Diskrete Strukturen und Logik

WiSe 2007/08 in Trier

Henning Fernau
Universität Trier
fernau@uni-trier.de

Diskrete Strukturen und Logik Gesamtübersicht

- Organisatorisches
- Einführung
- Logik & Mengenlehre
- Beweisverfahren
- Kombinatorik: Die Kunst des Zählens
- algebraische Strukturen

Diskrete Stochastik

. . . ist ein vornehmes Wort für die bekannte (?) Wahrscheinlichkeitsrechnung.

Im Gegensatz zur kontinuierlichen Stochastik legen wir fest:
Eine abzählbare Menge heißt auch *Ereignisraum*.
Die Elemente eines Ereignisraums heißen *elementare Ereignisse*.

Deutung: Ein *Zufallsexperiment* liefert ein Ergebnis, das nicht genau vorherzusagen ist. Jedes mögliche Ergebnis nennt man auch elementares Ereignis.

Diskrete Stochastik—Beispiele

Beispiel: Ein *Münzwurf* hat zwei elementare Ereignisse: “Kopf” (oben) oder “Zahl”.

Beispiel: Der *Würfelwurf* hat sechs elementare Ereignisse.

Beispiel: Unter Vernachlässigung der Zusatzzahl sind elementare Ereignisse des Lottos “6 aus 49” alle 6-elementigen Teilmenge der natürlichen Zahlen zwischen 1 und 49.

Der Ereignisraum in den Beispielen umfasst 2 bzw. 6 bzw. 13983816 Elemente.

Ereignisse sind Teilmengen des Ereignisraums.

Beispiel: Der Ereignisraum des Würfelns mit zwei Würfeln umfasst $36 = |\{1, \ldots, 6\}^2|$ Elemente.
Das Ereignis, die Augensumme 7 zu würfeln, ist die Teilmenge:

$$\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}.$$

Beispiel: Wie viele Elemente enthält das Ereignis, dass der Lotto-Tipp $\{2, 5, 6, 8, 12, 14\}$ genau fünf der gezogenen 6 Zahlen enthält ?

Wahrscheinlichkeitsdichte und -verteilung

Eine Funktion P, die jedem Elementarereignis $x \in S$ eine nichtnegative Zahl $P(x)$ zuordnet, heißt *Wahrscheinlichkeitsdichte* oder *Wahrscheinlichkeitsfunktion*, falls $\sum_{x \in S} P(x) = 1$ gilt.

P lässt sich leicht auf Ereignisse verallgemeinern durch $P(A) = \sum_{x \in A} P(x)$ und wird dann auch als *Wahrscheinlichkeitsverteilung* angesprochen.

Eigenschaften

$1 \geq P(A) \geq 0$ für jedes Ereignis A.

$P(A \cup B) = P(A) + P(B)$ für disjunkte Ereignisse A, B.

$P(B \setminus A) = P(B) - P(A)$ und $P(A) \leq P(B)$ für $A \subseteq B$.

$P(\bar{A}) = P(S \setminus A) = 1 - P(A)$.

$P(\emptyset) = 0$.

Axiomatischer Zugang

Das Paar (S, P) heißt *diskreter Wahrscheinlichkeitsraum*, wenn gilt:

1. S ist eine abzählbare Menge.
2. $P : 2^S \longrightarrow \mathbb{R}$ erfüllt:

(2a) $\forall A \subseteq S : P(A) \geq 0$ (*Nichtnegativität*)

(2b) $P(S) = 1$ (*Normiertheit*)

(2c) für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweise fremder Mengen aus S gilt:

$$P\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} P(A_n) \qquad (\sigma\text{-Additivität})$$

P heißt dann auch ein *(diskretes) Wahrscheinlichkeitsmaß*.

Beweise der Eigenschaften (ausgehend von der Axiomatik)

Betrachte $A_0 = S$ und $A_n = \emptyset$ für $n \geq 1$. $\leadsto$

$$1 = P(S) = P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} P(A_n) = P(S) + P(\emptyset) + P(\emptyset) + \cdots = 1 + \infty \cdot P(\emptyset)$$

$\leadsto P(\emptyset) = 0$.

Es seien A und B disjunkt. Betrachte $A_0 = A$, $A_1 = B$ und $A_n = \emptyset$ für $n \geq 2$. $\leadsto$

$$P(A \cup B) = P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} P(A_n) = P(A) + P(B) + P(\emptyset) + \cdots = P(A) + P(B) + 0$$

Also gilt: $P(A \cup B) = P(A) + P(B)$.

$1 = P(S) = P(A \cup \bar{A}) = P(A) + P(\bar{A}) \leadsto$ (1) $P(\bar{A}) = 1 - P(A) \leadsto$ (2) $P(A) \leq 1$ (wegen $P(\bar{A}) \geq 0$)

Für $A \subseteq B$ gilt: $B = (B \setminus A) \cup A$ (disjunkte Vereinigung)
$\rightsquigarrow$ (1) $P(B \setminus A) = P(B) - P(A)$ und (2) $P(A) \leq P(B)$ (wegen $P(B \setminus A) \geq 0$)

Wegen der σ-Additivität des Wahrscheinlichkeitsmaßes ist es durch die Festlegung der Wahrscheinlichkeiten der Elementarereignisse (also durch die Wahrscheinlichkeitsdichte) schon vollständig beschrieben.

Zusammenhang mit der Kombinatorik

Einfachster Fall: Alle elementaren Ereignisse sind gleichwahrscheinlich.
Dann gilt: $\forall x \in S : P(x) = 1/|S|$. (*Gleichverteilung*)

Für ein Ereignis $A \subseteq S$ gilt somit:

$$P(A) = \sum_{x \in A} P(x) = \frac{|A|}{|S|}.$$

Es kommt daher essentiell darauf an, die Elemente von A zu zählen.

Das Geburtstagsproblem

Wie viele Gäste muss man zu einer Party einladen, demit mit Wahrscheinlichkeit $\geq \frac{1}{2}$ zwei Gäste am selben Tag Geburtstag haben ?

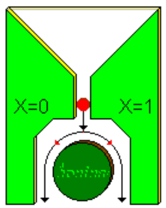
Vereinfachende Annahmen: Es gibt 365 Tage im Jahr und alle Tage seien für Geburten gleichwahrscheinlich; zudem sei unsere Gästeauswahl nicht z.B. auf Gäste konzentriert, die am 2.1. geboren wurden.

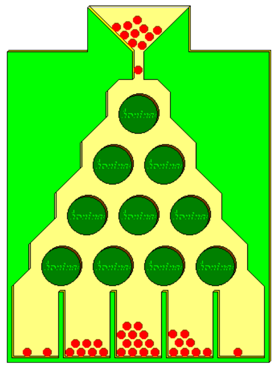
Der Ereignisraum S_k, der sich durch Einladen von k Gästen ergibt, lässt sich durch $[365]^k$ beschreiben, d.h., $|S_k| = 365^k$ nach der Potenzregel.
Das Ereignis E, dass alle Gäste an verschiedenen Tagen Geburtstag haben, lässt sich als Menge der k-Permutationen von $[365]$ ansehen, d.h., $|E| = \begin{bmatrix} 365 \\ k \end{bmatrix}$.

Also gilt: $P(E) = \frac{365 \cdot 364 \cdots (365-k+1)}{365^k}$.
$\leadsto$ 23 Gäste genügen !
PS: Auf dem Mars mit 669 Tagen müsste man 31 Gäste einladen.

Wiederholte Münzwürfe

Ein Münzwurf lässt sich auch durch den Weg einer Kugel in folgendem Diagramm (einfachstes *Galton-Brett*) von oben nach unten veranschaulichen.





Sind wir nur an der Zahl j der Zahl-oben-Ereignisse bei einem wiederholten Münzwurfexperiment interessiert, so entspricht dies der Zahl der Kugeln im j-ten Fach des n-fach kaskadierten Galtonbretts. n über j von insgesamt 2^n Möglichkeiten (elementaren Ereignissen) beinhaltet dieses Ereignis.

Bernoulli-Experimente: Münzwürfe mit Wahrscheinlichkeit p für “Zahl”

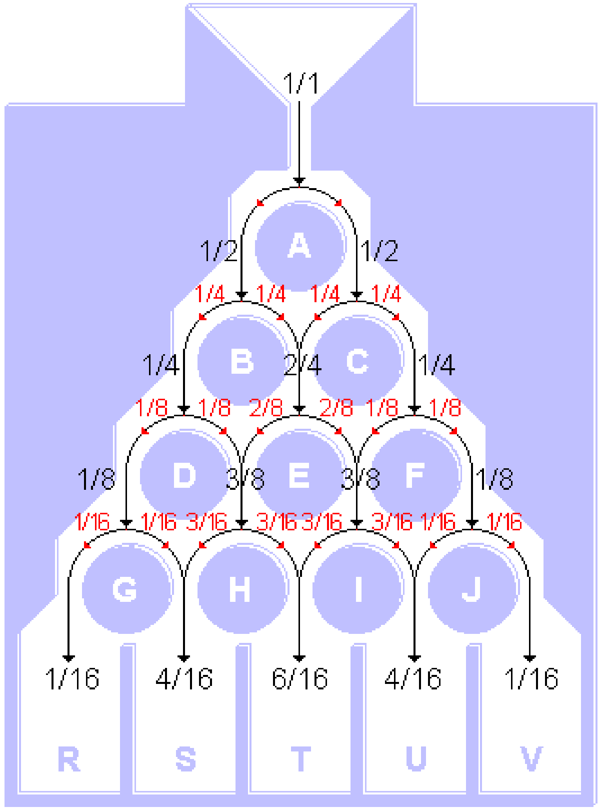
$b(k; n, p)$: Wahrscheinlichkeit, bei n unabhängigen Wiederholungen eines Bernoulli-Versuches mit Wahrscheinlichkeit p für “Zahl” genau k mal “Zahl” zu werfen.

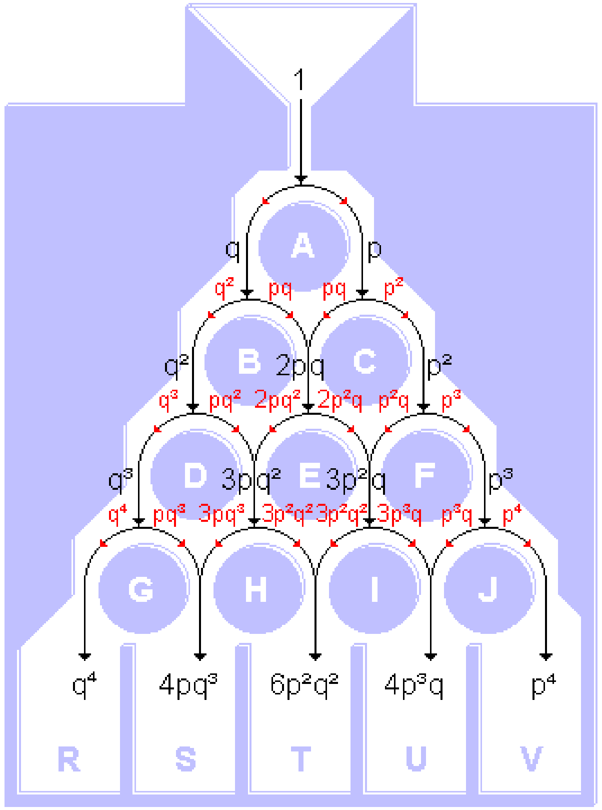
Satz: (*Binomialverteilung*) $b(k; n, p) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$.

Beweis: durch Induktion über n:
IA: $n = 1$ ✓
IV: Die Behauptung gelte für $n = m$.
IB: Die Behauptung gilt für $n = m + 1$.
Es gibt zwei Möglichkeiten, mit $m + 1$ Versuchen genau k Zahlen zu werfen.
1. Im $m + 1$. Versuch wird "Kopf" geworfen, aber in den vorigen m Versuchen genau k-mal Zahl.
2. Im $m + 1$. Versuch wird "Zahl" geworfen und in den vorigen m Versuchen genau $(k - 1)$-mal Zahl.

$$\begin{aligned}
b(k; m+1, p) &= b(k; m, p)(1-p) + b(k-1; m, p)p \\
&= \begin{pmatrix} m \\ k \end{pmatrix} \cdot p^k (1-p)^{m-k} (1-p) + \begin{pmatrix} m \\ k-1 \end{pmatrix} \cdot p^{k-1} (1-p)^{m-k+1} p \\
&= \left(\begin{pmatrix} m \\ k \end{pmatrix} + \begin{pmatrix} m \\ k-1 \end{pmatrix} \right) p^k (1-p)^{m-k+1} \\
&= \begin{pmatrix} m+1 \\ k \end{pmatrix} p^k (1-p)^{(m+1)-k}
\end{aligned}$$

Den Zusammenhang mit dem binomischen Lehrsatz (mit $q = 1 - p$) veranschaulicht das Galton-Brett:





Bedingte Wahrscheinlichkeit

Es sei S ein Ereignisraum mit Wahrscheinlichkeitsfunktion $P : S \to \mathbb{R}$.

Die *bedingte Wahrscheinlichkeit* von Ereignis $A \subseteq S$ unter der Voraussetzung, dass Ereignis $P(B) \neq 0$ eintritt, ist definiert als:

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

Beobachte: $P(A|B) = P(A'|B)$ gdw. $P(A \cap B) = P(A' \cap B)$; dies ist insbesondere dann der Fall, wenn $A \cap B = A' \cap B$ gilt.
$\leadsto$ nur die Wahrscheinlichkeitsverteilung "im Bereich B" ist von Interesse
Tatsächlich definiert $P_B(x) := P(x)/P(B)$ für $x \in B$ eine Wahrscheinlichkeitsfunktion auf dem Ereignisraum B.
Beweis: $\sum_{x \in B} P_B(x) = \sum_{x \in B} P(x)/P(B) = 1$.
Diese Deutung der bedingten Wahrscheinlichkeit ist oft bequem.

Ein Beispiel Münzwurf mit zwei (fairen) Münzen

Ereignisraum $S = \{(\text{Kopf}, \text{Kopf}), (\text{Kopf}, \text{Zahl}), (\text{Zahl}, \text{Kopf}), (\text{Zahl}, \text{Zahl})\}$

Wie groß ist die Wahrscheinlichkeit, dass zwei Münzwürfe beide “Kopf” ergeben, wenn man sicher weiß, dass mindestens einer der Würfe “Kopf” ergibt ?

$A = \{(\text{Kopf}, \text{Kopf})\}$
$B = \{(\text{Kopf}, \text{Kopf}), (\text{Kopf}, \text{Zahl}), (\text{Zahl}, \text{Kopf})\}$.

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{1/4}{3/4} = \frac{1}{3}$
oder: $P_B((\text{Kopf}, \text{Kopf})) = 1/3$ direkt, da $(\text{Kopf}, \text{Kopf})$ Elementarereignis der Gleichverteilung über B ist.

Unabhängigkeit

Zwei Ereignisse A und B heißen *unabhängig*, gdw. $P(A \cap B) = P(A) \cdot P(B)$.

Beispiel: Wir betrachten wieder zwei Münzwürfe.
A bedeute: Der erste Wurf ergibt “Kopf”.
B bedeute: Die beiden Würfe sind verschieden.
C bedeute: Beide Würfe ergeben “Kopf”.

$A \cap B$ heißt: Der erste Wurf ergibt Kopf und der zweite “Zahl”.
$A \cap C$ heißt: Beide Würfe ergeben “Kopf” (also wie C).
$B \cap C$ ist das leere Ereignis.

Daher sind A und B unabhängig, aber nicht A und C oder B und C.

Der Satz von Bayes

Hinweis: Gilt $P(B) \neq 0$, so folgt: A und B sind unabhängig gdw. $P(A|B) = P(A)$.
Dies motiviert: Setze $P(A|B) = P(A)$, falls $P(B) = 0$.

Dann gilt (auch für $P(B) = 0$): $P(A \cap B) = P(B) \cdot P(A|B)$
sowie symmetrisch: $P(A \cap B) = P(A) \cdot P(B|A)$

$\leadsto$ Satz: (Bayes):

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

In der Anwendung oft nützlich:

$$P(B) = P(B \cap A) + P(B \cap \bar{A}) = P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A}).$$

Eine Anwendung mit (un)fairen Münzen

Wir haben eine faire Münze und eine unfaire, die immer “Kopf” liefert. Das folgende Zufallsexperiment wird ausgeführt: “Wähle eine der beiden Münzen zufällig (fair) aus und wirf sie zweimal.”

Wie groß ist die Wahrscheinlichkeit dafür, die unfaire Münze ausgewählt zu haben, falls zweimal “Kopf” erscheint ?

A entspricht: die unfaire Münze wurde ausgewählt.
B heißt: Beide Münzwürfe liefern “Kopf”.
Wir fragen also nach: $P(A|B)$.
Bekannt: $P(A) = 1/2$; $P(B|A) = 1$; $P(B) = P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A}) = 5/8$.
Bayes $\rightsquigarrow P(A|B) = 4/5$.

Zufallsvariablen

Eine *Zufallsvariable* (ZV), auch *Zufallsgröße* genannt, ist eine Funktion aus einem Ereignisraum S in die reellen Zahlen.

Beispiel: Eine übliche ZV X_1 für ein Würfelexperiment mit einem Würfel liefert die Augenzahl. Beim Würfeln mit zwei Würfeln ergeben sich mehrere natürliche Möglichkeiten von ZV, z.B.: Minimum $X_{\min}$ / Maximum $X_{\max}$ / Summe der Augenzahlen $X_{\sum}$.

Es sei S ein Ereignisraum mit Wahrscheinlichkeitsfunktion P und einer ZV X. Die Wahrscheinlichkeit, dass X den Wert r annimmt, ist:

$$P[X = r] = \sum_{e \in X^{-}(r)} P(e).$$

Zufallsvariablen und Erwartungswert

Beispiel: Bestimme $P[X_{\max} = 3]$:
$X_{\max}(e) = 3$ für $e \in \{(1,3),(2,3),(3,3),(3,2),(3,1)\}$.
$\rightsquigarrow P[X_{\max} = 3] = \frac{5}{36}$.

Satz: $P_X(r) = P[X = r]$ definiert eine Wahrscheinlichkeitsdichte auf $X(S)$.

Der *Erwartungswert* von X ist gegeben durch:

$$E[X] = \sum_{r \in X(S)} r \cdot P[X = r].$$

Beispiel: $E[X_1] = 3,5$.
Beispiel: $E[X_{\max}] = \frac{1*1+2*3+3*5+4*7+5*9+6*11}{36} = \frac{1+6+15+28+45+66}{36} = \frac{40}{9} \approx 4,44$.

Ein längeres Beispiel: Multiple Choice

Bei einer Prüfung mit “Multiple-Choice-Fragen” werden drei Fragen gestellt, wobei für jede der drei Fragen zwei Antworten zur Auswahl vorliegen, von denen jeweils genau eine richtig ist. Die Antworten werden von einem nicht vorbereiteten Prüfling rein zufällig und unabhängig voneinander angekreuzt (Gleichverteilung). Sei Z die Zufallsvariable, welche die Anzahl der richtigen Antworten angibt.

Wie viele richtige Antworten liefert ein Prüfling “im Mittel”, wenn er “rein zufällig” seine Antworten wählt ?

Ein längeres Beispiel: Multiple Choice (Forts.)

Wie müssen wir den Wahrscheinlichkeitsraum wählen ?

$S = \{r, f\}^3$, wobei r für “richtig” und f für “falsch” stehe.

Ohne Vorkenntnisse gilt: $P(\{x\}) = 1/8$ für $x \in \{r, f\}^3$.

$Z(S) = \{0, 1, 2, 3\}$.

$P[Z = 0] = P((f, f, f)) = 1/8$

$P[Z = 1] = P(\{(r, f, f), (f, r, f), (f, f, r)\}) = 3/8$

$P[Z = 2] = P(\{(r, r, f), (f, r, r), (r, f, r)\}) = 3/8$

$P[Z = 3] = P((r, r, r)) = 1/8$

$E[Z] = \sum_{r\in[3]} rP[Z = r] = \frac{0*1+1*3+2*3+3*1}{8} = 1,5$

Erwartungswert Rechenregeln

Satz: Der Erwartungswert ist ein *lineares Funktional*.

Das heißt: $E[aX + bY] = aE[X] + bE[Y]$ für Zahlen a, b und ZV X, Y.

Beweis: elementar

Satz: (Ungleichung von Markoff)

Sei $c > 0$ und X eine ZV mit nichtnegativen Werten. $\rightsquigarrow$

$$P[X \geq c] \leq \frac{E[X]}{c}.$$

Beweis: $P[X \geq c] = \sum_{r \geq c} P[X = r] \leq \sum_{r \geq c} \frac{r}{c} \cdot P[X = r] \leq \sum_{r \geq 0} \frac{r}{c} \cdot P[X = r] = \frac{E[X]}{c}$.