Ereignisse, Wahrscheinlichkeiten

Ein Experiment (Messung, Beobachtung), unter **kontrollierten Bedingungen** (Werfen eines Würfels), das **beliebig oft wiederholbar** und dessen **Ergebnis** (unabhängig ob nominal, ordinal oder metrisch) **nicht vorhersehbar** ist (eine der Augenzahlen 1 bis 6 wird beim Würfeln erscheinen, welche ist ungewiss), heißt Zufallsexperiment.

Jede einzelne Augenzahl, die erscheint, ist ein Elementarereignis und die Menge aller möglichen Elementarereignisse eines Zufallsexperimentes heißt Ereignisraum **S**, beim Würfel S = {1,2,3,4,5,6}.

Jede Teilmenge von S heißt Ereignis. Zum Beispiel ist das Werfen ungerader Augenzahlen beim Würfeln das Ereignis A = {1,3,5}. Es tritt ein, wenn eines seiner Elementarereignisse eintritt, wenn also eine 1 oder eine 3 oder eine 5 gewürfelt wird. Ein anderes Ereignis B = {1,2} ist das Werfen einer Zahl kleiner als 3.

Das Ereignis **A U B** = {1,2,3,5} (entweder A oder B) tritt ein, wenn eine der Augenzahlen gewürfelt wird, die entweder in A oder in B vorkommen.

Das Ereignis **A ∩ B** = {1} (sowohl A als auch B) tritt ein, wenn eine Augenzahlen gewürfelt wird, die in A und in B vorkommt.

Bei disjunkten Ereignissen A und B ist A ∩ B = Ø (leere Menge).
Ist beim Würfeln beispielsweise A = {1,3,5} und B = {2,4,6}, so ist A ∩ B = Ø und A U B = S. Außerdem ist in diesem Fall B = $\bar{A}$ (nicht A), das komplementäre Ereignis zu A.

Eine grafische Darstellung der Zusammenhänge zwischen zwei Ereignissen liefern die Venn-Diagramme, die unter anderem bei: Hedderich, Sachs: Angewandte Statistik. 16. Aufl., Springer 2018 zu finden sind.

Durch Zuordnung von Wahrscheinlichkeiten (WSK) zu den Ereignissen lässt sich zum Beispiel die Frage beantworten, wie wahrscheinlich es ist, beim einmaligen Würfeln eine 6 zu erhalten. Ist A = {6}, dann ist die WSK W(A) natürlich W(A) = 1/6 für einen idealen Würfel. Die Anzahl der „günstigen Fälle" ist 1 , denn es gibt nur eine Seite des Würfels mit der Augenzahl 6. Die Anzahl aller „möglichen Fälle" ist 6, denn der Würfel hat bekanntlich 6 Seiten und jede Augenzahl tritt mit der gleichen WSK auf. Diese Überlegung führt zur Definition der WSK nach Laplace:

W = (Anzahl der günstigen Fälle) / (Anzahl aller möglichen Fälle)

Fragt man nach der WSK für die Blutgruppe 0 einer zufällig ausgewählen Person, so muss berücksichtigt werden, dass die Blutgruppen 0,A,B,AB mit unterschiedlicher WSK auftreten, d.h., W(0) ≠ 1/4. Hier hilft die Definition der WSK nach Richard von Mises. Man geht dabei im Beispiel der Blutgruppe 0 von einer großen Zahl n von Personen aus und zählt, mit welcher Häufigkeit $h_n(0)$ die Blutgruppe 0 erscheint. W(0) ist dann die relative Häufigkeit $h_n(0)$ / n für n --> unendlich.

$$W(0) = \lim_{n \to \infty} \frac{h_n(0)}{n}$$

Bei einer genügend großen Anzahl von Versuchen lässt sich die WSK durch die relative Häufigkeit schätzen.

Die WSK für die Blutgruppen in Deutschland betragen:
W(0) = 0,41 ; W(A) = 0,43 ; W(B) = 0,11 ; W(AB) = 0,05

Axiome von Kolmogorov:
Sind A und B Ereignisse, so ist
1. 0 ≤ W(A) ≤ 1
2. W(S) = 1
3. **W(A U B) = W(A) + W(B)** für disjunkte Ereignisse A und B (A ∩ B = Ø).

Wenn A und B nicht disjunkt sind, gilt der Additionssatz:
Additionssatz:
Sind A und B zwei beliebige Ereignisse, so gilt **W(A U B) = W(A) + W(B) - W(A ∩ B)**

Betrachten das obige Beispiel: Werfen ungerader Augenzahlen beim Würfeln sei A = {1,3,5} und Werfen einer Augenzahl kleiner als 3 sei B = {1,2}. Dann ist A U B = {1,2,3,5} und nach Laplace W(A U B) = 4/6. Das gleich Ergebnis erhält man wegen A ∩ B = {1} mit dem Additionssatz: W(A U B) = 3/6 + 2/6 - 1/6 = 4/6.

Beispiel zur bedingten WSK: Krankheit und Geschlecht
Die WSK für das Eintreten einer Krankheit kann beispielsweise vom Geschlecht (m,w) abhängen. In einer kleinen Gruppe von 79 Personen trat eine Krankheit insgesamt 53 mal auf. Betroffen waren 26 von 38 Männern und 27 von 41 Frauen. Die Daten lassen sich in einer Vierfeldertafel (2x2-Tafel) zusammenfassen.

	k	$\overline{k}$	
w	27	14	41
m	26	12	38
	53	26	79

Hier gibt es vier Ereignisse: w, m, k, $\overline{k}$ mit ihren Wahrscheinlichkeiten
W(w) = 41/79 , W(m) = 38/79 , W(k) = 53/79 und W($\overline{k}$) = 26/79

mit k = krank und $\overline{k}$ = nicht krank
Die bedingte WSK krank zu sein bei Frauen, also der Anteil der Kranken unter den Frauen beträgt:
W(k|w) = 27/41 = (27/79) / (41/79) = 0,659 = W(k ∩ w) / W(w) oder allgemein als

Multiplikationssatz:
W(A|B) = W(A ∩ B) / W(B) oder **W(A ∩ B) = W(B) · W(A|B)**

Für unabhängige Ereignisse ist W(A|B) = W(A) und damit **W(A ∩ B) = W(B) · W(A) = W(A) · W(B)**

In diesem Beispiel ist W(k ∩ w) = 27/79 = 0,342 und W(k) · W(w) = 53/79 · 41/79 = 0,348
Beide Werte unterscheiden sich nur wenig, sodass man Unabhängigkeit vermuten kann.
Interpretiert man W(k|w) = 27/41 als Risiko zu erkranken unter den Frauen und W(k|m) = 26/38 als Risiko zu erkranken unter den Männern, dann ergibt sich das **relative Risiko RR** = W(k|w) / W(k|m) = 0,96 mit einem Konfidenzintervall (Berechnung mit **OpenEpi**) von (0,71 ; 1,31). Da die „1" im Konfidenzintervall liegt, ist das Risiko der Erkrankung nicht vom Geschlecht abhängig, die zwei Ereignisse „Krankheit" und „Geschlecht" also unabhängig voneinander.
Schließlich liefert auch der Chi²-Test einen exakten p-Wert von 0,998 und damit Unabhängigkeit.

W(k|w) ist der Anteil der Kranken unter den Frauen. Den Anteil der Frauen unter den Kranken **W(w|k)** erhält man aus der Bayes - Formel, die im Falle zweier Ereignisse lautet:

$$W(A|B) = \frac{W(A)\cdot W(B|A)}{W(A)\cdot W(B|A) + W(\overline{A})\cdot W(B|\overline{A})}$$

In diesem Beispiel ist W(w|k) = 27/53 = {(41/79)·(27/41)} / {(41/79)·(27/41) + (38/79)·(26/38)} = 0,509

Beispiel zur bedingten WSK: Blutgruppen und Rhesusfaktor
Neben der Blutgruppe ist auch der Rhesusfaktor von Interesse. Eine repräsentative Studie in Deutschland ergab bei 100 Personen folgende Häufigkeiten:

Blutgruppe	0	A	B	AB	
Rh+	35	37	9	4	85
Zeilen%	41,2%	43,5%	10,6%	4,7%	100%
Rh-	6	6	2	1	15
Zeilen%	40,0%	40,0%	13,3%	6,7%	100%
	41	43	11	5	100

Blutgruppen und Rhesusfaktoren sind disjunkte Ereignisse. Wie groß ist die WSK für eine Person Blutgruppe A zu haben in der Subgruppe der Rh+ Personen, also W(A|Rh+) ?
W(A) = 43/100 , W(Rh+) = 85/100 , W(A ∩ Rh+) = 37/100 , dann ist W(A|Rh+) = 37/85 = 43,5%
Analog sind W(0|Rh+) = 35/85 = 41,2% , W(B|Rh+) = 9/85 = 10,6% und W(AB|Rh+) = 4/85 = 4,7% .

Fazit: Die bedingten WSK'n sind die Zeilen- oder Spaltenprozente der jeweiligen Subgruppen.