# Statistik und Wahrscheinlichkeit

für Ingenieurstudiengänge

Matthias Künzer

Universität Stuttgart

19. Mai 2025

## Inhalt

# Kapitel 1

# Deskriptive Statistik

## 1.1 Arithmetisches Mittel und Median

Seien $x_1, \dots, x_n \in \mathbb{R}$ gegeben.

**Definition.** Das *arithmetische Mittel* von $x_1, \dots, x_n$, auch *Mittelwert* genannt, ist

$$\overline{x} \;:=\; \frac{1}{n}\sum_{j=1}^{n} x_j \;=\; \frac{1}{n}(x_1 + \ldots + x_n)\;.$$

**Bemerkung.** Es gilt

$$\sum_{j=1}^{n}(x_j - \overline{x}) \;=\; 0\;.$$

Diese Gleichung nennt man auch *Schwerpunkteigenschaft*.

Denn es ist

$$\sum_{j=1}^{n}(x_j - \overline{x}) \;=\; \left(\sum_{j=1}^{n} x_j\right) - n\cdot\overline{x} \;=\; \left(\sum_{j=1}^{n} x_j\right) - n\cdot\frac{1}{n}\sum_{j=1}^{n} x_j \;=\; 0\;.$$

Falls man Ausreißer nicht einrechnen möchte, kann man auch wie folgt den "Wert an der mittleren Stelle" betrachten, genannt Median.

**Definition.** Seien $x_1 \leqslant x_2 \leqslant x_3 \leqslant \ldots \leqslant x_n$ angeordnet.

Der *Median* von $x_1, \dots, x_n$ ist

$$x_{\text{med}} \;:=\; \begin{cases} x_{\frac{n+1}{2}} & \text{falls } n \text{ ungerade} \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right) & \text{falls } n \text{ gerade}\;. \end{cases}$$

**Beispiel.**

(1). Sei $x_1 = 1$, $x_2 = 3$, $x_3 = 11$.

Das arithmetische Mittel ist

$$\overline{x} \;=\; \tfrac{1}{3}(x_1 + x_2 + x_3) \;=\; \tfrac{1}{3}(1 + 3 + 11) \;=\; 5\;.$$

Es ist $n = 3$ ungerade. Der Median ist also

$$x_{\mathrm{med}} \;=\; x_{\frac{3+1}{2}} \;=\; x_2 \;=\; 3\;.$$

(2) Sei $x_1 = 1$, $x_2 = 3$, $x_3 = 10$, $x_4 = 11$.

Es ist $n = 4$ gerade. Der Median ist also

$$x_{\mathrm{med}} \;=\; \tfrac{1}{2}(x_{\frac{4}{2}} + x_{\frac{4}{2}+1}) \;=\; \tfrac{1}{2}(x_2 + x_3) \;=\; 6{,}5\;.$$

## 1.2 Standardabweichung

Seien $x_1, \ldots, x_n \in \mathbb{R}$ gegeben.

Wir erinnern an den Mittelwert $\overline{x}$ von $x_1, \ldots, x_n$ aus §1.1.

**Definition.** Die *Varianz* ist

$$\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2\;.$$

Die *Standardabweichung* ist die Wurzel der Varianz, also

$$\sigma \;:=\; \left(\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2\right)^{1/2} \;=\; \sqrt{\frac{1}{n}\big((x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2\big)}\;.$$

Falls $x_j > 0$ für $1 \leqslant j \leqslant n$, dann definiert man auch den *Variationskoeffizienten* $v := \frac{\sigma}{\overline{x}}$. Dieser ist dimensionslos und maßstabsunabhängig.

Die Standardabweichung ist ein Maß für die Streuung der Werte $x_1, \ldots, x_n$ um ihren Mittelwert $\overline{x}$. Je größer die Standardabweichung, desto breiter gestreut sind die Werte. Umgekehrt ist die Standardabweichung gleich 0 genau dann, wenn alle Werte $x_1, \ldots, x_n$ gleich dem Mittelwert $\overline{x}$ sind, wenn also keine Streuung vorliegt.

**Bemerkung.** Alternativ kann die Standardabweichung auch berechnet werden als

$$\sigma \;=\; \left(\left(\frac{1}{n}\sum_{j=1}^{n}x_j^2\right) - \overline{x}^2\right)^{1/2}\;.$$

Denn unter Verwendung der Schwerpunkteigenschaft folgt

$$\begin{aligned}
\sigma^2 &= \tfrac{1}{n}\textstyle\sum_{j=1}^n (x_j - \overline{x})^2 \\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^n (x_j^2 - 2x_j\overline{x} + \overline{x}^2) \\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^n (x_j^2 - 2(x_j - \overline{x})\overline{x} - 2\overline{x}^2 + \overline{x}^2) \\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^n (x_j^2 - 2(x_j - \overline{x})\overline{x} - \overline{x}^2) \\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^n (x_j^2 - \overline{x}^2) \\
&= \left(\tfrac{1}{n}\textstyle\sum_{j=1}^n x_j^2\right) - \overline{x}^2\ .
\end{aligned}$$

**Beispiel.** Sei $x_1 = 1$, $x_2 = 3$, $x_3 = 11$; vgl. Beispiel (1) in §1.1.

Es ist $\overline{x} = 5$.

Die Variation ist

$$\sigma^2 \;=\; \frac{1}{3}\left((1-5)^2 + (3-5)^2 + (11-5)^2\right) \;=\; \frac{56}{3}\ .$$

Die Standardabweichung ist also

$$\sigma \;=\; \sqrt{\frac{56}{3}} \;=\; 2\sqrt{\frac{14}{3}} \;\approx\; 4{,}32\ .$$

Alternativ kann man rechnen:

$$\sigma \;=\; \sqrt{\frac{1}{3}(1^2 + 3^2 + 11^2) - 5^2} \;=\; \sqrt{\frac{56}{3}} \;\approx\; 4{,}32\ .$$

Der Variationskoeffizient ist

$$v \;=\; \frac{\sigma}{\bar{x}} \;=\; \frac{1}{5}\sqrt{\frac{56}{3}} \;\approx\; 0{,}86\ .$$

**Beispiel.** Eine Messung der Länge von 100 Werkstücken ergab folgendes.

| | | |
|---|---|---|
| 1 | Werkstück hat Länge | 35 cm |
| 4 | Werkstücke haben Länge | 38 cm |
| 23 | Werkstücke haben Länge | 39 cm |
| 60 | Werkstücke haben Länge | 40 cm (die Soll-Länge) |
| 12 | Werkstücke haben Länge | 41 cm |

Der Mittelwert ist

$$\bar{x} \;=\; \frac{1}{100}(1\cdot 35 + 4\cdot 38 + 23\cdot 39 + 60\cdot 40 + 12\cdot 41) \;=\; 39{,}76\ (\mathrm{cm})\ .$$

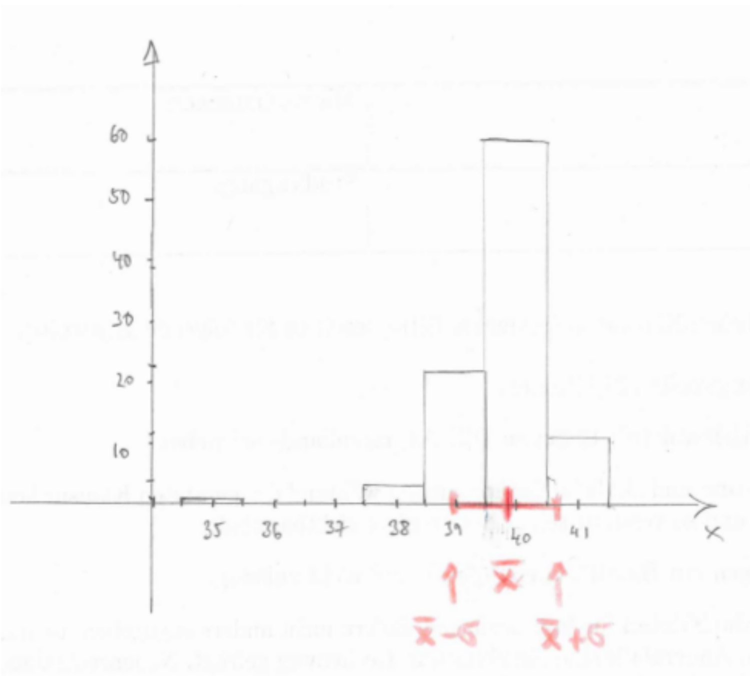
Der Median ist der Mittelwert aus der 50ten und der 51ten gemessenen Länge, wenn man sie der Größe nach anordnet. Diese Werte sind beide gleich 40 cm. Also wird

$$x_{\mathrm{med}} \;=\; \frac{1}{2}(40 + 40) \;=\; 40\ (\mathrm{cm})\ .$$

Die Standardabweichung wird

$$\begin{aligned}\sigma &= \left(\frac{1}{100}\left(1\cdot\underbrace{(35-\bar{x})^2}_{\approx 22{,}66}+4\cdot\underbrace{(38-\bar{x})^2}_{\approx 3{,}10}+23\cdot\underbrace{(39-\bar{x})^2}_{\approx 0{,}58}+60\cdot\underbrace{(40-\bar{x})^2}_{\approx 0{,}06}+12\cdot\underbrace{(41-\bar{x})^2}_{\approx 1{,}54}\right)\right)^{1/2} \\ &\approx 0{,}8381\ (\mathrm{cm})\ .\end{aligned}$$

Der Variationskoeffizient ist $v = \frac{\sigma}{\bar{x}} \approx \frac{0{,}8381}{39{,}76} \approx 0{,}0211$. Derselbe maßstabslose Wert für $v$ hätte sich ergeben, hätte man alles in mm statt in cm gerechnet.

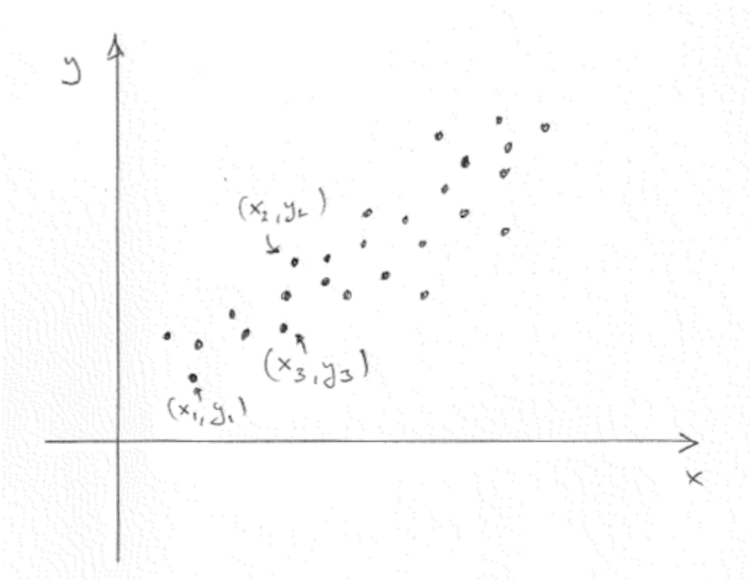



## 1.3 Ausgleichsgerade

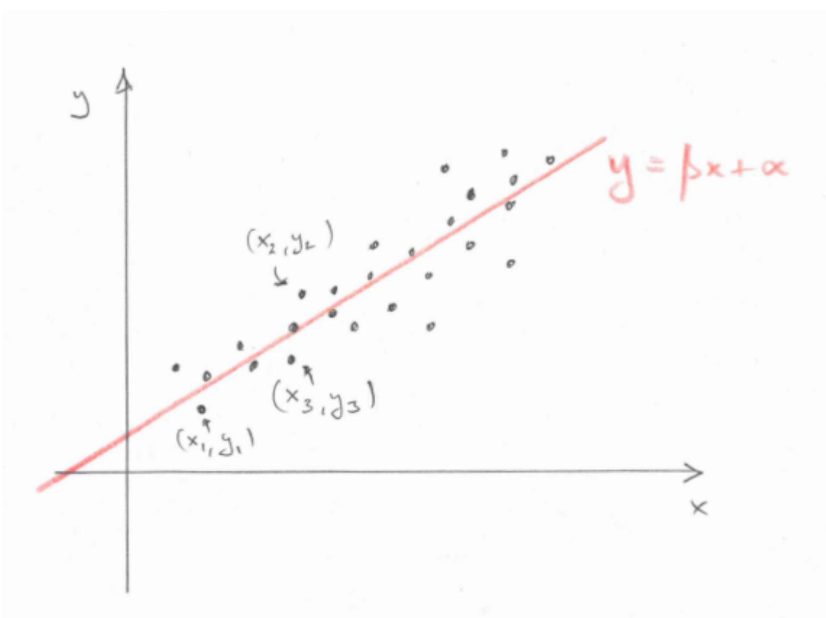
Seien $(x_1, y_1)$, $(x_2, y_2)$, ..., $(x_n, y_n)$ in $\mathbb{R}^2$ gegeben. Es soll dabei $(x_1, x_2, \ldots, x_n)$ nicht aus lauter gleichen Einträgen bestehen.

Man stelle sich unter $(x_j, y_j)$ ein Paar von zusammengehörenden Meßdaten vor. Wie z.B. die Temperatur eines Metallstabs und die bei dieser Temperatur gemessene Länge.

Zeichnet man diese Paare in der Ebene ein, so ergibt sich ein Bild wie folgt.




Gesucht ist nun eine Gerade der Form $y = f(x) = \beta x + \alpha$, die diese Sammlung von gemessenen Daten möglichst gut wiedergibt, genannt *Ausgleichsgerade*.




Folgende Bedingung ist dazu zu erfüllen:

$$e(\alpha, \beta) \; := \; \sum_{j=1}^{n} (f(x_j) - y_j)^2 \quad = \quad \text{Min!}$$

Mit anderen Worten, die Quadrate der vertikalen Abstände der einzelnen Punkte von der Ausgleichsgeraden sollen in der Summe minimal werden.

Um die kritische Stelle für $e(\alpha, \beta)$ zu bestimmen, setzen wir die partiellen Ableitungen

$e_\alpha(\alpha,\beta) = \frac{\partial}{\partial\alpha}e(\alpha,\beta)$ und $e_\beta(\alpha,\beta) = \frac{\partial}{\partial\beta}e(\alpha,\beta)$ gleich 0.
Wir rechnen. Es ist

$$\begin{aligned} e(\alpha,\beta) &= \textstyle\sum_{j=1}^n (f(x_j)-y_j)^2 \\ &= \textstyle\sum_{j=1}^n (\beta x_j + \alpha - y_j)^2 \\ &= \textstyle\sum_{j=1}^n (\beta^2 x_j^2 + \alpha^2 + y_j^2 + 2\beta x_j\alpha - 2\beta x_j y_j - 2\alpha y_j)\ . \end{aligned}$$

Somit soll sein

$$e_\alpha(\alpha,\beta) \;=\; \textstyle\sum_{j=1}^n (2\alpha + 2\beta x_j - 2y_j) \;=\; 2\alpha n + 2\beta\left(\sum_{j=1}^n x_j\right) - 2\left(\sum_{j=1}^n y_j\right) \;\stackrel{!}{=}\; 0$$

und

$$e_\beta(\alpha,\beta) \;=\; \textstyle\sum_{j=1}^n (2\beta x_j^2 + 2x_j\alpha - 2x_j y_j) \;=\; 2\beta\left(\sum_{j=1}^n x_j^2\right) + 2\alpha\left(\sum_{j=1}^n x_j\right) - 2\left(\sum_{j=1}^n x_j y_j\right) \;\stackrel{!}{=}\; 0$$

Wir kürzen ab: $\sum x_j := \sum_{j=1}^n x_j$ , $\sum y_j := \sum_{j=1}^n y_j$ , $\sum x_j^2 := \sum_{j=1}^n x_j^2$ , $\sum x_j y_j := \sum_{j=1}^n x_j y_j$ .
Zu erfüllen ist also das lineare Gleichungssystem

$$\begin{array}{lcl} \alpha n &+& \beta\sum x_j \;=\; \sum y_j \\ \alpha\sum x_j &+& \beta\sum x_j^2 \;=\; \sum x_j y_j\ . \end{array}$$

Aus der Forderung

$$\begin{pmatrix} n & \sum x_j \\ \sum x_j & \sum x_j^2 \end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \;=\; \begin{pmatrix}\sum y_j \\ \sum x_j y_j\end{pmatrix}$$

folgt durch Multiplikation mit der inversen Matrix

$$\begin{aligned} \begin{pmatrix}\alpha\\ \beta\end{pmatrix} &= \begin{pmatrix} n & \sum x_j \\ \sum x_j & \sum x_j^2 \end{pmatrix}^{-1}\begin{pmatrix}\sum y_j \\ \sum x_j y_j\end{pmatrix} \\ &= \frac{1}{n(\sum x_j^2)-(\sum x_j)^2}\begin{pmatrix} \sum x_j^2 & -\sum x_j \\ -\sum x_j & n \end{pmatrix}\begin{pmatrix}\sum y_j \\ \sum x_j y_j\end{pmatrix} \\ &= \frac{1}{n(\sum x_j^2)-(\sum x_j)^2}\begin{pmatrix} (\sum x_j^2)(\sum y_j)-(\sum x_j)(\sum x_j y_j) \\ -(\sum x_j)(\sum y_j)+n(\sum x_j y_j) \end{pmatrix}\ . \end{aligned}$$

Nebenrechnung 1. Es wird

$$\begin{aligned} n(\textstyle\sum x_j^2) - (\sum x_j)^2 &= n^2(\tfrac{1}{n}(\textstyle\sum x_j^2) - \bar{x}^2) \\ &\stackrel{\S 1.2}{=} n^2\sigma^2 \\ &= n^2\cdot\tfrac{1}{n}\textstyle\sum (x_j-\bar{x})^2 \\ &= n\textstyle\sum (x_j-\bar{x})^2\ . \end{aligned}$$

Nebenrechnung 2. Dank Schwerpunkteigenschaft für $\bar{y}$ ist $\sum (y_j - \bar{y}) = 0$; vgl. §1.1. Damit wird

$$\begin{aligned} n(\sum x_j y_j) - (\sum x_j)(\sum y_j) &= n(\sum x_j y_j) - n(\sum x_j)\bar{y} \\ &= n(\sum x_j y_j - x_j \bar{y}) \\ &= n(\sum x_j (y_j - \bar{y})) \\ &= n(\sum x_j (y_j - \bar{y})) - n(\sum \bar{x}(y_j - \bar{y})) \\ &= n(\sum (x_j - \bar{x})(y_j - \bar{y})) \; . \end{aligned}$$

Mit diesen beiden Nebenrechnungen erhalten wir aus dem unteren Vektoreintrag

$$\beta \;=\; \frac{n\sum(x_j - \bar{x})(y_j - \bar{y})}{n\sum(x_j - \bar{x})^2} \;=\; \frac{\sum(x_j - \bar{x})(y_j - \bar{y})}{\sum(x_j - \bar{x})^2} \; .$$

Aus der ersten Gleichung des linearen Gleichungssystems folgt $\alpha n + \beta n \bar{x} = n \bar{y}$, also

$$\beta \bar{x} + \alpha \;=\; \bar{y}$$

und somit $\alpha = \bar{y} - \beta \bar{x}$.

Nun haben wir die einzige kritische Stelle $(\alpha, \beta)$ von $e(\alpha, \beta)$ bestimmt. Man kann nun noch nachrechnen, daß an dieser tatsächlich ein globales Minimum vorliegt.

Wir fassen zusammen:

**Satz.** Mit

$$\beta \;=\; \frac{\sum_{j=1}^{n}(x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{n}(x_j - \bar{x})^2}$$

und

$$\alpha \;=\; \bar{y} - \beta \bar{x}$$

ist die Ausgleichsgerade gegeben durch

$$y \;=\; \beta x + \alpha \; .$$

**Bemerkung.** Die Gleichung $\beta \bar{x} + \alpha = \bar{y}$ besagt: Es liegt der Punkt $(\bar{x}, \bar{y})$ auf der Ausgleichsgeraden.

Hierbei sind $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ und $\bar{y} = \frac{1}{n}\sum_{j=1}^{n} y_j$ die jeweiligen Mittelwerte.

**Beispiel.** Sei $(x_1, y_1) = (2, 2)$, $(x_2, y_2) = (3, 4)$ und $(x_3, y_3) = (7, 3)$.

Durch dese drei Punkte soll eine Ausgleichsgerade gelegt werden.

Wir erhalten $\bar{x} = \frac{1}{3}(x_1+x_2+x_3) = \frac{1}{3}(2+3+7) = 4$ und $\bar{y} = \frac{1}{3}(y_1+y_2+y_3) = \frac{1}{3}(2+4+3) = 3$.

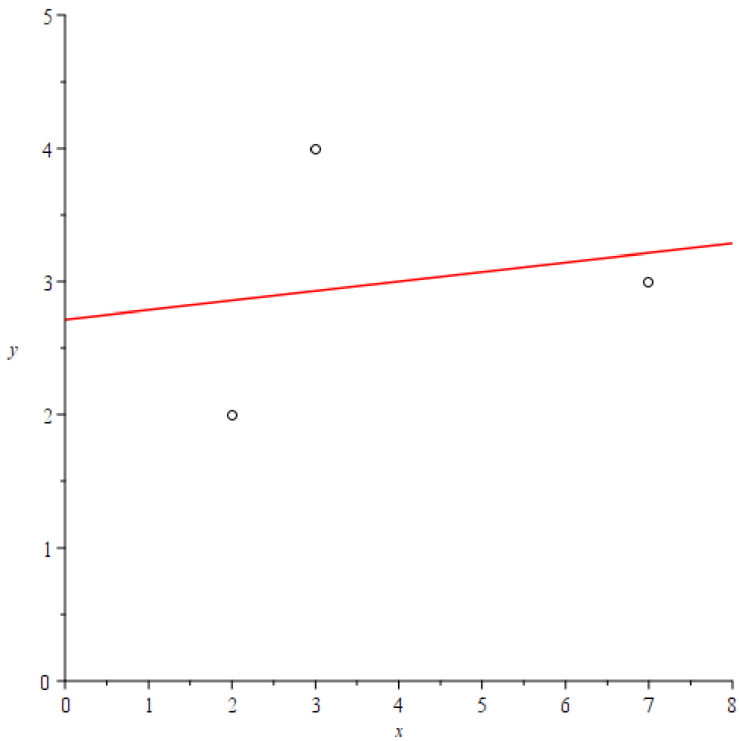
Es wird

$$\begin{aligned} \beta &= \tfrac{(x_1-\bar{x})(y_1-\bar{y})+(x_2-\bar{x})(y_2-\bar{y})+(x_3-\bar{x})(y_3-\bar{y})}{(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2} \\ &= \tfrac{(2-4)(2-3)+(3-4)(4-3)+(7-4)(3-3)}{(2-4)^2+(3-4)^2+(7-4)^2} \\ &= \tfrac{1}{14} \;\approx\; 0{,}0714 \; . \end{aligned}$$

Damit wird

$$\alpha \;=\; \bar{y} - \beta\bar{x} \;=\; 3 - \frac{1}{14}\cdot 4 \;=\; \frac{19}{7} \;\approx\; 2{,}7143\,.$$

Die Ausgleichsgerade ist also durch $y = \frac{1}{14}x + \frac{19}{7} \approx 0{,}0714 \cdot x + 2{,}7143$ gegeben.




**Beispiel.** Ein Experiment liefert folgende Daten. Zu einer Temperatur $x_j$ (in °C) eines Metallstabs wurde immer seine Länge $y_j$ (in cm) gemessen.

$$\begin{aligned}
(x_1, y_1) &= (0, 120)\\
(x_2, y_2) &= (11, 123)\\
(x_3, y_3) &= (24, 127)\\
(x_4, y_4) &= (50, 132)\\
(x_5, y_5) &= (68, 141)
\end{aligned}$$

Es ergibt sich $\bar{x} = \frac{1}{5}\sum_{j=1} x_j = 30{,}6$ und $\bar{y} = \frac{1}{5}\sum_{j=1} y_j = 128{,}6$.

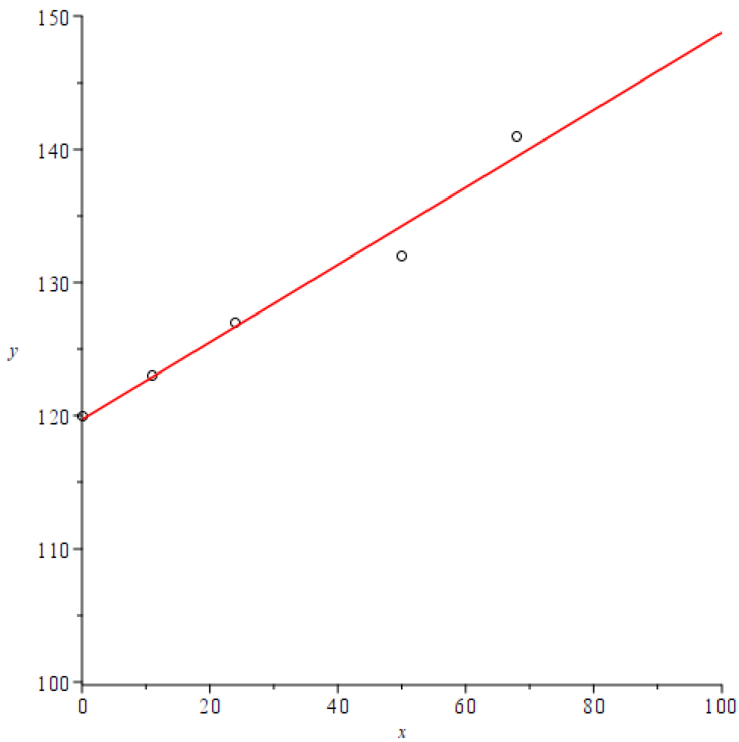
Damit wird

$$\beta \;=\; \frac{(0-30{,}6)(120-128{,}6)+(11-30{,}6)(123-128{,}6)+(24-30{,}6)(127-128{,}6)+(50-30{,}6)(132-128{,}6)+(68-30{,}6)(141-128{,}6)}{(0-30{,}6)^2+(11-30{,}6)^2+(24-30{,}6)^2+(50-30{,}6)^2+(68-30{,}6)^2}$$

$$\approx\; 0{,}2909\,.$$

und

$$\alpha \;=\; \bar{y} - \beta\bar{x} \;\approx\; 128{,}6 - 0{,}2909\cdot 30{,}6 \;\approx\; 119{,}69\,.$$




## 1.4 Empirischer Korrelationskoeffizient

Seien $(x_1, y_1)$, $(x_2, y_2)$, $\ldots$, $(x_n, y_n) \in \mathbb{R}^2$.

Es sollen dabei weder $\vec{x} := (x_1, x_2, \ldots, x_n)$ noch $\vec{y} := (y_1, y_2, \ldots, y_n)$ aus lauter gleichen Einträgen bestehen.

Gesucht ist ein Maß dafür, wie ausgeprägt ein linearer Zusammenhang zwischen den Größen $x_j$ und $y_j$ besteht.

**Definition.** Der *empirische Korrelationskoeffizient* ist gegeben durch

$$r_{\vec{x},\vec{y}} := \frac{\sum_{j=1}^{n}(x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\left(\sum_{j=1}^{n}(x_j - \bar{x})^2\right)\left(\sum_{j=1}^{n}(y_j - \bar{y})^2\right)}}$$

Kurz heißt $r_{\vec{x},\vec{y}}$ auch die *Korrelation*.

**Bemerkung.** Es kann $r_{\vec{x},\vec{y}}$ gelesen werden als der Cosinus des Winkels zwischen den Vektoren $(x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x})$ und $(y_1 - \bar{y}, y_2 - \bar{y}, \ldots, y_n - \bar{y})$. Also ist

$$-1 \leqslant r_{\vec{x},\vec{y}} \leqslant +1 .$$

Ansonsten braucht man die geometrische Interpretation dieser Vektoren aber nicht.

**Bemerkung.**

(1) Es ist $r_{\vec{x},\vec{y}} = r_{\vec{y},\vec{x}}$ .

(2) Für $c \in \mathbb{R}$ sei $\vec{x} + c := (x_1 + c, x_2 + c, \dots x_3 + c)$. Es ist

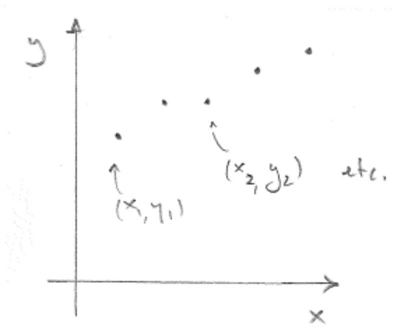
$$r_{\vec{x},\vec{y}+c} \;=\; r_{\vec{x},\vec{y}} \;.$$

(2) Für $a \in \mathbb{R} \setminus \{0\}$ sei $a\vec{x} := (ax_1, ax_2, \dots, ax_n)$. Es ist

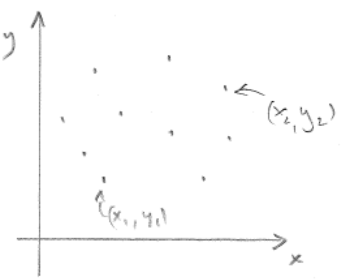
$$r_{\vec{x},a\vec{y}} \;=\; \begin{cases} r_{\vec{x},\vec{y}} & \text{falls } a > 0 \\ -r_{\vec{x},\vec{y}} & \text{falls } a < 0 \;. \end{cases}$$

**Beispiel** zur Vorstellung.

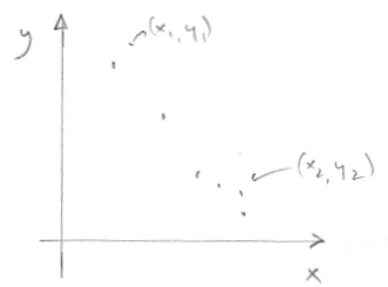
Korrelation $r_{\vec{x},\vec{y}}$ knapp unter 1:




Korrelation $r_{\vec{x},\vec{y}}$ nahe bei 0:




Korrelation $r_{\vec{x},\vec{y}}$ knapp über $-1$:




**Bemerkung.**

Ist die Korrelation knapp unter 1, so sagt man, die $x$-Werte und die $y$-Werte sind stark korreliert. Die Punkte liegen dann nahe an der Ausgleichsgeraden.

Ist die Korrelation nahe bei 0, so sagt man, die $x$-Werte und die $y$-Werte sind schwach korreliert. Die Punkte liegen dann weit entfernt von der Ausgleichsgeraden.

Ist die Korrelation knapp über $-1$, so sagt man, die $x$-Werte und die $y$-Werte sind stark negativ korreliert. Die Punkte liegen dann nahe an der Ausgleichsgeraden.

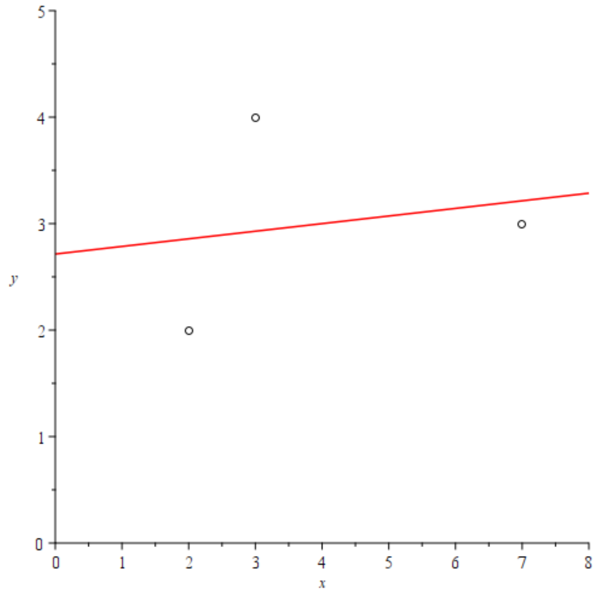
**Beispiel.** Sei $(x_1, y_1) = (2, 2)$, $(x_2, y_2) = (3, 4)$ und $(x_3, y_3) = (7, 3)$.

Vgl. erstes Beispiel in §1.3. Es ist $\bar{x} = 4$ und $\bar{y} = 3$.

Die Korrelation ergibt sich zu

$$\begin{aligned} r_{\vec{x},\vec{y}} &= \frac{(x_1-\bar{x})(y_1-\bar{y})+(x_2-\bar{x})(y_2-\bar{y})+(x_3-\bar{x})(y_3-\bar{y})}{((x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2)^{1/2} \cdot ((y_1-\bar{y})^2+(y_2-\bar{y})^2+(y_3-\bar{y})^2)^{1/2}} \\ &= \frac{(2-4)(2-3)+(3-4)(4-3)+(7-4)(3-3)}{((2-4)^2+(3-4)^2+(7-4)^2)^{1/2} \cdot ((2-3)^2+(4-3)^2+(3-3)^2)^{1/2}} \\ &= \frac{1}{\sqrt{14 \cdot 2}} \;=\; \frac{1}{2\sqrt{7}} \;\approx\; 0{,}1890 \;. \end{aligned}$$

Die $x$-Werte und die $y$-Werte sind nur schwach korreliert. Die Ausgleichsgerade hat großen Abstand von den Punkten.




**Beispiel.** Wir betrachten wieder das zweite Beispiel in §1.3. Dort lieferte ein Experiment folgende Daten. Zu einer Temperatur $x_j$ (in °C) eines Metallstabs wurde immer seine Länge $y_j$ (in cm) gemessen.

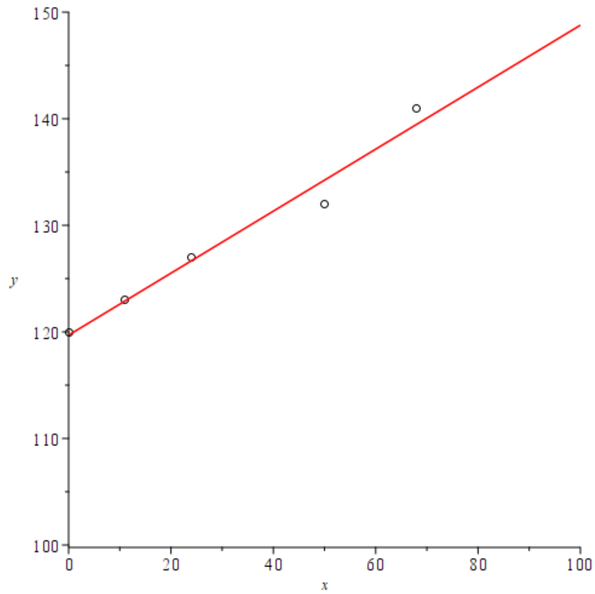
$$
\begin{aligned}
(x_1, y_1) &= (0, 120) \\
(x_2, y_2) &= (11, 123) \\
(x_3, y_3) &= (24, 127) \\
(x_4, y_4) &= (50, 132) \\
(x_5, y_5) &= (68, 141)
\end{aligned}
$$

Es ist $\bar{x} = 30{,}6$ und $\bar{y} = 128{,}6$.

Die Korrelation ergibt sich zu

$$
\begin{aligned}
r_{\vec{x},\vec{y}} &= \frac{(x_1-\bar{x})(y_1-\bar{y})+(x_2-\bar{x})(y_2-\bar{y})+(x_3-\bar{x})(y_3-\bar{y})+(x_4-\bar{x})(y_4-\bar{y})+(x_5-\bar{x})(y_5-\bar{y})}{((x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+(x_4-\bar{x})^2+(x_5-\bar{x})^2)^{1/2}\cdot((y_1-\bar{y})^2+(y_2-\bar{y})^2+(y_3-\bar{y})^2+(y_4-\bar{y})^2+(y_5-\bar{y})^2)^{1/2}} \\
&\approx 0{,}9861\ .
\end{aligned}
$$

Die $x$-Werte (jeweilige Temperatur des Metallstabs) und die $y$-Werte (jeweilige Länge des Metallstabs) sind stark korreliert. Die Ausgleichsgerade hat kleinen Abstand von den Punkten.


150
140
130
120
110
100
y
0
20
40
60
80
100
x

# Kapitel 2

# Wahrscheinlichkeitstheorie

## 2.1 Begriff der Wahrscheinlichkeit

**Definition.** Gegeben sei als *Ergebnisraum* eine Menge $\Omega$.

Es ist $\Omega$ die Menge der möglichen Ergebnisse eines Zufallsexperiments.

**Beispiel.** Beim Würfeln mit einem Würfel ist $\Omega = \{1, 2, 3, 4, 5, 6\}$.

**Beispiel.** Beim Messen der Temperatur (in °C) um 8:00 morgens ist $\Omega = \mathbb{R}_{\geqslant -273{,}15}$ .

**Bemerkung.** Aus mathematik-technischen Gründen muß man manchmal noch zwischen meßbaren und nicht meßbaren Teilmengen $A \subseteq \Omega$ unterscheiden. Nicht meßbare Teilmengen von $\Omega$ sind nicht ganz einfach zu finden. Die in der Praxis üblicherweise auftretenden Teilmengen sind in aller Regel meßbar. Wir werden dieses Problem daher weitgehend ignorieren.

Ist z.B. $\Omega$ endlich, so stellt sich dieses Problem nicht: Jede Teilmenge von $\Omega$ ist meßbar.

**Definition.** Eine meßbare Teilmenge $A \subseteq \Omega$ heißt *Ereignis*.

Diese Teilmenge stellt das Ereignis dar, daß das Ergebnis des Zufallsexperiment in $A$ liegt.

**Beispiel.** Sei wieder beim Würfeln $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Das Ereignis, eine ungerade Zahl zu würfeln, ist die Teilmenge $A = \{1, 3, 5\} \subseteq \Omega$.

**Beispiel.** Sei wieder beim Temperaturmessen $\Omega = \mathbb{R}_{\geqslant -273{,}15}$ .

Das Ereignis, Frost zu haben, also eine Temperatur $\leqslant 0$, ist die Teilmenge

$$\{\, x \in \Omega \,:\, x \leqslant 0 \,\} \;=\; [-273{,}15, 0] \;\subseteq\; \Omega \,.$$

**Definition.** Gegeben sei ein *Wahrscheinlichkeitsmaß* $P$, das jedem Ereignis $A \subseteq \Omega$ eine *Wahrscheinlichkeit* $P(A) \in \mathbb{R}$ zuordnet.

Dabei sollen die folgenden Axiome von Kolmogorov gelten.

(K 1) Es ist $P(A) \geqslant 0$ für jedes Ereignis $A \subseteq \Omega$.

(K 2) Es ist $P(\Omega) = 1$.

(K 3) Es ist $P(A \cup B) = P(A) + P(B)$ für Ereignisse $A,\ B \subseteq \Omega$ mit $A \cap B = \emptyset$.
Allgemeiner ist auch noch $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ für Ereignisse $A_1,\ A_2,\ \ldots \subseteq \Omega$ mit $A_i \cap A_j = \emptyset$ für $i, j \geqslant 1$ mit $i \neq j$.

Interpretation:

- Es ist $P(A)$ die Wahrscheinlichkeit, daß das Zufallsexperiment ein Ergebnis in $A$ liefert. ("$P$" für "probability".)
- $P(A) = 0$ heißt: Es ist fast unmöglich, ein Ergebnis in $A$ zu haben.
- $P(A) = 1$ heißt: Es ist fast sicher, ein Ergebnis in $A$ zu haben.

  Zum Beispiel ist $P(\Omega) = 1$, da es sicher ist, ein Ergebnis in $\Omega$ zu haben.
- Sei $A \subseteq \Omega$ ein Ereignis. Es heißt $\bar{A} := \Omega \setminus A \subseteq \Omega$ das zugehörige *Gegenereignis*.

  Dank (K 3) ist
  $$P(A) + P(\bar{A}) \;=\; P(A \cup \bar{A}) \;=\; P(\Omega) \;=\; 1\,.$$
  Insbesondere ist
  $$P(\bar{A}) \;=\; 1 - P(A)\,.$$
  Da $P(\bar{A}) \geqslant 0$ ist, folgt
  $$0 \;\leqslant\; P(A) \;\leqslant\; 1\,.$$
  Ferner folgt
  $$P(\emptyset) \;=\; P(\Omega \setminus \Omega) \;=\; P(\bar{\Omega}) \;=\; 1 - P(\Omega) \;=\; 0\,.$$
- (K 3) heißt: Wahrscheinlichkeiten disjunkter Ereignisse werden addiert.

**Beispiel.** Sei wieder beim Würfeln $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Es ist

$$P(\{1\}) \;=\; P(\{2\}) \;=\; P(\{3\}) \;=\; P(\{4\}) \;=\; P(\{5\}) \;=\; P(\{6\}) \;=\; \frac{1}{6}\,.$$

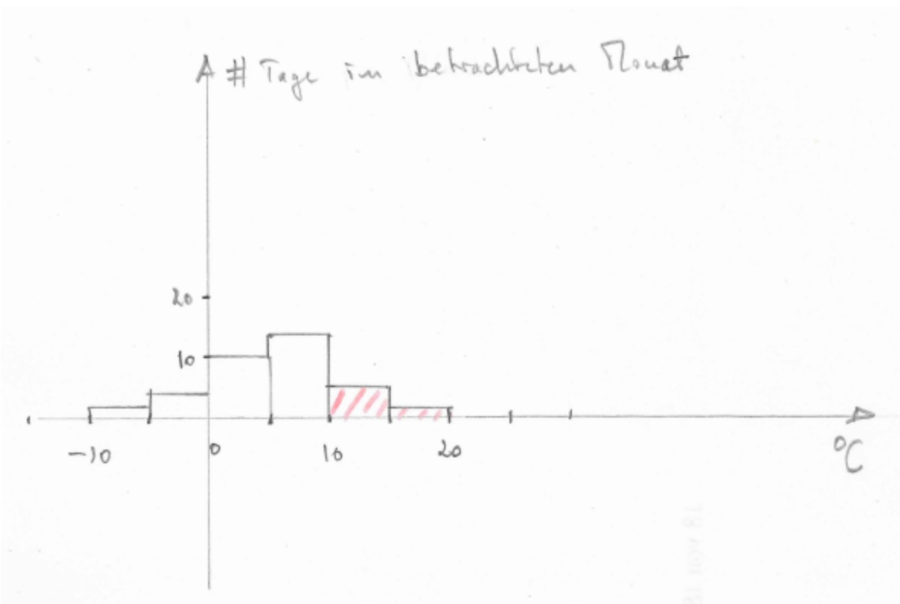
Die Wahrscheinlichkeit, eine ungerade Zahl zu würfeln, ist

$$P(\{1,3,5\}) \;=\; P(\{1\}) + P(\{1\}) + P(\{1\}) \;=\; 3 \cdot \frac{1}{6} \;=\; \frac{1}{2}\,.$$

Das Gegenereignis, eine gerade Zahl zu würfeln, hat also Wahrscheinlichkeit $1 - \frac{1}{2} = \frac{1}{2}$.

**Beispiel.** (Temperaturmessung)

Im Monat März sei folgende langjährige Temperaturverteilung bekannt, die wir unseren Wahrscheinlichkeitsbetrachtungen zugrunde legen wollen.




Es gebe also z.B. 10 Tage, an denen die Durchschnittstemperatur zwischen 0 und 5 Grad liegt.

Hier ist $\Omega = \mathbb{R}_{\geqslant -273{,}15}$ .

Es ist $P(\{10\}) = 0$, da es fast unmöglich ist, eine Temperatur von exakt 10 Grad zu erhalten.

Es ist $P([10,20])$ die Wahrscheinlichkeit, eine Temperatur zwischen 10 und 20 Grad zu erhalten.

Es ist

$$\begin{aligned} P([10,20]) \quad &= \quad \frac{\text{Anzahl der Tage mit Temperatur zwischen 10 und 20 Grad}}{\text{Anzahl der Tage im Monat}} \\ &\overset{\text{anschaulich}}{=} \quad \frac{\text{Flächeninhalt in rot}}{\text{Flächeninhalt unter Kurve insgesamt}}\,. \end{aligned}$$

An diesem Beispiel kann man erkennen, warum allgemein $P$ immer Ereignissen, also Teilmengen von $\Omega$, eine Wahrscheinlichkeit zuweist, und nicht einem einzelnen Ergebnis.

## 2.2 Laplace-Experimente

**Definition.** Ist $M$ eine endliche Menge, so schreiben wir $\#M$ für die Anzahl ihrer Elemente.

**Definition.** Ist $\Omega$ endlich und nichtleer und ist

$$P(\{x\}) = \frac{1}{\#\Omega}$$

für jedes $x \in \Omega$, so sprechen wir von einem *Laplace-Experiment*.

Diesenfalls gilt dank (K 3) für $A \subseteq \Omega$

$$P(A) \;=\; \frac{\#A}{\#\Omega}\;.$$

In Worten, die Wahrscheinlichkeit für ein Ereignis $A$ ist gleich der Anzahl der für $A$ günstigen Ergebnisse (also $\#A$), geteilt durch die Anzahl aller möglichen Ergebnisse (also $\#\Omega$).

**Beispiel.** Es werde eine Münze dreimal geworfen. Es stehe W für Wappen, Z für Zahl.

Es ist

$$\Omega \;=\; \{(\mathrm{W},\mathrm{W},\mathrm{W}),(\mathrm{W},\mathrm{W},\mathrm{Z}),(\mathrm{W},\mathrm{Z},\mathrm{W}),(\mathrm{W},\mathrm{Z},\mathrm{Z}),(\mathrm{Z},\mathrm{W},\mathrm{W}),(\mathrm{Z},\mathrm{W},\mathrm{Z}),(\mathrm{Z},\mathrm{Z},\mathrm{W}),(\mathrm{Z},\mathrm{Z},\mathrm{Z})\}\;.$$

Es liegt ein Laplace-Experiment vor, da jedes Ergebnis Wahrscheinlichkeit $\frac{1}{\#\Omega} = \frac{1}{8}$ hat.

Wir betrachten das Ereignis, daß genau zweimal Wappen geworfen wird, also

$$A \;=\; \{(\mathrm{W},\mathrm{W},\mathrm{Z}),(\mathrm{W},\mathrm{Z},\mathrm{W}),(\mathrm{Z},\mathrm{W},\mathrm{W})\}\;.$$

Dafür ist die Wahrscheinlichkeit

$$P(A) \;=\; \frac{\#A}{\#\Omega} \;=\; \frac{3}{8}\;.$$

**Beispiel.** Wir betrachten das Lottospiel "3 aus 9". Es werden aus den Zahlen $\{1, 2, \ldots, 9\}$ drei Zahlen gezogen.

Der Ergebnisraum ist also

$$\begin{aligned}
\Omega \;=\; \{&(1,2,3),(1,2,4),\ldots,(1,2,9),\\
&(1,3,2),(1,3,4),\ldots,(1,3,9),\\
&(1,4,2),(1,4,3),\ldots,(1,4,9),\\
&\vdots\\
&(9,8,1),(9,8,2),\ldots,(9,8,7)\}
\end{aligned}$$

Es ist $\#\Omega = 9 \cdot 8 \cdot 7$.

Jemand hat die Zahlen 1, 4 und 5 getippt.

Er will wissen, wie hoch die Wahrscheinlichkeit für drei Richtige ist. Das zugehörige Ereignis ist

$$A_3 \;=\; \{(1,4,5),(1,5,4),(4,1,5),(4,5,1),(5,1,4),(5,4,1)\}\;.$$

Es ist $\#A_3 = 3\cdot 2\cdot 1 = 6$. Also ist

$$P(A_3) \;=\; \frac{\#A_3}{\#\Omega} \;=\; \frac{3\cdot 2\cdot 1}{9\cdot 8\cdot 7} \;=\; \frac{1}{\binom{9}{3}} \;=\; \frac{1}{84}\;.$$

Er will auch noch wissen, wie hoch die Wahrscheinlichkeit für zwei Richtige ist. Das zugehörige Ereignis ist

$$A_2 \;=\; \{\;(1,4,x),\;(4,1,x),\;(1,5,x),\;(5,1,x),\;(4,5,x),\;(5,4,x),\;\;:\;x\in\{2,3,6,7,8,9\}\;\}\;.$$

Es ist $\#A_2 = 6\cdot(9-3)$. Also ist

$$P(A_2) \;=\; \frac{\#A_2}{\#\Omega} \;=\; \frac{6\cdot 6}{9\cdot 8\cdot 7} \;=\; \frac{1}{14}\;.$$

**Beispiel**. Beim Skatspiel gibt es 32 Karten, vier davon sind Buben. Jeder Spieler erhält 10 Karten.

Sei $A_4$ das Ereignis, daß ein Spieler 4 Buben unter seinen 10 Karten findet.

Wir ziehen nacheinander 10 Karten aus den 32 Karten. Sei $\Omega$ die Menge der Ziehungsprotokolle. Also z.B.

$$(\text{Herz-7},\text{Karo-9},\text{Kreuz-7},\text{Pik-9},\text{Pik-10},\text{Kreuz-Dame},\text{Pik-Dame},\text{Pik-König},\text{Pik-Bube},\text{Karo-Dame})\;\in\;\Omega$$

Die Anzahlen der beim jeweiligen Ziehungsschritt noch verfügbaren Karten geben

$$\#\Omega \;=\; 32\cdot 31\cdot 30\cdot 29\cdot 28\cdot 27\cdot 26\cdot 25\cdot 24\cdot 23\;.$$

Sei $A_4$ die Menge der Ziehungsprotokolle, bei denen 4 Buben gefunden werden. Wir haben $\#A_4$ zu bestimmen.

Bezeichnen wir mit B im Protokoll, daß ein Bube gezogen wurde. Bezeichnen wir mit N im Protokoll, daß ein Nicht-Bube gezogen wurde. Für die Stellen im Protokoll mit Buben gibt es dann die Möglichkeiten

$$(\mathrm{B,B,B,B,N,N,N,N,N,N}),\;(\mathrm{B,B,B,N,B,N,N,N,N,N}),\;\ldots,\;(\mathrm{N,N,N,N,N,N,B,B,B,B})$$

Das sind $\binom{10}{4}$ Möglichkeiten. Für jede dieser Möglichkeiten gibt es dann wiederum

$$4\cdot 3\cdot 2\cdot 1\cdot 28\cdot 27\cdot 26\cdot 25\cdot 24\cdot 23$$

mögliche Protokolle. Also ist

$$P(A_4) \;=\; \frac{\binom{10}{4} \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 28 \cdot 27 \cdot 26 \cdot 25 \cdot 24 \cdot 23}{32 \cdot 31 \cdot 30 \cdot 29 \cdot 28 \cdot 27 \cdot 26 \cdot 25 \cdot 24 \cdot 23} \;=\; \frac{\binom{10}{4}}{\binom{32}{4}} \;=\; \frac{21}{3594} \;\approx\; 0{,}005840$$

die Wahrscheinlichkeit, vier Buben zu finden.

Darf man übrigens als Alleinspieler noch zwei weitere Karten aufnehmen, dann ist dabei entsprechend die Wahrscheinlichkeit

$$\frac{\binom{12}{4}}{\binom{32}{4}} \;=\; \frac{99}{7192} \;\approx\; 0{,}01377\;,$$

vier Buben zu finden.

**Beispiel.** Wir würfeln.

(1) Sei $A_1$ das Ereignis, nach einem Wurf eine 6 zu würfeln. Es ist also

$$A_1 \;=\; \{1\} \;\subseteq\; \{1,2,3,4,5,6\} \;=\; \Omega\;.$$

Die Wahrscheinlichkeit, nach einem Wurf eine 6 zu würfeln, ist also

$$P(A_1) \;=\; \frac{\#A_1}{\#\Omega} \;=\; \frac{1}{6} \;\approx\; 0{,}1667\;.$$

(2) Sei $A_2$ das Ereignis, nach zwei Würfen mindestens eine 6 zu würfeln.

Hier gehen wir besser über das Gegenereignis $\bar{A}_2$, welches beinhaltet, nach zwei Würfen keine 6 zu würfeln. Dafür gibt es $\#\bar{A}_2 = 5 \cdot 5 = 25$ Möglichkeiten, aus einer Ergebnismenge mit $\#\Omega = 6 \cdot 6 = 36$. Also ist

$$P(\bar{A}_2) \;=\; \frac{\#A_2}{\#\Omega} \;=\; \frac{25}{36}\;.$$

Folglich ergibt sich die Wahrscheinlichkeit, bei zwei Würfen mindestens eine 6 zu würfeln, zu

$$P(A_2) \;=\; 1 - P(\bar{A}_2) \;=\; 1 - \frac{25}{36} \;=\; \frac{11}{36} \;\approx\; 0{,}3056\;.$$

(3) Sei $A_3$ das Ereignis, nach drei Würfen mindestens eine 6 zu würfeln. Analog zu (2) wird

$$P(A_3) \;=\; 1 - \frac{5^3}{6^3} \;=\; \frac{91}{216} \;\approx\; 0{,}4213\;.$$

## 2.3 Bedingte Wahrscheinlichkeit und Pfadregel

**Definition** (Bedingte Wahrscheinlichkeit).

Seien $A$, $B \subseteq \Omega$ Ereignisse, mit $P(B) \neq 0$. Es heißt

$$P(A|B) \;:=\; \frac{P(A \cap B)}{P(B)}$$

die *bedingte Wahrscheinlichkeit* von $A$ unter $B$.

Es gibt $P(A|B)$ die Wahrscheinlichkeit, daß $A$ eintritt, unter Voraussetzung, daß $B$ eintritt.

**Beispiel** (Alarmanlage). Wir wollen die Funktionalität einer Alarmanlage evaluieren.

Sei $A$ das Ereignis “Alarm wurde ausgelöst”. Sei $B$ das Ereignis “Ein Einbruch hat stattgefunden”.

Die Gegenereignisse sind $\bar{A}$, “Alarm wurde nicht ausgelöst”, und $\bar{B}$, “Ein Einbruch hat nicht stattgefunden”.

Wir erstellen eine Tafel wie folgt,

| $P$ | $A$ | $\bar{A}$ | |
|---|---|---|---|
| $B$ | $P(A \cap B)$ | $P(\bar{A} \cap B)$ | $P(B)$ |
| $\bar{B}$ | $P(A \cap \bar{B})$ | $P(\bar{A} \cap B)$ | $P(\bar{B})$ |
| | $P(A)$ | $P(\bar{A} \cap \bar{B})$ | $P(\Omega)$ |

Auf der Grundlage einer Versuchsreihe mit 1000 simulierten Verläufen einer Nacht ergab sich folgende Tafel.

| $P$ | $A$ | $\bar{A}$ | |
|---|---|---|---|
| $B$ | 0,08 | 0,02 | 0,1 |
| $\bar{B}$ | 0,1 | 0,8 | 0,9 |
| | 0,18 | 0,82 | 1 |

Hierbei ist die Summe in den Zeilen immer die rechts stehende Zahl. Z.B. ist ja

$$P(A \cap B) + P(\bar{A} \cap B) \;=\; P((A \cap B) \cup (\bar{A} \cap B)) \;=\; P(B)\,.$$

Ferner ist die Summe in den Spalten immer die unten stehende Zahl. Z.B. ist ja

$$P(A \cap B) + P(A \cap \bar{B}) \;=\; P((A \cap B) \cup (A \cap \bar{B})) \;=\; P(A)\,.$$

Wir entnehmen: Die Wahrscheinlichkeit für Fehlalarm beträgt $P(A \cap \bar{B}) = 0{,}1$. Die Wahrscheinlichkeit für ein Versagen der Alarmanlage beträgt $P(\bar{A} \cap B) = 0{,}02$.

Im Fall, daß ein Einbruch stattgefunden hat, war die bedingte Wahrscheinlichkeit, daß die Alarmanlage funktioniert,

$$P(A|B) \;=\; \frac{P(A \cap B)}{P(B)} \;=\; \frac{0{,}08}{0{,}1} \;=\; 0{,}8\;.$$

Das ist zufriedenstellend.

Im Fall, daß die Alarmanlage losging, war die Wahrscheinlichkeit, daß ein Fehlalarm vorlag, bei

$$P(\bar{B}|A) \;=\; \frac{P(A \cap \bar{B})}{P(A)} \;=\; \frac{0{,}1}{0{,}18} \;=\; \frac{5}{9} \;\approx\; 0{,}5556\;.$$

Das sollte dem Hersteller zu denken geben.

**Definition** (Unabhängigkeit).

Seien $A$, $B \subseteq \Omega$ Ereignisse. Es heißen $A$ und $B$ *unabhängig*, falls

$$P(A \cap B) \;=\; P(A) \cdot P(B)\;.$$

Falls $A$ und $B$ unabhängig sind und falls $P(B) \neq 0$ ist, folgt

$$P(A|B) \;=\; \frac{P(A \cap B)}{P(B)} \;=\; \frac{P(A) \cdot P(B)}{P(B)} \;=\; P(A)\;.$$

Mit anderen Worten, ob das Ereignis $B$ zur Bedingung gemacht wird oder auch nicht, ändert diesenfalls nichts an der Wahrscheinlichkeit für das Ereignis $A$.

Falls $A$ und $B$ unabhängig sind und falls $P(A) \neq 0$ ist, folgt analog

$$P(B|A) \;=\; P(B)\;.$$

**Beispiel** (Alarmanlage). Es sind im vorigen Beispiel die Ereignisse "Alarm wurde ausgelöst" und "Ein Einbruch hat stattgefunden" nicht unabhängig. Denn

$$P(A \cap B) \;=\; 0{,}08 \;\neq\; 0{,}18 \cdot 0{,}1 \;=\; P(A) \cap P(B)\;.$$

**Beispiel** (Kugeln). Es befinden sich drei Kugeln in einem Gefäß, mit den Nummern 1, 2, 3.

Es sollen nacheinander zwei Kugeln gezogen werden. Für eine Ziehung einer Kugel sollen alle Kugeln dieselbe Wahrscheinlichkeit haben.

Es gibt zwei Möglichkeiten des Vorgehens.

(1) Die Kugel wird nach dem ersten Zug wieder in das Gefäß zurückgelegt.

Es ist

$$\begin{aligned} \Omega \;=\; \{&(1,1),(1,2),(1,3),\\ &(2,1),(2,2),(2,3),\\ &(3,1),(3,2),(3,3)\} \end{aligned}$$

Sei $A$ das Ereignis, im ersten Zug eine 1 zu ziehen. Es ist $A = \{(1,1),(1,2),(1,3)\}$ und also $P(A) = \frac{3}{9} = \frac{1}{3}$.

Sei $B$ das Ereignis, im zweiten Zug eine 3 zu ziehen. Es ist $B = \{(1,3),(2,3),(3,3)\}$ und also $P(B) = \frac{3}{9} = \frac{1}{3}$.

Dann ist $A \cap B$ das Ereignis, im ersten Zug eine 1 und im zweiten Zug eine 3 zu ziehen. Es ist $A \cap B = \{(1,3)\}$ und also $P(A \cap B) = \frac{1}{9}$.

Die Ereignisse $A$ und $B$ sind unabhängig, da

$$P(A) \cdot P(B) \;=\; \frac{1}{3} \cdot \frac{1}{3} \;=\; \frac{1}{9} \;=\; P(A \cap B)$$

ist.

(2) Die Kugel wird nach dem ersten Zug nicht wieder in das Gefäß zurückgelegt.

Es ist

$$\begin{array}{rcllll} \Omega & = & \{ & & (1,2), & (1,3), \\ & & & (2,1), & & (2,3), \\ & & & (3,1), & (3,2) & \quad \} \end{array}$$

Sei $A$ das Ereignis, im ersten Zug eine 1 zu ziehen. Es ist $A = \{(1,2),(1,3)\}$ und also $P(A) = \frac{2}{6} = \frac{1}{3}$.

Sei $B$ das Ereignis, im zweiten Zug eine 3 zu ziehen. Es ist $B = \{(1,3),(2,3)\}$ und also $P(B) = \frac{2}{6} = \frac{1}{3}$.

Dann ist $A \cap B$ das Ereignis, im ersten Zug eine 1 und im zweiten Zug eine 3 zu ziehen. Es ist $A \cap B = \{(1,3)\}$ und also $P(A \cap B) = \frac{1}{6}$.

Die Ereignisse $A$ und $B$ sind nicht unabhängig, da

$$P(A) \cdot P(B) \;=\; \frac{1}{3} \cdot \frac{1}{3} \;=\; \frac{1}{9} \;\neq\; \frac{1}{6} \;=\; P(A \cap B)$$

ist.

Hier ist z.B. $P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{1/6}{1/3} = \frac{1}{2}$ die Wahrscheinlichkeit, beim zweiten Zug die Kugel 3 zu ziehen, unter der Bedingung, im ersten Zug die Kugel 1 gezogen zu haben. Insbesondere ist $P(B|A) = \frac{1}{2} \neq \frac{1}{3} = P(B)$.

**Bemerkung** (Pfadregel).

Sei $n \geqslant 1$. Seien $A_1, \ldots, A_n \subseteq \Omega$. Dann ist

$$\begin{array}{rl} & P(A_1 \cap A_2 \cap \ldots \cap A_n) \\ = & P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \,. \end{array}$$

Denn tatsächlich können wir durch Kürzen erreichen:

$$\begin{array}{rl} & P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \\ = & P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdot \ldots \cdot \frac{P(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n)}{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1})} \\ = & P(A_1 \cap A_2 \cap \ldots \cap A_n) \,. \end{array}$$

**Beispiel** (Streichhölzer ziehen).

Seien 4 Streichhölzer vorliegen. Sei unter diesen ein kurzes.

Es ziehen 4 Personen je ein Streichholz, ohne zu sehen, welches das kurze ist. Jeder behält nach dem Zug das Streichholz.

Wer das kurze Streichholz zieht, verliert.

Welche Position im Ablauf ist vorteilhaft? Die als erste Person, als zweite Person, als dritte Person oder als vierte Person?

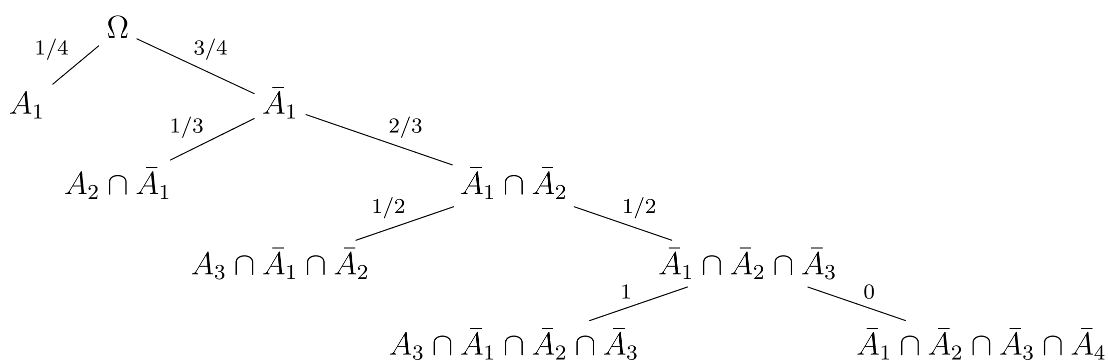
Sei $A_1$ das Ereignis, daß die erste Person ein kurzes Streichholz zieht. Dann ist $P(\bar{A}_1) = \frac{3}{4}$ und $P(A_1) = \frac{1}{4}$.

Sei $A_2$ das Ereignis, daß die zweite Person ein kurzes Streichholz zieht. Da dies nur eintritt, wenn die erste Person ein langes Streichholz gezogen hat und da danach noch 2 lange und ein kurzes Streichholz vorhanden sind, ist dann $P(\bar{A}_2|\bar{A}_1) = \frac{2}{3}$ und $P(A_2|\bar{A}_1) = \frac{1}{3}$.

Sei $A_3$ das Ereignis, daß die dritte Person ein kurzes Streichholz zieht. Da dies nur eintritt, wenn die erste und die zweite Person je ein langes Streichholz gezogen haben und da danach noch ein langes und ein kurzes Streichholz vorhanden sind, ist dann $P(\bar{A}_3|A_1 \cap A_2) = \frac{1}{2}$ und $P(A_3|A_1 \cap A_2) = \frac{1}{2}$.

Sei $A_4$ das Ereignis, daß die dritte Person ein kurzes Streichholz zieht. Da dies nur eintritt, wenn die erste und die zweite und die dritte Person je ein langes Streichholz gezogen haben und da danach noch ein kurzes Streichholz vorhanden ist, ist dann $P(\bar{A}_4|A_1 \cap A_2 \cap A_3) = 0$ und $P(A_4|A_1 \cap A_2 \cap A_3) = 1$.

Wir erstellen folgenden Baum.




An den Ästen notieren wir die bedingten Wahrscheinlichkeiten. Um zur Wahrscheinlichkeit eines Ereignisses weiter unten im Baum zu kommen, multiplizieren wir gemäß Pfadregel die bedingten Wahrscheinlichkeiten entlang dem Pfad, der dorthin führt.

Es wird

$$\begin{array}{rclcl}
P(A_1) & = & P(A_1) & = & \frac{1}{4} & = & \frac{1}{4} \\
P(A_2) & = & P(A_2 \cap \bar{A}_1) & = & \frac{3}{4} \cdot \frac{1}{3} & = & \frac{1}{4} \\
P(A_3) & = & P(A_3 \cap \bar{A}_1 \cap \bar{A}_2) & = & \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} & = & \frac{1}{4} \\
P(A_4) & = & P(A_4 \cap \bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3) & = & \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot 1 & = & \frac{1}{4}
\end{array}$$

Es ist also egal, ob man als erster, zweiter, dritter oder vierter zieht.

**Beispiel** (Alarmanlage).