Wahrscheinlichkeiten

Bestimmung der Wahrscheinlichkeit

Bei einem Zufallsexperiment kann man nicht voraussagen, welches Ereignis eintritt, aber manche Ereignisse treten naturgemäß mit einer größeren Wahrscheinlichkeit auf als andere bzw. kann man jedem Ereignis eine Wahrscheinlichkeit zuordnen. Nur nach welchem Prinzip? Wie kann man die Wahrscheinlichkeit quantitativ erfassen?

Wir betrachten eine Serie von n Beobachtungen, etwa den Wurf eines Würfels. Wir interessieren uns dafür, wie oft die 1 gewürfelt wurde. Sei f(1) die Anzahl der 1-er bei n Würfelwurfen. Wir bezeichnen diese Anzahl auch als *absolute Häufigkeit*.

$f(A)$ ist die absolute Häufigkeit für ein beliebiges Ereignis A

Die dazugehörige *relative Häufigkeit* ist $f(A)/n$

(bzw. $f(1)/n$ in unserem konkreten Fall).

In der Praxis kann man beobachten, dass nach n Würfen die relative Häufigkeit $f(1)/n$ etwa bei 1/6 liegt, falls die Anzahl der Würfe groß genug ist. Wenn man z.B. nur zweimal würfelt kann es sein dass man überhaupt keinen 1-er würfelt, die absolute und relative Häufigkeit vom Ereignis $\{1\}$ wäre dann 0. Für $n \to \infty$ aber stabilisiert sich die relative Häufigkeit bei 1/6.

Der Grenzert der relativen Häufigkeit wird als Wahrscheinlichkeit des Eintretens des dazugehörigen Ereignisses interpretiert.

D.h. die Wahrscheinlichkeit, dass eine "1" gewürfelt wird, ist

$$P(\text{"}1\text{"}) = \frac{1}{6} = 0.167(\text{gerundet}) \qquad (P = \textit{probability}).$$

Wenn die Wahrscheinlichkeit des Eintretens eines Ereignisses 0.167 beträgt, dann ist dies gleichbedeutend mit 16.7%. Man hat also die Wahl eine Wahrscheinlichkeit dezimal anzugeben (ein Wert zwischen 0 und 1) oder prozentuell (ein Wert zwischen 0% und 100%).

Auf diese Art und Weise wird jedem Ereignis eines Zufallsexperiments eine Wahrscheinlichkeit zugeordnet.

Eigenschaften einer Wahrscheinlichkeit

Durch folgende Eigenschaften wird eine Wahrscheinlichkeit charakterisiert:

1. $0 \leq P(A) \leq 1$ für jedes Ereignis A
2. $P(\Omega) = 1$
3. $P(\bigvee\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} P(A_i)$ für jede Folge paarweise disjunkter Ereignisse A_i.

Für Wahrscheinlichkeiten gelten unter anderem folgende Rechenregeln:

- $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$
- $P(A \vee B) = P(A) + P(B)$ für zwei disjunkte Ereignisse A und B
- $P(A) + P(\neg A) = 1$
- $P(B - A) = P(B) - P(A \wedge B)$

Für Zufallsexperimente, bei denen alle Ereignisse mit der gleichen Wahrscheinlichkeit auftreten (z.B. bei einem Würfelwurf), ergibt sich für ein beliebiges Ereignis A die Wahrscheinlichkeit

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der für } A \text{ “günstigen” Fälle}}{\text{Anzahl der möglichen Fälle}},$$

wobei $|A|$ die Anzahl der Elemente von A bezeichnet bzw. $|\Omega|$ die Anzahl der Elemente von Ω.

Solch ein Ereignis nennt man LAPLACE'sches (Zufalls-)EXPERIMENT, die dazugehörige Wahrscheinlichkeit ist die Laplace-Wahrscheinlichkeit.

Ein Beispiel: Auf 10 Karten werden je eine der Zahlen $1, \ldots, 10$ geschrieben. Wie groß ist die Wahrscheinlichkeit, von diesem (verdeckten) Kartenstapel eine Zahl > 5 zu ziehen? Es sind

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

also

$$\begin{aligned} |\Omega| &= 10 \\ |A| &= 5 \end{aligned}$$

d.h.

$$P(A) = \frac{|A|}{|\Omega|} = \frac{5}{10} = 0.5.$$

In Worten: Die Wahrscheinlichkeit, eine Karte mit einer Zahl größer als 5 zu ziehen, ist 0.5 bzw. 50%.

Die bedingte Wahrscheinlichkeit

Manche Ereignisse sind von anderen abhängig. Die Wahrscheinlichkeit solcher Ereignisse nennt man **bedingte Wahrscheinlichkeit**.

z.B. die Wahrscheinlichkeit vom Erkranken an Lungenkrebs abhängig vom Zigarettenkonsum.

Zur Illustration der bedingten Wahrscheinlichkeit betrachten wir ein Beispiel: Je 1000 Männer und Frauen werden befragt, ob sie Motorrad fahren. Folgendes Schema zeigt das Ergebnis (Tabelle 1):

Die einzelnen Ereignisse sind also:

Tabelle 1: Anteil von MotorradfahrerInnen

	Mann (M)	Frau (F)	
MotorradfahrerIn (J): Ja	125	23	148
MotorradfahrerIn: Nein (N)	875	977	1852
	1000	1000	2000(=n)

- J: MotorradfahrerIn
- N: Kein(e) MotorradfahrerIn
- M: Mann
- F: Frau

Dieser Tabelle können wir etwa entnehmen, daß die relative Häufigkeit von MotorradfahrerInnen $f(J)/n = 148/2000 = 7.4\%$ beträgt, die der Nicht-MotorradfahrerInnen $f(N)/n = 1852/2000 = 92.6\%$.

Wir interessieren uns nun für die **bedingte relative Häufigkeit**, etwa unter der Bedingung F, d.h wie groß ist die Wahrscheinlichkeit, daß eine Person (nicht) Motorrad fährt, falls sie weiblich ist.

Folgende Grössen sind zu bestimmen:

- $r(J|F)$
- $r(N|F)$

$r(A|B)$ ist die relative Häufigkeit vom Ereignis A unter der Bedingung des Eintretens von Ereignis B bezeichnet; die absolute bedingte Häufigkeit wird mit $f(A|B)$ symbolisiert.

Man bekommt

$$r(J|F) = \frac{23}{1000} = 0.023,$$

da die relative Häufigkeit der Quotient aus absoluter Häufigkeit (in diesem Fall 23, die Anzahl motorradfahrender Frauen) und Menge der Beobachtungen (in diesem Fall 1000 Frauen) ist.

Das heißt

$$\begin{aligned} r(J|F) &= \frac{23}{1000} = \frac{f(J \wedge F)}{f(F)} = \frac{r(J \wedge F)/n}{r(F)/n} = \frac{r(J \wedge F)}{n} * \frac{n}{r(F)} \\ &= \frac{r(J \wedge F)}{r(F)} \end{aligned}$$

Analog erhält man

$$r(N|F) = \frac{r(N \wedge F)}{r(F)}$$

bzw. ür allgemeine Ereignisse A und B gilt

$$r(A|B) = \frac{r(A \wedge B)}{r(B)}.$$

Durch diese Ergebnisse motiviert definieren wir die **bedingte Wahrscheinlichkeit eines Ereignisses A unter der Bedingung B** als

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

Man kann die bedingte Wahrscheinlichkeit dazu benutzen, um die statistische (Un-)Abhängigkeit von Ereignissen zu erfassen. Zwei Ereignisse sind dann unabhängig, wenn das Eintreten eines Ereignisses keinen Einfluss auf das Eintreten des anderen Ereignisses hat.

Wenn wir die Ereignisse J (MotorradfahrerIn) und M (Mann) unseres vorherigen Beispiels herannehmen, so sind diese Ereignisse *nicht* voneinander unabhängig, da die Wahrscheinlichkeit, ob eine Person Motorrad fährt, vom Geschlecht abhängig ist.

Falls man die Ereignisse

- A: Eine Person inskribiert Psychologie
- B: Es regnet mindestens 80 Tage im Jahr

betrachtet, so gilt klarerweise

$$P(A|B) = P(A) \qquad \text{bzw.} \qquad P(B|A) = P(B)$$

da die Anzahl der Regentage im Jahr keinen Einfluss auf die Wahl des Studiums einer Person hat (und umgekehrt). D.h. die Ereignisse sind unabhängig, wenn

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} = P(A) \qquad \text{bzw.} \qquad P(B|A) = \frac{P(B \wedge A)}{P(A)} = P(B).$$

Es muß also gelten

$$P(A \wedge B) = P(A)P(B).$$

Wir fassen zusammen:

Satz (Multiplikationsregel): *Zwei Ereignisse A und B sind genau dann unabhängig, wenn*

$$P(A \wedge B) = P(A)P(B),$$

wobei $P(A) \neq 0$ *und* $P(B) \neq 0$.