Wahrscheinlichkeitstheorie

Dr. C.J. Luchsinger

1 Wahrscheinlichkeit

1.1 Zufallsexperiment, Ereignisraum, Ereignisse

Um Zufallsexperimente zu modellieren, in der Sprache der Mathematik zu beschreiben, führen wir folgende Objekte ein: Ω; mathematisch ist dies einfach eine nichtleere Menge. Sie steht (aus Modellierungssicht) für die Menge der Versuchsausgänge; wir nennen sie auch Ereignisraum [engl Sample Space]. Es findet jeweils in einem Experiment genau ein sogenanntes Elementarereignis statt [engl (elementary) Outcome], z.B. $\omega_1 \in \Omega$ oder $\omega_2 \in \Omega$ etc. Ereignisse [engl Events] sind spezielle Teilmengen von Ω (Vorsicht: nicht *irgendeine* Teilmenge; wir müssen dem Ereignis auch eine Wahrscheinlichkeit zuordnen können - siehe später).

Die meisten Ereignisräume sind aus einer der folgenden Liste (wird in der Vorlesung ausgefüllt):

1) Endliche Mengen: ∘ Münzwurf $\Omega = \{K, Z\}$ od engl. $\{h, t\}$
$\Omega = \{0, 1\}$ } math. gleichwertig,

∘ Um uns das Leben einfach zu machen, wann immer möglich & sinnvoll, falls $|\Omega| = n$ wählen wir $\Omega = \{1, 2, \ldots, n\}$

∘ alternativ zB bei KW $\Omega = \{-1, 1\}$

2) Abzählbare Mengen: ∘ ~~Zeit~~ Zeit bis ersten Erfolg : $\Omega = \{1, 2, 3, \ldots\}$ od
Anzahl Misserfolge bis 1. Erfolg: $\Omega = \{0, 1, 2, \ldots\}$
(2 gleichwertige Ansätze; Ge(p) in WTS, AS)

∘ kommt in realer Welt nie (?) vor, aber keine obere Schranke macht wirklich Sinn!

3) $\mathbb{R}$ und $\mathbb{R}_+ := [0,\infty)$: ∘ $\Omega = \mathbb{R}$ f. Messfehler (auch hier: keine sinnvolle obere Schranke normalerweise)

∘ $\Omega = [0,\infty)$ „Zeit bis Atom zerfällt"

∘ $\Omega = [0,60]$ „Jemand kommt innert 60', keine Ahnung wann!"

4) Endliche kartesische Produkte (Replika):

∘ mehrfacher Münzwurf $\Omega_0 = \{0,1\}$; $\Omega = \Omega_0^n$, wo $\Omega_0^n = \{\omega = (\omega_1,\dots,\omega_n) : \omega_i \in \Omega_0 \ \forall i\}$

∘ auch $\Omega_0 = \{-1,1\}$ (RW)

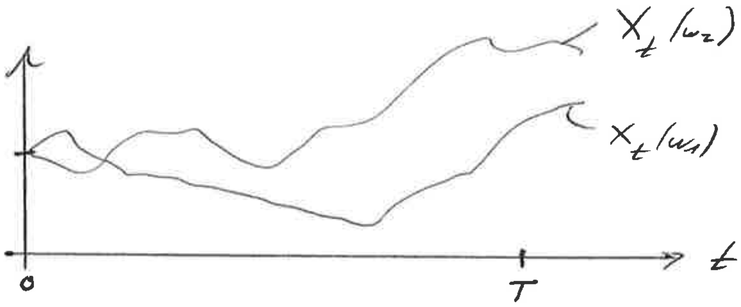
∘ auch $\Omega_0 = \mathbb{R}$ (Zeitreihenanalyse)

• ω ist n-Tupel; ω_i Ausgang des i-ten Experimentes

5) Unendliche kartesische Produkte (Replika):

∘ $\Omega = \Omega_0^{\mathbb{N}} = \{\omega = (\omega_1, \omega_2, \dots)\}$; zB unendl. RW in Vlg AS

6) Funktionen:



• falls stetige Funktionen: $\Omega = C[0,T]$

∘ Physik, El Eng, Oekonomie

∘ sog. Stochastische Prozesse → Vlg AS & Vlg SP (Zukunftsmusik an UBS)

Grobklassifikation SP:

Raum \ Zeit	diskret n	stetig t
diskret	MC	MP
stetig	ZRA	BB

+ Bilder (Realisation) jeder

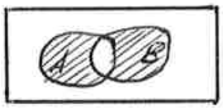
∘ Bemerkung: 1) – 6): alles einfach Menge, nichtleer.

Wir wenden uns jetzt den Ereignissen zu, also speziellen Teilmengen von Ω. Wir wollen ab 1.3 diesen Ereignissen auch eine Wahrscheinlichkeit zuordnen.

Nebenbemerkung: Wir müssen uns in einer Mathematikvorlesung mit der Frage auseinandersetzen, welche Verknüpfungsoperationen mit Mengen zugelassen sein sollen. Wenn wir hier nicht vorsichtig sind, können üble Sachen passieren; mehr dazu am Schluss dieses Kapitels. Wenn Sie jemals Serviceveranstaltungen für andere Studiengänge halten (v.a. Biologie, Medizin, Geographie, Psychologie, Soziologie), sollten Sie diese Diskussion nach Möglichkeit vermeiden; in den Ingenieurwissenschaften, Physik und quantitative Finance kann es notwendig sein, dass Sie dies kurz besprechen.

Da Sie bereits eine einführende Veranstaltung in diesem Gebiet gehört haben, können wir uns die elementaren Verknüpfungsoperationen wie $A \cap B$ und $A \cup B$ sparen und gleich zu den verbleibenden, für uns neuen Verknüpfungen schreiten, welche wir später da und dort benötigen:

1) $A \Delta B := (A \backslash B) \cup (B \backslash A)$ ist die sogenannte symmetrische Differenz; ein elementares Ereignis soll dabei in A oder B sein, nicht aber in beiden.



2) StudentInnen, welche bereits die Vorlesung angewandte Stochastik besucht haben, kennen die folgenden beiden Mengen:

$$\limsup_n A_n := \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n \qquad \oplus \qquad \text{(u.o. / i.o.)}$$

merken mit Suppenteller

und

$$\liminf_n A_n := \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n \qquad \oplus \qquad \text{(ult. = ultimately; „letzten Endes")}$$

Wir müssen uns darüber unterhalten, was diese Gebilde denn sind:

$\oplus$ Ereignis bestehend im Eintreten von ∞ vielen der Ereignisse $A_1, A_2, \ldots$

„ω muss immer wieder mal in einem A_n drin sein, um in limsup drin zu sein"

$$\left(\bigcup_{n=1}^{\infty} A_n\right) \cap \left(\bigcup_{n=2}^{\infty} A_n\right) \cap \left(\bigcup_{n=3}^{\infty} A_n\right) \cap \ldots$$

(zB RW kommt unendlich oft nach 0 ; $A_n := \{X_n = 0\}$)

$\oplus$ Ereignis, bestehend im Eintreten aller Ereignisse $A_1, A_2, \ldots$ mit eventueller Ausnahme einer endlichen Anzahl : $\left(\bigcap_{n \geq 1} A_n\right) \cup \left(\bigcap_{n \geq 2} A_n\right) \cup \left(\bigcap_{n \geq 3} A_n\right) \cup \ldots$

$\omega \notin \bigcap_{n=1}^{\infty} A_n$, $\omega \notin \bigcap_{n=2}^{\infty} A_n$, ..., $\omega \in \bigcap_{n=18}^{\infty} A_n$, ... (ab 18 immer)

(zB MC & Absorption in 0; $\{X_n = 0\} =: A_n$)

zu lim, limsup, liminf aus Analysis I ; FoI, §9
(von Zahlenfolge)

- $\lim\limits_{n\to\infty} a_n$ existiert oder gar nicht

- $\limsup\limits_{n\to\infty} a_n := \lim\limits_{n\to\infty} \underbrace{(\sup\{a_k \mid k \geq n\})}_{\text{monoton fallend}}$ $\exists$ immer eigentlich oder uneigentlich $(-\infty)$, weil monoton fallend

$a_n := (-1)^n\left(1+\frac{1}{n}\right)$

- $\liminf\limits_{n\to\infty} a_n := \lim\limits_{n\to\infty} \underbrace{(\inf\{a_k \mid k \geq n\})}_{\text{monoton wachsend}}$ ——— // ———
  ——— // ——— $(+\infty)$ ——— // ———
  monoton wachsend

- $a_n \longrightarrow a \iff \limsup\limits_n a_n = \liminf\limits_n a_n = a$
- $\limsup a_n \geq \liminf a_n$ (ausser bei $\emptyset$)
- damit praktische Beweisstrategie, wenn man nicht weiss, ob $\lim \exists$:
  (Studis suchen lassen)

$\limsup a_n \leq \ldots\ldots \leq b \leq \ldots\ldots \leq \liminf a_n$

$\Rightarrow \limsup a_n = \liminf a_n$ & $= \lim a_n$; da insbesondere existiert!
& $= b$

Jetzt retour WT bzw. Mengenlehre, Ereignisse

FS 14

- in WTS: falls $A_n \nearrow$, definiere $A := \lim A_n = \bigcup A_n$
  falls $B_n \searrow$, definiere $B := \lim B_n = \bigcap B_n$
- Spezialfall von p. 40

FS 10

Wir *definieren*, dass eine Folge von Ereignissen $A_1, A_2, \ldots$ gegen A konvergiert, notiert als

$$\lim_n A_n = A,$$

analog Analysis I bei Zahlenfolgen

wenn $\limsup_n A_n = \liminf_n A_n = A$. Sie zeigen in den Übungen, dass monotone Folgen von Mengen in diesem obigen Sinne konvergieren (wogegen?).

FS 11

Wir fassen die mengentheoretischen Ausdrücke und ihre Bedeutung für die Wahrscheinlichkeitstheorie in folgender Tabelle zusammen:

Symbol	**Mengentheorie / Bedeutung für die WT**
Ω	Menge / Ereignisraum, Menge der Versuchsausgänge
ω	Element von Ω / Elementarereignis, Versuchsausgang
A	Teilmenge von Ω / Ereignis; falls $\omega \in A$, sagt man, dass das Ereignis A eingetreten ist
A^c	Komplement von A / kein Elementarereignis aus A findet statt
$A \cap B$	Schnittmenge von A und B / ein Elementarereignis aus A und B findet statt
$A \cup B$	Vereinigung von A und B / ein Elementarereignis aus A oder B findet statt
$A \backslash B$	A ohne B / ein Elementarereignis aus A tritt ein, aber nicht aus B
$A \subset B$	A ist Teilmenge von B / Wenn ein Elementarereignis aus A stattfindet, dann immer auch ein Elementarereignis aus B
$\limsup_n A_n$	$\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$ / Ereignis, bestehend im Eintreten von unendlich vielen der Ereignisse $A_1, A_2 \ldots$
$\liminf_n A_n$	$\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$ / Ereignis, bestehend im Eintreten aller Ereignisse $A_1, A_2 \ldots$, mit eventueller Ausnahme einer endlichen Anzahl
ϕ	leere Menge / unmögliches Ereignis
Ω	ganze Menge / sicheres Ereignis (etwas muss passieren)

In der Literatur trifft man häufig folgende Notationen noch an: $\dot{\cup}$ für disjunkte Vereinigung, AB für die Schnittmenge, $A + B$ bzw $\sum_i A_i$ für disjunkte Vereinigungen.

Manchmal erlebt man den Umgang mit Funktionen einfacher als den mit Mengen. Weil wir Gott sei Dank eine 1 zu 1 Beziehung zwischen Mengen und Indikator Funktionen herstellen können, dürfen wir vieles auf der Ebene von Funktionen erledigen statt auf der Ebene von Mengen. Die 1 zu 1 Beziehung ist denn einfach die Indikatorfunktion einer Menge:

$$\mathbf{1}_A(\omega) := \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \notin A. \end{cases}$$

Wir wollen diese Funktion erstmal ein bisschen kennenlernen; in der Klasse: welche der folgenden Ausdrücke sind gleich?

$\mathbf{1}_{A\cup B}, \mathbf{1}_{A^c}, \min\{\mathbf{1}_A, \mathbf{1}_B\}, \mathbf{1}_{A\Delta B}, \mathbf{1}_{A\cap B}, \mathbf{1} - \mathbf{1}_A, \max\{\mathbf{1}_A, \mathbf{1}_B\}, \mathbf{1}_A\mathbf{1}_B, |\mathbf{1}_A - \mathbf{1}_B|$

∘ Farben benutzen & Graphen machen

- $\mathbf{1}_{A\cup B} = \max\{\mathbf{1}_A, \mathbf{1}_B\}$
- $\mathbf{1}_{A\cap B} = \min\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A \cdot \mathbf{1}_B$
- $\mathbf{1}_{A^c} = \mathbf{1}_\Omega - \mathbf{1}_A$
- $\mathbf{1}_{A\Delta B} = |\mathbf{1}_A - \mathbf{1}_B|$

Einschub: Konvergenz Fkt. — glm — pktw

von links
1. $\omega \in \liminf A_n$
0 gdw endlich

5' Zeit lassen

Überlegen Sie sich jetzt, dass

$$\text{(A)} \qquad \lim_{n\to\infty} A_n = A$$

genau dann wenn punktweise gilt

$$\text{(B)} \qquad \lim_{n\to\infty} \mathbf{1}_{A_n}(\omega) = \mathbf{1}_A(\omega).$$

∘ zuerst $\liminf \mathbf{1}_{A_n} = \mathbf{1}_{\liminf A_n}$ — sup — " — sup — was steht links, rechts

„immer wieder dran"

(A) ⇒ (B): $\liminf_n \mathbf{1}_{A_n} = \mathbf{1}_{\liminf_n A_n} \overset{(A)}{=} \mathbf{1}_{\limsup_n A_n} = \limsup_n \mathbf{1}_{A_n} \Rightarrow$ (B), da liminf = limsup pktw.

(B) ⇒ (A): ∃ lim ⇒ limsup & liminf & sind gleich

$\mathbf{1}_{\liminf A_n} = \liminf \mathbf{1}_{A_n} \overset{(B)}{=} \limsup \mathbf{1}_{A_n} = \mathbf{1}_{\limsup A_n} \Rightarrow \limsup A_n = \liminf A_n$

Zu 4. Punkt auf p 6

Zu jedem Teilmengensystem Z von Ω $\exists!$ σ-Algebra $\sigma(Z)$ (die von Z erzeugte σ-Algebra) derart, daß

a) $Z \subseteq \sigma(Z)$

b) $\forall$ σ-Algebra U mit $Z \subseteq U$ gilt $\sigma(Z) \subseteq U$ („kleinste")

Weshalb geht das? ∘ $\exists$ Kandidat, nämlich $\mathcal{P}(\Omega)$; die <u>ist</u> σ-Algebra & $Z \subseteq \mathcal{P}(\Omega)$

∘ Nehme jetzt <u>alle</u> σ-Algebren, welche Z enthalten & bilden den Durchschnitt $\Longrightarrow$ ist wieder σ-Algebra & Z ist darin enthalten

mündlich $\rightarrow$ $Z \subseteq$ ...
σ-Alg.
kleinste
eindeutig

(Prüfung)

1.2 Spezielle Mengen von Mengen (σ-Algebra, Dynkin- und π-Systeme)

1.2.1 σ-Algebren

Wir wollen den Ereignissen (z.B. A aus Ω) später eine Wahrscheinlichkeit ($P[A]$) zuordnen. Wenn wir mehrere Ereignisse vorgegeben haben, wollen wir auch die Wahrscheinlichkeiten von deren Vereinigungen, Durchschnitten oder Komplementen angeben können. An die Menge der Teilmengen von Ω, welche wir untersuchen, stellen wir also ein paar wenige Bedingungen:

Definition 1.1 [σ-Algebra] *Ein Teilmengensystem $\mathcal{A}$ von Ω heisst σ-Algebra, wenn folgende 3 Bedingungen erfüllt sind:*

a) $\Omega \in \mathcal{A}$

b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

c) $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \cup_{n \geq 1} A_n \in \mathcal{A}$.

1. Wieso muss ϕ immer in einer σ-Algebra enthalten sein?

wegen a) & b): $\phi = \Omega^c$

2. Welches ist die kleinste σ-Algebra überhaupt?

$\mathcal{A} = \{\phi, \Omega\}$

3. Wieso muss mit A und B immer auch $A \cap B$ in einer σ-Algebra enthalten sein?

$A \cap B = (A^c \cup B^c)^c$ &
- $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
- $B \in \mathcal{A} \Rightarrow B^c \in \mathcal{A}$
- $A^c, B^c \in \mathcal{A} \Rightarrow A^c \cup B^c \in \mathcal{A} \Rightarrow (A^c \cup B^c)^c \in \mathcal{A}$

4. Welches ist die kleinste σ-Algebra, welche Ereignis A enthält (von A erzeugte σ-Algebra)? $\mathcal{A} = \{\phi, \Omega, A, A^c\}$

$\sigma(Z)$ = von Z erzeugte σ-Algebra, was ist das genau, wenn Z ein Teilmengensystem von Ω ist?

Falls $|\Omega| = n < \infty$, so hat die Potenzmenge von Ω bekanntlich Kardinalität 2^n, ist also wiederum endlich. Man kann also im Fall $|\Omega| = n < \infty$ einfach als $\mathcal{A}$ die Potenzmenge von Ω wählen und muss sich dann nicht mehr sorgen, dass man allenfalls eine Menge untersucht, die gar nicht mehr in der σ-Algebra drin ist.

obschon $\mathcal{P}(\Omega)$ ist σ-Algebra!

Nebenbemerkung: Der naive Wunsch, im Fall $\Omega = \mathbb{R}$ als σ-Algebra einfach die Potenzmenge von $\mathbb{R}$ zu nehmen, ist zwar verständlich, führt aber zu unerwünschten Resultaten. Wir werden am Ende dieses Kapitels diesen Punkt kurz diskutieren (Satz 1.30 von Banach und Kuratowski). Wenn Sie also jemals in Service-Veranstaltungen Nicht-Mathematiker/innen unterrichten, sind Sie realistischerweise gezwungen, bei der Einführung normalverteilter Zufallsgrössen zu mogeln: Sie können *nicht* für jede x-beliebige Menge B aus $\mathbb{R}$ angeben, wie gross die Wahrscheinlichkeit ist, dass eine normalverteilte Zufallsgrösse X Werte in B annimmt. Es kommt dann nämlich vor, dass die normalverteilte Zufallsgrösse X einzelne Punkte mit Wahrscheinlichkeit grösser Null annimmt. Dies ist nicht das, was wir unter einer normalverteilten Zufallsgrösse verstehen wollen.

Wir müssen uns also einschränken; man nimmt statt der Potenzmenge von $\mathbb{R}$ die sogenannte *Borel-σ-Algebra* $\mathcal{B}(\mathbb{R})$. Sie ist per Definitionem die kleinste σ-Algebra auf $\mathbb{R}$, welche alle geschlossenen Intervalle enthält. Die Mengen aus $\mathcal{B}(\mathbb{R})$ nennen wir Borel-Mengen. Man sagt auch, $\mathcal{B}(\mathbb{R})$ wird von der Menge der geschlossenen Intervalle *erzeugt*; mehr dazu in den Übungen.

Wir wollen $\mathcal{B}(\mathbb{R})$ ein bisschen untersuchen; was ist darin alles enthalten?

$\sigma\{[a,b] / a<b\} \overset{\circledast}{=} \sigma\{(a,b] / a<b\} = \sigma\{[a,b) / a<b\} = \sigma\{(a,b) / a<b\}$
ist dann das Schlussresultat.

zB f $\circledast$: $(a,b] = \bigcup_n [a+\frac{1}{n}, b]$ & $[a,b] = \bigcap_n (a-\frac{1}{n}, b] \Rightarrow \circledast$

auch $\{a\} = \bigcap_n [a, a+\frac{1}{n}] \Rightarrow$ „fast alles drin"; aber p. 34: $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$, (Vitalimenge, vgl. Bandle)

insb N-Tabellen: $(-\infty, a]$ drin $\Rightarrow P[\mathcal{N}(\mu, \sigma^2) \leq a]$ immer unproblematisch

Was glauben Sie, wie ist die Kardinalität von $\mathcal{B}(\mathbb{R})$?

Bem. dazu: $|\mathbb{R}^2| = |\mathbb{R}| = |\mathcal{B}(\mathbb{R})|$

ohne Beweis

1.2.2 Dynkin- und π-Systeme

Wenn Sie ein komplexes, abstraktes Mengensystem dahingehend untersuchen müssen, ob es sich dabei um eine σ-Algebra handelt, kann dies auf direktem Weg sehr schwierig sein. Die folgenden Mengensysteme können hier helfen:

Definition 1.2 [Dynkin-System, auch d-System oder Monoton-System] *Ein Teilmengensystem $\mathcal{D}$ von Ω heisst Dynkin-System, wenn folgende 3 Bedingungen erfüllt sind:*

a) $\Omega \in \mathcal{D}$

b) $A \in \mathcal{D} \Rightarrow A^c \in \mathcal{D}$

c) $A_1, A_2, \ldots \in \mathcal{D}$, *paarweise disjunkt,* $\Rightarrow \cup_{n\geq 1} A_n \in \mathcal{D}$.

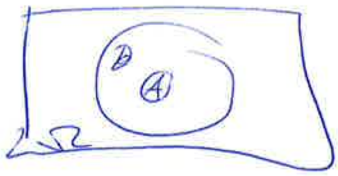
Untersuchen Sie den Zusammenhang zwischen Dynkin-System und σ-Algebra.

Jede σ-Algebra ist auch ein Dynkin-System.

Ein Beispiel eines Dynkin-Systems: FS12

- Ω so, daß $|\Omega| < \infty$ & $\exists p \in \mathbb{N} : |\Omega| = 2p$ (gerade # Elemente)
- $\mathcal{D} := \{D \subset \Omega \,/\, |D| = 2q : 0 \leq q \leq p\}$
- Dynkin-System? Ja
- σ-Algebra? Nein, da c) verletzt sobald $p > 1$.

FS14



Lemma 1.3 *Sei $\mathcal{D}$ ein Dynkin-System. Dann gelten:*

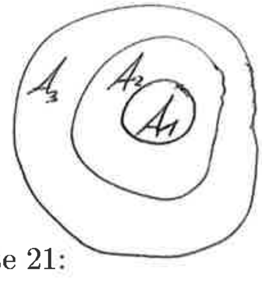
1. *$A, B \in \mathcal{D}$ und $A \subset B$, dann gilt auch $B \backslash A \in \mathcal{D}$ [Stabilität des Dynkin-Systems bei Bildung eigentlicher Komplemente]*

2. *$(A_n)_n$ eine monoton wachsende Folge aus $\mathcal{D}$, dann gilt $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$.* (Monoton-System)//

Beweis Lemma 1.3 (schnell's rechnen lassen)

① $A \dot\cup B^c \in \mathcal{D} \Rightarrow (A \dot\cup B^c)^c \in \mathcal{D} \Rightarrow A^c \cap B \in \mathcal{D} \Rightarrow B \cap A^c \in \mathcal{D} \Rightarrow B \setminus A \in \mathcal{D}$

② $\dot\bigcup_{n=1}^{\infty} A_n = \underbrace{\dot\bigcup_{n=1}^{\infty} \underbrace{(A_n \setminus A_{n-1})}_{\in \mathcal{D} \text{ wegen } ①}}_{\in \mathcal{D}, \text{ weil jetzt disjunkt (Ringe)}}$ wo $A_0 := \emptyset$



Wir ziehen hiermit gleich mit der Definition eines d-Systems aus Karr Seite 21:

Karr p. 21 d-system: $\Omega \in d$ (K1)
$A, B \in d,\ A \subseteq B \Rightarrow B \setminus A \in d$ (K2)
$(A_i) \nearrow \in d \Rightarrow \bigcup_{i=1}^{\infty} A_i \in d$ (K3)

Lachr p. 8: $\Omega \in \mathcal{D}$ (L1)
$A \in \mathcal{D} \Rightarrow A^c \in \mathcal{D}$ (L2)
A_i pw disjunkt $\Rightarrow \dot\bigcup_{i \geq 1} A_i \in \mathcal{D}$ (L3)

Beh.: $d \equiv \mathcal{D}$

- $\mathcal{D} \Rightarrow d$: • K1 ≡ L1; K2 folgt aus Lemma 1.3.1 (& wenn $A = B$, $A \setminus B = \emptyset \in d$ da L1 & L2 $\Rightarrow \emptyset \in d$); K3 folgt aus Lemma 1.3.2.

- $d \Rightarrow \mathcal{D}$: • L1 ≡ K1; L2: $A \in \mathcal{D}$, $\Omega \in \mathcal{D} \Rightarrow$ da $A \subset \Omega$ ~~ist~~ auch $\Omega \setminus A = A^c \in \mathcal{D}$ (K2) ⊕
  - L3: • $A_1 \in \mathcal{D}$, $A_2 \in \mathcal{D}$ & $A_1 \cap A_2 = \emptyset \Rightarrow$

$$A_1 \cup A_2 = \underbrace{\underbrace{(\underbrace{A_1^c}_{\in \mathcal{D} \text{ wegen } \oplus} \setminus A_2)}_{\in \mathcal{D}, \text{ da } A_2 \subseteq A_1^c \ (\text{da } A_1 \cap A_2 = \emptyset)}{}^c}_{\in \mathcal{D}, \text{ wegen } \oplus}$$

$\Longrightarrow$ also endliche, disjunkte Vereinigungen in $\mathcal{D}$

- $B_n := \dot\bigcup_{k=1}^{n} A_k \in \mathcal{D}$, wo A_i pw disjunkt

$(B_n \nearrow) \underset{K3}{\Longrightarrow} \bigcup_{n \geq 1} B_n \in \mathcal{D}$ aber $\dot\bigcup_{n=1}^{\infty} B_n = \dot\bigcup_{n=1}^{\infty} \dot\bigcup_{k=1}^{n} A_k = \dot\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$

Definition 1.4 [π-System, Durchschnittsstabilität] *Ein Teilmengensystem $\mathcal{C}$ von Ω heisst π-System oder durchschnittsstabil, wenn mit $A, B \in \mathcal{C}$ auch $A \cap B \in \mathcal{C}$.*

(z.B. $\{(a,b] \mid a<b\} \cup \{\emptyset\}$ ist π-System) $\rightarrow$ Bausteine $\mathcal{B}(\mathbb{R})$

Satz 1.5 *Ein Dynkin-System ist genau dann eine σ-Algebra, wenn es auch durchschnittsstabil ist.*

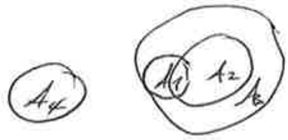
Beweis Satz 1.5

- σ-Algebra $\Rightarrow$ immer auch Dynkin-System
- z.z.: Dynkin & π-System $\Rightarrow$ σ-Algebra:
  - erste beide Eigenschaften klar
  - Seien $A, B \in \mathcal{D}$, dann gilt erstmal allgemein
    - $A \setminus B = \underbrace{A \setminus \underbrace{(A \cap B)}_{\in \mathcal{D}, \text{ da } \pi\text{-System}}}_{\in \mathcal{D} \text{ wegen Lemma 1.3.1.}}$
    - $A \cup B = (A \setminus B) \cup B$, wobei $(A \setminus B) \cap B = \emptyset \Rightarrow$
    - $A \cup B \in \mathcal{D}$, f. endl. Vereinigungen ‡
    - $\bigcup_{n=1}^{\infty} A_n = \underbrace{\bigcup_{n=0}^{\infty} (A'_{n+1} \setminus A'_n)}_{\text{elementfremde Vereinigung: immer nur das dazunehmen, was neu}}$, wobei $A'_0 := \emptyset$

      $A'_n := \bigcup_{k=1}^{n} A_k \in \mathcal{D}$ wegen ‡

      $\in \mathcal{D}$ wegen Lemma 1.3.1

      $\in \mathcal{D}$ wegen Definition 1.2c)



$4\times \Rightarrow A'_1 = A_1$

$A'_2 = A_1 \cup A_2$

$A'_3 = A_1 \cup A_2 \cup A_3$

$A'_4 = \ldots$

Wie bei den σ-Algebren, die von Mengensystemen erzeugt werden können, kann man auch Dynkin-Systeme von Mengen erzeugen; analog gilt hier per Definitionem nämlich: Sei $\mathcal{U}$ ein Teilmengensystem von Ω. Dann ist per Definitionem $\mathcal{D}(\mathcal{U})$ das kleinste Dynkin-System, welches $\mathcal{U}$ enthält. Es gilt dann der zentrale

Satz 1.6 [Monoton-Lemma für Mengen] *Sei C ein π-System. Dann gilt:*

$$\mathcal{D}(C) = \sigma(C).$$

[C in „Script"]

Beweis Satz 1.6 (darauf lange bis verstanden)

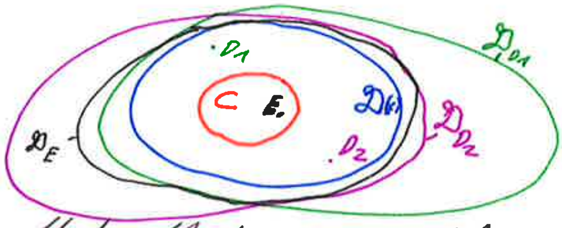
- Jede σ-Algebra ist auch ein Dynkin-System $\Rightarrow \sigma(C)$ ist ein Dynkin-System, welches C enthält $\Rightarrow \mathcal{D}(C) \subseteq \sigma(C)$ ist er nicht kleinstes
- Wir zeigen jetzt noch, daß $\mathcal{D}(C)$ eine σ-Algebra ist; dann gilt auch $\sigma(C) \subseteq \mathcal{D}(C)$ & wir sind fertig.
- Wir benutzen Satz 1.5: Sei für $D \in \mathcal{D}(C)$ $\mathcal{D}_D := \{U \in \mathcal{P}(\Omega) / U \cap D \in \mathcal{D}(C)\}$ ($\cap$-stabil) (w.H.d. ist Dynkin-system)
- ZZ: $\mathcal{D}(C) \subset \mathcal{D}_D \ \forall D \in \mathcal{D}(C)$. ($\pi$-System) $D_1 \cap D_2 \in \mathcal{D}(C)$ unten), weil $D_1 \in \mathcal{D}_{D_2} = \ldots$
- $\forall E \in C$ gilt (da π-system) $C \subset \mathcal{D}_E = \{U \in \mathcal{P}(\Omega) / U \cap E \in \mathcal{D}(C)\} \Rightarrow$
- $\mathcal{D}(C) \subset \mathcal{D}_E$, da $\mathcal{D}(C)$ kleinstes Dynkin-System mit C drin

1)

- $\forall D \in \mathcal{D}(C), \forall E \in C$ gilt also $D \cap E \in \mathcal{D}(C)$
- $\Rightarrow C \subset \mathcal{D}_D \Rightarrow \mathcal{D}(C) \subset \mathcal{D}_D \quad \forall D \in \mathcal{D}(C)$

2)

qed



FS10

2 Bemerkungen zur Bedeutung dieses Satzes:

- Wenn man zeigen will, daß ein System eine σ-Algebra ist, kann es ev. einfacher sein zu zeigen, daß es ein Dynkin-System ist (+ π-System als Erzeuger). Dann wendet man Satz 1.6 an
- Wieso ist dieser Satz so wichtig? Folgende Beweisstrategie: (zB bei Satz 1.12)

1. Man will eine bestimmte Eigenschaft für alle Elemente aus $\sigma(C)$ zeigen, wo C ein π-System.
2. Sei $C \subseteq G$, wo G die Menge, für die die Eigenschaft gilt.
3. Zeigen G ist Dynkin $\Rightarrow \mathcal{D}(C) \subseteq G$, da $\mathcal{D}(C)$ kleinste
4. $\sigma(C) = \mathcal{D}(C) \subseteq G \Rightarrow \sigma(C) \subseteq G$; d.h. alle Elemente von $\sigma(C)$ haben Eigenschaft (Satz 1.6)

FS12
14

1.3 Wahrscheinlichkeit $P[.]$

Definition 1.7 [Wahrscheinlichkeit P] *Eine Wahrscheinlichkeit P ist eine reellwertige Funktion auf den Mengen aus $\mathcal{A}$. Dabei müssen folgende 3 Bedingungen erfüllt sein:*

a) $A \in \mathcal{A} \Rightarrow P[A] \geq 0$

b) $P[\Omega] = 1$

c) Sei $\{A_i\}_{i=1}^{\infty}$ *eine abzählbare Folge von disjunkten Mengen aus* $\mathcal{A}$*, dann muss gelten:*

$$P[\cup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} P[A_i].$$

Man darf in Definition 1.7 c) z.B. auch $A_i = \phi, i \geq 3$ wählen!

Man nennt das Tripel $(\Omega, \mathcal{A}, P)$ auch Wahrscheinlichkeitsraum; auf englisch Probability Space. Eigenschaft c) nennen wir σ-Additivität.

Wir betrachten ein paar einfache Beispiele; mehr in den Übungen:

- Punktmass (Diracmass) in ω: δ_ω ; $\delta_\omega(A) = \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \notin A \end{cases}$ ähnlich (Indikator)
  (Maß 1 in Punkt ω konzentriert)
  (mündliche Prüfung der Eigenschaften)

- Uniformverteilung bei endl. Kardinalität von Ω: $|\Omega| = n < \infty$
  $(\Omega, \mathcal{P}(\Omega), P)$; wobei $\Omega = \{1, 2, \ldots, n\}$
  $\mathcal{P}(\Omega)$ = Potenzmenge von Ω
  $P[A] = \frac{|A|}{n}$
  → prüfen mündlich Def. 1.7.

- $\Omega = \{1, 2, \ldots, n\}$; $\mathcal{A} = \mathcal{P}(\Omega)$; alle W'keiten dieser Art:
  $\exists (p_i)_{i=1}^{n}$; $1 \leq i \leq n$; $p_i \geq 0$ & $\sum_{i=1}^{n} p_i = 1$ & $\forall A \in \mathcal{A}: P[A] = \sum_{i \in A} p_i$
  (vgl. WTS)

Aus Definition 1.7 lassen sich nützliche Eigenschaften ableiten, welche wir im folgenden Lemma zusammenfassen.

Lemma 1.8 [nützliche Eigenschaften von P] *Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Mit $A, B \in \mathcal{A}$, $(A_i)_{i=1}^{n}$ eine endliche und $(B_i)_{i=1}^{\infty}$ eine unendliche Folge von Ereignissen aus $\mathcal{A}$ gelten folgende Aussagen:*

a) $P[\phi] = 0$

b) [endliche Additivität] Sei $\{A_i\}_{i=1}^{n}$ eine endliche Folge von pw disjunkten Mengen aus $\mathcal{A}$, dann muss gelten:

$$P[\cup_{i=1}^{n} A_i] = \sum_{i=1}^{n} P[A_i].$$

Daraus folgt auch das "Prinzip der Gegenwahrscheinlichkeit": $P[A] = 1 - P[A^c]$.

c) $A \subseteq B \Rightarrow P[B] = P[A] + P[B \backslash A]$. Damit ist P insbesondere monoton in dem Sinne, dass $A \subseteq B \Rightarrow P[A] \leq P[B]$

nicht doppelt zählen

d) $P[A \cup B] = P[A] + P[B] - P[A \cap B]$. Damit ist P sogenannt (endlich) subadditiv: $P[A \cup B] \leq P[A] + P[B]$.

ähnlich in RM: „Risk Measures"

e) Sei $\{B_i\}_{i=1}^{\infty}$ eine abzählbare Folge von Mengen aus $\mathcal{A}$, dann muss gelten:

$$P[\cup_{i=1}^{\infty} B_i] \leq \sum_{i=1}^{\infty} P[B_i]. \qquad \text{(Boolesche Ungleichung; subadditiv)}$$

Beweis von Lemma 1.8 Diese Beweise haben wir zum Teil schon in der WTS in den Übungen besprochen. Sie sind jetzt in den WT-Übungen im "Must"-Teil angesiedelt. Im Gegensatz zum ersten Semester wird jetzt auf die strenge mathematische Beweisführung (jenseits von anschaulichen Venn-Diagrammen) Wert gelegt. Die obigen Aussagen sind so einleuchtend, dass man sich (als MathematikerIn) bewusst sein muss, dass sie trotzdem zu beweisen sind!

Satz 1.9 *Sei P eine nichtnegative, endlich additive Mengenfunktion auf $\mathcal{A}$ mit $P[\Omega] = 1$. Dann sind die folgenden 4 Aussagen äquivalent:*

a) P ist auch σ-additiv (und damit eine Wahrscheinlichkeit),

b) Mit $A_n \uparrow A$ in $\mathcal{A}$ gilt auch $P[A_n] \uparrow P[A]$, d.h. $A_n \nearrow$ & $\bigcup_{n=1}^{\infty} A_n = A$

c) Mit $A_n \downarrow A$ in $\mathcal{A}$ gilt auch $P[A_n] \downarrow P[A]$,

d) Mit $A_n \downarrow \phi$ in $\mathcal{A}$ gilt auch $P[A_n] \downarrow 0$.

Die Bedeutung dieses Satzes liegt in folgendem Punkt: endliche Additivität halten wir sofort für eine sinnvolle Anforderung an ein sinnvolles P. Schwierigkeiten hat man allenfalls mit der weitergehenden σ-Additivität. Obiger Satz sagt, dass dies die gleich starke Forderung ist wie Forderungen b), c) und d). Dies sind jedoch Forderungen nach einer (monotonen) Stetigkeit von P, welche wir eher akzeptieren können.

Beweis von Satz 1.9 a) $\Rightarrow$ b) $\Rightarrow$ c) $\Rightarrow$ d) $\Rightarrow$ a)

a) $\Rightarrow$ b): $A_n \nearrow A \rightarrow A = \bigcup_{i=1}^{\infty} (A_i \setminus A_{i-1})$ mit $A_0 := \emptyset$

$$P[A] = P\Big[\bigcup_{i=1}^{\infty} (A_i \setminus A_{i-1})\Big] \underset{a)}{=} \sum_{i=1}^{\infty} P[A_i \setminus A_{i-1}]$$

Def. $\sum_{i=1}^{\infty}$

$$\rightarrow = \lim_{n\to\infty} \sum_{i=1}^{n} P[A_i \setminus A_{i-1}] \underset{(L.1.8c)}{=} \lim_{n\to\infty} \sum_{i=1}^{n} (P[A_i] - P[A_{i-1}])$$

$$= \lim_{n\to\infty} P[A_n]$$, und zwar monoton wachsend (L. 1.8d))

b) $\Rightarrow$ c): mit Komplementbildung analog oben; wie Ue in WTS (vgl. HA)

c) $\Rightarrow$ d): d) ist Spezialfall von c)

d) $\Rightarrow$ a): • Sei $(A_i)_{i\geq 1}$ eine disjunkte Folge von Ereignissen

• $\Big(\bigcup_{i\geq n+1} A_i\Big)_n \searrow \emptyset$, da jedes ω in höchstens einem A_i ist.

• $$P\Big[\dot{\bigcup_{i=1}^{\infty}} A_i\Big] = P\Big[\Big(\dot{\bigcup_{i=1}^{n}} A_i\Big) \cup \Big(\dot{\bigcup_{i=n+1}^{\infty}} A_i\Big)\Big] \underset{\text{endl. Add.}}{=} \sum_{i=1}^{n} P[A_i] + \underbrace{P\Big[\dot{\bigcup_{i=n+1}^{\infty}} A_i\Big]}_{\to 0 \text{ wegen d)}} \quad \forall n$$

$$= \sum_{i=1}^{\infty} P[A_i].$$

Wo stehen wir im vgl. zu Analysis I.

(zu Blatt 15

Ziel, Wunsch, Hoffnung: Resultat analog!

„Angst :-)"

$x_n \rightarrow x_0 \Rightarrow f(x_n) \rightarrow f(x_0)$; wollen

$A_n \rightarrow A \Rightarrow P[A_n] \rightarrow P[A]$

wie genau?

Als Vorbereitung auf den kommenden Satz: konvergiert

$$A_n := \left[\frac{(-1)^n}{n}, 2 + \frac{(-1)^n}{n}\right]$$

und wenn ja, wogegen (vgl p 4 oben)?

$\limsup_n A_n = [0,2]$

$\neq$

$\liminf_n A_n = (0,2)$

$\Rightarrow$ keine Konvergenz

falls $P[\{0\}], P[\{2\}] > 0$

my [illegible] unten

$\exists \lim_n P[A_n]$

Satz 1.10 [Stetigkeit von P] *Es gelten*

$$P[\liminf_n A_n] \overset{\overline{II.}}{\leq} \liminf_n P[A_n] \overset{\overline{I.}}{\leq} \limsup_n P[A_n] \overset{\overline{III.}}{\leq} P[\limsup_n A_n]$$

und damit: falls $A_n \to A$, dann auch $P[A_n] \to P[A]$.

warum?

Beweis von Satz 1.10

$\underline{I.}$ Immer richtig

$\underline{II.}$ „$\Leftrightarrow$" $\underline{III.}$ mit Komplementbildung; beweisen nur $\underline{II.}$: $P[\liminf_n A_n] \leq \liminf_n P[A_n]$

(liminf von Mengen)

liminf von Zahlen

- $\liminf_n A_n = \bigcup_{k=1}^{\infty} \bigcap_{n \geq k} A_n$ & $\forall k$: $\bigcap_{n \geq k} A_n \subseteq A_k$.
- Damit wegen Lemma 1.8 c):

$$P\left[\bigcap_{n \geq k} A_n\right] \leq P[A_k] \quad (\forall k)$$

$\overset{\text{Analysis I}}{\Rightarrow}$ $\liminf_k P\left[\bigcap_{n \geq k} A_n\right] \leq \liminf_k P[A_k]$

- $\left(\bigcap_{n \geq k} A_n\right)_k \nearrow \liminf_k A_k = \bigcup_{k=1}^{\infty}\left(\bigcap_{n \geq k} A_n\right)$
- Satz 1.9 b): Konvergenz gegen $P[\liminf_n A_n]$

Satz 1.11 [Borel-Cantelli I - wichtig für Konvergenzaussagen]

$$\sum_{n=1}^{\infty} P[A_n] < \infty \Rightarrow P[\limsup_n A_n] = 0. \qquad \text{(BC − I)}$$

Es folgt wegen Satz 1.10 automatisch auch $\limsup_{n\to\infty} P[A_n] = 0$ und damit auch $\lim_{n\to\infty} P[A_n] = 0$; spätestens jetzt sollte dies an ein Resultat aus der Analysis I erinnern! Die Hauptaussage (BC-I) ist jedoch flexibler einsetzbar, da der limsup sehr umfassend ist.

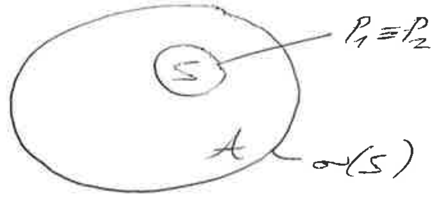
Beweis von Satz 1.11

$$P[\limsup_n A_n] := P\Big[\bigcap_{n=1}^{\infty}\bigcup_{k\geq n} A_k\Big] \underbrace{=}_{\text{da } (\bigcup_{k\geq n} A_k)_n \searrow \bigcap_{n=1}^{\infty}\bigcup_{k\geq n} A_k \text{ \& Satz 1.9 c)}} \lim_{n\to\infty} P\Big[\bigcup_{k\geq n} A_k\Big]$$

$$\underbrace{\leq}_{\text{Boole}} \lim_{n\to\infty} \sum_{k\geq n} P[A_k] \underbrace{=}_{\text{da } \sum_{k=1}^{\infty} P[A_k] < \infty} 0$$

FS 10
12

Satz 1.12 [Eindeutigkeit von P] *Sei $\mathcal{A}$ eine σ-Algebra auf Ω und S ein π-System derart, dass $\sigma(S) = \mathcal{A}$. Seien nun P_1, P_2 Wahrscheinlichkeiten auf $(\Omega, \mathcal{A})$ derart, dass $P_1 = P_2$ auf S, dann $P_1 = P_2$ auf $\mathcal{A}$.*



Beweis von Satz 1.12

- zielen auf Monoton-Lemma f. Mengen (Satz 1.6)
- $S \subseteq \mathcal{M} := \{A \in \mathcal{A} \mid P_1(A) = P_2(A)\}$ „Menge wo $P_1 = P_2$"
- Zeigen, daß $\mathcal{M}$ Dynkin, dann $\mathcal{A} = \sigma(S) \underset{\text{Satz 1.6.}}{=} \mathcal{D}(S) \subseteq \mathcal{M}$ ↳ $\mathcal{D}(S)$ kleinste....

FS 14

- $\mathcal{M}$ ist dynkin: (5' lassen)
  - $P_1(\Omega) = P_2(\Omega) = 1 \Rightarrow \Omega \in \mathcal{M}$
  - $A \in \mathcal{M} \Rightarrow P_1(A) = P_2(A) \Rightarrow \underbrace{1 - P_1(A)}_{P_1(A^c)} = \underbrace{1 - P_2(A)}_{= P_2(A^c)}$

    $\Rightarrow A^c \in \mathcal{M}$
  - vgl. p 80

    $A_n \in \mathcal{M}$; paarw. disjunkt [hohe Ebene: hier einfacher als σ-Alg.]

    $$P_1\Big[\dot{\bigcup_{n=1}^{\infty}} A_n\Big] \underset{P_1 \text{ ist W'keit}}{=} \sum_{n=1}^{\infty} P_1[A_n] \underset{A_n \in \mathcal{M}\ \forall n}{=} \sum_{n=1}^{\infty} P_2[A_n] \underset{P_2 \text{ ist W'keit}}{=} P_2\Big[\dot{\bigcup_{n=1}^{\infty}} A_n\Big]$$

    $\Rightarrow \dot{\bigcup_{n=1}^{\infty}} A_n \in \mathcal{M}$

qed

~~**Wichtige Folgerung für die Anwendungen:**~~

- ~~$\mathcal{N}$-Tabellen führt eindeutig auf ganz $\mathcal{B}(\mathbb{R})$ zu W'keit. (Bitte ☺)~~
- ~~$P[(a,b]]$; $a, b \in \mathbb{R}$, $a < b$ (↳ π-System) echt reicht; erzeugt $\mathcal{B}(\mathbb{R})$~~
- ~~Unterschied zu NatWiss, Psychologen, Soziologen: Gelernt, Hypothesen unproblemat.~~

1.4 Wahrscheinlichkeiten auf $\mathbb{R}$: $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$

Satz 1.17

liefern uns

Aus der Vorlesung WTS kennen wir bereits die Zufallsgrössen, welche wir in Kapitel 2 intensiv studieren werden. Deren Verteilungsfunktionen ~~sind~~ Wahrscheinlichkeiten auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ - und dies ist Grund genug, bereits jetzt in Kapitel 1 die Wahrscheinlichkeiten auf $\mathbb{R}$ ein bisschen genauer unter die Lupe zu nehmen. Bevor wir dies tun, wollen wir noch sogenannte Null-Mengen einführen:

Definition 1.13 [P-Nullmenge, P-fast sicher, (P-f.s., P-fs, fs)] *Ein Ereignis A gilt P-fast sicher, wenn $P[A] = 1$. Hingegen ist A eine P-Nullmenge, wenn $P[A] = 0$.*

<u>Paar kleine Bemerkungen:</u>

Gilt zwingend $A = \Omega$ bzw $A = \phi$? Nein!

Bsp I.: $\Omega = \{0,1,2\}$; $P[\{0\}] = 0$ & $P[\{i\}] = \frac{1}{2}$ $\forall i \in \{1,2\}$

/ diskret

$\mathcal{A} = \mathcal{P}(\Omega)$; $A := \{1,2\} \Rightarrow P[A] = 1$ (P-f.s.)

$B := \{0\} \Rightarrow P[B] = 0$ (P-Nullmenge) (wichtig, angeben)

aber $A \neq \Omega$ & $B \neq \emptyset$

berechtigte Entgegnung: „dumme" Wahl von Ω, P; wie wenn man x^2 als $\frac{x^4}{x^2}$ angibt & dann sagt, es sei bei 0 nicht definiert

Bsp II.: Vorgriff auf Kapitel 2 (& auf Teil 1.6); vgl. WTS:

/ stetig

$X \sim N(0,1)$; $P[X=0] := P[\{\omega \mid X(\omega) = 0\}] = 0$, aber $\{\omega \mid X(\omega) = 0\} \neq \emptyset$ & auch $P[X \in \mathbb{R} \setminus \mathbb{Q}] = 1$; aber $\{\omega \mid X(\omega) \in \mathbb{R} \setminus \mathbb{Q}\} \neq \Omega$. „$X$ f.s. $\in \mathbb{R} \setminus \mathbb{Q}$".

„Lsg": $\Omega' := \Omega \setminus \{\omega \mid X(\omega) \in \mathbb{Q}\}$, dann gilt $\forall \omega \in \Omega' : X(\omega) \in \mathbb{R} \setminus \mathbb{Q}$

(nicht nur f.s.

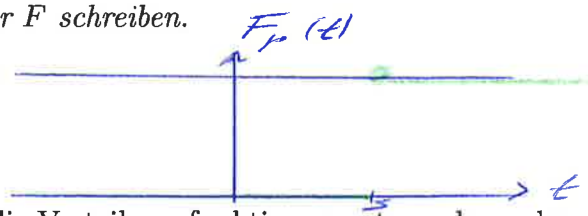
ind ∅ — Zuordnung: $X \mapsto F_X$, P_X

Von Satz 1.12 wissen wir, dass jede Wahrscheinlichkeit P auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ durch die Werte auf den Intervallen der Art $(-\infty, t]$ eindeutig determiniert ist. Es lohnt sich deshalb, diese (aus der WTS bekannten) Gebilde genauer zu untersuchen. Dazu definieren wir erstmals:

Definition 1.14 [Verteilungsfunktion von P] *Die Verteilungsfunktion von P ist die Funktion $F_P : \mathbb{R} \longrightarrow [0,1]$, definiert als $F_P(t) := P[(-\infty, t]]$. Wenn es klar ist, können wir die Indexierung in F_P auch lassen und nur F schreiben.*

BSP: $P \equiv \delta_s$; Punktmass auf s:

$$F_P(t) = \begin{cases} 0 & t < s \\ 1 & t \geq s \end{cases}$$



Achten Sie bitte darauf, dass wir in Kapitel 1 die Verteilungsfunktionen untersuchen, ohne Zufallsgrössen zu erwähnen (ausser zur Motivation)! Wir lernen jetzt die Verteilungsfunktionen ein bisschen kennen. Es gilt

Satz 1.15 [Eindeutigkeit F, P] *Wenn $F_{P_1} = F_{P_2}$, dann gilt $P_1 = P_2$ auf $\mathcal{B}(\mathbb{R})$.* (gdw ☺)

Beweis Satz 1.15
- $\mathcal{J} := \{(a,b] / a \leq b\}$ ist π-System ($\cup \emptyset$)
- $\sigma(\mathcal{J}) = \mathcal{B}(\mathbb{R})$
- Satz 1.12: reicht zu zeigen, dass $P_1 = P_2$ auf $\mathcal{J}$
- $(a,b] = (-\infty, b] \setminus (-\infty, a]$ & $(-\infty, a] \subseteq (-\infty, b]$

$$P_1[(a,b]] = P_1[(-\infty,b] \setminus (-\infty,a]] \overset{L\,1.8.c)}{=} P_1[(-\infty,b]] - P_1[(-\infty,a]]$$
$$=: F_{P_1}(b) - F_{P_1}(a) \underset{\text{Annahme}}{=} F_{P_2}(b) - F_{P_2}(a) := P_2[(-\infty,b]] - P_2[(-\infty,a]]$$
$$= P_2[(a,b]]$$

Wichtige Folgerung für die Anwendungen: In der Vlsg WTS und in der Ausbildung anderer Studiengänge lernen die StudentInnen zum Beispiel die Wahrscheinlichkeiten der Normalverteilung über die Normalverteilungstabelle (meist hinten in Statistik-Büchern) kennen. Man könnte sich fragen, ob durch diese Tabelle (abgesehen von der Maschenweite des Gitters; beachten Sie auch die Monotonie von F) P eindeutig festgelegt ist. Satz 1.15 bejaht dies auf $\mathcal{B}(\mathbb{R})$. $\mathcal{B}(\mathbb{R})$ enthält aber alles, was AnwenderInnen ausrechnen wollen: Komplemente, Vereinigungen, Schnitte.

Der folgende Satz ist bereits aus der WTS bekannt; wir formulieren ihn nochmals und beweisen ihn unter Einsatz der bisherigen Resultate.

Satz 1.16 [Elementare Eigenschaften von F_P] *Sei F_P die Verteilungsfunktion von P. Dann gelten:*

a) F_P ist monoton wachsend; $\underbrace{\text{damit existieren jeweils die Limiten von links und von rechts}}_{\text{Analysis I \& II}}$

b) F_P ist rechtsstetig; a) und b) heissen zusammen vom Französischen: "càdlàg"

↳ „Böppeli' oben"

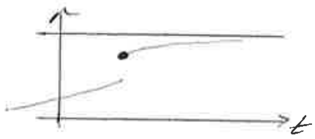
continuant à droite, limite à gauche

c) $\lim_{t\to-\infty} F_P(t) = 0$ und $\lim_{t\to\infty} F_P(t) = 1$.

Beweis von Satz 1.16

a) $s \le t \Rightarrow (-\infty, s] \subseteq (-\infty, t] \overset{\text{Lemma 1.8c)}}{\Longrightarrow} P[(-\infty, s]] \le P[(-\infty, t]] \Rightarrow F_P(s) \le F_P(t)$.

b) da Limiten existieren, darf ich spezielle Folge nehmen: $t_n \searrow t$



, fallend, ∩

$\Rightarrow (-\infty, t_n] \searrow (-\infty, t] \underset{\text{Satz 1.9c)}}{\Longrightarrow} \lim_{n\to\infty} F_P(t_n) := \lim_{n\to\infty} P[(-\infty, t_n]] = P[(-\infty, t]] = F_P(t)$

c) $(-\infty, -n] \searrow \emptyset$ & $(-\infty, n] \nearrow \mathbb{R}$ (wieder spezielle Folge, da $\exists$ lim)

zB $\lim_{t\nearrow\infty} F_P(t) \ge \lim_{t\nearrow\infty} F_P(\lfloor t \rfloor) = \lim_{n\nearrow\infty} F_P(n) = 1$ (Satz 1.9b)

qed

<u>Nachbetrachtung zum Beweis von Teil b)</u>

Vorsicht: $t_n \nearrow t$ & $t_n < t \not\Rightarrow (-\infty, t_n] \nearrow (-\infty, t]$, sondern nur

$\Rightarrow (-\infty, t_n] \nearrow (-\infty, t)$ & deshalb

können wir linksstetigkeit <u>nicht</u> so beweisen (sie stimmt ja auch nicht):

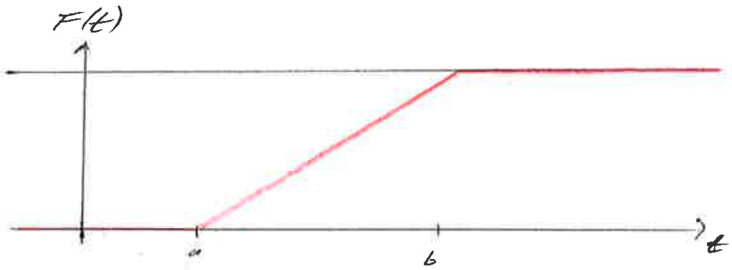
$(-\infty, t_n] \nearrow (-\infty, t) \underset{\text{Satz 1.9b)}}{\Longrightarrow} \lim_{n\to\infty} F_P(t_n) := \lim_{n\to\infty} P[(-\infty, t_n]] = P[(-\infty, t)] =: F_P(t-) \ne F_P(t)$

Ohne Beweis fügen wir noch an, dass *jede* Funktion auf $\mathbb{R}$, welche die Eigenschaften aus Satz 1.16 besitzt, eine Verteilungsfunktion einer Wahrscheinlichkeit P ist. Damit lassen sich beinahe beliebige Wahrscheinlichkeiten entwickeln.

Satz 1.17 *Sei $F : \mathbb{R} \to \mathbb{R}$ monoton wachsend und rechtsstetig mit $F(-\infty) = 0$ und $F(\infty) = 1$. Dann existiert ein eindeutiges P auf $\mathcal{B}(\mathbb{R})$ so, dass $F_P = F$.*

Beispiel zu Satz 1.17

- Sei $a < b$
- $F(t) = \begin{cases} 0 & t < a \\ \frac{t-a}{b-a} & a \leq t \leq b \\ 1 & t > b \end{cases}$



- Bedingungen von Satz 1.17 erfüllt; $\exists P$ (aus WTS: $U[a,b]$)
- $P[[a,b]] := F(b) - F(a)$ ist W'keit (vgl. Def 1.7) auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

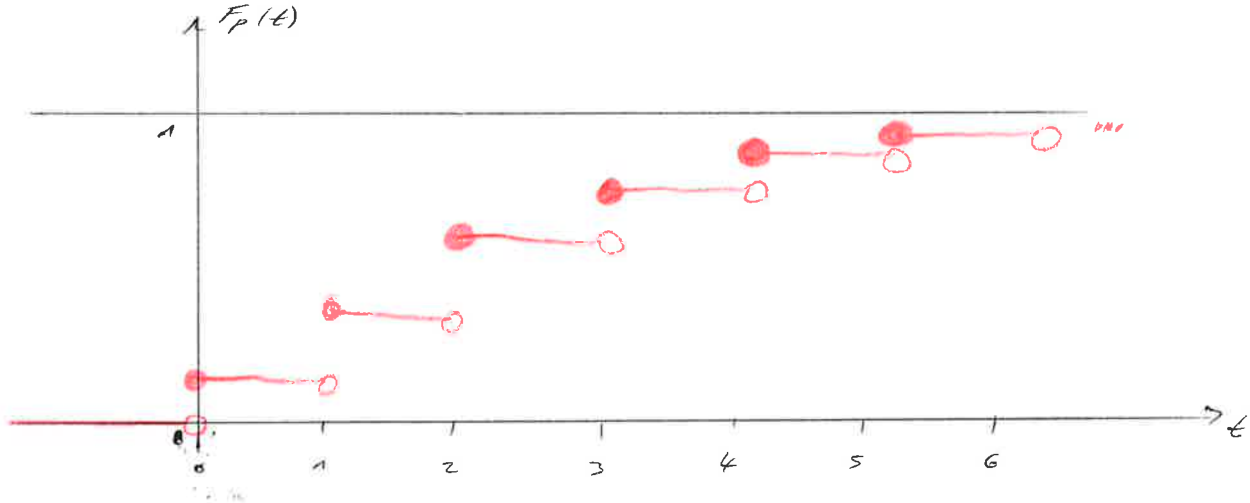
wicht wegen Satz 1.12.

Wir haben in der WTS 2 Typen von Zufallsgrössen kennengelernt: *diskret* und *stetig*. Mittels der Verteilungsfunktionen dieser Zufallsgrössen erhalten wir mit Satz 1.17 also damit auch 2 Typen von Wahrscheinlichkeiten auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Schon in der WTS haben Sie sich vielleicht die Frage gestellt, ob das denn alles sei. Mit wenig Nachdenken kommt man schnell auf die Idee, dass man ja auch Linearkombinationen solcher Wahrscheinlichkeiten nehmen kann (siehe auch Übungsblatt 3). Haben wir damit alles? Die Antwort folgt erst in 1.6 (Vollständige Klassifikation der Wahrscheinlichkeiten auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Wir wollen jedoch kurz, halb zur Repetition, die beiden bisherigen Arten von Wahrscheinlichkeiten nochmals anschauen.

Definition 1.18 [Diskrete Wahrscheinlichkeit] *Eine Wahrscheinlichkeit P auf $\mathbb{R}$ ist diskret, wenn es eine höchstens abzählbare Menge C gibt, sodass $P(C) = 1$.*

Beispiel zu Definition 1.18, von der $Po(\lambda)$-ZG aus WTS motiviert.

- $P[\{k\}] := e^{-\lambda}\frac{\lambda^k}{k!}$; $k \in \mathbb{N}_0$, $\lambda > 0$ $\rightarrow P[\mathbb{N}_0] = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = 1$
- $C := \mathbb{N}_0$, abzählbar



FS14

Der folgende Satz ist derart anschaulich, dass er in der WTS bereits unbewiesen (und vielleicht auch unausgesprochen) benutzt wurde. Er besagt, dass diskrete Wahrscheinlichkeiten endliche oder abzählbar unendliche konvexe Linearkombinationen von Dirac-Massen (Punktmassen) sind. Die Verteilungsfunktionen wachsen nur durch "Sprünge".

(Teil b) (Teil c)

wo stehen wir?

$\rightarrow$ 1.3: P allg auf $(\Omega, \mathcal{A}, P)$

$\rightarrow$ 1.4: P auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$; speziell

$\rightarrow$ $P[A]$, $A \in \mathcal{B}(\mathbb{R})$; $P[(-\infty, t]] =: F_P(t)$ (VF)

Satz 1.19 [Charakterisierung von diskreten Wahrscheinlichkeiten] *Für Wahrscheinlichkeiten auf $\mathbb{R}$ sind die folgenden Aussagen äquivalent:*

a) P ist diskret

b) Es existiert eine reelle Folge (t_i) und Zahlen $p_i \geq 0$ mit $\sum_i p_i = 1$ derart, dass $P = \sum_i p_i \delta_{t_i}$. — formal wie bei Funktionen: $\int f + \int g = \int f+g$ (kein (x); kein (A) bla, ... Borel)

c) Es existiert eine reelle Folge (t_i) und Zahlen $p_i \geq 0$ mit $\sum_i p_i = 1$ derart, dass $F_P(t) = \sum_i p_i \mathbf{1}_{\{t_i \leq t\}}$.

Sprungstelle Sprunghöhe

Wir lassen oben im Satz und unten im Beweis beide Fälle zu: endliche oder abzählbar unendliche Folgen bzw. Reihen.

Beweis Satz 1.19

a) ⇒ b) $C := \{t_i \,/\, i \in I\}$ wo $t_i \neq t_j$ $\forall i \neq j$ & I endlich oder $\mathbb{N}$ & $P[C] = 1$.

$\forall$ Borel-Menge B gilt dann

$$P[B] \overset{HA}{=} P[B \cap C] = \sum_i P[B \cap \{t_i\}] = \sum_{t_i \in B} P[\{t_i\}] = \sum_i \underbrace{P[\{t_i\}]}_{=: p_i} \delta_{t_i}(B)$$

b) ⇒ c) $\forall t \in \mathbb{R}: F_P(t) := P[(-\infty, t]] \overset{b)}{=} \sum_i p_i \delta_{t_i}[(-\infty, t]] = \sum_i p_i \mathbf{1}_{\{t_i \leq t\}}$

c) ⇒ a) $P[\{t_i \,/\, i \in I\}] = \sum_{i \in I} P[\{t_i\}] = \sum_{i \in I} [F_P(t_i) - F_P(t_i -)] \overset{c)}{=} \sum_{i \in I} p_i = 1$

$F_P(t) = P[(-\infty, t]]$

$F_P(t-) = P[(-\infty, t)]$

Bemerkung/Warnung zum Wort "diskret" in der WT und der restlichen Mathematik (zB diskrete Menge): M diskret in $\mathbb{R}$ $\overset{\text{def}}{\Leftrightarrow}$ $\forall x \in M \; \exists \varepsilon > 0: (x - \varepsilon, x + \varepsilon) \cap M = \{x\}$ (isoliert)

Gegensatz: $\mathbb{Q}$ ist dicht in $\mathbb{R}$, also klar nicht isoliert in $\mathbb{R}$; aber:

$\mathbb{Q}$ abzählbar $\Rightarrow (p_i)_{i=1}^{\infty} \geq 0$ mit $\sum_{i=1}^{\infty} p_i = 1$; $P[\{q_i\}] = p_i$ $\forall i \in \mathbb{N}_0$

$\Rightarrow P[\mathbb{Q}] = 1$. $\Rightarrow$ diskrete W'keit

Sprechweise. Gegeben seien zwei Funktionen $f, g: \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Wir nennen f und g *(Lebesgue-) fast überall gleich,* wenn die Menge

$$\{x \in \mathbb{R}^n : f(x) \neq g(x)\}$$

eine Nullmenge ist. $f = g$; $f.\ddot{u}.$ (a.e.) almost everywhere

Bemerkung. Sei $f = g$ fast überall. Falls im Punkt $a \in \mathbb{R}^n$ beide Funktionen endlich und stetig sind, gilt $f(a) = g(a)$. Denn in jeder Umgebung U von a gibt es Punkte $x \in U$ mit $f(x) = g(x)$.

Satz 2. *Seien $f, g: \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ zwei Funktionen, so daß*

$$f = g \quad \textit{fast überall.}$$

Ist f integrierbar, so ist auch g integrierbar und es gilt

$$\int f(x)\,dx = \int g(x)\,dx.$$

Beweis. Sei $\varphi_\nu \in \mathscr{C}_c(\mathbb{R}^n)$, $\nu \in \mathbb{N}$, eine Folge mit

$$\lim_{\nu \to \infty} \|f - \varphi_\nu\|_{L_1} = 0.$$

Sei $A := \{x \in \mathbb{R}^n : f(x) \neq g(x)\}$ und u_A die in Hilfssatz 2 definierte Funktion. Damit gilt

$$|g - \varphi_\nu| \leqq |f - \varphi_\nu| + u_A,$$

also

$$\|g - \varphi_\nu\|_{L_1} \leqq \|f - \varphi_\nu\|_{L_1} + \|u_A\| = \|f - \varphi_\nu\|_{L_1}.$$

Es folgt $\lim \|g - \varphi_\nu\|_{L_1} = 0$; daher ist g integrierbar und

$$\int g(x)\,dx = \lim_{\nu \to \infty} \int \varphi_\nu(x)\,dx = \int f(x)\,dx, \quad \text{q.e.d.}$$

10) **Satz 3.** *Sei $f: \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ eine Funktion mit $\|f\|_{L_1} < \infty$. Dann ist die Menge*

$$A := \{x \in \mathbb{R}^n : f(x) = \pm\infty\}$$

eine Nullmenge.

Die Voraussetzung $\|f\|_{L_1} < \infty$ ist wegen § 6, Satz 4, insbesondere für jede integrierbare Funktion erfüllt.

Beweis. Sei

$$M := \|f\|_{L_1} = \int^* |f(x)|\,dx < \infty.$$

Für jedes $\epsilon > 0$ gilt $\chi_A \leqq \epsilon |f|$, also

$$\int^* \chi_A(x)\,dx \leqq \epsilon M.$$

Dies ist nur möglich, wenn $\int^* \chi_A(x)\,dx = 0$, also A eine Nullmenge ist.

p 71

Satz 2 und Satz 3 [zu]sammen sagen, daß man eine integrierbare Funktion $f: \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ ohne Abänderung des Integrals durch eine fast überall gleiche Funktion ersetzen kann, die nur endliche Werte annimmt.

Satz 4. *Sei $f: \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ eine Funktion. Dann gilt*

$$\|f\|_{L_1} = \int^* |f(x)|\,dx = 0$$

genau dann, wenn $f = 0$ fast überall.

Beweis. Sei $A := \{x \in \mathbb{R}^n : f(x) \neq 0\}$. Falls A Nullmenge ist, folgt aus Hilfssatz 2, daß

$$\int^* |f(x)|\,dx = 0.$$

Sei umgekehrt vorausgesetzt, daß $\int^* |f(x)|\,dx = 0$. Setzt man $f_k := |f|$ für alle $k \in \mathbb{N}$, so ist

$$\chi_A \leqq \sum_{k=0}^{\infty} f_k,$$

daher nach § 6, Satz 2,

$$\|\chi_A\|_{L_1} \leqq \sum_k \|f_k\|_{L_1} = 0,$$

d.h. A ist eine Nullmenge.

Der Raum $L_1(\mathbb{R}^n)$

Im Vektorraum $\mathscr{L}_1^e(\mathbb{R}^n)$ aller integrierbaren Funktionen $f: \mathbb{R}^n \to \mathbb{R}$ sei $\mathscr{N}$ die Teilmenge aller $f \in \mathscr{L}_1^e(\mathbb{R}^n)$ mit

$$\|f\|_{L_1} = \int |f(x)|\,dx = 0.$$

$\mathscr{N}$ ist ein Untervektorraum von $\mathscr{L}_1(\mathbb{R}^n)$. Für $f, g \in \mathscr{L}_1(\mathbb{R}^n)$ gilt nach Satz 4

$$f - g \in \mathscr{N} \iff f = g \quad \text{fast überall.}$$

Wir definieren $L_1(\mathbb{R}^n)$ als den Quotienten-Vektorraum

$$L_1(\mathbb{R}^n) := \mathscr{L}_1(\mathbb{R}^n)/\mathscr{N}. = \{f + N \mid f \in \mathscr{L}_1^e(\mathbb{R}^n)\}$$

$L_1(\mathbb{R}^n)$ besteht aus Äquivalenzklassen integrierbarer Funktionen modulo der Relation „fast überall gleich". Wegen Satz 3 induziert das Integral eine lineare Abbildung

$$\int : L_1(\mathbb{R}^n) \to \mathbb{R}.$$

Ebenso induziert die Pseudonorm $\|\ \|_{L_1}$ eine Abbildung

$$\|\ \|_{L_1} : L_1(\mathbb{R}^n) \to \mathbb{R}_+.$$

somit:

$f = g$ $f.\ddot{u}.$

$\Updownarrow$

Wir wenden uns jetzt den stetigen Wahrscheinlichkeiten zu und präzisieren gleich mal: ab jetzt heissen die stetigen Wahrscheinlichkeiten bzw stetigen Zufallsgrössen aus der WTS *absolut* stetige Wahrscheinlichkeiten (bzw. Zufallsgrössen).

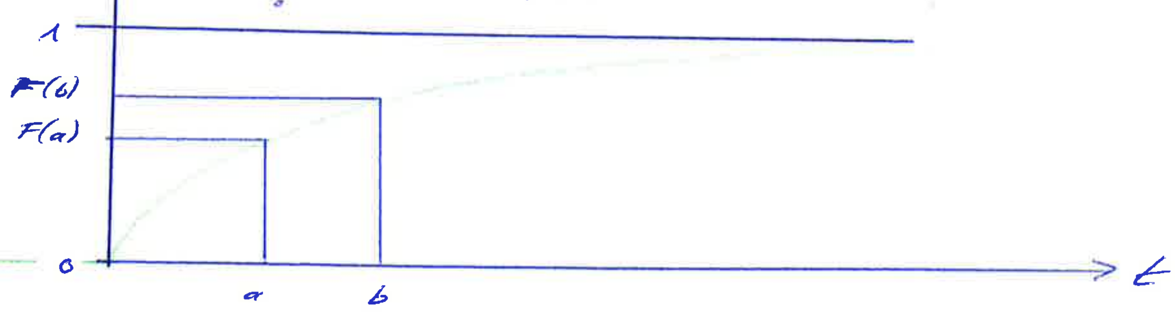
Definition 1.20 [absolut stetige Wahrscheinlichkeit] *Eine Wahrscheinlichkeit P auf $\mathbb{R}$ nennen wir absolut stetig, wenn es eine nichtnegative Funktion f_P (Dichte von P) auf $\mathbb{R}$ derart gibt, dass für alle $(a, b]$*

$$P[(a,b]] = L - \int_a^b f_P(t)dt.$$

Beispiel aus der WTS: von der $\exp(\lambda)$-ZV motiviert; $0 \leq a < b$:

$P[(a,b]] = \int_a^b \lambda e^{-\lambda t} dt$; genauer $f_P(t) = 1_{[0,\infty)}(t)\lambda e^{-\lambda t}$

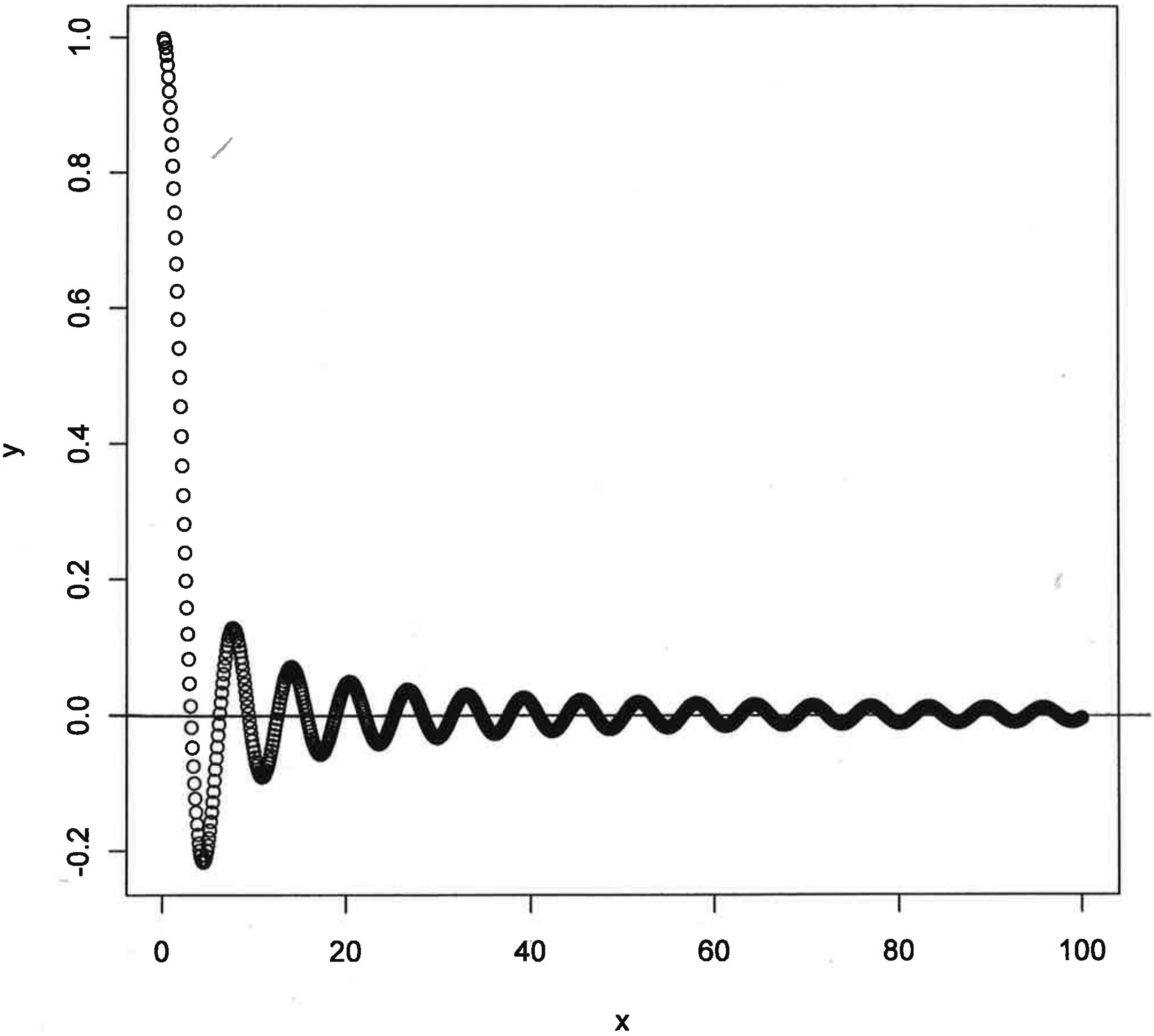
$F_P(t) = \int_0^t \lambda e^{-\lambda s} ds$; $t \geq 0$



Bemerkung zur Dichtefunktion: $f_P(t)$ **ist nicht eindeutig:**

1) Das obige Integral ist ein Lebesgue-Integral (L-$\int$, vgl Vlsg "Reelle Analysis"; siehe auch kommende Seite); aber schon bei einem "normalen" *Riemann-Integral*, R-$\int$, kann man solch eine Dichtefunktion mindestens an *endlich* vielen Punkten ändern.
2) Bei *Lebesgue-Integralen* gilt das sowieso (vgl Vlsg "Reelle Analysis").
3) Die Differenzen bei den diversen denkbaren f_p's betreffen aber lediglich Lebesgue-Nullmengen (Forster Analysis III, Sätze 2-4 in § 7). (Copy & studb)
4) Man spricht deshalb auch von einer "Version" der Dichtefunktion (und wählt dann mit Vorteil zum Beispiel eine stetige Version).

$$\frac{\sin x}{x}$$



p 25 (R-$\int$ immer L-$\int$?)

- falls $f \geq 0 \implies$ R-$\int$ immer auch L-$\int$.
- wegen unterschiedlichen Einführungsansätze
  - * R-$\int$ über Ober & Untersumme
  - * L-$\int$ über f_+, f_- & $|f|$
- zB $\displaystyle\int_1^\infty \frac{\sin x}{x}\,dx$ $\nearrow \exists$ R-$\int$ $\searrow \nexists$ L-$\int$
- auf Kompakten & falls R-$\int |f| \exists \rightarrow$ R-$\int$ ist L-$\int$ & gleich

Bemerkungen zur Integrationsart:

1) Das Integral in Definition 1.20 ist im allgemeinen Fall ein L-$\int$.

2) Wenn der Integrand nichtnegativ ist (zum Beispiel bei einer Dichte), ist ein R-$\int$ immer auch ein L-$\int$ (die Umkehrung gilt nicht - damit ist das L-$\int$ allgemeiner als das R-$\int$).

3) Was wenn der Integrand auch negativ sein darf? $\longrightarrow$ Beiblätter

4) In den Vlsg'en WTS, AS, SM und WT sind konkrete Integrale de facto immer R-$\int$, ausser es wird speziell erwähnt. In den Beweisen sind es aber oft L-$\int$. StudentInnen, welche das L-$\int$ noch nicht kennen, stellen sich ohne Nachteil einfach immer ein R-$\int$ vor. Falls $f_P(t)$ stückweise stetig ist (endliche Unterteilung), ist ein L-$\int$ immer ein R-$\int$.

5) Kontrastbeispiel: L-$\int_0^1 1_{\mathbb{Q}}(s)ds =$ (L-$\int$ aber nicht R-$\int$).

6) Schema Integrationsarten, falls Integrand nicht-negativ:

$$\{Tf\} \subset \{f \mid \text{R-}\!\int \exists\} \subset \{f \mid \text{L-}\!\int \exists\} \subset \{f \mid \text{LS-}\!\int \exists\} \subset \{f \mid \text{stoch.}\!\int \exists\}$$

7) "Stieltjes"-Integrale (Riemann-Stieltjes und Lebesgue-Stieltjes-Integrale) haben auf der Basis (x-Achse) im Allgemeinen keine gleichmässige Gewichtung. Riemann- und Lebesgue-Integrale schon. Mehr dazu in Kapitel 4.

Sie beweisen noch im Must-Teil von Übungsblatt 4 folgende kleine Umformulierung:

Korollar 1.21 [absolut stetige Wahrscheinlichkeit und F_P] *Eine Wahrscheinlichkeit P auf $\mathbb{R}$ ist genau dann absolut stetig, wenn es eine nichtnegative Funktion f_P (Dichte von P) auf $\mathbb{R}$ gibt mit $\int_{-\infty}^{\infty} f_P(s)ds = 1$, so dass*

$$F_P(t) = \int_{-\infty}^{t} f_P(s)ds.$$

Damit können wir also jede beliebige nichtnegative Funktion f mit $\int_{-\infty}^{\infty} f(s)ds = 1$ als Dichte einer Wahrscheinlichkeit P auffassen - dies ergibt uns also ein grosses Universum von Wahrscheinlichkeiten!

• 7 durchnehmen → HA & Ue

1.5 Bedingte Wahrscheinlichkeit $P[A|B]$; Produktformel, Bayes und FTW

Diese Konzepte kamen schon in der Vlsg WTS (und AS) sehr ausführlich zum Einsatz, so dass wir nur als Repetition die Definition und die drei zentralen Regeln angeben. Kleine Aufgaben dazu sind auf Übungsblatt 4 zu lösen.

Definition 1.22 [Bedingte Wahrscheinlichkeit $P[A|B]$]

$$P[A|B] := \frac{P[A \cap B]}{P[B]},$$

falls $P[B] > 0$. Man nennt $P[A|B]$ die bedingte Wahrscheinlichkeit von A gegeben B.

Es gilt die sogenannte Produktformel:

$$P[A|B]P[B] = P[A \cap B] = P[B|A]P[A].$$

Der Leser / die Leserin zeige: $P[.|B]$ ist selber auch eine Wahrscheinlichkeit.

Formel von Bayes: (Umgekehrte Frage: von $P[B|A]$ zu $P[A|B]$)

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B|A]P[A]}{P[B|A]P[A] + P[B|A^c]P[A^c]}.$$

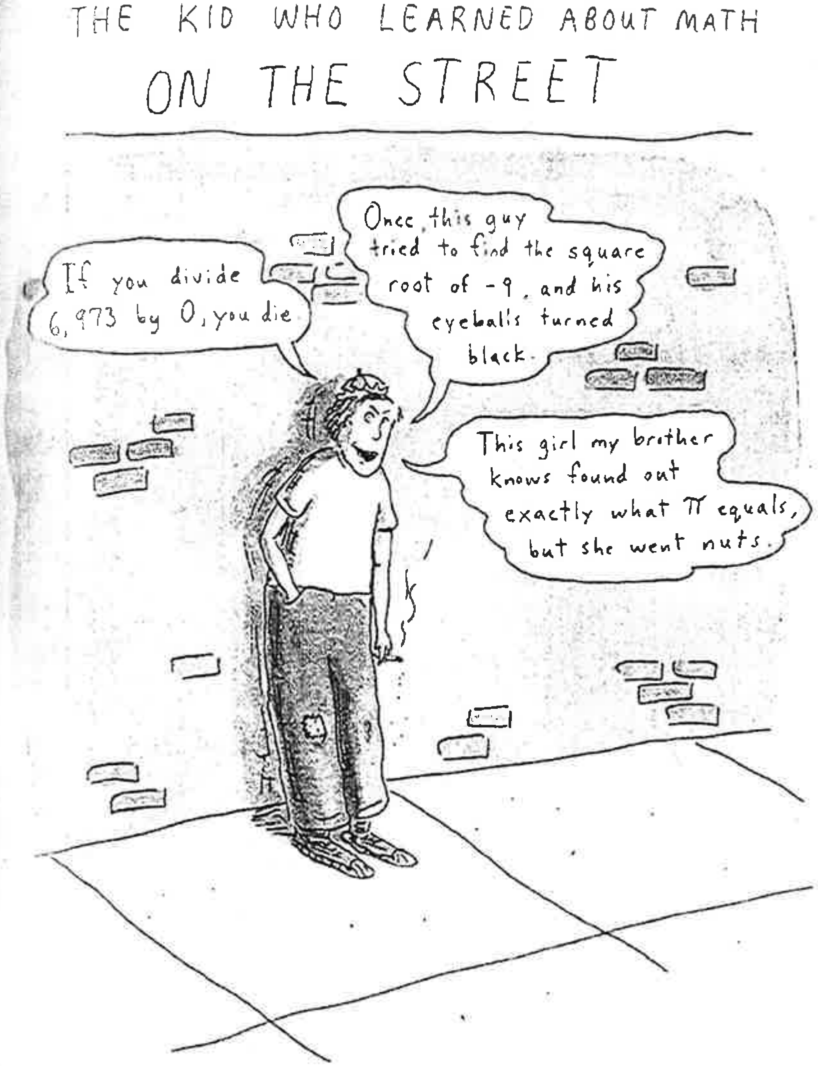
Lemma 1.23 [Formel von der totalen Wahrscheinlichkeit FTW] *$B_1, B_2, \ldots$ sei eine Partition von Ω (die B_i's sind disjunkt und $\cup_{i=1}^{\infty} B_i = \Omega$). Weiter sei für alle $B_i, i \geq 1$, $P[B_i] > 0$ erfüllt. Dann gilt für jedes $A \in \mathcal{A}$:*

$$P[A] = \sum_{i=1}^{\infty} P[A|B_i]P[B_i]. \qquad (FTW)$$

Ein analoges Resultat gilt auch für eine endliche Partition.

machen mit Vermögen & Wl. Partner

Beyond the Real Numbers



Only real numbers greater than or equal to zero have real number square roots. The square root of -1, $\sqrt{-1}$, is not a real number. This is because there is no real number that can be multiplied by itself that results in -1. Multiplying any real number by itself can never give a negative product. In the sixteenth century, mathematician Girolamo Cardano wrote that square roots of negative numbers would cause "mental tortures." In spite of these "tortures," mathematicians invented a new number, called i, to represent $\sqrt{-1}$. The number i is not a real number; it is called an **imaginary number**. Thus, $\sqrt{9} = 3$, $-\sqrt{9} = -3$, but $\sqrt{-9}$ is not a real number. However, $\sqrt{-9}$ is an imaginary number, represented by $3i$. The adjective *real* as a way of describing what we now call the real numbers was first used by the French mathematician and philosopher René Descartes (1596–1650) in response to the concept of imaginary numbers.

zu WT p. 27

1.6 Miscellanea; Sie finden hier Bemerkungen zu:

1.6.1 $\overline{\mathbb{R}}$

1.6.2 **Allgemeine Masse**

1.6.3 **Lebesgue Mass**

1.6.4 **Singulär stetige Wahrscheinlichkeit auf $\mathbb{R}$ - Cantorsches Diskontinuum**

1.6.5 **Vollständige Klassifikation der Wahrscheinlichkeiten auf** $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

1.6.6 **Warum σ-Algebren? Warum P auf $\mathcal{B}(\mathbb{R})$ und nicht auf $\mathcal{P}(\mathbb{R})$?**

1.6.7 **Das Banach-Tarski-Paradoxon**

1.6.8 **Wichtige, nicht behandelte Probleme**

1.6.1 $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$

Mutti hat Ihnen mal gesagt, dass Sie nicht "durch 0 teilen" dürfen. Auch war es Ihnen verboten, $\infty + \infty = \infty$ zu schreiben, obschon Sie dies immer gereizt hätte. Hier dürfen Sie solch schlimme Sachen endlich machen - vorausgesetzt, Sie beachten einige wenige Regeln. Wir erlauben dies, weil sich dann einige Sätze eleganter formulieren lassen. Ayatollah's aus der reinen Mathematik sei versichert: unteres ist ganz koscher.

$$x + y := \begin{cases} \infty & \text{falls} \quad x \in \mathbb{R} \quad \text{und} \quad y = \infty \quad \text{oder umgekehrt} \\ \infty & \text{falls} \quad x = y = \infty \\ -\infty & \text{falls} \quad x \in \mathbb{R} \quad \text{und} \quad y = -\infty \quad \text{oder umgekehrt} \\ -\infty & \text{falls} \quad x = y = -\infty. \end{cases}$$

$$xy := \begin{cases} \infty & \text{falls} \quad x > 0 \quad \text{und} \quad y = \infty \quad \text{oder umgekehrt} \\ \infty & \text{falls} \quad x < 0 \quad \text{und} \quad y = -\infty \quad \text{oder umgekehrt} \\ \infty & \text{falls} \quad x = y = \infty \quad \text{oder} \quad x = y = -\infty \\ -\infty & \text{falls} \quad x > 0 \quad \text{und} \quad y = -\infty \quad \text{oder umgekehrt} \\ -\infty & \text{falls} \quad x < 0 \quad \text{und} \quad y = \infty \quad \text{oder umgekehrt} \\ -\infty & \text{falls} \quad x = \infty \quad \text{und} \quad y = -\infty \quad \text{oder umgekehrt} \\ 0 & \text{falls} \quad x = 0 \quad \text{oder} \quad y = 0. \end{cases}$$

$$\frac{x}{y} := \begin{cases} \infty & \text{falls} \quad x > 0 \quad \text{und} \quad y = 0 \\ -\infty & \text{falls} \quad x < 0 \quad \text{und} \quad y = 0 \\ 0 & \text{falls} \quad x \in \mathbb{R} \quad \text{und} \quad y = \pm\infty. \end{cases}$$

$\infty - \infty$ dürfen Sie nach wie vor nicht machen; ebenso nicht $\pm\infty$ durch $\pm\infty$ teilen.

Falls Sie Schwierigkeiten haben, sich etwas unter $-\infty$ und $+\infty$ vorzustellen, ersetzen Sie einfach $-\infty$ durch "Velo" und $+\infty$ durch "Maschendrahtzaun". Es geht topologisch genau so gut - aber $-\infty$ und $+\infty$ sind anschaulicher.

Das Massproblem und seine Paradoxien

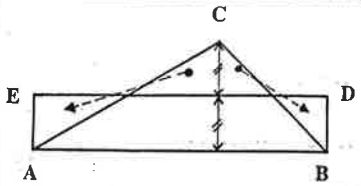
Von Catherine Bandle

Die Bestimmung der Flächen und Volumina von geometrischen Figuren hat zu grundlegenden mathematischen Theorien geführt: Integralrechnung, abstrakte Masstheorie, Wahrscheinlichkeitsrechnung. Sie hat auch zu unglaublichen Ergebnissen geführt wie das Banach-Tarski-Paradoxon, das besagt, dass sich eine Kugel so zerlegen und wieder zusammensetzen lässt, dass daraus zwei Kugeln vom selben Volumen wie das der ursprünglichen entstehen.

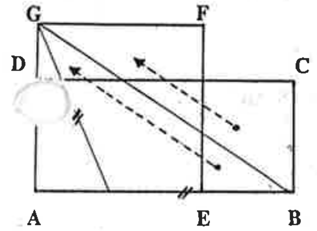
Der Inhalt von elementaren Figuren

Das Problem, den Inhalt einer geometrischen Figur in der Ebene oder im Raum zu messen, beschäftigt die Mathematik seit ihren Anfängen. Die Länge einer Strecke ist bestimmt durch den Abstand ihrer Endpunkte; das Rechteck mit den Seiten a und b hat den *Flächeninhalt* ab, und durch *Zerlegung* (Abb.1) lassen sich auch Inhalte von Dreiecken und Polygonen bestimmen.

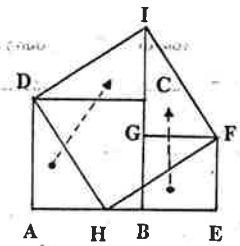
Im Raum ist die *Volumenberechnung* bereits bei einfachen Körpern wesentlich schwieriger. Bekanntlich ist es nicht möglich, ein Tetraeder



Dreieck ABC in Rechteck ABDE



Rechteck ABCD in Quadrat AEFG



Quadrate ABCD und BEFG zu Quadrat HFID

Abb. 1. Durch Zerlegung lassen sich Flächeninhalte von Dreiecken und Polygonen bestimmen.

in Polyeder zu zerlegen und aus diesen durch geschicktes Umordnen einen Würfel zu bilden. Die Frage nach einer solchen Zerlegung wurde 1900 v[illegible] *Hilbert* anlässlich des internationalen [illegible]natikerkongresses in Paris aufgeworfen [illegible] kurz danach von *Dehn* im negativen Sinn beantwortet.

Die Inhalte von Figuren, die von krummen, jedoch glatten Flächen oder Linien begrenzt werden, können mit Hilfe der sehr leistungsfähigen Methode der *Infinitesimalrechnung* berechnet werden. Die Grundidee, die dieser Methode zugrunde liegt, wurde bereits von *Archimedes* 287–212 v. Chr. benützt. Sie beruht darin, dass die Figur durch elementare Mengen ausgeschöpft wird. Archimedes führte dieses Verfahren für den Kreis und den Parabelabschnitt durch (Abb. 2). Mit dem Kalkül der Integration lassen sich die Inhalte der meisten in der Praxis auftretenden Mengen berechnen oder numerisch beliebig genau approximieren.

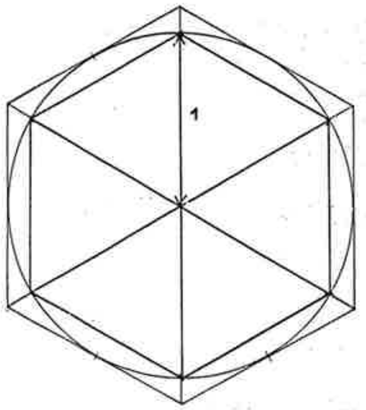


Abb. 2. Vergleich der Kreisfläche durch ein- und umgeschriebenes Sechseck ergibt für den Einheitskreis die Abschätzung $3<\pi<3.46$. Je feiner die Einteilung, desto genauer der Näherungswert.

Mit den grossen Fortschritten in der Funktionentheorie und der Funktionalanalysis am Ende des letzten Jahrhunderts wurde es notwendig, auch sehr allgemeinen Mengen einen Inhalt zuzuordnen.

Abstraktes Mass

Bei der in Abb. 3 abgebildeten Menge ist es naheliegend, ihr als Inhalt den Wert der *unendlichen Reihe* $F=1+(1/2)^2+(1/4)^2+(1/8)^2+\ldots=4/3$ zuzuschreiben. Während wir uns in diesem Fall auf die Anschauung stützen konnten, versagt diese im nächsten Beispiel gänzlich.

Wie gross ist der Inhalt der Menge R_1, die aus den rationalen Zahlen, d. h. denjenigen Zahlen, die sich als Quotient zweier ganzer Zahlen darstellen lassen, im Intervall (0,1) besteht? Ohne Präzisierung des Inhaltsbegriffes lässt sich diese Frage nicht beantworten.

Um den Inhalt einer Menge – heute ist es üblicher, vom *Mass* einer Menge zu sprechen – im mathematischen Sinn zu definieren, müssen wir uns zunächst im klaren sein, welchen Forderungen er genügen soll. Der Einfachheit halber beschränken wir uns auf Punktmengen auf Geraden, in der Ebene und im Raum. Wenn er unserer naiven Vorstellung von Länge, Fläche und Volumen bei elementargeometrischen Figuren entsprechen soll, muss er folgende Eigenschaften haben:

(E–0) Das Mass einer Menge A – kurz m(A) – ist eine nichtnegative Zahl.

(E–1) Das Mass ist dimensionsabhängig. Das 1-dimensionale Mass der Einheitsstrecke ist 1. Entsprechend ist das 2-(3-)dimensionale Mass des Einheitsquadrates bzw. -würfels 1.

(E–2) Sind zwei Mengen A und B deckungsgleich, so haben sie dasselbe Mass.

(E–3) Ist die Menge A die Vereinigung von höchstens abzählbaren, paarweise punktfremden Mengen A_k $(k=1, 2, 3\ldots)$, so ist ihr Mass die Summe der Masse von A_k, d. h. $m(A)=m(A_1)+m(A_2)+m(A_3)+\ldots$

Namhafte Mathematiker haben sich mit der Konstruktion von Massen auseinandergesetzt, darunter *Jordan, E. Borel* und *Lebesgue*. Lebesgue entwickelte sein Mass im Zusammenhang mit einer genialen Integrationstheorie, die zu den *Eckpfeilern* der modernen *höheren Analysis* gehört und die sich auch in den Anwendungen wie zum Beispiel in der Quantenmechanik bewährt hat. Es stellt sich heraus, dass das Lebesguesche Mass der zu Beginn dieses Abschnittes betrachteten Menge R_1 gleich *null* ist. Erstaunlich ist die Tatsache, dass die Menge aller rationalen Zahlen das Lebesguesche Mass Null hat. Das liegt daran – wie *Cantor* gezeigt hat –, dass diese Zahlen numeriert werden können und dass ihre Gesamtheit vernachlässigbar klein ist, verglichen mit allen übrigen Zahlen. Lebesgue war es gelungen, sehr allgemeine Mengen zu messen.

Die klare Einsicht in die Notwendigkeit, den elementaren Massbegriff seiner Dinglichkeit zu entkleiden und nur die Grundregeln genau zu beschreiben, hat sich als eine der fruchtbarsten Wendungen in der modernen Masstheorie erwiesen. Dieser Weg wurde auch in der *Wahrscheinlichkeitsrechnung* begangen, wo es darum geht, Ereignissen ein Mass für ihre Wahrscheinlichkeit zuzuordnen.

Nichtmessbare Mengen

Bald stellte sich die Frage nach der Existenz eines Masses, das den Bedingungen (E–0)–(E–3) genügt und auf *allen* Mengen definiert ist. *Vitali* war der erste, der zeigen konnte, dass dies sogar auf der Geraden unmöglich ist. Er konstruierte eine nichtmessbare Menge, indem er wesentlich von der Forderung (E–3) Gebrauch machte.

Schwächt man diese Forderung ab, indem man sie nur auf *endlich* viele Mengen bezieht, so kann man erneut fragen, ob es ein Mass gibt, das auf *allen* Mengen definiert ist und den Bedingungen (E–0)–(E–3) in der abgeschwächten Form genügt. *S. Banach* hat gezeigt, dass es auf der Geraden und in der Ebene unendlich viele solcher Masse gibt, während im Raum – wie *F. Hausdorff* bewies – keines vorhanden ist. Beide Aufgaben sind mit Fragen der *Grundlagen der Mathematik* verknüpft.

Die Vitalische Konstruktion und das Auswahlaxiom

Um eine nichtmessbare Menge zu konstruieren, denken wir uns die Strecke der Länge Eins zu einem Kreis zusammengebogen. Zu jedem Punkt x auf der Kreislinie betrachten wir die Menge aller Punkte, die man erhält, wenn man x um den Winkel $2\pi r$ (Bogenmass) dreht, wobei *r eine rationale Zahl* zwischen null und eins ist. Die ganze Kreislinie zerfällt so in Klassen, die keine gemeinsamen Punkte besitzen.

Aus jeder Klasse wählen wir einen Repräsentanten, welche zusammen die Klasse Z bilden. Wir numerieren die rationalen Zahlen $r_1, r_2, r_3 \ldots$ und betrachten die Klassen n, die durch Drehung von Z um den Winkel $2\pi r_i$ entstehen. Die Kreislinie wird nun eingeteilt in eine abzählbare Folge von Mengen $Z, Z_1, Z_2, Z_3 \ldots$, die paarweise disjunkt sind, und jeder Punkt des Kreises gehört einer dieser Klassen an. Ferner sind alle Mengen $Z, Z_1, Z_2 \ldots$ *deckungsgleich*. Gäbe es ein Mass, das auf Z definiert ist, so müsste deswegen gelten: $m(Z)=m(Z_i)$ für alle $i=1, 2, 3\ldots$ Wegen (E–3) wäre $1=m(Z)+m(Z_1)+m(Z_2)+\ldots$

Da es unendlich viele dieser Klassen gibt, erhalten wir einen *Widerspruch*. Die Annahme, dass m(Z) existiert, war infolgedessen falsch. Die Menge Z kann insofern nicht konstruiert werden, als es keine Maschine gibt, die die Auswahl eines Repräsentanten vornimmt. Die Klassen sind nicht explizit genug, um ein Verfahren vorschreiben zu können, das die Auswahl besorgen könnte.

Anfang des 20. Jahrhunderts wurde diese Konstruktion von *Zermelo* legitimiert, indem er ein *Auswahlaxiom* einführte: *Zu jedem Mengensystem, das nur nichtleere, paarweise disjunkte Mengen M enthält, existiert eine Auswahlmenge A, die von jedem M aus dem Mengensystem genau ein Element m aus M enthält.*

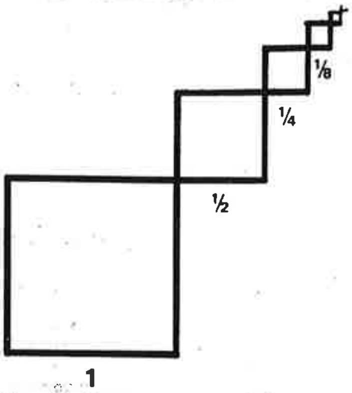


Abb. 3. Die Menge besteht aus unendlich vielen Quadraten, deren Seiten jeweils um die Hälfte verkürzt sind.

Das Axiom scheint einleuchtend und wird auch nur von denjenigen Mathematikern angefochten, die reine *Existenzaussagen* ablehnen und nur konstruktive Aussagen zulassen. In vielen Beweisen der klassischen Analysis wurde dieses Axiom schon längst benützt, ohne dass es als Grundlage des Schliessens ausdrücklich formuliert worden war. Und dennoch führt es zu Paradoxien, die Zweifel über seine Zulässigkeit aufkommen lassen.

Banach-Tarski-Paradoxien

Unter Verwendung des Auswahlaxioms und gestützt auf Untersuchungen von *Hausdorff*, haben *Banach* und *Tarski* das folgende *verblüffende Ergebnis* erhalten: Eine Kugeloberfläche lässt sich derart in eine endliche Anzahl von paarweise disjunkten Mengen $A_1, A_2 \ldots A_n$ zerlegen, dass diese bei geeigneter Drehung die Kugeloberfläche zweimal überdecken. Daraus lässt sich ohne allzu grosse Schwierigkeiten schliessen, dass eine Orange so in endlich viele paarweise disjunkte Teile zerlegt werden kann, dass aus diesen durch Umordnen 1000 neue Orangen derselben Grösse entstehen.

Es ist klar, dass diese Mengen äusserst kompliziert sind und dass die Zerlegung nicht mit dem Messer durchführbar ist. Sonst hätte dieser *mathematische Sachverhalt* verheerende Folgen für die *Orangenproduktion*. Ferner können die Mengen wegen der Eigenschaft (E–3) nicht messbar sein.

Diese *paradoxen* Resultate sind nicht falsch im logischen Sinn, sie scheinen uns nur unglaubwürdig, wenn wir von den für die Körper der Elementargeometrie gültigen Eigenschaften ausgehen. Das Auswahlaxiom muss deswegen nicht aufgegeben werden. Wenn wir es als gültige Regel anerkennen, bleibt uns *keine andere Wahl*, als Sätze, die unserer Intuition vollkommen entgegengesetzt sind – wie diejenigen von Banach und Tarski –, in Kauf zu nehmen.

Zerlegungen in der Ebene

In der *Ebene* sind «paradoxe» Zerlegungen *nicht möglich*. Banach zeigte, dass bei jeder Umordnung eine Menge vom gleichen Mass entsteht. In diesem Gebiet der Mathematik gibt es eine Reihe von interessanten *Problemen*, die teilweise bis heute *ungelöst* sind. Eine hübsche Frage stammt von *Tarski* (1925): Kann ein Kreis in endlich viele Teilmengen zerlegt werden, welche zu einem Quadrat der gleichen Fläche umgeordnet werden können? Dieses Problem ist nicht zu verwechseln mit der *Quadratur des Kreises*, bei der es darum geht, mit Zirkel und Lineal ein Quadrat zu konstruieren, das denselben Inhalt wie ein vorgegebener Kreis aufweist. 1882 hat *Lindemann* gezeigt, dass eine solche Konstruktion nicht durchführbar ist.

Kürzlich wurde von einem ungarischen Mathematiker eine Lösung des *Tarski*-Problems vorgeschlagen. Bevor sie in das *mathematische Allgemeingut* aufgenommen wird, muss sie sich jedoch einer gründlichen Prüfung der Spezialisten unterziehen.

Literatur: S. Wagon: The Banach-Tarski Paradox, Cambridge University Press 1985. R. M. French: The Banach-Tarski Theorem, The Mathematical Intelligencer Vol. 10, No. 4, 1988, 21–28.

Adresse der Autorin: Mathematisches Institut, Universität Basel, Rheinsprung 21, 4051 Basel.

Verantwortlich für «Forschung und Technik»
Lucien Trueb

Der höchste Berg ist ein Abfallberg

tr. Das *Everest-Basislager* auf einer Höhe von 5334 m ü. M. wird seit Jahrzehnten von allem Expeditionen zum höchsten Berg der Welt benutzt. Es ist auch das Ziel zahlreicher Trekkergruppen. Aus diesem Grund wurde es mit der Zeit zu einer «wilden Kehrichtdeponie», wo sich Tonnen von Konservendosen, Flaschen, Plasticfolien, Papier und Medikamenten ansammelten. Vor kurzem wurde dieser Abfallberg von Mitgliedern des britischen *Survival Club* abgetragen. Diese Organisation setzt sich weltweit aktiv für den Schutz der Umwelt ein. Nachdem sich die Leute zwei Wochen lang in Nepal akklimatisiert hatten, stiegen sie zum Basislager auf und säuberten dieses gründlich. Alles brennbare Material wurde verbrannt, der Rest wurde in grossen Kunststoffsäcken verstaut und in eine tiefe Felsspalte geworfen. Ein grösserer Vorrat solcher Säcke wurde zurückgelassen, damit die Bergsteiger von jetzte an ihr Lager selbst entsorgen können.

1.6.2 Allgemeine Masse

Wir kommen jetzt zu einer Verallgemeinerung des Konzeptes der Wahrscheinlichkeit, zu den Massen. Masse sind nichtnegativ und σ-additiv; hingegen muss das Mass nicht 1 sein; nicht mal endlich. Despektierlich ist die Wahrscheinlichkeit ein Spezialfall der Masstheorie, bei der das Mass endlich, genauer von Mass 1 ist. Aber diese Sicht ist polemisch, despektierlich und vor allem ignorant.

Definition 1.24 [**Mass**] *Sei E eine Menge und $\mathcal{E}$ eine σ-Algebra auf E. Dann definieren wir:*

a) $(E, \mathcal{E})$ heisst Messraum; die Mengen aus $\mathcal{E}$ nennen wir messbare Mengen

b) Ein Mass μ auf $(E, \mathcal{E})$ ist eine Mengenfunktion $\mu : \mathcal{E} \rightarrow \overline{\mathbb{R}}_+$ derart, dass $\mu(\phi) = 0$ und wir verlangen auch, dass μ σ-additiv ist:

$$\mu(\cup_n A_n) = \sum_n \mu(A_n),$$

wo die Folge A_n disjunkt aus $\mathcal{E}$.

zB.: $\mu_2(\emptyset)=0$, $\mu_2(A)=\infty$ $\forall A \in \mathcal{E} \setminus \{\emptyset\}$

$\mu_1(A)=0$ $\forall A \in \mathcal{E}$

Lebesgue, P

c) μ ist endlich wenn $\mu(E) < \infty$.

d) μ ist σ-endlich, wenn eine aufsteigende Folge E_i aus $\mathcal{E}$ existiert derart, dass $\cup E_i = E$ und $\mu(E_i) < \infty$ für alle $i \geq 1$.

e) Das Tripel $(E, \mathcal{E}, \mu)$ bezeichnen wir als Massraum.

Wir sehen sofort, dass unsere Wahrscheinlichkeitsräume immer auch Massräume sind (vgl Definition 1.7). Die endlichen Massräume sind insofern nahe verwandt mit Wahrscheinlichkeiten, als dass jedes endliche Mass μ mit Hilfe einer Wahrscheinlichkeit P geschrieben werden kann:

$$\mu(A) = \mu(E)P(A).$$

$\rightarrow P[A] = \frac{\mu(A)}{\mu(E)}$; „Anteil"

Es gibt vor allem ein zentral wichtiges, unendliches Mass, welches wir für diese Vorlesung brauchen. Wenn wir dieses haben, können wir weitere nicht-triviale Beispiele zu Definition 1.24 anschauen und den obigen Begriffen ein bisschen Leben einhauchen. Es handelt sich dabei um das Lebesgue-Mass.

1.6.3 Lebesgue-Mass

In der Analysis wird das Lebesgue-Mass eingeführt. Die saubere Einführung des Lebesgue-Masses dauert mehrere Stunden. Deshalb verzichten wir in dieser Vorlesung darauf. Es ist auch so, dass man durch die saubere Einführung des Lebesgue-Masses nicht unbedingt besser damit zu rechnen versteht...

Artikel NZZ Bandle

Wir kennen den Messraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Der folgende Satz garantiert uns ein Mass λ auf diesem Messraum, der dadurch zum Massraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ wird. Das Mass λ, dessen Existenz dort garantiert wird, *erweitert* unseren bisherigen Begriff der Länge eines Intervalls. Auf normalen Intervallen $I = [a, b]$ mit $a \leq b$ gilt $\lambda(I) = b - a$. a darf übrigens $-\infty$ sein und b darf ebenso $+\infty$ sein. Die Länge wird dann $+\infty$.

Satz 1.25 [Existenz des Lebesgue-Masses λ] *Auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ existiert ein eindeutiges σ-endliches Mass λ - das Lebesgue-Mass - derart, dass für jedes Intervall $I := [a, b]$, $a \leq b$ gilt: $\lambda(I) = b - a$.*

warum Lebesgue σ-endlich?

Wir haben in Definition 1.13 die P-Nullmengen kennengelernt. Analog definiert man jetzt

Definition 1.26 [Lebesgue-Nullmengen] *Eine Menge $A \in \mathcal{B}(\mathbb{R})$ heisst Lebesgue-Nullmenge, wenn $\lambda(A) = 0$.*

Bemerkungen zu **fs/as (WT)** und **fü/ae (Analysis)**

almost everywhere

fast überall

fast sicher almost surely

$P[A] = 1$

$\lambda(\mathbb{R}) = \infty$; f.ü. falls $\exists A \in \mathcal{B}(\mathbb{R})$: Eigenschaft gilt überall ausser auf A wo $\lambda(A) = 0$ (zB 2 fkt gleich)

man muss nichts unterschiedlich formulieren

andere Bezeichnung für analoge Sachverhalte

FS12

Lemma 1.27 [$\mathbb{Q}$ ist eine Lebesgue-Nullmenge]

Beweis Lemma 1.27

∘ Frage an Publikum: wie geht Beweis?

∘ $\mathbb{Q}$ abzählbar $\Longrightarrow (r_i)_{i\geq 1}$

∘ $\forall \varepsilon > 0$: $r_i \in \left(r_i - \frac{\varepsilon}{2^{i+1}}, r_i + \frac{\varepsilon}{2^{i+1}}\right]$; d.h.: $\pm\frac{\varepsilon}{4}, \pm\frac{\varepsilon}{8}, \pm\frac{\varepsilon}{16}$ & jeweils $\times 2$

∘ $$\lambda[\mathbb{Q}] = \lambda\Big[\bigcup_{i=1}^{\infty}\{r_i\}\Big] \underset{(1.24\,D)}{=} \sum_{i=1}^{\infty}\lambda[\{r_i\}] \overset{(HA)}{\leq} \sum_{i=1}^{\infty}\lambda\Big[\Big(r_i - \frac{\varepsilon}{2^{i+1}}, r_i + \frac{\varepsilon}{2^{i+1}}\Big]\Big]$$
$$= \sum_{i=1}^{\infty}\varepsilon\left(\frac{1}{2}\right)^i = \varepsilon\,;\ \varepsilon > 0 \text{ frei gewählt} \rightarrow \lambda[\mathbb{Q}] = 0 \qquad \text{qed}$$

∘ Nebenbemerkungen:

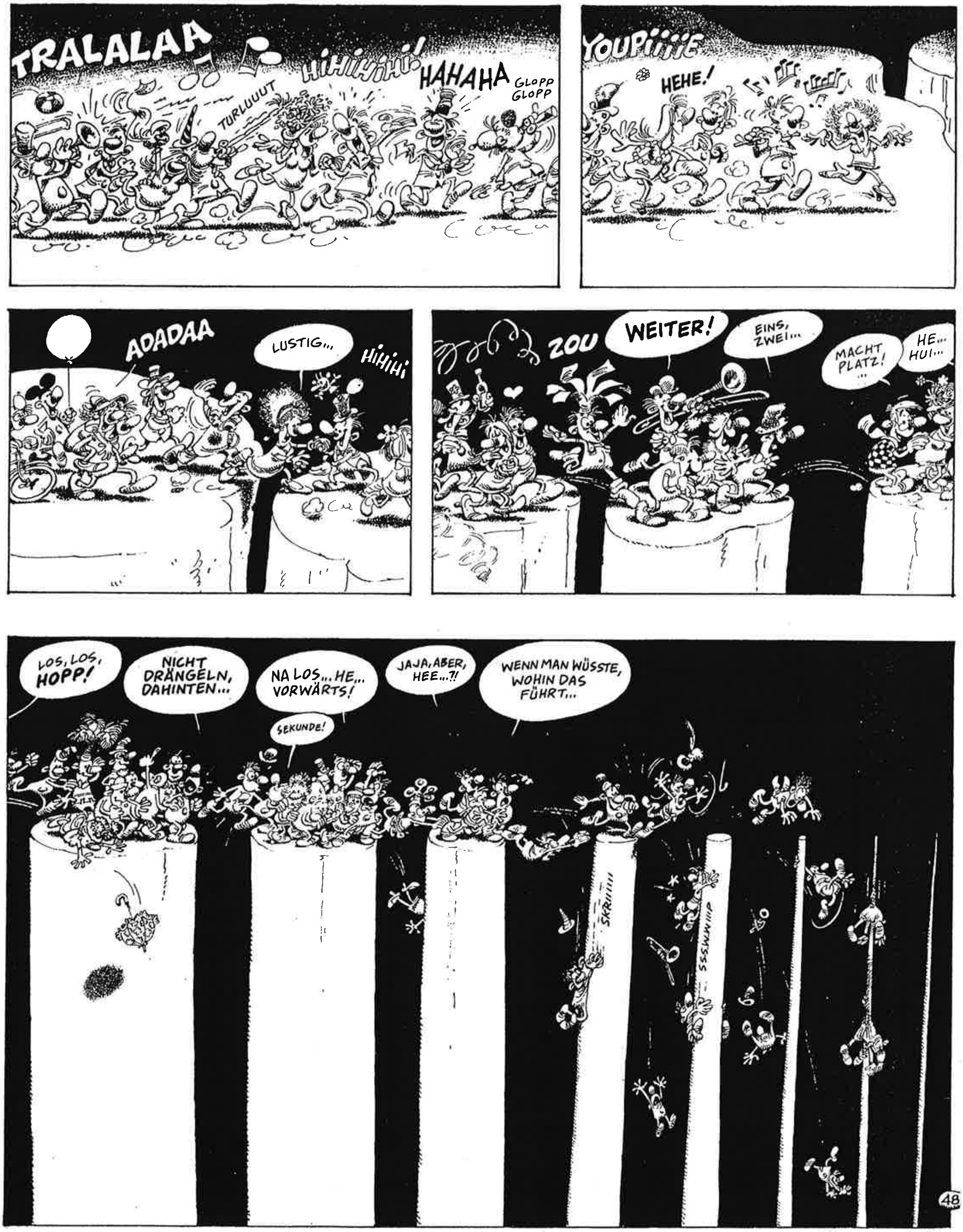
- NB I: $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$
- NB II: Korollar, gilt $\forall$ abzählbaren Mengen ($\mathbb{N}$, $\mathbb{Z}$)
- NB IV: „$\mathbb{R}$ besteht fast überall aus Irrationalzahlen", sogar:
- NB III: $\mathbb{Q}$ ist abzählbar & damit Lebesgue-Nullmenge & dicht in $\mathbb{R}$; damit kann man viel in WT & Analysis beweisen 
- NB V: Menge der alg. Zahlen ebenfalls abzählbar & damit Lebesgue-Nullmenge (d.h.: transzendente Zahlen in $[0,1]$ haben Maß 1; einzige nicht algebr. Zahlen, welche allg. bekannt: e & π)

Fangfrage zum Beweis von Lemma 1.27: $\mathbb{Q}$ ist dicht in $\mathbb{R}$; haben wir damit nicht auch FS14 bewiesen, dass $\mathbb{R}$ eine Lebesgue-Nullmenge ist? (d.h.: $\forall$ ε-Umgebung von jedem $x \in \mathbb{R}$ gilt: $\exists\, q \in \mathbb{Q}: q \in (x-\varepsilon, x+\varepsilon)$)

∘ Noch ein „Argument": Jede Irrationalzahl lässt sich durch Folge von rationalen Zahlen approximieren: $\forall x \in \mathbb{R}\ \exists\, (q_i)_{i=1}^{\infty} \in \mathbb{Q}: \lim_{i\to\infty} q_i = x$

∘ Aber ich habe im Beweis nur 1 Folge zur Verfügung & kann damit nicht gleichzeitig alle irrationalen Zahlen x_0 ~~approximieren~~. abzudecken!

∘ Es gilt ja auch in der Tat: $\lambda(\mathbb{R}) = +\infty$, nämlich $+\infty - (-\infty) = \infty$ (vgl. 1.6.1)



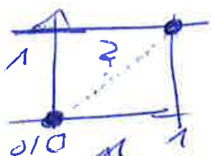
TRALALAA
TURLUUUT
HIHIHIHI!
HAHAHA
GLOPP GLOPP
YOUPIIIIE
HEHE!
ADADAA
LUSTIG...
HIHIHI
ZOU
WEITER!
EINS, ZWEI...
MACHT PLATZ! ...
HE... HUI...
LOS, LOS, HOPP!
NICHT DRÄNGELN, DAHINTEN...
NA LOS... HE... VORWÄRTS!
SEKUNDE!
JAJA, ABER, HEE...?!
WENN MAN WÜSSTE, WOHIN DAS FÜHRT...
SKRIIII
SSSWWIIP
48

Expedition in eine Cantormenge

1.6.4 Singulär stetige Wahrscheinlichkeit auf $\mathbb{R}$ - Cantorsches Diskontinuum

Wir haben bisher die beiden Typen von Wahrscheinlichkeiten "diskret" und "absolut stetig" kennengelernt ("absolut stetig" war in der einführenden Vlsg WTS einfach "stetig"). Es war uns intuitiv sofort klar, dass man auch konvexe Linearkombinationen dieser Wahrscheinlichkeiten bilden kann. Jetzt kommt ein dritter Typ, dessen Existenz nicht offensichtlich ist:

Definition 1.28 [singulär stetige Wahrscheinlichkeiten] *Sei P eine Wahrscheinlichkeit und F_P deren Verteilungsfunktion. Falls F_P stetig ist und die Menge der Wachstumspunkte von F_P Lebesgue-Mass 0 haben, nennen wir P singulär stetig.*



Bemerkungen zu **Wachstumspunkte von F_P**

- x ist Wachstumspunkt von $F_P \iff F(x+\varepsilon)-F(x-\varepsilon)>0 \ \forall\, \varepsilon>0$
- bei $\mathcal{N}(0,1)$: jeder Punkt ist WP; $\lambda(\mathbb{R})=\infty\neq 0$ (Bild)
- also: müssen von 0 nach 1 und zwar stetig, dürfen nur bei Nullmenge wachsen!

Man könnte Zweifel haben, dass so was überhaupt existiert. Zudem ist man versucht zu formulieren, dass abzählbare Mengen immer Lebesgue-Mass 0 haben (das stimmt sogar (unformuliertes Korollar zu Lemma 1.27)) und überabzählbare Mengen nicht mehr Mass 0 haben. Zum Kontrast führen wir jetzt das Cantorsche Diskontinuum ein:

- $\Omega=[0,1]$; nachfolgende Konstruktion mit farbigen Bildern:
- $J_{11} := (1/3,\ 2/3)$

  $J_{21} := (1/9,\ 2/9)$, $J_{22} := (7/9,\ 8/9)$

  $J_{31} := (1/27,\ 2/27)$, $J_{32} := (7/27,\ 8/27)$, $J_{33} := (19/27,\ 20/27)$, $J_{34} := (25/27,\ 26/27)$

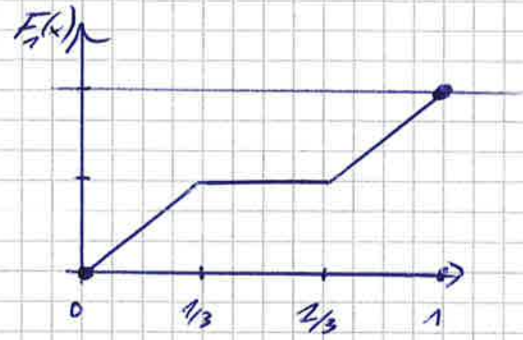
  $\vdots$
- nach m Schritten sind (2^m-1) Intervalle entfernt worden; es bleiben noch 2^m disjunkte Intervalle der Länge 3^{-m}.
- definieren jetzt (im Kurs leicht anders): $C_n := \left(\bigcup_{m=1}^{n}\bigcup_{k=1}^{2^{m-1}} J_{mk}\right)^c$ (Intervalle)

$$C := \left(\bigcup_{m=1}^{\infty}\bigcup_{k=1}^{2^{m-1}} J_{mk}\right)^c$$ gegenüber $[0,1]$

(was am Schluss übrig bleibt (fast nichts))

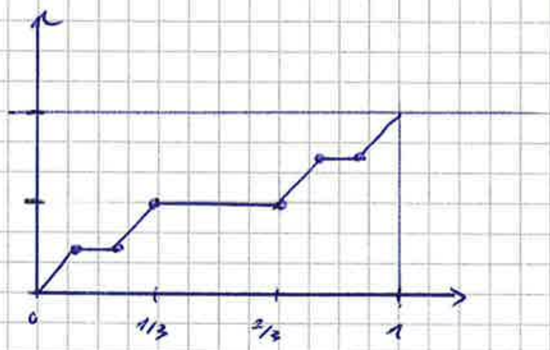
$$F_1(x) = \begin{cases} 0 & x = 0 \\ 1/2 & x \in (1/3, 2/3) \\ 1 & x = 1 \end{cases}$$

; Rest durch Interpolation! ($F_0 := \mathrm{id}$)



$$F_2(x) = \begin{cases} 0 & x = 0 \\ 1/4 & x \in (1/9, 2/9) \\ 1/2 & x \in (1/3, 2/3) \\ 3/4 & x \in (7/9, 8/9) \\ 1 & x = 1 \end{cases}$$

; Rest durch Interpolation!



⋮

$F_\infty(x) := \lim_{n \to \infty} F_n(x)$ punktweise definiert

F_∞ ist monoton wachsend, (glm) stetig (& damit rechtsstetig), (Ü. Ue) $F(\infty) = 1$ & $F(-\infty) = 0$ } ist Verteilungsfkt. wegen Satz 1.17

• Menge der Wachstumspunkte hat Lebesgue-Maß 0 ⟹ C

∃ singulär-stetige Wahrscheinlichkeit!

• „Länge" von F_∞ = ? 2 ; wie?

• allg.: falls stetig, monoton von $(0,0) \to (1,1) \Rightarrow$ Länge $\in [\sqrt{2}, 2]$; falls f' stetig $\Rightarrow < 2$; limes Länge cantor $= 2$!

Welche Eigenschaften hat dieses Diskontinuum? ∘ $C \in \mathcal{B}(\mathbb{R})$; da „$\cup$" & „$c$"

∘ $\lambda[[0,1]] = 1 = \lambda[C \cup C^c] = \lambda[C] + \lambda[C^c]$ Def. 1.24 b)

∘ $\lambda[C^c] = \lambda\left[\bigcup_{m=1}^{\infty} \bigcup_{k=1}^{2^{m-1}} I_{mk}\right] \underset{\text{disjunkt}}{=} \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \frac{8}{81} + \dots = \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n =$

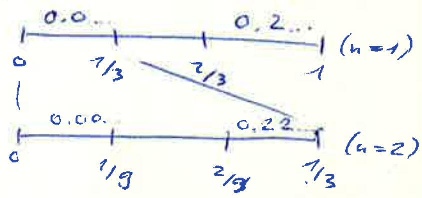
$= \frac{1}{3} \cdot \frac{1}{1-\frac{2}{3}} = 1 \implies \underline{\underline{\lambda[C] = 0}}$ (C hat Lebesgue-Maß 0)

∘ C^c ist offen, da abzählbare Vereinigung offener Mengen (elementare top.)
$\implies$ <u>C abgeschlossen</u>

∘ <u>C ist überabzählbar:</u> [illegible]
∘ rep.: binäre Entwicklung von $[0,1]$: $\{y \mid y = \sum_{n=1}^{\infty} \beta_n \left(\frac{1}{2}\right)^n ; \beta_n \in \{0,1\}\}$
∘ jetzt <u>triadische</u> Entwicklung:
∘ $C = \{x \mid x = \sum_{n=1}^{\infty} \alpha_n \cdot \frac{1}{3^n} ; \alpha_n \in \{0,2\}\} =: R$ Randpunkte von C

z.B.: $1/3 = 0.1_{(3)} = 0{,}0222\overline{2}_{(3)}$
$2/3 = 0.2_{(3)} = 0{,}2000\overline{0}_{(3)}$

∘ 0 → links, 2 → rechts:



NB: $C \equiv R$ (brauchen nicht)
∘ $|R| = |\{0,2\}^{\mathbb{N}}| = |[0,1]| = |\mathbb{R}|$ $\implies$ überabzählbar

dazu:
∘ kein innerer Punkt.
∘ Jeder Punkt ist HP.

∘ <u>Vorsicht:</u> $|\mathbb{N}| < |C| = |[0,1]| = |\mathbb{R}|$; aber $\lambda(\mathbb{N}) = 0$, $\lambda(C) = 0$, $\lambda[[0,1]] = 1$, $\lambda(\mathbb{R}) = \infty$
– abzählbar $\implies \lambda(\cdot) = 0$ & damit $\lambda(\cdot) > 0 \implies$ überabzählbar; $\boxed{\text{Cantor-Fkt-Blatt}}$
– aber C überabz. Nullmenge!

1.6.5 Vollständige Klassifikation der Wahrscheinlichkeiten auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Wir geben noch ohne Beweis an:

Satz 1.29 [Vollständige Klassifikation der Wahrscheinlichkeiten] *Jede Verteilungsfunktion F kann als konvexe Linearkombination $F = aF_d + bF_a + cF_s$ dargestellt werden. Dabei sind F_d eine diskrete, F_a eine absolut stetige und F_s eine singulär stetige Verteilungsfunktion.*

~~∘ def. absolut stetig, wenn ... (1.29)~~
~~∘ def.: Lebesgue-Maß~~
~~Lebesgue~~
~~abs. stetig~~
~~Gegenbeispiel~~

Welche Eigenschaften hat dieses Diskontinuum?

1.6.5 Vollständige Klassifikation der Wahrscheinlichkeiten auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, ohne Beweis

Satz 1.29 [Vollständige Klassifikation der Wahrscheinlichkeiten] *Jede Verteilungsfunktion F kann als konvexe Linearkombination $F = aF_d + bF_a + cF_s$ dargestellt werden. Dabei sind F_d eine diskrete, F_a eine absolut stetige und F_s eine singulär stetige Verteilungsfunktion.*

Worin liegt die tiefere Bedeutung dieses Satzes? Definition 1.18 (diskrete Wahrscheinlichkeit) scheint einleuchtend (mit der Ausnahme, dass eine Wahrscheinlichkeit auf $\mathbb{Q}$ diskret ist (obschon $\mathbb{Q}$ dicht in $\mathbb{R}$)). Danach folgt jedoch die irritierend indirekte Definition von absolut stetigen Wahrscheinlichkeiten über die Verteilungsfunktion mit Hilfe des Lebesgue-Integrals! Man muss also das Lebesgue-Mass kennen und das Lebesgue-Integral, um die Definition von absolut stetigen Wahrscheinlichkeiten zu verstehen - ist das nicht Willkür, muss das so sein? Satz 1.29 ist dann jedoch so einfach und elegant, dass das wohl der kanonische Weg ist, den eine höhere Instanz vorgesehen hat!

FS10
12
14

1.6.6 Warum σ-Algebren? Warum P auf $\mathcal{B}(\mathbb{R})$ und nicht auf $\mathcal{P}(\mathbb{R})$?

Man fragt sich als jungeR StudentIn zu recht, weshalb wir diese σ-Algebren ~~so~~ einführen und nicht einfach ein P auf $\mathcal{P}(\mathbb{R})$ definieren. Dazu ein paar Gründe:

1. Es funktioniert so, wie wir es gemacht haben (siehe bisheriges Kapitel). Dies tönt defensiv-hilflos, in Anbetracht der Schwierigkeiten, welche sonst auftreten, ist es eine sehr gute Antwort.

2. In der Finanzmathematik (allgemein in einer Vorlesung "Stochastische Prozesse, Martingaltheorie") betrachtet man nicht nur einzelne Zufallsgrössen X (siehe Kapitel 2), sondern ganze sogenannte Stochastische Prozesse X_t. Bereits in der WTS haben wir am Rand darauf hingewiesen, dass Zufallsgrössen X nicht beliebige Funktionen $X : \Omega \to \mathbb{R}$ sind, sondern sogenannt messbare Abbildungen sein müssen (das Urbild muss in der σ-Algebra sein). Dies wird auch bei den Stochastischen Prozessen der Fall sein. Dort wird man aber nicht nur eine σ-Algebra haben, sondern eine ganze Folge von solchen σ-Algebren. Diese stehen in der Finanzwelt für die Informationsmenge - und die ist gerade dort sehr wichtig!

3. Der Hammer ist dann der folgende

gibt nichts zwischen $|\mathbb{N}|$ & $|\mathbb{R}|$

Satz 1.30 [von Banach und Kuratowski (1929)] *Unter Annahme der Gültigkeit der Kontinuumshypothese gibt es keine auf ganz $\mathcal{P}(\mathbb{R})$ definierte, σ-additive Funktion P so, dass $P[\mathbb{R}] = 1$ und für jedes $x \in \mathbb{R}$ gilt $P[\{x\}] = 0$.*

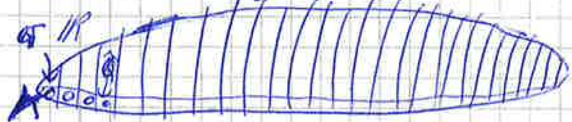
Damit scheiden die absolut stetigen Wahrscheinlichkeiten schon mal aus; diese geben einzelnen Punkten immer Wahrscheinlichkeit 0. Damit haben wir auf ganz $\mathcal{P}(\mathbb{R})$ auch keine Normalverteilung. Die Service-Vorlesungen in Statistik für IngenieurInnen, NaturwissenschaftlerInnen, OekonomInnen, SoziologInnen und PsychologInnen sind in diesem Punkt also regelmässig falsch. Hingegen wird es kaum jemals Probleme geben, da die Mengen zwischen $\mathcal{B}(\mathbb{R})$ und $\mathcal{P}(\mathbb{R})$ meines Wissens NIE in der Praxis auftreten.

Was, wenn wir die Kontinuumshypothese nicht annehmen wollen? Solange es zwischen $\mathbb{N}$ und $\mathbb{R}$ nur eine endliche Kaskade von verschiedenen Mächtigkeiten gibt, gilt ein analoger Satz.

$\overline{\mathcal{B}}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$; Vitali-Menge

(zu p. 34

- ∃ diverse analoge Beweise; hier ähnlich NZZ-Bandle
- $x, y \in \mathbb{R}$; $x \sim y \overset{d}{\Leftrightarrow} x - y \in \mathbb{Q}$ ist Äquivalenzrelation
- $\mathbb{R}$ zerfällt damit in Äquivalenzklassen der Art $(x + \mathbb{Q})$, wo $x \in \mathbb{R}$
- $x \sim y \Leftrightarrow x + \mathbb{Q} = y + \mathbb{Q}$
- $\forall$ Äquivalenzklasse $\exists$ Punkt $\in [0, 1]$



- Auswahlaxiom: nehme $K \subset [0, 1]$: K hat mit jeder Äquivalenzklasse genau 1 Element gemeinsam.

$\Rightarrow$

- I. $\mathbb{R} = \bigcup_{y \in \mathbb{Q}} (y + K)$; $(y+K) := \{y + u \mid u \in K\}$

nur 1 rationaler Vektor!

II. $\underset{0,2}{y_1} \neq \underset{0,3}{y_2} \Rightarrow (\underset{0,2}{y_1} + K) \cap (\underset{0,3}{y_2} + K) = \emptyset$ wenn $\underset{0,2}{y_1}, \underset{0,3}{y_2} \in \mathbb{Q}$ (sonst ↯)

- Zeigen jetzt: $K \notin \overline{\mathcal{B}}(\mathbb{R})$:

  - $\mathbb{Q}$ abzählbar

  III. - $+\infty = \lambda(\mathbb{R}) \overset{I}{=} \lambda\Big(\bigcup_{y \in \mathbb{Q}} (y + K)\Big) \overset{II}{=} \sum_{y \in \mathbb{Q}} \lambda(y + K)$

  IV. - λ translationsinvariant: $\lambda(y + K) = \lambda(K) \quad \forall y \in \mathbb{Q}$

  vgl. Vlg reelle Analysis

  - da $K \subset [0, 1]$

$$\bigcup_{\substack{y \in [0,1] \\ y \in \mathbb{Q}}} (y + K) \subset [0, 2]$$

$$\sum_{\substack{y \in [0,1] \\ y \in \mathbb{Q}}} \lambda(y + K) = \lambda\Big(\bigcup_{\substack{y \in [0,1] \\ y \in \mathbb{Q}}} (y + K)\Big) \leq \lambda\big([0,2]\big) = 2 < \infty$$

$\Rightarrow$ wegen IV.: $\lambda(K) = 0$ ↯ zu III. ; also

K hat kein Lebesgue-Maß ; vgl Satz 1.25;

$\Rightarrow K \notin \mathcal{B}(\mathbb{R})$ (& auch $\notin \overline{\mathcal{B}}(\mathbb{R})$).

□

1.6.7 Das Banach-Tarski-Paradoxon

Der folgende Satz benötigt im Beweis das Auswahlaxiom und sonst lediglich die akzeptierten Axiome der Mathematik. Dann ist der Satz mathematisch richtig, aber schwer nachvollziehbar (vgl. auch Artikel in der NZZ von Frau Prof. Bandle):

Satz 1.31 [Banach-Tarski-Paradoxon] *Sei K eine Kugel im $\mathbb{R}^3$. Dann existiert eine Zerlegung*

$$K = A_1 \cup A_2 \cup \ldots \cup A_m \cup B_1 \cup B_2 \cup \ldots \cup B_n$$

von K in paarweise disjunkte Teilmengen A_i, B_j derart, dass wir damit 2 neue Kugeln K gleicher Grösse zusammensetzen können:

$$K = A_1' \cup A_2' \cup \ldots \cup A_m'$$

und

$$K = B_1' \cup B_2' \cup \ldots \cup B_n',$$

wo A_i kongruent zu A_i' ist und B_j kongruent zu B_j'. Die A_i' bzw B_j' sind wieder disjunkt.

1.6.8 Wichtige, nicht behandelte Probleme

"Sandwich"

1. Man kann sich fragen, ob es zwischen $\mathcal{B}(\mathbb{R})$ und $\mathcal{P}(\mathbb{R})$ etwas "relevantes" gibt? Die Antwort ist klar JA: Wir haben in Satz 1.25 die Existenz des Lebesgue-Masses lediglich auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ garantiert erhalten. Wir sagen jetzt, dass eine Menge Λ zum System $\overline{\mathcal{B}}(\mathbb{R})$ gehört, falls 2 Borel-Mengen A, B derart existieren, dass $A \subset \Lambda \subset B$ mit $\lambda(B \backslash A) = 0$. Das System $\overline{\mathcal{B}}(\mathbb{R})$ heisst das System der *Lebesgue-Mengen* und ist eine σ-Algebra (kleine Übungsaufgabe). Damit kann man also das Lebesgue-Mass natürlich auf $\overline{\mathcal{B}}(\mathbb{R})$ fortsetzen; man spricht dann von einer *Vervollständigung* von λ. Wir haben damit folgende Kaskade von Systemen:

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

Vitali

Die im Artikel von Frau Bandle erwähnte Vitali-Menge ist ein Grund für das letzte "$\subsetneq$" (Beweis in Vlsg falls Zeit). Wir werden $\overline{\mathcal{B}}(\mathbb{R})$ in dieser Vorlesung kaum benutzen, aber in der höheren Stochastik und Masstheorie ist es notwendig, sich damit auseinanderzusetzen.

Mächtigkeit $= 2^{|\mathbb{R}|}$

Mächtigkeit $= |\mathbb{R}|$

- Cantor: C
- $\mathcal{P}(C)$: alle Maß 0, $\in \overline{\mathcal{B}}(\mathbb{R})$
- $|C| = |\mathcal{B}(\mathbb{R})| = |\mathbb{R}| < |\mathcal{P}(C)|$

$\Rightarrow$ echte Teilmenge $\to$ muss Elemente in $\mathcal{P}(C)$ geben, welche $\neg$ in $\mathcal{B}(\mathbb{R}) \Rightarrow$ in $\overline{\mathcal{B}}(\mathbb{R})$

Algebra

2. Wir haben - nicht nur beim Lebesgue-Mass - die Existenzfrage von Massen und vor allem Wahrscheinlichkeiten ausgeklammert. Dabei ist es meist einfach, Wahrscheinlichkeiten auf einfachen Systemen zu definieren und deren Existenz und Vereinbarkeit mit den Axiomen der Wahrscheinlichkeit zu beweisen. Dass diese Wahrscheinlichkeiten dann aber zum Beispiel sinnvoll auf ganz $\mathcal{B}(\mathbb{R})$ erweitert/fortgesetzt werden können - das ist langwierig. Die Beweise (Fortsetzungssätze) gehören in eine Vorlesung über Masstheorie.