Michael Buhlmann

Mathematik-Formelsammlung
> Stochastik/Wahrscheinlichkeitsrechnung
> Grundbegriffe (diskrete Stochastik)

Stochastik oder Wahrscheinlichkeitsrechnung ist die Beschäftigung mit Wahrscheinlichkeiten, also mit mathematischen Größen p (0≤p≤1, p reell), die im Rahmen von Zufallsexperimenten als Maß für die Sicherheit bzw. Unsicherheit eines Ergebnisses bzw. Ereignisses in Erscheinung treten. Zufallsexperimente (Zufallsversuche, Zufallsvorgänge) sind mathematisch modellierte Prozesse, die auf der (endlichen) Wiederholung (Mehrstufigkeit) einer gleichen festgelegten Versuchssituation (Merkmale, Versuchsausgänge) beruhen, wobei die (abzählbar-endlichen) möglichen Ergebnisse einer solchen Versuchsdurchführung ebenso wie die Ergebniswahrscheinlichkeiten (als relative Häufigkeiten [Gesetz der großen Zahlen]) bekannt sind. Zufallsexperimente lassen sich durch sog. Wahrscheinlichkeitsbäume (aus Knoten, Verzweigungen [Ausgänge, Merkmalsausprägungen], Kanten [Zweige] und Pfaden [Äste]) darstellen, die Ergebnisse und Wahrscheinlichkeiten anzeigen. Zufallsexperimente, die auf Ergebnisse mit immer derselben Wahrscheinlichkeit hinführen, heißen Laplace-Experimente. Ergebnisse sind Elementarereignisse, Ereignisse sind Zusammenfassungen von Ergebnissen (Mengenlehre der Ereignisse), die Wahrscheinlichkeit eines Ereignisses errechnet sich gemäß den Pfadregeln (Multiplikation von Wahrscheinlichkeiten innerhalb eines Pfades, Addition von Wahrscheinlichkeiten verschiedener Pfade); die Wahrscheinlichkeit aller Ergebnisse eines Zufallsexperiments stellt eine Wahrscheinlichkeitsverteilung dar; die Summe der Wahrscheinlichkeiten aller Ergebnisse ergibt das sichere Ereignis. Aus diesen Sachverhalten folgt die Axiomatik der Wahrscheinlichkeiten mit den daraus abgeleiteten Formeln des Additionssatzes und des Gegenereignisses (De Morgansche Regeln). Eine Zufallsvariable ist eine Funktion, die Ereignissen des Zufallsexperiments eine reelle Zahl zuordnet. Bzgl. der Zufallsvariablen lassen sich Aussagen zu Erwartungswert und Standardabweichung treffen.

Ist S die Menge der Ergebnisse eines Zufallsexperiments, so stellen Teilmengen A, B, … von S (A, B, … ⊂ S) Ereignisse dar. Hinsichtlich eines Ergebnisses aεS ergibt sich als Wahrscheinlichkeit des dazugehörigen Elementarereignisses {a} ⊂ S als p({a}), die Wahrscheinlichkeit p(A) eines Ereignisses A als Summe der Wahrscheinlichkeiten der Ergebnisse, die im Ereignis enthalten sind.
Im Fall eines Laplace-Experiments errechnet sich die Wahrscheinlichkeit als:

$$p(A) = \frac{g}{n}$$

mit: g = Anzahl der das Ereignis A ausmachenden „günstigen“ Versuchsergebnisse, n = Anzahl aller möglichen Versuchsergebnisse. Zur Ermittlung der Anzahlen sind mitunter kombinatorische Überlegungen notwendig:

Kombinatorik	
Voraussetzungen	**Permutationen**
n Elemente werden angeordnet, alle Elemente sind verschieden	Anzahl der Möglichkeiten: n!
n Elemente werden angeordnet, jeweils n_i Elemente sind gleich (i = 1, 2, … r; $n_1 + n_2 + \ldots + n_r = n$)	Anzahl der Möglichkeiten: $\frac{n!}{n_1! \cdot \ldots \cdot n_r!}$
Voraussetzungen	**Kombinationen**
k Elemente werden aus n Elementen ausgewählt, mit Zurücklegen, mit Berücksichtigung der Reihenfolge	Anzahl der Möglichkeiten: n^k

k Elemente werden aus n Elementen ausgewählt, mit Zurücklegen, ohne Berücksichtigung der Reihenfolge	Anzahl der Möglichkeiten: $\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-k)!}$
k Elemente werden aus n Elementen ausgewählt, ohne Zurücklegen, mit Berücksichtigung der Reihenfolge	Anzahl der Möglichkeiten: $\binom{n}{k} \cdot k! = \frac{n!}{(n-k)!}$
k Elemente werden aus n Elementen ausgewählt, ohne Zurücklegen, ohne Berücksichtigung der Reihenfolge	Anzahl der Möglichkeiten: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
	Fakultät: $n! = 1 \cdot 2 \cdot \ldots \cdot n$, $0! = 1$ Binomialkoeffizient: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ (für natürliche Zahlen k, n einschließlich 0)

Für <u>Wahrscheinlichkeiten</u> p <u>und Ereignisse</u> A, B, … $\subset$ S eines Zufallsexperiments gilt:

Wahrscheinlichkeiten und Ereignisse	
Ereignisse	**Wahrscheinlichkeiten**
Leeres/unmögliches Ereignis Ø	p(Ø) = 0
Sicheres Ereignis S	p(S) = 1
Ereignis A	$0 \leq p(A) \leq 1$
Ereignisse A, B disjunkt -> $A \cup B$ als Ereignis „A oder B“	$p(A \cup B) = p(A) + p(B)$
Ereignisse A, B -> $A \cup B$ als Ereignis „A oder B“, $A \cap B$ als Ereignis „A und B“	Additionssatz für Wahrscheinlichkeiten: $p(A \cup B) = p(A) + p(B) - p(A \cap B)$
Ereignisse A, B, $B \subset A$ -> $A \backslash B$ als Ereignis „A ohne B“	$p(A \backslash B) = p(A) - p(B)$
Ereignis A -> Gegenereignis $\bar{A} = S \backslash A$ als Ereignis „nicht A“	$p(A) + p(\bar{A}) = 1$, $p(\bar{A}) = 1 - p(A)$

Eine (diskrete) <u>Zufallsvariable</u> X ist eine reelle Abbildung zur Beschreibung eines Zufallsexperiments. Sie ordnet daher den (abzählbar-endlichen) möglichen Ergebnissen des Versuchs eine reelle Zahl zu, also: $X(a_1) = x_1$, $X(a_2) = x_2$, … $X(a_n) = x_n$ mit Ergebnissen $a_1, a_2, \ldots, a_n$. Es ergibt sich die durch die Zufallsvariable abgeleitete <u>Wahrscheinlichkeitsverteilung</u>:

$$x_1: p(X=x_1),\ x_2: p(X=x_2),\ \ldots,\ x_n: p(X=x_n)$$

mit $p(X=x_1)+p(X=x_2)+\ldots+p(X=x_n) = 1$ und daraus der <u>Erwartungswert</u>

$$E(X) = \mu = x_1 p(X=x_1) + x_2 p(X=x_2) + \ldots + x_n p(X=x_n)$$

(als durchschnittlicher [mittlerer] Wert der Zufallsvariablen) sowie die <u>Standardabweichung</u>

$$\sigma = \sqrt{Var(X)} = \sqrt{(x_1-\mu)^2 p(X=x_1) + (x_2-\mu)^2 p(X=x_2) + \ldots + (x_n-\mu)^2 p(X=x_n)}$$

(als Maß für die Abweichung vom Erwartungswert). Zufallsvariablen X sind durch ihren Erwartungswert μ und ihre Standardabweichung σ charakterisierbar. Aufbauend auf der Wahrscheinlichkeitsverteilung und der daraus abgeleiteten Wahrscheinlichkeitsfunktion $f(x_i) = P(X=x_i)$ (i=1, …n) ergibt sich die <u>Verteilungsfunktion</u> $F(x) = p(X \leq x)$ als Summe der Wahrscheinlichkeiten der Ausprägungen der Zufallsvariablen X, die kleiner gleich einer vorgegebenen reellen Zahl x sind.

Ein <u>Wahrscheinlichkeitsbaum</u> dient der grafischen Aufbereitung und Beschreibung eines n-stufigen Zufallsexperiments mit $k_1, k_2, \ldots$ Verzweigungen von Merkmalen (Ausgänge, Merkmalsausprägungen, je Stufe) (n, $k_1, k_2, \ldots$ als natürliche Zahlen) und besteht aus einer dem Zufallsversuch und dessen (kausaler) Entscheidungshierarchie entsprechenden (gerichteten) Anordnung von Knoten und Kanten.

Wahrscheinlichkeitsbaum (Merkmale: Merkmal A (Ausgänge a, c), Merkmal B (Ausgänge b, d); zweistufig):

	Merkmal A	Merkmal B	◀ *Merkmale*					
		p(b) **b**	>	p(**a; b**) =	p(a)p(b)	**1**	◀ *Pfad (a;b)*	◀
	p(a) **a**							
		1-p(b) **d**	>	p(**a; d**) =	p(a)(1-p(b))	**2**	◀ *Pfad (a;d)*	*Ereignis E:* p(A) = p(a;b)+p(c;d) = p(a)p(b)+ p(c)p(d)
		1-p(d) **b**	>	p(**c; b**) =	p(c)(1-p(d))	**3**	◀ *Pfad (c;b)*	
	p(c) **c**							
		p(d) **d**	>	p(**c; d**) =	p(c)p(d)	**4**	◀ *Pfad (c;d)*	◀
	Wurzel ▲ Knoten, Kanten ▲ Blätter							
	1. Stufe	*2. Stufe*	Summe:		1			
	▲ *Wahrscheinlichkeiten, Merkmalsausprägungen*				▲ *Wahrscheinlichkeitsverteilung*		▲ *Ergebnisse*	

Die Knoten stellen die (Zwischen-) Ergebnisse des Zufallsversuchs dar gemäß der jeweiligen Durchführung des Experiments, die Kanten sind mit den dem jeweiligen Ergebnis entsprechenden Wahrscheinlichkeiten versehen. Die Wahrscheinlichkeiten im Wahrscheinlichkeitsbaum ergeben sich aus kombinatorischen Überlegungen (Urnenmodell mit/ohne Zurücklegen u.a.). Die Summe der Wahrscheinlichkeiten an den Kanten, die von einem Knoten ausgehen, ist 1.
Die Aufeinanderfolge von Kanten zwischen dem Anfangsknoten des Wahrscheinlichkeitsbaums (Wurzel) und den Endknoten (Blättern) heißt Pfad. Für ein Ereignis A gelten hinsichtlich der das Ereignis charakterisierenden Pfade und deren Wahrscheinlichkeiten die Pfadregeln:

1. Entlang eines zum Ereignis E gehörenden Pfades werden die Wahrscheinlichkeiten der Pfadkanten multipliziert. Es ergibt die Wahrscheinlichkeit des Pfades (Pfad: O -> p(a) -> a -> p(b) -> b …: p(a;b;…) = p(a)·p(b)·…).
2. Die multiplizierten Wahrscheinlichkeiten von allen zum Ereignis E gehörenden Pfaden werden addiert. Es ergibt sich die Wahrscheinlichkeit des Ereignisses E (Pfade: O -> p(a) -> a -> p(b) -> b …; O -> p(c) -> c -> p(d) -> d …; …: p(E) = p(a;b;…) + p(c;d;…) + …).

Ein Pfad in einem Wahrscheinlichkeitsbaum stellt ein Elementarereignis dar. Die Summe der Wahrscheinlichkeiten aller Elementarereignisse ist folglich 1. Wird ein Ereignis E durch mehr Pfade charakterisiert als sein Gegenereignis E^{-}, so empfiehlt sich, die Wahrscheinlichkeit des Gegenereignisses zu berechnen, so dass die Formel $p(E) = 1 – p(E^{-})$ zum Tragen kommt.
Je nach Fragestellung können vorgegebene Wahrscheinlichkeitsbäume umorganisiert, reduziert und „eingeklappt“ werden, indem bestimmte Merkmalsausprägungen zusammengefasst und somit die Zahl der Kanten vermindert werden. Allgemein ist zur Anzahl der Pfade in einem Wahrscheinlichkeitsbaum zu sagen, dass bei n-maligem Durchführen eines Zufallsversuchs mit gleichen k Ausgängen je Stufe die Pfadanzahl k^n beträgt; bei unterschiedlicher Anzahl von Ausgängen k_1, k_2, …, k_n ist die Pfadanzahl $k_1 \cdot k_2 \cdot \ldots \cdot k_n$.