten.
- Bestimmen Sie die Gleichung der linearen Regressionsgeraden.

### Aufgabe 3

Der Zusammenhang zwischen dem Merkmal X (Körpergröße eines Neugeborenen in cm) und dem Merkmal Y (Körpergewicht eines Neugeborenen in kg) wird untersucht. Zu diesem Zweck werden 8 Neugeborene bei ihrer Geburt gemessen und gewogen. Die daraus resultierenden Wertepaare  $(x_i, y_i)$  können der nachfolgenden Tabelle entnommen werden.

$x_i$	44	47	45	49	50	53	46	50
$y_i$	5,1	5,7	5,4	6,1	6,2	6,6	5,3	6,1

- Bestimmen Sie die arithmetischen Mittelwerte  $\bar{X}$ ,  $\bar{Y}$ ,  $\overline{X^2}$ ,  $\overline{Y^2}$  und  $\overline{XY}$ .
- Bestimmen Sie die Varianzen  $\sigma_x^2$  und  $\sigma_y^2$  sowie die Standardabweichungen  $\sigma_x$  und  $\sigma_y$ .
- Bestimmen Sie die Kovarianz und den Korrelationskoeffizienten.
- Bestimmen Sie die Gleichung der linearen Regressionsgeraden.
- Schätzen Sie aufgrund der linearen Regressionsgeraden aus d) das Körpergewicht eines Neugeborenen mit einer Körpergröße von 58 cm.

#### Aufgabe 4

Sie sind als Geschäftsführer eines Unternehmens für die Preisgestaltung zuständig und wollen den Zusammenhang zwischen dem Merkmal X (Absatzmenge des betreffenden Produkts in ME) und dem Merkmal Y (Absatzpreis des betreffenden Produkts in GE / ME) möglichst genau durch eine Preis-Absatz-Funktion beschreiben. Diese Preis-Absatz-Funktion kann anschließend zur Grundlage einer gewinnmaximierenden Absatzpolitik des Unternehmens gemacht werden und Ihnen somit viele neue Freunde unter den Aktionären verschaffen. Sie unterstellen eine lineare Beziehung zwischen der Absatzmenge und dem Absatzpreis des betreffenden Produkts. Zur Quantifizierung des linearen Zusammenhangs haben Sie in den vergangenen Wochen an 8 verschiedenen Tagen Wertepaare  $(x_i, y_i)$  beobachtet, die der Tabelle entnommen werden können.

i	$x_i$	$y_i$	$x_i - \bar{X}$	$(x_i - \bar{X})^2$	$y_i - \bar{Y}$	$(y_i - \bar{Y})^2$	$(x_i - \bar{X}) \cdot (y_i - \bar{Y})$
1	21	88					
2	23	86					
3	24	81					
4	26	74					
5	31	62					
6	34	59					
7	39	58					
8	42	52					
$\Sigma$							
$\frac{1}{N} \cdot \Sigma$							

- Vervollständigen Sie die angegebene Tabelle.
- Benutzen Sie die vervollständigte Tabelle aus a), um die arithmetischen Mittelwerte, Varianzen und Standardabweichungen der Merkmale X und Y anzugeben.
- Benutzen Sie die vervollständigte Tabelle aus a), um die Kovarianz anzugeben und bestimmen Sie auch den Korrelationskoeffizienten. Welche Aussagekraft hat der Korrelationskoeffizient im vorliegenden Kontext? Welche Richtung und Intensität des linearen Zusammenhangs zwischen den Merkmalen X und Y können Sie aus dem Wert des Korrelationskoeffizienten ableiten?
- Bestimmen Sie zunächst die Gleichung der Regressionsgeraden. Verwenden Sie anschließend die ermittelte Regressionsgerade um zu schätzen, welcher Absatzpreis bei einer Absatzmenge von 55 ME zu erwarten ist.

#### Aufgabe 5

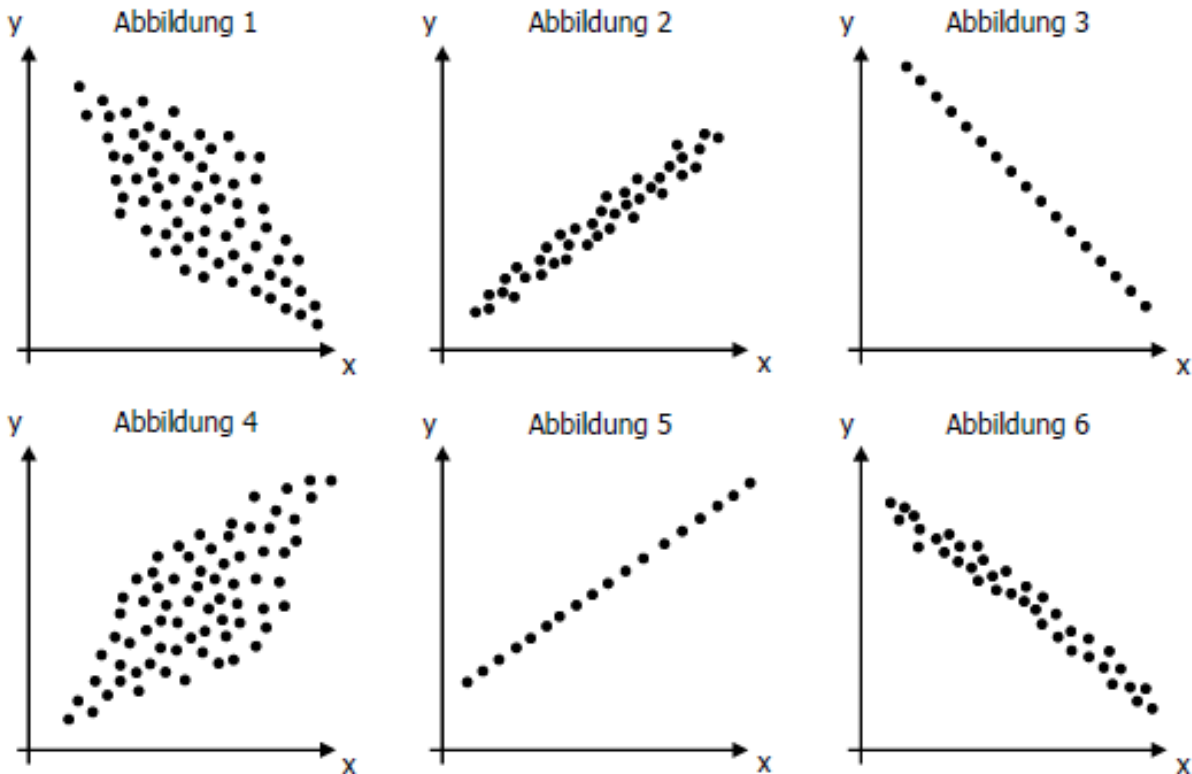
Der Zusammenhang zwischen den beiden Merkmalen X und Y wird untersucht. Zu diesem Zweck werden 10 Wertepaare  $(x_i, y_i)$  betrachtet, die der nachfolgenden Tabelle zu entnehmen sind.

$x_i$	7,8	8,2	9,4	9,9	10,8	12,1	13,3	13,7	15,2	16,6
$y_i$	8,1	7,4	6,8	6,2	5,9	5,6	4,9	4,2	3,6	3,3

- Bestimmen Sie die arithmetischen Mittelwerte  $\bar{X}$ ,  $\bar{Y}$ ,  $\overline{X^2}$ ,  $\overline{Y^2}$  und  $\overline{XY}$ .
- Bestimmen Sie die Varianzen  $\sigma_x^2$  und  $\sigma_y^2$  sowie die Standardabweichungen  $\sigma_x$  und  $\sigma_y$ .
- Bestimmen Sie die Kovarianz und den Korrelationskoeffizienten.
- Bestimmen Sie die Gleichung der linearen Regressionsgeraden.

## Aufgabe 6

Betrachten Sie die folgenden Streudiagramme der Merkmale X und Y in den Abbildungen 1 bis 6.



Betrachten Sie darüber hinaus die folgenden "verbale" Aussagen A bis F.

- A Zwischen X und Y besteht ein schwacher positiver linearer Zusammenhang.
- B Zwischen X und Y besteht ein schwacher negativer linearer Zusammenhang.
- C Zwischen X und Y besteht ein starker negativer linearer Zusammenhang.
- D Zwischen X und Y besteht ein perfekter negativer linearer Zusammenhang.
- E Zwischen X und Y besteht ein perfekter positiver linearer Zusammenhang.
- F Zwischen X und Y besteht ein starker positiver linearer Zusammenhang.

Betrachten Sie schließlich die folgenden "formalen" Aussagen G bis L.

- G Der Korrelationskoeffizient zwischen X und Y nimmt den Wert  $-1$  an.
- H Der Korrelationskoeffizient zwischen X und Y ist sehr nahe bei  $-1$ , aber nicht gleich  $-1$ .
- I Der Korrelationskoeffizient zwischen X und Y ist negativ, aber nicht nahe bei  $-1$ .
- J Der Korrelationskoeffizient zwischen X und Y ist positiv, aber nicht nahe bei  $1$ .
- K Der Korrelationskoeffizient zwischen X und Y ist sehr nahe bei  $1$ , aber nicht gleich  $1$ .
- L Der Korrelationskoeffizient zwischen X und Y nimmt den Wert  $1$  an.

Vervollständigen Sie die angegebene Tabelle, indem Sie zu jeder Abbildung (Streudiagramm) genau eine korrekte "verbale" Aussage und zudem genau eine zutreffende "formale" Aussage zuordnen. Jede der Aussagen A bis L muss dabei genau einmal zugeordnet werden.

Abbildung (Streudiagramm)	1	2	3	4	5	6
"Verbale" Aussage						
"Formale" Aussage						

## Aufgabe 7

Ein Schalke-Fan und ein Dortmund-Fan müssen die nachfolgend aufgelisteten 9 Musikalben aus dem 20. Jahrhundert bewerten und jeweils getrennt eine persönliche Rangordnung erstellen.

- A Cat Stevens (1970): "Tea for the Tillerman"
- B The Who (1969): "Tommy"
- C Bob Dylan (1965): "Highway 61 Revisited"
- D The Beatles (1967): "Sgt. Pepper's Lonely Hearts Club Band"
- E David Bowie (1972): "The Rise and Fall of Ziggy Stardust and the Spiders from Mars"
- F Pink Floyd (1973): "Dark Side of the Moon"
- G The Rolling Stones (1968): "Beggars Banquet"
- H U2 (1991): "Achtung Baby"
- I The Doors (1971): "L.A. Woman"

Nach einer Befragung der beiden Fußball-Fans können die Rangordnungen angegeben werden.

Rangfolge Schalke-Fan:                    B I G C F D A H E

Rangfolge Dortmund-Fan:                H C A I B G F E D

Führen Sie einen Korrelationsvergleich mit Hilfe des Spearmanschen Rangkorrelationskoeffizienten durch, wobei Sie die Gemeinsamkeiten im Musikgeschmack der Fußball-Fans untersuchen.

## Aufgabe 8

Insgesamt 4 namhafte Kunstkritiker müssen die nachfolgend aufgelisteten 7 berühmten Bilder bewerten und jeweils getrennt eine persönliche Rangordnung erstellen.

- A Pablo Picasso (1937): "Guernica"
- B Michelangelo (1508 bis 1512): "Die Erschaffung Adams"
- C Vincent van Gogh (1888): "Das Nachtcafé"
- D Salvador Dali (1931): "Die Beständigkeit der Erinnerung"
- E Edvard Munch (1892 bis 1895): "Der Schrei"
- F Leonardo da Vinci (1495 bis 1498): "Das Abendmahl"
- G Rembrandt (1632): "Die Anatomie des Dr. Tulp"

Nach einer Befragung der Kunstkritiker können die Rangordnungen angegeben werden.

Rangfolge Kunstkritiker 1:                F D B A G C E

Rangfolge Kunstkritiker 2:                D F A C G E B

Rangfolge Kunstkritiker 3:                C E B D F A G

Rangfolge Kunstkritiker 4:                E D A B G C F

Führen Sie zwei Korrelationsvergleiche mit Hilfe des Spearmanschen Rangkorrelationskoeffizienten durch, wobei Sie in einem 1. Vergleich die Gemeinsamkeiten von Kunstkritiker 1 und 2, sowie in einem 2. Vergleich die Gemeinsamkeiten von Kunstkritiker 3 und 4 untersuchen. Geben Sie für beide Vergleiche den Spearmanschen Rangkorrelationskoeffizienten an. Bei welchem Vergleich nimmt der Rangkorrelationskoeffizient den größeren Wert an und welche Aussagekraft haben die ermittelten Werte für die Rangkorrelationskoeffizienten?