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:

2) Abzählbare Mengen:

3) $\mathbb{R}$ und $\mathbb{R}_+ := [0, \infty)$:

4) Endliche kartesische Produkte (Replika):

5) Unendliche kartesische Produkte (Replika):

6) Funktionen:

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:

$$\limsup_n A_n := \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$$

und

$$\liminf_n A_n := \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$$

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

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

$$\lim_n A_n = A,$$

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?).

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 $A_1, A_2 ...$
$\liminf_n A_n$	$\bigcup_{k=1}^\infty \bigcap_{n=k}^\infty A_n$ / Ereignis, bestehend im Eintreten aller Ereignisse $A_1, A_2 ...$, 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 Indikatorfunktionen herstellen können, dürfen wir vieles auf der Ebene von Funktionen erledigen statt auf der Ebene von Mengen. Die 1 zu 1 Beziehung ist dann 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?

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

Überlegen Sie sich jetzt, dass

$$\lim_{n\to\infty} A_n = A$$

genau dann wenn punktweise gilt

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

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?

2. Welches ist die kleinste σ-Algebra überhaupt?

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

4. Welches ist die kleinste σ-Algebra, welche Ereignis A enthält (von A erzeugte σ-Algebra)?

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.

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?

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

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.

Ein Beispiel eines Dynkin-Systems:

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 $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$ (deshalb "Monoton-System").

Beweis Lemma 1.3

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

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* $A \cap B \in \mathcal{C}$.

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

Beweis Satz 1.5

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 $\mathcal{C}$ ein π-System. Dann gilt:*

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

Beweis Satz 1.6

2 Bemerkungen zur Bedeutung dieses Satzes:

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. In Vorlesung und Übungen sei Ω immer nichtleer (später auch die Grundmenge bei allgemeinen Massen) - ausser wir sprechen es explizit an.

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

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* $A \subseteq B \Rightarrow P[A] \leq P[B]$.

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

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 $P[A_n] \uparrow P[A]$,

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

Als Vorbereitung auf den kommenden Satz: konvergiert

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

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

Satz 1.10 [Stetigkeit von P] *Es gelten*

$$P[\liminf_n A_n] \leq \liminf_n P[A_n] \leq \limsup_n P[A_n] \leq P[\limsup_n A_n]$$

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

Beweis von Satz 1.10

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

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

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

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

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

Aus der Vorlesung WTS kennen wir bereits die Zufallsgrössen, welche wir in Kapitel 2 intensiv studieren werden. Deren Verteilungsfunktionen liefern uns Wahrscheinlichkeiten auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, siehe Satz 1.17. 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 $P[A] = 0$.*

Ein paar kleine Bemerkungen:

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

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.*

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 $\mathcal{B}(\mathbb{R})$.*

Beweis Satz 1.15

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; damit existieren jeweils die Limiten von links und von rechts

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

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

Beweis von Satz 1.16

Wir *definieren* noch $F(\infty) := \lim_{t\to\infty} F(t)$ und $F(-\infty) := \lim_{t\to-\infty} F(t)$ - nach obigem sind diese Definitionen sinnvoll!

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

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

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".

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* $P = \sum_i p_i \delta_{t_i}$.

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

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

Beweis Satz 1.19

Bemerkung/Warnung zum Wort "diskret" in der WT und der restlichen Mathematik (zB diskrete Menge):

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:

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).
4) Man spricht deshalb auch von einer "Version" der Dichtefunktion (und wählt dann mit Vorteil zum Beispiel eine stetige Version).

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?
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:

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!

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:

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

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 $A \in \mathcal{A}$:*

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

Ein analoges Resultat gilt auch für eine endliche Partition.

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.

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

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

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

$\infty - \infty$ 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.

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 $\mu : \mathcal{E} \to \overline{\mathbb{R}}_+$ 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 $\mathcal{E}$.

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:

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

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...

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: $\lambda(I) = b - a$.*

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)**

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

Beweis Lemma 1.27

Fangfrage zum Beweis von Lemma 1.27: $\mathbb{Q}$ ist dicht in $\mathbb{R}$; haben wir damit nicht auch bewiesen, dass $\mathbb{R}$ eine Lebesgue-Nullmenge ist?

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

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:

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!

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 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

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.

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

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:

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

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.

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.