

# Mathematikaufgaben

## > Daten, Statistik

## > Korrelation, lineare Regression

**Aufgabe:** In einem elektrischen Stromkreis gilt das Ohmsche Gesetz  $R = \frac{U}{I}$  mit Widerstand R ( $\Omega$  = Ohm), Spannung U (V = Volt) und Stromstärke I (A = Ampere). Spannung und Stromstärke werden gemessen, die Paare der Messwerte lauten:

|                      |     |     |     |     |     |      |
|----------------------|-----|-----|-----|-----|-----|------|
| Spannung (Volt)      | 5   | 10  | 20  | 50  | 100 | 200  |
| Stromstärke (Ampere) | 0.4 | 1.1 | 2.0 | 4.5 | 8.3 | 16.7 |

Berechne Korrelationskoeffizient und Bestimmtheitsmaß. Ermittle die Gleichung der Geraden, die als lineare Regression den Zusammenhang zwischen Spannung und Stromstärke mit dem Widerstand als Proportionalitätsfaktor wiedergibt.

**Lösung:** I. Allgemein gilt: Mit Daten und Statistik verbunden sind die zahlenmäßige Erfassung von Sachverhalten und Beobachtungen und deren grafische Darstellung. Daten ergeben damit ein mathematisches Modell, das mit Hilfe mathematischer Methoden analysiert und bewertet werden kann. Sind zwei Zufallsvariablen X, Y gegeben, so lässt sich ein stochastischer Zusammenhang (Korrelation) zwischen diesen Zufallsgrößen nachweisen mittels des sog. (Bravais-Pearson-) Korrelationskoeffizienten r, der als Maßzahl ( $-1 \leq r \leq 1$ ) Stärke und Richtung einer linearen Beziehung zwischen X und Y charakterisiert. Der Korrelationskoeffizient errechnet sich als:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{bzw.} \quad r = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sqrt{\sum_{i=1}^n x_i^2 - n \bar{x}^2} \cdot \sqrt{\sum_{i=1}^n y_i^2 - n \bar{y}^2}},$$

wobei die Wertepaare  $(x_i, y_i)$  Stichprobenpaare von Ausprägungen  $x_i, y_i$  der Zufallsvariablen X und Y sind,  $i = 1, 2, \dots, n$ . Ausprägungen sind Werte eines Merkmals; Merkmale sind geeignete, charakteristische Eigenschaften eines Untersuchungsobjekt innerhalb der Statistik und können als Zufallsvariablen aufgefasst werden.

Mit dem Korrelationskoeffizienten lässt sich die Stärke des stochastischen Zusammenhangs zwischen den Zufallsvariablen wie folgt einschätzen:

| Zusammenhang | Korrelationskoeffizient |
|--------------|-------------------------|
| kein         | $r = 0$                 |
| gering       | $0 <  r  \leq 0,3$      |
| schwach      | $0,3 <  r  \leq 0,5$    |
| mittel       | $0,5 <  r  \leq 0,8$    |
| stark        | $0,8 <  r  < 1$         |
| funktional   | $ r  = 1$               |

Das Bestimmtheitsmaß  $R^2$  errechnet sich als Quadrat des Korrelationskoeffizienten  $r$  mit:  $R^2 = r^2$  ( $0 \leq R^2 \leq 1$ ) und charakterisiert damit die Größe der Abhängigkeit zwischen den Zufallsvariablen  $X$  und  $Y$ .

II. Unter der Voraussetzung einer z.B. starken Korrelation  $r$  zwischen zwei Merkmalen  $X$  und  $Y$  mit deren jeweiligen Ausprägungen  $x_i, y_i, i = 1, 2, \dots, n$  gibt die lineare Regression einen linearen Zusammenhang zwischen den vorgegebenen Zufallsvariablen  $X$  als Einfluss- und  $Y$  als Zielgröße wieder. Es entsteht eine Regressionsgerade  $y = a + bx$  durch die Wertepaare  $(x_i, y_i), i = 1, 2, \dots, n$ ,

auch auf der Grundlage der Mittelwerte  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  vermöge der Berechnungen:

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{bzw.} \quad b = \frac{n \cdot \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \cdot \left( \sum_{i=1}^n y_i \right)}{n \cdot \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad (\text{Geradensteigung})$$

$$a = \bar{y} - b \cdot \bar{x} \quad \text{bzw.} \quad a = \frac{\left( \sum_{i=1}^n x_i^2 \right) \cdot \left( \sum_{i=1}^n y_i \right) - \left( \sum_{i=1}^n x_i \right) \cdot \left( \sum_{i=1}^n x_i y_i \right)}{n \cdot \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad (\text{y-Achsenabschnitt der Geraden}).$$

III. Korrelationskoeffizient und Regressionsgerade in Zusammenhang mit den Merkmalen Stromstärke und Spannung (gemäß dem Ohmschen Gesetz  $R = \frac{U}{I} \Rightarrow U = RI$ ) berechnen sich (wahlweise) über die folgenden Tabellen:

Merkmal Stromstärke (Ampere)  $[x]$ , Merkmal Spannung (Volt)  $[y]$ ; Tabelle (I):

| i                                                         | Wert $x_i$ | Wert $y_i$ | Differenz $x_i - \bar{x}$ | Quadrat $(x_i - \bar{x})^2$ | Differenz $y_i - \bar{y}$ | Quadrat $(y_i - \bar{y})^2$ | Produkt $(x_i - \bar{x}) \cdot (y_i - \bar{y})$ |
|-----------------------------------------------------------|------------|------------|---------------------------|-----------------------------|---------------------------|-----------------------------|-------------------------------------------------|
| 1                                                         | 0.4        | 5          | -5.1                      | 26.01                       | -59.1667                  | 3500.6944                   | 301.75                                          |
| 2                                                         | 1.1        | 10         | -4.4                      | 19.36                       | -54.1667                  | 2934.0278                   | 238.3333                                        |
| 3                                                         | 2.0        | 20         | -3.5                      | 12.25                       | -44.1667                  | 1950.6944                   | 154.5833                                        |
| 4                                                         | 4.5        | 50         | -1                        | 1                           | -14.1667                  | 200.6944                    | 14.1667                                         |
| 5                                                         | 8.3        | 100        | 2.8                       | 7.84                        | 35.8333                   | 1284.0278                   | 100.3333                                        |
| 6                                                         | 16.7       | 200        | 11.2                      | 125.44                      | 135.8333                  | 18450.6944                  | 1521.3333                                       |
| <b>Summe:</b>                                             | 33         | 385        | -                         | 191.9                       | -                         | 28320.8333                  | 2330.5                                          |
| <b>Mittelwert:</b>                                        | 5.5        | 64.1667    | -                         | -                           | -                         | -                           | -                                               |
| <b>Standardabweichung:</b>                                |            |            | -                         | 5.6554                      | -                         | 68.7033                     | -                                               |
| <b>Regressionskoeffizient <math>r =</math></b>            |            |            |                           |                             |                           |                             | 0.9997                                          |
| <b>Steigung/Regressionsgerade <math>b =</math></b>        |            |            |                           |                             |                           |                             | 12.1443                                         |
| <b>Achsenabschnitt/Regressionsgerade <math>a =</math></b> |            |            |                           |                             |                           |                             | -2.6272                                         |
| <b>Bestimmtheitsmaß <math>R^2 =</math></b>                |            |            |                           |                             |                           |                             | 0.9993                                          |

Merkmal Stromstärke (Ampere) [x], Merkmal Spannung (Volt) [y]; Tabelle (II):

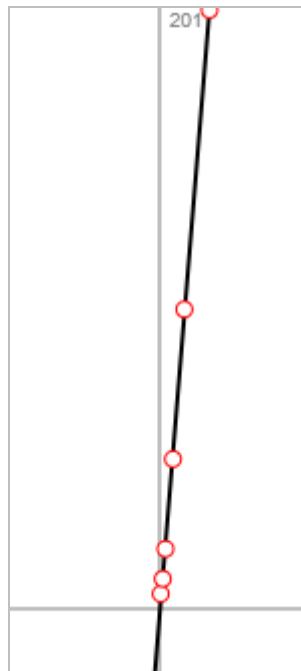
| i                                                         | Wert<br>$x_i$ | Quadrat<br>$x_i^2$ | Wert<br>$y_i$ | Quadrat<br>$y_i^2$ | Produkt<br>$x_i \cdot y_i$ |
|-----------------------------------------------------------|---------------|--------------------|---------------|--------------------|----------------------------|
| 1                                                         | 0.4           | 0.16               | 5             | 25                 | 2                          |
| 2                                                         | 1.1           | 1.21               | 10            | 100                | 11                         |
| 3                                                         | 2.0           | 4                  | 20            | 400                | 40                         |
| 4                                                         | 4.5           | 20.25              | 50            | 2500               | 225                        |
| 5                                                         | 8.3           | 68.89              | 100           | 10000              | 830                        |
| 6                                                         | 16.7          | 278.89             | 200           | 40000              | 3340                       |
| <b>Summe:</b>                                             | 33            | 373.4              | 385           | 53025              | 4448                       |
| <b>Mittelwert:</b>                                        | 5.5           | -                  | 64.1667       | -                  | -                          |
| <b>Regressionskoeffizient <math>r =</math></b>            |               |                    |               |                    | 0.9997                     |
| <b>Steigung/Regressionsgerade <math>b =</math></b>        |               |                    |               |                    | 12.1443                    |
| <b>Achsenabschnitt/Regressionsgerade <math>a =</math></b> |               |                    |               |                    | -2.6272                    |
| <b>Bestimmtheitsmaß <math>R^2 =</math></b>                |               |                    |               |                    | 0.9993                     |

Mit  $r = 0,9997$  als Korrelationskoeffizienten liegt damit ein starker, schon durch physikalische Überlegungen begründeter Zusammenhang zwischen Spannung und Stromstärke vor. Die Regressionsgerade lautet:

$$y = -2.6272 + 12.1443x \quad (x \text{ [A]}, y \text{ [V]}).$$

D.h.: Der Widerstand im Stromkreis beträgt:  $R \approx 12 \, \Omega$ .

IV. Graph von Regressionsgerade und Punktwolke ergeben nun folgendes Bild:



www.michael-buhlmann.de / 07.2025 / Aufgabe 2473