Mitschrift der Vorlesung Diskrete Strukturen und Wahrscheinlichkeitsrechnung

14.4.2010 - 14.7.2010

Inhaltsverzeichnis

Kapitel 1

Mathematische Modellierung des Zufalls

1.1 Zufallsexperimente

1.1.1 Das Würfeln

Vorlesung vom 14.4.2010
y $\Omega = \{1, ..., n\}, n \in \mathbb{N}$ Wir möchten zufällig genau eine Zahl aus Ω ziehen. Eine Möglichkeit: n-seitiger Würfel
Ansatz: $Pr[i] = \frac{1}{n} \forall i \in \Omega$
$Pr \mathrel{\widehat{=}}$ "$Probability$"
Sei $A = \{a_1, ..., a_k\} \subset \Omega$, dann ist $Pr[A] = \frac{|A|}{n} = \frac{k}{n}$ die Wahrscheinlichkeit, dass $a_1, ..., a_{k-1}$ oder a_k ausgewählt werden.
A nennt man *Ereignis* — Ereignis
Wenn alle $Pr[i]$ gleich sind so spricht man von einer *Gleichverteilung* . — Gleichverteilung
Bei Spielen: n=6 Es herrscht Unabhängigkeit der Würfe, d.h. Ergebnisse beeinflussen sich nicht.

1.1.2 Das Lottospiel

Es werden 6 Zahlen aus 49 gezogen, sagen wir $a_1, ..a_6$. Wir nehmen an, dass wir diese schon geordnet haben: $a_1 < ... < a_6$. Eine Ziehung ist ein Vektor $(a_1, ..., a_6)$ mit $a_1 < ... < a_6$. Ergebnisse sind diese Vektoren. Man fasst die Ergebnisse zu einem *Grundraum* zusammen, den wir üblicherweise Ω nennen. — Grundraum

$$\Omega = \{\{a_1, ..., a_6\} \mid a_i \in \{1, ..., 49\} \forall i = 1...6\}$$

$$|\Omega| = \binom{49}{6} = 13983816$$

Wie hoch ist die Wahrscheinlichkeit, dass ein Tipp 6 Richtige hat? Allgemeiner: k Richtige? Welche $\{a_1, ..., a_6\} \in \Omega$ haben k Stellen gemeinsam mit dem Tipp? Die bezeichnen wir als *günstige Ereignisse* . — günstige Ereignisse

$$A = \{\{a_1, ..., a_6\} \in \Omega \mid \left|\{a_1, ..., a_6\} \bigcup \{b_1, ..., b_6\}\right| = k\}$$

Wenn ein Element aus A gezogen wird haben wir k Richtige.

$$Pr[k\text{Richtige}] = \frac{|A|}{|\Omega|} = \frac{\binom{6}{k} \cdot \binom{43}{6-k}}{\binom{49}{6}}$$

$$\Rightarrow Pr[6\text{Richtige}] = \frac{\binom{6}{6}\binom{43}{0}}{\binom{49}{6}} = \binom{1}{49}$$

Vorlesung vom 19.4.2010

1.1.3 Das Geburtstagsproblem

Wie hoch ist die Wahrscheinlichkeit, dass zwei Personen in einer Gruppe von m Personen am gleichen Tag Geburtstag haben? z.B.: $m = 50$

1.1.4 Das Ziegenproblem

Bei einem Fernsehquiz gibt es drei Türen, wobei hinter zwei Türen eine Ziege steht und hinter der dritten Tür ein Hauptgewinn. Der Showmaster fragt den Kandidaten, welche Tür er öffnen soll. Nachdem der Kandidat einen Tipp abgegeben hat, bietet der Showmaster seine Hilfe an: Er schlägt vor eine Tür zu öffnen und danach soll der Kandidat die Möglichkeit haben seinen Tipp zu verändern.Würden Sie sein Angebot annehmen?

> In diesem Fall ist es nicht so einfach, da die günstigen Ereignisse nicht unabhängig sind.

1.2 Algebra und Maße

Die axiomatische Fundierung der Wahrscheinlichkeitsrechnung wurder von Kolmogoroff Anfang des 20. Jahrhunderts gegeben. Erst dadurch wurde Wahrscheinlichkeitsrechnung eine wissenschaftliche und quantitative Disziplin.

1.2.1 Die Axiome von Komolgoroff

Ereignisse

- Ergebnisse eines Zufallsexperimentes fassen wir in einer Menge Ω zusammen.
- *Ereignisse* sind gewisse Teilmengen von Ω.
- Wahrscheinlichkeiten werden auf Ereignisse bezogen und sind Zahlen zwischen 0 und 1. Eine Zahl $p \in [0, 1]$,z.B. $p = \frac{3}{4}$ besagt dass das entsprechende Ereigniss in $\frac{3}{4}$ der Zufallsexperimente auftrit, oder die Wahrscheinlichkeit ist 75 Prozent.

  $p = 0$: Das Ereignis tritt nicht auf.

  $p = 1$: Das Ereignis tritt stets auf.

Beispiel: Wir werfen eine faire Münze. Kopf und Zahl kommen mit der gleichen Wahrscheinlichkeit.

$$\text{Menge } \Omega = \{\text{Kopf}, \text{Zahl}\} = \{0, 1\}$$

$$\text{Ereignisse} = \{\{0\}, \{1\}, \emptyset, \{0, 1\}\} = \mathcal{P}(\Omega) = \text{Potenzmenge von } \Omega$$

$$P(\{0\}) = \frac{1}{2} = P(\{1\}), P(\emptyset) = 0, P(\{0, 1\}) = 1$$

1.2.2 Erweiterung der Kolmogoroff-Axiome

Wie muss die Menge der Ereignisse aussehen?

- A Ereignis mit Wahrscheinlichkeit p. Wir möchten das Gegenereignis ("A tritt nicht auf" $= A^c$) ebenfalls als Ereignis ansehen mit der Wahrscheinlichkeit $1-p$.
- Ereignisse A,B mit den Wahrscheinlichkeiten p und q . Was ist das Ereignis "A oder B" bzw. ?
  "A oder B" $= A \cup B$
  "A und B" $= A \cap B$
  Mithin sollten auch diese Ereignisse in der Ereignismenge drin sein.
- Was ist $P(A \cup B)$?
  Situation: $A \cap B = \emptyset$ dann ist $P(A \cup B) = P(A) + P(B)$

1.2.3 Sigma-Algebra

Definition 1.1. *Es sei Ω eine Menge und $\Sigma \subseteq \mathcal{P}(\Omega)$. Σ heißt σ-Algebraüber Ω fallse folgende Eigenschaften gelten:*

> Das σ steht für die Abzählbarkeit der Vereinigung. Für bestimmte Modellierungen ist die Potenzmenge zu groß, daher beschränkt man sich auf eine Teilmenge.

- $\Omega, \emptyset \in \Sigma$
- $A \in \Sigma \Rightarrow A^c \in \Sigma$
- Für eine Folge $A_1, A_2, ... \in \Sigma$ gilt:

$$\bigcup_{i=1}^{\infty} A_i \in \Sigma$$

(Σ-Abgeschlossenheit)

Ereignisse über einer Menge Ω werden in der Wahrscheinlichkeitstheorie durch eine geeignete σ-Algebra beschrieben wird.

Was ist mit $A \cap B$ falls $A, B \in \Sigma$? Ist $A \cap B \in \Sigma$?

$$A \cap B = (A^c \cup B^c)^c$$

$$\Rightarrow A^c, B^c \in \Sigma \Rightarrow A^c \cup B^c \in \Sigma \Rightarrow (A^c \cup B^c)^c \in \Sigma$$

$\{\emptyset, \Omega\}$ ist die kleinste und $\mathcal{P}(\Omega)$ ist die größte σ-Algebra über Ω. Vorlesung vom 21.4.2010

Beispiel Sei $\mathcal{F}$ die Menge der abgeschlossenen Intervalle der Form $[a, b] \in \mathbb{R}^d; a, b \in \mathbb{R}^d$

$$[a, b] = [a_1, b_1] \times ... \times [a_d, b_d], a = (a_1, ..., a_d), b = (b_1, ...b_d)$$

Für d=1: $[a, b]$
Für d=2:
Für d=3:

$\mathcal{B}^d$ sei die kleinste σ-Algebra über $\mathbb{R}^d$, die $\mathcal{F}$ enthält. Unter kleinste verstehen wir, dass jede σ-Umgebung Σ mit $\Sigma \supset \mathcal{F}$ und $\Sigma \subseteq \mathcal{B}^d$ $\Sigma = \mathcal{B}^d$ erfüllt.

Warum existiert $\mathcal{B}^d$?

> Dann hat es auch keinen weiteren Grund, dass ich diese Vorlesung halte.

$\mathcal{B}^d$ existiert, da mit $\mathcal{P}(\mathbb{R}^d) \supseteq \mathcal{F}$ eine σ-Algebra existiert. Somit existiert auch eine minimale σ-Algebra.

Sei $M = \{\Sigma | \Sigma ist \sigma - \text{Algebra über} \mathbb{R}^d, \mathcal{F} \subseteq \Sigma\}$. Suche die kleinste σ-Algebra aus M. Das geht wenn $M \neq \emptyset$. Da $\mathcal{P}(\mathbb{R}^d) \in M$ ist dies gegeben. Definition 1.2 $\mathcal{B}^d$ nennt man die *Borelsche σ-Algebra* über $\mathbb{R}^d$. Wie sehen die Mengen in $\mathcal{B}^d$ aus? Borelsche σ-Algebra

Für eine Menge $X \subseteq \mathcal{P}(\Omega)$ verstehen wir unter $\sigma(X)$ die kleinster σ-Algebra, die X enthält. Sie $\mathcal{O}$ die Menge der offenen Intervall in $\mathbb{R}^d$. Betrachte $\sigma(\mathcal{O})$.

Satz 1.2.

$$X \subseteq \mathcal{P}(\Omega); Y \subseteq \mathcal{P}(\Omega), X \subseteq Y \Rightarrow$$
$$\sigma(X) \subseteq \sigma(Y)$$

Beweis. Nehme $Z \in \sigma(X) \Rightarrow Z \in \sigma(Y)$.

Ausgangssituation: Ω ist gegeben. X ist Menge von Teilmengen von Ω d.h. $X \subseteq \mathcal{P}(\Omega)$. Dann ist $\sigma(X)$ die kleinste σ-Algebra über Ω, die X enthält, d.h. $X \in \sigma(X)$. □

Vorlesung vom 26.4.2010

$\sigma(\mathcal{O})$ ist die kleinste σ-Algebra, die die Menge $\mathcal{O}$ der offenen Intervalle in $\mathbb{R}^d$ enthält.

Satz 1.3. $\mathcal{B}^d = \sigma(\mathcal{O})$

Beweis. Zur Vereinfachung sei $d = 1$.

Zunächst $\mathcal{B}^d \subseteq \sigma(\mathcal{O})$: Wir wissen, dass $\mathcal{B}^d = \sigma(\mathcal{F})$ ist. $\mathcal{F}$ ist Erzeuger von $\mathcal{B}^d$.

Bei der σ-Algebra reicht es zu Zeigen, dass der Erzeuger in der Menge enthalten ist.

Falls $\mathcal{F} \subseteq^{(*)} \sigma(\mathcal{O})$ dann gilt $\sigma(\mathcal{F}) \subseteq \sigma(\sigma(\mathcal{O})) = \sigma((O))$.

$\sigma(\sigma(M)) = \sigma(M)$ da $\sigma(M)$ bereits eine σ-Algebra ist.

Das heißt: es reicht $(*)$ zu zeigen. Sei $[a, b] in \mathcal{F}$ ein beliebiges geschlossenes Interval. Wie können wir $[a, b]$ mit Hilfe von offenen Intervallen oder anderen Elementen aus $\sigma(\mathcal{O})$ erzeugen?

Vorschlag:

$$[a, b] = (]-\infty, a[\cup]b, \infty[)^c$$

Schon nicht schlecht, aber die Frage ist, ob $]-\infty, a[\in \sigma(\mathcal{O})$ ist. Denn bisher haben wir die offenen Intervalle die Form $]x, y[; x, y \in \mathbb{R}$ und $\infty \notin \mathbb{R}$.

Vorschlag2:

$$\sigma(\mathcal{O}) \ni \bigcup_{n=1}^{\infty}]b, b+n[=]b, \infty[$$

Da dies eine abzählbare Vereinigung ist von Mengen aus $\sigma(\mathcal{O})$ ist auch die Vereinigung wieder aus $\sigma(\mathcal{O})$. Somit ist der Vorschlag brauchbar.

Folglich ist jedes Element auf $\mathcal{F}$ auch in $\sigma(\mathcal{O})$.

Nun $\sigma(\mathcal{O}) \subseteq \mathcal{B}^d$: Es reicht zu zeigen, dass $\mathcal{O} \subseteq \mathcal{B}^d$ (wie eben). Sei $]a, b[\in \mathcal{O}$ beliebig und $a, b \in \mathbb{R}$. Zu zeigen $]a, b[\in \mathcal{B}^d$.

Vorschlag:

$$\mathcal{B}^d \ni \bigcup_{n=1}^{\infty} [a + \frac{1}{n}, b - \frac{1}{n}] =]a, b[$$

Wieder handelt es sich um eine abzählbare Vereinigung von Elementen aus $\mathcal{B}^d$. Die Grenzen a und b sind natürlich nicht in der Vereinigung, aber man kann für jede Zahl c zwischen a und b ein n finden sodass c in der Vereinigung ist.

Wir wissen:

$$\mathcal{B}^d \subseteq \mathcal{P}(\mathbb{R}^d)$$

Frage: Gibt es Mengen $A \subseteq \mathbb{R}^d$ mit $A \neq \mathcal{B}^d$? bzw. $\mathcal{B}^d = \mathcal{P}(\mathbb{R}^d)$?

Vorschlag:

$$\mathbb{R}^d \supset A = \bigcup_{a \in A} [a] \in \mathcal{B}^d$$

Die Vereinigung ist nur in $\mathcal{B}^d$ falls A abzählbar ist.

Es gibt Teilmengen des $\mathbb{R}^d$ die keine Borelmengen sind!

$$\mathcal{B}^d \subsetneq \mathcal{P}(\mathbb{R}^d)$$

Diese nennt man *Cantor-Mengen* . □ Cantor-Mengen

Um die Wahrscheinlichkeit eines Ereignisses zu definieren möchten wir - wenn Ereignisse durch Elemente einer σ-Algebra beschreiben werde, also Teilmengen einer Grundmenge Ω sind - diesen Teilmengen eine Maßzahl zuordnen, d.h. sie messen.

1.Fall

$$|\Omega| < \infty, A \subseteq \Omega$$

$|A|$ ist ein Maß. Bei Gleichverteilung ergibt $\frac{|A|}{|\Omega|}$.

2.Fall

$$|\Omega| = \infty, A \subseteq \Omega$$

Falls $|A| = \infty$ so ist ein Maß der Form $|A|$ zu undifferenziert.

1.3 Wahrscheinlichkeitsraum

Ist Σ eine σ-Algebra über Ω, so nennt man (Ω, Σ) einen *Meßraum* . Meßraum

Definition 1.4. *Sei (Ω, Σ) ein Meßraum.*

1. *Eine Funktion $\mu : \Sigma \to [0, \infty]$ heißt Maß falls folgendes gilt:* Maß
   (a) $\mu(\emptyset) = 0$
   (b) *für jede Folge von paarweise disjunkten Mengen $A_1, A_2, ... \in \Sigma$ gilt:*
   $$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$
2. *Ein Maß $P : \Sigma \to [0, 1]$ mit $P(\Omega) = 1$ heißt Wahrscheinlichkeitsmaß .* Wahrscheinlichkeitsmaß
3. *Das Tripel (Ω, Σ, P) nennt man Wahrscheinlichkeitsraum* Wahrscheinlichkeitsraum

Zusammenhang zur "experimentellen" oder "intuitiven" Wahrscheinlichkeit:

1.3.1 Komolgoroff-Axiome

Ein Zufallsexperiment wird im mathematischen Sinne durch einen Wahrscheinlichkeitsraum (Ω, Σ, P) beschrieben. Dabei gilt:

1. Ω umfasst alle Ergebnisse $\omega \in \Omega$ des Zufallsexperimentes.
2. Die Ereignisse sind die Mengen aus Σ.
3. Für eine Menge $A \in \Sigma$ ist $P(A)$ die Wahrscheinlichkeit des Ereignisses A.

Vorlesung vom 28.4.2010

1.3.2 Eigenschaften eines Wahrscheinlichkeitsmaßes

Proposition 1.5. *Sei (Ω, Σ, P) ein Wahrscheinlichkeitsraum. Es seien $A, B, A_1, A_2, ... \in \Sigma$. Dann gilt:*

1. $P(A^c) = 1 - P(A)$
2. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$
3. *Aus* $A \subseteq B$ *folgt* $P(B \setminus A) = P(B) - P(A)$
4. *Aus* $A \subseteq B$ *folgt* $P(A) \leq P(B)$
5. *Für* $A_1, ...A_n \in \Sigma$ *gilt:*
$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1} P(A_i)$$
(Union-Bound)

Beweis. 1. Wir wissen schon, dass $P(\Omega) = 1$ gilt. Folglich ist zu zeigen, dass $P(A^c) + P(A) = 1 = P(\Omega)$ gilt.

$$P(\Omega) = P(A \cup A^c) = P(A) + P(A^c) \tag{1.1}$$

Der letzte Schritt ist möglich, da A und A^c paarweise disjunkt sind.

2.
$$A \cup B = A \cup (B \setminus A)$$
Daraus folgt aufgrund der σ-Additivität:
$$P(A \cup B) = P(A) + P(B \setminus A)$$
Weiterhin gilt:
$$B = B \setminus A \cup (A \cap B) \Rightarrow P(B) = P(B \setminus A) + P(A \cap B)$$
$$\Rightarrow P(B \setminus A) = P(B) - P(A \cap B)$$
$$\Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

3.
$$B = A \dot{\cup} (B \setminus A)$$
$$\Rightarrow P(B) = P(A) + P(B \setminus A)$$
$$\Rightarrow P(B \setminus A) = P(B) - P(A)$$

4.
$$B = A \dot{\cup} (B \setminus A)$$
$$\Rightarrow P(B) = P(A) + P(B \setminus A) \geq P(A)$$

5. Für $n = 2$ ist die Behauptung schon gezeigt.
Versuch Darstellung als:
$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{l} B_i$$
mit B_i paarweise disjunkt.
$$B_1 = A_1$$
$$B_2 = A_2 \setminus A_1$$
$$B_3 = A_3 \setminus (A_1 \cup A_2)$$
$$B_n = A_n \setminus (A_1 \cup ... \cup A_{n-1})$$

Die B_i sind immer paarweise disjunkt und $B_i \subseteq A_i; i \in 1, ..., n$ so wie

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i$$

$$\Rightarrow P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n} B_i\right) = \sum_{i=1}^{n} P(B_i)$$

Nun wissen wir aber nach Konstruktion, dass jedes $B_i \subseteq A_i$ und somit $P(B_i) \leq P(A_i)$. Daraus folgt:

$$\sum_{i=1}^{n} P(B_i) \leq \sum_{i=1}^{n} P(A_i)$$

□

1.3.3 Der Laplace-Raum

Sei Ω endllich. Sei $\Sigma = \mathcal{P}(\Omega)$ und $P : \Sigma \to [0,1], P(A) := \frac{|A|}{|\Omega|}$ für alle $A \in \Omega$. (Ω, Σ) ist ein Meßraum, weil $\Sigma = \mathcal{P}(\Omega)$ eine σ-Algebra. P ist ein Wahrscheinlichkeitsmaß, da:

1. $P(A) \in [0,1] \forall A \subseteq \Omega$
2. $P(\emptyset) = \frac{|A|}{|\Omega|} = \frac{0}{|\Omega|} = 0$
3. $P(\Omega) = \frac{|\Omega|}{|\Omega|} = \frac{|\Omega|}{|\Omega|} = 1$
4. $A_1, A_2, ...A_n \subseteq \Omega$ paarweise disjunkt.

$$P(A_1 \cup ... \cup A_n) = \frac{|A_1 \cup ... \cup A_n|}{|\Omega|} = \frac{|A_1| + ...|A_n|}{|\Omega|} =$$

$$\frac{|A_1|}{|\Omega|} + ... + \frac{|A_n|}{|\Omega|} = P(A_1) + ... + P(A_n) = \sum_{i=1}^{n} P(A_i)$$

Da Ω endlich ist, gibt es auch nur endlich viele verschieden Teilmengen.

Mit "Übergang" aus $n \to \infty$ ist P σ-Additiv, also ein Wahrscheinlichkeitsmaß.

Also ist (Ω, Σ, P) ein Wahrscheinlichkeitsraum.

1. Gibt es ein Wahrscheinlichkeitsmaß P auf $\mathbb{N}$ mit $\mathcal{P}(\mathbb{N})$ als σ-Algebra, so dass $P(\{i\}) > 0 \forall i \in \mathbb{N}$?
2. Gibt es ein Wahrscheinlichkeitsmaß P auf $\mathbb{R}$ mit geeigneter σ-Algebra, die alle einelementigen Mengen $\{x\}; x \in \mathbb{R}$ enthält und $P(\{x\}) > 0$?

Zu 1: $P(i) = \frac{1}{2^i}; i = 1, 2, ...$ (geometrische Folge). Dann ist

$$P(\mathbb{N}) = P\left(\bigcup_{i=1}^{\infty} \{i\}\right) = \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{1 - \frac{1}{2}} - 1 = 1$$

Wie sieht dann $P(A)$ für $A \subseteq \mathbb{N}$ aus?

$$P(A) = \sum_{i \in A} P(\{i\})$$

$$\Rightarrow P(\emptyset) = 0 \text{ da die Leere Summe 0 ist.}$$

Wir überzeugen uns, dass ein P ein Wahrscheinlichkeitsmaß ist:

1. $P(\Omega) = P(\mathbb{N}) = 1$

2. Sei $A_1, A_2, ...A_n \subseteq \Omega$ paarweise disjunkt dann gilt:

$$\begin{aligned} P(A_1 \cup ... \cup A_n) &= \sum_{i \in A_1 \cup ... \cup A_n} P(\{i\}) \\ &= \sum_{i \in A_1} P(\{i\}) + ... + \sum_{i \in A_n} P(\{i\}) \\ &= P(A_1) + ... + P(A_n) \end{aligned}$$

Folglich ist P ein Wahrscheinlichkeitsmaß.

1.3.4 Lebesquemaß

Lebesguemaß

Vorlesung vom 3.5.2010
Als wichtiges Beispiel betrachten wir das *Lebesguemaß* , das einerseits unser natürliches Verständnis des Messens" von regulären Mengeninhalten in $\mathbb{R}^d$ widerspigelt, aber gleichzeitig durch die Nullmengeneigenschaft unser Intuition zuwiderläuft. Diese Nullmengeneigenschaft ist aber charackteristisch und grundlegend zum Verständnis der Wahrscheinlichkeitstheorie, auch im diskreten Fall. Sei $[a, b] \in \mathbb{R}^d$ ein Intervall, für $a = (a_1, ..., a_d)$ und $b = (b_1, ..., b_d)$ sei

$$[a, b,] = [a_1, b_1] \times ... \times [a_d, b_d]$$

Das Volumen von $[a, b]$ ist offenbar $(b_1 - a_1) \cdot ... \cdot (b_d - a_d)$. Wir schreiben $Vol([a, b])$ für das Volumen von $[a, b]$.

Eine grundl Frage der Maptheorie ist die Frage: Gibt es ein Maß für Borelmengen $\mathcal{B}^d$ in $\mathbb{R}^d$, das für $[a, b]$ in $\mathbb{R}^d$ genau das Volumen angibt? Diese Frage kann positiv beantwortet werden und ist ein grundlegendes Resultat der Maßtheorie. Wir notieren (ohne Beweiß):

Satz 1.6. *Für jedes $d \in \mathbb{N}$ gibt es ein Maß $\lambda^d, \lambda^d : \mathcal{B}^d \to [0, \infty[$ mit der Eigenschaft $\lambda^d([a, b]) = Vol([a, b])$ für jedes Intervall $[a, b] \in \mathbb{R}^d$.*

Definition 1.7. *λ^d heißt d-dimensionales Lebesguemaß. Wir nenne $\lambda = \lambda^d$ Lebesguemaß.*

Definition 1.8. *Sie (Ω, Σ, μ) ein Maßraum. Eine Menge $A \in \Sigma$ heißt μ-Nullmenge, wenn $\mu(A) = 0$.*

Für das Lebesguemaß haben wir:

Satz 1.9. *Jede abzählbare Menge in $\mathbb{R}^d$ ist eine Borelmenge und sogar eine Nullmenge bezüglich λ.*

Beweis. Sei $A \subseteq \mathbb{R}^d$ und $A = \{a_1, a_2, ...\}$. Wegen $A = \bigcup_{i=1}^{\infty} \{a_i\}$ ist A aus $\mathcal{B}^d$ Wegen $\lambda^d(\{a_i\}) = 0$ folgt

$$\lambda^d(A) = \sum_{i=1}^{\infty} \lambda^d(\{a_i\}) = 0$$

$\lambda^d(\{a_i\}) = 0$ gilt, weil eine einzelne Zahl kein Volumen hat.

□

Wenn man beachtet, dass $\mathcal{B}^d_{[0,1]}$ eine σ-Algebra über $[0, 1]$ ist, so erhält man:

Proposition 1.10. *$([0, 1], \mathcal{B}^d_{[0,1]}, \lambda^d)$ ist ein Wahrscheinlichkeitsraum.*

Nillmengen spielen in der Wahrscheinlichkeitstheorie eine grundlegende Rolle, denn ein Ereignis mit der Wahrscheinlichkeit 0 tritt nicht auf.

1.4 Bedingte Wahrscheinlichkeiten

Ereignisse können sich in vielfacher Weise beeinflussen. Es ist daher von grundlegender Bedeutung, der Einfluss eines Ereignisses auf das auftreten eines anderen Ereignisses mathematisch zu formulieren.

Definition 1.11 (bedingte Wahrscheinlichkeit)**.** *Sei (Ω, Σ, P) ein Wahrscheinlichkeitsraum und $A, B \in \Sigma$ mit $P(B) > 0$. Der Ausdruck*

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

ist die bedingte Wahrscheinlichkeit von A gegeben B.

bedingte Wahrscheinlichkeit

Ein besonderer Fall liegt vor, wenn $P(A \cap B) = P(A) \cdot P(B)$ ist. Dann gilt:

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

. d.h. B hat keinen Einfluss auf A. Umgangssprachlich sagen wir, dass A von B unabhängig ist. Der Begriff der *Unabhängigkeit* ist ein grundlegendes Konzept, das wir später definieren werde. Mithilfe der bedingten Wahrscheinlichkeit können wir $P(A \cap B)$ für beliebige Ereignisse $A, B \in \Sigma$ (mit $P(B) > 0$) berechnen. Gemäß Def(1.12) gilt:

Unabhängigkeit

$$P(A \cap B) = P(A|B) \cdot P(B) 1.4$$

Falls A und B unabhängig sind, vereinfacht sich diese Gleichung wieder zu $P(A \cap B) = P(A) \cdot P(B)$, das heißt sie enthält den Fall ünabhängige Ereignisseäls wichtigen Spezialfall. Gleichung 1.4 lässt sich für mehr als zwei Ereignisse verallgemeinern (siehe nächster Satz).

Wenn $P(A \cap B) = P(A) \cdot P(B)$, dann $P(A|B) = P(A)$. Dann ist A ünabhängig" von B.

Wenn $A \cap B = \emptyset$, dann ist $P(A \cap B) = P(\emptyset) = 0$.

Proposition 1.12 (Multiplikationssatz)**.** *Seien $A_1, ..., A_n$ Ereignisse mit $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$. Dann gilt:*

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot ... \cdot P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right)$$

Beweis. Wegen $P(A_1) => P(A_1 \cap A_2) => .. => P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$ sind alle bedingten Wahrscheinlichkeiten wohldefiniert. Es gilt

$$\begin{aligned} P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot ... \cdot P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right) &= \frac{P(A_1)}{1} \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot ... \cdot \frac{P(A_1 \cap ... \cap A_n)}{P(A_1 \cap ... \cap A_{n-1})} \\ &= P(A_1 \cap ... \cap A_n) \end{aligned}$$

□

Satz 1.13 (Satz der totalen Wahrscheinlichkeit)**.** *Die Ereignisse $B_1, ..., B_n$ sind paarweise disjunkt mit $P(B_i) > 0$ für $i = 1, ..., n$ und es gelte $\Omega = B_1 \cup ... \cup B_n$. Dann folgt:*

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i) \forall A \in \Sigma$$

Motivation: man möchte $P(A)$ berechnen. A liegt jedoch in einer Weise vor, dass wir $P(A)$ nicht direkt berechnen können. Falls alle $P(A \cap B_i)$ leichter zu berechnen wären als $P(A)$, so könnte man mit Satz 1.15 $P(A)$ leicht berechnen.

Vorlesung vom 5.5.2010

Beweis.

$$
\begin{aligned}
\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i) &= \sum_{i=1}^{n} P(A \cap B_i) \\
&= P\left(\bigcup_{i=1}^{n} (A \cap B_i)\right) \\
&= P\left(A \cap \left(\bigcup_{i=1}^{n} B_i\right)\right) \\
&= P(A \cap \Omega) \\
&= P(A)
\end{aligned}
$$

□

1.4.1 Lösung des Ziegenproblems

Betrachte geeignete Ereignisse:

$$
\begin{aligned}
A &:= \text{Kandidatin hat bei erster Wahl das Auto gewählt} \\
G &:= \text{Kandidatin gewinnt nach dem Wechsel der Tür das Auto}
\end{aligned}
$$

Uns interessiert $P(G)$ im Vergleich zu $P(A)$. Falls der Tip ohne Hilfe abgegeben wurde, so ist $P(A) = \frac{1}{3}$.

> Nun haben wir eine Situation in der wir $P(G)$ nicht direkt berechnen können. Aber wir können folgendes mit Satz 1.15 machen:

$$
P(G) = P(G|A) \cdot P(A) + P(G|A^c) \cdot P(A^c)
$$

Nun ist $P(G|A) = 0$ und somit $P(G|A) \cdot P(A) = 0$. Weiterhin ist $P(G|A^c) = 1$ und $P(A^c) = \frac{2}{3}$ und somit $P(G|A^c) \cdot P(A^c) = \frac{2}{3}$. Folglich ist $P(G) = \frac{2}{3} > P(A)$.

1.4.2 Lösung des Geburtstagsproblems

Dieses Problem gehen wir mit der Proposition 1.14 an.

Modelliere das Problem mit Hilfe von Urnen. Wir haben n Urnen (z.B. $n = 365$). Wir werfen jeden der m Bälle (z.B. Personen) zufällig unter Gleichverteilung in die Urnen. Wir setzen vorraus $m \leq n$. Die Wahrscheinlichkeit, dass ein Ball in die i-te Urne fällt ist $\frac{1}{n}$.

> Wir nehmen an, dass die Wahrscheinlichkeit an einem bestimmten Tag Geburtstag zu haben nicht vom Tag abhängt.

A ist das Ereignis, dass alle Bälle in unterschiedlichen Urnen landen. A^c ist das gewünschte Ereignis, nämlich dass in mindestens einer Urne mehr als ein Ball landet. Wir berechnen $P(A^c)$ über $P(A^c) = 1 - P(A)$ und einer Abschätzung des schlechten Ereignisses $P(A) \leq \alpha$, was zu $P(A^c) = 1 - P(A) \geq 1 - \alpha$ führt. Sei:

$$
A_i := \text{Ball } i \text{ landet in einer noch leeren Urne}
$$

dann ist:

$$
P(A) = P\left(\bigcap_{i=1}^{m} A_i\right)
$$

Dann können wir den multiplikationssatz anwenden.

$$P\left(\bigcap_{i=1}^{m} A_i\right) = P(A_i) \cdot P(A_2|A_1) \cdot ... \cdot P(A_m| \bigcap_{i=1}^{n-1} A_i)$$

Schätze jeden Faktor nach oben ab:

Vorlesung vom 10.5.2010

1.5 Stochastische Unabhängigkeit

Definition 1.14. *Seien $A_1, ..., A_n$ Ereignisse.*

1. *Für ein festes $k \in \mathbb{N}, 2 \leq k \leq n$ heißen $A_1, ..., A_n$ k-weise unabhängig, fallls für jede Auswahl $A_{i_1}, ..., A_{i_k}$ von k Ereignissen aus den $A_1, ..., A_n$ gilt:*

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

2. *Die Ereignisse $A_1, ..A_n$ heißen unabhängig (stochastisch unabhängig), falls für jede Teilmenge $J \subseteq \{1, ...n\}, J \neq \emptyset$ gilt:* stochastisch unabhängig

$$P\left(\bigcap_{j\in J} A_{i_j}\right) = \prod_{j\in J} P(A_{i_j})$$

Wenn $A_1, ..., A_n$ unabhängig ist, dann auch k-weise unabhängig für jedes $k \in \{2, ..., n\}$.

Für $n = 2$ spricht man von paarweiser Unabhängigkeit. Das bedeutet, dass aus $A_1, A_2, ...A_n$ je zwei unabhängig sind. Sind dann alle voneinander unabhängig?

Beispiel Sei $\Omega = \{1, 2, 3, 4\}$. Betrachte das Experiment, aus Ω eine Zahl zufällig mit Wahrscheinlichkeit $\frac{1}{4}$ auszuwählen, unter Gleichverteilung.

- A sei das Ereignis: 1 oder 2 wird gewählt
- B sei das Ereignis: 1 oder 3 wird gewählt
- C sei das Ereignis: 1 oder 4 wird gewählt

2-weise Unabhängigkeit

$$\begin{aligned} P(A)P(B) &= \left(\frac{1}{4} + \frac{1}{4}\right)\left(\frac{1}{4} + \frac{1}{4}\right) \\ &= \frac{1}{4} \\ &= P(A \cap B) \end{aligned}$$

analog zerlegt man $P(A)P(C)$ und $P(B)P(C)$. Also ist paarweise unabhängigkeit gegeben.

3-weise Unabhängigkeit

$$\begin{aligned} P(A)P(B)P(C) &= \left(\frac{1}{4}+\frac{1}{4}\right)\left(\frac{1}{4}+\frac{1}{4}\right)\left(\frac{1}{4}+\frac{1}{4}\right) \\ &= \frac{1}{8} \\ &\neq \frac{1}{4} = P(A\cap B\cap C) \end{aligned}$$

Also keine 3-weise Unabhängigkeit.

Hieraus folgt: Aus k-weiser Unabhängigkeit für ein festes k folgt **nicht** notwendigerweise volle Unabhängigkeit.

Beispiel Wir werfen einen Würfel zweimal hintereinander. Sei

- A das Ereigniss, dass beim ersten Wurf eine gerade Zahl auftritt
- B das Ereigniss, dass beim zweiten Wurf eine gerade Zahl auftritt

A und B sind nach dem intuitiven Verständnis ünabhängig". Wir zeigen, dass sie auch im stochastischen Sinne unabhängig sind. Dazu muss das Experiment durch einen geeigneten Wahrscheinlichkeitsraum modelliert werden. Wähle $\Omega_1 = \{1,2,3,4,5,6\}$ für den ersten Wurf und Ω_2 ebenso für den zweiten Wurf.

> Diese reichen nicht aus. Wir wollen ja nicht nur zwei einzelne Würfe betrachten, sondern eine Kombination aus zwei Würfen.

Dann ist $\Omega = \Omega_1 \times \Omega_2$. $(i,j) \in \Omega$ repräsentiert, dass i das Ergebnis des ersten Wurfes ist und j das Ergebnis des zweiten Wurfes ist. Weiterhin gilt:

$$\Omega = \{(i,j) : 1 \leq i,j \leq 6\}$$

Das Wahrscheinlichkeitsmaß P wählen wir als die Gleichverteilung über Ω d.h.:

$$\mathbb{P}(\{(i,j)\}) = \frac{1}{|\Omega|} = \frac{1}{36} \forall (i,j) \in \Omega$$

Wähle $\Sigma = \mathcal{P}(\Omega)$. Dann haben wir den Laplaceraum $(\Omega, \Sigma, \mathbb{P})$. Daraus folgt:

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(\{(i,j) : 1 \leq i,j \leq 6, i \text{ gerade}\}) \\ &= \frac{|A|}{|\Omega|} \\ &= \frac{18}{36} = \frac{1}{2} \end{aligned}$$

Analog dazu gilt $\mathbb{P}(B) = \frac{1}{2}$. D.h.:

$$\mathbb{P}(A)\mathbb{P}(B) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

Betrachte weiterhin:

$$\begin{aligned} A\cap B = \{(i,j) : 1 \leq i,j \leq 6, i,j \text{ gerade}\} |A\cap B| &= 9 \mathbb{P}(A\cap B) = \frac{|A\cap B|}{|\Omega|} \\ = \frac{9}{36} = \frac{1}{4} \end{aligned}$$

Also gilt:$\mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A\cap B)$, d.h. A und B sind unabhängig.

1.5.1 Informationstheoretischer Aspekt

Betrachte n Würfe einer fairen Münze. Dann gibt es unterschiedliche Möglichkeiten:

1. Die Würfe sind unabhängig voneinander
2. Je zwei Würfe sind unabhängig

   Der erste Möglichkeit ist einfach zu realisieren.

Wieviele zufällige Bits benötigt man um diese Experimente zu beschreiben?

1. Man benötigt genau n Bits.

Vorlesung vom 12.5.2010
Ein endliches Problem Π mit Instanzen $x \in \{0,1\}*$, Kodierungslänge $|x|$ = Anzahl der Komponenten von x, $|x| < \infty$. Ein Polynomzeitalgorithmus benötigt $\mathcal{O}(|x|^k)$ Schritte , $k \in \mathbb{N}$ fest abhängig von Π, um das Problem Π mit Instanz x zu lösen. Lösen bedeutet hier: Entscheiden, ob Π lösbar ist obder nicht, dabei ist Π also „Ja", „Nein" Entscheidungsproblem gegeben. z.B.: $p \in \mathbb{N}$, Frage: Ist p eine Primzahl?

Zahlen $r \in \mathbb{N}$ bzw $r \in \mathbb{Q}$ werden binär kodiert : $|r| = \lceil log_2 r \rceil$ Für $r \in \mathbb{Q}, r = \frac{a}{b}, |r| := |a| + |b|$.

Beispiel Problem: Sei $G = (V, E)$ ein Graph mit Kantengewichten $w : E \to \mathbb{N}$. Frage: Besitzt G einen aufspannenden Baum mit Gewicht höchstens $W \in \mathbb{N}$? Sei Π obiges Problem. Wie ist die Kodierungslänge diese Problems. x sei Instanz mit $G = (V, E)$ und w.

$$|x| = |V| + |E| + \sum_{\varphi \in E} \lceil log_2 w(\varphi) \rceil$$

Anmerkung Probleme werden eingeteilt in:

Die Klasse $\mathcal{P}$ Die Klasse der endlichen Probleme, die in Polynomialzeit lösbar sind.

Die Klasse $\mathcal{NP}$ Die Klasse der endlichen Probleme, deren Lösungen in Polynomialzeit verifizierbar sind

Schematische Darstellung von P und NP

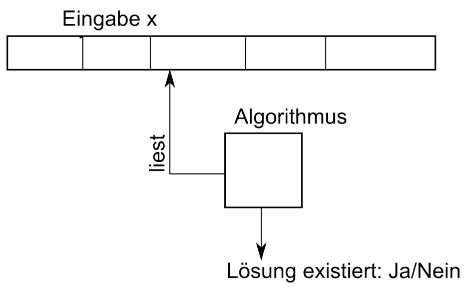

Abbildung 1.1: Darstellung eines Algorithmus aus $\mathcal{P}$

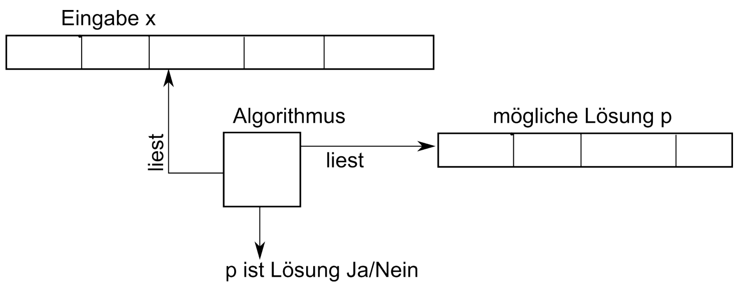

Abbildung 1.2: Darstellung eines Algorithmus aus $\mathcal{N}P$

Einbeziehung des Zufalls

Sei Π ein Problem in $\mathcal{N}P$. Das Orakel gibt uns einen Beweis/Lösung π. Nehmen wir mal an, dass π eine wirkliche Lösung ist. π habe l Stellen. Wieviele Stellen von π muss man lesen um in polynomieller Zeit in $|x|$ zu verifizieren, dass π wirklich eine Lösung ist? Im Allgemeinen müssen wir alle l Einträge von π lesen! Dabei sollte gelten: $l = \mathcal{O}(|x|^k), k$ konstant.

Wieviele Stellen von π muss ich lesen, wenn ich noch Random Bits τ, $|\tau| = r$ zur Verfügung habe?

Satz 1.15 (Arora, Motwani, Lund, Szegdi '94; PCP-Theorem)**.** *Man kann mit $log(|x|)$ Stellen von π und konstant vielen Random Bits verifizieren, ob π eine Lösung ist. (Neue Characterisierung von $\mathcal{NP}$!*

New York Times: ,,Short-Cut to long proofs“.

Anwendung auf den Spannbaum. Eingabe x mit Beweis Π seien $2^1 00$ bzw. $2^8 0$ Bit groß. Diese Eingabe passt in keinen Speicher. Nach dem Theorem reicht es 100 bzw. 80 Bits und eine konstante Menge an zufälligen Bits zu lesen!

1.6 Diskrete Wahrscheinlichkeitsverteilung

Wahrscheinlichkeitsverteilung

Seien $(\Omega, \Sigma, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Man nennt $\mathbb{P}$ auch *Wahrscheinlichkeitsverteilung* Die Begriffe Wahrscheinlichkeitsverteilung und Wahrscheinlichkeitsmaß werden synonym verwendet.

1.6.1 Die Binomialverteilung

Sei $\Omega = \{0, 1, 2, ..., n\}$. Sei $0 \leq p \leq 1$. Für $\omega \in \Omega$ setzen wir

$$p(\omega) := \binom{n}{k} \cdot p^k \cdot (1-p)^(n-k)$$

stochastische Vektor

$\vec{p}(p(0), p(1), p(2), ..., p(n))$ heißt der *stochastische Vektor* . Sei $\Sigma := \mathcal{P}(\Omega)$ und definiere ein Wahrscheinlichkeitsmaß auf Σ wie folgt: Für alle $A \in \Sigma$ sei

$$\begin{aligned} B_{(n,p)}(A) &:= \sum_{\omega \in A} p(\omega) \\ &= \sum_{\omega \in A} \binom{n}{\omega} p^\omega (1-p)^{n-\omega} \end{aligned}$$

eine Funktion $B_{(n,p)} : \Sigma \to [0, 1]$

Satz 1.16. *$B_{(n,p)}$ ist ein Wahrscheinlichkeitsmaß.*

Beweis. Es sind drei Eigenschaften zu zeigen:

1. $B_{(n,p)}(\Omega) = 1$*:*

$$\begin{aligned} B_{(n,p)}(\Omega) &= \sum_{\omega \in \Omega} p(\omega) \\ &= \sum_{\omega=0}^{n} \binom{n}{\omega} p^\omega (1-p)^{n-\omega} \qquad = (p + (1-p))^n = 1^n = 1 \end{aligned}$$

2. $0 \leq B_{(n,p)}(A) \leq 1 \forall A \subseteq \Omega$: siehe vorheriger Punkt
3. *σ-Additivität:* Für $A_1, A_2, ... \in \Sigma$ paarweise disjunkt gilt, da es höchstens $l \leq n+1$ nicht-leere Teilmengen von Ω gibt:

$$
\begin{aligned}
B_{(n,p)}\left(\bigcup_{i=1}^{\infty} A_i\right) &= B_{(n,p)}\left(\bigcup_{k=1}^{l} A_{i_k}\right) \\
A &= \bigcup_{k=1}^{l} A_{i_k} \\
B_{(n,p)}(A) &= \sum_{\omega \in A} p(\omega) \\
&= \sum_{\omega \in A_{i_1}} p(\omega) + ... + \sum_{\omega \in A_{i_l}} p(\omega) \\
&= B_{(n,p)}(A_{i_1}) + ... + B_{(n,p)}(A_{i_l}) \\
&= \sum_{k=1}^{l} B_{(n,p)}(A_{i_k}) \\
&= \sum_{i=1}^{\infty} B_{(n,p)}(A_i)
\end{aligned}
$$

□

Das Wahrscheinlichkeitsmaß $B_{(n,p)}$ nennt man *Binomialverteilung* . Binomialverteilung

Definition 1.17. *Sei Ω eine abzählbar endlichen oder unendlichen Menge. Eine Funktion $p : \Omega \to [0,1]$ heißt stochastischer Vektor falls $\sum_{\omega \in \Omega} p(\omega) = 1$.* stochastischer Vektor

Satz 1.18. *Sei p ein stochastischer Vektor über Ω. Dann ist $(\Omega, \Sigma, \mathbb{P})$ mit*

$$
\begin{aligned}
\Sigma &:= \mathcal{P}(\Omega) \\
\mathbb{P} &: \Sigma \to [0,1] \\
P(A) &:= \sum_{\omega \in A} p(\omega), A \in \Sigma
\end{aligned}
$$

ein Wahrscheinlichkeitsraum und $\mathbb{P}$ ein Wahrscheinlichkeitsmaß.

Beweis. Es sind drei Eigenschaften zu zeigen:

1. $\mathbb{P}(\Omega) = 1$: Nach Definition:

$$\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} p(\omega) \qquad = 1$$

2. $0 \leq B_{(n,p)}(A) \leq 1 \forall A \subseteq \Omega$: siehe vorheriger Punkt
3. *σ-Additivität:* Sei $A_1, A_2, ...$ eine Folge aus $\Sigma = \mathcal{P}(\Omega)$ von paarweise disjunkten Mengen. Dann gilt nach Definition von p:

$$
\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{\omega \in A_1} p(\omega) + \sum_{\omega \in A_1} p(\omega) + ... \\
&= \mathbb{P}(A_1) + \mathbb{P}(A_2) + ... \\
&= \sum_{i=1}^{\infty} P(A_i)
\end{aligned}
$$

□

1.6.2 Konstruktionsprinzip für endliche Wahrscheinlichkeitsmaße

1. Ω endlich gegeben
2. Definiere $p(\omega), \omega \in \Omega$.
3. Zeige $\sum\limits_{\omega \in \Omega} p(\omega) = 1$.

Dann erhalten wir mit Satz **??** einen endlichen Wahrscheinlichkeitsraum $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Dies gilt auch für abzählbar unendliche Ω.

1.6.3 Bernoulli-Verteilung

$\Omega = \{0, 1\}$, $\Sigma = \mathcal{P}(\Omega)$ und $p(\omega) := p^{\omega}(1-p)^{1-\omega}, \omega \in \Omega$ mit

Index