

# Das Disko-Problem

## bedingte Wahrscheinlichkeiten

## Mengendiagramme

## Baumdiagramme

und

## Vierfeldertafeln

### 1 Das Disko-Problem: Problemstellung

Jens geht samstags abends mit der Wahrscheinlichkeit von  $p$  Prozent in seine Lieblings-Disko. Die Disko hat 4 Räume. Jens hält sich mit derselben Wahrscheinlichkeit in jedem dieser Räume auf. Eines Tages versucht Mona, Jens in der Disko zu treffen. Sie sucht ihn in den Räumen 1, 2 und 3, findet ihn aber nicht. Wie groß ist danach die Wahrscheinlichkeit, Jens in Raum 4 zu finden?

Mit anderen Worten: Wie groß ist die Wahrscheinlichkeit dafür, dass Jens in Raum 4 ist, wenn man bereits weiß, dass er sich nicht in einem der Räume 1, 2 oder 3 aufhält?

Im Folgenden werden zunächst einige Werkzeuge (bedingte Wahrscheinlichkeiten, Mengendiagramme, Baumdiagramme, Vierfeldertafeln) diskutiert, die geeignet sind, die Darstellung derartiger Probleme leichter zugänglich zu machen und zu Lösungen zu gelangen.

## 2 Zum Konzept der bedingten Wahrscheinlichkeit

Wir stützen uns im Folgenden auf die aus der Wahrscheinlichkeitsrechnung stammenden Grundbegriffe: *Wahrscheinlichkeitsraum*, *Ereignis* (bzw. *Ereignismenge*), *Elementarereignis* und *Wahrscheinlichkeit* im Sinne des Wahrscheinlichkeitsbegriffs von *Kolmogorov*<sup>1</sup>. In diesem Text werden wir es aber nur mit endlichen Wahrscheinlichkeitsräumen zu tun haben. Die "Trägermenge" des Wahrscheinlichkeitsraums wird im Folgenden mit dem Symbol  $\Omega$  bezeichnet und wenn keine Verwechslungsgefahr besteht, werden wir  $\Omega$  auch kurz als Wahrscheinlichkeitsraum bezeichnen – man vergleiche hierzu Abschnitt 2.2.3 mit dem Beispiel der Trägermenge (2.6).

### 2.1 Definition des Begriffs

Sind  $A$  und  $B$  Ereignisse eines Wahrscheinlichkeitsraums  $\Omega$ , so ist die *bedingte Wahrscheinlichkeit*  $P(A|B)$  von  $A$  unter der Bedingung  $B$  definiert als

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad (2.1)$$

Man kann sich den Übergang von der Wahrscheinlichkeit des Ereignisses  $A$  zur bedingten Wahrscheinlichkeit  $P(A|B)$  so vorstellen, dass man vom ursprünglich gegebenen Wahrscheinlichkeitsraum  $\Omega$  übergeht zu einem neuen Wahrscheinlichkeitsraum  $\Omega_{\text{neu}} := B$ , wobei jedes Ereignis  $X \subseteq \Omega$  übergeht in  $X \cap B$ .

Auf eine Veranschaulichung dieser Begriffsbildung, wird in Abschnitt 2.3 im Zusammenhang mit der Darstellung von Ereignissen in Mengendiagrammen eingegangen.

## 2.2 Unabhängigkeit von Ereignissen

### 2.2.1 Einige Vorbetrachtungen

Fast jedermann dürfte wohl eine intuitive Vorstellung davon haben, was es bedeutet, dass zwei Ereignisse unabhängig voneinander sind. Zum Beispiel wird man geneigt sein, zu sagen "das Ereignis  $A$  ist unabhängig von Ereignis  $B$ ", wenn das Eintreten von  $B$  das Eintreten von  $A$  in keiner Weise beeinflusst.

Nehmen wir z.B. an, wir würfeln (gleichzeitig) mit zwei normalen Würfeln (Standard-

---

<sup>1</sup> A. N. Kolmogorov, 1903–1987, sowjetischer Mathematiker; auf Kolmogorov geht u.a. die heute übliche Axiomatisierung der Wahrscheinlichkeitsrechnung zurück.

Würfeln, oder auch: "Laplace"<sup>2</sup>-Würfeln); einem roten und einem blauen. Dann wird das Ereignis "die Augenzahl des roten Würfels ist gerade" unabhängig vom Ereignis "die Augenzahl des blauen Würfels ist eine Primzahl" sein müssen; falls nicht, dann wäre der Begriff der Unabhängigkeit vermutlich schlecht definiert.

In ähnlicher Weise muss der Begriff der Unabhängigkeit so gefasst sein, dass jedes Ereignis, das nur über den roten Würfel definiert ist, stets unabhängig von jedem Ereignis ist, das nur über den blauen Würfel definiert ist (kurz: *komponentenweise* definierte Ereignisse sollten unabhängig voneinander sein).

So weit, so gut, aber was ist z.B. mit den Ereignissen "die Augensumme der beiden Würfel ist gerade" oder "die Augensumme ist eine Primzahl" oder "das Produkt der Augenzahlen ist durch 4 teilbar" oder "die Augenzahlen sind teilerfremd" oder so ähnlich. Wir werden derartige Beispiele im Abschnitt 2.2.3 ausführlich behandeln.

### 2.2.2 Möglichkeiten der formalen Definition

Die (stochastische) Unabhängigkeit zweier Ereignisse  $A$  und  $B$  (mit  $P(A) > 0$  und  $P(B) > 0$ ) wird manchmal definiert über die Gültigkeit der Gleichung

$$P(A \cap B) = P(A) \cdot P(B) \quad (2.2)$$

Diese formale Definition ist jedoch gar nicht (und damit schlecht) motiviert. Natürlicher im Sinne der obigen Vorbetrachtungen wäre die Fassung: Die Ereignisse  $A$  und  $B$  sind unabhängig voneinander, wenn ihre bedingten Wahrscheinlichkeiten gleich ihren (nicht-bedingten, "absoluten", "totalen" ) Wahrscheinlichkeiten sind; genauer: Das Ereignis  $A$  ist (stochastisch) *unabhängig* vom Ereignis  $B$ , wenn Folgendes gilt:

$$P(A|B) = P(A) \quad (2.3)$$

Ist Bedingung (2.3) erfüllt, so folgt

$$P(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.4)$$

und somit

$$P(A \cap B) = P(A) \cdot P(B) \quad (2.5)$$

Die Bedingung (2.5) ist wegen  $P(B) > 0$  gleichwertig zu (2.3). Als *Definition* für die Unabhängigkeit ist sie zwar nicht geeignet, aber als *Kriterium* zur Überprüfung der Unabhängigkeit von Ereignissen ist sie sehr nützlich. Sie macht im Übrigen

---

<sup>2</sup> Pierre-Simon Laplace, 1749–1827, französischer Mathematiker, Physiker und Astronom

auch die *Symmetrie* des Begriffs der stochastischen Unabhängigkeit deutlich (man vergleiche dazu die folgende Aufgabe).

*Aufgabe:* Zeigen Sie, dass für beliebige Ereignisse  $A$  und  $B$  gilt: Aus  $P(A|B) = P(A)$  folgt  $P(B|A) = P(B)$ .

Der Begriff der stochastischen Unabhängigkeit ist somit *symmetrisch* in Bezug auf die beiden Ereignisse. Wenn das Ereignis  $A$  unabhängig von  $B$  ist, dann ist auch  $B$  unabhängig von  $A$ , und umgekehrt. Die Aussagen

Das Ereignis  $A$  ist unabhängig von  $B$ .

Das Ereignis  $B$  ist unabhängig von  $A$ .

Die Ereignisse  $A$  und  $B$  sind voneinander unabhängig.

sind also gleichwertig.

Die Gleichung  $P(A) = \frac{P(A \cap B)}{P(B)}$  lässt sich auch folgendermaßen deuten: Das "Gewicht" (die Repräsentanz) von  $A$  in  $\Omega$  ist gleich groß wie das Gewicht von  $A \cap B$  in  $B$ .

Noch anders ausgedrückt: Die Bedingung  $B$  oder die Information " $B$  ist der Fall" ist ohne Relevanz für die Wahrscheinlichkeit, mit der das Ereignis  $A$  eintritt. Dies wird besonders deutlich im Zusammenhang mit der Darstellung anhand von Baumdiagrammen (siehe Abschnitt 2.4).

Man beachte jedoch: In [Henze, 2013] wird anhand eines treffenden Beispiels darauf aufmerksam gemacht, dass sich *stochastische Unabhängigkeit* und *reale Beeinflussung* nicht notwendigerweise gegenseitig ausschließen.

*Aufgabe:* Geben Sie Beispiele für abhängige und für unabhängige Ereignisse an.

*Aufgabe* (Diskussion einiger Sonderfälle): Wie verhält es sich mit der Unabhängigkeit der Ereignisse  $A$  und  $B$ ,

- (a) wenn  $A$  und  $B$  disjunkt (elementfremd) sind (m.a.W. wenn:  $A \cap B = \emptyset$ )
- (b) wenn  $A = \Omega$  oder  $B = \Omega$  ist
- (c) wenn  $A = B$  ist
- (d) wenn  $A \subseteq B$  ist

Geben Sie je eine Antwort

\* auf der Basis des umgangssprachlichen Gebrauchs des Begriffs "unabhängig"

\* auf der Basis der formalen Definition

und vergleichen Sie, ob und in wie weit die Ergebnisse übereinstimmen.

### 2.2.3 Ein ausführliches Beispiel zum Begriff der Unabhängigkeit von Ereignissen

Wir betrachten das Experiment "(gleichzeitiges) Würfeln mit zwei (Standard-) Würfeln; einer rot, der andere blau". Das jeweilige Ergebnis werde als Zahlenpaar  $(a, b)$  dargestellt, wobei  $a$  die Augenzahl des roten und  $b$  die Augenzahl des blauen Würfels ist. Es gibt 36 gleichwahrscheinliche Elementarereignisse  $(x, y)$ ,  $(x = 1, \dots, 6, y = 1, \dots, 6)$ . Jedes dieser Elementarereignisse hat die Wahrscheinlichkeit  $\frac{1}{36}$ .

Die Ergebnismenge ("Grundmenge") dieses Wahrscheinlichkeitsraums, also die Gesamtheit aller Elementarereignisse, lässt sich folgendermaßen darstellen:

$$\begin{aligned} \Omega := \{ & (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), \\ & (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), \\ & (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), \\ & (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\ & (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), \\ & (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \} \end{aligned} \quad (2.6)$$

$\Omega$  ist die *Trägermenge* dieses Wahrscheinlichkeitsraums. Wenn keine Verwechslungsgefahr vorliegt, wird  $\Omega$ , wie in der Einleitung zu diesem Abschnitt beschrieben, gelegentlich auch selbst als Wahrscheinlichkeitsraum bezeichnet.

Man sollte meinen, dass bei irgendwelchen "aus der Luft gegriffenen" Ereignissen eine gute Chance dafür besteht, dass sie unabhängig voneinander sind. Solche Ereignisse könnten z.B. sein:

A1: Die Summe der Augenzahlen ist gerade.

In der aufzählenden Form:

$$\begin{aligned} A1 = \{ & (1, 1), (1, 3), (1, 5), \\ & (2, 2), (2, 4), (2, 6), \\ & (3, 1), (3, 3), (3, 5), \\ & (4, 2), (4, 4), (4, 6), \\ & (5, 1), (5, 3), (5, 5), \\ & (6, 2), (6, 4), (6, 6) \} \end{aligned} \quad (2.7)$$

A2: Das Produkt der Augenzahlen ist gerade.

In der aufzählenden Form:

$$\begin{aligned}
 A2 = \{ & (1, 2), (1, 4), (1, 6), \\
 & (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), \\
 & (3, 2), (3, 4), (3, 6), \\
 & (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\
 & (5, 2), (5, 4), (5, 6), \\
 & (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \}
 \end{aligned} \tag{2.8}$$

A3: Die Summe der Augenzahlen ist eine Primzahl.

In der aufzählenden Form:

$$\begin{aligned}
 A3 = \{ & (1, 1), (1, 2), (1, 4), (1, 6), (2, 1), (2, 3), (2, 5), \\
 & (3, 2), (3, 4), (4, 1), (4, 3), (5, 2), (5, 6), (6, 1), (6, 5) \}
 \end{aligned} \tag{2.9}$$

A4: Die Summe der Augenzahlen ist eine Quadratzahl.

In der aufzählenden Form:

$$\begin{aligned}
 A4 = \{ & (1, 3), (2, 2), (3, 1), (3, 6), \\
 & (4, 5), (5, 4), (6, 3) \}
 \end{aligned} \tag{2.10}$$

A5: Die Summe der Augenzahlen ist quadratfrei

In der aufzählenden Form:

$$\begin{aligned}
 A5 = \{ & (1, 1), (1, 2), (1, 4), (1, 5), (1, 6), \\
 & (2, 1), (2, 3), (2, 4), (2, 5), \\
 & (3, 2), (3, 3), (3, 4), \\
 & (4, 1), (4, 2), (4, 3), (4, 6), \\
 & (5, 1), (5, 2), (5, 5), (5, 6), \\
 & (6, 1), (6, 4), (6, 5) \}
 \end{aligned} \tag{2.11}$$

A6: Das Produkt der Augenzahlen ist quadratfrei

In der aufzählenden Form:

$$\begin{aligned}
 A6 = \{ & (1, 1), (1, 2), (1, 3), (1, 5), (1, 6), \\
 & (2, 1), (2, 3), (2, 5), \\
 & (3, 1), (3, 2), (3, 5),
 \end{aligned} \tag{2.12}$$

$$(5, 1), (5, 2), (5, 3), (5, 6), \\ (6, 1), (6, 5)\}$$

A7: Die Augenzahlen sind teilerfremd

In der aufzählenden Form:

$$\begin{aligned} A7 = \{ & (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), \\ & (2, 1), (2, 3), (2, 5), \\ & (3, 1), (3, 2), (3, 4), (3, 5), \\ & (4, 1), (4, 3), (4, 5), \\ & (5, 1), (5, 2), (5, 3), (5, 4), (5, 6), \\ & (6, 1), (6, 5)\} \end{aligned} \quad (2.13)$$

A8: Die Augenzahlen sind "pythagoreisch", d.h.  $rot^2 + blau^2$  ist eine Quadratzahl.

In der aufzählenden Form:

$$A8 = \{(3, 4), (4, 3)\} \quad (2.14)$$

A9: (experimentell)  $rot^{blau} + blau^{rot} \equiv 2 \pmod{5}$

In der aufzählenden Form:

$$\begin{aligned} A9 = \{ & (1, 1), (1, 6), \\ & (2, 3), (2, 4), (2, 5), (3, 2), \\ & (4, 2), (4, 4), (4, 6), \\ & (5, 2), (6, 1), (6, 4), (6, 6)\} \end{aligned} \quad (2.15)$$

A10: Die Summe der Augenzahlen ist kleiner als 7.

In der aufzählenden Form:

$$\begin{aligned} A10 = \{ & (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), \\ & (2, 1), (2, 2), (2, 3), (2, 4), \\ & (3, 1), (3, 2), (3, 3), \\ & (4, 1), (4, 2), \\ & (5, 1)\} \end{aligned} \quad (2.16)$$

Es mag durchaus als überraschend erscheinen, dass alle diese Ereignisse paarweise (stochastisch) abhängig voneinander sind.

Wir "buchstabieren" dies im Folgenden an einem konkreten Fall durch: Sind die

Ereignisse  $A1$  und  $A2$  voneinander unabhängig oder nicht? Die Ereignismengen  $A1$  und  $A2$  sind oben in der Aufzählungsform beschrieben (siehe 2.7 und 2.8). Weiterhin ist

$$A1 \cap A2 = \{(2, 2), (2, 4), (2, 6), (4, 2), (4, 4), (4, 6), (6, 2), (6, 4), (6, 6)\} \quad (2.17)$$

Da jedes Elementarereignis die Wahrscheinlichkeit  $\frac{1}{36}$  hat, ist:

$$P(A1 \cap A2) = \frac{9}{36} = \frac{1}{4}$$

$$P(A1) = \frac{18}{36} = \frac{1}{2}$$

$$P(A2) = \frac{27}{36} = \frac{3}{4}$$

$$P(A1) \cdot P(A2) = \frac{3}{8}$$

Das Kriterium (2.5) ist also nicht erfüllt und die Ereignisse  $A1$  (die Summe der Augenzahlen ist gerade) und  $A2$  (das Produkt der Augenzahlen ist gerade) sind nicht voneinander unabhängig.

*Aufgabe:* Zeigen Sie für jedes der restlichen obigen Ereignispaare  $(Ai, Aj)$  mit  $i, j = 1, \dots, 10$ , dass die beiden Ereignisse nicht unabhängig voneinander sind. Schreiben Sie dazu nach Möglichkeit ein Programm in einer geeigneten Programmiersprache<sup>3</sup>, an das Sie gewisse Routine-Teilaufgaben (insbesondere: Beschreibung der Ereignisse in Listen- oder Mengenform und Bestimmung des Durchschnitts der beiden Ereignismengen) delegieren können.

Man sollte meinen, dass man ein Ereignis nur hinreichend "skurril" zu definieren braucht, damit es unabhängig von anderen Ereignissen ist. Aber auch darin kann man sich täuschen. Betrachten wir z.B. die Ereignisse

$B1: \sin(\text{rot}) + \cos(\text{blau}) < r$  ( $r$ : reelle Zahl)

In der aufzählenden Form sei (für  $r = 0$ ):

$$\begin{aligned} B1 := \{ & (1, 5), (2, 5), (3, 4), (3, 5), (3, 6), \\ & (4, 3), (4, 4), (4, 5), (4, 6), \\ & (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), \\ & (6, 3), (6, 4), (6, 5), (6, 6) \} \end{aligned} \quad (2.18)$$

Es stellt sich heraus, dass keines der Ereignisse  $A1, \dots, A10$  unabhängig von  $B1$  ist.

---

<sup>3</sup>das könnte auch ein Computeralgebra System sein



Dasselbe gilt für  $r = 1$ . In der aufzählenden Form sei:

$$\begin{aligned} B2 := \{ & (1, 3), (1, 4), (1, 5), (1, 6), (2, 4), (2, 5), (2, 6), \\ & (3, 1), (3, 3), (3, 4), (3, 5), (3, 6), \\ & (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\ & (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), \\ & (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \} \end{aligned} \quad (2.19)$$

Weiterhin sei für  $r = 0.5$ :

$$\begin{aligned} B3 := \{ & (1, 4), (1, 5), (2, 4), (2, 5), \\ & (3, 3), (3, 4), (3, 5), (3, 6), \\ & (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\ & (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), \\ & (6, 3), (6, 4), (6, 5), (6, 6) \} \end{aligned} \quad (2.20)$$

Dann sind  $B1$ ,  $B2$  und  $B3$  paarweise voneinander abhängig.  $B1$  und  $B2$  sind von allen  $A_i$  ( $i = 1, \dots, 10$ ) abhängig.  $B3$  ist unabhängig von  $A1$  und  $A2$ , nicht aber von  $A3, \dots, A10$ .

Die oben betrachteten Ereignisse  $A(i)$  ( $i = 1, \dots, 10$ ) und  $B(i)$  ( $i = 1, \dots, 3$ ) waren "symmetrisch" in dem Sinne, dass man die Bestandteile "roter Würfel" und "blauer Würfel" vertauschen konnte (so z.B. bei der Eigenschaft "die Augensumme ist gerade"). Dadurch wurde zwar die Gefahr einer "komponentenweisen" Beschreibung der Ereignisse vermieden, vielleicht aber verdirbt gerade diese Symmetrie-Eigenschaft die Unabhängigkeit? (Wir werden weiter unten sehen, dass dieser Eindruck ebenfalls täuscht.)

Als nächstes betrachten wir eine Eigenschaft, die nicht mehr symmetrisch ist.

$C1$ : Das Würfelergebnis ( $rot, blau$ ), gedeutet als zweistellige Dezimalzahl, ist eine Primzahl.

Mit anderen Worten:  $10 \cdot rot + blau$  ist eine Primzahl.

In der aufzählenden Form ist:

$$C1 = \{(1, 1), (1, 3), (2, 3), (3, 1), (4, 1), (4, 3), (5, 3), (6, 1)\} \quad (2.21)$$

Dann ist  $C1$  unabhängig von  $A1$  und abhängig von  $A2, \dots, A10, B1, B2, B3$ .

Die folgenden Beispiele zeigen, dass die Eigenschaften der Symmetrie und der Unabhängigkeit durchaus miteinander verträglich sein können. Wir betrachten:

$D1$ : Die Augensumme ist durch 2 teilbar.

Mit anderen Worten:  $D1 = A1$  (vgl. 2.7).

$D2$ : Die Augensumme ist durch 3 teilbar.

In der aufzählenden Form:

$$D2 = \{(1, 2), (1, 5), (2, 1), (2, 4), (3, 3), (3, 6), \\ (4, 2), (4, 5), (5, 1), (5, 4), (6, 3), (6, 6)\} \quad (2.22)$$

Dann sind  $D1$  und  $D2$  voneinander unabhängig.

$D3$ : Die Augensumme ist durch 4 teilbar.

In der aufzählenden Form:

$$D3 = \{(1, 3), (2, 2), (2, 6), (3, 1), (3, 5), \\ (4, 4), (5, 3), (6, 2), (6, 6)\} \quad (2.23)$$

$D4$ : Die Augensumme ist durch 5 teilbar.

In der aufzählenden Form:

$$D4 = \{(1, 4), (2, 3), (3, 2), \\ (4, 1), (4, 6), (5, 5), (6, 4)\} \quad (2.24)$$

$D5$ : Die Augensumme ist durch 6 teilbar.

In der aufzählenden Form:

$$D5 = \{(1, 5), (2, 4), (3, 3), \\ (4, 2), (5, 1), (6, 6)\} \quad (2.25)$$

$D6$ : Das Produkt der Augenzahlen ist durch 2 teilbar.

In der aufzählenden Form:

$$D6 = \{(1, 2), (1, 4), (1, 6), \\ (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), \\ (3, 2), (3, 4), (3, 6), \\ (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\ (5, 2), (5, 4), (5, 6), \\ (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\} \quad (2.26)$$

$D7$ : Das Produkt der Augenzahlen ist durch 3 teilbar.

In der aufzählenden Form:

$$\begin{aligned}
 D7 = \{ & (1, 3), (1, 6), (2, 3), (2, 6), \\
 & (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), \\
 & (4, 3), (4, 6), (5, 3), (5, 6), \\
 & (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \}
 \end{aligned} \tag{2.27}$$

D8: Das Produkt der Augenzahlen ist durch 4 teilbar.

In der aufzählenden Form:

$$\begin{aligned}
 D8 = \{ & (1, 4), (2, 2), (2, 4), (2, 6), (3, 4), \\
 & (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\
 & (5, 4), (6, 2), (6, 4), (6, 6) \}
 \end{aligned} \tag{2.28}$$

D9: Das Produkt der Augenzahlen ist durch 5 teilbar.

In der aufzählenden Form:

$$\begin{aligned}
 D9 = \{ & (1, 5), (2, 5), (3, 5), (4, 5), \\
 & (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), \\
 & (6, 5) \}
 \end{aligned} \tag{2.29}$$

D10: Das Produkt der Augenzahlen ist durch 6 teilbar.

In der aufzählenden Form:

$$\begin{aligned}
 D10 = \{ & (1, 6), (2, 3), (2, 6), (3, 2), (3, 4), (3, 6), (4, 3), (4, 6), \\
 & (5, 6), (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \}
 \end{aligned} \tag{2.30}$$

Insgesamt stellt sich heraus, dass von den obigen Beispielen nur die folgenden Ereignisse jeweils paarweise unabhängig voneinander sind:

$$\begin{aligned}
 & (A1, B3), (A1, C1), (A1, D2), (A1, D7), (A2, B3), (A2, D2), (A2, D7), \\
 & (B2, D2), (B3, D1), (B3, D6), \\
 & (C1, D1), \\
 & (D1, D2), (D1, D7), (D2, D6), (D2, D8), (D6, D7)
 \end{aligned}$$

Bei 24 relativ willkürlich herausgegriffenen Eigenschaften (manifestiert durch die Ereignismengen A1, ..., A10, B1, B2, B3, C1, D1, ..., D10) ergeben sich formal  $576 (= 24 \cdot 24)$  Ereignispaare. Wenn man davon die "identischen" Paare vom Typ (X, X) abzieht und die "symmetrischen" Paare (Y, Z) und (Z, Y) nur einmal zählt,

bleiben  $(24 \cdot 24 - 24)/2$  ( $= 276$ ) echt verschiedene Paare<sup>4</sup> übrig. Und von diesen sind nur 16 Paare (also weniger als 6%) voneinander unabhängig.

*Aufgabe:* Wenden Sie auf jedes der restlichen obigen Ereignispaare das Kriterium (2.5) für stochastische Unabhängigkeit an und entscheiden Sie, ob die jeweiligen Ereignisse unabhängig voneinander sind oder nicht. Benutzen Sie dazu nach Möglichkeit das bereits geschriebene Programm – u.U. mit gewissen Anpassungen oder Erweiterungen.

### 2.2.4 Weitere Beispiele zum Begriff der Unabhängigkeit von Ereignissen

Man könnte meinen, dass die Abhängigkeit bzw. Unabhängigkeit von Ereignissen ausschließlich von den die Ereignisse definierenden Eigenschaften abhängt – und nicht z.B. vom Umfang des jeweiligen Wahrscheinlichkeitsraumes. Aber auch das trifft nicht zu.

Um dies zu illustrieren verwenden wir an Stelle von zwei Würfeln zwei Laplace'sche "Glücksräder" mit jeweils  $n$  gleichgroßen Sektoren. Jeder Sektor sei mit einer der Zahlen  $1, 2, 3, \dots, n$  bezeichnet. Die Elementarereignisse seien (ähnlich wie im Würfel-Beispiel) als Paare der Form  $(i, j)$  mit  $i, j \in \{1, 2, 3, \dots, n\}$  dargestellt. Die Wahrscheinlichkeit für jedes der Elementarereignisse ist  $\frac{1}{n^2}$ . Entsprechend dem Würfel-Beispiel sei ein Glücksrad blau (1. Komponente) und eines rot (2. Komponente).

Wir betrachten die Ereignisse:

\* ASP: Die Summe der Glücksrad-Zahlen ist eine Primzahl.

\* AD3: Die Differenz der Glücksrad-Zahlen (*rot* minus *blau*) ist durch 3 teilbar.

Eine Analyse für Glücksräder von jeweils  $n = 2, 3, \dots, 40$  Sektoren liefert die Ergebnisse: Beim Glücksrädern mit 6, 12, 21, 27, 36, und 39 Sektoren sind die Ereignisse unabhängig; in den anderen Fällen sind sie abhängig.

*Aufgabe:* Überprüfen Sie stichprobenartig die Fälle  $n = 6, 13, 24, 27$  anhand des Kriteriums (2.5) auf stochastische Unabhängigkeit.

*Fazit:* Im Hinblick auf die Einschätzung der *stochastischen Unabhängigkeit* kann die "natürliche" Intuition trügen; sie hat zudem relativ wenig mit den "intrinsischen" Eigenschaften der jeweiligen Ereignisse zu tun.

---

<sup>4</sup> Eine andere Art der Berechnung:  $1 + 2 + 3 + \dots + 23 = 276$

## 2.3 Darstellung der bedingten Wahrscheinlichkeit im Mengendiagramm

(Venn<sup>5</sup>- bzw. Euler<sup>6</sup>-Diagramm)

Als Venn-Diagramme (oder Euler-Diagramme) bezeichnet man schematische flächenhafte Darstellungen von Mengen, insbesondere zur Veranschaulichung der Teilmengenrelation, sowie von Durchschnitt, Vereinigung, Komplement und Differenz von Mengen. Diese Darstellungsform eignet sich insbesondere auch für die Veranschaulichung von Sachverhalten im Zusammenhang mit der bedingten Wahrscheinlichkeit.

Man kann sich das Konzept der bedingten Wahrscheinlichkeit so vorstellen, dass man vom Wahrscheinlichkeitsraum  $\Omega$  übergeht zum Wahrscheinlichkeitsraum  $\Omega_{neu} := \Omega \cap B = B$ , wobei jedes Ereignis  $X \subseteq \Omega$  durch  $X \cap B$  ersetzt wird. Der Übergang von  $\Omega$  zu  $\Omega_{neu}$  ist also eine Art "Neu-Focussierung" des Wahrscheinlichkeitsraums  $\Omega$  auf  $\Omega \cap B$  (vgl. folgende Abbildung).

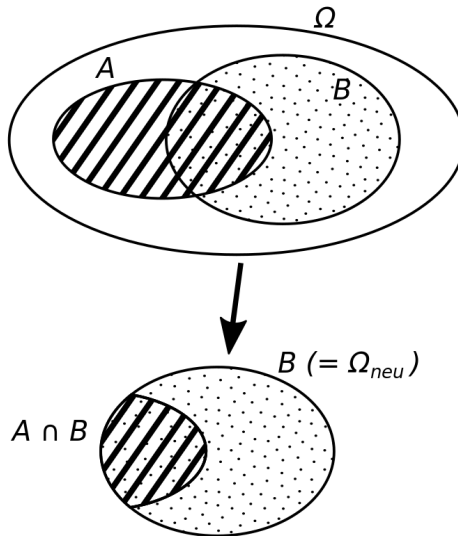


Abbildung 2.1: Veranschaulichung zur bedingten Wahrscheinlichkeit  $P(A|B)$

In den folgenden Abbildungen sind die "Lagemöglichkeiten" der Ereignismengen  $A, B, \bar{A}$  und  $\bar{B}$  dargestellt; den Wahrscheinlichkeiten ("Gewichten") der Ereignisse mögen dabei die jeweiligen Flächenanteile entsprechen ( $a := P(A)$ ,  $b := P(B)$ ,  $c := P(A \cap B)$ ).

<sup>5</sup> John Venn, englischer Mathematiker 1834–1923

<sup>6</sup> Leonhard Euler, schweizer Mathematiker 1707–1783

Die Berechnung der anderen Wahrscheinlichkeiten erfolgt mit Hilfe des Axioms, dass für die Wahrscheinlichkeit disjunkter (überschneidungsfreier) Ereignisse  $X$  und  $Y$  gilt:  $P(X \cup Y) = P(X) + P(Y)$ .

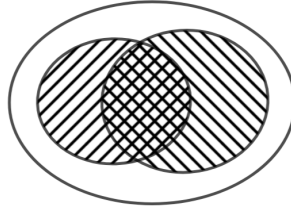


Abbildung 2.2: Venn-Diagramm zur Menge  $A \cap B$   
Gewicht =  $c$

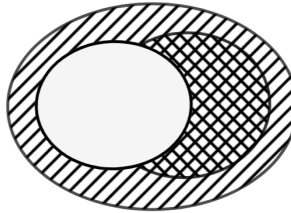


Abbildung 2.3: Venn-Diagramm zur Menge  $\bar{A} \cap B$   
Gewicht =  $b - c$

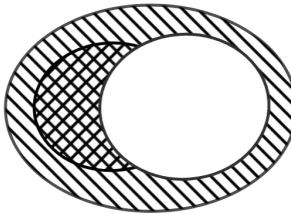


Abbildung 2.4: Venn-Diagramm zur Menge  $A \cap \bar{B}$   
Gewicht =  $a - c$

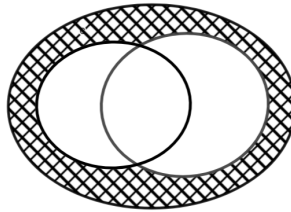


Abbildung 2.5: Venn-Diagramm zur Menge  $\bar{A} \cap \bar{B} (= \overline{A \cup B})$   
Gewicht =  $1 - (a + b) + c$

Abbildung 2.5 lässt sich auch als Veranschaulichung des *Gesetzes von De Morgan*<sup>7</sup> deuten:

$$\bar{A} \cap \bar{B} = \overline{A \cup B} \quad (2.31)$$

<sup>7</sup> Augustus De Morgan, englischer Mathematiker, 1806–1871

## 2.4 Darstellung im Baumdiagramm

Der Begriff "Baumdiagramm" ist hochgradig anschaulich und dürfte praktisch selbsterklärend sein. Man kann sich ein Baumdiagramm wie ein System zusammenhängender Kanäle vorstellen, in die "oben" eine Menge von Kugeln oder eine Flüssigkeit hineingeschüttet wird, die sich entsprechend der angedeuteten "Kanalkapazitäten" verteilen. Eine wohlbekannte, historisch bedeutsame, mechanische Form solcher Baumdiagramme ist das sogenannte *Galton<sup>8</sup>-Brett*.

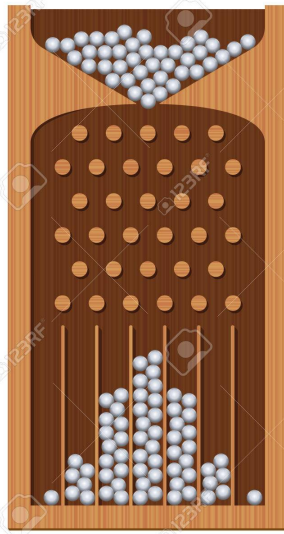


Abbildung 2.6: Galton-Brett

Das Konzept der bedingten Wahrscheinlichkeit lässt sich als Teil eines von der Wurzel bis zu einem Endpunkt ("Blatt") verlaufenden Pfades im folgenden Baumdiagramm deuten (vgl. Abbildung 2.7). Konkret kommt die bedingte Wahrscheinlichkeit  $P(A|B)$  in dem am weitesten links liegenden Pfad vor. Die anderen Pfade sind wegen des Gesamtüberblicks mit eingezeichnet. Dies ermöglicht auch gewisse Plausibilitätsbetrachtungen wie z.B.: Die Summe aller Pfadwahrscheinlichkeiten ist gleich 1. Die Begründung der (berechneten) Wahrscheinlichkeiten erfolgt mit Hilfe der Mengendiagramme 2.2, 2.3, 2.4 und 2.5 und der Tatsache, dass für die Wahrscheinlichkeit disjunkter Ereignisse  $X$  und  $Y$  gilt:  $P(X \cup Y) = P(X) + P(Y)$ . Den Endpunkten des Baumdiagramms entsprechen (grob gesagt) die doppelt schraffierten Mengen in den Abbildungen von Abschnitt 2.3.

<sup>8</sup> Sir Francis Galton, 1822–1911, englischer Naturforscher

Das Baumdiagramm verdeutlicht die

*Pfadregel für Baumdiagramme:* Die Wahrscheinlichkeit für einen Pfad im Baumdiagramm ist gleich dem Produkt der Einzelwahrscheinlichkeiten entlang des Pfades.

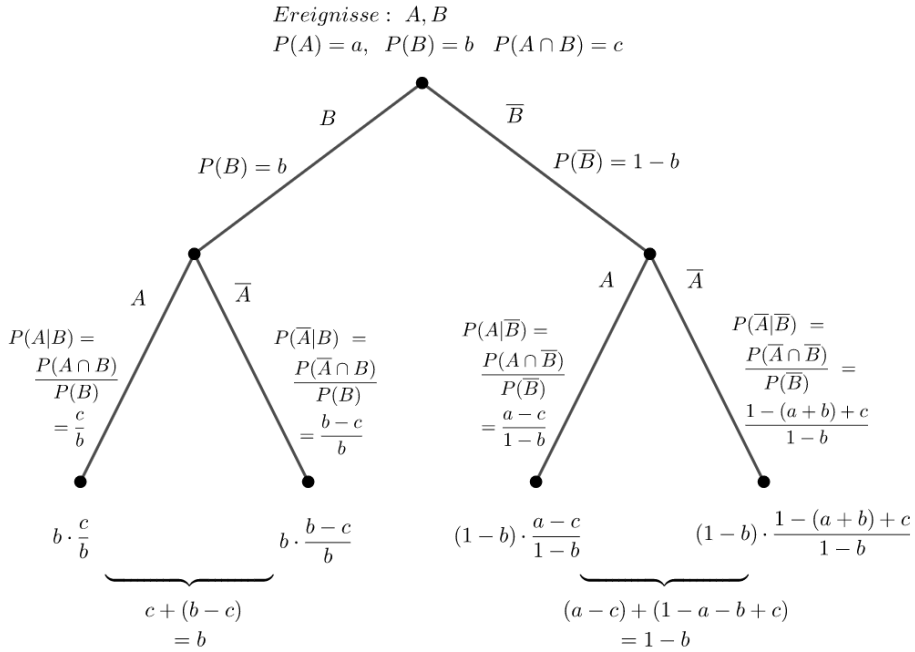


Abbildung 2.7: Baumdiagramm zur bedingten Wahrscheinlichkeit  $P(A|B)$

Die Darstellung im Baumdiagramm suggeriert eine zeitliche Reihenfolge (Ereignis  $B$  vor Ereignis  $A$ ), die aber nicht immer gegeben sein muss. Beim Würfeln mit zwei Standard-Würfeln (vgl. das Beispiel in Abschnitt 2.2.3) finden die dort beschriebenen Ereignis-Paare jeweils gleichzeitig statt.



## 2.5 Die Vierfeldertafel

Die Zusammenhänge zwischen zwei Ereignissen und ihren Gegenereignissen werden oft auch mit Hilfe des Darstellungsmittels der *Vierfeldertafel* beschrieben. Um dies zu demonstrieren, betrachten wir wieder die Ereignisse  $A$ ,  $B$  und ihre Gegenereignisse  $\bar{A}$  und  $\bar{B}$ . Die zugehörige Vierfeldertafel sieht folgendermaßen aus.

Vierfeldertafel	$B$	$\bar{B}$	Zeilensummen
$A$	$P(A \cap B)$	$P(A \cap \bar{B})$	$P(A)$
$\bar{A}$	$P(\bar{A} \cap B)$	$P(\bar{A} \cap \bar{B})$	$P(\bar{A})$
Spaltensummen	$P(B)$	$P(\bar{B})$	1

Tabelle 1: Vierfeldertafel zu den Ereignissen  $A$  und  $B$

Bei dieser Darstellung stellt sich die Frage, wo die bedingte Wahrscheinlichkeit  $P(A|B)$  zu sehen ist. Der Übergang von  $\Omega$  auf  $B = \Omega_{neu}$  bedeutet "Normierung auf  $B$ ":  $P(X|B) = \frac{P(X \cap B)}{P(B)}$  für  $X \subseteq \Omega$  und insbesondere  $P(B|B) = \frac{P(B \cap B)}{P(B)} = 1$ . Für  $X = \bar{B}$  ist  $X \cap B = \bar{B} \cap B = \emptyset$  und in der Vierfeldertafel spielt nur noch die Spalte mit "Kopf"  $B$  eine Rolle; die Spalte mit Kopf  $\bar{B}$  wird ausgeblendet.

Vierfeldertafel	$B \cap B = B = \Omega_{neu}$	$\bar{B} \cap B = \emptyset$	Zeilensummen
$A$	$P(A B) = \frac{P(A \cap B)}{P(B)}$	0	$P(A B)$
$\bar{A}$	$P(\bar{A} B) = \frac{P(\bar{A} \cap B)}{P(B)}$	0	$P(\bar{A} B)$
Spaltensummen	$\frac{P(A \cap B) + P(\bar{A} \cap B)}{P(B)} = 1$	0	1

Tabelle 2: Vierfeldertafel zu bedingten Wahrscheinlichkeiten

Mit den Bezeichnungen von Abbildung 2.7 lässt sich die Vierfeldertafel auch wie folgt darstellen.

Vierfeldertafel	$B$	$\overline{B}$	Zeilensummen
$A$	$c$	$a - c$	$a$
$\overline{A}$	$b - c$	$1 - (a + b) + c$	$1 - a$
Spaltensummen	$b$	$1 - b$	$1$

Tabelle 3: Vierfeldertafel vereinfacht

In manchen Situationen (wir werden einige im Abschnitt 2.6 kennenlernen) ist zunächst mal eine Vierfeldertafel in der Art von Tabelle 3 gegeben. Es wird sich die Frage stellen, wie ein zugehöriges Baumdiagramm dazu aussehen könnte.

## 2.6 Von der Vierfeldertafel zum Baumdiagramm

A unter B oder B unter A ?  
(bzw.  $P(A|B)$  oder  $P(B|A)$  ? )

Baumdiagramme sind vom Design her mehrstufig (in den hier diskutierten Fällen meist zweistufig). Diese Form der Darstellung passt z.B. dann besonders gut zu den diskutierten Ereignissen, wenn diese in einer zeitlichen Reihenfolge stehen.

*Beispiel:* Zweimaliges Würfeln mit einem Würfel.

Ereignis  $A$ : der erste Wurf ergibt eine Primzahl,

Ereignis  $B$ : der zweite Wurf ergibt eine gerade Zahl.

Bei diesem Beispiel ist ziemlich klar, dass man wohl das Ergebnis des ersten Wurfs in der ersten Ebene und das Ergebnis des zweiten Wurfs in der zweiten Ebene des Baumdiagramms darstellen wird.

Anders ist es beim Beispiel: Gleichzeitiger Wurf mit zwei Würfeln; einer blau, einer rot (vgl. Abschnitt 2.2.3). Da ist nicht von vorn herein klar, welchen Würfel man in welcher Ebene des Baumdiagramms darstellen sollte.

Die Darstellung im Baumdiagramm ist also fundamental asymmetrisch. Diese Asymmetrie ist bei der Darstellung im Mengendiagramm oder in der Vierfeldertafel nicht unmittelbar gegeben.

Wir betrachten nochmals die in Tabelle 1 gegebene Vierfeldertafel. Sie könnte im Prinzip jedem der beiden folgenden Baumdiagramme entsprechen.

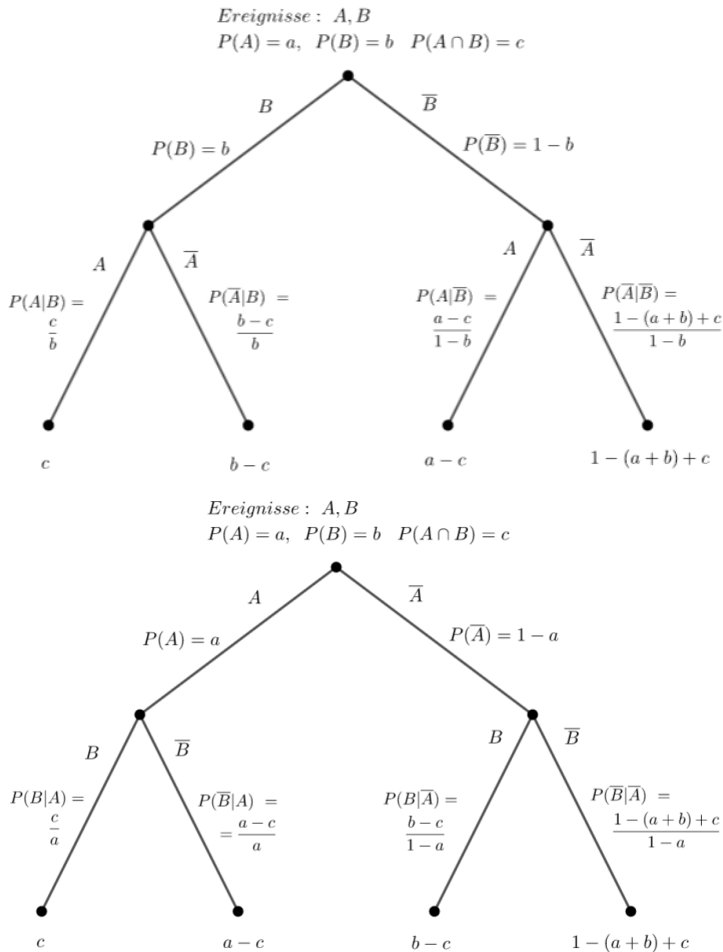


Abbildung 2.8: Baumdiagramme zu bedingten Wahrscheinlichkeiten

*Aufgabe:* Diskutieren Sie, unter welchen Bedingungen  $P(A|B) = P(B|A)$  ist, und zwar

- (1.) allgemein
- (2.) im Kontext der beiden Baumdiagramme in Abbildung 2.8

## 2.7 Das "Simpson-Paradoxon"

Vierfeldertafeln werden oft auch im Zusammenhang mit statistischen Fragestellungen eingesetzt. In diesen Fällen verwendet man in der Regel *relative* oder *absolute Häufigkeiten* an Stelle von *Wahrscheinlichkeiten*.

Statistische Daten werden oft "aggregiert" (d.h. zu "Klumpen" zusammengefasst). Dass und wie dies zu überraschenden, der Intuition zuwider laufenden ("paradoxen") Situationen führen kann, sei am folgenden Beispiel erläutert. Wir werden sehen, dass das "Simpson<sup>9</sup>-Paradoxon" kein eigentliches Paradoxon ist. Es lässt sich durch eine mathematische Analyse problemlos erklären. Wir wollen trotzdem aus historischen Gründen im Folgenden weiter vom Simpson-Paradoxon sprechen. Simpson beschrieb das Phänomen im Jahre 1951. Er war aber nicht der Erste, der sich damit beschäftigte. Bereits 1899 hatte Karl Pearson<sup>10</sup> einen ähnlichen Sachverhalt beschrieben.

### *Fahrschul-Prüfung*

Die Schülerinnen und Schüler einer Fahrschule legen die Prüfung an zwei aufeinanderfolgenden Tagen ab. Dabei ergeben sich die folgenden Prüfungsergebnisse:

1. Tag:

Fahrschul-Prüfung	Männer	Frauen
bestanden	15	22
durchgefallen	2	4
Spaltensummen	17	26

*Tabelle 4: Fahrschul-Prüfung 1. Tag*

2. Tag:

Fahrschul-Prüfung	Männer	Frauen
bestanden	13	5
durchgefallen	6	3
Spaltensummen	19	8

*Tabelle 5: Fahrschul-Prüfung 2. Tag*

Die Durchfallquoten sind somit:

<sup>9</sup> Edward Hugh Simpson, 1922–2019, englischer Statistiker

<sup>10</sup> Karl Pearson, 1857–1936, englischer Mathematiker und Statistiker

Durchfallquoten	Männer	Frauen
1. Tag	0,1176	0,1538
2. Tag	0,3158	0,3750

Tabelle 6: Fahrschul-Prüfung Durchfallquoten

Die Durchfallquoten der Männer sind also an beiden Tagen niedriger als die Durchfallquoten der Frauen. Legt man die Daten der beiden Tage zusammen, so ergibt sich das folgende Bild.

Fahrschul-Prüfung	Männer	Frauen
bestanden	0,7778	0,7941
durchgefallen	0,2222	0,2059

Tabelle 7: Fahrschul-Prüfung: beide Tage zusammen

*Fazit:* An jedem der beiden Tage ist die Durchfallquote der Männer niedriger als die der Frauen. Legt man jedoch die Ergebnisse der beiden Tage zusammen, so ist die Durchfallquote der Männer höher als die der Frauen. Dieser Sachverhalt ist ein typisches Beispiel für das *Simpson-Paradoxon*.

*Aufgabe:* Erstellen Sie ein Tabellenkalkulationsblatt, um das Simpson-Paradoxon interaktiv nachzuspielen und Zusammenhänge zu erkunden.

*Diskussion des Simpson-Paradoxons:* Wir verwenden im Folgenden die Bezeichnungen

$dm1$ : Anzahl der durchgefallenen Männer am Tag 1

$am1$ : Anzahl aller geprüften Männer am Tag 1

$dw1$ : Anzahl der durchgefallenen Frauen am Tag 1

$aw1$ : Anzahl aller geprüften Frauen am Tag 1

$dm2$ : Anzahl der durchgefallenen Männer am Tag 2

$am2$ : Anzahl aller geprüften Männer am Tag 2

$dw2$ : Anzahl der durchgefallenen Frauen am Tag 2

$aw2$ : Anzahl aller geprüften Frauen am Tag 2

Dann sind  $\frac{dm1}{am1}$ ,  $\frac{dw1}{aw1}$ ,  $\frac{dm2}{am2}$ ,  $\frac{dw2}{aw2}$ , die entsprechenden Durchfallquoten. Die Durchfallquoten für beide Tage zusammen sind gegeben durch:

$\frac{dm1+dm2}{am1+am2}$ : aggregierte Durchfallquote der Männer

$\frac{dw1+dw2}{aw1+aw2}$ : aggregierte Durchfallquote der Frauen.

Die Daten  $(am1 + am2, dm1 + dm2)$  bzw.  $(aw1 + aw2, dw1 + dw2)$  lassen sich deuten als die komponentenweisen Summen ("Vektorsummen") von  $(am1, dm1)$  und  $(am2, dm2)$  bzw.  $(aw1, dw1)$  und  $(aw2, dw2)$ .

Wir wollen im Folgenden die Durchfallquote  $\frac{y}{x}$  im Koordinatensystem durch die Strecke  $\overline{(0,0), (x,y)}$  vom Ursprung zum Punkt  $(x,y)$  darstellen. Die Aussage "die Durchfallquote  $\frac{y}{x}$  ist grösser als die Durchfallquote  $\frac{v}{u}$ " stellt sich dann so dar, dass die erste Strecke steiler verläuft als die zweite Strecke.

In der folgenden Abbildung sind die Daten aus der Fahrschul-Prüfung graphisch dargestellt.

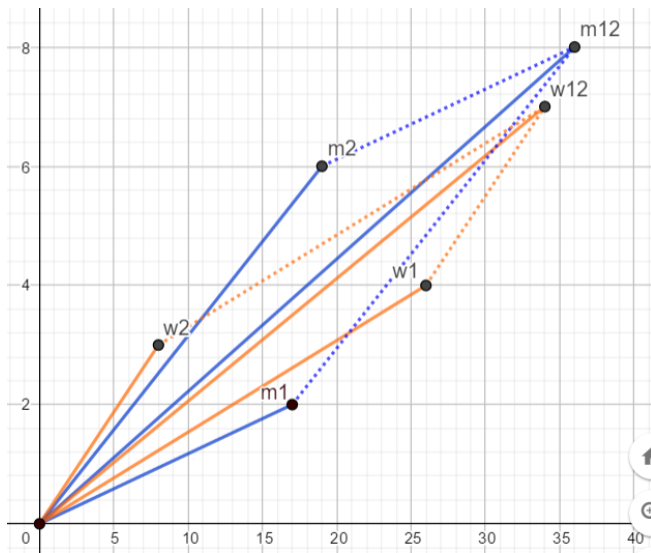


Abbildung 2.9: Graphische Interpretation der Fahrschul-Prüfungs-Daten

Die blauen Strecken gehören zu den Daten der Männer; die roten zu denen der Frauen. Bei Tag 1 ( $m1, w1$ ) und Tag 2 ( $m2, w2$ ) sind die Durchfallquoten der Männer kleiner als die der Frauen (die Steigung der blauen Linien ist kleiner als die der zugehörigen roten Linien). Bei den aggregierten Daten  $m12$  und  $w12$  ist es umgekehrt. Die gestrichelten Linien dienen der Verdeutlichung der Vektoraddition ("Parallelogramm der Kräfte").

*Aufgabe:* Erstellen Sie mit Hilfe geeigneter Software zur "Dynamischen Geometrie" ein Arbeitsblatt, um das Simpson-Paradoxon ähnlich wie in Abbildung 2.9 interaktiv nachzuspielen und Zusammenhänge zu erkunden.

### 3 Anwendung auf das Disko-Problem

#### 3.1 Formalisierung und Lösung des Problems

Beim Disko-Problem betrachten wir die folgenden Ereignisse:

$R1$ : Jens ist in Raum 1.

$R2$ : Jens ist in Raum 2.

$R3$ : Jens ist in Raum 3.

$R4$ : Jens ist in Raum 4.

$N$ : Jens ist nicht in der Disko.

Ist  $p$  die Wahrscheinlichkeit dafür, dass Jens in der Disko ist, so ist  $P(N) = 1 - p$  und wegen der vorausgesetzten Gleichwahrscheinlichkeit der Disko-Räume ist

$$P(R1) = P(R2) = P(R3) = P(R4) = \frac{p}{4}$$

Das von Mona erwünschte Ereignis lässt sich beschreiben als  $R4$  unter der Bedingung "(nicht  $R1$ ) und (nicht  $R2$ ) und (nicht  $R3$ )". Bezeichnet man mit  $\overline{X}$  das Komplement des Ereignisses  $X$ , und mit  $B$  das Ereignis

$$B := \overline{R1} \cap \overline{R2} \cap \overline{R3}$$

so ist also die bedingte Wahrscheinlichkeit

$$P(R4 | B)$$

gesucht. Nach den De Morgan'schen Gesetzen ist

$$\overline{R1} \cap \overline{R2} \cap \overline{R3} = \overline{R1 \cup R2 \cup R3}$$

und somit gilt (mit  $B = \overline{R1} \cap \overline{R2} \cap \overline{R3} = \overline{R1 \cup R2 \cup R3}$  )

$$\overline{B} = \overline{\overline{R1 \cup R2 \cup R3}} = R1 \cup R2 \cup R3$$

Da letzteres eine disjunkte Vereinigung ist, gilt:

$$P(R1 \cup R2 \cup R3) = P(R1) + P(R2) + P(R3) = 3 \cdot \frac{p}{4}$$

Für die gesuchte bedingte Wahrscheinlichkeit gilt

$$P(R4|B) = \frac{P(R4 \cap B)}{P(B)}$$

Da sich  $R4$  und  $R1$  gegenseitig ausschließen, ist  $R4$  in  $\overline{R1}$  und ebenso in  $\overline{R2}$  und  $\overline{R3}$ , also in  $\overline{R1} \cap \overline{R2} \cap \overline{R3}$  ( $= B$ ) enthalten. Das heißt,  $R4 \cap B = R4$  und  $P(R4 \cap B) = P(R4) = \frac{p}{4}$ . Weiterhin ist  $P(B) = 1 - P(\overline{B}) = 1 - P(R1 \cup R2 \cup R3) = 1 - 3 \cdot \frac{p}{4}$ . Somit ist schließlich

$$P(R4|B) = \frac{\frac{p}{4}}{1 - 3 \cdot \frac{p}{4}} = \frac{p}{4 - 3p} \quad (3.1)$$

### 3.2 Darstellung im Mengendiagramm

Es ist eine reizvolle Aufgabe, die Überlegungen in Abschnitt 3.1 durch anschauliche Darstellungen zu ergänzen. Dies soll im Folgenden mit Hilfe von Mengendiagrammen und Baumdigrammen geschehen, die an diese spezielle Situation angepasst sind.

Im folgenden Diagramm ist der gesamte Wahrscheinlichkeitsraum  $\Omega$  aus darstellungstechnischen Gründen als Quadrat dargestellt. Der neue Wahrscheinlichkeitsraum  $\Omega_{neu} := B = \Omega \setminus (R1 \cup R2 \cup R3)$  ist schraffiert (Schraffur von links oben nach rechts unten) und das Ereignis  $A = R4$  ist durch eine Schraffur (von links unten nach rechts oben) dargestellt. Da  $A$  ganz in  $B$  enthalten ist, erscheint es in der Abbildung doppelt schraffiert.  $\overline{B} = R1 \cup R2 \cup R3$  (weiß dargestellt) gehört nicht mehr zum neuen Wahrscheinlichkeitsraum  $\Omega_{neu}$ .

*Aufgabe:* Es bietet sich an, die Graphik zu dynamisieren. Setzen Sie dies mit einem geeigneten Programm zur "Dynamischen Geometrie"<sup>11</sup> um. (Dies ist auch einer der Gründe dafür, dass die einschlägigen Mengen in der Abbildung 3.1 quadratisch bzw. rechteckig dargestellt sind.)

Im folgenden Beispiel ist  $p = 0.6$ .

<sup>11</sup>Dies sind Programme wie z.B. Cabri Géomètre, Euklid - DynaGeo, Cinderella, GeoGebra u.v.m.



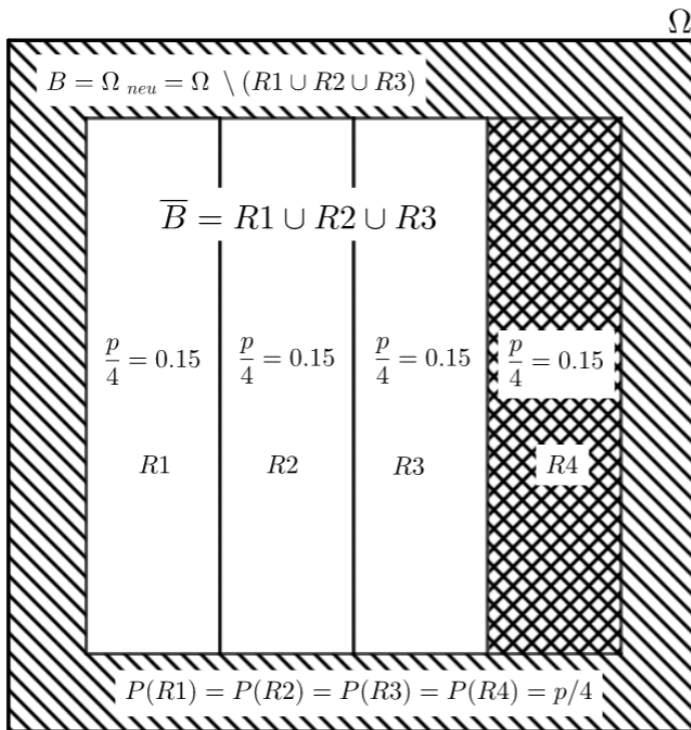


Abbildung 3.1: Venn-Diagramm zum Disko-Problem

Deutet man im obigen Diagramm die jeweiligen Wahrscheinlichkeiten als die entsprechenden Flächenanteile, so bestätigt die Dynamisierung die folgende (korrekte) intuitive Einschätzung: Ist  $p = 1$ , so ist  $P(R4) = \frac{1}{4}$  (denn  $R4 \cap B = R4$  und  $P(R4|B) = \frac{P(R4 \cap B)}{P(B)} = \frac{1/4}{1-3/4} = 1$ ).

*Aufgabe:* Erstellen Sie für  $p = 0,1 / 0,2 / 0,4 / 0,8 / 0,9$  zu Abbildung 3.1 analoge Graphiken, wenn möglich mit Hilfe eines geeigneten Programms zur Dynamischen Geometrie.

### 3.3 Darstellung im Baumdiagramm

Im Folgenden werden wir in Analogie zu den Abbildungen (2.2), (2.3), (2.4) und (2.5) neben  $A \cap B$  auch die Mengen  $\overline{A} \cap B$ ,  $A \cap \overline{B}$  und  $\overline{A} \cap \overline{B}$  genauer betrachten müssen.

*Aufgabe:* Zeichnen Sie im Zusammenhang mit dem Disko-Problem in Analogie zu Abbildung (3.1) die Venn-Diagramme zu den Mengen  $\overline{A} \cap B$ ,  $A \cap \overline{B}$  und  $\overline{A} \cap \overline{B}$ .

Wir müssen für die Darstellung des Disko-Problems als Baumdiagramm nun nur noch das allgemeine Baumdiagramm (2.7) an die Vorgaben des Disko-Problems anpassen. In diesem Fall ist

$$A = R4, \quad a = P(A) = \frac{p}{4}$$

$$B = \overline{R1} \cap \overline{R2} \cap \overline{R3} = \overline{R1 \cup R2 \cup R3}$$

$$b = P(B) = 1 - 3 \cdot \frac{p}{4}$$

$$\text{wegen } \overline{B} = R1 \cup R2 \cup R3 \quad \text{und} \quad P(\overline{B}) = 3 \cdot \frac{p}{4}$$

(Vereinigung disjunkter Mengen)

$$A \cap B = A = R4, \quad c = P(A \cap B) = P(R4) = \frac{p}{4}$$

$$\overline{A} \cap B = \overline{R4} \cap \overline{R1} \cap \overline{R2} \cap \overline{R3} = \overline{R4 \cup R1 \cup R2 \cup R3} \quad (\text{nach De Morgan})$$

$$P(\overline{A} \cap B) = 1 - p$$

$$A \cap \overline{B} = \emptyset, \quad P(A \cap \overline{B}) = 0$$

$$\overline{A} \cap \overline{B} = \overline{B}, \quad (\text{denn } \overline{B} \subseteq \overline{A})$$

$$P(\overline{A} \cap \overline{B}) = 3 \cdot \frac{p}{4}$$

*Plausibilitätsbetrachtungen* (“Proben“): Mit den Bezeichnung  $a, b$  und  $c$  aus Abbildung 2.7 gilt

$$1. \ P(\overline{A} \cap B) = b - c = (1 - 3 \cdot \frac{p}{4}) - \frac{p}{4} = 1 - p$$

$$2. \ P(A \cap \overline{B}) = a - c = \frac{p}{4} - \frac{p}{4} = 0$$

$$3. \ P(\overline{A} \cap \overline{B}) = 1 - (a + b) + c = 1 - (\frac{p}{4} + (1 - 3 \cdot \frac{p}{4})) + \frac{p}{4} = 3 \cdot \frac{p}{4}$$

$$4. \ P(A \cap B) + P(\overline{A} \cap B) + P(A \cap \overline{B}) + P(\overline{A} \cap \overline{B}) = \frac{p}{4} + (1 - p) + 0 + 3 \cdot \frac{p}{4} = 1$$

Mit diesen Werten sieht das allgemeine Baumdiagramm 2.7 nun folgendermaßen aus:

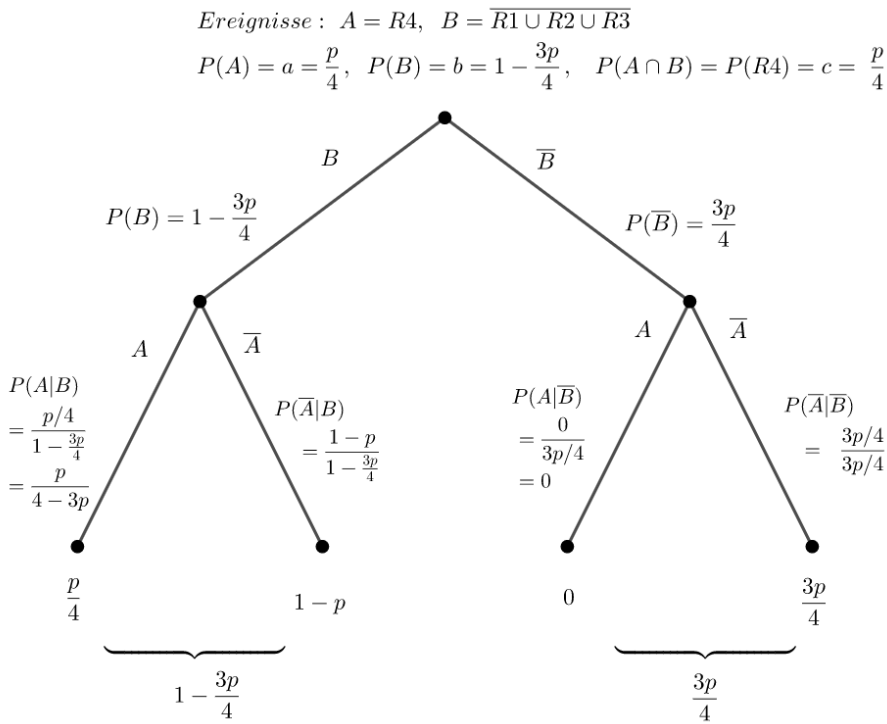


Abbildung 3.2: Baumdiagramm zum Disko-Problem

### 3.4 Darstellung mit der Vierfeldertafel

Wir behalten die Bezeichnungen aus den beiden vorhergehenden Abschnitten bei. Mit den Werten  $P(A) = \frac{p}{4}$ ,  $P(B) = 1 - 3 \cdot \frac{p}{4}$ ,  $P(A \cap B) = \frac{p}{4}$ ,  $P(A \cap \overline{B}) = 0$ ,  $P(\overline{A} \cap B) = 1 - p$  und  $P(\overline{A} \cap \overline{B}) = 3 \cdot \frac{p}{4}$  erhalten wir die folgende Vierfeldertafel

Vierfeldertafel	$B$	$\overline{B}$	Zeilensummen
$A$	$\frac{p}{4}$	0	$\frac{p}{4}$
$\overline{A}$	$1 - p$	$3 \cdot \frac{p}{4}$	$1 - \frac{p}{4}$
Spaltensummen	$1 - 3 \cdot \frac{p}{4}$	$3 \cdot \frac{p}{4}$	1

Für die gesuchte bedingte Wahrscheinlichkeit gilt (vgl. Gleichung (3.1)):

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{p}{4}}{1 - 3 \cdot \frac{p}{4}} = \frac{p}{4 - 3p}$$

## 4 Simulation

Die Rolle der Simulation in Problemlösungs- und Erkenntnisprozessen ist in der einschlägigen Literatur ausführlich dargestellt. An dieser Stelle sei exemplarisch auf das Buch *Algorithmen – von Hammurapi bis Gödel* (2016), besonders Abschnitt 4.5, beschrieben, auf das hier grundsätzlich verwiesen wird.

Eine besondere Rolle spielt die Simulation *stochastischer* (also zufallsbedingter) Prozesse. Gerade diese Prozesse laufen oft der Intuition zuwider, selbst bei erfahrenen Mathematikern. Eine der bemerkenswertesten Forscherpersönlichkeiten des 20. Jahrhunderts war der aussergewöhnlich produktive und sehr kooperationsfreudige Mathematiker Pál Erdős<sup>12</sup>. Aber sogar Erdős hat sich schwer damit getan, die korrekte Lösung des wohlbekannten *Ziegenproblems*<sup>13</sup> zu akzeptieren und konnte erst durch eine Simulation dazu gebracht werden, seine Position zu revidieren (vgl. von Randow 1992, Vazsonyi 1999, oder Ziegenbalg 2016 / Seite 150).

Deshalb sollen die obigen theoretischen Überlegungen noch durch eine Computersimulation ergänzt werden. Dies geschieht im Anhang zu diesem Text mit Hilfe des (open source) Computeralgebra Systems *Maxima*. Das dargestellte Simulationsprogramm arbeitet weitgehend mit bekannten und üblichen Konstrukten der Programmierung; eine gewisse Vertrautheit mit dem von der Programmiersprache LISP herrührenden *Listenkonzept* ist jedoch hilfreich für das Verständnis.

*Bemerkung:* Die durch Simulation gewonnenen Ergebnisse sind überzeugend. Dennoch muss man Simulationen (wie auch allen anderen Begründungsversuchen) immer wieder kritisch gegenüberstehen. Es ist insbesondere zu prüfen, ob die Simulation problemgerecht durchgeführt wurde; mit anderen Worten, ob sie genau das gestellte Problem (und nicht etwa ein mehr oder weniger ähnliches Problem) beschreibt.

Minimale Voraussetzung für die Überprüfbarkeit von Simulationen sind, dass das jeweilige Simulationsprogramm selbst offengelegt wird und dass als Hilfs-Software nur solche aus der "open source Welt" verwendet wird.

---

<sup>12</sup> Pál Erdős, ungarischer Mathematiker, 1913–1996

<sup>13</sup> im englischsprachigen Raum als *Monty Hall Problem* bekannt

## 5 Literaturhinweise

Behrends E.: Elementare Stochastik, Springer Spektrum, Wiesbaden 2013

Dürr R. / Ziegenbalg J.: Mathematik für Computeranwendungen: Dynamische Prozesse und ihre Mathematisierung durch Differenzengleichungen; Ferdinand Schöningh Verlag, Paderborn 1984  
2. Auflage: Mathematik für Computeranwendungen; Ferdinand Schöningh Verlag, Paderborn 1989

Henze N.: Stochastik für Einsteiger (10. Auflage), Springer Spektrum, Wiesbaden 2013

Vazsonyi A.: Which Door Has the Cadillac?, atlanta Verlag, Decision Line, December / January 1999, 17–19  
Link zum Artikel (pdf-Datei):  
[http://en.wikipedia.org/wiki/Monty\\_Hall\\_problem#refVazsonyi1999/](http://en.wikipedia.org/wiki/Monty_Hall_problem#refVazsonyi1999/)

von Randow G.: Das Ziegenproblem – Denken in Wahrscheinlichkeiten; rororo science, Hamburg 1992

Ziegenbalg J. / Ziegenbalg O. / Ziegenbalg B.: Algorithmen – von Hammurapi bis Gödel (4., überarbeitete und erweiterte Auflage); Springer Spektrum, Wiesbaden 2016

## 6 Anhang: Simulationsprogramm

Disko-Simulation-Maxima.wxmx

1 / 4

# Disko-Problem: Simulation mit Maxima

Einige Hinweise:

Das folgende Maxima-Dokument enthält Maxima-Code-Zellen, Text-Zellen und Graphik-Zellen.

Maxima-Code ist in einer Festpunktschrift, reiner Text ist in einer Proportionalchrift (meist in "Times Roman") dargestellt.

Kommentare sind syntaktisch durch das "Klammerpaar" `/* Kommentar-Beispiel */` gekennzeichnet.

`/* Kommentare können sich auch über  
mehrere Zeilen erstrecken. */`

Das `$` - Zeichen am Ende einer Eingabezeile verhindert, dass die Ausgabe ausgedruckt wird.

Die folgenden beiden Befehle dienen dazu, um verschiedene Darstellungsarten einzurichten. Sie haben nichts mit der eigentlichen Simulation zu tun.

```
(%i26) set_display('ascii)$ /* macht Zeilenumbruch am Seitenrand möglich */
(%i2) set_display('xml)$ /* ermöglicht Graphik-Darstellung im Dokument */
```

Die folgenden Kommandos `make_random_state` und `set_random_state` dienen der Initiierung des Zufallszahlen-Generators von Maxima

```
(%i5) randomstate1 : make_random_state (654321) $
      set_random_state (randomstate1) $
      make_random_state(true) $

(%i6) random(1000);
(%o6) 768
```

Die folgende Funktion `RaumJ` ermittelt, in welchem Raum sich Jens aufhält. Der Eingabeparameter `p` gibt an, mit welcher Wahrscheinlichkeit Jens in der Disko ist.

Die Werte der Variablen `r1`, `r2`, `r3`, `r4`, `rn` sind jeweils gleich 0 oder 1, je nachdem, ob Jens sich in dem Raum befindet oder nicht (`r1=1`: Jens ist in Raum 1, ..., `rn=1`: Jens ist nicht in der Disko).

Variable `rog`: Obergrenze für die `random`-Funktion

Die Ausgabe erfolgt als Liste: `[r1, r2, r3, r4, rn]`, wobei genau ein Wert gleich 1 ist und alle anderen Werte gleich Null sind.

```
(%i7) RaumJ(p) :=
  block([r1:0, r2:0, r3:0, r4:0, rn:0, rog : 1000000],
    r : random(rog),
    if r < (1-p)*rog then rn : 1          /* J nicht in Disko */
    elseif r < (1-p)*rog + rog*p/4 then r1 : 1      /* J in Raum 1 */
    elseif r < (1-p)*rog + 2*rog*p/4 then r2 : 1    /* J in Raum 2 */
    elseif r < (1-p)*rog + 3*rog*p/4 then r3 : 1    /* J in Raum 3 */
    else r4 : 1,                                  /* J in Raum 4 */
    [r1, r2, r3, r4, rn] ) $
```

Einige Test-Aufrufe:

```
(%i8) RaumJ(0.75);
(%o8) [0,0,0,0,1]

(%i9) RaumJ(1);
(%o9) [1,0,0,0,0]

(%i10) RaumJ(0);
(%o10) [0,0,0,0,1]

(%i11) RaumJ(0.5);
(%o11) [0,0,0,0,1]
```

Die folgende Funktion SucheM simuliert einen einzigen Such-Vorgang. Zunächst wird durch den Aufruf von RaumJ(p) der Raum (mit Hilfe des Zufallszahlengenerators) festgelegt, in dem sich Jens befindet.

Variable n123: Jens befindet sich nicht in Raum 1 und nicht in Raum 2 und nicht in Raum 3.

Variable n123r4: Jens befindet sich nicht in Raum 1 und nicht in Raum 2 und nicht in Raum 3, sondern in Raum 4.

Ausgabe: Die Zweier-Liste [n123r4, n123] für die weitere Verarbeitung (Aggregation der Daten).

```
(%i12) SucheM(p) :=
  block( [RJ : [0, 0 ,0 ,0, 0], n123:0, n123r4:0],
    RJ : RaumJ(p),
    if (RJ[1]=0 and RJ[2]=0 and RJ[3]=0)
      then n123 : n123+1,
    if (RJ[1]=0 and RJ[2]=0 and RJ[3]=0 and RJ[4]=1)
      then n123r4 : n123r4+1,
    [n123r4, n123] ) $

(%i13) SucheM(0.6);
(%o13) [0,0]
```

Der Einzel-Suchvorgang SucheM(p) wird im Folgenden L-mal hintereinander ausgeführt.

Die Ergebnisse von SucheM(p) werden in der Liste SumL aufaddiert.

SumL: 2-elementige Liste von der Struktur der Ausgabe von SucheM(p).

Ausgabe: Anteil von "Raum4" in der Menge

"(nicht Raum 1) und (nicht Raum 2) und (nicht Raum 3) und (Raum 4)"

Disko-Simulation-Maxima.wxmx

3 / 4

```
(%i14) SimulationSuche(p, L) :=
  block( [SumL : [0,0]] ,
    for i:1 thru L do
      SumL : SumL + SucheM(p),
    if notequal(SumL[2],0) then return(float(SumL[1]/SumL[2]))
    else return("Bitte Simulation mit mehr Läufen wiederholen") ) $

(%i15) SimulationSuche(0.7,1000);
(%o15) 0.3622559652928417
```

Im Folgenden werden je 1000 Testläufe für die Wahrscheinlichkeiten von 0 bis 1 mit Schrittweite 0.1 durchgeführt.

```
(%i16) LT : makelist([0.1*i, SimulationSuche(0.1*i, 1000)], i, 0, 10, 1) $

(%i27) LT;
(%o27) [[0, 0.0], [0.1, 0.02939587160920782], [0.2, 0.05722359590250795],
[0.3, 0.09514013885317563], [0.4, 0.1370909615107395],
[0.5, 0.1921404412948945], [0.6000000000000001, 0.2641441441441442],
[0.7000000000000001, 0.3603622577927548], [0.8, 0.4960449094156673],
[0.9, 0.6942325297528227], [1.0, 1.0]]

(%i18) TabellenAusdruck(L) :=
  (print("          p          n123r4 / n123"),
  for K in L do
    printf(true, "~10,2h ~12,4h ~%",first(K), last(K) ) ) $
```

In der folgenden Tabelle sind in Abhängigkeit von der Wahrscheinlichkeit  $p$  die Simulationsergebnisse (d.h. die bedingten Wahrscheinlichkeiten  $P(R4 \mid \text{nicht}(R1) \text{ und nicht}(R2) \text{ und nicht}(R3))$ ) dargestellt.

```
(%i19) TabellenAusdruck(LT) $

      p      n123r4/n123
0.00      0.0000
0.10      0.0236
0.20      0.0590
0.30      0.1139
0.40      0.1312
0.50      0.1864
0.60      0.2956
0.70      0.3664
0.80      0.4600
0.90      0.7021
1.00      1.0000
```

Die (diskreten) Werte der obigen Tabelle sollen im Folgenden mit den Werten der (stetigen) Funktion  $f(p)$  verglichen werden. Zu diesem Zweck wird beides in einem gemeinsamen Schaubild eingetragen.

```
(%i20) f(p) := p/(4-3*p) $ /* Definition der Funktion f */
```



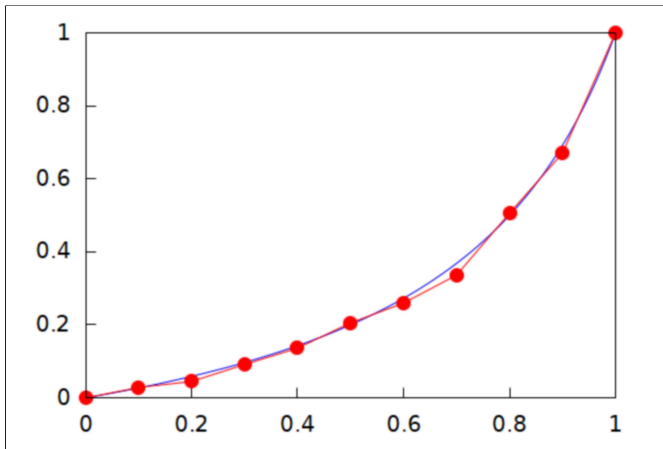
```
(%i21) load(draw);
```

```
(%o21) C:/maxima-5.42.2/share/maxima/5.42.2/share/draw/draw.lisp
```

Im Folgenden: Graphik-Darstellung einer Simulation mit jeweils 1.000 Läufen

```
(%i23) LT : makelist([0.1*i, SimulationSuche(0.1*i, 1000)], i, 0, 10, 1) $ ;
DrawGraphik()
(G1 : [color=blue, explicit(f(p), p,0,1)],
 G2 : [color=red, point_type=filled_circle, point_size=2,
 points_joined=true, points(LT)],
 wxdraw2d(G1, G2) ) $
```

```
(%t23)
```



Im Folgenden: Graphik-Darstellung einer Simulation mit jeweils 10.000 Läufen

```
(%i25) LT : makelist([0.1*i, SimulationSuche(0.1*i, 10000)], i, 0, 10, 1) $ ;
DrawGraphik()
(G1 : [color=blue, explicit(f(p), p,0,1)],
 G2 : [color=red, point_type=filled_circle, point_size=2,
 points_joined=true, points(LT)],
 wxdraw2d(G1, G2)) $ ;
```

```
(%t25)
```

