

## Übung 6

1. Eine Studie wurde durchgeführt, um die Wirksamkeit von Aspirin bei der Behandlung von Hirnschlag zu untersuchen. Patienten wurden rein zufällig in eine Behandlungs- und eine Kontrollgruppe (d.h. mit Placebo behandelt) eingeteilt. Nach einer sechsmonatigen Behandlung wurden die Patienten beurteilt und deren Entwicklung als günstig oder ungünstig bezeichnet.

Von den 78 Patienten in der Aspirin-Gruppe hatten 63 eine günstige Entwicklung, von den 77 Patienten in der Kontrollgruppe hatten 43 eine günstige Entwicklung.

- a) Stelle die Daten in der Form einer Vierfeldertafel dar.
- b) Führe auf dem 5%-Niveau einen geeigneten Test durch, um zu entscheiden, ob es einen signifikanten Unterschied zwischen der Behandlungs- und der Kontrollgruppe gibt.

2. Im untenstehenden Streudiagramm sind die beiden Variablen

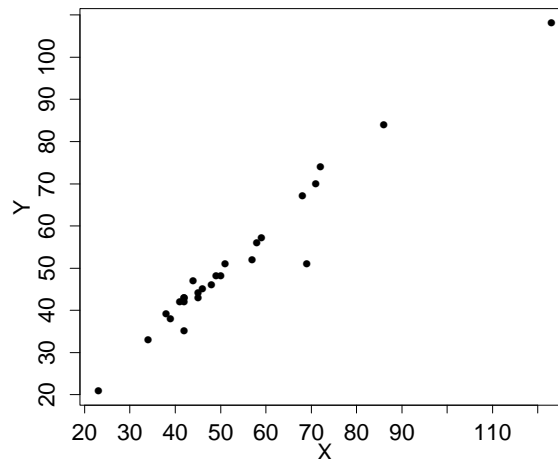
$X_i$  = Partikelkonzentration in der Stadt

$Y_i$  = Partikelkonzentration 16 km ausserhalb der Stadt

an 26 verschiedenen Tagen dargestellt.

$$\bar{x} = 53.23, \quad \sum_{i=1}^{26} (x_i - \bar{x})^2 = 9852.62 \quad \bar{y} = 51.04, \quad \sum_{i=1}^{26} (y_i - \bar{y})^2 = 7836.96$$

$$\sum_{i=1}^{26} (y_i - \bar{y})(x_i - \bar{x}) = 8569.77$$



- a) Welche Voraussage würdest Du für die Partikelkonzentration ausserhalb der Stadt machen, wenn der Wert innerhalb der Stadt 80 beträgt?
- b) Beeinflusst die Partikelkonzentration innerhalb der Stadt die Partikelkonzentration ausserhalb der Stadt signifikant?

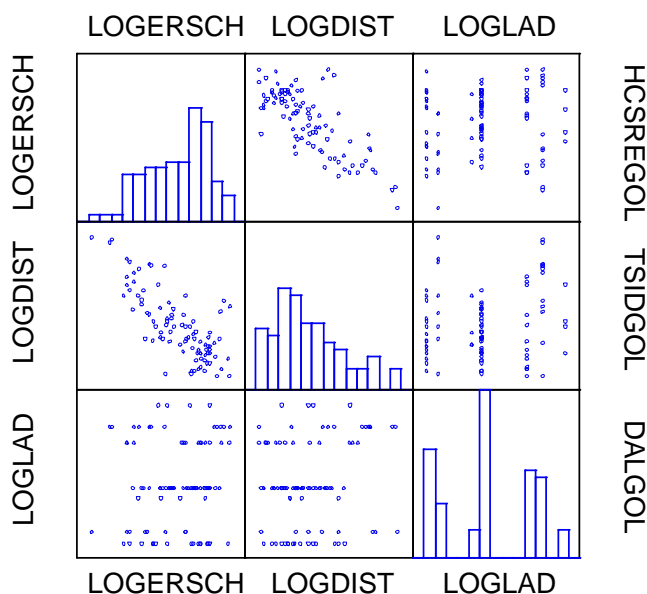
3. Beim Bau eines Strassentunnels zur Unterfahrung der Stadt Schaffhausen wurden im Siedlungsgebiet Sprengungen durchgeführt. Bei der Durchführung der Sprengarbeiten musste darauf geachtet werden, dass die Erschütterung der Häuser einen bestimmten Wert nicht überschritt.

Die Erschütterung der Häuser bei einer Sprengung ist abhängig von der Sprengladung, von der Distanz zwischen dem Spreng- und dem Messort sowie von weiteren, hier nicht näher betrachteten Gegebenheiten.

Im folgenden interessieren wir uns für die Abhängigkeit der Erschütterung (LOGERSCH) von der Distanz zur Sprengladung (LOGDIST) und von der Stärke der Sprengladung (LOGLAD). Die Messgrößen wurden alle logarithmiert. Wir betrachten das Modell

$$\text{LOGERSCH}_i = \alpha + \beta_1 \cdot \text{LOGDIST}_i + \beta_2 \cdot \text{LOGLAD}_i + \epsilon_i \quad \text{mit } \epsilon_i \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.}$$

Im folgenden sind die paarweisen Streudiagramme der betrachteten Größen dargestellt.



Die Analyse mit Systat ergibt folgenden Output:

```

Dep Var: LOGERSCH  N: 93  Multiple R: 0.81  Squared multiple R: 0.65
Adjusted squared multiple R: 0.64  Standard error of estimate: 0.47
Effect      Coefficient  Std Error  Std Coef  Tolerance  t      P(2 Tail)
CONSTANT    6.53           0.54       0.0       .          12.01  0.00
LOGDIST     -1.57          0.12       -0.81     0.97      -12.84  0.00
LOGLAD      1.07           0.29       0.23     0.97       3.65   0.00

```

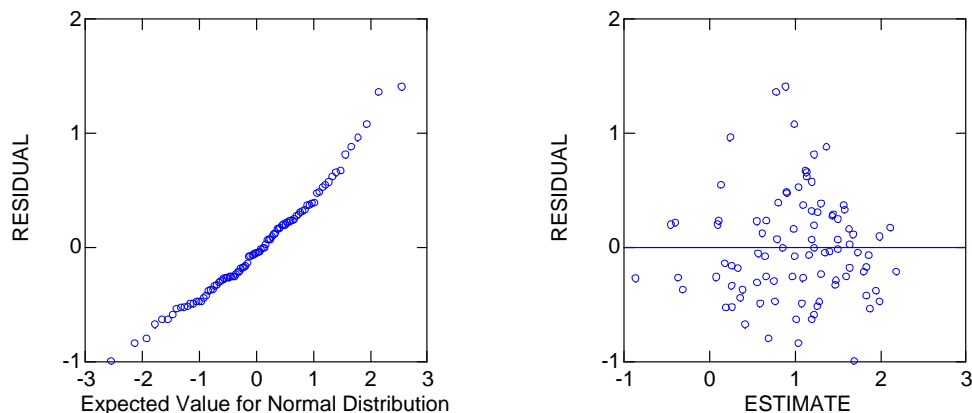
```

              Analysis of Variance
Source        Sum-of-Squares  df  Mean-Square  F-ratio  P
Regression    36.82           2   18.41        83.45    0.00
Residual      19.86           90   0.22

```

- Kommentiere die paarweisen Streudiagramme im Hinblick auf die Abhängigkeit der Erschütterung von der Distanz zur Sprengladung und von der Stärke der Sprengladung.
- Lies die Gleichung des Regressionsmodelles aus dem Systat-Output und schreibe sie auf.
- Mit welchem Mass beurteilst Du die Güte der Modellanpassung? Wie gross ist dieses Mass hier? Passt das Modell gut?

- d) Wir untersuchen im Folgenden den Einfluss einer einzelnen erklärenden Variablen (unter Berücksichtigung des Einflusses der anderen erklärenden Variablen). Hat die Distanz gemäss obigen Output einen signifikanten Einfluss auf die Erschütterung? Die Stärke der Sprengladung? Woraus siehst Du das?
- e) Zur Überprüfung der Modellvoraussetzungen erhältst Du einen Normalplot der Residuen sowie einen Tukey-Anscombe-Plot vorgelegt. Gibt es Hinweise auf die Verletzung von Modellvoraussetzungen. Falls ja, welche?  
*(Anmerkung: Der Tukey-Anscombe-Plot ist nichts anderes, als das in Kapitel 11.3 vorgestellte Verfahren, wobei die Residuen gegen die berechneten y-Werte aufgetragen werden)*



4. Leonardo da Vinci sammelte folgende Daten über die Flügelspannweite und die Körpermasse von Vögeln. Die Grössen sind in den heutigen Einheiten Meter und Kilogramm angegeben:

Vogelart	Masse $X_i$ [kg]	Spannweite $Y_i$ [m]
Amsel	0.32	0.17
Eichelhäher	0.48	0.42
Blesshuhn	0.95	0.92
Stockente	1.10	1.95
Graugans	1.85	4.80
Storch	1.95	6.60

(Daten aus *Gerthsen Physik, 18. Auflage*)

- a) Trage die Werte  $X_i$ ,  $Y_i$  in ein Koordinatensystem ein und berechne die Regressionsgerade  $y = a + b \cdot x$ .
- b) Zeichne die Regressionsgerade aus a) in das Bild ein und überlege, ob dies ein sinnvolles Modell ist. Gäbe es eine Kurve, die besser zu den Daten passt? Skizziere sie von Hand.
- c) Trage nun die Werte  $Z_i = \log_{10}(Y_i)$  gegen  $X_i$  auf, und berechne nochmals die Regressionsgerade  $z = c + d \cdot x$ . Ist dieses Modell besser?
- d) Welche Flügelspannweite bräuchte ein 80kg schwerer Mensch? (Zum Vergleich: Die Boeing B747-436 "City of London" der British Airways hat eine Spannweite von 64.92m und kann ein maximales Startgewicht von 396'893kg haben).

**Vorbesprechung:** Freitag, 23.01.2004 in der Übungsstunde

**Abgabe:** Donnerstag, 29.01.2004 in der Vorlesung

**Präsenz:** Montag, 12-13 Uhr im LEO C12.1