978-3-7910-3394-5 Schuldenzucker, Prüfungstraining Induktive Statistik

SCHÄFFER
POESCHEL

Einführung

Methoden der induktiven Statistik sind häufig eine wesentliche Voraussetzung, um Entscheidungen fundiert treffen zu können: Nicht aufgrund eines Bauchgefühls, sondern aufgrund von Wissen über bestimmte Werte. Da man relativ häufig über interessierende Größen ein zu geringes Wissen hat, versucht man dieses Defizit aufgrund erhobener Daten zu beheben. Induktive Statistik hilft, hinreichend vertrauenswürdige Daten zu erheben und ihre Aussagekraft zu beurteilen.

Induktive Statistik befasst sich mit der Aufgabe, prinzipielle Aussagen über Grundgesamtheiten von Größen zu treffen: So kann etwa die Wahrscheinlichkeit, mit der ein Ereignis eintrifft, berechnet oder geschätzt werden; man möchte für jeden möglichen Wert oder jeden Zahlenbereich, in dem die Größe Werte annehmen kann, ermitteln können, mit welcher Wahrscheinlichkeit diese Größe dort bestimmte Werte annimmt.

Solche Wahrscheinlichkeiten können auf zwei Arten bestimmt werden: Entweder hat man verlässliche Vorkenntnisse, aus denen Wahrscheinlichkeiten exakt berechnet werden können, oder man muss Daten erheben, um die gesuchten Wahrscheinlichkeiten abschätzen zu können. In diesem zweiten Fall wird man Aussagen aufgrund der Datenlage mit einer gewissen Sicherheit treffen können, muss aber eine Restunsicherheit hinnehmen.

Sind Wahrscheinlichkeiten über eine interessierende Größe ermittelt, so können sie zur Berechnung oder Eingrenzung weiterer Maßzahlen benutzt werden, die im Kontext mit dieser Grundgesamtheit zu erwarten sind. Zum Beispiel kann die Frage, welches Ergebnis bei häufigem Wiederholen desselben Versuchs im Mittel zu erwarten ist, beantwortet oder näherungsweise geklärt werden. Der Anteil, zu dem man bei der Durchführung vieler identischer Versuche vermutlich Erfolg haben wird, kann ermittelt werden. Die Frage, ob zwei Größen sich prinzipiell miteinander oder gegeneinander entwickeln – oder ob kein solcher Zusammenhang feststellbar ist – kann mit einer vorgegebenen Sicherheit beantwortet werden.

Beispiele für die Anwendung von Methoden der induktiven Statistik finden sich in allen Alltagsbereichen: Wenn aufgrund erhobener Daten Schlussfolgerungen über eine größere Grundgesamtheit gezogen werden, ist induktive Statistik beteiligt. Dies können etwa Fragestellungen aus der Psychologie zu Vorlieben oder Verhaltensweisen sein, zu erwartende Kosten, Gewinne oder Kursentwicklungen, zu erwartende Marktanteile oder auch die Häufigkeit, mit der eine Straßenbahn zu spät kommt und vieles mehr.

1 Wahrscheinlichkeitsrechnung

Die Wahrscheinlichkeit, mit der ein Ereignis A eintritt, spiegelt den erwarteten Anteil der Ergebnisse eines häufig wiederholten Zufallsexperiments wider, bei denen A eintritt. Damit ist die Wahrscheinlichkeit eines Ereignisses ein Maß dafür, dass dieses Ereignis beim Durchführen eines Versuchs eintritt. Man erwartet, dass Ereignisse mit hoher Wahrscheinlichkeit relativ häufig eintreten, Ereignisse mit niedriger Wahrscheinlichkeit eher selten.

Im einfachsten Fall ist nur eine begrenzte Anzahl von Ergebnissen des Zufallsexperiments möglich, die alle gleich wahrscheinlich sind. Dann spricht man von einer Laplace-Wahrscheinlichkeit und die Wahrscheinlichkeit eines Ereignisses ist gleich dem Anteil der Ergebnisse, die zu diesem Ereignis gehören, an der gesamten Ergebnismenge (Formel 1.2 in der Formelsammlung).
Für das prinzipielle Errechnen von Wahrscheinlichkeiten steht eine Anzahl von Regeln zur Verfügung, die sich aus logischen Überlegungen ergeben (Formeln 1.2 bis 1.6).

1.1 Zufallsexperiment und Ereignis

Unter einem *Zufallsexperiment* versteht man ein Experiment mit mehreren möglichen Ergebnissen, bei dem vor Durchführung nicht feststeht, welches Ergebnis eintreten wird.
Ein Ereignis zu einem Zufallsexperiment kann als eine Teilmenge der Ergebnismenge aufgefasst werden. Dadurch ist es möglich, das Ermitteln von Wahrscheinlichkeiten auf einfache Mengenlehre zurückzuführen.

Beispiele:
von Zufallsexperimenten:

(1) Werfen einer Münze
(2) Werfen eines Würfels
(3) Werfen von zwei Würfeln
(4) Lotto spielen
(5) medizinische Messungen

Definitionen:

Die *Ergebnismenge* eines Zufallsexperiments ist die Menge aller möglichen Ergebnisse.

Sie wird üblicherweise mit dem großen grichischen Omega Ω bezeichnet.

Im Fall endlich oder abzählbar vieler möglicher Ergebnisse kann Ω durch Aufzählen angegeben werden in der Form

$\Omega = \{a, b, \dots\}$

Häufig werden die einzelnen möglichen Ergebnisse, also die Elemente von Ω, mit dem kleinen griechischen Omega bezeichnet: $\omega \in \Omega$.

Im Folgenden werden wir als häufige Beispiele das Werfen eines oder zweier Würfel nutzen. Damit sie einfach wiedererkannt werden können, wird das Werfen eines Würfels mit **Beispiel (1)**, das zweier Würfel mit **Beispiel (2)** bezeichnet.

Beispiel (1):
Ein Würfel: $\Omega = \{1, 2, 3, 4, 5, 6\}$
$\omega = 3 \in \Omega$

Beispiel (2):
Zwei Würfel:
$\Omega = \{(1,1), (1,2), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,1), \dots, (6,6)\}$
Der Sechserpasch $\omega = (6,6)$ ist ein mögliches Ergebnis: $(6,6) \in \Omega$

Ein *Zufallsereignis* kann als Teilmenge der Ergebnismenge aufgefasst werden.
Zufallsereignisse werden häufig mit lateinischen Großbuchstaben bezeichnet:
$A \subset \Omega$ ist ein Ereignis.

Beispiel (1):
Ein Würfel: $A = \{6\}$ ist das Ereignis, dass eine 6 geworfen wird.
$B = \{2, 4, 6\}$ ist das Ereignis, dass eine gerade Zahl geworfen wird.

Beispiel (2):
Zwei Würfel: $A = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$ ist das Ereignis, dass zwei gleiche Zahlen geworfen werden.

Unter einem *Elementarereignis* versteht man ein Ereignis, das nur aus einem Element besteht.

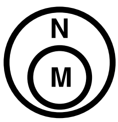
Beispiel (1):
Ein Würfel: Die Elementarereignisse sind die Mengen $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$

Beispiel (2):
Zwei Würfel: Jedes mögliche Wertepaar definiert ein Elementarereignis.
Z. B. ist $\{(1,1)\}$ eins der insgesamt 36 Elementarereignisse.

Beziehungen zwischen Mengen:

Seien M und N zwei Mengen.

(a) $M \subset N$ bedeutet: M ist *Teilmenge* von N.



Das heißt, dass alle Elemente von M auch in N liegen.

Beispiel:

$M = \{\text{Hunde}\}$, $N = \{\text{Tiere}\}$

Jeder Hund ist ein Tier: $M \subset N$

Gleichbedeutend ist:

$N \supset M$ *N umfasst M*

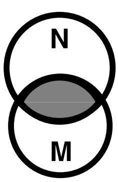
(b) $M = N$ bedeutet: M umfasst N und N umfasst M.

Beispiel:

$M = \{\text{Segelfulgzeuge}\}$

$N = \{\text{gliders}\}$

(c) $M \cap N$ $= \{x \mid x \in M \text{ und } x \in N\}$



$=$ $\boxed{\textit{Durchschnitt}}$ von M und N,

$=$ Schnittmenge von M und N

Beispiel:

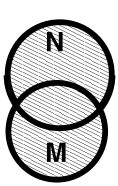
$N = \{\text{Brötchen mit Butter}\}$

$M = \{\text{Brötchen mit Käse}\}$

$M \cap N = \{\text{Brötchen mit Butter } \boxed{\textit{und}} \text{ Käse}\}$

Der Durchschnitt zweier Mengen M und N besteht aus den Elementen, die in M *und* in N liegen.

(d) $M \cup N$ $= \{x \mid x \in M \text{ oder } x \in N\}$



$=$ $\boxed{\textit{Vereinigung}}$ von M und N

Beispiel:

$N = \{\text{Brötchen mit Butter}\}$

$M = \{\text{Brötchen mit Käse}\}$

$M \cup N =$

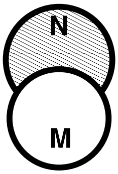
$\{\text{Brötchen mit Butter } \boxed{\textit{oder}} \text{ Käse}\}$

Die Vereinigung zweier Mengen M und N besteht aus den Elementen, die in M *oder* in N (oder in beiden) liegen.

(e) $N \backslash M$ $= \{x \in N \mid x \notin M\}$

$=$ Differenzmenge von N und M

$= N$ außer M



Beispiel:

$N = \{\text{Brötchen mit Butter}\}$

$M = \{\text{Brötchen mit Käse}\}$

$N \backslash M = \{\text{Brötchen mit Butter ohne Käse}\}$

(f) $\overline{M}$ in N ist das *Komplement* von M in N.

Es besteht aus allen Elementen von N, die nicht in M liegen.

Beispiel:

$M = \{\text{Brötchen mit Käse}\}$

In N: $\overline{M} = N \backslash M$

(g) $\emptyset$ ist das Symbol für die *leere Menge*:

Sie enthält kein Element.

Beispiel:

$N = \{\text{Studierende in Köln}\}$

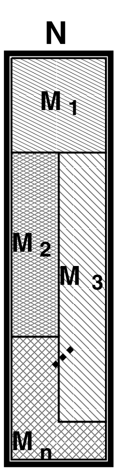
$M = \{\text{Menschen}\}$

$N \backslash M = \emptyset$

(h) $|N|$ ist die *Kardinalzahl* oder *Mächtigkeit* von N.

Dies ist die Anzahl der Elemente von N.

(i) Es sei $N =$ $M_1 \cup \cdots \cup M_n = \bigcup_{i=1}^n M_i$,
wobei $M_i \cap M_j = \emptyset$ für alle $i \neq j \in \{1, \ldots, n\}$



Eine *vollständige Disjunktion der Menge N* ist eine *Zerlegung* von N in Teilmengen, die einander *nicht schneiden*:

$N = M_1 \cup \cdots \cup M_n = \bigcup_{i=1}^n M_i$, wobei $M_i \cap M_j = \emptyset$ für alle $i \neq j \in \{1, \ldots, n\}$

Beispiel:

$$\begin{aligned} \{1, 2, \ldots, 10\} &= \{1, 2, 3, 4, 5\} \cup \{6, 7, 8, 9, 10\} \\ &= \{1, 2\} \cup \{3, 4, 5\} \cup \\ &\quad \{6, 7\} \cup \{8, 9, 10\} \end{aligned}$$

sind zwei vollständige Disjunktionen der Menge $\{1, 2, \ldots, 10\}$.

1.2 Wahrscheinlichkeit

1.2.1 Einführung

Die Wahrscheinlichkeit eines Ereignisses ist der bei häufiger Wiederholung eines Zufallsexperiments erwartete Anteil der Ergebnisse, bei denen das Ereignis eintritt. Damit ist die Wahrscheinlichkeit die erwartete relative Häufigkeit bei vielen Versuchen.

Wahrscheinlichkeitsbegriff nach *Laplace*:
Die Ergebnismenge Ω besitze m Elemente.
Alle Elementarereignisse mögen dieselbe Wahrscheinlichkeit p haben. Dann ist die Wahrscheinlichkeit jedes Elementarereignisses $p = \frac{1}{m}$.

Beispiel (1):
Ein Würfel: $p = \frac{1}{6}$

$p(1) + \cdots + p(6) = 6 \cdot \frac{1}{6} = 1$

Die Wahrscheinlichkeit eines beliebigen Ereignisses $A \subset \Omega$ ist damit

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der möglichen Ergebnisse in } A}{\text{Anzahl aller möglichen Ergebnisse}}$$

Beispiel (1):
Ein Würfel: $A = \{2, 4, 6\}$, $\quad P(A) = P(\text{gerade Zahl}) = \frac{3}{6}$

1.2.2 Definition von Wahrscheinlichkeit

Sei Ω die Ergebnismenge eines Zufallsexperiments.

Eine *Wahrscheinlichkeit* ordnet jedem möglichen Ereignis eine Zahl zu. Sie muss dabei bestimmten Regeln genügen:

(a) Nichtnegativität:
Wahrscheinlichkeiten liegen immer zwischen 0 und 1.
$0 \leq P(A) \leq 1$ für jedes Ereignis A

(b) Normierung:
Die Wahrscheinlichkeit der ganzen Ergebnismenge liegt bei 1. $P(\Omega) = 1$

(c) Additivität:
Falls A und B unvereinbare Ereignisse sind, d. h. $A \cap B = \emptyset$, ist die Wahrscheinlichkeit, dass Ereignis A oder B stattfindet, gleich der Summe der Einzelwahrscheinlichkeiten.
$P(A \cup B) = P(A) + P(B), \textit{falls } A \cap B = \emptyset$

»σ-Additivität«:
Diese Additivität gilt sogar für abzählbar unendlich viele Ereignisse.
$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ für paarweise unvereinbare Ereignisse, also $A_i \cap A_k = \emptyset$ für alle Indizes $i \neq k$

Folgerungen aus den Regeln:

Wahrscheinlichkeit, dass kein Ergebnis erzielt wird:
$P(\emptyset) = 0$

Wahrscheinlichkeit, dass das Gegenteil von A geschieht.
$P(\overline{A}) = 1 - P(A)$

Beispiel (1):
Ein Würfel: $A =$ Menge der geraden Würfelergebnisse:
$P(\text{ungerades Würfelergebnis}) \quad = \quad 1 - P(\text{gerades Würfelergebnis})$

Wahrscheinlichkeit, dass »höchstens A« geschieht:
$P(A) \leq P(B)$ wenn $A \subset B$

Beispiel (1):
Ein Würfel: $A =$ Menge der Würfelergebnisse ≤ 2, $B =$ Menge der Würfelergebnisse ≤ 4:
$P(\text{Würfelergebnis} \leq 2) \quad \leq \quad P(\text{Würfelergebnis} \leq 4)$

Wahrscheinlichkeit, dass A oder B geschieht:
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Beispiel (1):
Ein Würfel: $A =$ Menge der geraden Würfelergebnisse, $B =$ Menge der Würfelergebnisse ≤ 4:

$$P(\text{gerade Zahl oder} \leq 4) \quad = \quad P(\text{gerade Zahl}) + P(\text{Zahl } \leq 4) - P(\text{gerade Zahl } \leq 4)$$

Wahrscheinlichkeit, dass A und nicht B stattfindet:
$P(A - B) = P(A) - P(A \cap B)$

Beispiel (1):
Ein Würfel: $A =$ Menge der geraden Würfelergebnisse, $B =$ Menge der Würfelergebnisse ≤ 4:
$P(\text{gerades Würfelergebnis, das nicht} \leq 4 \text{ ist}) =$
$P(\text{gerades Würfelergebnis}) - P(\text{gerades Würfelergebnis, das} \leq 4 \text{ ist})$

$P(A - B) \quad = \quad P(A) - P(B) \quad$ wenn $B \subset A$

Beispiel (1):
Ein Würfel: $A =$ Menge der Würfelergebnisse ≤ 5, $B =$ Menge der Würfelergebnisse ≤ 3:

$$P(\text{Zahl } \leq 5 \text{ und nicht} \leq 3) \quad = \quad P(\text{Zahl } \leq 5) - P(\text{Zahl } \leq 3)$$

Übergang zur Wahrscheinlichkeitsdichte:
Für ein stetiges Merkmal füllt die Ergebnismenge Ω ein ganzes Intervall aus.

Beispiel:
Körpergröße Erwachsener: $\Omega = [0.50 - 2.50]$ m

Dann ist die Wahrscheinlichkeit eines Elementarereignisses $= 0$:
$P(\{\omega\}) = 0$ für $\omega \in \Omega$

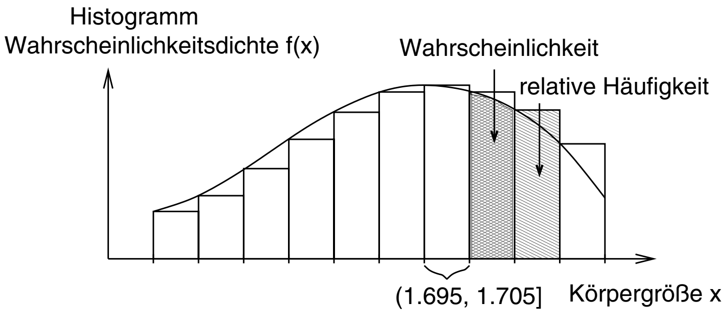
Beispiel:
Die Wahrscheinlichkeit, dass jemand genau 1.70 m groß ist, ist $= 0$.

Von Interesse sind stattdessen Wahrscheinlichkeiten von Intervallen.

Beispiel:
Die Wahrscheinlichkeit, dass eine zufällig ausgewählte Person zwischen 1.695 m und 1.705 m groß ist.

An die Stelle von Summen von Wahrscheinlichkeiten treten Integrale über eine *Wahrscheinlichkeitsdichte*-Funktion:

Das Konzept von Histogrammen aus der deskriptiven Statistik wird so stark verfeinert, dass die oberen Begrenzungslinien der Histogramme in eine stetige Kurve übergehen. Da bei Histogrammen relative Häufigkeiten Rechteckflächen sind, ergeben sich nun Wahrscheinlichkeiten als Flächen unter der Wahrscheinlichkeitsdichte.



1.3 Kontingenztafel

Sollen bei einem Zufallsexperiment zwei Merkmale gemeinsam untersucht werden, so stellt man die zugehörige zweidimensionale Wahrscheinlichkeitsverteilung in einer *Kontingenztafel* zusammen:

Mögliche Werte des Merkmals X seien $a_1, \dots, a_m$, mögliche Werte des Merkmals Y seien $b_1, \dots, b_l$.

Merkmal Y / Merkmal X	b_1	$\dots$	b_l	Summe
a_1	p_{11}	$\dots$	p_{1l}	$p_{1\bullet}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
a_m	p_{m1}	$\dots$	p_{ml}	$p_{m\bullet}$
Summe	$p_{\bullet 1}$	$\dots$	$p_{\bullet l}$	1

mit
p_{ij} = Wahrscheinlichkeit der Kombination der Werte a_i und b_j
$p_{i\bullet}$ = Wahrscheinlichkeit des Werts a_i
$p_{\bullet j}$ = Wahrscheinlichkeit des Werts b_j

Insbesondere bei der *Vierfeldertafel* hat jedes Merkmal zwei mögliche Werte.

1.4 Bedingte Wahrscheinlichkeit

Unterliegt ein Zufallsexperiment bestimmten Einschränkungen oder sind schon Teilinformationen über den Ausgang bekannt, so müssen diese in die Berechnung von Wahrscheinlichkeiten eingehen.

Beispiel:
Grippeimpfung

Der Zusammenhang einer bestimmten Grippeimpfung und der Erkrankung an dieser Grippe kann der folgenden Kontingenztafel entnommen werden:

	Erkrankung (E)	Keine Erkrankung $(\bar{E})$	Summe
Geimpft (G)	0.016	0.174	$p_{1\bullet} = 0.190$
Nicht geimpft $(\bar{G})$	0.121	0.689	$p_{2\bullet} = 0.810$
Summe	$p_{\bullet 1} = 0.137$	$p_{\bullet 2} = 0.863$	1.0

So ist etwa eine zufällig ausgewählte Person mit einer Wahrscheinlichkeit von 17.4 % geimpft und erkrankt nicht.

Wenn diese Zusammenstellung zur Beurteilung eines Impfstoffs genutzt werden soll, stellt sich etwa die Frage, wie wahrscheinlich es ist, dass jemand, der geimpft wurde, erkrankt. Diese bedingte Wahrscheinlichkeit berechnet sich aus der Größe der Teilgruppen der Geimpften und der Erkrankten unter den Geimpften:
Die Wahrscheinlichkeit, dass jemand erkrankt unter der Bedingung, dass er geimpft ist, ist

$$\begin{aligned} P(\text{Erkrankt}|\text{Geimpft}) &= \frac{\text{Anzahl aller Personen, die geimpft sind und erkranken}}{\text{Anzahl Geimpfter}} \\ &= \frac{\text{Anteil aller Personen, die geimpft sind und erkranken}}{\text{Anteil Geimpfter}} \\ &= \frac{P(E \cap G)}{P(G)} \\ &= \frac{0.016}{0.190} \quad = 0.084 \end{aligned}$$

Allgemein ergibt sich für Ereignisse A und B einer Ergebnismenge Ω:
Die *bedingte Wahrscheinlichkeit* des Ereignisses A unter der Bedingung B ist

$$P(A \mid B) \quad = \frac{P(A \cap B)}{P(B)} \qquad \text{für } P(B) > 0.$$

1.5 Gesetze zur Wahrscheinlichkeit

1.5.1 Additionsgesetze

(s. S. 7 und 8)

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) && \text{falls } A \text{ und } B \text{ unvereinbar sind} \\ P(A \cup B) &= P(A) + P(B) - P(A \cap B) && \text{allgemein} \end{aligned}$$

1.5.2 Satz von der totalen Wahrscheinlichkeit

Aus der Formel für die bedingte Wahrscheinlichkeit

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{für } P(B) > 0 \quad \text{(s. S. 11)}$$

ergibt sich zunächst der

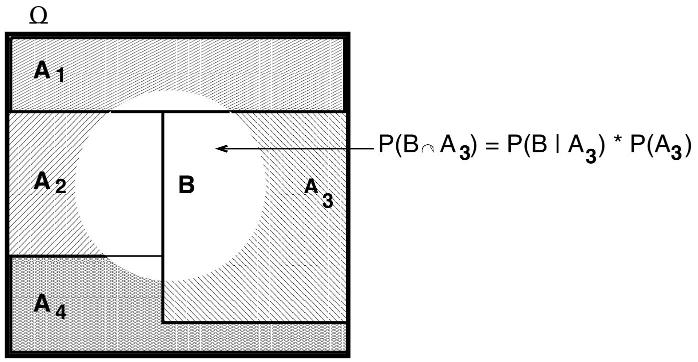
(I) **Multiplikationssatz**
$$P(A \cap B) \quad = P(A|B) \cdot P(B) \qquad \text{für } P(B) > 0$$

Beispiel (1):
$$\begin{aligned} P(\{2,4\}) &= P(\text{gerade Zahl}|\text{Zahl } \leq 4) \cdot P(\text{Zahl } \leq 4) \\ \tfrac{1}{3} &= \tfrac{1}{2} \cdot \tfrac{2}{3} \end{aligned}$$

Ist nun eine vollständige Ereignisdisjunktion gegeben, d. h. Ereignisse $A_1, \ldots, A_n$, die miteinander unvereinbar sind ($A_i \cap A_j = \emptyset$ für $i \neq j$) und die gesamte Ergebnismenge abdecken ($\Omega = A_1 \cup A_2 \cup \cdots \cup A_n$), so kann die Wahrscheinlichkeit eines Ereignisses B aus den bedingten Wahrscheinlichkeiten $P(B|A_i)$ und den Wahrscheinlichkeiten $P(A_i)$ zusammengesetzt werden:

(II) **Satz von der totalen Wahrscheinlichkeit**

$P(B) \quad = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$



Beispiel:
Es werden Mobiltelefone verlost.

Der Anteil der türkisfarbenen liegt bei 20 %, derjenige der magentafarbenen bei 30 %, 15 % sind grün und 35 % rot. Unter den türkisfarbenen haben 50 % eine Kamera, unter den magentafarbenen sind es zwei Drittel, unter den grünen nur ein Drittel und unter den roten vier Siebtel.

Mit welcher Wahrscheinlichkeit erhält ein Gewinner ein Telefon mit Kamera?

Lösung:

$P(\text{türkis}) = 0.2$

$P(\text{magentafarben}) = 0.3$

$P(\text{grün}) = 0.15$

$P(\text{rot}) = 0.35$

$P(\text{Kamera}|\text{türkis}) = 0.5$

$P(\text{Kamera}|\text{magentafarben}) = 0.\bar{6}$

$P(\text{Kamera}|\text{grün}) = 0.\bar{3}$

$P(\text{Kamera}|\text{rot}) = 0.57$

$$\begin{aligned}
P(\text{Kamera}) &= P(\text{Kamera} \cap \text{ türkis}) + P(\text{Kamera} \cap \text{magentafarben}) \\
&\quad + P(\text{Kamera} \cap \text{grün}) + P(\text{Kamera} \cap \text{rot}) \\
&= P(\text{Kamera}|\text{türkis}) \cdot P(\text{türkis}) \\
&\quad + P(\text{Kamera}|\text{magentafarben}) \cdot P(\text{magentafarben}) \\
&\quad + P(\text{Kamera}|\text{grün}) \cdot P(\text{grün}) + P(\text{Kamera}|\text{rot}) \cdot P(\text{rot}) \\
&= 0.5 \cdot 0.2 + 0.\bar{6} \cdot 0.3 + 0.\bar{3} \cdot 0.15 + 0.57 \cdot 0.35 \\
&= 0.55
\end{aligned}$$

Die Wahrscheinlichkeit, dass ein Gewinner ein Telefon mit Kamera erhält, liegt bei 55 %.

1.5.3 Formel von Bayes

Häufig sucht man eine bedingte Wahrscheinlichkeit $P(A|B)$, kennt aber statt dessen die umgekehrt bedingte Wahrscheinlichkeit $P(B|A)$. Die Formel von Bayes schafft eine Verbindung zwischen diesen beiden Wahrscheinlichkeiten:
Setzt man in die Formel für die bedingte Wahrscheinlichkeit

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

für die Wahrscheinlichkeit im Zähler, dass beide Ereignisse A und B eintreten, die Darstellung des Multiplikationssatzes mit vertauschten Rollen ein:

$$P(A \cap B) = P(B|A) \cdot P(A),$$

so ergibt sich die

Formel von Bayes

Für B mit $P(B) > 0$ gilt:

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

Steht eine vollständige Ereignisdisjunktion $A_1, \ldots, A_n$ zur Verfügung, so lässt sich die Wahrscheinlichkeit $P(B)$ im Nenner weiter zerlegen:

$$\begin{aligned}
P(A_k|B) &= \frac{P(B|A_k) \cdot P(A_k)}{P(B)} \\
&= \frac{P(B|A_k) \cdot P(A_k)}{\sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)}
\end{aligned}$$