Klaus D. Schmidt

Maß und Wahrscheinlichkeit

$$\Big(E^{\mathcal{G}}(h \circ X)\Big)(\omega) = \int_{\mathbb{R}} h(x)\, dP_X^{\mathcal{G}}(x,\omega)$$

2., durchgesehene Auflage

Springer

Springer-Lehrbuch

Klaus D. Schmidt

Maß und Wahrscheinlichkeit

Zweite, durchgesehene Auflage

Springer

Prof. Dr. Klaus D. Schmidt
Technische Universität Dresden
Fachrichtung Mathematik
Lehrstuhl für Versicherungsmathematik
Zellescher Weg 12-14
01062 Dresden
Deutschland
klaus.d.schmidt@tu-dresden.de

ISSN 0937-7433
ISBN 978-3-642-21025-9 e-ISBN 978-3-642-21026-6
DOI 10.1007/978-3-642-21026-6
Springer Heidelberg Dordrecht London New York

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Mathematics Subject Classification (2010): 28-01, 60-01

Einbandentwurf: WMXDesign GmbH, Heidelberg

Gedruckt auf säurefreiem Papier

Springer ist Teil der Fachverlagsgruppe Springer Science+Business Media (www.springer.com)

Vorwort zur 2. Auflage

Das Manuskript wurde sorgfältig durchgesehen und an vielen Stellen korrigiert oder in der Darstellung verbessert. Dabei wurde ich von Georg Berschneider, Sebastian Fuchs, Klaus Th. Hess, Elisabeth Löser, Wilfried Schenk und nicht zuletzt von meinen Studenten mit vielen wertvollen Hinweisen unterstützt. Elisabeth Löser und Christiane Weber haben wieder bei der Erstellung des Symbolverzeichnisses und des Sachverzeichnisses mitgewirkt. Ihnen allen sei herzlich gedankt.

Dresden, im Mai 2011 Klaus D. Schmidt

Vorwort zur 1. Auflage

Jede Zeit erfordert ihre eigene Sicht der Dinge. Das Anliegen dieses Buches ist es, in einer Zeit des Übergangs von einer Vielfalt von Diplomstudiengängen zu einer noch größeren Vielfalt von Bachelor– und Master–Studiengängen und der damit verbundenen Tendenz zur Verlagerung der Studieninhalte von der Theorie zur Anwendung, die Grundlagen der Wahrscheinlichkeitstheorie im Spannungsfeld zwischen Theorie und Anwendung darzustellen.

Als theoretische Grundlage der Wahrscheinlichkeitstheorie ist die Maß– und Integrationstheorie unverzichtbar, und zu einem gewissen Grad gilt dies auch für die Topologie, auf der unter anderem der Begriff der Borelschen σ–Algebra, die Konstruktion des Lebesgue–Maßes und die Konstruktion stochastischer Prozesse beruht.

Auf der anderen Seite erfordern Anwendungen der Wahrscheinlichkeitstheorie ein umfangreiches Repertoire an Methoden zur Konstruktion wahrscheinlichkeitstheoretischer Modelle. Grundlegend sind hier zum einen der Begriff der Unabhängigkeit und zum anderen der Begriff der bedingten Erwartung und der davon abgeleitete Begriff der bedingten Verteilung. In Anwendungen der Wahrscheinlichkeitstheorie ist schließlich auch die Kenntnis der Eigenschaften spezieller univariater und multivariater Verteilungen erforderlich.

Neben der Darstellung der Grundlagen der Wahrscheinlichkeitstheorie liefert dieses Buch mit zahlreichen Aufgaben auch Ansatzpunkte für das Studium spezieller Fragestellungen, von denen einige theoretisch orientiert sind und andere sich aus Anwendungen insbesondere im Bereich der Statistik und der Versicherungsmathematik ergeben.

Ein Autor, der stärker der angewandten reinen Mathematik als der reinen angewandten Mathematik verhaftet ist, ist geneigt, mathematische Aussagen unter möglichst allgemeinen Voraussetzungen zu beweisen. Die Aufgabe, ein Lehrbuch zu schreiben, setzt dieser Versuchung natürliche Grenzen, und so habe ich mich bemüht, zwischen dem Streben nach Allgemeinheit und der Beschränkung auf das Wesentliche ein Gleichgewicht zu finden.

Bei der Arbeit an diesem Buch habe ich vielfältige Unterstützung erhalten:
- Klaus Th. Hess und Mathias Zocher haben die Entstehung des Buches begleitet und mir als anregende Gesprächspartner zur Seite gestanden.
- Lothar Partzsch und Wilfried Schenk haben Teile des Manuskriptes durchgesehen und wertvolle Hinweise gegeben.
- Elisabeth Löser, Alexander Ludwig, Andreas Ringel und viele andere Studenten haben zu wesentlichen Verbesserungen beigetragen.
- Mandy Karzig hat fast alle Beispiele und Aufgaben zu univariaten und multivariaten Verteilungen überprüft.
- Christiane Weber hat Teile des Manuskriptes mit der ihr eigenen unübertrefflichen Sorgfalt korrekturgelesen und zusammen mit Elisabeth Löser bei der Erstellung des Symbolverzeichnisses und des Sachverzeichnisses mitgewirkt.

Ihnen allen sei herzlich gedankt.

Schließlich danke ich dem Springer–Verlag und insbesondere Clemens Heine für die angenehme Zusammenarbeit.

Dresden, im November 2008 Klaus D. Schmidt

Inhaltsverzeichnis

Anhang

Einleitung

Das vorliegende Buch bietet eine Einführung in die Maß– und Integrationstheorie und die Wahrscheinlichkeitstheorie. Diese Teilgebiete der Mathematik sind eng miteinander verbunden: Einerseits bildet die Maß– und Integrationstheorie die unverzichtbare Grundlage für die Wahrscheinlichkeitstheorie, und andererseits ist die Wahrscheinlichkeitstheorie neben der Funktionalanalysis das wichtigste Anwendungsgebiet der Maß– und Integrationstheorie. Das Buch ist in fünf Teile gegliedert, deren Inhalte wir im folgenden skizzieren:

Teil I – Mengensysteme: Mengensysteme sind Familien von Teilmengen einer Grundmenge und werden nach ihren Stabilitätseigenschaften unter mengentheoretischen Operationen klassifiziert. Die grundlegenden Mengensysteme in der Maß– und Integrationstheorie und damit auch in der Wahrscheinlichkeitstheorie sind die σ–Algebren, die in vielen wichtigen Fällen durch eine Topologie erzeugt werden. Daher werden neben den Eigenschaften spezieller Mengensysteme auch die Analogien zwischen Topologien und σ–Algebren und zwischen topologischen und messbaren Räumen sowie zwischen stetigen und messbaren Abbildungen dargestellt.

Teil II – Maßtheorie: Gegenstand der Maßtheorie sind Abbildungen von einem Mengensystem in die Menge der erweiterten reellen Zahlen. Derartige Abbildungen werden allgemein als Mengenfunktionen bezeichnet und messen die Mengen des Mengensystems, auf dem sie definiert sind. Die wichtigsten Mengenfunktionen sind die Maße, und das Maßproblem ist das Problem der Fortsetzung eines Maßes von einem kleinen Mengensystem auf ein möglichst großes Mengensystem. Im Fall des $\mathbb{R}^n$ führt die Lösung des Maßproblems für das Volumen von Intervallen auf das Lebesgue–Maß.

Teil III – Integrationstheorie: Das Maß einer Menge der σ–Algebra, auf der das Maß definiert ist, kann als Integral ihrer Indikatorfunktion aufgefasst werden. Damit ist ein allgemeines Integral bezüglich einem Maß definiert, das als Lebesgue–Integral bezeichnet wird und zunächst auf positive einfache Funktionen und dann auf positive messbare Funktionen und auf eine Klasse beliebiger messbarer Funktionen fortgesetzt wird. Dieser Fortsetzungsprozess, der auch als algebraische Induktion bezeichnet wird, bildet gleichzeitig die Grundlage für den Beweis vieler Ergebnisse über das Lebesgue–Integral. Die Berechnung des Lebesgue–Integrals kann in vielen Fällen durch die Berücksichtigung der besonderen Eigenschaften des integrierenden Maßes vereinfacht werden. Im Fall der Integration nach dem Lebesgue–Maß liefert insbesondere der Zusammenhang mit dem Riemann–Integral ein wichtiges Hilfsmittel für die Berechnung des Lebesgue–Integrals.

Teil IV – Wahrscheinlichkeitstheorie: Die Wahrscheinlichkeitstheorie hat ihren Ursprung in der Untersuchung von Glücksspielen. Es ist Kolmogorov zu verdanken, dass die Wahrscheinlichkeitstheorie durch ihre Einbettung in die Maß– und Integrationstheorie zu einer mathematischen Disziplin geworden ist, ohne die beispielsweise die moderne Versicherungsmathematik undenkbar wäre. Gegenüber der allgemeinen Maß– und Integrationstheorie ergeben sich in der Wahrscheinlichkeitstheorie Besonderheiten daraus, dass Wahrscheinlichkeitsmaße auf Eins normiert sind. Aufgrund dieser Normierung ist der für die Wahrscheinlichkeitstheorie typische und in ihr zentrale Begriff der Unabhängigkeit sinnvoll. Auf diesem Begriff beruht insbesondere das Gesetz der Großen Zahlen, das für eine unabhängig und identisch verteilte Folge von Zufallsvariablen die Konvergenz der Stichprobenmittel gegen den gemeinsamen Erwartungswert der Zufallsvariablen gewährleistet und damit eine Grundlage für die Schätzung ihres Erwartungswertes liefert.

Teil V – Vertiefung der Wahrscheinlichkeitstheorie: Das Gesetz der Großen Zahlen ist nur ein erstes Ziel der Wahrscheinlichkeitstheorie und eine erste Grundlage für die Mathematische Statistik. Da neben der Konvergenz der Stichprobenmittel gegen den Erwartungswert auch die Abweichung der Stichprobenmittel vom Erwartungswert von Interesse ist, stellt sich die Frage nach der näherungsweisen Berechnung ihrer Verteilungen. Diese Frage wird mit dem Zentralen Grenzwertsatz beantwortet, der eine weitere Grundlage für die Mathematische Statistik bildet. Die Annahme der Unabhängigkeit einer Familie von Zufallsvariablen, die in den starken Gesetzen der Großen Zahlen und im Zentralen Grenzwertsatz getroffen wird, ist ein wichtiges Hilfsmittel bei der Konstruktion wahrscheinlichkeitstheoretischer Modelle, aber sie ist in vielen Fällen nicht vertretbar. In derartigen Fällen ist es oft angemessen, für eine Familie von Zufallsvariablen Annahmen über ihre bedingte Verteilung bezüglich einem zufälligen Parameter und Annahmen über die unbedingte Verteilung des Parameters zu treffen. Diese zweistufige Modellierung ist daher ein weiteres Hilfsmittel bei der Konstruktion wahrscheinlichkeitstheoretischer Modelle. Ein solches Hilfsmittel ist schließlich auch der Satz von Kolmogorov, der die Grundlage für die Theorie der stochastischen Prozesse bildet.

Ein Buch wie dieses steht im Spannungsfeld zwischen Theorie und Handwerk: Zum einen beruhen Teile der Maß– und Integrationstheorie und der

Wahrscheinlichkeitstheorie auf Ergebnissen der Topologie und zum anderen ist für Anwendungen der Wahrscheinlichkeitstheorie eine gewisse Vertrautheit mit speziellen univariaten und multivariaten Verteilungen sowie die Fähigkeit, mit Verteilungen zu rechnen, unerlässlich. Im Hinblick auf Anwendungen der Wahrscheinlichkeitstheorie ist es außerdem wichtig, nicht nur die wesentlichen Ergebnisse der Theorie zu kennen, sondern auch die Beweismethoden zu einem gewissen Grad zu beherrschen. Dies ist die Voraussetzung für die Fähigkeit, bei Bedarf neue Ergebnisse zu entdecken und zu beweisen, die zunächst sehr anwendungsspezifisch sein mögen, aber auch die allgemeine Theorie bereichern und sich in ganz anderen Anwendungen als nützlich erweisen können.

Es ist daher nicht das Ziel dieses Buches, möglichst schnell zu möglichst wichtigen Ergebnissen hinzuführen. Statt dessen wird ein mathematischer Begriff zusammen mit den typischen Beweismethoden zunächst ausgeleuchtet, bevor weitere Begriffe eingeführt werden. Die Darstellung der allgemeinen Theorie wird durch zahlreiche Beispiele und Aufgaben begleitet. Dies gilt insbesondere für die Wahrscheinlichkeitstheorie, in der neben univariaten Verteilungen auch multivariate Verteilungen eingehend behandelt werden und Ansatzpunkte zu speziellen Themen und Anwendungen der Wahrscheinlichkeitstheorie gegeben werden, die über den Rahmen dieses Buches hinausgehen.

Konventionen: Wir verwenden die in der Maß– und Integrationstheorie und damit auch in der Wahrscheinlichkeitstheorie übliche Konvention

$$\frac{0}{0} := 0$$

sowie die Konvention

$$0^0 := 1$$

Wir bezeichnen eine reelle Zahl $x \in \mathbb{R}$ und eine Funktion $f : \Omega \to \bar{\mathbb{R}}$ als *positiv*, wenn $x \geq 0$ bzw. $f(\omega) \geq 0$ für alle $\omega \in \Omega$ gilt. Des weiteren setzen wir

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\}$$

und für $m \in \mathbb{N}$ setzen wir

$$\mathbb{N}(m) := \{n \in \mathbb{N} \mid m \leq n\}$$

wobei $\mathbb{N}$ die Menge der natürlichen Zahlen $\{1, 2, \dots\}$ bezeichnet. Für weitere Bezeichnungen und Begriffe verweisen wir auf das Symbolverzeichnis und das Sachverzeichnis.

Teil I

Mengensysteme und Abbildungen

1

Mengensysteme

Ein Mengensystem ist eine Familie von Teilmengen einer Grundmenge und damit eine Teilmenge der Potenzmenge der Grundmenge. In diesem Kapitel untersuchen wir Mengensysteme, die unter bestimmten mengentheoretischen Operationen stabil sind. Die wichtigsten Mengensysteme sind die Topologien und die σ–Algebren, wobei in der Maßtheorie vor allem die σ–Algebren von Bedeutung sind.

Zwischen Topologien und σ–Algebren bestehen Analogien, die es als sinnvoll erscheinen lassen, diese Mengensysteme parallel zu untersuchen. So lassen sich zum Beispiel die Mengen, die einer Topologie oder σ–Algebra angehören, nur in seltenen Fällen explizit beschreiben. Topologien und σ–Algebren werden daher oft indirekt definiert, indem man ein kleineres Mengensystem angibt, das die Topologie oder die σ–Algebra in einer noch zu präzisierenden Weise erzeugt.

Da bestimmte σ–Algebren durch eine Topologie erzeugt werden, beginnen wir die Untersuchung von Mengensystemen mit Topologien (Abschnitt 1.1) und führen erst dann σ–Algebren ein (Abschnitt 1.2). Die nächsten Abschnitte betreffen Mengensysteme, die für die Erzeugung von σ–Algebren von Bedeutung sind. Dies sind vor allem Dynkin–Systeme (Abschnitt 1.3) und $\cap$–stabile Mengensysteme (Abschnitt 1.4) sowie Halbringe und Ringe (Abschnitt 1.5).

Im gesamten Kapitel sei Ω eine nichtleere Menge. Wir bezeichnen die Potenzmenge von Ω mit

$$2^\Omega$$

Für $A \in 2^\Omega$ bezeichnen wir die Mächtigkeit von A mit

$$|A|$$

und das Komplement von A mit

$$\overline{A}$$

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_1, © Springer-Verlag Berlin Heidelberg 2011

Für disjunkte Mengen $A, B \in 2^\Omega$ setzen wir

$$A + B := A \cup B$$

Eine Familie $\{A_i\}_{i \in I} \subseteq 2^\Omega$ heißt *disjunkt*, wenn $A_i \cap A_j = \emptyset$ für alle $i, j \in I$ mit $i \neq j$ gilt, und in diesem Fall setzen wir

$$\sum_{i \in I} A_i := \bigcup_{i \in I} A_i$$

Für eine Folge $\{A_n\}_{n \in \mathbb{N}} \subseteq 2^\Omega$ heißt die Menge

$$\liminf_{n \to \infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

der *Limes inferior* der Folge $\{A_n\}_{n \in \mathbb{N}}$ und die Menge

$$\limsup_{n \to \infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

heißt der *Limes superior* der Folge $\{A_n\}_{n \in \mathbb{N}}$. Es gilt

$$\liminf_{n \to \infty} A_n = \Big\{ \omega \in \Omega \;\Big|\; \omega \in A_n \text{ für alle außer endlich viele } n \in \mathbb{N} \Big\}$$

und

$$\limsup_{n \to \infty} A_n = \Big\{ \omega \in \Omega \;\Big|\; \omega \in A_n \text{ für unendlich viele } n \in \mathbb{N} \Big\}$$

und damit $\liminf_{n \to \infty} A_n \subseteq \limsup_{n \to \infty} A_n$.

Eine Menge heißt *abzählbar*, wenn sie endlich oder abzählbar unendlich ist.

Jede Teilmenge der Menge 2^Ω heißt *Mengensystem* auf Ω.

1.1 Topologien

Ein Mengensystem $\mathcal{T} \subseteq 2^\Omega$ heißt *Topologie* auf Ω, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\Omega \in \mathcal{T}$ und $\emptyset \in \mathcal{T}$.
(ii) Für jede Familie $\{Q_i\}_{i \in I} \subseteq \mathcal{T}$ gilt $\bigcup_{i \in I} Q_i \in \mathcal{T}$.
(iii) Für jede endliche Familie $\{Q_i\}_{i \in I} \subseteq \mathcal{T}$ gilt $\bigcap_{i \in I} Q_i \in \mathcal{T}$.
Ist $\mathcal{T} \subseteq 2^\Omega$ eine Topologie, so heißt eine Menge $A \in 2^\Omega$
– *offen* bezüglich $\mathcal{T}$, wenn $A \in \mathcal{T}$ gilt.
– *abgeschlossen* bezüglich $\mathcal{T}$, wenn $\overline{A} \in \mathcal{T}$ gilt.
– *kompakt* bezüglich $\mathcal{T}$, wenn es zu jeder Familie $\{Q_i\}_{i \in I} \subseteq \mathcal{T}$ mit $A \subseteq \bigcup_{i \in I} Q_i$ eine endliche Menge $J \subseteq I$ gibt mit $A \subseteq \bigcup_{i \in J} Q_i$.

1.1.1 Beispiele.
(1) Die Potenzmenge 2^Ω ist eine Topologie auf Ω.
(2) Das Mengensystem $\{\emptyset, \Omega\}$ ist eine Topologie auf Ω.

1.1.2 Beispiel (Topologie eines normierten Raumes). Sei $(E\,,\|\,.\,\|)$ ein normierter Raum. Für $x \in E$ und $\varepsilon \in (0,\infty)$ sei

$$B_{\|\,.\,\|}(x,\varepsilon) := \Bigl\{ y \in E \Bigm| \|y-x\| < \varepsilon \Bigr\}$$

Die Menge $B_{\|\,.\,\|}(x,\varepsilon)$ heißt *offene Kugel* um x mit Radius ε bezüglich der Norm $\|\,.\,\|$. Sei

$$\mathcal{T}_{\|\,.\,\|} := \Bigl\{ Q \in 2^E \Bigm| \text{für alle } x \in Q \text{ gibt es ein } \varepsilon \in (0,\infty) \text{ mit } B_{\|\,.\,\|}(x,\varepsilon) \subseteq Q \Bigr\}$$

Dann ist $\mathcal{T}_{\|\,.\,\|}$ eine Topologie auf E und jede offene Kugel bezüglich der Norm $\|\,.\,\|$ ist offen bezüglich der Topologie $\mathcal{T}_{\|\,.\,\|}$.
In der Tat: Wir zeigen zunächst, dass $\mathcal{T}_{\|\,.\,\|}$ eine Topologie auf E ist:
(i) Offenbar gilt $E \in \mathcal{T}_{\|\,.\,\|}$ und $\emptyset \in \mathcal{T}_{\|\,.\,\|}$.
(ii) Sei $\{Q_i\}_{i\in I}$ eine beliebige Familie von Mengen in $\mathcal{T}_{\|\,.\,\|}$ und sei

$$Q := \bigcup_{i\in I} Q_i$$

Sei $x \in Q$. Dann gibt es ein $i \in I$ mit $x \in Q_i$. Wegen $Q_i \in \mathcal{T}_{\|\,.\,\|}$ gibt es ein $\varepsilon \in (0,\infty)$ mit $B_{\|\,.\,\|}(x,\varepsilon) \subseteq Q_i$ und wegen $Q_i \subseteq Q$ folgt daraus $B_{\|\,.\,\|}(x,\varepsilon) \subseteq Q$. Daher gilt $Q \in \mathcal{T}_{\|\,.\,\|}$.
(iii) Sei $\{Q_i\}_{i\in I}$ eine endliche Familie von Mengen in $\mathcal{T}_{\|\,.\,\|}$ und sei

$$Q := \bigcap_{i\in I} Q_i$$

Sei $x \in Q$. Für alle $i \in I$ gilt $x \in Q_i$ und wegen $Q_i \in \mathcal{T}_{\|\,.\,\|}$ gibt es für alle $i \in I$ ein $\varepsilon_i \in (0,\infty)$ mit $B_{\|\,.\,\|}(x,\varepsilon_i) \subseteq Q_i$. Sei $\varepsilon := \min_{i\in I} \varepsilon_i$. Da I endlich ist, gilt $\varepsilon \in (0,\infty)$ und für alle $i \in I$ ergibt sich $B_{\|\,.\,\|}(x,\varepsilon) \subseteq B_{\|\,.\,\|}(x,\varepsilon_i) \subseteq Q_i$. Daraus folgt $B_{\|\,.\,\|}(x,\varepsilon) \subseteq \bigcap_{i\in I} Q_i = Q$. Daher gilt $Q \in \mathcal{T}_{\|\,.\,\|}$.

Daher ist $\mathcal{T}_{\|\,.\,\|}$ eine Topologie auf E.
Wir zeigen nun, dass jede offene Kugel bezüglich der Norm $\|\,.\,\|$ offen bezüglich der Topologie $\mathcal{T}_{\|\,.\,\|}$ ist: Sei $x \in E$ und $\varepsilon \in (0,\infty)$. Für $y \in B_{\|\,.\,\|}(x,\varepsilon)$ sei $\eta_y := \|y-x\|$ und $\varepsilon_y := \varepsilon - \eta_y$. Wegen $\eta_y < \varepsilon$ gilt $\varepsilon_y \in (0,\infty)$. Für alle $z \in B_{\|\,.\,\|}(y,\varepsilon_y)$ gilt daher

$$\|z-x\| \le \|z-y\| + \|y-x\| < \varepsilon_y + \eta_y = (\varepsilon-\eta_y) + \eta_y = \varepsilon$$

und damit $z \in B_{\|\,.\,\|}(x,\varepsilon)$. Da $z \in B_{\|\,.\,\|}(y,\varepsilon_y)$ beliebig war, ergibt sich

$$B_{\|\,.\,\|}(y,\varepsilon_y) \subseteq B_{\|\,.\,\|}(x,\varepsilon)$$

Da $y \in B_{\|\,.\,\|}(x,\varepsilon)$ beliebig war, ergibt sich

$$B_{\|\,.\,\|}(x,\varepsilon) \in \mathcal{T}_{\|\,.\,\|}$$

Daher ist jede offene Kugel bezüglich der Norm $\|\,.\,\|$ offen bezüglich der Topologie $\mathcal{T}_{\|\,.\,\|}$.

Für einen normierten Raum $(E,\|\,.\,\|)$ wird die in Beispiel 1.1.2 definierte Topologie $\mathcal{T}_{\|\,.\,\|}$ als die von der Norm $\|\,.\,\|$ *erzeugte Topologie* oder als die *Normtopologie* bezüglich der Norm $\|\,.\,\|$ bezeichnet.

Wir betrachten nun Durchschnitte von Topologien:

1.1.3 Satz. *Sei H eine nichtleere Indexmenge und sei $\{\mathcal{T}_h\}_{h\in H}$ eine Familie von Topologien auf Ω. Dann ist das Mengensystem*

$$\mathcal{T} := \bigcap_{h\in H} \mathcal{T}_h$$

eine Topologie auf Ω.

Beweis. Wir zeigen, dass $\mathcal{T}$ die Axiome einer Topologie erfüllt:
(i) Für alle $h\in H$ gilt $\Omega, \emptyset \in \mathcal{T}_h$. Daher gilt $\Omega, \emptyset \in \bigcap_{h\in H} \mathcal{T}_h = \mathcal{T}$.
(ii) Sei $\{Q_i\}_{i\in I}$ eine beliebige Familie von Mengen in $\mathcal{T} = \bigcap_{h\in H} \mathcal{T}_h$. Für alle $h \in H$ und für alle $i \in I$ gilt dann $Q_i \in \mathcal{T}_h$. Für alle $h \in H$ folgt daraus $\bigcup_{i\in I} Q_i \in \mathcal{T}_h$. Daher gilt $\bigcup_{i\in I} Q_i \in \bigcap_{h\in H} \mathcal{T}_h = \mathcal{T}$.
(iii) Sei $\{Q_i\}_{i\in I}$ eine endliche Familie von Mengen in $\mathcal{T} = \bigcap_{h\in H} \mathcal{T}_h$. Für alle $h \in H$ und für alle $i \in I$ gilt dann $Q_i \in \mathcal{T}_h$. Für alle $h \in H$ folgt daraus $\bigcap_{i\in I} Q_i \in \mathcal{T}_h$. Daher gilt $\bigcap_{i\in I} Q_i \in \bigcap_{h\in H} \mathcal{T}_h = \mathcal{T}$.
Daher ist $\mathcal{T}$ eine Topologie auf Ω. □

Aus Satz 1.1.3 ergibt sich eine wichtige Folgerung:

1.1.4 Folgerung. *Zu jedem Mengensystem $\mathcal{E} \subseteq 2^\Omega$ gibt es eine kleinste Topologie auf Ω, die $\mathcal{E}$ enthält.*

Beweis. Wir betrachten die Familie aller Topologien auf Ω, die $\mathcal{E}$ enthalten. Diese Familie ist nichtleer, denn die Potenzmenge 2^Ω ist eine Topologie mit $\mathcal{E} \subseteq 2^\Omega$. Nach Satz 1.1.3 ist der Durchschnitt aller Topologien, die $\mathcal{E}$ enthalten, ebenfalls eine Topologie. Diese Topologie enthält $\mathcal{E}$ und ist offenbar die kleinste Topologie, die $\mathcal{E}$ enthält. □

Für ein Mengensystem $\mathcal{E} \subseteq 2^\Omega$ bezeichnen wir die kleinste Topologie auf Ω, die $\mathcal{E}$ enthält, mit

$$\tau(\mathcal{E})$$

Diese Topologie wird als die *von $\mathcal{E}$ erzeugte Topologie* bezeichnet und das Mengensystem $\mathcal{E}$ wird als *Erzeuger* der Topologie $\tau(\mathcal{E})$ bezeichnet.

1.1.5 Beispiel (Topologie eines normierten Raumes). Sei $(E, \|\,.\,\|)$ ein normierter Raum. Sei ferner

$$\mathcal{E}_{\|\,.\,\|} := \Big\{ B_{\|\,.\,\|}(x,\varepsilon) \,\Big|\, x \in E,\, \varepsilon \in (0,\infty) \Big\}$$

Dann gilt

$$\mathcal{T}_{\|\,.\,\|} = \tau(\mathcal{E}_{\|\,.\,\|})$$

In der Tat: Nach Beispiel 1.1.2 gilt

$$\mathcal{E}_{\|\,.\,\|} \subseteq \mathcal{T}_{\|\,.\,\|}$$

und aus der Definition von $\tau(\mathcal{E}_{\|\,.\,\|})$ folgt nun

$$\tau(\mathcal{E}_{\|\,.\,\|}) \subseteq \mathcal{T}_{\|\,.\,\|}$$

Sei nun $\mathcal{T}$ eine beliebige Topologie auf E mit $\mathcal{E}_{\|\,.\,\|} \subseteq \mathcal{T}$. Sei $Q \in \mathcal{T}_{\|\,.\,\|}$. Für alle $x \in Q$ gibt es ein $\varepsilon_x \in (0,\infty)$ mit $B_{\|\,.\,\|}(x,\varepsilon_x) \subseteq Q$. Daraus folgt

$$Q = \bigcup_{x \in Q} \{x\} \subseteq \bigcup_{x \in Q} B_{\|\,.\,\|}(x,\varepsilon_x) \subseteq Q$$

und damit $Q = \bigcup_{x \in Q} B_{\|\,.\,\|}(x,\varepsilon_x)$. Wegen $B_{\|\,.\,\|}(x,\varepsilon_x) \in \mathcal{E}_{\|\,.\,\|} \subseteq \mathcal{T}$ gilt daher $Q \in \mathcal{T}$. Daraus folgt zunächst

$$\mathcal{T}_{\|\,.\,\|} \subseteq \mathcal{T}$$

und aus der Definition von $\tau(\mathcal{E}_{\|\,.\,\|})$ folgt nun

$$\mathcal{T}_{\|\,.\,\|} \subseteq \tau(\mathcal{E}_{\|\,.\,\|})$$

Daher gilt $\tau(\mathcal{E}_{\|\,.\,\|}) = \mathcal{T}_{\|\,.\,\|}$.

Auf einem Vektorraum können mehrere Normen definiert sein; insbesondere ist jedes strikt positive Vielfache einer Norm wieder eine Norm. Interessanter ist das folgende Beispiel:

1.1.6 Beispiel (Normen auf $\mathbb{R}^n$). Für jedes $p \in [1,\infty]$ ist die Abbildung $\|\,.\,\|_p : \mathbb{R}^n \to \mathbb{R}_+$ mit

$$\|\mathbf{x}\|_p := \begin{cases} \left(\sum_{i \in \{1,\ldots,n\}} |x_i|^p \right)^{1/p} & \text{falls } p \in [1,\infty) \\ \max_{i \in \{1,\ldots,n\}} |x_i| & \text{falls } p = \infty \end{cases}$$

eine Norm auf $\mathbb{R}^n$.

Damit stellt sich die Frage, unter welchen Umständen die von verschiedenen Normen erzeugten Topologien übereinstimmen.

1.1.7 Beispiel (Topologien äquivalenter Normen). Sei E ein Vektorraum und seien $\|\,.\,\|' : E \to \mathbb{R}_+$ und $\|\,.\,\|'' : E \to \mathbb{R}_+$ Normen. Die Normen $\|\,.\,\|'$ und $\|\,.\,\|''$ heißen *äquivalent*, wenn es reelle Zahlen $c', c'' \in (0,\infty)$ gibt derart, dass für alle $x \in E$

$$\begin{aligned} \|x\|' &\leq c'' \cdot \|x\|'' \\ \|x\|'' &\leq c' \cdot \|x\|' \end{aligned}$$

gilt. Sind $\|\,.\,\|' : E \to \mathbb{R}_+$ und $\|\,.\,\|'' : E \to \mathbb{R}_+$ äquivalente Normen, so gilt

$$\mathcal{T}_{\|\,.\,\|'} = \mathcal{T}_{\|\,.\,\|''}$$

In der Tat: Zum Nachweis der behaupteten Identität genügt es, eine der Inklusionen $\mathcal{T}_{\|\,.\,\|'} \subseteq \mathcal{T}_{\|\,.\,\|''}$ und $\mathcal{T}_{\|\,.\,\|''} \subseteq \mathcal{T}_{\|\,.\,\|'}$ zu zeigen.
Sei $Q \in \mathcal{T}_{\|\,.\,\|'}$. Für alle $x \in Q$ gibt es ein $\varepsilon_x \in (0,\infty)$ mit $B_{\|\,.\,\|'}(x,\varepsilon_x) \subseteq Q$. Daher gilt

$$Q = \bigcup_{x\in Q}\{x\} \subseteq \bigcup_{x\in Q} B_{\|\,.\,\|''}(x,\varepsilon_x/c'') \subseteq \bigcup_{x\in Q} B_{\|\,.\,\|'}(x,\varepsilon_x) \subseteq Q$$

und damit $Q = \bigcup_{x\in Q} B_{\|\,.\,\|''}(x,\varepsilon_x/c'') \in \mathcal{T}_{\|\,.\,\|''}$.

Ist E ein endlichdimensionaler Vektorraum, so sind alle Normen auf E äquivalent. In diesem Fall stimmen nach Beispiel 1.1.7 alle von einer Norm erzeugten Topologien auf E überein und die von einer beliebigen Norm auf E erzeugte Topologie

$$\mathcal{T}(E)$$

wird als *natürliche Topologie* auf E bezeichnet.

Ist $\mathcal{T}$ eine Topologie und $\mathcal{E} \subseteq \mathcal{T}$ ein Mengensystem mit der Eigenschaft, dass jede Menge in $\mathcal{T}$ als Vereinigung von Mengen in $\mathcal{E}$ dargestellt werden kann, so heißt $\mathcal{E}$ *Basis* der Topologie $\mathcal{T}$; in diesem Fall gilt insbesondere $\mathcal{T} = \tau(\mathcal{E})$.

Das folgende Beispiel zeigt, dass die natürliche Topologie auf $\mathbb{R}^n$ eine abzählbare Basis besitzt:

1.1.8 Beispiel (Natürliche Topologie auf $\mathbb{R}^n$). Sei $\|\,.\,\| : \mathbb{R}^n \to \mathbb{R}$ eine beliebige Norm und sei

$$\mathcal{E}_{\|\,.\,\|,\mathbb{Q}} := \Bigl\{ B_{\|\,.\,\|}(\mathbf{x},\varepsilon) \Bigm| \mathbf{x} \in \mathbb{Q}^n,\ \varepsilon \in \mathbb{Q}\cap(0,\infty) \Bigr\}$$

Dann ist $\mathcal{E}_{\|\,.\,\|,\mathbb{Q}}$ abzählbar und eine Basis der natürlichen Topologie $\mathcal{T}(\mathbb{R}^n)$.
In der Tat: Offenbar ist $\mathcal{E}_{\|\,.\,\|,\mathbb{Q}}$ abzählbar und nach Beispiel 1.1.2 gilt

$$\mathcal{E}_{\|\,.\,\|,\mathbb{Q}} \subseteq \mathcal{T}_{\|\,.\,\|} = \mathcal{T}(\mathbb{R}^n)$$

Sei nun $Q \in \mathcal{T}_{\|\,.\,\|}$. Für $\mathbf{x} \in \mathbb{R}^n$ gibt es ein $\varepsilon_{\mathbf{x}} \in (0,\infty)$ mit $B_{\|\,.\,\|}(\mathbf{x},\varepsilon_{\mathbf{x}}) \subseteq Q$, und damit gibt es auch ein $\eta_{\mathbf{x}} \in \mathbb{Q}\cap(0,\infty)$ mit

$$B_{\|\,.\,\|}(\mathbf{x},\eta_{\mathbf{x}}) \subseteq Q$$

Außerdem gibt es ein $\mathbf{z}_{\mathbf{x}} \in \mathbb{Q}^n$ mit $\|\mathbf{z}_{\mathbf{x}}-\mathbf{x}\| < \eta_{\mathbf{x}}/2$ und damit

$$\mathbf{x} \in B_{\|\,.\,\|}(\mathbf{z}_{\mathbf{x}},\eta_{\mathbf{x}}/2)$$

Für alle $\mathbf{y} \in B_{\|\,.\,\|}(\mathbf{z}_{\mathbf{x}},\eta_{\mathbf{x}}/2)$ gilt $\|\mathbf{y}-\mathbf{x}\| \le \|\mathbf{y}-\mathbf{z}_{\mathbf{x}}\| + \|\mathbf{z}_{\mathbf{x}}-\mathbf{x}\| < \eta_{\mathbf{x}}/2 + \eta_{\mathbf{x}}/2 = \eta_{\mathbf{x}}$ und damit $\mathbf{y} \in B_{\|\,.\,\|}(\mathbf{x},\eta_{\mathbf{x}})$. Daraus folgt

$$B_{\|\,.\,\|}(\mathbf{z}_{\mathbf{x}},\eta_{\mathbf{x}}/2) \subseteq B_{\|\,.\,\|}(\mathbf{x},\eta_{\mathbf{x}})$$

Daher gilt

$$Q = \bigcup_{\mathbf{x} \in Q} \{\mathbf{x}\} \subseteq \bigcup_{\mathbf{x} \in Q} B_{\|\,.\,\|}(\mathbf{z_x}, \eta_{\mathbf{x}}/2) \subseteq \bigcup_{\mathbf{x} \in Q} B_{\|\,.\,\|}(\mathbf{x}, \eta_{\mathbf{x}}) \subseteq Q$$

und damit

$$Q = \bigcup_{\mathbf{x} \in Q} B_{\|\,.\,\|}(\mathbf{z_x}, \eta_{\mathbf{x}}/2)$$

Daher ist $\mathcal{E}_{\|\,.\,\|,\mathbb{Q}}$ eine Basis von $\mathcal{T}_{\|\,.\,\|} = \mathcal{T}(\mathbb{R}^n)$.

Ausgehend von der natürlichen Topologie auf $\mathbb{R}$ lässt sich eine Topologie auf der Menge $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ definieren:

1.1.9 Beispiel (Natürliche Topologie auf $\bar{\mathbb{R}}$). Das Mengensystem

$$\mathcal{T}(\bar{\mathbb{R}}) := \left\{ Q \in 2^{\bar{\mathbb{R}}} \,\middle|\, Q \cap \mathbb{R} \in \mathcal{T}(\mathbb{R}) \right\}$$

ist eine Topologie auf $\bar{\mathbb{R}}$ und wird als *natürliche Topologie* auf $\bar{\mathbb{R}}$ bezeichnet. Die natürliche Topologie auf $\bar{\mathbb{R}}$ besteht offenbar genau aus den Mengen der Form Q, $Q \cup \{-\infty\}$, $Q \cup \{\infty\}$, $Q \cup \{-\infty, \infty\}$ mit $Q \in \mathcal{T}(\mathbb{R})$.

Aufgaben

1.1.A Sei $\mathcal{T}$ eine Topologie auf Ω. Zu jeder Menge $A \in 2^\Omega$ gibt es eine größte offene Menge A° mit $A^\circ \subseteq A$ und eine kleinste abgeschlossene Menge $A^\bullet$ mit $A \subseteq A^\bullet$. Die Menge A° heißt das *Innere* von A, die Menge $A^\bullet$ heißt der *Abschluss* von A, und die Menge $\partial A := A^\bullet \setminus A^\circ$ heißt der *Rand* von A. Der Rand von A ist abgeschlossen.

1.1.B Sei $\mathcal{T}$ eine Topologie auf Ω. Ist $K \in 2^\Omega$ kompakt und $A \in 2^\Omega$ abgeschlossen, so ist $K \cap A$ kompakt.

1.1.C Mindestens eine der Topologien $2^\mathbb{R}$ und $\{\emptyset, \mathbb{R}\}$ kann nicht durch eine Norm auf $\mathbb{R}$ erzeugt werden. Welche?

1.1.D **Topologie eines metrischen Raumes:** Sei (E, d) ein metrischer Raum. Für $x \in E$ und $\varepsilon \in (0, \infty)$ sei

$$B_d(x, \varepsilon) := \left\{ y \in E \,\middle|\, d(y, x) < \varepsilon \right\}$$

Die Menge $B_d(x, \varepsilon)$ heißt *offene Kugel* um x mit Radius ε bezüglich der Metrik d. Sei

$$\mathcal{T}_d := \left\{ Q \in 2^E \,\middle|\, \text{für alle } x \in Q \text{ gibt es ein } \varepsilon \in (0, \infty) \text{ mit } B_d(x, \varepsilon) \subseteq Q \right\}$$

Dann ist $\mathcal{T}_d$ eine Topologie auf E. Die Topologie $\mathcal{T}_d$ wird als die von der Metrik d *erzeugte Topologie* bezeichnet. Sei ferner

$$\mathcal{E}_d := \left\{ B_d(x, \varepsilon) \,\middle|\, x \in E,\, \varepsilon \in (0, \infty) \right\}$$

Dann gilt $\mathcal{T}_d = \tau(\mathcal{E}_d)$.

1.1.E **Hausdorff–Räume:** Ein topologischer Raum $(E, \mathcal{T})$ heißt *Hausdorff–Raum*, wenn es für alle $x, y \in E$ mit $x \neq y$ offene Mengen $U, V \in \mathcal{T}$ gibt mit $x \in U$ und $y \in V$ sowie $U \cap V = \emptyset$.

(1) Für jeden metrischen Raum (E, d) ist $(E, \mathcal{T}_d)$ ein Hausdorff–Raum.

(2) In einem Hausdorff–Raum ist jede kompakte Menge abgeschlossen.

1.2 σ–Algebren

Ein Mengensystem $\mathcal{F} \subseteq 2^\Omega$ heißt *σ–Algebra* auf Ω, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\Omega \in \mathcal{F}$.
(ii) Für jede Menge $A \in \mathcal{F}$ gilt $\overline{A} \in \mathcal{F}$.
(iii) Für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ gilt $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.
Ist $\mathcal{F} \subseteq 2^\Omega$ eine σ–Algebra, so heißt eine Menge $A \in 2^\Omega$ *messbar* bezüglich $\mathcal{F}$, wenn $A \in \mathcal{F}$ gilt.

1.2.1 Beispiele.
(1) Die Potenzmenge 2^Ω ist eine σ–Algebra auf Ω.
(2) Das Mengensystem

$$\mathcal{F} := \Bigl\{A \in 2^\Omega \Bigm| A \text{ oder } \overline{A} \text{ ist abzählbar}\Bigr\}$$

ist eine σ–Algebra auf Ω.
(3) Für jede Menge $A \in 2^\Omega$ mit $A \notin \{\emptyset, \Omega\}$ ist das Mengensystem

$$\mathcal{F} := \{\emptyset, A, \overline{A}, \Omega\}$$

eine σ–Algebra auf Ω.
(4) Das Mengensystem $\{\emptyset, \Omega\}$ ist eine σ–Algebra auf Ω.

1.2.2 Lemma. *Sei $\mathcal{F}$ eine σ–Algebra auf Ω. Dann gilt:*
(1) $\emptyset \in \mathcal{F}$.
(2) *Für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ gilt $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}$.*
(3) *Für jede endliche Familie $\{A_i\}_{i\in I} \subseteq \mathcal{F}$ gilt $\bigcup_{i\in I} A_i \in \mathcal{F}$ und $\bigcap_{i\in I} A_i \in \mathcal{F}$.*

Beweis. Es gilt $\emptyset = \overline{\Omega} \in \mathcal{F}$ und für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ gilt

$$\bigcap_{n\in\mathbb{N}} A_n = \overline{\bigcup_{n\in\mathbb{N}} \overline{A_n}}$$

und damit $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}$. Damit sind (1) und (2) gezeigt. Schließlich folgt (3) aus (1) und (2), da man jede endliche Familie von Mengen in $\mathcal{F}$ mit Hilfe der Mengen $\emptyset$ bzw. Ω zu einer Folge in $\mathcal{F}$ erweitern kann, ohne die Vereinigung bzw. den Durchschnitt zu verändern. □

Den nächsten Satz und die anschließende Folgerung beweist man genau wie im Fall von Topologien:

1.2.3 Satz. *Sei H eine nichtleere Indexmenge und sei $\{\mathcal{F}_h\}_{h\in H}$ eine Familie von σ–Algebren auf Ω. Dann ist das Mengensystem*

$$\mathcal{F} := \bigcap_{h\in H} \mathcal{F}_h$$

eine σ–Algebra auf Ω.

1.2.4 Folgerung. *Zu jedem Mengensystem $\mathcal{E} \subseteq 2^\Omega$ gibt es eine kleinste σ–Algebra auf Ω, die $\mathcal{E}$ enthält.*

Für ein Mengensystem $\mathcal{E} \subseteq 2^\Omega$ bezeichnen wir die kleinste σ–Algebra auf Ω, die $\mathcal{E}$ enthält, mit

$$\sigma(\mathcal{E})$$

Diese σ–Algebra wird als die *von $\mathcal{E}$ erzeugte σ–Algebra* bezeichnet und das Mengensystem $\mathcal{E}$ wird als *Erzeuger* der σ–Algebra $\sigma(\mathcal{E})$ bezeichnet. Für eine Menge $A \in 2^\Omega$ schreiben wir auch

$$\sigma(A)$$

anstelle von $\sigma(\{A\})$ und bezeichnen $\sigma(A)$ als die *von A erzeugte σ–Algebra*.

Von Interesse sind unter anderem σ–Algebren auf Ω, die von einer Topologie auf Ω erzeugt werden. Ist $\mathcal{T}$ eine Topologie auf Ω, so wird die σ–Algebra

$$\mathcal{B}(\Omega, \mathcal{T}) := \sigma(\mathcal{T})$$

als *Borelsche σ–Algebra* bezüglich $\mathcal{T}$ bezeichnet. Jede Menge $B \in \mathcal{B}(\Omega, \mathcal{T})$ heißt *Borel–Menge* von Ω bezüglich $\mathcal{T}$.

1.2.5 Beispiel (Borelsche σ–Algebra auf $\mathbb{R}^n$). Die σ–Algebra

$$\mathcal{B}(\mathbb{R}^n) := \sigma(\mathcal{T}(\mathbb{R}^n))$$

wird als *Borelsche σ–Algebra* (bezüglich der natürlichen Topologie) auf $\mathbb{R}^n$ bezeichnet; vgl. Aufgabe 1.2.A.

Von Interesse ist auch die Borelsche σ–Algebra auf $\bar{\mathbb{R}}$:

1.2.6 Beispiel (Borelsche σ–Algebra auf $\bar{\mathbb{R}}$). Die σ–Algebra

$$\mathcal{B}(\bar{\mathbb{R}}) := \sigma(\mathcal{T}(\bar{\mathbb{R}}))$$

wird als *Borelsche σ–Algebra* auf $\bar{\mathbb{R}}$ bezeichnet. Es gilt

$$\mathcal{B}(\bar{\mathbb{R}}) = \Bigl\{ B \in 2^{\bar{\mathbb{R}}} \Bigm| B \cap \mathbb{R} \in \mathcal{B}(\mathbb{R}) \Bigr\}$$

Die Borelsche σ–Algebra $\mathcal{B}(\bar{\mathbb{R}})$ besteht offenbar genau aus den Mengen der Form B, $B \cup \{-\infty\}$, $B \cup \{\infty\}$, $B \cup \{-\infty, \infty\}$ mit $B \in \mathcal{B}(\mathbb{R})$.

Aufgaben

1.2.A **Borelsche σ–Algebra auf $\mathbb{R}^n$:** Für jede Norm $\|\,.\,\| : \mathbb{R}^n \to \mathbb{R}_+$ gilt

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{E}_{\|\,.\,\|,\mathbb{Q}})$$

Insbesondere besitzt die Borelsche σ–Algebra auf $\mathbb{R}^n$ einen abzählbaren Erzeuger.

1.2.B **Erzeugte σ–Algebra:** Bestimmen Sie für $A \in 2^\Omega$ die σ–Algebra $\sigma(A)$.

1.3 Dynkin–Systeme

Ein Mengensystem $\mathcal{D} \subseteq 2^\Omega$ heißt *Dynkin–System* auf Ω, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\Omega \in \mathcal{D}$.
(ii) Für jede Menge $A \in \mathcal{D}$ gilt $\overline{A} \in \mathcal{D}$.
(iii) Für jede disjunkte Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{D}$ gilt $\sum_{n\in\mathbb{N}} A_n \in \mathcal{D}$.

1.3.1 Lemma. *Jede σ–Algebra ist ein Dynkin–System.*

Die Umkehrung der Implikation von Lemma 1.3.1 ist jedoch falsch; vgl. Beispiel 1.3.6.

1.3.2 Lemma. *Sei $\mathcal{D}$ ein Dynkin–System. Dann gilt für alle $A, B \in \mathcal{D}$ mit $B \subseteq A$*

$$A \setminus B \in \mathcal{D}$$

Beweis. Wegen $B \subseteq A$ gilt $\overline{A} \cap B = \emptyset$ und damit $A \setminus B = A \cap \overline{B} = \overline{\overline{A} \cup B} = \overline{\overline{A} + B} \in \mathcal{D}$. □

Den nächsten Satz und die anschließende Folgerung beweist man genau wie im Fall von Topologien:

1.3.3 Satz. *Sei H eine nichtleere Indexmenge und sei $\{\mathcal{D}_h\}_{h\in H}$ eine Familie von Dynkin–Systemen auf Ω. Dann ist das Mengensystem*

$$\mathcal{D} := \bigcap_{h\in H} \mathcal{D}_h$$

ein Dynkin–System auf Ω.

1.3.4 Folgerung. *Zu jedem Mengensystem $\mathcal{E} \subseteq 2^\Omega$ gibt es ein kleinstes Dynkin–System auf Ω, das $\mathcal{E}$ enthält.*

Für ein Mengensystem $\mathcal{E} \subseteq 2^\Omega$ bezeichnen wir das kleinste Dynkin–System auf Ω, das $\mathcal{E}$ enthält, mit

$$\delta(\mathcal{E})$$

Dieses Dynkin–System wird als das *von $\mathcal{E}$ erzeugte Dynkin–System* bezeichnet und das Mengensystem $\mathcal{E}$ wird als *Erzeuger* des Dynkin–Systems $\delta(\mathcal{E})$ bezeichnet.

1.3.5 Folgerung. *Sei $\mathcal{E} \subseteq 2^\Omega$ ein Mengensystem. Dann gilt*

$$\delta(\mathcal{E}) \subseteq \sigma(\mathcal{E})$$

Beweis. Die Behauptung folgt aus Lemma 1.3.1. □

Das folgende Beispiel zeigt, dass das von einem Mengensystem erzeugte Dynkin–System keine σ–Algebra zu sein braucht:

1.3.6 Beispiel. Sei $\Omega := \{1,2,3,4\}$ und sei $A := \{1,2\}$ und $B := \{1,3\}$. Für das Mengensystem $\mathcal{E} := \{A,B\}$ gilt dann $\delta(\mathcal{E}) = \{\emptyset, A, \overline{A}, B, \overline{B}, \Omega\} \neq 2^\Omega = \sigma(\mathcal{E})$.

Unter einer zusätzlichen Bedingung an das Mengensystem $\mathcal{E}$, die wir im nächsten Abschnitt behandeln, gilt jedoch $\delta(\mathcal{E}) = \sigma(\mathcal{E})$.

1.4 $\cap$–stabile Mengensysteme

Ein Mengensystem $\mathcal{C} \subseteq 2^\Omega$ heißt $\cap$*–stabil*, wenn für alle $A, B \in \mathcal{C}$

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

gilt. Offenbar ist jede Topologie und jede σ–Algebra $\cap$–stabil. Andererseits muss ein Dynkin–System nicht $\cap$–stabil sein; vgl. Beispiel 1.3.6. Der folgende Satz klärt die Beziehung zwischen σ–Algebren und Dynkin–Systemen:

1.4.1 Satz. *Für ein Mengensystem $\mathcal{F} \subseteq 2^\Omega$ sind folgende Aussagen äquivalent*:
(a) *$\mathcal{F}$ ist eine σ–Algebra.*
(b) *$\mathcal{F}$ ist ein $\cap$–stabiles Dynkin–System.*

Beweis. Wegen Lemma 1.2.2 und Lemma 1.3.1 ist jede σ–Algebra ein $\cap$–stabiles Dynkin–System.
Sei nun $\mathcal{F}$ ein $\cap$–stabiles Dynkin–System.
(i) Es gilt $\Omega \in \mathcal{F}$.
(ii) Für jede Menge $A \in \mathcal{F}$ gilt $\overline{A} \in \mathcal{F}$.
(iii) Sei $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ und

$$A := \bigcup_{n=1}^{\infty} A_n$$

Um zu zeigen, dass $A \in \mathcal{F}$ gilt, konstruieren wir eine disjunkte Folge $\{B_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ mit

$$A = \sum_{n=1}^{\infty} B_n$$

Die Folge $\{C_n\}_{n\in\mathbb{N}_0}$ mit

$$C_n := \bigcup_{k=1}^{n} A_k$$

(und damit $C_0 = \emptyset$) ist monoton wachsend und die Folge $\{B_n\}_{n\in\mathbb{N}}$ mit

$$B_n := C_n \setminus C_{n-1}$$

ist disjunkt mit

$$C_n = \sum_{k=1}^{n} B_k$$

Daher gilt

$$A = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{n} A_k = \bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} \sum_{k=1}^{n} B_k = \sum_{n=1}^{\infty} B_n$$

Außerdem gilt für alle $n \in \mathbb{N}$

$$\begin{aligned} B_n &= C_n \setminus C_{n-1} \\ &= \left(\bigcup_{k=1}^{n} A_k \right) \setminus \bigcup_{k=1}^{n-1} A_k \\ &= A_n \setminus \bigcup_{k=1}^{n-1} A_k \\ &= A_n \cap \overline{\bigcup_{k=1}^{n-1} A_k} \\ &= A_n \cap \bigcap_{k=1}^{n-1} \overline{A_k} \end{aligned}$$

Da $\mathcal{F}$ ein $\cap$–stabiles Dynkin–System ist, folgt daraus $\{B_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ und damit $A \in \mathcal{F}$.

Damit ist gezeigt, dass $\mathcal{F}$ eine σ–Algebra ist. □

Der folgende Satz zeigt, dass für einen $\cap$–stabilen Erzeuger das erzeugte Dynkin–System mit der erzeugten σ–Algebra übereinstimmt:

1.4.2 Satz. *Sei $\mathcal{E} \subseteq 2^\Omega$ ein $\cap$–stabiles Mengensystem. Dann ist $\delta(\mathcal{E})$ $\cap$–stabil und es gilt*

$$\delta(\mathcal{E}) = \sigma(\mathcal{E})$$

Beweis. Ist $\mathcal{E}$ leer, so gilt $\delta(\mathcal{E}) = \{\emptyset, \Omega\} = \sigma(\mathcal{E})$. Sei also $\mathcal{E}$ nichtleer. Nach Folgerung 1.3.5 gilt

$$\delta(\mathcal{E}) \subseteq \sigma(\mathcal{E})$$

Zum Beweis der Inklusion

$$\sigma(\mathcal{E}) \subseteq \delta(\mathcal{E})$$

genügt es zu zeigen, dass das Dynkin–System $\delta(\mathcal{E})$ $\cap$–stabil ist, denn dann ist $\delta(\mathcal{E})$ nach Satz 1.4.1 eine σ–Algebra, die $\mathcal{E}$ und damit auch $\sigma(\mathcal{E})$ enthält.

Für $B \in \delta(\mathcal{E})$ sei

$$\mathcal{D}_B := \Big\{ C \in 2^\Omega \;\Big|\; C \cap B \in \delta(\mathcal{E}) \Big\}$$

Dann ist $\mathcal{D}_B$ ein Dynkin–System:
(i) Es gilt $\Omega \cap B = B \in \delta(\mathcal{E})$. Daher gilt $\Omega \in \mathcal{D}_B$.
(ii) Sei $C \in \mathcal{D}_B$. Dann gilt $C \cap B \in \delta(\mathcal{E})$ und $C \cap B \subseteq B \in \delta(\mathcal{E})$. Aus Lemma 1.3.2 folgt $\overline{C} \cap B = B \setminus C = B \setminus (C \cap B) \in \delta(\mathcal{E})$. Daher gilt $\overline{C} \in \mathcal{D}_B$.
(iii) Sei $\{C_n\}_{n\in\mathbb{N}} \subseteq \mathcal{D}_B$ disjunkt. Dann ist auch die Folge $\{C_n \cap B\}_{n\in\mathbb{N}} \subseteq \delta(\mathcal{E})$ disjunkt und es gilt $(\sum_{n\in\mathbb{N}} C_n) \cap B = \sum_{n\in\mathbb{N}}(C_n \cap B) \in \delta(\mathcal{E})$. Daher gilt $\sum_{n\in\mathbb{N}} C_n \in \mathcal{D}_B$.

Sei $E \in \mathcal{E}$. Für alle $F \in \mathcal{E}$ gilt, da $\mathcal{E}$ $\cap$–stabil ist, $F \cap E \in \mathcal{E} \subseteq \delta(\mathcal{E})$ und damit $F \in \mathcal{D}_E$. Daraus folgt zunächst $\mathcal{E} \subseteq \mathcal{D}_E$ und sodann

$$\delta(\mathcal{E}) \subseteq \mathcal{D}_E$$

Sei $D \in \delta(\mathcal{E})$. Für alle $E \in \mathcal{E}$ ergibt sich aus der letzten Inklusion $D \in \mathcal{D}_E$ und damit $E \in \mathcal{D}_D$. Daraus folgt zunächst $\mathcal{E} \subseteq \mathcal{D}_D$ und sodann

$$\delta(\mathcal{E}) \subseteq \mathcal{D}_D$$

Für alle $C, D \in \delta(\mathcal{E})$ ergibt sich aus der letzten Inklusion $C \in \mathcal{D}_D$ und damit $C \cap D \in \delta(\mathcal{E})$. Daher ist das Dynkin–System $\delta(\mathcal{E})$ $\cap$–stabil. □

Aufgaben

1.4.A Der Durchschnitt einer nichtleeren Familie von $\cap$–stabilen Mengensystemen auf Ω ist ein $\cap$–stabiles Mengensystem.

1.4.B Ein Mengensystem $\mathcal{C} \subseteq 2^\Omega$ heißt $\cup$–*stabil*, wenn für alle $A, B \in \mathcal{C}$

$$A \cup B \in \mathcal{C}$$

gilt. Jede Topologie und jede σ–Algebra ist ein $\cup$–stabiles Mengensystem.

1.4.C Ein Mengensystem $\mathcal{C} \subseteq 2^\Omega$ heißt *Verband*, wenn es $\cup$–stabil und $\cap$–stabil ist. Jede Topologie und jede σ–Algebra ist ein Verband.

1.5 Halbringe und Ringe

Ein Mengensystem $\mathcal{H} \subseteq 2^\Omega$ heißt *Halbring* auf Ω, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\emptyset \in \mathcal{H}$.
(ii) Für alle $A, B \in \mathcal{H}$ gilt $A \cap B \in \mathcal{H}$.
(iii) Für alle $A, B \in \mathcal{H}$ gibt es eine endliche disjunkte Familie $\{C_j\}_{j\in J} \subseteq \mathcal{H}$ mit $A \setminus B = \sum_{j\in J} C_j$.

Der Begriff des Halbringes ist durch das folgende Beispiel motiviert:

1.5.1 Beispiel (Halbring der halboffenen Intervalle auf $\mathbb{R}^n$). Wir setzen

$$
\begin{aligned}
(\mathbf{a},\mathbf{b}) &:= \Big\{\mathbf{x} \in \mathbb{R}^n \;\Big|\; \mathbf{a} < \mathbf{x} < \mathbf{b}\Big\} \\
(\mathbf{a},\mathbf{b}] &:= \Big\{\mathbf{x} \in \mathbb{R}^n \;\Big|\; \mathbf{a} < \mathbf{x} \leq \mathbf{b}\Big\} \\
[\mathbf{a},\mathbf{b}] &:= \Big\{\mathbf{x} \in \mathbb{R}^n \;\Big|\; \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}\Big\}
\end{aligned}
$$

und bezeichnen
- $(\mathbf{a},\mathbf{b})$ als *offenes Intervall*,
- $(\mathbf{a},\mathbf{b}]$ als *halboffenes Intervall* und
- $[\mathbf{a},\mathbf{b}]$ als *abgeschlossenes Intervall*.

Das Mengensystem

$$
\mathcal{J}(\mathbb{R}^n) := \Big\{(\mathbf{a},\mathbf{b}] \;\Big|\; \mathbf{a},\mathbf{b} \in \mathbb{R}^n \text{ mit } \mathbf{a} \leq \mathbf{b}\Big\}
$$

ist ein Halbring und wird als *Halbring der halboffenen Intervalle* auf $\mathbb{R}^n$ bezeichnet.

Das folgende Beispiel zeigt, dass die Borelsche σ–Algebra auf $\mathbb{R}^n$ auch durch den Halbring der halboffenen Intervalle des $\mathbb{R}^n$ erzeugt wird:

1.5.2 Beispiel (Borelsche σ–Algebra auf $\mathbb{R}^n$). Es gilt

$$
\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{J}(\mathbb{R}^n))
$$

Außerdem enthält die Borelsche σ–Algebra $\mathcal{B}(\mathbb{R}^n)$ alle offenen und alle abgeschlossenen Intervalle und jede abzählbare Teilmenge des $\mathbb{R}^n$.
In der Tat: Sei $\mathbf{x} \in \mathbb{R}^n$ und $\varepsilon \in (0,\infty)$. Dann gilt

$$
B_{\|\,.\,\|_\infty}(\mathbf{x},\varepsilon) = (\mathbf{x}-\varepsilon\mathbf{1}, \mathbf{x}+\varepsilon\mathbf{1}) = \bigcup_{n\in\mathbb{N}} \left(\mathbf{x}-\varepsilon\mathbf{1}, \mathbf{x}+\varepsilon\mathbf{1} - \frac{1}{n}\,\mathbf{1}\right]
$$

und damit $B_{\|\,.\,\|_\infty}(\mathbf{x},\varepsilon) \in \sigma(\mathcal{J}(\mathbb{R}^n))$. Daher gilt $\mathcal{E}_{\|\,.\,\|_\infty} \subseteq \sigma(\mathcal{J}(\mathbb{R}^n))$ und insbesondere

$$
\mathcal{E}_{\|\,.\,\|_\infty,\mathbb{Q}} \subseteq \sigma(\mathcal{J}(\mathbb{R}^n))
$$

Nach Beispiel 1.1.8 ist das Mengensystem $\mathcal{E}_{\|\,.\,\|_\infty,\mathbb{Q}}$ eine abzählbare Basis der Topologie $\mathcal{T}_{\|\,.\,\|_\infty}$. Daraus ergibt sich nun $\mathcal{T}_{\|\,.\,\|_\infty} \subseteq \sigma(\mathcal{J}(\mathbb{R}^n))$ und damit

$$
\sigma(\mathcal{T}_{\|\,.\,\|_\infty}) \subseteq \sigma(\mathcal{J}(\mathbb{R}^n))
$$

Sei nun $J \in \mathcal{J}(\mathbb{R}^n)$. Dann gibt es $\mathbf{a},\mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} \leq \mathbf{b}$ und

$$
J = (\mathbf{a},\mathbf{b}] = \bigcap_{n\in\mathbb{N}} \left(\mathbf{a}, \mathbf{b} + \frac{1}{n}\,\mathbf{1}\right)
$$

Da jedes offene Intervall in $\mathbb{R}^n$ eine Menge in $\mathcal{T}_{\|\,.\,\|_\infty}$ ist, folgt daraus $J \in \sigma(\mathcal{T}_{\|\,.\,\|_\infty})$. Daher gilt $\mathcal{J}(\mathbb{R}^n) \subseteq \sigma(\mathcal{T}_{\|\,.\,\|_\infty})$ und damit

$$
\sigma(\mathcal{J}(\mathbb{R}^n)) \subseteq \sigma(\mathcal{T}_{\|\,.\,\|_\infty})
$$

Daher gilt

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{T}_{\|\,.\,\|_\infty}) = \sigma(\mathcal{J}(\mathbb{R}^n))$$

Außerdem gilt für alle $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} \le \mathbf{b}$

$$[\mathbf{a}, \mathbf{b}] = \bigcap_{n \in \mathbb{N}} \left(\mathbf{a} - \frac{1}{n}\mathbf{1}, \mathbf{b} \right]$$

und damit $[\mathbf{a}, \mathbf{b}] \in \mathcal{B}(\mathbb{R}^n)$, und für alle $\mathbf{x} \in \mathbb{R}^n$ ergibt sich dann $\{\mathbf{x}\} = [\mathbf{x}, \mathbf{x}] \in \mathcal{B}(\mathbb{R}^n)$. Damit ist auch die abschließende Behauptung gezeigt.

Wir betrachten nun Intervalle auf $\bar{\mathbb{R}}$:

1.5.3 Beispiel (Halbring der halboffenen Intervalle auf $\bar{\mathbb{R}}$). Für $a, b \in \bar{\mathbb{R}}$ setzen wir

$$(a, b) := \left\{ x \in \bar{\mathbb{R}} \;\middle|\; a < x < b \right\}$$
$$(a, b] := \left\{ x \in \bar{\mathbb{R}} \;\middle|\; a < x \le b \right\}$$
$$[a, b] := \left\{ x \in \bar{\mathbb{R}} \;\middle|\; a \le x \le b \right\}$$

und bezeichnen
– (a, b) als *offenes Intervall*,
– $(a, b]$ als *halboffenes Intervall* und
– $[a, b]$ als *abgeschlossenes Intervall*.

Das Mengensystem

$$\mathcal{J}(\bar{\mathbb{R}}) := \left\{ (a, b] \;\middle|\; a, b \in \bar{\mathbb{R}} \text{ mit } a \le b \right\}$$

ist ein Halbring und wird als *Halbring der halboffenen Intervalle* auf $\bar{\mathbb{R}}$ bezeichnet.

1.5.4 Beispiel (Borelsche σ–Algebra auf $\bar{\mathbb{R}}$). Es gilt

$$\mathcal{B}(\bar{\mathbb{R}}) = \sigma(\mathcal{J}(\bar{\mathbb{R}}))$$

In der Tat: Sei $(a, b] \in \mathcal{J}(\bar{\mathbb{R}})$.
– Im Fall $b = -\infty$ gilt $(a, b] \cap \mathbb{R} = \emptyset \in \mathcal{B}(\mathbb{R})$ und damit $(a, b] \in \mathcal{B}(\bar{\mathbb{R}})$.
– Im Fall $b \in \mathbb{R}$ gilt $(a, b] \cap \mathbb{R} = (a, b] \in \mathcal{B}(\mathbb{R})$ und damit $(a, b] \in \mathcal{B}(\bar{\mathbb{R}})$.
– Im Fall $b = \infty$ gilt $(a, b] \cap \mathbb{R} = (a, \infty) \in \mathcal{B}(\mathbb{R})$ und damit $(a, b] \in \mathcal{B}(\bar{\mathbb{R}})$.

Daraus folgt zunächst $\mathcal{J}(\bar{\mathbb{R}}) \subseteq \mathcal{B}(\bar{\mathbb{R}})$ und sodann

$$\sigma(\mathcal{J}(\bar{\mathbb{R}})) \subseteq \mathcal{B}(\bar{\mathbb{R}})$$

Sei nun $Q \in \mathcal{B}(\bar{\mathbb{R}})$. Dann gilt $Q \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})$ und damit
– Im Fall $Q \in \mathcal{B}(\mathbb{R})$ gilt $Q \in \sigma(\mathcal{J}(\mathbb{R})) \subseteq \sigma(\mathcal{J}(\bar{\mathbb{R}}))$.
– Im Fall $Q = \{\infty\}$ gilt $Q = (-\infty, \infty] \setminus \mathbb{R} \in \sigma(\mathcal{J}(\bar{\mathbb{R}}))$.
– Im Fall $Q = \{-\infty\}$ gilt $Q = \bar{\mathbb{R}} \setminus (-\infty, \infty] \in \sigma(\mathcal{J}(\bar{\mathbb{R}}))$.

Daraus folgt zunächst $\mathcal{B}(\mathbb{R}) \cup \{\infty\} \cup \{-\infty\} \subseteq \sigma(\mathcal{J}(\bar{\mathbb{R}}))$ und sodann

$$\mathcal{B}(\bar{\mathbb{R}}) \subseteq \sigma(\mathcal{J}(\bar{\mathbb{R}}))$$

Damit ist die Behauptung gezeigt.

Wir betrachten abschließend eine weitere Klasse von Mengensystemen, die mit Halbringen verwandt sind:

Ein Mengensystem $\mathcal{R} \subseteq 2^\Omega$ heißt *Ring* auf Ω, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\emptyset \in \mathcal{R}$.
(ii) Für alle $A, B \in \mathcal{R}$ gilt $A \setminus B \in \mathcal{R}$.
(iii) Für alle $A, B \in \mathcal{R}$ gilt $A \cup B \in \mathcal{R}$.

1.5.5 Lemma. *Jeder Ring ist ein Halbring.*

Beweis. Wegen $A \cap B = (A \cup B) \setminus ((A \setminus B) \cup (B \setminus A))$ ist jeder Ring $\cap$–stabil. Daraus folgt die Behauptung. $\square$

Das folgende Beispiel zeigt, dass die Umkehrung der Implikation von Lemma 1.5.5 im allgemeinen falsch ist:

1.5.6 Beispiel. Sei $\Omega := \{1, 2, 3\}$ und $\mathcal{H} := \{\emptyset, \{1\}, \{2\}, \{3\}, \Omega\}$. Dann ist $\mathcal{H}$ ein Halbring, aber kein Ring.

Der Durchschnitt einer nichtleeren Familie von Ringen auf Ω ist ein Ring; vgl. Aufgabe 1.5.F. Daher gibt es zu jedem Mengensystem $\mathcal{E} \subseteq 2^\Omega$ einen kleinsten Ring

$$\varrho(\mathcal{E})$$

auf Ω, der $\mathcal{E}$ enthält. Dieser Ring wird als der *von $\mathcal{E}$ erzeugte Ring* bezeichnet und das Mengensystem $\mathcal{E}$ wird als *Erzeuger* des Ringes $\varrho(\mathcal{E})$ bezeichnet. Der folgende Satz gibt eine explizite Darstellung des von einem Halbring erzeugten Ringes:

1.5.7 Satz (Der von einem Halbring erzeugte Ring). *Sei $\mathcal{H} \subseteq 2^\Omega$ ein Halbring und sei*

$$\mathcal{R} := \left\{ A \in 2^\Omega \;\middle|\; A = \sum_{i \in I} H_i \text{ mit } I \text{ endlich und } \{H_i\}_{i \in I} \subseteq \mathcal{H} \text{ disjunkt} \right\}$$

Dann gilt $\mathcal{R} = \varrho(\mathcal{H})$.

Beweis. Es gilt $\mathcal{H} \subseteq \mathcal{R} \subseteq \varrho(\mathcal{H})$. Daher genügt es zu zeigen, dass $\mathcal{R}$ ein Ring ist.
(i) Wegen $\emptyset \in \mathcal{H}$ und $\mathcal{H} \subseteq \mathcal{R}$ gilt $\emptyset \in \mathcal{R}$.
(ii) Sei $A, B \in \mathcal{R}$. Dann gibt es endliche disjunkte Familien $\{G_i\}_{i \in I} \subseteq \mathcal{H}$ und $\{H_j\}_{j \in J} \subseteq \mathcal{H}$ mit $A = \sum_{i \in I} G_i$ und $B = \sum_{j \in J} H_j$, und es gilt

$$A \setminus B = \left(\sum_{i \in I} G_i \right) \setminus \left(\sum_{j \in J} H_j \right)$$

$$
\begin{aligned}
&= \sum_{i\in I} G_i \cap \bigcap_{j\in J} \overline{H_j} \\
&= \sum_{i\in I} \bigcap_{j\in J} G_i \cap \overline{H_j} \\
&= \sum_{i\in I} \bigcap_{j\in J} G_i \setminus H_j
\end{aligned}
$$

Da $\mathcal{H}$ ein Halbring ist, gilt $G_i \setminus H_j \in \mathcal{R}$. Da $\mathcal{H}$ $\cap$–stabil ist, ist auch $\mathcal{R}$ $\cap$–stabil und es gilt $\bigcap_{j\in J} G_i \setminus H_j \in \mathcal{R}$. Da die Familie $\{G_i\}_{i\in I}$ disjunkt ist, gilt sogar $A \setminus B = \sum_{i\in I} \bigcap_{j\in J} G_i \setminus H_j \in \mathcal{R}$.

(iii) Sei $A, B \in \mathcal{R}$. Da $\mathcal{R}$ $\cap$–stabil ist, gilt $A \cap B \in \mathcal{R}$, und aus (ii) folgt $A \setminus B \in \mathcal{R}$ und $B \setminus A \in \mathcal{R}$. Wegen $A \cup B = (A \setminus B) + (A \cap B) + (B \setminus A)$ gilt dann auch $A \cup B \in \mathcal{R}$.

Daher ist $\mathcal{R}$ ein Ring. □

Aufgaben

1.5.A Sei $\mathcal{H} \subseteq 2^\Omega$ ein Halbring und sei $\{A_i\}_{i\in I} \subseteq \mathcal{H}$ eine endliche Familie. Dann gibt es eine endliche disjunkte Familie $\{C_j\}_{j\in J} \subseteq \mathcal{H}$ und eine Familie $\{J_i\}_{i\in I} \subseteq 2^J$ mit $\bigcup_{i\in I} J_i = J$ derart, dass

$$\bigcup_{i\in I} A_i = \sum_{j\in J} C_j$$

und für alle $i \in I$

$$A_i = \sum_{j\in J_i} C_j$$

gilt.

1.5.B Ist der Durchschnitt einer nichtleeren Familie von Halbringen auf Ω ein Halbring?

1.5.C Jedes offene Intervall des $\mathbb{R}^n$ ist eine offene Menge bezüglich der natürlichen Topologie, und jedes abgeschlossene Intervall des $\mathbb{R}^n$ ist eine abgeschlossene Menge bezüglich der natürlichen Topologie.

1.5.D **Borelsche σ–Algebra auf $\mathbb{R}^n$:** Sei

$$\mathcal{J}(\mathbb{Q}^n) := \Big\{ (\mathbf{a}, \mathbf{b}] \,\Big|\, \mathbf{a}, \mathbf{b} \in \mathbb{Q}^n \Big\}$$

Dann gilt $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{J}(\mathbb{Q}^n))$.

1.5.E Jeder Ring ist ein Verband.

1.5.F Der Durchschnitt einer nichtleeren Familie von Ringen auf Ω ist ein Ring.

1.5.G Für ein Mengensystem $\mathcal{R} \subseteq 2^\Omega$ sind folgende Aussagen äquivalent:
(a) $\mathcal{R}$ ist ein Ring.
(b) $\mathcal{R}$ ist ein $\cup$–stabiler Halbring.

1.5.H **Ideale:** Sei $\mathcal{R} \subseteq 2^\Omega$ ein Ring. Ein Mengensystem $\mathcal{I} \subseteq \mathcal{R}$ heißt *Ideal* in $\mathcal{R}$, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\emptyset \in \mathcal{I}$.
(ii) Für alle $A \in \mathcal{I}$ und alle $B \in \mathcal{R}$ mit $B \subseteq A$ gilt $B \in \mathcal{I}$.
(iii) Für alle $A, B \in \mathcal{I}$ gilt $A \cup B \in \mathcal{I}$.
Jedes Ideal in einem Ring ist ein Ring.

1.5.I **Algebren:** Ein Mengensystem $\mathcal{R} \subseteq 2^\Omega$ heißt *Algebra* auf Ω, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\Omega \in \mathcal{R}$.
(ii) Für alle $A \in \mathcal{R}$ gilt $\overline{A} \in \mathcal{R}$.
(iii) Für alle $A, B \in \mathcal{R}$ gilt $A \cup B \in \mathcal{R}$.
Jede σ–Algebra ist eine Algebra, und jede Algebra ist ein Ring.

1.5.J σ–**Ringe:** Ein Mengensystem $\mathcal{R} \subseteq 2^\Omega$ heißt σ*–Ring* auf Ω, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\emptyset \in \mathcal{R}$.
(ii) Für alle $A, B \in \mathcal{R}$ gilt $A \setminus B \in \mathcal{R}$.
(iii) Für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{R}$ gilt $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{R}$.
Jede σ–Algebra ist ein σ–Ring, und jeder σ–Ring ist ein Ring.

1.5.K σ–**Ideale:** Sei $\mathcal{R} \subseteq 2^\Omega$ ein σ–Ring. Ein Mengensystem $\mathcal{I} \subseteq \mathcal{R}$ heißt σ*–Ideal* in $\mathcal{R}$, wenn es die folgenden Eigenschaften besitzt:
(i) Es gilt $\emptyset \in \mathcal{I}$.
(ii) Für alle $A \in \mathcal{I}$ und für alle $B \in \mathcal{R}$ mit $B \subseteq A$ gilt $B \in \mathcal{I}$.
(iii) Für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{I}$ gilt $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{I}$.
Jedes σ–Ideal in einem σ–Ring ist ein σ–Ring.

1.5.L Stellen Sie die Inklusionen zwischen den in diesem Kapitel betrachteten Klassen von Mengensystemen in einem Diagramm dar.

1.5.M **Borelsche σ–Algebra auf $\mathbb{R}$:** Sei

$$\mathcal{E} := \Big\{ (-\infty, c] \,\Big|\, c \in \mathbb{R} \Big\}$$

Dann gilt $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{E})$.

1.5.N **Borelsche σ–Algebra auf $\bar{\mathbb{R}}$:** Sei

$$\bar{\mathcal{E}} := \Big\{ [-\infty, c] \,\Big|\, c \in \mathbb{R} \Big\}$$

Dann gilt $\mathcal{B}(\bar{\mathbb{R}}) = \sigma(\bar{\mathcal{E}})$.

2

Topologische Räume und messbare Räume

In diesem Kapitel betrachten wir Abbildungen zwischen Mengen, die beide mit einer Topologie oder beide mit einer σ–Algebra ausgestattet sind. In beiden Fällen sind die wichtigsten Abbildungen durch Eigenschaften ihres Urbildes definiert.

Wir erinnern zunächst an die Eigenschaften des Urbildes einer Abbildung und definieren das Urbild eines Mengensystems (Abschnitt 2.1). Wir betrachten sodann topologische Räume und stetige Abbildungen zwischen topologischen Räumen (Abschnitt 2.2) sowie messbare Räume und messbare Abbildungen zwischen messbaren Räumen (Abschnitt 2.3).

Die Beweise der allgemeinen Ergebnisse über topologische oder messbare Räume und über stetige oder messbare Abbildungen verlaufen analog und werden daher nur für topologische Räume und stetige Abbildungen ausgeführt.

Für Abbildungen zwischen topologischen Räumen ergibt sich ein interessanter Zusammenhang zwischen der Stetigkeit bezüglich den gegebenen Topologien und der Messbarkeit bezüglich den zugehörigen Borelschen σ–Algebren.

2.1 Urbilder von Mengensystemen

Im gesamten Abschnitt seien Ω und Ω' nichtleere Mengen und sei $f : \Omega \to \Omega'$ eine Abbildung.

Für $A \in 2^\Omega$ heißt die Menge

$$f(A) := \left\{\omega' \in \Omega' \,\middle|\, \omega' = f(\omega) \text{ für ein } \omega \in A\right\}$$

das *Bild* von A unter f.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_2, © Springer-Verlag Berlin Heidelberg 2011

Die Abbildung $f^{-1} : 2^{\Omega'} \to 2^{\Omega}$ mit

$$f^{-1}(A') := \left\{ \omega \in \Omega \,\middle|\, f(\omega) \in A' \right\}$$

heißt *Urbild* von f und für $A' \in 2^{\Omega'}$ heißt die Menge $f^{-1}(A')$ das *Urbild* von A' unter f. Das Urbild einer Abbildung kommutiert mit den grundlegenden Operationen der Mengenlehre:

2.1.1 Lemma.
(1) *Es gilt* $f^{-1}(\Omega') = \Omega$ *und* $f^{-1}(\emptyset) = \emptyset$.
(2) *Für alle* $A' \in 2^{\Omega'}$ *gilt*

$$f^{-1}(\overline{A'}) = \overline{f^{-1}(A')}$$

(3) *Für jede Familie* $\{A'_i\}_{i \in I} \subseteq 2^{\Omega'}$ *gilt*

$$f^{-1}\left(\bigcup_{i \in I} A'_i \right) = \bigcup_{i \in I} f^{-1}(A'_i)$$

und

$$f^{-1}\left(\bigcap_{i \in I} A'_i \right) = \bigcap_{i \in I} f^{-1}(A'_i)$$

Insbesondere sind die Urbilder disjunkter Mengen disjunkt.

Für ein Mengensystem $\mathcal{E}' \subseteq 2^{\Omega'}$ heißt das Mengensystem

$$f^{-1}(\mathcal{E}') := \left\{ A \in 2^{\Omega} \,\middle|\, A = f^{-1}(A') \text{ mit } A' \in \mathcal{E}' \right\}$$

das *Urbild* von $\mathcal{E}'$ unter f.

2.1.2 Folgerung.
(1) *Das Urbild einer Topologie ist eine Topologie.*
(2) *Das Urbild einer σ–Algebra ist eine σ–Algebra.*

Aufgabe

2.1.A Für alle $A', B' \in 2^{\Omega'}$ gilt $f^{-1}(A' \setminus B') = f^{-1}(A') \setminus f^{-1}(B')$.

2.1.B Für welche der in Kapitel 1 eingeführten Klassen von Mengensystemen besitzt das Urbild $f^{-1}(\mathcal{E}')$ eines Mengensystems $\mathcal{E}' \subseteq 2^{\Omega'}$ dieselben Eigenschaften wie $\mathcal{E}'$?

2.1.C Sei $\Omega := \{1, 2, 3\}$ und $\Omega' := \{1, 2, 3, 4\}$ und sei $f : \Omega \to \Omega'$ gegeben durch

$$f(\omega) := \omega$$

Dann gibt es ein Dynkin–System $\mathcal{D}'$ auf Ω' derart, dass $f^{-1}(\mathcal{D}')$ kein Dynkin–System ist.

2.2 Topologische Räume und stetige Abbildungen

Ist Ω eine nichtleere Menge und $\mathcal{T}$ eine Topologie auf Ω, so wird das Paar $(\Omega, \mathcal{T})$ als *topologischer Raum* bezeichnet.

Sind $(\Omega, \mathcal{T})$ und $(\Omega', \mathcal{T}')$ topologische Räume, so heißt eine Abbildung $f : \Omega \to \Omega'$ *$\mathcal{T}$–$\mathcal{T}'$–stetig* oder *$\mathcal{T}$–stetig* oder *stetig für $\mathcal{T}$* oder kurz *stetig*, wenn

$$f^{-1}(\mathcal{T}') \subseteq \mathcal{T}$$

gilt.

Ist Ω eine nichtleere Menge und $(\Omega', \mathcal{T}')$ ein topologischer Raum, so ist jede Abbildung $f : \Omega \to \Omega'$ stetig für die Topologie 2^Ω und aus Folgerung 1.1.4 ergibt sich nun, dass es eine kleinste Topologie

$$\tau(f)$$

auf Ω gibt, für die f stetig ist; diese Topologie wird als die von f *erzeugte Topologie* bezeichnet. Nach Folgerung 2.1.2 ist auch das Mengensystem $f^{-1}(\mathcal{T}')$ eine Topologie auf Ω, für die f offenbar stetig ist. Wir erhalten damit den folgenden Satz:

2.2.1 Satz. *Sei Ω eine nichtleere Menge und sei $(\Omega', \mathcal{T}')$ ein topologischer Raum. Dann gilt für jede Abbildung $f : \Omega \to \Omega'$*

$$\tau(f) = f^{-1}(\mathcal{T}')$$

Sind also $(\Omega, \mathcal{T})$ und $(\Omega', \mathcal{T}')$ topologische Räume, so ist eine Abbildung $f : \Omega \to \Omega'$ genau dann stetig, wenn die von f erzeugte Topologie $\tau(f)$ in der Topologie $\mathcal{T}$ enthalten ist.

2.2.2 Lemma. *Sei $\mathcal{E}' \subseteq 2^{\Omega'}$ ein Mengensystem und sei $f : \Omega \to \Omega'$ eine Abbildung. Dann gilt*

$$\tau(f^{-1}(\mathcal{E}')) = f^{-1}(\tau(\mathcal{E}'))$$

Beweis. Wegen $\mathcal{E}' \subseteq \tau(\mathcal{E}')$ gilt $f^{-1}(\mathcal{E}') \subseteq f^{-1}(\tau(\mathcal{E}'))$ und aus Folgerung 2.1.2 folgt, dass $f^{-1}(\tau(\mathcal{E}'))$ eine Topologie ist. Daher gilt

$$\tau(f^{-1}(\mathcal{E}')) \subseteq f^{-1}(\tau(\mathcal{E}'))$$

Zum Nachweis der umgekehrten Inklusion betrachten wir das Mengensystem

$$\mathcal{T}_0' := \left\{ A' \in 2^{\Omega'} \;\middle|\; f^{-1}(A') \in \tau(f^{-1}(\mathcal{E}')) \right\}$$

Dann gilt $\mathcal{E}' \subseteq \mathcal{T}_0'$ und aus Lemma 2.1.1 folgt, dass $\mathcal{T}_0'$ eine Topologie ist. Daher gilt $\tau(\mathcal{E}') \subseteq \mathcal{T}_0'$ und damit

$$f^{-1}(\tau(\mathcal{E}')) \subseteq f^{-1}(\mathcal{T}_0') \subseteq \tau(f^{-1}(\mathcal{E}'))$$

Damit ist das Lemma bewiesen. □

Der folgende Satz erleichtert den Nachweis der Stetigkeit einer Abbildung zwischen topologischen Räumen:

2.2.3 Satz (Stetige Abbildungen). *Seien $(\Omega, \mathcal{T})$ und $(\Omega', \mathcal{T}')$ topologische Räume und sei $\mathcal{E}'$ ein Erzeuger von $\mathcal{T}'$ und $f : \Omega \to \Omega'$ eine Abbildung. Dann sind folgende Aussagen äquivalent*:
(a) *f ist stetig.*
(b) *Es gilt $f^{-1}(\mathcal{E}') \subseteq \mathcal{T}$.*

Beweis. Wir nehmen zunächst an, dass f stetig ist. Dann gilt $f^{-1}(\mathcal{T}') \subseteq \mathcal{T}$, und wegen $\mathcal{E}' \subseteq \tau(\mathcal{E}') = \mathcal{T}'$ ergibt sich daraus $f^{-1}(\mathcal{E}') \subseteq f^{-1}(\mathcal{T}') \subseteq \mathcal{T}$. Daher folgt (b) aus (a).
Wir nehmen nun an, dass $f^{-1}(\mathcal{E}') \subseteq \mathcal{T}$ gilt. Da $\mathcal{T}$ eine Topologie ist, gilt dann auch $\tau(f^{-1}(\mathcal{E}')) \subseteq \mathcal{T}$, und mit Lemma 2.2.2 erhalten wir $f^{-1}(\mathcal{T}') = f^{-1}(\tau(\mathcal{E}')) = \tau(f^{-1}(\mathcal{E}')) \subseteq \mathcal{T}$. Daher folgt (a) aus (b). □

2.2.4 Lemma (Komposition). *Seien $(\Omega, \mathcal{T})$, $(\Omega', \mathcal{T}')$ und $(\Omega'', \mathcal{T}'')$ topologische Räume und seien $f : \Omega \to \Omega'$ und $g : \Omega' \to \Omega''$ stetig. Dann ist auch die Komposition $g \circ f$ stetig.*

Beweis. Es gilt $(g \circ f)^{-1}(\mathcal{T}'') = f^{-1}(g^{-1}(\mathcal{T}'')) \subseteq f^{-1}(\mathcal{T}') \subseteq \mathcal{T}$. □

Wir betrachten nun stetige Abbildungen zwischen normierten Räumen:

2.2.5 Beispiel (Stetige Abbildungen zwischen normierten Räumen). Seien $(E, \|\,.\,\|)$ und $(E', \|\,.\,\|')$ normierte Räume. Eine Abbildung $f : E \to E'$ heißt *stetig in* $x \in E$, wenn es für alle $\varepsilon \in (0, \infty)$ ein $\delta \in (0, \infty)$ gibt derart, dass für alle $y \in E$ mit $\|y - x\| < \delta$ auch $\|f(y) - f(x)\|' < \varepsilon$ gilt. Für eine Abbildung $f : E \to E'$ sind folgende Aussagen äquivalent:
(a) f ist stetig.
(b) Für alle $x \in E$ ist f stetig in x.
In der Tat: Sei zunächst f stetig und $x \in E$. Für alle $\varepsilon \in (0, \infty)$ gilt $B_{\|.\|'}(f(x), \varepsilon) \in \mathcal{T}_{\|.\|'}$. Aus der Stetigkeit von f folgt $x \in f^{-1}(B_{\|.\|'}(f(x), \varepsilon)) \in \mathcal{T}_{\|.\|}$. Daher gibt es ein $\delta \in (0, \infty)$ mit $B_{\|.\|}(x, \delta) \subseteq f^{-1}(B_{\|.\|'}(f(x), \varepsilon))$. Für alle $y \in E$ mit $\|y - x\| < \delta$ gilt daher $y \in B_{\|.\|}(x, \delta) \subseteq f^{-1}(B_{\|.\|'}(f(x), \varepsilon))$ und damit $f(y) \in B_{\|.\|'}(f(x), \varepsilon)$, also $\|f(y) - f(x)\|' < \varepsilon$. Daher ist f für alle $x \in E$ stetig in x.
Sei nun f für alle $x \in E$ stetig in x und sei $Q' \in \mathcal{T}_{\|.\|'}$. Für alle $x \in f^{-1}(Q')$ gilt $f(x) \in Q'$. Daher gibt es ein $\varepsilon \in (0, \infty)$ mit $B_{\|.\|'}(f(x), \varepsilon) \subseteq Q'$. Da f stetig in x ist, gibt es ein $\delta \in (0, \infty)$ mit $f(B_{\|.\|}(x, \delta)) \subseteq B_{\|.\|'}(f(x), \varepsilon) \subseteq Q'$ und damit $B_{\|.\|}(x, \delta) \subseteq f^{-1}(Q')$. Daraus folgt $f^{-1}(Q') \in \mathcal{T}_{\|.\|}$. Daher ist f stetig.
Damit ist die Behauptung bewiesen.

Aufgaben

2.2.A **Spurtopologie:** Sei $(\Omega, \mathcal{T})$ ein topologischer Raum. Dann ist für jede nichtleere Menge $C \in 2^\Omega$ das Mengensystem

$$\mathcal{T}(C) := \Big\{ D \in 2^\Omega \;\Big|\; D = A \cap C \text{ für eine Menge } A \in \mathcal{T} \Big\}$$

eine Topologie auf C. Die Topologie $\mathcal{T}(C)$ heißt *Spurtopologie* oder kurz *Spur* von $\mathcal{T}$ auf C.

2.2.B **Restriktion:** Seien $(\Omega, \mathcal{T})$ und $(\Omega', \mathcal{T}')$ topologische Räume und sei $f : \Omega \to \Omega'$ stetig. Dann ist für jede nichtleere Menge $C \in 2^\Omega$ die Abbildung $f|_C : C \to \Omega'$ mit

$$f|_C(c) := f(c)$$

$\mathcal{T}(C)$–$\mathcal{T}'$–stetig. Die Abbildung $f|_C$ heißt *Restriktion* von f auf C.

2.2.C Seien $(\Omega, \mathcal{T})$ und $(\Omega', \mathcal{T}')$ topologische Räume und sei $f : \Omega \to \Omega'$ stetig. Ist $K \in 2^\Omega$ kompakt, so ist auch $f(K)$ kompakt.

2.2.D **Bildtopologie:** Sei $(\Omega, \mathcal{T})$ ein topologischer Raum, sei Ω' eine nichtleere Menge und sei $f : \Omega \to \Omega'$ eine Abbildung. Dann ist das Mengensystem

$$\mathcal{T}_f := \Big\{ A' \in 2^{\Omega'} \;\Big|\; f^{-1}(A') \in \mathcal{T} \Big\}$$

eine Topologie auf Ω', für die f stetig ist. Die Topologie $\mathcal{T}_f$ ist die größte Topologie auf Ω', für die f stetig ist. Die Topologie $\mathcal{T}_f$ heißt *Bildtopologie* von $\mathcal{T}$ unter f.

2.2.E **Stetige Abbildungen zwischen metrischen Räumen:** Seien (E, d) und (E', d') metrische Räume. Eine Abbildung $f : E \to E'$ heißt *stetig in* $x \in E$, wenn es für alle $\varepsilon \in (0, \infty)$ ein $\delta \in (0, \infty)$ gibt derart, dass für alle $y \in E$ mit $d(y, x) < \delta$ auch $d'(f(y), f(x)) < \varepsilon$ gilt. Für eine Abbildung $f : E \to E'$ sind folgende Aussagen äquivalent:
(a) f ist stetig.
(b) Für alle $x \in E$ ist f stetig in x.

2.3 Messbare Räume und messbare Abbildungen

Ist Ω eine nichtleere Menge und $\mathcal{F}$ eine σ–Algebra auf Ω, so wird das Paar $(\Omega, \mathcal{F})$ als *messbarer Raum* oder als *Messraum* bezeichnet.

Sind $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ messbare Räume, so heißt eine Abbildung $f : \Omega \to \Omega'$ *$\mathcal{F}$–$\mathcal{F}'$–messbar* oder *$\mathcal{F}$–messbar* oder *messbar für $\mathcal{F}$* oder kurz *messbar*, wenn

$$f^{-1}(\mathcal{F}') \subseteq \mathcal{F}$$

gilt.

Ist Ω eine nichtleere Menge und $(\Omega', \mathcal{F}')$ ein messbarer Raum, so ist jede Abbildung $f : \Omega \to \Omega'$ messbar für die σ–Algebra 2^Ω und aus Folgerung 1.2.4 ergibt sich nun, dass es eine kleinste σ–Algebra

$$\sigma(f)$$

auf Ω gibt, für die f messbar ist; diese σ–Algebra wird als die von f *erzeugte σ–Algebra* bezeichnet. Nach Folgerung 2.1.2 ist auch das Mengensystem $f^{-1}(\mathcal{F}')$ eine σ–Algebra auf Ω, für die f offenbar messbar ist. Wir erhalten damit den folgenden Satz:

2.3.1 Satz. *Sei Ω eine nichtleere Menge und sei $(\Omega', \mathcal{F}')$ ein messbarer Raum. Dann gilt für jede Abbildung $f : \Omega \to \Omega'$*

$$\sigma(f) = f^{-1}(\mathcal{F}')$$

Sind also $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ messbare Räume, so ist eine Abbildung $f : \Omega \to \Omega'$ genau dann messbar, wenn die von f erzeugte σ–Algebra $\sigma(f)$ in der σ–Algebra $\mathcal{F}$ enthalten ist.

2.3.2 Lemma. *Sei $\mathcal{E}' \subseteq 2^{\Omega'}$ ein Mengensystem und sei $f : \Omega \to \Omega'$ eine Abbildung. Dann gilt*

$$\sigma(f^{-1}(\mathcal{E}')) = f^{-1}(\sigma(\mathcal{E}'))$$

Der folgende Satz erleichtert den Nachweis der Messbarkeit einer Abbildung zwischen messbaren Räumen:

2.3.3 Satz (Messbare Abbildungen). *Seien $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ messbare Räume und sei $\mathcal{E}'$ ein Erzeuger von $\mathcal{F}'$ und $f : \Omega \to \Omega'$ eine Abbildung. Dann sind folgende Aussagen äquivalent*:
(a) *f ist messbar.*
(b) *Es gilt $f^{-1}(\mathcal{E}') \subseteq \mathcal{F}$.*

2.3.4 Lemma (Komposition). *Seien $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ und $(\Omega'', \mathcal{F}'')$ messbare Räume und seien $f : \Omega \to \Omega'$ und $g : \Omega' \to \Omega''$ messbar. Dann ist auch die Komposition $g \circ f$ messbar.*

Wir zeigen nun, dass stetige Abbildungen zwischen topologischen Räumen bezüglich den zugehörigen Borelschen σ–Algebren messbar sind:

2.3.5 Satz (Stetigkeit und Messbarkeit). *Sei $\mathcal{T}$ eine Topologie auf Ω und sei $\mathcal{T}'$ eine Topologie auf Ω'. Dann ist jede $\mathcal{T}$–$\mathcal{T}'$–stetige Abbildung $\Omega \to \Omega'$ $\mathcal{B}(\Omega, \mathcal{T})$–$\mathcal{B}(\Omega', \mathcal{T}')$–messbar.*

Beweis. Sei $f : \Omega \to \Omega'$ $\mathcal{T}$–$\mathcal{T}'$–stetig. Dann gilt $f^{-1}(\mathcal{T}') \subseteq \mathcal{T} \subseteq \sigma(\mathcal{T})$ und damit

$$\sigma(f^{-1}(\mathcal{T}')) \subseteq \sigma(\mathcal{T})$$

Aus Lemma 2.3.2 folgt nun

$$f^{-1}(\mathcal{B}(\Omega', \mathcal{T}')) = f^{-1}(\sigma(\mathcal{T}')) = \sigma(f^{-1}(\mathcal{T}')) \subseteq \sigma(\mathcal{T}) = \mathcal{B}(\Omega, \mathcal{T})$$

Daher ist f $\mathcal{B}(\Omega, \mathcal{T})$–$\mathcal{B}(\Omega', \mathcal{T}')$–messbar. □

2.3.6 Beispiel. Jede (bezüglich der natürlichen Topologien auf $\mathbb{R}^m$ und $\mathbb{R}^n$) stetige Abbildung $f : \mathbb{R}^m \to \mathbb{R}^n$ ist (bezüglich der Borelschen σ–Algebren auf $\mathbb{R}^m$ und $\mathbb{R}^n$) messbar.

Andererseits zeigen die folgenden Beispiele, dass eine messbare Abbildung zwischen topologischen Räumen nicht notwendigerweise stetig ist:

2.3.7 Beispiele.
(1) **Dirichlet–Funktion:** Die Dirichlet–Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

ist messbar, aber nicht stetig.
In der Tat: Da $\mathbb{Q}$ abzählbar ist, gilt $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$. Wegen $f^{-1}(\{1\}) = \mathbb{Q}$ und $f^{-1}(\{0\}) = \mathbb{R} \setminus \mathbb{Q}$ gilt für alle $B \in \mathcal{B}(\mathbb{R})$

$$f^{-1}(B) = \begin{cases} \mathbb{R} & \text{falls } 0 \in B \text{ und } 1 \in B \\ \mathbb{R} \setminus \mathbb{Q} & \text{falls } 0 \in B \text{ und } 1 \notin B \\ \mathbb{Q} & \text{falls } 0 \notin B \text{ und } 1 \in B \\ \emptyset & \text{falls } 0 \notin B \text{ und } 1 \notin B \end{cases}$$

und damit $f^{-1}(B) \in \mathcal{B}(\mathbb{R})$. Daher ist f messbar.
(2) **Monotone Funktionen:** Jede monotone Funktion $f : \mathbb{R} \to \mathbb{R}$ ist messbar.
In der Tat: Wir nehmen zunächst an, dass f monoton wachsend ist, und betrachten das Mengensystem

$$\mathcal{E} := \Big\{ (-\infty, c] \;\Big|\; c \in \mathbb{R} \Big\}$$

Dann gilt $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{E})$. Da f monoton wachsend ist, gilt für alle $B \in \mathcal{E}$ entweder $f^{-1}(B) \in \{\emptyset, \mathbb{R}\}$ oder es gibt ein $a \in \mathbb{R}$ mit $f^{-1}(B) \in \{(-\infty, a), (-\infty, a]\}$; in beiden Fällen gilt daher $f^{-1}(B) \in \mathcal{B}(\mathbb{R})$. Daher gilt $f^{-1}(\mathcal{E}) \subseteq \mathcal{B}(\mathbb{R})$, und aus Satz 2.3.3 folgt nun, dass f messbar ist.
Wir nehmen nun an, dass f monoton fallend ist, und betrachten die Funktionen $g, h : \mathbb{R} \to \mathbb{R}$ mit $g(x) := -f(x)$ und $h(x) := -x$. Dann ist g monoton wachsend und damit messbar, h ist stetig und damit messbar, und es gilt $f = h \circ g$. Aus Lemma 2.3.4 folgt nun, dass f messbar ist.

Aufgaben

2.3.A **Spur–σ–Algebra:** Sei $(\Omega, \mathcal{F})$ ein messbarer Raum. Dann ist für jede nichtleere Menge $C \in 2^\Omega$ das Mengensystem

$$\mathcal{F}(C) := \left\{ D \in 2^\Omega \;\middle|\; D = A \cap C \text{ für eine Menge } A \in \mathcal{F} \right\}$$

eine σ–Algebra auf C. Die σ–Algebra $\mathcal{F}(C)$ heißt *Spur–σ–Algebra* oder kurz *Spur* von $\mathcal{F}$ auf C.

2.3.B **Restriktion:** Seien $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ messbare Räume und sei $f : \Omega \to \Omega'$ messbar. Dann ist für jede nichtleere Menge $C \in 2^\Omega$ die Abbildung $f|_C : C \to \Omega'$ mit

$$f|_C(c) := f(c)$$

$\mathcal{F}(C)$–$\mathcal{F}'$–messbar. Die Abbildung $f|_C$ heißt *Restriktion* von f auf C.

2.3.C **Bild–σ–Algebra:** Sei $(\Omega, \mathcal{F})$ ein messbarer Raum, sei Ω' eine nichtleere Menge und sei $f : \Omega \to \Omega'$ eine Abbildung. Dann ist das Mengensystem

$$\mathcal{F}_f := \left\{ A' \in 2^{\Omega'} \;\middle|\; f^{-1}(A') \in \mathcal{F} \right\}$$

eine σ–Algebra auf Ω', für die f messbar ist. Die σ–Algebra $\mathcal{F}_f$ ist die größte σ–Algebra auf Ω', für die f messbar ist. Die σ–Algebra $\mathcal{F}_f$ heißt *Bild–σ–Algebra* von $\mathcal{F}$ unter f.

3

Produkträume

In diesem Kapitel führen wir ein allgemeines Prinzip zur Konstruktion von topologischen oder messbaren Räumen auf dem Produkt der Grundmengen einer Familie von topologischen oder messbaren Räumen ein.

Wir definieren zunächst das Produkt einer Familie von Mengen und das Produkt einer Familie von Mengensystemen (Abschnitt 3.1). Die Forderung der Stetigkeit bzw. der Messbarkeit aller Projektionen des Produktes der Familie der Grundmengen auf seine Koordinaten führt dann in natürlicher Weise auf die Definition des Produktes einer Familie topologischer Räume (Abschnitt 3.2) und auf die Definition des Produktes einer Familie messbarer Räume (Abschnitt 3.3).

Die Beweise der allgemeinen Ergebnisse über Produkte von topologischen oder messbaren Räumen verlaufen analog und werden daher nur für Produkte von topologischen Räumen ausgeführt.

Sei I eine nichtleere Indexmenge und sei $\mathcal{H}(I)$ die Familie der endlichen nichtleeren Teilmengen von I.

3.1 Produkte und Projektionen

Sei $\{\Omega_i\}_{i\in I}$ eine Familie von nichtleeren Mengen. Die Menge aller Abbildungen $\omega : I \to \bigcup_{i\in I} \Omega_i$ mit $\omega(i) \in \Omega_i$ für alle $i \in I$ heißt das *Produkt* der Familie $\{\Omega_i\}_{i\in I}$ und wird mit

$$\prod_{i\in I} \Omega_i$$

bezeichnet. Für $\omega \in \prod_{i\in I} \Omega_i$ und $i \in I$ setzen wir $\omega_i := \omega(i)$ und nennen ω_i die i–te *Koordinate* von ω. Außerdem identifizieren wir jede Abbildung $\omega \in \prod_{i\in I} \Omega_i$ mit der Familie $\{\omega_i\}_{i\in I}$ ihrer Koordinaten. Dann lässt sich das Produkt der Familie $\{\Omega_i\}_{i\in I}$ in der Form

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_3, © Springer-Verlag Berlin Heidelberg 2011

$$\prod_{i \in I} \Omega_i = \left\{ \{\omega_i\}_{i \in I} \,\middle|\, \omega_i \in \Omega_i \text{ für alle } i \in I \right\}$$

darstellen; im Fall $\Omega_i = \Omega$ für alle $i \in I$ setzen wir

$$\Omega^I := \prod_{i \in I} \Omega$$

Für $j \in I$ heißt die Abbildung $\pi_j : \prod_{i \in I} \Omega_i \to \Omega_j$ mit

$$\pi_j(\omega) := \omega_j$$

die *Projektion* auf die j–te Koordinate. Für $j \in I$ und $A_j \in 2^{\Omega_j}$ erhält man für das Urbild von A_j unter π_j die Darstellung

$$\pi_j^{-1}(A_j) = \prod_{i \in I} B_i$$

mit

$$B_i := \begin{cases} A_i & \text{falls } i = j \\ \Omega_i & \text{sonst} \end{cases}$$

und damit eine Darstellung als Produkt von Mengen.

Sei nun $\{\mathcal{E}_i\}_{i \in I}$ eine Familie von Mengensystemen mit $\mathcal{E}_i \subseteq 2^{\Omega_i}$ für alle $i \in I$. Dann heißt das Mengensystem

$$\prod_{i \in I} \mathcal{E}_i = \left\{ \prod_{i \in I} A_i \,\middle|\, A_i \in \mathcal{E}_i \text{ für alle } i \in I \right\}$$

das *Produkt* der Familie $\{\mathcal{E}_i\}_{i \in I}$. Es gilt

$$\prod_{i \in I} \mathcal{E}_i \subseteq 2^{\prod_{i \in I} \Omega_i}$$

Erfüllen die Mengensysteme für alle $i \in I$ die Bedingung $\Omega_i \in \mathcal{E}_i$, so gilt für alle $j \in I$

$$\pi_j^{-1}(\mathcal{E}_j) \subseteq \prod_{i \in I} \mathcal{E}_i$$

Diese Bedingung ist für jede Familie von Topologien oder σ–Algebren erfüllt.

Wir betrachten abschließend den Fall einer abzählbaren Indexmenge I. In diesem Fall existiert eine Menge $N \subseteq \mathbb{N}$ mit $|I| = |N|$ sowie eine Bijektion $I \to N$ und die natürliche Ordnungsrelation auf N induziert eine vollständige Ordnungsrelation auf I, sodass das Produkt der Familie $\{\Omega_i\}_{i \in I}$ durch ein *kartesisches Produkt* dargestellt werden kann, in dem die Koordinaten von $\omega \in \prod_{i \in I} \Omega_i$ vollständig geordnet sind; entsprechend kann auch das Produkt der Familie $\{\mathcal{E}_i\}_{i \in I}$ durch ein kartesisches Produkt dargestellt werden.

3.1.1 Beispiel. Sei $I = \{\circ, \bullet\}$. Dann kann das Produkt $\prod_{i \in I} \Omega_i$ der Familie $\{\Omega_i\}_{i \in I}$ durch die kartesischen Produkte

$$\Omega_\circ \times \Omega_\bullet := \Bigl\{(\omega_\circ, \omega_\bullet) \Bigm| \omega_\circ \in \Omega_\circ, \omega_\bullet \in \Omega_\bullet\Bigr\}$$

und

$$\Omega_\bullet \times \Omega_\circ := \Bigl\{(\omega_\bullet, \omega_\circ) \Bigm| \omega_\bullet \in \Omega_\bullet, \omega_\circ \in \Omega_\circ\Bigr\}$$

dargestellt werden, und das Produkt $\prod_{i \in I} \mathcal{E}_i$ der Familie $\{\mathcal{E}_i\}_{i \in I}$ kann durch die entsprechenden kartesischen Produkte

$$\mathcal{E}_\circ \times \mathcal{E}_\bullet := \Bigl\{A_\circ \times A_\bullet \Bigm| A_\circ \in \mathcal{E}_\circ, A_\bullet \in \mathcal{E}_\bullet\Bigr\}$$

bzw.

$$\mathcal{E}_\bullet \times \mathcal{E}_\circ := \Bigl\{A_\bullet \times A_\circ \Bigm| A_\bullet \in \mathcal{E}_\bullet, A_\circ \in \mathcal{E}_\circ\Bigr\}$$

dargestellt werden.

Das Beispiel zeigt, dass das Produkt einer abzählbaren Familie von Mengen oder Mengensystemen im allgemeinen auf mehrere Weisen durch ein kartesisches Produkt dargestellt werden kann. Dessen ungeachtet werden wir im Fall zweier Faktoren im allgemeinen $\Omega_1 \times \Omega_2$ anstelle von $\prod_{i \in \{1,2\}} \Omega_i$ und $\mathcal{E}_1 \times \mathcal{E}_2$ anstelle von $\prod_{i \in \{1,2\}} \mathcal{E}_i$ schreiben.

Sind $K, L \subseteq I$ nichtleer und disjunkt mit $K + L = I$, so ist die Abbildung

$$\left(\prod_{i \in K} \Omega_i\right) \times \left(\prod_{i \in L} \Omega_i\right) \to \prod_{i \in I} \Omega_i$$

mit

$$\Bigl(\{\omega_i\}_{i \in K}, \{\omega_i\}_{i \in L}\Bigr) \mapsto \{\omega_i\}_{i \in I}$$

eine Bijektion, die es gestattet, diese Produkte miteinander zu identifizieren. Dadurch wird die Bildung von Produkten von Mengen oder Mengensystemen zu einer assoziativen Operation.

Sind Ω und Ω' nichtleere Mengen und ist $f : \Omega \to \Omega'$ eine Abbildung, so heißt die Abbildung $F : \Omega \to \Omega \times \Omega'$ mit

$$F(\omega) := (\omega, f(\omega))$$

der *Graph* von f.

Aufgabe

3.1.A **Halbringe:** Das Produkt einer endlichen Familie von Halbringen ist ein Halbring. Insbesondere gilt $\mathcal{J}(\mathbb{R}^n) = \prod_{i \in \{1,\dots,n\}} \mathcal{J}(\mathbb{R})$.

3.2 Produkte von topologischen Räumen

Sei $\{(\Omega_i, \mathcal{T}_i)\}_{i \in I}$ eine Familie topologischer Räume. Dann gibt es eine kleinste Topologie

$$\bigotimes_{i \in I} \mathcal{T}_i$$

auf $\prod_{i \in I} \Omega_i$, für die alle Projektionen stetig sind. Diese Topologie wird als *Produkttopologie* der Familie $\{\mathcal{T}_i\}_{i \in I}$ bezeichnet und der topologische Raum

$$\bigotimes_{i \in I} (\Omega_i, \mathcal{T}_i) := \left(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{T}_i \right)$$

wird als *topologisches Produkt* der Familie $\{(\Omega_i, \mathcal{T}_i)\}_{i \in I}$ oder als *Produkt der topologischen Räume* $(\Omega_i, \mathcal{T}_i)$ bezeichnet.

Das folgende Beispiel zeigt, dass die natürliche Topologie auf $\mathbb{R}^n$ als Produkttopologie dargestellt werden kann:

3.2.1 Beispiel (Natürliche Topologie auf $\mathbb{R}^n$). Für alle $n \in \mathbb{N}$ gilt

$$\mathcal{T}(\mathbb{R}^n) = \bigotimes_{i \in \{1, \dots, n\}} \mathcal{T}(\mathbb{R})$$

In der Tat: Für alle $\mathbf{x} \in \mathbb{R}^n$ und $\varepsilon \in (0, \infty)$ gilt

$$B_{\|\,.\,\|_\infty}(\mathbf{x}, \varepsilon) = \prod_{i \in \{1, \dots, n\}} B_{|\,.\,|}(x_i, \varepsilon) = \bigcap_{i \in \{1, \dots, n\}} \pi_i^{-1}(B_{|\,.\,|}(x_i, \varepsilon))$$

Da alle Projektionen bezüglich der Produkttopologie stetig sind, ergibt sich daraus $B_{\|\,.\,\|_\infty}(\mathbf{x}, \varepsilon) \in \bigotimes_{i \in \{1, \dots, n\}} \mathcal{T}(\mathbb{R})$. Daher gilt $\mathcal{E}_{\|\,.\,\|_\infty} \subseteq \bigotimes_{i \in \{1, \dots, n\}} \mathcal{T}(\mathbb{R})$ und damit

$$\mathcal{T}_{\|\,.\,\|_\infty} \subseteq \bigotimes_{i \in \{1, \dots, n\}} \mathcal{T}(\mathbb{R})$$

Sei nun $i \in \{1, \dots, n\}$. Für alle $\mathbf{x} \in \mathbb{R}^n$ und für alle $\varepsilon, \delta \in (0, \infty)$ mit $\delta \leq \varepsilon$ gilt

$$\pi_i(B_{\|\,.\,\|_\infty}(\mathbf{x}, \delta)) \subseteq \pi_i(B_{\|\,.\,\|_\infty}(\mathbf{x}, \varepsilon)) = B_{|\,.\,|}(\pi_i(\mathbf{x}), \varepsilon)$$

Daher ist π_i für alle $\mathbf{x} \in \mathbb{R}^n$ stetig in $\mathbf{x}$, und damit stetig bezüglich der Topologie $\mathcal{T}_{\|\,.\,\|_\infty}$. Damit ist gezeigt, dass alle Projektionen bezüglich der Topologie $\mathcal{T}_{\|\,.\,\|_\infty}$ stetig sind, und daraus folgt

$$\bigotimes_{i \in \{1, \dots, n\}} \mathcal{T}(\mathbb{R}) \subseteq \mathcal{T}_{\|\,.\,\|_\infty}$$

Die Behauptung folgt nun aus $\mathcal{T}(\mathbb{R}^n) = \mathcal{T}_{\|\,.\,\|_\infty}$.

Die Darstellung einer Topologie als Produkttopologie kann den Nachweis der Stetigkeit einer Abbildung erleichtern:

3.2.2 Satz (Stetige Abbildungen in ein topologisches Produkt). *Sei* $(\Omega_0, \mathcal{T}_0)$ *ein topologischer Raum und sei* $f : \Omega_0 \to \prod_{i\in I} \Omega_i$ *eine Abbildung. Dann sind folgende Aussagen äquivalent*:
(a) f *ist stetig bezüglich der Produkttopologie auf* $\prod_{i\in I} \Omega_i$.
(b) *Für alle* $i \in I$ *ist* $\pi_i \circ f$ *stetig.*

Beweis. Wir nehmen zunächst an, dass f bezüglich der Produkttopologie auf $\prod_{i\in I} \Omega_i$ stetig ist. Da alle Projektionen bezüglich der Produkttopologie stetig sind, ist dann auch für alle $i \in I$ die Komposition $\pi_i \circ f$ stetig. Daher folgt (b) aus (a).
Wir nehmen nun an, dass für alle $i \in I$ die Komposition $\pi_i \circ f$ stetig ist. Dann gilt

$$f^{-1}\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{T}_i)\right) = \bigcup_{i\in I} f^{-1}\Bigl(\pi_i^{-1}(\mathcal{T}_i)\Bigr) = \bigcup_{i\in I} (\pi_i \circ f)^{-1}(\mathcal{T}_i) \subseteq \mathcal{T}_0$$

und damit

$$f^{-1}\left(\bigotimes_{i\in I} \mathcal{T}_i\right) \subseteq f^{-1}\left(\tau\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{T}_i)\right)\right) = \tau\left(f^{-1}\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{T}_i)\right)\right) \subseteq \mathcal{T}_0$$

Daher folgt (a) aus (b). □

3.2.3 Satz. *Für alle* $i \in I$ *sei* $\mathcal{E}_i$ *ein Mengensystem mit* $\mathcal{T}_i = \tau(\mathcal{E}_i)$. *Dann gilt*

$$\bigotimes_{i\in I} \mathcal{T}_i = \tau\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i)\right)$$

Beweis. Für alle $j \in I$ gilt

$$\pi_j^{-1}(\mathcal{T}_j) = \pi_j^{-1}(\tau(\mathcal{E}_j)) = \tau(\pi_j^{-1}(\mathcal{E}_j)) \subseteq \tau\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i)\right)$$

Daraus folgt, dass alle Projektionen bezüglich der Topologie $\tau\bigl(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i)\bigr)$ stetig sind. Daher gilt

$$\bigotimes_{i\in I} \mathcal{T}_i \subseteq \tau\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i)\right)$$

Die umgekehrte Inklusion ist klar. □

Wir betrachten abschließend das Produkt einer endlichen Familie von Topologien:

3.2.4 Satz. *Sei I endlich. Dann gilt*

$$\bigotimes_{i\in I} \mathcal{T}_i = \tau\left(\prod_{i\in I} \mathcal{T}_i\right)$$

Beweis. Für alle $i \in I$ gilt $\Omega_i \in \mathcal{T}_i$. Daher gilt für alle $j \in I$

$$\pi_j^{-1}(\mathcal{T}_j) \subseteq \prod_{i\in I} \mathcal{T}_i \subseteq \tau\left(\prod_{i\in I} \mathcal{T}_i\right)$$

Daraus folgt, dass alle Projektionen bezüglich der Topologie $\tau\big(\prod_{i\in I} \mathcal{T}_i\big)$ stetig sind. Daher gilt

$$\bigotimes_{i\in I} \mathcal{T}_i \subseteq \tau\left(\prod_{i\in I} \mathcal{T}_i\right)$$

Sei nun $\{A_i\}_{i\in I}$ eine Familie von Mengen mit $A_i \in \mathcal{T}_i$ für alle $i \in I$. Da alle Projektionen bezüglich der Produkttopologie stetig sind, gilt für alle $j \in I$

$$\pi_j^{-1}(A_j) \in \bigotimes_{i\in I} \mathcal{T}_i$$

Da I endlich und jede Topologie $\cap$–stabil ist, folgt daraus

$$\prod_{i\in I} A_i = \bigcap_{i\in I} \pi_i^{-1}(A_i) \in \bigotimes_{i\in I} \mathcal{T}_i$$

Daraus ergibt sich zunächst

$$\prod_{i\in I} \mathcal{T}_i \subseteq \bigotimes_{i\in I} \mathcal{T}_i$$

und sodann

$$\tau\left(\prod_{i\in I} \mathcal{T}_i\right) \subseteq \bigotimes_{i\in I} \mathcal{T}_i$$

Daraus folgt die Behauptung. □

Aufgaben

3.2.A Sei $\{(\Omega_i, \mathcal{T}_i)\}_{i\in I}$ eine Familie topologischer Räume. Sind $K, L \subseteq I$ nichtleer mit $K + L = I$, so gilt

$$\bigotimes_{i\in I} \mathcal{T}_i = \left(\bigotimes_{i\in K} \mathcal{T}_i\right) \otimes \left(\bigotimes_{i\in L} \mathcal{T}_i\right)$$

Hinweis: Verwenden Sie Satz 3.2.2.

3.2.B **Graph:** Seien $(\Omega, \mathcal{T})$ und $(\Omega', \mathcal{T}')$ topologische Räume und sei $f : \Omega \to \Omega'$ stetig. Dann ist der Graph von f $\mathcal{T}$–$(\mathcal{T} \otimes \mathcal{T}')$–stetig.

3.3 Produkte von messbaren Räumen

Sei $\{(\Omega_i, \mathcal{F}_i)\}_{i\in I}$ eine Familie messbarer Räume. Dann gibt es eine kleinste σ–Algebra

$$\bigotimes_{i\in I} \mathcal{F}_i$$

auf $\prod_{i\in I} \Omega_i$, für die alle Projektionen messbar sind. Diese σ–Algebra wird als *Produkt–σ–Algebra* oder als *Kolmogorovsche σ–Algebra* der Familie $\{\mathcal{F}_i\}_{i\in I}$ bezeichnet und der messbare Raum

$$\bigotimes_{i\in I} (\Omega_i, \mathcal{F}_i) := \left(\prod_{i\in I} \Omega_i, \bigotimes_{i\in I} \mathcal{F}_i \right)$$

wird als *messbares Produkt* der Familie $\{(\Omega_i, \mathcal{F}_i)\}_{i\in I}$ oder als *Produkt der messbaren Räume* $(\Omega_i, \mathcal{F}_i)$ bezeichnet.

Das folgende Beispiel zeigt, dass die Borelsche σ–Algebra auf $\mathbb{R}^n$ als Produkt–σ–Algebra dargestellt werden kann:

3.3.1 Beispiel (Borelsche σ–Algebra auf $\mathbb{R}^n$). Für alle $n \in \mathbb{N}$ gilt

$$\mathcal{B}(\mathbb{R}^n) = \bigotimes_{i\in\{1,\dots,n\}} \mathcal{B}(\mathbb{R})$$

In der Tat: Für alle $\mathbf{a}, \mathbf{c} \in \mathbb{R}^n$ gilt

$$(\mathbf{a}, \mathbf{c}] = \prod_{i\in\{1,\dots,n\}} (a_i, c_i] = \bigcap_{i\in\{1,\dots,n\}} \pi_i^{-1}((a_i, c_i])$$

Da alle Projektionen bezüglich der Produkt–σ–Algebra messbar sind, ergibt sich daraus $(\mathbf{a}, \mathbf{c}] \in \bigotimes_{i\in\{1,\dots,n\}} \mathcal{B}(\mathbb{R})$. Daher gilt $\mathcal{J}(\mathbb{R}^n) \subseteq \bigotimes_{i\in\{1,\dots,n\}} \mathcal{B}(\mathbb{R})$, und damit

$$\mathcal{B}(\mathbb{R}^n) \subseteq \bigotimes_{i\in\{1,\dots,n\}} \mathcal{B}(\mathbb{R})$$

Sei nun $i \in \{1, \dots, n\}$ und $(a, b] \in \mathcal{J}(\mathbb{R})$. Für alle $j \in \{1, \dots, n\}$ und $k \in \mathbb{N}$ sei

$$H_{j,k} := \begin{cases} (a, b] & \text{falls } j = i \\ (-k, k] & \text{sonst} \end{cases}$$

Dann gilt

$$\pi_i^{-1}((a, b]) = \bigcup_{k\in\mathbb{N}} \prod_{j\in\{1,\dots,n\}} H_{j,k} \in \sigma(\mathcal{J}(\mathbb{R}^n)) = \mathcal{B}(\mathbb{R}^n)$$

Damit ist gezeigt, dass alle Projektionen bezüglich der σ–Algebra $\mathcal{B}(\mathbb{R}^n)$ messbar sind, und daraus folgt

$$\bigotimes_{i\in\{1,\dots,n\}} \mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}(\mathbb{R}^n)$$

Damit ist die Behauptung gezeigt.

Die Darstellung einer σ–Algebra als Produkt–σ–Algebra kann den Nachweis der Messbarkeit einer Abbildung erleichtern:

3.3.2 Satz (Messbare Abbildungen in ein messbares Produkt). *Sei* $(\Omega_0, \mathcal{F}_0)$ *ein messbarer Raum und sei* $f : \Omega_0 \to \prod_{i\in I} \Omega_i$ *eine Abbildung. Dann sind folgende Aussagen äquivalent*:
(a) f *ist messbar bezüglich der Produkt–σ–Algebra auf* $\prod_{i\in I} \Omega_i$.
(b) *Für alle* $i \in I$ *ist* $\pi_i \circ f$ *messbar.*

3.3.3 Satz. *Für alle* $i \in I$ *sei* $\mathcal{E}_i$ *ein Mengensystem mit* $\mathcal{F}_i = \sigma(\mathcal{E}_i)$. *Dann gilt*

$$\bigotimes_{i\in I} \mathcal{F}_i = \sigma\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i)\right)$$

Wir betrachten abschließend das Produkt einer abzählbaren Familie von σ–Algebren:

3.3.4 Satz. *Sei* I *abzählbar. Dann gilt*

$$\bigotimes_{i\in I} \mathcal{F}_i = \sigma\left(\prod_{i\in I} \mathcal{F}_i\right)$$

Die unterschiedlichen Voraussetzungen an die Mächtigkeit der Indexmenge in den Sätzen 3.2.4 und 3.3.4 haben ihre Ursache darin, dass Topologien unter der Bildung von endlichen Durchschnitten stabil sind, während σ–Algebren unter der Bildung von abzählbaren Durchschnitten stabil sind.

Aufgaben

3.3.A Sei $\{(\Omega_i, \mathcal{F}_i)\}_{i\in I}$ eine Familie messbarer Räume. Sind $K, L \subseteq I$ nichtleer mit $K + L = I$, so gilt

$$\bigotimes_{i\in I} \mathcal{F}_i = \left(\bigotimes_{i\in K} \mathcal{F}_i\right) \otimes \left(\bigotimes_{i\in L} \mathcal{F}_i\right)$$

3.3.B **Graph:** Seien $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ messbare Räume und sei $f : \Omega \to \Omega'$ messbar. Dann ist der Graph von f $\mathcal{F}$–$(\mathcal{F}\otimes\mathcal{F}')$–messbar.

Teil II

Maßtheorie

4

Mengenfunktionen

In diesem Kapitel betrachten wir Abbildungen von einem nichtleeren Mengensystem nach $[0, \infty]$ oder $[-\infty, \infty]$. Eine derartige Abbildung wird als *Mengenfunktion* bezeichnet.

Die wichtigsten Mengenfunktionen sind Inhalte (Abschnitt 4.1) und Maße (Abschnitt 4.2). Die Bildung der Differenz zwischen zwei Maßen, von denen mindestens eines endlich ist, führt auf den Begriff des signierten Maßes (Abschnitt 4.3).

Im gesamten Kapitel sei Ω eine nichtleere Menge und $\mathcal{C} \subseteq 2^\Omega$ ein Mengensystem mit $\emptyset \in \mathcal{C}$.

Für eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ und ein nichtleeres Mengensystem $\mathcal{D} \subseteq \mathcal{C}$ bezeichnen wir die Abbildung $\mu|_\mathcal{D} : \mathcal{D} \to [0, \infty]$ mit

$$\mu|_\mathcal{D}[D] := \mu[D]$$

als *Restriktion* von μ auf $\mathcal{D}$.

Für Mengenfunktionen $\mu, \nu : \mathcal{C} \to [0, \infty]$ und $\alpha \in \mathbb{R}_+$ definieren wir Mengenfunktionen $\mu + \nu : \mathcal{C} \to [0, \infty]$ und $\alpha\mu : \mathcal{C} \to [0, \infty]$ durch

$$\begin{aligned}(\mu+\nu)[C] &:= \mu[C] + \nu[C] \\ (\alpha\mu)[C] &:= \alpha\mu[C]\end{aligned}$$

und wir schreiben

$$\mu \leq \nu$$

wenn für alle $C \in \mathcal{C}$

$$\mu[C] \leq \nu[C]$$

gilt.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_4, © Springer-Verlag Berlin Heidelberg 2011

4.1 Inhalte

Eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ heißt
– *additiv*, wenn für je zwei Mengen $A, B \in \mathcal{C}$ mit $A \cap B = \emptyset$ und $A + B \in \mathcal{C}$

$$\mu[A + B] = \mu[A] + \mu[B]$$

gilt.
– *endlich additiv*, wenn für jede endliche disjunkte Familie $\{A_i\}_{i \in I} \subseteq \mathcal{C}$ mit $\sum_{i \in I} A_i \in \mathcal{C}$

$$\mu\left[\sum_{i \in I} A_i\right] = \sum_{i \in I} \mu[A_i]$$

gilt.
– *Inhalt*, wenn $\mu[\emptyset] = 0$ gilt und μ endlich additiv ist.

Jede endlich additive Mengenfunktion ist additiv.

4.1.1 Beispiele. Sei $\Omega := \mathbb{N}$ und $\mathcal{C} := \{ A \in 2^\Omega \mid A \text{ oder } \Omega \setminus A \text{ ist endlich}\}$.
(1) Die Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit

$$\mu[A] := \infty$$

ist endlich additiv, aber kein Inhalt.
(2) Die Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit

$$\mu[A] := \begin{cases} 0 & \text{falls } A \text{ endlich ist} \\ \infty & \text{sonst} \end{cases}$$

ist ein Inhalt.
(3) Die Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit

$$\mu[A] := |A|$$

ist ein Inhalt.

Das folgende Lemma erleichtert den Nachweis der Additivität einer Mengenfunktion auf einem Ring; vgl. Aufgabe 4.1.A:

4.1.2 Lemma. *Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0, \infty]$ eine Mengenfunktion. Dann sind äquivalent:*
(a) *μ ist additiv.*
(b) *μ ist endlich additiv.*

Beweis. Da $\mathcal{C}$ ein Ring ist, enthält $\mathcal{C}$ mit jeder endlichen Familie von Mengen auch deren Vereinigung.
Wir nehmen an, dass μ additiv ist, und zeigen durch vollständige Induktion, dass für alle $n \in \mathbb{N}$ und jede disjunkte Familie $\{A_i\}_{i \in \{1,\dots,n\}} \subseteq \mathcal{C}$

$$\mu\left[\sum_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \mu[A_i]$$

gilt:

- $n = 1$: In diesem Fall ist nichts zu zeigen.
- $n \to n+1$: Wir nehmen an, die Behauptung sei für n bereits bewiesen, und betrachten eine disjunkte Familie $\{A_i\}_{i\in\{1,\dots,n+1\}} \subseteq \mathcal{C}$. Aus der Additivität von μ folgt dann

$$\begin{aligned}\mu\left[\sum_{i=1}^{n+1} A_i\right] &= \mu\left[\sum_{i=1}^{n} A_i + A_{n+1}\right] \\ &= \mu\left[\sum_{i=1}^{n} A_i\right] + \mu[A_{n+1}] \\ &= \sum_{i=1}^{n} \mu[A_i] + \mu[A_{n+1}] \\ &= \sum_{i=1}^{n+1} \mu[A_i]\end{aligned}$$

Daher ist μ endlich additiv. □

Das folgende Beispiel zeigt, dass die Äquivalenz aus Lemma 4.1.2 für Mengenfunktionen auf einem Halbring im allgemeinen nicht gilt:

4.1.3 Beispiel. Sei $\Omega := \{1,2,3\}$ und $\mathcal{C} := \{\emptyset, \{1\}, \{2\}, \{3\}, \Omega\}$. Dann ist $\mathcal{C}$ ein Halbring, aber kein Ring, und die Mengenfunktion $\mu : \mathcal{C} \to [0,\infty]$ mit

$$\mu[A] := \begin{cases} 0 & \text{falls } A = \emptyset \\ 1 & \text{falls } A = \{\omega\} \text{ für ein } \omega \in \Omega \\ 4 & \text{falls } A = \Omega \end{cases}$$

ist additiv, aber nicht endlich additiv.

Die Untersuchung von Inhalten auf einem Halbring wird dadurch erschwert, dass Halbringe im allgemeinen nicht stabil unter der Bildung von Vereinigungen und relativen Komplementen sind; vgl. Beispiel 4.1.3. Hier hilft der folgende Fortsetzungssatz:

4.1.4 Satz. *Sei $\mathcal{C}$ ein Halbring. Dann besitzt jeder Inhalt $\mathcal{C} \to [0,\infty]$ eine eindeutige Fortsetzung zu einem Inhalt $\varrho(\mathcal{C}) \to [0,\infty]$.*

Beweis. Sei $\mu : \mathcal{C} \to [0,\infty]$ ein Inhalt. Nach Satz 1.5.7 gilt

$$\varrho(\mathcal{C}) = \left\{ A \in 2^\Omega \;\middle|\; A = \sum_{i\in I} H_i \text{ mit } I \text{ endlich und } \{H_i\}_{i\in I} \subseteq \mathcal{C} \text{ disjunkt} \right\}$$

Ist $\tilde{\mu} : \varrho(\mathcal{C}) \to [0,\infty]$ ein Inhalt mit $\tilde{\mu}|_{\mathcal{C}} = \mu$, so gilt für alle $A \in \varrho(\mathcal{C})$ und jede endliche disjunkte Familie $\{H_i\}_{i\in I} \subseteq \mathcal{C}$ mit $A = \sum_{i\in I} H_i$

$$\tilde{\mu}[A] = \sum_{i\in I} \tilde{\mu}[H_i] = \sum_{i\in I} \mu[H_i]$$

Andererseits gilt für endliche disjunkte Familien $\{G_i\}_{i\in I} \subseteq \mathcal{C}$ und $\{H_j\}_{j\in J} \subseteq \mathcal{C}$

mit $\sum_{i\in I} G_i = \sum_{j\in J} H_j$

$$\sum_{i\in I} \mu[G_i] = \sum_{i\in I} \mu\left[G_i \cap \sum_{j\in J} H_j\right] = \sum_{i\in I}\sum_{j\in J} \mu[G_i \cap H_j]$$

und

$$\sum_{j\in J} \mu[H_j] = \sum_{j\in J} \mu\left[H_j \cap \sum_{i\in I} G_i\right] = \sum_{j\in J}\sum_{i\in I} \mu[H_j \cap G_i]$$

und damit

$$\sum_{i\in I} \mu[G_i] = \sum_{j\in J} \mu[H_j]$$

Daher ist die Abbildung $\bar{\mu} : \varrho(\mathcal{C}) \to [0,\infty]$ mit

$$\bar{\mu}[A] := \sum_{i\in I} \mu[H_i]$$

und einer endlichen disjunkten Familie $\{H_i\}_{i\in I} \subseteq \mathcal{C}$ mit $A = \sum_{i\in I} H_i$ wohldefiniert und für alle $H \in \mathcal{C}$ gilt $\bar{\mu}[H] = \mu[H]$. Daher ist $\bar{\mu}$ eine Fortsetzung von μ und es ist klar, dass $\bar{\mu}$ ein Inhalt ist. □

Eine Mengenfunktion $\mu : \mathcal{C} \to [0,\infty]$ heißt

– *monoton*, wenn für alle $A, B \in \mathcal{C}$ mit $B \subseteq A$
$$\mu[B] \leq \mu[A]$$
gilt.

– *subadditiv*, wenn für alle $A, B \in \mathcal{C}$ mit $A \cup B \in \mathcal{C}$
$$\mu[A \cup B] \leq \mu[A] + \mu[B]$$
gilt.

– *endlich subadditiv*, wenn für jede endliche Familie $\{A_i\}_{i\in I} \subseteq \mathcal{C}$ mit $\bigcup_{i\in I} A_i \in \mathcal{C}$
$$\mu\left[\bigcup_{i\in I} A_i\right] \leq \sum_{i\in I} \mu[A_i]$$
gilt.

Jede endlich subadditive Mengenfunktion ist subadditiv, und jede subadditive Mengenfunktion auf einem Ring ist endlich subadditiv.

4.1.5 Lemma. *Sei $\mathcal{C}$ ein Halbring und sei $\mu : \mathcal{C} \to [0,\infty]$ ein Inhalt. Dann ist μ monoton und endlich subadditiv.*

Beweis. Aufgrund von Satz 4.1.4 können wir annehmen, dass $\mathcal{C}$ ein Ring ist. Für alle $A, B \in \mathcal{C}$ mit $B \subseteq A$ gilt

$$\mu[B] \leq \mu[B] + \mu[A \setminus B] = \mu[B + (A \setminus B)] = \mu[A]$$

Daher ist μ monoton.
Sei nun $\{A_k\}_{k \in \{1,\ldots,n\}} \subseteq \mathcal{C}$ eine endliche Familie und für alle $k \in \{1, \ldots, n\}$ sei

$$B_k := A_k \setminus \bigcup_{j=1}^{k-1} A_j$$

Dann ist $\{B_k\}_{k \in \{1,\ldots,n\}}$ eine endliche disjunkte Familie in $\mathcal{C}$ mit $\sum_{k=1}^n B_k = \bigcup_{k=1}^n A_k$ und aus der endlichen Additivität und der Monotonie von μ ergibt sich nun

$$\mu\left[\bigcup_{k=1}^n A_k\right] = \mu\left[\sum_{k=1}^n B_k\right] = \sum_{k=1}^n \mu[B_k] \leq \sum_{k=1}^n \mu[A_k]$$

Daher ist μ endlich subadditiv. □

Eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ heißt *endlich*, wenn für alle $A \in \mathcal{C}$

$$\mu[A] < \infty$$

gilt. Eine endliche Mengenfunktion ist genau dann ein Inhalt, wenn sie endlich additiv ist.

Aufgaben

4.1.A Sei $\mathcal{C}$ $\cup$–stabil und sei $\mu : \mathcal{C} \to [0, \infty]$ eine Mengenfunktion. Dann sind äquivalent:
(a) μ ist additiv.
(b) μ ist endlich additiv.

4.1.B Sei $\mathcal{C}$ $\cup$–stabil und sei $\mu : \mathcal{C} \to [0, \infty]$ eine Mengenfunktion. Dann sind äquivalent:
(a) μ ist subadditiv.
(b) μ ist endlich subadditiv.

4.1.C Eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ heißt *subtraktiv*, wenn für alle $A, B \in \mathcal{C}$ mit $B \subseteq A$ und $A \setminus B \in \mathcal{C}$ sowie $\mu[B] < \infty$

$$\mu[A \setminus B] = \mu[A] - \mu[B]$$

gilt. Jede additive Mengenfunktion ist subtraktiv.

4.1.D Eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ heißt *modular*, wenn für alle $A, B \in \mathcal{C}$ mit $A \cup B \in \mathcal{C}$ und $A \cap B \in \mathcal{C}$

$$\mu[A \cup B] + \mu[A \cap B] = \mu[A] + \mu[B]$$

gilt. Jede modulare Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit $\mu[\emptyset] = 0$ ist additiv, und jede endlich additive Mengenfunktion auf einem Halbring ist modular.

4.1.E **Einschluss–Ausschluss–Formel (Poincaré):** Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0,\infty]$ additiv und endlich. Dann gilt für jede endliche Familie $\{A_k\}_{k\in\{1,\dots,n\}} \subseteq \mathcal{C}$

$$\mu\left[\bigcup_{k\in\{1,\dots,n\}} A_k\right] = \sum_{i=1}^{n} (-1)^{i-1} \sum_{I\subseteq\{1,\dots,n\},\,|I|=i} \mu\left[\bigcap_{k\in I} A_k\right]$$

4.1.F **Nullmengen:** Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0,\infty]$ ein Inhalt. Dann ist das Mengensystem

$$\mathcal{N}_\mu := \Big\{ N \in \mathcal{C} \;\Big|\; \mu[N] = 0 \Big\}$$

ein Ideal in $\mathcal{C}$. Jede Menge $N \in \mathcal{N}_\mu$ heißt *μ–Nullmenge*.

4.1.G Eine Mengenfunktion $\mu : \mathcal{C} \to [0,\infty]$ heißt *vollständig*, wenn für alle $A \in \mathcal{C}$ mit $\mu[A] = 0$ und alle $B \subseteq A$

$$B \in \mathcal{C}$$

gilt. Ist $\mathcal{C}$ ein Ring und $\mu : \mathcal{C} \to [0,\infty]$ ein vollständiger Inhalt, so ist jede Teilmenge einer μ–Nullmenge selbst eine μ–Nullmenge.

4.1.H **Vervollständigung:** Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0,\infty]$ ein Inhalt. Sei ferner

$$\mathcal{M}_\mu := \Big\{ A \in 2^\Omega \;\Big|\; A \subseteq N \text{ für ein } N \in \mathcal{N}_\mu \Big\}$$

und

$$\mathcal{C}_\mu := \Big\{ A \in 2^\Omega \;\Big|\; A = C + D \text{ mit } C \in \mathcal{C} \text{ und } D \in \mathcal{M}_\mu \text{ disjunkt} \Big\}$$

Dann gilt

$$\begin{aligned} \mathcal{C}_\mu &= \Big\{ A \in 2^\Omega \;\Big|\; A = C \cup D \text{ mit } C \in \mathcal{C} \text{ und } D \in \mathcal{M}_\mu \Big\} \\ &= \Big\{ A \in 2^\Omega \;\Big|\; A = C \triangle D \text{ mit } C \in \mathcal{C} \text{ und } D \in \mathcal{M}_\mu \Big\} \\ &= \varrho(\mathcal{C} \cup \mathcal{M}_\mu) \end{aligned}$$

Sei ferner $\widetilde{\mu} : \mathcal{C}_\mu \to [0,\infty]$ gegeben durch

$$\widetilde{\mu}[A] := \mu[C]$$

mit $C \in \mathcal{C}$ sodass $A = C + D$ für ein $D \in \mathcal{M}_\mu$ gilt. Dann gilt:
(1) $\widetilde{\mu}$ ist wohldefiniert und ein vollständiger Inhalt mit $\widetilde{\mu}|_\mathcal{C} = \mu$.
(2) $\widetilde{\mu}$ ist der einzige Inhalt auf $\mathcal{C}_\mu$, der μ fortsetzt.
(3) Ist $\mathcal{C}_0$ ein Ring auf Ω mit $\mathcal{C} \subseteq \mathcal{C}_0$ und ist $\mu_0 : \mathcal{C}_0 \to [0,\infty]$ ein vollständiger Inhalt mit $\mu_0|_\mathcal{C} = \mu$, so gilt $\mathcal{C}_\mu \subseteq \mathcal{C}_0$ und $\mu_0|_{\mathcal{C}_\mu} = \widetilde{\mu}$.

Der Inhalt $\widetilde{\mu}$ heißt die *Vervollständigung* von μ.

4.1.I Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0,\infty]$ ein Inhalt. Zwei Mengen $A, B \in \mathcal{C}$ heißen *μ–äquivalent*, wenn $\mu[A \triangle B] = 0$ gilt, und in diesem Fall schreiben wir

$$A =_\mu B$$

Dann ist $=_\mu$ eine Äquivalenzrelation auf $\mathcal{C}$ und $\mathcal{C}/_{=_\mu}$ ist ein Ring.

4.1.J Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0, \infty]$ ein endlicher Inhalt. Dann ist die Abbildung $d_\mu : \mathcal{C}/_{=_\mu} \times \mathcal{C}/_{=_\mu} \to \mathbb{R}_+$ mit

$$d_\mu(A, B) := \mu[A \triangle B]$$

(wohldefiniert und) eine Metrik.

4.1.K Sei $\mathcal{C}$ ein Halbring und sei $\mu : \mathcal{C} \to [0, \infty]$ ein endlicher Inhalt. Dann ist der eindeutig bestimmte Inhalt $\bar{\mu} : \varrho(\mathcal{C}) \to [0, \infty]$ mit $\bar{\mu}|_\mathcal{C} = \mu$ endlich.

4.2 Maße

Eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ heißt

– *σ–additiv*, wenn für jede disjunkte Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ mit $\sum_{k=1}^\infty A_k \in \mathcal{C}$

$$\mu\left[\sum_{k=1}^\infty A_k\right] = \sum_{k=1}^\infty \mu[A_k]$$

gilt.

– *Maß*, wenn $\mu[\emptyset] = 0$ gilt und μ σ–additiv ist.

Die folgende Charakterisierung von Maßen ist offensichtlich:

4.2.1 Lemma. *Sei $\mu : \mathcal{C} \to [0, \infty]$ eine Mengenfunktion. Dann sind äquivalent*:

(a) *μ ist ein Maß.*

(b) *μ ist ein σ–additiver Inhalt.*

4.2.2 Beispiele.

(1) Sei $\Omega := \mathbb{N}$ und $\mathcal{C} := \{\, A \in 2^\Omega \mid A \text{ oder } \Omega \setminus A \text{ ist endlich}\}$. Dann ist die Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit

$$\mu[A] := \begin{cases} 0 & \text{falls } A \text{ endlich ist} \\ \infty & \text{sonst} \end{cases}$$

ein Inhalt, aber kein Maß.

(2) Sei $\Omega := \mathbb{R}$ und $\mathcal{C} := \{\, A \in 2^\Omega \mid A \text{ oder } \Omega \setminus A \text{ ist abzählbar}\}$. Dann ist die Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit

$$\mu[A] := \begin{cases} 0 & \text{falls } A \text{ abzählbar ist} \\ \infty & \text{sonst} \end{cases}$$

ein Maß.

(3) **Dirac–Maß:** Für jedes $\omega \in \Omega$ ist die Mengenfunktion $\boldsymbol{\delta}_\omega : \mathcal{C} \to [0, \infty]$ mit

$$\boldsymbol{\delta}_\omega[A] := \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{sonst} \end{cases}$$

ein endliches Maß. Das Maß $\boldsymbol{\delta}_\omega$ heißt *Dirac–Maß* bezüglich $\omega \in \Omega$.

(4) **Zählmaß:** Die Mengenfunktion $\boldsymbol{\zeta} : 2^\Omega \to [0,\infty]$ mit

$$\boldsymbol{\zeta}[A] := |A|$$

ist ein Maß. Das Maß $\boldsymbol{\zeta}$ heißt *Zählmaß*.

(5) **Lokales Zählmaß:** Sei $C \in 2^\Omega$ abzählbar. Dann ist die Mengenfunktion $\boldsymbol{\zeta}_C : 2^\Omega \to [0,\infty]$ mit

$$\boldsymbol{\zeta}_C[A] := |A \cap C|$$

ein Maß mit

$$\boldsymbol{\zeta}_C[A] = \sum_{\omega \in C} \boldsymbol{\delta}_\omega[A]$$

Das Maß $\boldsymbol{\zeta}_C$ heißt *lokales Zählmaß bezüglich* C.

Der folgende Fortsetzungssatz ist ein Analogon zu Satz 4.1.4:

4.2.3 Satz. *Sei $\mathcal{C}$ ein Halbring. Dann besitzt jedes Maß $\mathcal{C} \to [0,\infty]$ eine eindeutige Fortsetzung zu einem Maß $\varrho(\mathcal{C}) \to [0,\infty]$.*

Beweis. Sei $\mu : \mathcal{C} \to [0,\infty]$ ein Maß. Dann ist μ ein Inhalt und nach Satz 4.1.4 gibt es einen eindeutig bestimmten Inhalt $\bar{\mu} : \varrho(\mathcal{C}) \to [0,\infty]$ mit $\bar{\mu}|_\mathcal{C} = \mu$. Wir zeigen nun, dass $\bar{\mu}$ σ–additiv ist:
Sei $A \in \varrho(\mathcal{C})$ und sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \varrho(\mathcal{C})$ eine disjunkte Folge mit $A = \sum_{k=1}^\infty A_k$. Dann gibt es eine endliche disjunkte Familie $\{B_i\}_{i\in I} \subseteq \mathcal{C}$ mit $A = \sum_{i\in I} B_i$ und für alle $k \in \mathbb{N}$ gibt es eine endliche disjunkte Familie $\{B_{k,j}\}_{j\in J(k)} \subseteq \mathcal{C}$ mit $A_k = \sum_{j\in J(k)} B_{k,j}$. Es gilt

$$B_i = B_i \cap A = B_i \cap \sum_{k=1}^\infty A_k = B_i \cap \sum_{k=1}^\infty \sum_{j\in J(k)} B_{k,j} = \sum_{k=1}^\infty \sum_{j\in J(k)} B_i \cap B_{k,j}$$

und

$$B_{k,j} = B_{k,j} \cap A = B_{k,j} \cap \sum_{i\in I} B_i = \sum_{i\in I} B_i \cap B_{k,j}$$

Aus der Definition von $\bar{\mu}$ und der σ–Additivität von μ ergibt sich nun

$$\bar{\mu}[A] = \sum_{i\in I} \mu[B_i] = \sum_{i\in I} \sum_{k=1}^\infty \sum_{j\in J(k)} \mu[B_i \cap B_{k,j}]$$

und

$$\sum_{k=1}^\infty \bar{\mu}[A_k] = \sum_{k=1}^\infty \sum_{j\in J(k)} \mu[B_{k,j}] = \sum_{k=1}^\infty \sum_{j\in J(k)} \sum_{i\in I} \mu[B_i \cap B_{k,j}]$$

und damit $\bar{\mu}[A] = \sum_{k=1}^\infty \bar{\mu}[A_k]$. Daher ist $\bar{\mu}$ σ–additiv. □

Eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ heißt

– *stetig von unten*, wenn für jede monoton wachsende Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ mit $\bigcup_{k\in\mathbb{N}} A_k \in \mathcal{C}$
$$\mu\left[\bigcup_{k\in\mathbb{N}} A_k\right] = \sup_{k\in\mathbb{N}} \mu[A_k]$$
gilt.

– *stetig von oben*, wenn für jede monoton fallende Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ mit $\bigcap_{k\in\mathbb{N}} A_k \in \mathcal{C}$ und $\mu[A_1] < \infty$
$$\mu\left[\bigcap_{k\in\mathbb{N}} A_k\right] = \inf_{k\in\mathbb{N}} \mu[A_k]$$
gilt.

– $\emptyset$*–stetig*, wenn für jede monoton fallende Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ mit $\bigcap_{k\in\mathbb{N}} A_k = \emptyset$ und $\mu[A_1] < \infty$
$$\inf_{k\in\mathbb{N}} \mu[A_k] = 0$$
gilt.

4.2.4 Lemma. *Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0, \infty]$ ein Inhalt. Dann sind äquivalent*:

(a) *μ ist σ–additiv.*

(b) *μ ist stetig von unten.*

Beweis. Sei zunächst μ σ–additiv und sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine monoton wachsende Folge mit $\bigcup_{k\in\mathbb{N}} A_k \in \mathcal{C}$. Sei $A_0 := \emptyset$ und für alle $k \in \mathbb{N}$ sei
$$B_k := A_k \setminus A_{k-1}$$
Dann ist $\{B_k\}_{k\in\mathbb{N}}$ eine disjunkte Folge in $\mathcal{C}$ mit $\sum_{k=1}^\infty B_k = \bigcup_{k\in\mathbb{N}} A_k$ und für alle $n \in \mathbb{N}$ gilt $\sum_{k=1}^n B_k = A_n$. Daraus folgt
$$\begin{aligned}\mu\left[\bigcup_{k\in\mathbb{N}} A_k\right] &= \mu\left[\sum_{k=1}^\infty B_k\right] \\ &= \sum_{k=1}^\infty \mu[B_k] \\ &= \sup_{n\in\mathbb{N}} \sum_{k=1}^n \mu[B_k] \\ &= \sup_{n\in\mathbb{N}} \mu\left[\sum_{k=1}^n B_k\right] \\ &= \sup_{n\in\mathbb{N}} \mu[A_n]\end{aligned}$$
Daher ist μ stetig von unten.

Sei nun μ stetig von unten und sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine disjunkte Folge mit $\sum_{k=1}^\infty A_k \in \mathcal{C}$. Für alle $n \in \mathbb{N}$ sei

$$B_n := \sum_{k=1}^n A_k$$

Dann ist $\{B_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge in $\mathcal{C}$ mit $\bigcup_{n\in\mathbb{N}} B_n = \sum_{k=1}^\infty A_k$. Daraus folgt

$$\begin{aligned}\mu\left[\sum_{k=1}^\infty A_k\right] &= \mu\left[\bigcup_{n\in\mathbb{N}} B_n\right] \\ &= \sup_{n\in\mathbb{N}} \mu[B_n] \\ &= \sup_{n\in\mathbb{N}} \mu\left[\sum_{k=1}^n A_k\right] \\ &= \sup_{n\in\mathbb{N}} \sum_{k=1}^n \mu[A_k] \\ &= \sum_{k=1}^\infty \mu[A_k]\end{aligned}$$

Daher ist μ σ–additiv. □

4.2.5 Lemma. *Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0,\infty]$ ein Inhalt. Ist μ stetig von unten, so ist μ stetig von oben.*

Beweis. Sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine monoton fallende Folge mit $\bigcap_{k\in\mathbb{N}} A_k \in \mathcal{C}$ und $\mu[A_1] < \infty$. Dann ist $\{A_1 \setminus A_k\}_{k\in\mathbb{N}}$ eine monoton wachsende Folge in $\mathcal{C}$ mit $\bigcup_{k\in\mathbb{N}}(A_1 \setminus A_k) = A_1 \setminus \bigcap_{k\in\mathbb{N}} A_k \in \mathcal{C}$. Daraus folgt

$$\begin{aligned}\mu[A_1] - \inf_{k\in\mathbb{N}} \mu[A_k] &= \mu[A_1] + \sup_{k\in\mathbb{N}}(-\mu[A_k]) \\ &= \sup_{k\in\mathbb{N}}(\mu[A_1] - \mu[A_k]) \\ &= \sup_{k\in\mathbb{N}} \mu[A_1 \setminus A_k] \\ &= \mu\left[\bigcup_{k\in\mathbb{N}}(A_1 \setminus A_k)\right] \\ &= \mu\left[A_1 \setminus \bigcap_{k\in\mathbb{N}} A_k\right] \\ &= \mu[A_1] - \mu\left[\bigcap_{k\in\mathbb{N}} A_k\right]\end{aligned}$$

und aus $\mu[A_1] < \infty$ folgt nun

$$\mu\left[\bigcap_{k\in\mathbb{N}} A_k\right] = \inf_{k\in\mathbb{N}} \mu[A_k]$$

Daher ist μ stetig von oben. □

Wir notieren eine offensichtliche Folgerung aus Lemma 4.2.4 und Lemma 4.2.5:

4.2.6 Folgerung. *Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0,\infty]$ ein Maß. Dann ist μ stetig von oben.*

Für *endliche* Inhalte auf einem Ring gilt auch die Umkehrung der Implikation von Lemma 4.2.5:

4.2.7 Lemma. *Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0,\infty]$ ein endlicher Inhalt. Ist μ stetig von oben, so ist μ stetig von unten.*

Beweis. Sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine monoton wachsende Folge mit $\bigcup_{k\in\mathbb{N}} A_k \in \mathcal{C}$ und sei $A := \bigcup_{k\in\mathbb{N}} A_k$. Dann ist $\{A \setminus A_k\}_{k\in\mathbb{N}}$ eine monoton fallende Folge in $\mathcal{C}$ mit $\bigcap_{k\in\mathbb{N}}(A \setminus A_k) = \emptyset$. Nach Voraussetzung gilt $\sup_{k\in\mathbb{N}} \mu[A_k] \leq \mu[A] < \infty$ und damit

$$\begin{aligned}
\mu[A] - \sup_{k\in\mathbb{N}} \mu[A_k] &= \mu[A] + \inf_{k\in\mathbb{N}}(-\mu[A_k]) \\
&= \inf_{k\in\mathbb{N}}(\mu[A] - \mu[A_k]) \\
&= \inf_{k\in\mathbb{N}} \mu[A \setminus A_k] \\
&= \mu[\emptyset] \\
&= 0
\end{aligned}$$

Daher ist μ stetig von unten. □

Das folgende Beispiel zeigt, dass im letzten Lemma die Forderung der Endlichkeit des Inhaltes wesentlich ist:

4.2.8 Beispiel. Sei $\Omega := \mathbb{N}$ und $\mathcal{C} := \{\, A \in 2^\Omega \mid A \text{ oder } \Omega \setminus A \text{ ist endlich}\}$. Dann ist die Mengenfunktion $\mu : \mathcal{C} \to [0,\infty]$ mit

$$\mu[A] := \begin{cases} 0 & \text{falls } A \text{ endlich ist} \\ \infty & \text{sonst} \end{cases}$$

ein Inhalt und stetig von oben, aber sie ist nicht stetig von unten.

Wir klären nun noch den Zusammenhang zwischen der Stetigkeit von oben und der $\emptyset$–Stetigkeit:

4.2.9 Lemma. *Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0, \infty]$ ein Inhalt. Dann sind äquivalent*:
(a) *μ ist stetig von oben.*
(b) *μ ist $\emptyset$–stetig.*

Beweis. Es ist klar, dass (b) aus (a) folgt.
Wir nehmen nun an, dass (b) gilt. Sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine monoton fallende Folge mit $\bigcap_{k\in\mathbb{N}} A_k \in \mathcal{C}$ und $\mu[A_1] < \infty$, und sei $A := \bigcap_{k\in\mathbb{N}} A_k$. Dann ist $\{A_k \setminus A\}_{k\in\mathbb{N}}$ eine monoton fallende Folge in $\mathcal{C}$ mit $\bigcap_{k\in\mathbb{N}}(A_k \setminus A) = \emptyset$ und $\mu[A_1 \setminus A] < \infty$. Daraus folgt

$$\begin{aligned}
\mu\left[\bigcap_{k\in\mathbb{N}} A_k\right] &= \mu[A] \\
&= \mu[A] + \mu\left[\bigcap_{k\in\mathbb{N}} (A_k \setminus A)\right] \\
&= \mu[A] + \inf_{k\in\mathbb{N}} \mu[A_k \setminus A] \\
&= \inf_{k\in\mathbb{N}} (\mu[A] + \mu[A_k \setminus A]) \\
&= \inf_{k\in\mathbb{N}} \mu[A_k]
\end{aligned}$$

Daher folgt (a) aus (b). □

Für *endliche* Inhalte auf einem Ring können wir die letzten Ergebnisse wie folgt zusammenfassen:

4.2.10 Folgerung. *Sei $\mathcal{C}$ ein Ring und sei $\mu : \mathcal{C} \to [0, \infty]$ ein endlicher Inhalt. Dann sind äquivalent*:
(a) *μ ist σ–additiv.*
(b) *μ ist stetig von unten.*
(c) *μ ist stetig von oben.*
(d) *μ ist $\emptyset$–stetig.*

Eine Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ heißt *σ–subadditiv*, wenn für jede Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ mit $\bigcup_{k=1}^\infty A_k \in \mathcal{C}$

$$\mu\left[\bigcup_{k=1}^\infty A_k\right] \le \sum_{k=1}^\infty \mu[A_k]$$

gilt. Wegen $\emptyset \in \mathcal{C}$ ist jede σ–subadditive Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit $\mu[\emptyset] = 0$ endlich subadditiv.

4.2.11 Lemma. *Sei $\mathcal{C}$ ein Halbring und sei $\mu : \mathcal{C} \to [0, \infty]$ ein Maß. Dann ist μ monoton und σ–subadditiv.*

Beweis. Aufgrund von Satz 4.2.3 können wir annehmen, dass $\mathcal{C}$ ein Ring ist. Da μ ein Maß ist, ist μ auch ein Inhalt, und aus Lemma 4.1.5 folgt nun, dass μ monoton und endlich subadditiv ist. Nach Lemma 4.2.4 ist μ außerdem stetig von unten.
Sei nun $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine Folge mit $\bigcup_{k=1}^\infty A_k \in \mathcal{C}$. Dann ist $\{\bigcup_{k=1}^n A_k\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge in $\mathcal{C}$ mit $\bigcup_{n=1}^\infty \bigcup_{k=1}^n A_k = \bigcup_{k=1}^\infty A_k \in \mathcal{C}$ und es gilt

$$\mu\left[\bigcup_{k=1}^\infty A_k\right] = \mu\left[\bigcup_{n=1}^\infty \bigcup_{k=1}^n A_k\right] = \sup_{n\in\mathbb{N}} \mu\left[\bigcup_{k=1}^n A_k\right] \le \sup_{n\in\mathbb{N}} \sum_{k=1}^n \mu[A_k] = \sum_{k=1}^\infty \mu[A_k]$$

Daher ist μ σ–subadditiv. □

Für Mengenfunktionen auf einem Halbring erhalten wir die folgende Variante von Lemma 4.2.1:

4.2.12 Lemma. *Sei $\mathcal{C}$ ein Halbring und sei $\mu : \mathcal{C} \to [0,\infty]$ eine Mengenfunktion. Dann sind äquivalent*:
(a) *μ ist ein Maß.*
(b) *μ ist ein σ–subadditiver Inhalt.*

Beweis. Sie zunächst μ ein Maß. Dann ist μ ein Inhalt. Da $\mathcal{C}$ ein Halbring ist, ist μ nach Lemma 4.2.11 auch σ–subadditiv. Daher folgt (b) aus (a).
Sei nun μ ein σ–subadditiver Inhalt und sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine disjunkte Folge mit $\sum_{k=1}^\infty A_k \in \mathcal{C}$. Nach Satz 4.1.4 besitzt μ eine eindeutige Fortsetzung zu einem Inhalt $\overline{\mu} : \varrho(\mathcal{C}) \to [0,\infty]$ und nach Lemma 4.1.5 ist $\overline{\mu}$ monoton. Daher gilt für alle $n \in \mathbb{N}$

$$\sum_{k=1}^n \mu[A_k] = \sum_{k=1}^n \overline{\mu}[A_k] = \overline{\mu}\left[\sum_{k=1}^n A_k\right] \le \overline{\mu}\left[\sum_{k=1}^\infty A_k\right] = \mu\left[\sum_{k=1}^\infty A_k\right]$$

und daraus folgt

$$\sum_{k=1}^\infty \mu[A_k] \le \mu\left[\sum_{k=1}^\infty A_k\right]$$

Da μ σ–subadditiv ist, gilt außerdem

$$\mu\left[\sum_{k=1}^\infty A_k\right] \le \sum_{k=1}^\infty \mu[A_k]$$

Aus den letzten beiden Ungleichungen folgt nun, dass μ σ–additiv ist. Daher folgt (a) aus (b). □

Eine Mengenfunktion $\mu : \mathcal{C} \to [0,\infty]$ heißt σ–*endlich*, wenn es eine Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ gibt mit $\bigcup_{k\in\mathbb{N}} A_k = \Omega$ und $\mu[A_k] < \infty$ für alle $k \in \mathbb{N}$.

Aufgaben

4.2.A Ist jede σ–additive Mengenfunktion additiv?

4.2.B Sei $\mathcal{C}$ ein σ–Ring und sei $\mu : \mathcal{C} \to [0, \infty]$ ein Maß. Dann ist das Mengensystem

$$\mathcal{N}_\mu := \Big\{ N \in \mathcal{C} \;\Big|\; \mu[N] = 0 \Big\}$$

ein σ–Ideal in $\mathcal{C}$.

4.2.C **Vervollständigung:** Sei $\mathcal{C}$ eine σ–Algebra und sei $\mu : \mathcal{C} \to [0, \infty]$ ein Maß. Sei ferner

$$\mathcal{M}_\mu := \Big\{ A \in 2^\Omega \;\Big|\; A \subseteq N \text{ für ein } N \in \mathcal{N}_\mu \Big\}$$

und

$$\mathcal{C}_\mu := \Big\{ A \in 2^\Omega \;\Big|\; A = C + D \text{ mit } C \in \mathcal{C} \text{ und } D \in \mathcal{M}_\mu \text{ disjunkt} \Big\}$$

Dann gilt

$$\mathcal{C}_\mu = \sigma(\mathcal{C} \cup \mathcal{M}_\mu)$$

Sei ferner $\widetilde{\mu} : \mathcal{C}_\mu \to [0, \infty]$ gegeben durch

$$\widetilde{\mu}[A] := \mu[C]$$

mit $C \in \mathcal{C}$ derart, dass $A = C + D$ für ein $D \in \mathcal{M}_\mu$ mit $C \cap D = \emptyset$ gilt. Dann gilt
(1) $\widetilde{\mu}$ ist wohldefiniert und ein vollständiges Maß mit $\widetilde{\mu}|_\mathcal{C} = \mu$.
(2) $\widetilde{\mu}$ ist das einzige Maß auf $\mathcal{C}_\mu$, das μ fortsetzt.
(3) Ist $\mathcal{C}_0$ eine σ–Algebra auf Ω mit $\mathcal{C} \subseteq \mathcal{C}_0$ und ist $\mu_0 : \mathcal{C}_0 \to [0, \infty]$ ein vollständiges Maß mit $\mu_0|_\mathcal{C} = \mu$, so gilt $\mathcal{C}_\mu \subseteq \mathcal{C}_0$ und $\mu_0|_{\mathcal{C}_\mu} = \widetilde{\mu}$.
Das Maß $\widetilde{\mu}$ ist die Vervollständigung von μ.

4.2.D Sei $\mathcal{C}$ $\cap$–stabil, sei $\mu : \mathcal{C} \to [0, \infty]$ ein Maß und sei $C \in \mathcal{C}$. Dann ist die Mengenfunktion $\mu_C : \mathcal{C} \to [0, \infty]$ mit

$$\mu_C[A] := \mu[A \cap C]$$

ein Maß.

4.2.E Sei $\{\mu_n\}_{n\in\mathbb{N}}$ eine Folge von Maßen $\mathcal{C} \to [0, \infty]$ und sei $\{a_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$. Dann ist die Mengenfunktion $\mu : \mathcal{C} \to [0, \infty]$ mit

$$\mu[A] := \sum_{n=1}^{\infty} a_n \, \mu_n[A]$$

ein Maß.

4.2.F Ist jede endliche Mengenfunktion σ–endlich?

4.2.G Sei $\mathcal{C}$ eine σ–Algebra und sei $\mu : \mathcal{C} \to [0, \infty]$ ein σ–endliches Maß. Dann ist jede disjunkte Familie $\{A_i\}_{i\in I} \subseteq \mathcal{C}$ mit $\mu[A_i] > 0$ für alle $i \in I$ abzählbar.

4.3 Signierte Maße

Eine Mengenfunktion $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ heißt *signiertes Maß*, wenn sie die folgenden Eigenschaften besitzt:
(i) ν nimmt höchstens einen der Werte $-\infty$ und $+\infty$ an.
(ii) Es gilt $\nu[\emptyset] = 0$.
(iii) Für jede disjunkte Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ mit $\sum_{k=1}^\infty A_k \in \mathcal{C}$ gilt

$$\nu\left[\sum_{k=1}^\infty A_k\right] = \sum_{k=1}^\infty \nu[A_k]$$

Diese Bedingung ist so zu verstehen, dass die Reihe $\sum_{k=1}^\infty \nu[A_k]$ in $\bar{\mathbb{R}}$ unbedingt konvergent ist und den Grenzwert $\nu\bigl[\sum_{k=1}^\infty A_k\bigr]$ besitzt.

Jedes Maß ist ein signiertes Maß, aber nicht jedes signierte Maß ist ein Maß. Ein signiertes Maß $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ist genau dann ein Maß, wenn $\nu(\mathcal{C}) \subseteq [0,\infty]$ gilt. Das folgende Lemma liefert eine einfache Möglichkeit, signierte Maße zu erzeugen, die keine Maße sind:

4.3.1 Lemma. *Seien $\varphi, \psi : \mathcal{C} \to [0,\infty]$ Maße, von denen mindestens eines endlich ist. Dann ist $\varphi - \psi$ ein signiertes Maß.*

Wir wollen nun zeigen, dass sich jedes signierte Maß auf einer σ–Algebra als Differenz zweier Maße, von denen mindestens eines endlich ist, darstellen lässt. Wir benötigen das folgende Analogon zu Folgerung 4.2.6:

4.3.2 Lemma. *Sei $\mathcal{C}$ ein Ring und sei $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ein signiertes Maß. Dann gilt für jede monoton fallende Folge $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ mit $\bigcap_{k\in\mathbb{N}} A_k \in \mathcal{C}$ und $|\nu[A_1]| < \infty$*

$$\lim_{k\to\infty} \nu[A_k] = \nu\left[\bigcap_{k\in\mathbb{N}} A_k\right]$$

Der Beweis von Lemma 4.3.2 ergibt sich durch Wiederholung der in den Beweisen von Lemma 4.2.4 und Lemma 4.2.5 verwendeten Argumente, wobei alle auftretenden Suprema und Infima durch Limites zu ersetzen sind.

Für $A \in \mathcal{C}$ setzen wir

$$\mathcal{C}(A) := \Bigl\{B \in \mathcal{C} \Bigm| B \subseteq A\Bigr\}$$

und für eine Mengenfunktion $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ definieren wir Mengenfunktionen $\nu^+ : \mathcal{C} \to [0,\infty]$ und $\nu^- : \mathcal{C} \to [0,\infty]$ durch

$$\nu^+[A] := \sup_{B\in\mathcal{C}(A)} \nu[B]$$

$$\nu^-[A] := \sup_{B\in\mathcal{C}(A)} (-\nu[B])$$

Die Mengenfunktion ν^+ heißt die *positive Variation* oder der *Positivteil* von ν und die Mengenfunktion ν^- heißt die *negative Variation* oder der *Negativteil* von ν. Für signierte Maße auf einer σ–Algebra erhalten wir das folgende Ergebnis:

4.3.3 Satz (Jordan–Zerlegung). *Sei $\mathcal{C}$ eine σ–Algebra und sei $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ein signiertes Maß. Dann sind ν^+ und ν^- Maße, von denen mindestens eines endlich ist, und es gilt $\nu = \nu^+ - \nu^-$.*

Beweis. Wir führen den Beweis in mehreren Schritten.
(1) Wir zeigen zunächst, dass die Mengenfunktionen ν^+ und ν^- Maße sind:
(i) Es gilt $\nu^+[\emptyset] = \nu[\emptyset] = 0$.
(ii) Sei nun $A \in \mathcal{C}$ und $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine disjunkte Folge mit $\sum_{k=1}^{\infty} A_k = A$. Für jede Menge $B \in \mathcal{C}(A)$ gilt dann

$$\nu[B] = \sum_{k=1}^{\infty} \nu[B \cap A_k] \leq \sum_{k=1}^{\infty} \nu^+[A_k]$$

und damit

$$\nu^+[A] \leq \sum_{k=1}^{\infty} \nu^+[A_k]$$

Sei nun $\{B_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ eine Folge mit $B_k \in \mathcal{C}(A_k)$ für alle $k \in \mathbb{N}$. Dann ist die Folge $\{B_k\}_{k\in\mathbb{N}}$ disjunkt und für alle $n \in \mathbb{N}$ gilt $\sum_{k=1}^{n} B_k \in \mathcal{C}(A)$ und damit

$$\sum_{k=1}^{n} \nu[B_k] = \nu\left[\sum_{k=1}^{n} B_k\right] \leq \nu^+[A]$$

Durch Variation über alle derartigen Folgen erhält man zunächst

$$\sum_{k=1}^{n} \nu^+[A_k] \leq \nu^+[A]$$

und sodann

$$\sum_{k=1}^{\infty} \nu^+[A_k] \leq \nu^+[A]$$

Daher gilt

$$\sum_{k=1}^{\infty} \nu^+[A_k] = \nu^+[A]$$

Damit ist gezeigt, dass ν^+ ein Maß ist. Da mit ν auch $-\nu$ ein signiertes Maß ist, ist auch $\nu^- = (-\nu)^+$ ein Maß.

(2) Wir zeigen nun, dass mindestens eines der Maße ν^+ und ν^- endlich ist. Dazu können wir ohne Beschränkung der Allgemeinheit annehmen, dass für alle $A \in \mathcal{C}$

$$\nu[A] < \infty$$

gilt, und zeigen, dass in diesem Fall ν^+ endlich ist. Wir führen den Beweis durch Widerspruch und nehmen an, dass

$$\nu^+[\Omega] = \infty$$

gilt.
Wir konstruieren nun induktiv eine monoton fallende Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{C}$ mit

$$\nu^+[A_n] = \infty$$

für alle $n \in \mathbb{N}$:

- $n = 1$: Sei

$$A_1 := \Omega$$

  Dann gilt $A_1 \in \mathcal{C}$ und $\nu^+[A_1] = \infty$.
- $n \to n+1$: Sei $A_n \in \mathcal{C}$ mit $\nu^+[A_n] = \infty$ gegeben. Dann gibt es eine Menge $B_n \in \mathcal{C}(A_n)$ mit $\nu[B_n] \geq n$ und $\max\{\nu^+[B_n], \nu^+[A_n \setminus B_n]\} = \infty$ und wir setzen

$$A_{n+1} := \begin{cases} B_n & \text{falls } \nu^+[B_n] = \infty \\ A_n \setminus B_n & \text{sonst} \end{cases}$$

  Dann gilt $A_{n+1} \in \mathcal{C}$ und $\nu^+[A_{n+1}] = \infty$ sowie $A_{n+1} \subseteq A_n$.

Für die Folgen $\{A_n\}_{n\in\mathbb{N}}$ und $\{B_n\}_{n\in\mathbb{N}}$ gilt eine der folgenden Alternativen:

– Es gibt unendlich viele $n \in \mathbb{N}$ mit $A_{n+1} = A_n \setminus B_n$
– Es gibt ein $m \in \mathbb{N}$ mit $A_{n+1} = B_n$ für alle $n \in \mathbb{N}(m)$.

Wir untersuchen diese Alternativen getrennt:

– Im ersten Fall besitzt die Folge $\{B_n\}_{n\in\mathbb{N}}$ eine disjunkte Teilfolge $\{B_{n_k}\}_{k\in\mathbb{N}}$ und es gilt

$$\nu\left[\sum_{k=1}^{\infty} B_{n_k}\right] = \sum_{k=1}^{\infty} \nu[B_{n_k}] \geq \sum_{k=1}^{\infty} n_k = \infty$$

– Im zweiten Fall ist die Folge $\{B_n\}_{n\in\mathbb{N}(m)}$ monoton fallend. In diesem Fall gibt es entweder ein $n \in \mathbb{N}(m)$ mit

$$\nu[B_n] = \infty$$

  oder es gilt $\nu[B_n] < \infty$ für alle $n \in \mathbb{N}(m)$ und aus Lemma 4.3.2 folgt

$$\nu\left[\bigcap_{n\in\mathbb{N}(m)} B_n\right] = \lim_{n\to\infty} \nu[B_n] \geq \lim_{n\to\infty} n = \infty$$

Daher gibt es in jedem Fall ein $B \in \mathcal{C}$ mit $\nu[B] = \infty$. Dies widerspricht der Annahme an ν. Daher ist ν^+ endlich.

(3) Zum Beweis der letzten Behauptung können wir wieder ohne Beschränkung der Allgemeinheit annehmen, dass für alle $A \in \mathcal{C}$

$$\nu[A] < \infty$$

gilt. Dann ist ν^+ endlich. Sei nun $A \in \mathcal{C}$.

– Im Fall $\nu[A] = -\infty$ gilt $\nu^-[A] = \infty$ und damit

$$\nu[A] = \nu^+[A] - \nu^-[A]$$

– Im Fall $\nu[A] \in \mathbb{R}$ gilt

$$\begin{aligned}
\nu^+[A] - \nu[A] &= \sup_{B \in \mathcal{C}(A)} \nu[B] - \nu[A] \\
&= \sup_{B \in \mathcal{C}(A)} (\nu[B] - \nu[A]) \\
&= \sup_{B \in \mathcal{C}(A)} (-\nu[A \setminus B]) \\
&= \sup_{C \in \mathcal{C}(A)} (-\nu[C]) \\
&= \nu^-[A]
\end{aligned}$$

und damit ebenfalls

$$\nu[A] = \nu^+[A] - \nu^-[A]$$

Damit ist auch die letzte Behauptung des Satzes gezeigt. □

Der folgende Satz präzisiert die Jordan–Zerlegung eines signierten Maßes:

4.3.4 Satz (Hahn–Zerlegung). *Sei $\mathcal{C}$ eine σ–Algebra und sei $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ein signiertes Maß. Dann gibt es disjunkte Mengen $\Omega^+, \Omega^- \in \mathcal{C}$ mit $\Omega = \Omega^+ + \Omega^-$ sowie*

$$\nu^+[A] = \nu[A \cap \Omega^+]$$

und

$$\nu^-[A] = -\,\nu[A \cap \Omega^-]$$

für alle $A \in \mathcal{C}$.

Beweis. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass ν^+ endlich ist. Dann gibt es eine Folge $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{C}$ mit

$$\nu^+[\Omega] \le \nu[A_n] + 1/2^n$$

und wir setzen

$$\Omega^+ := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$
$$\Omega^- := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} (\Omega \setminus A_k)$$

Dann gilt $\Omega^+, \Omega^- \in \mathcal{C}$ sowie $\Omega^+ \cap \Omega^- = \emptyset$ und $\Omega = \Omega^+ + \Omega^-$. Aus der Jordan–Zerlegung erhalten wir $\nu^-[A_k] = \nu^+[A_k] - \nu[A_k] \le \nu^+[\Omega] - \nu[A_k] \le 1/2^k$, und aus Lemma 4.2.12 folgt nun

$$\nu^-\left[\bigcup_{k=n}^{\infty} A_k\right] \le \sum_{k=n}^{\infty} \nu^-[A_k] \le \sum_{k=n}^{\infty} \frac{1}{2^k} \le \frac{1}{2^{n-1}}$$

Insbesondere gilt $\nu^-[\bigcup_{k=1}^{\infty} A_k] \le 1 < \infty$ und aus Folgerung 4.2.6 ergibt sich nun

$$\nu^-[\Omega^+] = \nu^-\left[\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right] = \lim_{n\to\infty} \nu^-\left[\bigcup_{k=n}^{\infty} A_k\right] = 0$$

Ferner gilt $\nu^+[A_n] + \nu^+[\Omega \setminus A_n] = \nu^+[\Omega] \le \nu[A_n] + 1/2^n \le \nu^+[A_n] + 1/2^n$. Da ν^+ endlich ist, folgt daraus zunächst

$$\nu^+\left[\bigcap_{k=n}^{\infty} (\Omega \setminus A_k)\right] \le \nu^+[\Omega \setminus A_n] \le 1/2^n$$

und sodann mit Lemma 4.2.4

$$\nu^+[\Omega^-] = \nu^+\left[\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} (\Omega \setminus A_k)\right] = \lim_{n\to\infty} \nu^+\left[\bigcap_{k=n}^{\infty} (\Omega \setminus A_k)\right] = 0$$

Aus der Jordan–Zerlegung erhalten wir nun für alle $A \in \mathcal{C}$

$$\begin{aligned}
\nu^+[A] &= \nu^+[A \cap \Omega^+] + \nu^+[A \cap \Omega^-] \\
&= \nu^+[A \cap \Omega^+] \\
&= \nu^+[A \cap \Omega^+] - \nu^-[A \cap \Omega^+] \\
&= \nu[A \cap \Omega^+]
\end{aligned}$$

und analog zeigt man $\nu^-[A] = -\nu[A \cap \Omega^-]$. □

Eine Mengenfunktion $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ heißt

– *endlich*, wenn für alle $A \in \mathcal{C}$
$$|\nu[A]| < \infty$$
gilt.

– *beschränkt*, wenn
$$\sup_{A \in \mathcal{C}} |\nu[A]| < \infty$$
gilt.

Jede beschränkte Mengenfunktion ist endlich.

Aufgaben

4.3.A Sei $\mathcal{C}$ eine σ–Algebra und seien $\varphi, \psi : \mathcal{C} \to [0, \infty]$ Maße. Dann ist jede der Mengenfunktionen $\varphi \vee \psi, \varphi \wedge \psi : \mathcal{C} \to [0, \infty]$ mit

$$\begin{aligned}(\varphi \vee \psi)[A] &:= \sup_{B \in \mathcal{C}(A)} \Big(\varphi[B] + \psi[A \setminus B] \Big) \\ (\varphi \wedge \psi)[A] &:= \inf_{B \in \mathcal{C}(A)} \Big(\varphi[B] + \psi[A \setminus B] \Big)\end{aligned}$$

ein Maß.

4.3.B Formulieren und beweisen Sie für Inhalte Analoga zu den Aussagen von Aufgabe 4.3.A.

4.3.C Formulieren und beweisen Sie für endliche signierte Maße Analoga zu den Aussagen von Aufgabe 4.3.A.

4.3.D **Jordan–Zerlegung:** Sei $\mathcal{C}$ eine σ–Algebra und sei $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ein signiertes Maß. Dann ist die Mengenfunktion $|\nu| : \mathcal{C} \to [0, \infty]$ mit

$$|\nu|[A] := \sup \left\{ \sum_{k=1}^{n} |\nu[A_k]| \;\middle|\; n \in \mathbb{N} \text{ und } \{A_k\}_{k \in \{1,\ldots,n\}} \subseteq \mathcal{C}(A) \text{ disjunkt} \right\}$$

ein Maß und es gilt $|\nu| = \nu^+ + \nu^-$ sowie $\nu^+ \vee \nu^- = |\nu|$ und $\nu^+ \wedge \nu^- = 0$. Das Maß $|\nu|$ heißt die *totale Variation* von ν.

4.3.E **Jordan–Zerlegung:** Sei $\mathcal{C}$ eine σ–Algebra und $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ein signiertes Maß. Sind $\varphi, \psi : \mathcal{C} \to [0, \infty]$ Maße mit $\nu = \varphi - \psi$, so gilt $\nu^+ \le \varphi$ und $\nu^- \le \psi$; im Fall $\varphi \wedge \psi = 0$ gilt sogar $\nu^+ = \varphi$ und $\nu^- = \psi$.

4.3.F **Hahn–Zerlegung:** Sei $\mathcal{C}$ eine σ–Algebra und $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ein signiertes Maß. Für $i \in \{1, 2\}$ seien $P_i, N_i \in \mathcal{C}$ Mengen mit $P_i \cap N_i = \emptyset$ und $P_i + N_i = \Omega$ sowie

$$\begin{aligned}\nu^+[A] &= \nu[A \cap P_i] \\ \nu^-[A] &= -\nu[A \cap N_i]\end{aligned}$$

für alle $A \in \mathcal{C}$. Dann gilt $|\nu|(P_1 \triangle P_2) = 0 = |\nu|(N_1 \triangle N_2)$.

4.3.G Sei $\mathcal{C}$ eine σ–Algebra und $\nu : \mathcal{C} \to \bar{\mathbb{R}}$ ein signiertes Maß. Dann sind folgende Aussagen äquivalent:
(a) ν ist endlich.
(b) ν ist beschränkt.

5

Fortsetzung von Maßen

In diesem Kapitel untersuchen wir für ein Maß auf einem Mengensystem, das die leere Menge enthält, die Existenz und die Eindeutigkeit einer Fortsetzung zu einem Maß auf einer σ–Algebra, die das gegebene Mengensystem enthält.

Wir beginnen mit einem Eindeutigkeitssatz (Abschnitt 5.1). Als Vorbereitung für den Beweis eines Fortsetzungssatzes führen wir dann äußere Maße ein und zeigen, dass es zu jedem äußeren Maß eine σ–Algebra gibt, für die die Restriktion des äußeren Maßes ein Maß ist (Abschnitt 5.2). Wir beweisen sodann den Satz von Caratheodory über die Existenz und Eindeutigkeit einer Fortsetzung eines σ–endlichen Maßes von einem Halbring auf die von ihm erzeugte σ–Algebra (Abschnitt 5.3). Aus dem Fortsetzungssatz erhalten wir einen nützlichen Approximationssatz (Abschnitt 5.4) und das Lebesgue–Maß auf der Borelschen σ–Algebra des Euklidischen Raumes (Abschnitt 5.5).

Im gesamten Kapitel sei Ω eine nichtleere Menge.

5.1 Eindeutigkeitssatz

In diesem Abschnitt zeigen wir, dass ein σ–endliches Maß auf einem nichtleeren $\cap$–stabilen Mengensystem höchstens eine Fortsetzung auf die erzeugte σ–Algebra besitzt:

5.1.1 Satz (Eindeutigkeitssatz). *Sei $\mathcal{F} \subseteq 2^\Omega$ eine σ–Algebra und sei $\mathcal{E} \subseteq 2^\Omega$ ein $\cap$–stabiles Mengensystem mit $\sigma(\mathcal{E}) = \mathcal{F}$. Sind $\mu, \nu : \mathcal{F} \to [0, \infty]$ Maße mit $\mu|_\mathcal{E} = \nu|_\mathcal{E}$ und ist $\mu|_\mathcal{E}$ σ–endlich, so gilt $\mu = \nu$ und μ ist σ–endlich.*

Beweis. Da $\mu|_\mathcal{E}$ σ–endlich ist, gibt es eine Folge $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathcal{E}$ mit

$$\bigcup_{n\in\mathbb{N}} E_n = \Omega$$

und $\mu[E_n] < \infty$ für alle $n \in \mathbb{N}$.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_5, © Springer-Verlag Berlin Heidelberg 2011

(1) Für $E \in \mathcal{E}$ mit $\mu[E] = \nu[E] < \infty$ sei

$$\mathcal{D}_E := \Bigl\{ A \in \mathcal{F} \Bigm| \mu[E \cap A] = \nu[E \cap A] \Bigr\}$$

Dann ist das Mengensystem $\mathcal{D}_E$ ein Dynkin–System:
(i) Es gilt $\Omega \in \mathcal{D}_E$.
(ii) Sei $A \in \mathcal{D}_E$. Dann gilt

$$\mu[E \cap \overline{A}] = \mu[E] - \mu[E \cap A] = \nu[E] - \nu[E \cap A] = \nu[E \cap \overline{A}]$$

und damit $\overline{A} \in \mathcal{D}_E$.
(iii) Sei $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{D}_E$ disjunkt. Dann gilt

$$\mu\left[E \cap \sum_{n=1}^\infty A_n\right] = \sum_{n=1}^\infty \mu[E \cap A_n] = \sum_{n=1}^\infty \nu[E \cap A_n] = \nu\left[E \cap \sum_{n=1}^\infty A_n\right]$$

und damit $\sum_{n\in\mathbb{N}} A_n \in \mathcal{D}_E$.
Da $\mathcal{E}$ $\cap$–stabil ist, gilt $\mathcal{E} \subseteq \mathcal{D}_E$ und damit

$$\delta(\mathcal{E}) \subseteq \mathcal{D}_E$$

Da $\mathcal{E}$ $\cap$–stabil ist, gilt außerdem $\sigma(\mathcal{E}) = \delta(\mathcal{E})$, und damit

$$\mathcal{F} = \sigma(\mathcal{E}) = \delta(\mathcal{E}) \subseteq \mathcal{D}_E$$

Damit ist gezeigt, dass für alle $E \in \mathcal{E}$ mit $\mu[E] < \infty$ und für alle $A \in \mathcal{F}$

$$\mu[E \cap A] = \nu[E \cap A]$$

gilt.
(2) Wir betrachten nun die Folge $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathcal{E}$. Für alle $n \in \mathbb{N}_0$ sei

$$C_n := \bigcup_{k=1}^n E_k$$

und für alle $n \in \mathbb{N}$ sei

$$B_n := C_n \setminus C_{n-1}$$

Dann ist $\{B_n\}_{n\in\mathbb{N}}$ eine disjunkte Folge in $\mathcal{F}$ mit $\sum_{n=1}^\infty B_n = \Omega$, und für alle $n \in \mathbb{N}$ gilt $B_n \subseteq E_n$.
(3) Für alle $A \in \mathcal{F}$ gilt wegen (2)

$$A = \sum_{n=1}^\infty B_n \cap A = \sum_{n=1}^\infty E_n \cap (B_n \cap A)$$

und aus (1) folgt nun wegen $\mu[E_n] < \infty$ und $B_n \cap A \in \mathcal{F} \subseteq \mathcal{D}_{E_n}$

$$\mu[A] = \sum_{n=1}^\infty \mu[E_n \cap (B_n \cap A)] = \sum_{n=1}^\infty \nu[E_n \cap (B_n \cap A)] = \nu[A]$$

Damit ist gezeigt, dass $\mu = \nu$ gilt. □

Aufgabe

5.1.A Sei $\mathcal{F} \subseteq 2^\Omega$ eine σ–Algebra und sei $\mathcal{A} \subseteq 2^\Omega$ eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{F}$. Sind $\mu, \nu : \mathcal{F} \to [0, \infty]$ endliche Maße mit $\mu[\Omega] = \nu[\Omega]$ und

$$\mu[A] \leq \nu[A]$$

für alle $A \in \mathcal{A}$, so gilt $\mu = \nu$.

5.2 Äußere Maße

Eine Mengenfunktion $\varrho : 2^\Omega \to [0, \infty]$ heißt *äußeres Maß*, wenn sie folgende Eigenschaften besitzt:
(i) Es gilt $\varrho[\emptyset] = 0$.
(ii) ϱ ist monoton.
(iii) ϱ ist σ–subadditiv.
Jedes Maß $2^\Omega \to [0, \infty]$ ist ein äußeres Maß, und jedes äußere Maß ist endlich subadditiv.

Für ein äußeres Maß $\varrho : 2^\Omega \to [0, \infty]$ bestimmen wir nun eine σ–Algebra $\mathcal{F}(\varrho)$ derart, dass die Restriktion von ϱ auf $\mathcal{F}(\varrho)$ ein Maß ist. Da ϱ σ–subadditiv ist, genügt es nach Lemma 4.2.12 zu zeigen, dass die Restriktion von ϱ auf $\mathcal{F}(\varrho)$ ein Inhalt ist, und wegen $\varrho[\emptyset] = 0$ genügt es nach Lemma 4.1.2 zu zeigen, dass die Restriktion von ϱ auf $\mathcal{F}(\varrho)$ additiv ist. Diese Überlegung legt die folgende Definition nahe:

Für ein äußeres Maß $\varrho : 2^\Omega \to [0, \infty]$ sei

$$\mathcal{F}(\varrho) := \Bigl\{ A \in 2^\Omega \Bigm| \text{für alle } E \in 2^\Omega \text{ gilt } \varrho[E] = \varrho[E \cap A] + \varrho[E \setminus A] \Bigr\}$$

Wir bezeichnen das Mengensystem $\mathcal{F}(\varrho)$ als *System der additiven Zerleger* bezüglich ϱ. Da jedes äußere Maß subadditiv ist, gilt

$$\mathcal{F}(\varrho) = \Bigl\{ A \in 2^\Omega \Bigm| \text{für alle } E \in 2^\Omega \text{ gilt } \varrho[E] \geq \varrho[E \cap A] + \varrho[E \setminus A] \Bigr\}$$

Der folgende Satz zeigt, dass das Mengensystem $\mathcal{F}(\varrho)$ eine σ–Algebra ist und dass die Restriktion $\varrho|_{\mathcal{F}(\varrho)} : \mathcal{F}(\varrho) \to [0, \infty]$ von ϱ auf $\mathcal{F}(\varrho)$ ein Maß ist:

5.2.1 Satz. *Sei $\varrho : 2^\Omega \to [0, \infty]$ ein äußeres Maß. Dann ist $\mathcal{F}(\varrho)$ eine σ–Algebra und $\varrho|_{\mathcal{F}(\varrho)}$ ist ein Maß.*

Beweis. Wir untersuchen zunächst das Mengensystem $\mathcal{F}(\varrho)$.
Als erstes zeigen wir, dass $\mathcal{F}(\varrho)$ $\cap$–stabil ist. Sei $A, B \in \mathcal{F}(\varrho)$ und $E \in 2^\Omega$. Wegen $E \setminus (A \cap B) = (E \cap A) \setminus B + (E \setminus A) \cap B + (E \setminus A) \setminus B$ folgt aus der endlichen Subadditivität von ϱ

$$\begin{aligned}\varrho[E] &\le \varrho[E \cap (A \cap B)] + \varrho[E \setminus (A \cap B)] \\ &\le \varrho[(E \cap A) \cap B] + \varrho[(E \cap A) \setminus B] + \varrho[(E \setminus A) \cap B] + \varrho[(E \setminus A) \setminus B] \\ &= \varrho[E \cap A] + \varrho[E \setminus A] \\ &= \varrho[E]\end{aligned}$$

und damit

$$\varrho[E] = \varrho[E \cap (A \cap B)] + \varrho[E \setminus (A \cap B)]$$

Daher gilt $A \cap B \in \mathcal{F}(\varrho)$. Damit ist gezeigt, dass $\mathcal{F}(\varrho)$ $\cap$–stabil ist.
Als nächstes zeigen wir, dass $\mathcal{F}(\varrho)$ ein Dynkin–System ist:

(i) Es gilt $\Omega \in \mathcal{F}(\varrho)$.

(ii) Für alle $A \in \mathcal{F}(\varrho)$ gilt $\overline{A} \in \mathcal{F}(\varrho)$.

(iii) Sei $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{F}(\varrho)$ disjunkt und $E \in 2^\Omega$. Wir betrachten zunächst die Menge $A_1 + A_2$. Da $\mathcal{F}(\varrho)$ $\cap$–stabil ist, ergibt sich aus (ii) und den Gesetzen von DeMorgan $A_1 + A_2 \in \mathcal{F}(\varrho)$ und es gilt

$$\begin{aligned}\varrho[E \cap (A_1 + A_2)] &= \varrho[E \cap (A_1 + A_2) \cap A_1] + \varrho[E \cap (A_1 + A_2) \setminus A_1] \\ &= \varrho[E \cap A_1] + \varrho[E \cap A_2]\end{aligned}$$

Durch vollständige Induktion ergibt sich sodann $\sum_{k=1}^n A_k \in \mathcal{F}(\varrho)$ und

$$\varrho\left[E \cap \sum_{k=1}^n A_k\right] = \sum_{k=1}^n \varrho[E \cap A_k]$$

und aus der Monotonie von ϱ folgt nun

$$\begin{aligned}\sum_{k=1}^n \varrho[E \cap A_k] + \varrho\left[E \setminus \sum_{k=1}^\infty A_k\right] &= \varrho\left[E \cap \sum_{k=1}^n A_k\right] + \varrho\left[E \setminus \sum_{k=1}^\infty A_k\right] \\ &\le \varrho\left[E \cap \sum_{k=1}^n A_k\right] + \varrho\left[E \setminus \sum_{k=1}^n A_k\right] \\ &= \varrho[E]\end{aligned}$$

Daher gilt

$$\sum_{k=1}^\infty \varrho[E \cap A_k] + \varrho\left[E \setminus \sum_{k=1}^\infty A_k\right] \le \varrho[E]$$

Aus der σ–Subadditivität von ϱ und der letzten Ungleichung folgt nun

$$\varrho[E] \le \varrho\left[E \cap \sum_{k=1}^\infty A_k\right] + \varrho\left[E \setminus \sum_{k=1}^\infty A_k\right]$$

$$
\begin{aligned}
&= \varrho\left[\sum_{k=1}^{\infty} E \cap A_k\right] + \varrho\left[E \setminus \sum_{k=1}^{\infty} A_k\right] \\
&\le \sum_{k=1}^{\infty} \varrho[E \cap A_k] + \varrho\left[E \setminus \sum_{k=1}^{\infty} A_k\right] \\
&\le \varrho[E]
\end{aligned}
$$

und damit

$$
\varrho[E] = \varrho\left[E \cap \sum_{k=1}^{\infty} A_k\right] + \varrho\left[E \setminus \sum_{k=1}^{\infty} A_k\right]
$$

Daher gilt $\sum_{k=1}^{\infty} A_k \in \mathcal{F}(\varrho)$.
Damit ist gezeigt, dass $\mathcal{F}(\varrho)$ ein Dynkin–System ist.
Da $\mathcal{F}(\varrho)$ ein $\cap$–stabiles Dynkin–System ist, ist $\mathcal{F}(\varrho)$ eine σ–Algebra.
Wir zeigen nun, dass $\varrho|_{\mathcal{F}(\varrho)}$ ein Inhalt ist:
(i) Es gilt $\varrho|_{\mathcal{F}(\varrho)}[\emptyset] = \varrho[\emptyset] = 0$.
(ii) Für alle $A, B \in \mathcal{F}(\varrho)$ mit $A \cap B = \emptyset$ gilt $A + B \in \mathcal{F}(\varrho)$ und damit

$$
\begin{aligned}
\varrho|_{\mathcal{F}(\varrho)}[A+B] &= \varrho[A+B] \\
&= \varrho[(A+B) \cap A] + \varrho[(A+B) \setminus A] \\
&= \varrho[A] + \varrho[B] \\
&= \varrho|_{\mathcal{F}(\varrho)}[A] + \varrho|_{\mathcal{F}(\varrho)}[B]
\end{aligned}
$$

Daher ist $\varrho|_{\mathcal{F}(\varrho)}$ ein Inhalt.
Da $\varrho|_{\mathcal{F}(\varrho)}$ ein Inhalt ist und mit ϱ auch $\varrho|_{\mathcal{F}(\varrho)}$ σ–subadditiv ist, ist $\varrho|_{\mathcal{F}(\varrho)}$ nach Lemma 4.2.12 ein Maß. □

Aufgabe

5.2.A Für jedes äußere Maß $\varrho : 2^{\Omega} \to [0, \infty]$ ist das Maß $\varrho|_{\mathcal{F}(\varrho)} : \mathcal{F}(\varrho) \to [0, \infty]$ vollständig.

5.3 Existenzsatz

Sei $\mathcal{E} \subseteq 2^{\Omega}$ ein Mengensystem mit $\emptyset \in \mathcal{E}$ und sei $\mu : \mathcal{E} \to [0, \infty]$ eine Mengenfunktion mit $\mu[\emptyset] = 0$. Dann ist die Abbildung $\mu^* : 2^{\Omega} \to [0, \infty]$ mit

$$
\mu^*[A] := \inf_{\{E_k\}_{k\in\mathbb{N}} \subseteq \mathcal{E},\, A \subseteq \bigcup_{k\in\mathbb{N}} E_k} \sum_{k=1}^{\infty} \mu[E_k]
$$

eine Mengenfunktion. Der folgende Satz zeigt, dass μ^* ein äußeres Maß ist:

5.3.1 Lemma (Konstruktion eines äußeren Maßes). *Sei $\mathcal{E} \subseteq 2^{\Omega}$ ein Mengensystem mit $\emptyset \in \mathcal{E}$ und sei $\mu : \mathcal{E} \to [0, \infty]$ eine Mengenfunktion mit $\mu[\emptyset] = 0$. Dann ist μ^* ein äußeres Maß.*

Beweis. Wir zeigen, dass μ^* die Axiome eines äußeren Maßes erfüllt:
(i) Es gilt $\mu^*[\emptyset] = 0$.
(ii) Für alle $A, B \in 2^\Omega$ mit $B \subseteq A$ gilt $\mu^*[B] \leq \mu^*[A]$.
(iii) Sei $\{A_n\}_{n\in\mathbb{N}} \subseteq 2^\Omega$. Wir nehmen zunächst an, dass für alle $n \in \mathbb{N}$

$$\mu^*[A_n] < \infty$$

gilt. Sei $\varepsilon \in (0,\infty)$. Dann gibt es für alle $n \in \mathbb{N}$ eine Folge $\{E_{n,k}\}_{k\in\mathbb{N}} \subseteq \mathcal{E}$ mit $A_n \subseteq \bigcup_{k\in\mathbb{N}} E_{n,k}$ und

$$\sum_{k=1}^{\infty} \mu[E_{n,k}] \leq \mu^*[A_n] + \varepsilon/2^n$$

Daher gilt

$$\begin{aligned} \mu^*\left[\bigcup_{n\in\mathbb{N}} A_n\right] &\leq \mu^*\left[\bigcup_{n\in\mathbb{N}}\bigcup_{k\in\mathbb{N}} E_{n,k}\right] \\ &\leq \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mu[E_{n,k}] \\ &\leq \sum_{n=1}^{\infty}\Big(\mu^*[A_n] + \varepsilon/2^n\Big) \\ &= \sum_{n=1}^{\infty} \mu^*[A_n] + \varepsilon \end{aligned}$$

Da $\varepsilon \in (0,\infty)$ beliebig war, folgt daraus

$$\mu^*\left[\bigcup_{n\in\mathbb{N}} A_n\right] \leq \sum_{n=1}^{\infty} \mu^*[A_n]$$

Diese Ungleichung gilt auch dann, wenn die Bedingung $\mu^*[A_n] < \infty$ für ein $n \in \mathbb{N}$ verletzt ist. Daher ist μ^* σ–subadditiv.

Damit ist gezeigt, dass μ^* ein äußeres Maß ist. □

Lemma 5.3.1 enthält keine Aussage über die Beziehung zwischen der Mengenfunktion $\mu : \mathcal{E} \to [0,\infty]$ und der Restriktion $\mu^*|_\mathcal{E} : \mathcal{E} \to [0,\infty]$ des von ihr erzeugten äußeren Maßes. Unter einer zusätzlichen Bedingung an das Mengensystem $\mathcal{E}$ erhält man jedoch die Übereinstimmung von μ und $\mu^*|_\mathcal{E}$:

5.3.2 Satz (Existenzsatz). *Sei $\mathcal{E} \subseteq 2^\Omega$ ein Halbring und sei $\mu : \mathcal{E} \to [0,\infty]$ ein Maß. Dann gilt $\mathcal{E} \subseteq \mathcal{F}(\mu^*)$ und $\mu^*|_\mathcal{E} = \mu$. Insbesondere ist $\mu^*|_{\mathcal{F}(\mu^*)}$ ein Maß und eine Fortsetzung von μ.*

Beweis. Aufgrund von Satz 4.2.3 können wir annehmen, dass $\mathcal{E}$ ein Ring ist.

(1) Wir zeigen zunächst, dass $\mathcal{E} \subseteq \mathcal{F}(\mu^*)$ gilt. Sei $A \in \mathcal{E}$ und $E \in 2^\Omega$. Da μ^* ein äußeres Maß ist, gilt

$$\mu^*[E] \leq \mu^*[E \cap A] + \mu^*[E \setminus A]$$

Im Fall $\mu^*[E] = \infty$ folgt daraus

$$\mu^*[E] = \mu^*[E \cap A] + \mu^*[E \setminus A]$$

Wir nehmen nun an, dass $\mu^*[E] < \infty$ gilt. Dann gibt es zu jedem $\varepsilon \in (0, \infty)$ eine Folge $\{E_k\}_{k \in \mathbb{N}} \subseteq \mathcal{E}$ mit $E \subseteq \bigcup_{k \in \mathbb{N}} E_k$ und

$$\sum_{n=1}^{\infty} \mu[E_k] \leq \mu^*[E] + \varepsilon$$

Da $\mathcal{E}$ ein Ring ist, sind $\{E_k \cap A\}_{k \in \mathbb{N}}$ und $\{E_k \setminus A\}_{k \in \mathbb{N}}$ Folgen in $\mathcal{E}$, und wegen $E \cap A \subseteq \bigcup_{k \in \mathbb{N}} E_k \cap A$ und $E \setminus A \subseteq \bigcup_{k \in \mathbb{N}} E_k \setminus A$ erhalten wir

$$\begin{aligned}
\mu^*[E] &\leq \mu^*[E \cap A] + \mu^*[E \setminus A] \\
&\leq \sum_{k=1}^{\infty} \mu[E_k \cap A] + \sum_{k=1}^{\infty} \mu[E_k \setminus A] \\
&= \sum_{k=1}^{\infty} \Big(\mu[E_k \cap A] + \mu[E_k \setminus A] \Big) \\
&= \sum_{k=1}^{\infty} \mu[E_k] \\
&\leq \mu^*[E] + \varepsilon
\end{aligned}$$

Da $\varepsilon \in (0, \infty)$ beliebig war, ergibt sich daraus

$$\mu^*[E] = \mu^*[E \cap A] + \mu^*[E \setminus A]$$

Daher gilt für alle $E \in 2^\Omega$

$$\mu^*[E] = \mu^*[E \cap A] + \mu^*[E \setminus A]$$

und damit $A \in \mathcal{F}(\mu^*)$. Damit ist gezeigt, dass $\mathcal{E} \subseteq \mathcal{F}(\mu^*)$ gilt.
(2) Wir zeigen nun, dass $\mu^*|_{\mathcal{E}} = \mu$ gilt. Sei $A \in \mathcal{E}$. Dann gilt

$$\mu^*[A] \leq \mu[A]$$

Sei nun $\{E_k\}_{k \in \mathbb{N}} \subseteq \mathcal{E}$ eine Folge mit $A \subseteq \bigcup_{k \in \mathbb{N}} E_k$. Da $\mathcal{E}$ ein Ring ist, ist $\{A \cap E_k\}_{k \in \mathbb{N}}$ eine Folge in $\mathcal{E}$ mit $A = \bigcup_{k \in \mathbb{N}} A \cap E_k$, und da μ σ–subadditiv und monoton ist, erhalten wir

$$\mu[A] \leq \sum_{k=1}^{\infty} \mu[A \cap E_k] \leq \sum_{k=1}^{\infty} \mu[E_k]$$

Daher gilt

$$\mu[A] \leq \mu^*[A]$$

und damit

$$\mu[A] = \mu^*[A]$$

Damit ist gezeigt, dass $\mu^*|_{\mathcal{E}} = \mu$ gilt. □

Aus dem Existenzsatz und dem Eindeutigkeitssatz ergibt sich nun der wichtige Existenz– und Eindeutigkeitssatz von Caratheodory:

5.3.3 Satz (Caratheodory). *Sei $\mathcal{E} \subseteq 2^\Omega$ ein Halbring und sei $\mu : \mathcal{E} \to [0, \infty]$ ein σ–endliches Maß. Dann gibt es genau ein Maß $\widetilde{\mu} : \sigma(\mathcal{E}) \to [0, \infty]$ mit $\widetilde{\mu}|_{\mathcal{E}} = \mu$. Das Maß $\widetilde{\mu}$ ist σ–endlich und es gilt $\widetilde{\mu} = \mu^*|_{\sigma(\mathcal{E})}$.*

Beweis. Aus dem Existenzsatz folgt, dass $\mathcal{E} \subseteq \mathcal{F}(\mu^*)$ und damit auch $\sigma(\mathcal{E}) \subseteq \mathcal{F}(\mu^*)$ gilt und dass $\mu^*|_{\mathcal{F}(\mu^*)}$ ein Maß und eine Fortsetzung von μ auf $\mathcal{F}(\mu^*)$ ist. Wegen $\sigma(\mathcal{E}) \subseteq \mathcal{F}(\mu^*)$ ist das Maß $\mu^*|_{\sigma(\mathcal{E})}$ eine Fortsetzung von μ auf $\sigma(\mathcal{E})$. Da μ σ–endlich ist, folgt nun aus dem Eindeutigkeitssatz, dass $\mu^*|_{\sigma(\mathcal{E})}$ die einzige Fortsetzung von μ auf $\sigma(\mathcal{E})$ ist und dass $\mu^*|_{\sigma(\mathcal{E})}$ σ–endlich ist. □

Aufgaben

5.3.A **Caratheodory:** Sei $\mathcal{E} \subseteq 2^\Omega$ ein Halbring und sei $\mu : \mathcal{E} \to [0, \infty]$ ein σ–endliches Maß. Dann gibt es genau ein Maß $\widetilde{\mu} : \mathcal{F}(\mu^*) \to [0, \infty]$ mit $\widetilde{\mu}|_{\mathcal{E}} = \mu$. Das Maß $\widetilde{\mu}$ ist σ–endlich und es gilt $\widetilde{\mu} = \mu^*|_{\mathcal{F}(\mu^*)}$.

5.3.B Sei $\mathcal{E} \subseteq 2^\Omega$ ein Halbring und sei $\mu : \mathcal{E} \to [0, \infty]$ ein Maß. Dann ist das Maß $\mu^*|_{\mathcal{F}(\mu^*)} : \mathcal{F}(\mu^*) \to [0, \infty]$ vollständig.

5.3.C **Vervollständigung:** Sei $\mathcal{E} \subseteq 2^\Omega$ ein Halbring und sei $\mu : \mathcal{E} \to [0, \infty]$ ein σ–endliches Maß. Sei ferner $\mathcal{F}_0 \subseteq 2^\Omega$ eine σ–Algebra mit $\mathcal{E} \subseteq \mathcal{F}_0$ und sei $\mu_0 : \mathcal{F}_0 \to [0, \infty]$ ein vollständiges Maß mit $\mu_0|_{\mathcal{E}} = \mu$. Dann gilt $\mathcal{F}(\mu^*) \subseteq \mathcal{F}_0$ und $\mu_0|_{\mathcal{F}(\mu^*)} = \mu^*|_{\mathcal{F}(\mu^*)}$.

5.4 Approximationssatz

Aus dem Satz von Caratheodory erhält man den folgenden Approximationssatz für endliche Maße:

5.4.1 Satz (Approximationssatz). *Sei $\mathcal{F} \subseteq 2^\Omega$ eine σ–Algebra und sei $\mathcal{A} \subseteq 2^\Omega$ eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{F}$. Ist $\mu : \mathcal{F} \to [0, \infty]$ ein endliches Maß, so gibt es für jede Menge $A \in \mathcal{F}$ eine Folge $\{C_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ mit*

$$\lim_{n \to \infty} \mu[A \triangle C_n] = 0$$

Beweis. Da $\mathcal{A}$ eine Algebra und μ endlich ist, ist insbesondere $\mathcal{A}$ ein Halbring und $\mu|_{\mathcal{A}}$ σ–endlich. Nach dem Satz von Caratheodory gilt daher

$$\mu = (\mu|_{\mathcal{A}})^*|_{\mathcal{F}}$$

Sei $A \in \mathcal{F}$ und $\varepsilon \in (0, \infty)$. Nach Konstruktion des äußeren Maßes $(\mu|_{\mathcal{A}})^*$ gibt es eine Folge $\{E_k\}_{k\in\mathbb{N}} \subseteq \mathcal{A}$ mit $A \subseteq \bigcup_{k\in\mathbb{N}} E_k \in \mathcal{F}$ und

$$\sum_{k=1}^{\infty} \mu|_{\mathcal{A}}[E_k] \leq (\mu|_{\mathcal{A}})^*[A] + \varepsilon/2$$

Daher gilt

$$\sum_{k=1}^{\infty} \mu[E_k] \leq \mu[A] + \varepsilon/2$$

Sei nun $E := \bigcup_{k\in\mathbb{N}} E_k$. Da μ σ–subadditiv ist, ergibt sich nun

$$\mu[A] + \mu[E \setminus A] = \mu[E] \leq \sum_{k=1}^{\infty} \mu[E_k] \leq \mu[A] + \varepsilon/2$$

und aus der Endlichkeit von μ folgt dann

$$\mu[E \setminus A] \leq \varepsilon/2$$

Für alle $n \in \mathbb{N}$ sei $C_n := \bigcup_{k=1}^{n} E_k$. Dann ist $\{C_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge in $\mathcal{A}$ mit $\bigcup_{n\in\mathbb{N}} C_n = E$, und damit ist $\{E\setminus C_n\}_{n\in\mathbb{N}}$ eine monoton fallende Folge in $\mathcal{F}$ mit $\bigcap_{n\in\mathbb{N}} E \setminus C_n = \emptyset$. Da μ endlich ist, gilt $\lim_{n\to\infty} \mu[E \setminus C_n] = 0$. Für alle hinreichen großen $n \in \mathbb{N}$ gilt daher

$$\mu[E \setminus C_n] \leq \varepsilon/2$$

Wegen $A \cup C_n \subseteq E$ gilt dann für alle hinreichen großen $n \in \mathbb{N}$

$$\mu[A \triangle C_n] = \mu[A \setminus C_n] + \mu[C_n \setminus A] \leq \mu[E \setminus C_n] + \mu[E \setminus A] \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

Daraus folgt die Behauptung. □

5.4.2 Folgerung. *Sei $\mathcal{F} \subseteq 2^\Omega$ eine σ–Algebra und sei $\mathcal{A} \subseteq 2^\Omega$ eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{F}$. Ist $\mu : \mathcal{F} \to [0, \infty]$ ein endliches Maß, so gibt es für jede Menge $A \in \mathcal{F}$ eine Folge $\{C_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ mit*

$$\lim_{n\to\infty} \mu[C_n] = \mu[A]$$

Beweis. Für alle $A, C \in \mathcal{F}$ gilt

$$\mu[A] - \mu[C] = \mu[A \setminus C] - \mu[C \setminus A] \leq \mu[A \setminus C] + \mu[C \setminus A] = \mu[A \triangle C]$$

und damit

$$|\mu[A] - \mu[C]| \leq \mu[A \triangle C]$$

Die Behauptung folgt nun aus dem Approximationssatz. □

Aufgabe

5.4.A Sei $\mathcal{F} \subseteq 2^\Omega$ eine σ–Algebra und sei $\mathcal{A} \subseteq 2^\Omega$ eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{F}$. Ist $\mu : \mathcal{F} \to [0, \infty]$ ein endliches Maß, so gibt es für jede Menge $A \in \mathcal{F}$ eine Folge $\{C_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ mit

$$\lim_{n\to\infty} d_\mu(A, C_n) = 0$$

5.5 Lebesgue–Maß

Auf dem Halbring $\mathcal{J}(\mathbb{R}^n)$ der halboffenen Intervalle des $\mathbb{R}^n$ wird in natürlicher Weise durch

$$\boldsymbol{\lambda}^n[(\mathbf{a}, \mathbf{b}]] := \begin{cases} \prod_{i=1}^{n} (b_i - a_i) & \text{falls } \mathbf{a} < \mathbf{b} \\ 0 & \text{sonst} \end{cases}$$

eine Mengenfunktion $\boldsymbol{\lambda}^n : \mathcal{J}(\mathbb{R}^n) \to [0, \infty]$ definiert. Ziel dieses Abschnitts ist es zu zeigen, dass diese Mengenfunktion ein Maß ist und eine eindeutige Fortsetzung zu einem Maß auf der Borelschen σ–Algebra $\mathcal{B}(\mathbb{R}^n)$ besitzt.

Wir benötigen die folgende Definition:

Eine Menge $H \subseteq \mathbb{R}^n$ heißt *Halbraum*, wenn es ein $j \in \{1, \ldots, n\}$ und ein $z \in \mathbb{R}$ gibt mit

$$H = \Big\{ \mathbf{x} \in \mathbb{R}^n \;\Big|\; x_j \leq z \Big\}$$

Das folgende Lemma zeigt, dass je zwei disjunkte halboffene Intervalle durch einen Halbraum getrennt werden:

5.5.1 Lemma. *Für alle $J_1, J_2 \in \mathcal{J}(\mathbb{R}^n)$ mit $J_1 \cap J_2 = \emptyset$ gibt es einen Halbraum $H \subseteq \mathbb{R}^n$ derart, dass entweder $J_1 \cap H = J_1$ und $J_2 \cap H = \emptyset$ oder $J_2 \cap H = J_2$ und $J_1 \cap H = \emptyset$ gilt.*

Beweis. Im Fall $J_1 = \emptyset$ oder $J_2 = \emptyset$ ist nichts zu zeigen. Sei nun $J_1 \neq \emptyset$ und $J_2 \neq \emptyset$. Dann gibt es $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$ und $\mathbf{c} < \mathbf{d}$ sowie

$$\begin{aligned} J_1 &= (\mathbf{a}, \mathbf{b}] \\ J_2 &= (\mathbf{c}, \mathbf{d}] \end{aligned}$$

und wegen $J_1 \cap J_2 = \emptyset$ gibt es ein $j \in \{1, \ldots, n\}$ mit $b_j \leq c_j$ oder $d_j \leq a_j$. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $b_j \leq c_j$ und damit $[b_j, c_j] \neq \emptyset$ gilt. Wir wählen nun ein $z \in [b_j, c_j]$ und setzen

$$H := \Big\{ \mathbf{x} \in \mathbb{R}^n \;\Big|\; x_j \leq z \Big\}$$

Dann ist H ein Halbraum und es gilt $J_1 \cap H = J_1$ und $J_2 \cap H = \emptyset$. □

5.5.2 Lemma. *Für jedes Intervall $J \in \mathcal{J}(\mathbb{R}^n)$ und jeden Halbraum $H \subseteq \mathbb{R}^n$ gilt $J \cap H \in \mathcal{J}(\mathbb{R}^n)$ und $J \setminus H \in \mathcal{J}(\mathbb{R}^n)$ sowie*

$$\boldsymbol{\lambda}^n[J] = \boldsymbol{\lambda}^n[J \cap H] + \boldsymbol{\lambda}^n[J \setminus H]$$

Beweis. Im Fall $J \cap H = \emptyset$ oder $J \setminus H = \emptyset$ ist nichts zu zeigen. Sei nun $J \cap H \neq \emptyset$ und $J \setminus H \neq \emptyset$. Dann gibt es $\mathbf{a}, \mathbf{d} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{d}$ und

$$J = (\mathbf{a}, \mathbf{d}]$$

sowie ein $j \in \{1, \ldots, n\}$ und ein $z \in \mathbb{R}$ mit

$$H = \left\{\mathbf{x} \in \mathbb{R}^n \,\middle|\, x_j \leq z\right\}$$

Wegen $J \cap H \neq \emptyset$ und $J \setminus H \neq \emptyset$ gilt $a_j < z < d_j$. Wir definieren nun $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ durch

$$b_i := \begin{cases} z & \text{falls } i = j \\ d_i & \text{sonst} \end{cases}$$
$$c_i := \begin{cases} z & \text{falls } i = j \\ a_i & \text{sonst} \end{cases}$$

Dann gilt $J \cap H = (\mathbf{a}, \mathbf{b}] \in \mathcal{J}(\mathbb{R}^n)$ und $J \setminus H = (\mathbf{c}, \mathbf{d}] \in \mathcal{J}(\mathbb{R}^n)$ und damit

$$\begin{aligned}
\boldsymbol{\lambda}^n[J] &= \boldsymbol{\lambda}^n[(\mathbf{a}, \mathbf{d}]] \\
&= \prod_{i=1}^n (d_i - a_i) \\
&= (d_j - a_j) \prod_{i \in \{1,\ldots,n\} \setminus \{j\}} (d_i - a_i) \\
&= (z - a_j) \prod_{i \in \{1,\ldots,n\} \setminus \{j\}} (d_i - a_i) + (d_j - z) \prod_{i \in \{1,\ldots,n\} \setminus \{j\}} (d_i - a_i) \\
&= (b_j - a_j) \prod_{i \in \{1,\ldots,n\} \setminus \{j\}} (b_i - a_i) + (d_j - c_j) \prod_{i \in \{1,\ldots,n\} \setminus \{j\}} (d_i - c_i) \\
&= \prod_{i=1}^n (b_i - a_i) + \prod_{i=1}^n (d_i - c_i) \\
&= \boldsymbol{\lambda}^n[(\mathbf{a}, \mathbf{b}]] + \boldsymbol{\lambda}^n[(\mathbf{c}, \mathbf{d}]] \\
&= \boldsymbol{\lambda}^n[J \cap H] + \boldsymbol{\lambda}^n[J \setminus H]
\end{aligned}$$

Damit ist die Behauptung gezeigt. □

5.5.3 Lemma. *Die Mengenfunktion $\boldsymbol{\lambda}^n$ ist ein σ–endliches Maß.*

Beweis. Wir zeigen zunächst, dass $\boldsymbol{\lambda}^n$ ein Inhalt ist.

(i) Es gilt $\boldsymbol{\lambda}^n[\emptyset] = 0$.

(ii) Für alle $m \in \mathbb{N}$ und für jede disjunkte Familie $\{J_k\}_{k\in\{1,\dots,m\}} \subseteq \mathcal{J}(\mathbb{R}^n)$ mit $\sum_{k=1}^m J_k \in \mathcal{J}(\mathbb{R}^n)$ gilt

$$\boldsymbol{\lambda}^n\left[\sum_{k=1}^m J_k\right] = \sum_{k=1}^m \boldsymbol{\lambda}^n[J_k]$$

Wir führen den Beweis dieser Behauptung durch vollständige Induktion:

- $m = 1$: In diesem Fall ist nichts zu zeigen.
- $m \to m+1$: Wir nehmen an, die Behauptung sei für m bereits bewiesen, und betrachten eine disjunkte Familie $\{J_k\}_{k\in\{1,\dots,m+1\}} \subseteq \mathcal{J}(\mathbb{R}^n)$ mit $\sum_{k=1}^{m+1} J_k \in \mathcal{J}(\mathbb{R}^n)$. Nach Lemma 5.5.1 gibt es einen Halbraum $H \subseteq \mathbb{R}^n$ mit $J_m \cap H = J_m$ und $J_{m+1} \cap H = \emptyset$ oder $J_{m+1} \cap H = J_{m+1}$ und $J_m \cap H = \emptyset$. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $J_m \cap H = J_m$ und $J_{m+1} \cap H = \emptyset$ gilt. Dann gilt

$$\begin{aligned}\sum_{k=1}^{m+1} J_k \cap H &= \sum_{k=1}^{m-1} (J_k \cap H) + J_m \\ \sum_{k=1}^{m+1} J_k \setminus H &= \sum_{k=1}^{m-1} (J_k \setminus H) + J_{m+1}\end{aligned}$$

Aus Lemma 5.5.2 und der Voraussetzung folgt nun

$$\begin{aligned}&\boldsymbol{\lambda}^n\left[\sum_{k=1}^{m+1} J_k\right] \\ &= \boldsymbol{\lambda}^n\left[\sum_{k=1}^{m+1} J_k \cap H\right] + \boldsymbol{\lambda}^n\left[\sum_{k=1}^{m+1} J_k \setminus H\right] \\ &= \boldsymbol{\lambda}^n\left[\sum_{k=1}^{m-1} J_k \cap H + J_m\right] + \boldsymbol{\lambda}^n\left[\sum_{k=1}^{m-1} J_k \setminus H + J_{m+1}\right] \\ &= \sum_{k=1}^{m-1} \boldsymbol{\lambda}^n[J_k \cap H] + \boldsymbol{\lambda}^n[J_m] + \sum_{k=1}^{m-1} \boldsymbol{\lambda}^n[J_k \setminus H] + \boldsymbol{\lambda}^n[J_{m+1}] \\ &= \sum_{k=1}^{m-1} \Big(\boldsymbol{\lambda}^n[J_k \cap H] + \boldsymbol{\lambda}^n[J_k \setminus H]\Big) + \boldsymbol{\lambda}^n[J_m] + \boldsymbol{\lambda}^n[J_{m+1}] \\ &= \sum_{k=1}^{m+1} \boldsymbol{\lambda}^n[J_k]\end{aligned}$$

Damit ist gezeigt, dass $\boldsymbol{\lambda}^n$ ein Inhalt ist.

Wir zeigen nun, dass $\boldsymbol{\lambda}^n$ sogar ein Maß ist. Da $\mathcal{J}(\mathbb{R}^n)$ ein Halbring und $\boldsymbol{\lambda}^n$ ein Inhalt ist, genügt es nach Lemma 4.2.12 zu zeigen, dass $\boldsymbol{\lambda}^n$ σ–subadditiv ist. Sei $J \in \mathcal{J}(\mathbb{R}^n)$ und sei $\{J_k\}_{k\in\mathbb{N}} \subseteq \mathcal{J}(\mathbb{R}^n)$ eine Folge mit

$$J = \bigcup_{k\in\mathbb{N}} J_k$$

Wegen $\boldsymbol{\lambda}^n[\emptyset] = 0$ genügt es, den Fall zu betrachten, in dem $J \neq \emptyset$ und $J_k \neq \emptyset$ für alle $k \in \mathbb{N}$ gilt. Dann gibt es $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$ und

$$J = (\mathbf{a}, \mathbf{b}]$$

und für alle $k \in \mathbb{N}$ gibt es $\mathbf{a}_k, \mathbf{b}_k \in \mathbb{R}^n$ mit $\mathbf{a}_k < \mathbf{b}_k$ und

$$J_k = (\mathbf{a}_k, \mathbf{b}_k]$$

Sei $\varepsilon \in (0,\infty)$. Dann gibt es ein $\mathbf{a}^\varepsilon \in (\mathbf{a}, \mathbf{b})$ mit

$$\boldsymbol{\lambda}^n[(\mathbf{a}, \mathbf{b}]] \leq \boldsymbol{\lambda}^n[(\mathbf{a}^\varepsilon, \mathbf{b}]] + \varepsilon$$

und für jedes $k \in \mathbb{N}$ gibt es ein $\mathbf{b}_k^\varepsilon \in (\mathbf{b}_k, \boldsymbol{\infty})$ mit

$$\boldsymbol{\lambda}^n[(\mathbf{a}_k, \mathbf{b}_k^\varepsilon]] \leq \boldsymbol{\lambda}^n[(\mathbf{a}_k, \mathbf{b}_k]] + \varepsilon/2^k$$

Es gilt

$$[\mathbf{a}^\varepsilon, \mathbf{b}] \subseteq (\mathbf{a}, \mathbf{b}] = J = \bigcup_{k\in\mathbb{N}} J_k = \bigcup_{k\in\mathbb{N}} (\mathbf{a}_k, \mathbf{b}_k] \subseteq \bigcup_{k\in\mathbb{N}} (\mathbf{a}_k, \mathbf{b}_k^\varepsilon)$$

Da einerseits das Intervall $[\mathbf{a}^\varepsilon, \mathbf{b}]$ abgeschlossen und beschränkt und damit nach dem Satz von Heine/Borel kompakt ist und andererseits jedes der Intervalle $(\mathbf{a}_k, \mathbf{b}_k^\varepsilon)$ offen ist, gibt es eine endliche Menge $K \subseteq \mathbb{N}$ mit

$$[\mathbf{a}^\varepsilon, \mathbf{b}] \subseteq \bigcup_{k\in K} (\mathbf{a}_k, \mathbf{b}_k^\varepsilon)$$

Daher gilt $(\mathbf{a}^\varepsilon, \mathbf{b}] \subseteq [\mathbf{a}^\varepsilon, \mathbf{b}] \subseteq \bigcup_{k\in K} (\mathbf{a}_k, \mathbf{b}_k^\varepsilon) \subseteq \bigcup_{k\in K} (\mathbf{a}_k, \mathbf{b}_k^\varepsilon]$ und damit

$$(\mathbf{a}^\varepsilon, \mathbf{b}] = \bigcup_{k\in K} (\mathbf{a}_k, \mathbf{b}_k^\varepsilon] \cap (\mathbf{a}^\varepsilon, \mathbf{b}]$$

Da $\boldsymbol{\lambda}^n$ ein Inhalt und damit endlich subadditiv und monoton ist, ergibt sich nun

$$\begin{aligned} \boldsymbol{\lambda}^n[J] - \varepsilon &= \boldsymbol{\lambda}^n[(\mathbf{a}, \mathbf{b}]] - \varepsilon \\ &\leq \boldsymbol{\lambda}^n[(\mathbf{a}^\varepsilon, \mathbf{b}]] \\ &= \boldsymbol{\lambda}^n\left[\bigcup_{k\in K} (\mathbf{a}_k, \mathbf{b}_k^\varepsilon] \cap (\mathbf{a}^\varepsilon, \mathbf{b}]\right] \end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{k \in K} \boldsymbol{\lambda}^n \Big[(\mathbf{a}_k, \mathbf{b}_k^\varepsilon] \cap (\mathbf{a}^\varepsilon, \mathbf{b}] \Big] \\
&\leq \sum_{k \in K} \boldsymbol{\lambda}^n [(\mathbf{a}_k, \mathbf{b}_k^\varepsilon]] \\
&\leq \sum_{k \in K} \Big(\boldsymbol{\lambda}^n [(\mathbf{a}_k, \mathbf{b}_k]] + \varepsilon/2^k \Big) \\
&\leq \sum_{k=1}^\infty \boldsymbol{\lambda}^n [(\mathbf{a}_k, \mathbf{b}_k]] + \varepsilon \\
&= \sum_{k=1}^\infty \boldsymbol{\lambda}^n [J_k] + \varepsilon
\end{aligned}
$$

Da $\varepsilon \in (0, \infty)$ beliebig war, folgt daraus

$$\boldsymbol{\lambda}^n [J] \leq \sum_{k=1}^\infty \boldsymbol{\lambda}^n [J_k]$$

Damit ist gezeigt, dass der Inhalt $\boldsymbol{\lambda}^n$ σ–subadditiv und damit ein Maß ist. Wir zeigen abschließend, dass $\boldsymbol{\lambda}^n$ σ–endlich ist. Es gilt $\bigcup_{k \in \mathbb{N}} (-k\mathbf{1}, k\mathbf{1}] = \mathbb{R}^n$ und für alle $k \in \mathbb{N}$ gilt $(-k\mathbf{1}, k\mathbf{1}] \in \mathcal{J}(\mathbb{R}^n)$ und $\boldsymbol{\lambda}^n[(-k\mathbf{1}, k\mathbf{1}]] = (2k)^n < \infty$. Damit ist gezeigt, dass $\boldsymbol{\lambda}^n$ σ–endlich ist. □

5.5.4 Satz. *Die Mengenfunktion $\boldsymbol{\lambda}^n$ besitzt eine eindeutige Fortsetzung zu einem Maß auf $\mathcal{B}(\mathbb{R}^n)$ und die Fortsetzung ist σ–endlich.*

Beweis. Nach Lemma 5.5.3 ist die Mengenfunktion $\boldsymbol{\lambda}^n : \mathcal{J}(\mathbb{R}^n) \to [0, \infty]$ ein σ–endliches Maß. Da das Mengensystem $\mathcal{J}(\mathbb{R}^n)$ ein Halbring ist, folgt die Behauptung aus dem Satz von Caratheodory. □

Das nach dem letzten Satz eindeutig bestimmte Maß $\mathcal{B}(\mathbb{R}^n) \to [0, \infty]$, das die Mengenfunktion $\boldsymbol{\lambda}^n : \mathcal{J}(\mathbb{R}^n) \to [0, \infty]$ fortsetzt, heißt (*n–dimensionales*) *Lebesgue–Maß* und wird wieder mit

$$\boldsymbol{\lambda}^n$$

bezeichnet. Im Fall $n = 1$ schreiben wir

$$\boldsymbol{\lambda}$$

anstelle von $\boldsymbol{\lambda}^1$. Das Lebesgue–Maß ist σ–endlich, aber es ist nicht endlich, denn es gilt $\boldsymbol{\lambda}^n[\mathbb{R}^n] = \infty$.

5.5.5 Folgerung. *Für alle $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ und für jede Menge $B \in \mathcal{B}(\mathbb{R}^n)$ mit $(\mathbf{a}, \mathbf{b}) \subseteq B \subseteq [\mathbf{a}, \mathbf{b}]$ gilt $\boldsymbol{\lambda}^n[B] = \boldsymbol{\lambda}^n[(\mathbf{a}, \mathbf{b}]]$.*

Beweis. Im Fall $(\mathbf{a},\mathbf{b}) = \emptyset$ gilt $(\mathbf{a},\mathbf{b}] = \emptyset$ und damit

$$\boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b})] = 0 = \boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b}]]$$

Im Fall $(\mathbf{a},\mathbf{b}) \neq \emptyset$ gibt es eine monoton wachsende Folge $\{\mathbf{b}_k\}_{k\in\mathbb{N}} \subseteq (\mathbf{a},\mathbf{b})$ mit $\mathbf{b} = \sup_{k\in\mathbb{N}} \mathbf{b}_k$. Dann ist $\{(\mathbf{a},\mathbf{b}_k]\}_{k\in\mathbb{N}}$ eine monoton wachsende Folge in $\mathcal{B}(\mathbb{R}^n)$ mit $(\mathbf{a},\mathbf{b}) = \bigcup_{k\in\mathbb{N}}(\mathbf{a},\mathbf{b}_k]$ und es gilt

$$\boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b})] = \sup_{k\in\mathbb{N}} \boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b}_k]] = \sup_{k\in\mathbb{N}} \prod_{i=1}^n (b_{ki} - a_i) = \prod_{i=1}^n (b_i - a_i) = \boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b}]]$$

In beiden Fällen gibt es eine monoton wachsende Folge $\{\mathbf{a}_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}$ mit $\mathbf{a} = \sup_{k\in\mathbb{N}} \mathbf{a}_k$. Dann ist $\{(\mathbf{a}_k,\mathbf{b}]\}_{k\in\mathbb{N}}$ eine monoton fallende Folge in $\mathcal{B}(\mathbb{R}^n)$ mit $[\mathbf{a},\mathbf{b}] = \bigcap_{k\in\mathbb{N}}(\mathbf{a}_k,\mathbf{b}]$ und wegen $\boldsymbol{\lambda}^n[(\mathbf{a}_1,\mathbf{b}]] < \infty$ gilt

$$\boldsymbol{\lambda}^n[[\mathbf{a},\mathbf{b}]] = \inf_{k\in\mathbb{N}} \boldsymbol{\lambda}^n[(\mathbf{a}_k,\mathbf{b}]] = \inf_{k\in\mathbb{N}} \prod_{i=1}^n (b_i - a_{ki}) = \prod_{i=1}^n (b_i - a_i) = \boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b}]]$$

Daher gilt

$$\boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b})] = \boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b}]] = \boldsymbol{\lambda}^n[[\mathbf{a},\mathbf{b}]]$$

und die Behauptung folgt aus der Monotonie von $\boldsymbol{\lambda}^n$. □

Die folgenden Ergebnisse sind nun evident:

5.5.6 Folgerung. *Für alle* $\mathbf{x} \in \mathbb{R}^n$ *gilt* $\boldsymbol{\lambda}^n[\{\mathbf{x}\}] = 0$.

5.5.7 Folgerung. *Für jede abzählbare Menge* $B \subseteq \mathbb{R}^n$ *gilt* $\boldsymbol{\lambda}^n[B] = 0$.

Wir geben abschließend ein Beispiel für eine $\boldsymbol{\lambda}$–Nullmenge, die nicht abzählbar ist:

5.5.8 Beispiel (Cantor–Menge). Sei $\mathcal{Z}$ das System aller Teilmengen von $[0,1]$, die als Vereinigung von endlich vielen disjunkten abgeschlossenen (und nichtleeren) Intervallen dargestellt werden können, und sei $\Psi : \mathcal{Z} \to \mathcal{Z}$ gegeben durch

$$\Psi\left(\sum_{i=1}^m [a_i,b_i]\right) := \sum_{i=1}^m \left(\left[a_i, \frac{2a_i+b_i}{3}\right] + \left[\frac{a_i+2b_i}{3}, b_i\right]\right)$$

Dann gilt $\mathcal{Z} \subseteq \mathcal{B}(\mathbb{R})$ und die Abbildung Ψ ist wohldefiniert. Für alle $A \in \mathcal{Z}$ gilt $\Psi(A) \subseteq A$ und

$$\boldsymbol{\lambda}[\Psi(A)] = \frac{2}{3}\,\boldsymbol{\lambda}[A]$$

Für alle $n \in \mathbb{N}$ sei

$$C_n := \Psi^n([0,1])$$

Dann ist $\{C_n\}_{n\in\mathbb{N}}$ eine monoton fallende Folge in $\mathcal{B}(\mathbb{R})$ und es gilt

$$\boldsymbol{\lambda}[C_n] = \left(\frac{2}{3}\right)^n$$

Sei

$$C := \bigcap_{n\in\mathbb{N}} C_n$$

Dann gilt $C \in \mathcal{B}(\mathbb{R})$ sowie $C \neq \emptyset$ und

$$\boldsymbol{\lambda}[C] = 0$$

Die Menge C heißt *Cantor–Menge*. Wir zeigen in Beispiel 12.1.14, dass die Cantor–Menge nicht abzählbar ist.

Aufgaben

5.5.A **Vervollständigung:** Das n–dimensionale Lebesgue–Maß besitzt eine eindeutige Fortsetzung zu einem vollständigen Maß auf $\mathcal{F}((\boldsymbol{\lambda}^n)^*)$. Die Fortsetzung ist σ–endlich.

5.5.B **Cantor–Menge:**. Die Cantor–Menge ist kompakt.

6

Transformation von Maßen

In diesem Kapitel behandeln wir die Transformation von Maßen unter messbaren Abbildungen und insbesondere die Transformation des Lebesgue–Maßes unter einer affinen Abbildung im Euklidischen Raum.

Als erstes zeigen wir, dass eine messbare Abbildung jedem Maß auf ihrem Definitionsbereich ein Maß auf ihrem Bildbereich zuordnet; dies ist ein grundlegendes Ergebnis, das gleichzeitig die Bedeutung der Messbarkeit beleuchtet (Abschnitt 6.1). Als nächstes betrachten wir die Translationsinvarianz von Maßen auf der Borelschen σ–Algebra des Euklidischen Raumes und erhalten einerseits eine Charakterisierung des Lebesgue–Maßes unter den translationsinvarianten Maßen und andererseits die Aussage, dass das Lebesgue–Maß nicht zu einem translationsinvarianten Maß auf der Potenzmenge des Euklidischen Raumes fortgesetzt werden kann (Abschnitt 6.2). Abschließend bestimmen wir das Bild des Lebesgue–Maßes unter einer linearen oder affinen Abbildung des Euklidischen Raumes (Abschnitt 6.3).

6.1 Bildmaße

Sind $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ Messräume und ist $f : \Omega \to \Omega'$ messbar, so ist wegen $f^{-1}(\mathcal{F}') \subseteq \mathcal{F}$ für jede Mengenfunktion $\mu : \mathcal{F} \to [0, \infty]$ die Mengenfunktion $\mu_f : \mathcal{F}' \to [0, \infty]$ mit

$$\mu_f[A'] := \mu[f^{-1}(A')]$$

wohldefiniert. Ist μ ein Maß, so heißt μ_f das *Bildmaß* von μ unter f; diese Bezeichnung wird durch das folgende Lemma gerechtfertigt:

6.1.1 Lemma. *Seien $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ Messräume und sei $f : \Omega \to \Omega'$ messbar. Dann ist für jedes Maß $\mu : \mathcal{F} \to [0, \infty]$ auch μ_f ein Maß.*

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_6, © Springer-Verlag Berlin Heidelberg 2011

Beweis. Es gilt $\mu_f[\emptyset] = \mu[f^{-1}(\emptyset)] = \mu[\emptyset] = 0$ und für jede disjunkte Folge $\{A'_k\}_{k\in\mathbb{N}} \subseteq \mathcal{F}'$ gilt

$$\begin{aligned}\mu_f\left[\sum_{k=1}^{\infty} A'_k\right] &= \mu\left[f^{-1}\left(\sum_{k=1}^{\infty} A'_k\right)\right]\\ &= \mu\left[\sum_{k=1}^{\infty} f^{-1}(A'_k)\right]\\ &= \sum_{k=1}^{\infty} \mu[f^{-1}(A'_k)]\\ &= \sum_{k=1}^{\infty} \mu_f[A'_k]\end{aligned}$$

Daher ist μ_f ein Maß. □

Lemma 6.1.1 zeigt, dass die Messbarkeit einer Abbildung zwischen zwei Messräumen es der Abbildung ermöglicht, ein Maß von ihrem Definitionsbereich in ihren Bildbereich zu transportieren. Das folgende Lemma zeigt, dass das Bildmaß einer Komposition von messbaren Abbildungen iterativ bestimmt werden kann:

6.1.2 Lemma. *Seien $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ und $(\Omega'', \mathcal{F}'')$ Messräume und seien $f : \Omega \to \Omega'$ und $g : \Omega' \to \Omega''$ messbar. Dann gilt für jedes Maß $\mu : \mathcal{F} \to [0, \infty]$*

$$\mu_{(g\circ f)} = (\mu_f)_g$$

Beweis. Mit f und g ist auch $g \circ f$ messbar und für alle $A'' \in \mathcal{F}''$ gilt

$$\begin{aligned}\mu_{(g\circ f)}[A''] &= \mu[(g \circ f)^{-1}(A'')]\\ &= \mu[f^{-1}(g^{-1}(A''))]\\ &= \mu_f[g^{-1}(A'')]\\ &= (\mu_f)_g[A'']\end{aligned}$$

Damit ist die Behauptung gezeigt. □

6.2 Translationsinvariante Maße auf $\boldsymbol{\mathcal{B}}(\mathbb{R}^{\boldsymbol{n}})$

Für $\mathbf{c} \in \mathbb{R}^n$ heißt die Abbildung $T_{\mathbf{c}} : \mathbb{R}^n \to \mathbb{R}^n$ mit

$$T_{\mathbf{c}}(\mathbf{x}) := \mathbf{c} + \mathbf{x}$$

Translation bezüglich $\mathbf{c}$. Jede Translation ist bijektiv mit $T_{\mathbf{c}}^{-1} = T_{-\mathbf{c}}$ und sie ist stetig und damit messbar.

Ein Maß $\mu : \mathcal{B}(\mathbb{R}^n) \to [0,\infty]$ heißt *translationsinvariant*, wenn für alle $\mathbf{c} \in \mathbb{R}^n$

$$\mu_{T_\mathbf{c}} = \mu$$

gilt.

6.2.1 Lemma. *Für ein Maß $\mu : \mathcal{B}(\mathbb{R}^n) \to [0,\infty]$ sind folgende Aussagen äquivalent*:
(a) *μ ist translationsinvariant.*
(b) *Für alle $\mathbf{c} \in \mathbb{R}^n$ und $B \in \mathcal{B}(\mathbb{R}^n)$ gilt $\mu[\mathbf{c}+B] = \mu[B]$.*

Beweis. Sei μ translationsinvariant. Dann gilt für alle $\mathbf{c} \in \mathbb{R}^n$ und $B \in \mathcal{B}(\mathbb{R}^n)$

$$\mu[\mathbf{c}+B] = \mu[T_\mathbf{c}(B)] = \mu[T_{-\mathbf{c}}^{-1}(B)] = \mu_{T_{-\mathbf{c}}}[B] = \mu[B]$$

Gilt andererseits für alle $\mathbf{c} \in \mathbb{R}^n$ und $B \in \mathcal{B}(\mathbb{R}^n)$ die Gleichung $\mu[\mathbf{c}+B] = \mu[B]$, so erhält man

$$\mu_{T_\mathbf{c}}[B] = \mu[T_\mathbf{c}^{-1}(B)] = \mu[T_{-\mathbf{c}}(B)] = \mu[-\mathbf{c}+B] = \mu[B]$$

und damit $\mu_{T_\mathbf{c}} = \mu$ für alle $\mathbf{c} \in \mathbb{R}^n$. □

Wir wenden uns nun dem Lebesgue–Maß zu:

6.2.2 Satz. *Das Lebesgue–Maß ist translationsinvariant.*

Beweis. Das Mengensystem $\mathcal{J}(\mathbb{R}^n)$ ist ein $\cap$–stabiler Erzeuger der Borelschen σ–Algebra $\mathcal{B}(\mathbb{R}^n)$ und die Restriktion $\boldsymbol{\lambda}^n|_{\mathcal{J}(\mathbb{R}^n)}$ des Lebesgue–Maßes auf $\mathcal{J}(\mathbb{R}^n)$ ist σ–endlich. Wir zeigen nun, dass für alle $\mathbf{c} \in \mathbb{R}^n$

$$\boldsymbol{\lambda}^n_{T_\mathbf{c}}|_{\mathcal{J}(\mathbb{R}^n)} = \boldsymbol{\lambda}^n|_{\mathcal{J}(\mathbb{R}^n)}$$

gilt. Aus dem Eindeutigkeitssatz folgt dann $\boldsymbol{\lambda}^n_{T_\mathbf{c}} = \boldsymbol{\lambda}^n$.
Sei also $\mathbf{c} \in \mathbb{R}^n$. Dann gilt $\boldsymbol{\lambda}^n_{T_\mathbf{c}}[\emptyset] = 0 = \boldsymbol{\lambda}^n[\emptyset]$, und für alle $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$ gilt

$$\begin{aligned}
\boldsymbol{\lambda}^n_{T_\mathbf{c}}[(\mathbf{a},\mathbf{b}]] &= \boldsymbol{\lambda}^n[T_\mathbf{c}^{-1}((\mathbf{a},\mathbf{b}])] \\
&= \boldsymbol{\lambda}^n[(\mathbf{a}-\mathbf{c},\mathbf{b}-\mathbf{c}]] \\
&= \prod_{i=1}^n \Big((b_i-c_i) - (a_i-c_i)\Big) \\
&= \prod_{i=1}^n (b_i-a_i) \\
&= \boldsymbol{\lambda}^n[(\mathbf{a},\mathbf{b}]]
\end{aligned}$$

Damit ist die Behauptung gezeigt. □

Wir zeigen nun, dass das Lebesgue–Maß unter allen translationsinvarianten Maßen auf $\mathcal{B}(\mathbb{R}^n)$ durch eine einfache Normierungsbedingung ausgezeichnet ist.

6.2.3 Lemma. *Sei $\mu : \mathcal{B}(\mathbb{R}^n) \to [0,\infty]$ ein translationsinvariantes Maß mit $\mu[(\mathbf{0},\mathbf{1}]] < \infty$. Dann gilt*

$$\mu = \mu[(\mathbf{0},\mathbf{1}]] \cdot \boldsymbol{\lambda}^n$$

Beweis. Für $m \in \mathbb{N}$ betrachten wir die Familie

$$\mathcal{W}_m := \left\{ \prod_{i=1}^n \left(\frac{k_i - 1}{m}, \frac{k_i}{m} \right] \,\middle|\, k_1, \dots, k_n \in \{1, \dots, m\} \right\}$$

Da μ translationsinvariant ist und jeder der m^n Würfel der Familie $\mathcal{W}_m$ durch Translation aus jedem anderen Würfel der Familie $\mathcal{W}_m$ gewonnen werden kann, gibt es ein $c_m \in [0,\infty]$ derart, dass für alle $W \in \mathcal{W}_m$

$$\mu[W] = c_m$$

gilt. Andererseits gilt für alle $W \in \mathcal{W}_m$

$$\boldsymbol{\lambda}^n[W] = \frac{1}{m^n}$$

Da μ ein Maß ist und die disjunkte Vereinigung aller Würfel aus $\mathcal{W}_m$ mit dem Intervall $(\mathbf{0},\mathbf{1}]$ übereinstimmt, erhält man

$$m^n \, c_m = \sum_{W \in \mathcal{W}_m} \mu[W] = \mu\left[\sum_{W \in \mathcal{W}_m} W \right] = \mu[(\mathbf{0},\mathbf{1}]] = c_1$$

Daher gilt für alle $W \in \mathcal{W}_m$

$$\mu[W] = c_m = c_1 \, \frac{1}{m^n} = c_1 \, \boldsymbol{\lambda}^n[W]$$

Für $m \in \mathbb{N}$ betrachten wir nun die Familie

$$\mathcal{V}_m := \left\{ (\mathbf{a},\mathbf{b}] \in \mathcal{J}(\mathbb{R}^n) \,\middle|\, b_i - a_i = \frac{1}{m} \text{ für alle } i \in \{1, \dots, n\} \right\}$$

Unter nochmaliger Verwendung der Translationsinvarianz von μ erhalten wir für alle $V \in \mathcal{V}_m$

$$\mu[V] = c_1 \, \boldsymbol{\lambda}^n[V]$$

Da sich jede Menge aus $\mathcal{J}(\mathbb{Q}^n)$ als Vereinigung einer endlichen disjunkten Familie von Mengen aus $\mathcal{V}_m$ mit hinreichend großem $m \in \mathbb{N}$ darstellen lässt, gilt auch für alle $J \in \mathcal{J}(\mathbb{Q}^n)$

$$\mu[J] = c_1 \,\boldsymbol{\lambda}^n[J]$$

Es gilt also

$$\mu|_{\mathcal{J}(\mathbb{Q}^n)} = c_1 \,\boldsymbol{\lambda}^n|_{\mathcal{J}(\mathbb{Q}^n)}$$

Außerdem ist $\boldsymbol{\lambda}^n|_{\mathcal{J}(\mathbb{Q}^n)}$ σ–endlich und das Mengensystem $\mathcal{J}(\mathbb{Q}^n)$ ist ein Halbring mit $\sigma(\mathcal{J}(\mathbb{Q}^n)) = \mathcal{B}(\mathbb{R}^n)$; vgl. Aufgabe 1.5.D. Aus dem Eindeutigkeitssatz folgt nun

$$\mu|_{\mathcal{B}(\mathbb{R}^n)} = c_1 \,\boldsymbol{\lambda}^n|_{\mathcal{B}(\mathbb{R}^n)}$$

Damit ist das Lemma bewiesen. □

Ein analoges Ergebnis erhält man, wenn man den halboffenen Würfel $(\mathbf{0}, \mathbf{1}]$ durch den abgeschlossenen Würfel $[\mathbf{0}, \mathbf{1}]$ ersetzt:

6.2.4 Folgerung. *Sei $\mu : \mathcal{B}(\mathbb{R}^n) \to [0, \infty]$ ein translationsinvariantes Maß mit $\mu[[\mathbf{0}, \mathbf{1}]] < \infty$. Dann gilt*

$$\mu = \mu[[\mathbf{0}, \mathbf{1}]] \,\boldsymbol{\lambda}^n$$

Beweis. Wegen $\mu[(\mathbf{0}, \mathbf{1}]] \le \mu[[\mathbf{0}, \mathbf{1}]] < \infty$ folgt aus Lemma 6.2.3

$$\mu = \mu[(\mathbf{0}, \mathbf{1}]] \,\boldsymbol{\lambda}^n$$

Aus Folgerung 5.5.5 ergibt sich ferner $\boldsymbol{\lambda}^n[[\mathbf{0}, \mathbf{1}]] = \boldsymbol{\lambda}^n[(\mathbf{0}, \mathbf{1}]] = 1$. Dann gilt aber $\mu[[\mathbf{0}, \mathbf{1}]] = \mu[(\mathbf{0}, \mathbf{1}]] \cdot \boldsymbol{\lambda}^n[[\mathbf{0}, \mathbf{1}]] = \mu[(\mathbf{0}, \mathbf{1}]]$ und damit

$$\mu = \mu[[\mathbf{0}, \mathbf{1}]] \,\boldsymbol{\lambda}^n$$

Damit ist die Folgerung bewiesen. □

Damit gelangen wir zu der angekündigten Charakterisierung des Lebesgue–Maßes:

6.2.5 Satz. *Das Lebesgue–Maß ist das einzige translationsinvariante Maß $\mu : \mathcal{B}(\mathbb{R}^n) \to [0, \infty]$ mit $\mu[[\mathbf{0}, \mathbf{1}]] = 1$.*

Wir zeigen abschließend, dass das Lebesgue–Maß keine translationsinvariante Fortsetzung auf der Potenzmenge des $\mathbb{R}^n$ besitzt:

6.2.6 Satz. *Das Lebesgue–Maß besitzt keine Fortsetzung zu einem translationsinvarianten Maß auf $2^{\mathbb{R}^n}$.*

Beweis. Für $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ sei $\mathbf{x} \sim \mathbf{y}$ falls $\mathbf{x} - \mathbf{y} \in \mathbb{Q}^n$. Dann ist $\sim$ eine Äquivalenzrelation auf $\mathbb{R}^n$ und wir bezeichnen die Menge aller Äquivalenzklassen von $\mathbb{R}^n$

bezüglich $\sim$ mit $\mathbb{R}^n/_\sim$. Da jede Äquivalenzklasse in $\mathbb{R}^n/_\sim$ ein Element in $(\mathbf{0},\mathbf{1}]$ enthält, bilden die Durchschnitte der Äquivalenzklassen in $\mathbb{R}^n/_\sim$ mit $(\mathbf{0},\mathbf{1}]$ eine disjunkte Familie von nichtleeren Teilmengen von $(\mathbf{0},\mathbf{1}]$, deren Vereinigung gleich $(\mathbf{0},\mathbf{1}]$ ist.
Nach dem Auswahlaxiom gibt es eine Menge $Z \subseteq (\mathbf{0},\mathbf{1}]$ derart, dass Z aus jeder Äquivalenzklasse von $\mathbb{R}^n/_\sim$ (bzw. aus ihrem Durchschnitt mit $(\mathbf{0},\mathbf{1}]$) genau ein Element enthält. Dann gilt

$$\mathbb{R}^n = \sum_{\mathbf{z}\in Z}(\mathbf{z}+\mathbb{Q}^n) = \sum_{\mathbf{z}\in Z}\sum_{\mathbf{q}\in\mathbb{Q}^n}\{\mathbf{z}+\mathbf{q}\} = \sum_{\mathbf{q}\in\mathbb{Q}^n}\sum_{\mathbf{z}\in Z}\{\mathbf{q}+\mathbf{z}\} = \sum_{\mathbf{q}\in\mathbb{Q}^n}(\mathbf{q}+Z)$$

und wegen $Z \subseteq (\mathbf{0},\mathbf{1}]$ gilt

$$\sum_{\mathbf{q}\in\mathbb{Q}^n\cap\,(\mathbf{0},\mathbf{1}]}(\mathbf{q}+Z) \subseteq (0,2]^n$$

Wir nehmen nun an, dass es ein translationsinvariantes Maß $\mu : 2^{\mathbb{R}^n} \to [0,\infty]$ mit $\mu|_{\mathcal{B}(\mathbb{R}^n)} = \boldsymbol{\lambda}^n$ gibt. Dann gilt für alle $\mathbf{q}\in\mathbb{Q}^n$

$$\mu[\mathbf{q}+Z] = \mu[Z]$$

Aus

$$\sum_{\mathbf{q}\in\mathbb{Q}^n\cap\,(\mathbf{0},\mathbf{1}]}\mu[Z] = \sum_{\mathbf{q}\in\mathbb{Q}^n\cap\,(\mathbf{0},\mathbf{1}]}\mu[\mathbf{q}+Z] \le \mu[(0,2]^n] = \boldsymbol{\lambda}^n[(0,2]^n] < \infty$$

folgt $\mu[Z] = 0$ und damit

$$\boldsymbol{\lambda}^n[\mathbb{R}^n] = \mu[\mathbb{R}^n] = \sum_{\mathbf{q}\in\mathbb{Q}^n}\mu[\mathbf{q}+Z] = \sum_{\mathbf{q}\in\mathbb{Q}^n}\mu[Z] = 0$$

Dies ist ein Widerspruch. □

Da das Lebesgue–Maß nach Satz 6.2.2 ein translationsinvariantes Maß auf der Borelschen σ–Algebra $\mathcal{B}(\mathbb{R}^n)$ ist, ergibt sich aus dem letzten Satz die folgende offensichtliche Folgerung:

6.2.7 Folgerung. *Es gilt* $\mathcal{B}(\mathbb{R}^n) \neq 2^{\mathbb{R}^n}$.

Die letzten Ergebnisse zeigen, dass es nicht immer möglich ist, ein Maß mit gewissen zusätzlichen Eigenschaften auf der Potenzmenge der Grundmenge zu definieren. Dies ist einer der Gründe dafür, Maße zu betrachten, die auf einer σ–Algebra oder einem allgemeineren Mengensystem definiert sind.

Aufgabe

6.2.A **Hyperebenen:** Eine Menge $H \subseteq \mathbb{R}^n$ heißt *Hyperebene* des $\mathbb{R}^n$, wenn es einen Unterraum $E \subseteq \mathbb{R}^n$ und einen Vektor $\mathbf{d}\in\mathbb{R}^n$ gibt mit $\dim(E) \le n-1$ und $H = \mathbf{d}+E$. Für jede Hyperebene $H \subseteq \mathbb{R}^n$ gilt $\boldsymbol{\lambda}^n[H] = 0$.

6.3 Lineare Abbildungen des Lebesgue–Maßes

Das folgende Ergebnis zeigt, dass das Bildmaß des Lebesgue–Maßes $\boldsymbol{\lambda}^n$ unter einer invertierbaren linearen Abbildung $\mathbb{R}^n \to \mathbb{R}^n$ sich nur um einen strikt positiven Faktor vom Lebesgue–Maß $\boldsymbol{\lambda}^n$ unterscheidet.

6.3.1 Satz. *Sei $T : \mathbb{R}^n \to \mathbb{R}^n$ gegeben durch*

$$T(\mathbf{x}) := \mathbf{D}\mathbf{x}$$

mit einer invertierbaren Matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$. Dann gilt

$$\boldsymbol{\lambda}^n_T = \frac{1}{|\det(\mathbf{D})|}\,\boldsymbol{\lambda}^n$$

Beweis. Wir zeigen, dass das Maß $\boldsymbol{\lambda}^n_T$ translationsinvariant ist mit

$$\boldsymbol{\lambda}^n_T[(\mathbf{0},\mathbf{1}]] = \frac{1}{|\det(\mathbf{D})|}$$

Aus Lemma 6.2.3 folgt dann die Behauptung des Satzes.
(1) *Das Maß $\boldsymbol{\lambda}^n_T$ ist translationsinvariant*:
Da $\boldsymbol{\lambda}^n$ translationsinvariant ist, gilt nach Lemma 6.2.1 für alle $\mathbf{c} \in \mathbb{R}^n$ und $B \in \mathcal{B}(\mathbb{R}^n)$

$$\begin{aligned}
\boldsymbol{\lambda}^n_T[\mathbf{c}+B] &= \boldsymbol{\lambda}^n[T^{-1}(\mathbf{c}+B)] \\
&= \boldsymbol{\lambda}^n[T^{-1}(\mathbf{c})+T^{-1}(B)] \\
&= \boldsymbol{\lambda}^n[T^{-1}(B)] \\
&= \boldsymbol{\lambda}^n_T[B]
\end{aligned}$$

und aus demselben Lemma folgt nun, dass $\boldsymbol{\lambda}^n_T$ translationsinvariant ist.
(2) *Es gilt $\boldsymbol{\lambda}^n_T[(\mathbf{0},\mathbf{1}]] = |\det(\mathbf{D})|^{-1}$*:
Sei $n \geq 3$ und für $k \in \mathbb{N}$ sei

$$W^k := (0,1]^k$$

Wir zeigen, dass für jede invertierbare Matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$

$$\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] = \frac{1}{|\det(\mathbf{D})|}$$

gilt. Wir betrachten zunächst drei Spezialfälle:

– Sei $\mathbf{D}$ eine Permutationsmatrix. Dann gilt einerseits $|\det(\mathbf{D})| = 1$ und andererseits $\mathbf{D}^{-1}(W^n) = W^n$ und damit $\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] = \boldsymbol{\lambda}^n[W^n] = 1$. Daher gilt

$$\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] = \frac{1}{|\det(\mathbf{D})|}$$

– Sei $\mathbf{D}$ eine Elementarmatrix der Form

$$\mathbf{D} = \begin{pmatrix} \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$$

mit $c \in \mathbb{R} \setminus \{0\}$. Dann gilt $|\det(\mathbf{D})| = |c| \neq 0$ und wegen

$$\mathbf{D}^{-1} = \begin{pmatrix} \begin{pmatrix} c^{-1} & 0 \\ 0 & 1 \end{pmatrix} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$$

gilt

$$\mathbf{D}^{-1}(W^n) = \begin{cases} (0, c^{-1}] \times W^{n-1} & \text{falls } c > 0 \\ [c^{-1}, 0) \times W^{n-1} & \text{falls } c < 0 \end{cases}$$

und damit $\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] = |c^{-1}| = |c|^{-1}$. Daher gilt

$$\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] = \frac{1}{|\det(\mathbf{D})|}$$

– Wir betrachten abschließend die Elementarmatrix

$$\mathbf{D} = \begin{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$$

Dann gilt $|\det(\mathbf{D})| = 1$ und

$$\mathbf{D}^{-1} = \begin{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$$

Zur Bestimmung von $\mathbf{D}^{-1}(W^n)$ betrachten wir die Matrix

$$\mathbf{D}_2 := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

und ihre Inverse

$$\mathbf{D}_2^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

Dann gilt

$$\mathbf{D}^{-1}(W^n) = \mathbf{D}^{-1}(W^2 \times W^{n-2}) = \mathbf{D}_2^{-1}(W^2) \times W^{n-2}$$

Zur Bestimmung von $\mathbf{D}_2^{-1}(W^2)$ betrachten wir die Vektoren

$$\mathbf{a}_1 := \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
$$\mathbf{a}_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$\mathbf{a}_3 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$\mathbf{a}_4 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\mathbf{a}_5 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

und die Mengen

$$A_1 := \Big(\mathrm{cx}(\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4\}) \setminus \mathrm{cx}(\{\mathbf{a}_1, \mathbf{a}_4\})\Big)$$
$$A_2 := \Big(\mathrm{cx}(\{\mathbf{a}_2, \mathbf{a}_4, \mathbf{a}_5\}) \setminus \Big(\mathrm{cx}(\{\mathbf{a}_2, \mathbf{a}_4\}) \cup \mathrm{cx}(\{\mathbf{a}_4, \mathbf{a}_5\})\Big)\Big)$$
$$A_3 := \Big(\mathrm{cx}(\{\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_5\}) \setminus \mathrm{cx}(\{\mathbf{a}_2, \mathbf{a}_5\})\Big)$$
$$A_4 := \Big(\mathrm{cx}(\{\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}) \setminus \mathrm{cx}(\{\mathbf{a}_2, \mathbf{a}_4\})\Big)$$
$$A_5 := \Big(\mathrm{cx}(\{\mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5\}) \setminus \Big(\mathrm{cx}(\{\mathbf{a}_3, \mathbf{a}_4\}) \cup \mathrm{cx}(\{\mathbf{a}_4, \mathbf{a}_5\})\Big)\Big)$$

(wobei $\mathrm{cx}(A)$ die konvexe Hülle einer Menge A bezeichnet). Dann gilt $A_1, A_2, A_3, A_4, A_5 \in \mathcal{B}(\mathbb{R}^2)$ sowie $A_1 \cap A_2 = A_2 \cap A_3 = A_4 \cap A_5 = \emptyset$ und

$$A_2 + A_3 = W^2 = A_4 + A_5$$

Wegen

$$\mathbf{D}_2^{-1}\mathbf{a}_2 = \mathbf{a}_1$$
$$\mathbf{D}_2^{-1}\mathbf{a}_3 = \mathbf{a}_2$$
$$\mathbf{D}_2^{-1}\mathbf{a}_4 = \mathbf{a}_4$$
$$\mathbf{D}_2^{-1}\mathbf{a}_5 = \mathbf{a}_5$$

gilt

$$\mathbf{D}_2^{-1}(A_4) = A_1$$
$$\mathbf{D}_2^{-1}(A_5) = A_2$$

und damit

$$\mathbf{D}_2^{-1}(W^2) = \mathbf{D}_2^{-1}(A_4) + \mathbf{D}_2^{-1}(A_5) = A_1 + A_2$$

Daher gilt

$$\mathbf{D}^{-1}(W^n) = \mathbf{D}_2^{-1}(W^2) \times W^{n-2} = A_1 \times W^{n-2} + A_2 \times W^{n-2}$$

Da A_1 durch Translation aus A_3 hervorgeht, geht auch $A_1 \times W^{n-2}$ durch Translation aus $A_3 \times W^{n-2}$ hervor. Aus der letzten Gleichung und der Translationsinvarianz des Lebesgue–Maßes ergibt sich nun

$$\begin{aligned}\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] &= \boldsymbol{\lambda}^n[A_1 \times W^{n-2}] + \boldsymbol{\lambda}^n[A_2 \times W^{n-2}] \\ &= \boldsymbol{\lambda}^n[A_3 \times W^{n-2}] + \boldsymbol{\lambda}^n[A_2 \times W^{n-2}] \\ &= \boldsymbol{\lambda}^n[W^2 \times W^{n-2}] \\ &= \boldsymbol{\lambda}^n[W^n] \\ &= 1\end{aligned}$$

Daher gilt auch in diesem Fall

$$\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] = \frac{1}{|\det(\mathbf{D})|}$$

Nach Proposition C.1.1 ist jede invertierbare Matrix ein Produkt von Permutationsmatrizen und Elementarmatrizen der vorher betrachteten Art. Aus dem Gezeigten folgt daher, dass für jede invertierbare Matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$

$$\boldsymbol{\lambda}^n[\mathbf{D}^{-1}(W^n)] = \frac{1}{|\det(\mathbf{D})|}$$

gilt. Damit ist (2) für $n \geq 3$ gezeigt.
Für $n = 2$ verläuft der Beweis von (2) analog, und für $n = 1$ ist (2) trivial. □

Die Sätze 6.2.2 und 6.3.1 lassen sich nun wie folgt zusammenfassen:

6.3.2 Folgerung. *Sei $T : \mathbb{R}^n \to \mathbb{R}^n$ gegeben durch*

$$T(\mathbf{x}) := \mathbf{c} + \mathbf{D}\mathbf{x}$$

mit $\mathbf{c} \in \mathbb{R}^n$ und einer invertierbaren Matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$. Dann gilt

$$\boldsymbol{\lambda}^n_T = \frac{1}{|\det(\mathbf{D})|}\,\boldsymbol{\lambda}^n$$

Ist $\mathbf{D}$ eine Orthogonalmatrix, so gilt $\boldsymbol{\lambda}^n_T = \boldsymbol{\lambda}^n$.

Beweis. Sei $S : \mathbb{R}^n \to \mathbb{R}^n$ gegeben durch $S(\mathbf{x}) := \mathbf{D}\mathbf{x}$ und sei $T_\mathbf{c}$ die Translation bezüglich $\mathbf{c}$. Dann gilt $T = T_\mathbf{c} \circ S$. Aus Satz 6.3.1 und der Translationsinvarianz des Lebesgue–Maßes folgt nun

$$\boldsymbol{\lambda}^n_T = \boldsymbol{\lambda}^n_{T_\mathbf{c}\circ S} = \left(\boldsymbol{\lambda}^n_S\right)_{T_\mathbf{c}} = \left(\frac{1}{|\det(\mathbf{D})|}\,\boldsymbol{\lambda}^n\right)_{T_\mathbf{c}} = \frac{1}{|\det(\mathbf{D})|}\,\boldsymbol{\lambda}^n_{T_\mathbf{c}} = \frac{1}{|\det(\mathbf{D})|}\,\boldsymbol{\lambda}^n$$

Damit ist die Folgerung bewiesen. □

Die letzte Aussage der Folgerung wird als *Bewegungsinvarianz* des Lebesgue–Maßes bezeichnet.

Teil III

Integrationstheorie

7

Messbare Funktionen

Sei Ω eine nichtleere Menge. Eine Abbildung $f : \Omega \to \mathbb{R}$ heißt *reelle Funktion* und eine Abbildung $f : \Omega \to \bar{\mathbb{R}}$ heißt *numerische Funktion*. Jede reelle Funktion ist eine numerische Funktion.

Das eigentliche Interesse gilt reellen Funktionen. Da aber das Infimum und das Supremum einer Folge von reellen Funktionen im allgemeinen keine reelle Funktion sondern nur eine numerische Funktion ist, ist es erforderlich, auch numerische Funktionen zu betrachten.

Neben den üblichen Rechenregeln auf der Menge $\bar{\mathbb{R}}$ der erweiterten reellen Zahlen verwenden wir die in der Integrationstheorie üblichen Konventionen $(\pm\infty) \cdot 0 := 0$ und $0 \cdot (\pm\infty) := 0$.

Für numerische Funktionen betrachten wir im folgenden stets die punktweise definierte Addition, die punktweise definierte Multiplikation und die punktweise definierte Ordnungsrelation.

Für numerische Funktionen f und g setzen wir

$$\begin{aligned} f \vee g &:= \max\{f, g\} \\ f \wedge g &:= \min\{f, g\} \end{aligned}$$

und

$$\begin{aligned} f^+ &:= f \vee 0 \\ f^- &:= (-f) \vee 0 \\ |f| &:= f \vee (-f) \end{aligned}$$

Dann gilt $f = f^+ - f^-$ und $|f| = f^+ + f^-$.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_7, © Springer-Verlag Berlin Heidelberg 2011

Für eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von numerischen Funktionen heißt die Funktion

$$\liminf_{n\to\infty} f_n := \sup_{n\in\mathbb{N}} \inf_{k\in\mathbb{N}(n)} f_k$$

der *Limes inferior* der Folge $\{f_n\}_{n\in\mathbb{N}}$ und die Funktion

$$\limsup_{n\to\infty} f_n := \inf_{n\in\mathbb{N}} \sup_{k\in\mathbb{N}(n)} f_k$$

heißt der *Limes superior* der Folge $\{f_n\}_{n\in\mathbb{N}}$, und wir nennen die Folge (*punktweise*) *konvergent*, wenn $\liminf_{n\to\infty} f_n = \limsup_{n\to\infty} f_n$ gilt.

In diesem Kapitel betrachten wir zunächst verschiedene Klassen messbarer Funktionen auf einem Messraum und untersuchen ihre Struktureigenschaften (Abschnitt 7.1). Wir gehen dann zu einem Maßraum über, der durch die Festlegung eines Maßes auf dem Messraum entsteht, und betrachten fast überall bestehende Eigenschaften messbarer Funktionen und die Bildung von Äquivalenzklassen bezüglich des Maßes (Abschnitt 7.2).

7.1 Messbare Funktionen auf einem Messraum

Im gesamten Abschnitt sei $(\Omega, \mathcal{F})$ ein Messraum.

Wir betrachten zunächst numerische Funktionen. Eine numerische Funktion heißt *$\mathcal{F}$–messbar* oder kurz *messbar*, wenn sie $\mathcal{F}$–$\mathcal{B}(\bar{\mathbb{R}})$–messbar ist.

7.1.1 Beispiele.

(1) **Konstante Funktion:** Für $c \in \bar{\mathbb{R}}$ ist die konstante Funktion $f : \Omega \to \bar{\mathbb{R}}$ mit

$$f(\omega) := c$$

messbar.
In der Tat: Für alle $B \in \mathcal{B}(\bar{\mathbb{R}})$ gilt

$$f^{-1}(B) = \begin{cases} \Omega & \text{falls } c \in B \\ \emptyset & \text{sonst} \end{cases}$$

und damit $f^{-1}(B) \in \mathcal{F}$.

(2) **Indikatorfunktion:** Für $A \in 2^\Omega$ heißt die Funktion $\chi_A : \Omega \to \bar{\mathbb{R}}$ mit

$$\chi_A(\omega) := \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{sonst} \end{cases}$$

Indikatorfunktion von A. Die Indikatorfunktion χ_A ist genau dann messbar, wenn $A \in \mathcal{F}$ gilt.
In der Tat: Für alle $B \in \mathcal{B}(\bar{\mathbb{R}})$ gilt

$$\chi_A^{-1}(B) = \begin{cases} \Omega & \text{falls } 0 \in B \text{ und } 1 \in B \\ A & \text{falls } 0 \notin B \text{ und } 1 \in B \\ \Omega \setminus A & \text{falls } 0 \in B \text{ und } 1 \notin B \\ \emptyset & \text{falls } 0 \notin B \text{ und } 1 \notin B \end{cases}$$

Daraus folgt die Behauptung.

Für eine numerische Funktion f und für $B \in \mathcal{B}(\bar{\mathbb{R}})$ und $a, b, c \in \bar{\mathbb{R}}$ setzen wir

$$\begin{aligned}\{f \in B\} &:= \{\omega \in \Omega \mid f(\omega) \in B\} \\ \{f \notin B\} &:= \{\omega \in \Omega \mid f(\omega) \notin B\}\end{aligned}$$

und

$$\begin{aligned}\{a < f < b\} &:= \{\omega \in \Omega \mid a < f(\omega) < b\} \\ \{a < f \le b\} &:= \{\omega \in \Omega \mid a < f(\omega) \le b\} \\ \{a \le f < b\} &:= \{\omega \in \Omega \mid a \le f(\omega) < b\} \\ \{a \le f \le b\} &:= \{\omega \in \Omega \mid a \le f(\omega) \le b\}\end{aligned}$$

sowie

$$\begin{aligned}\{f < c\} &:= \{\omega \in \Omega \mid f(\omega) < c\} \\ \{f \le c\} &:= \{\omega \in \Omega \mid f(\omega) \le c\} \\ \{f = c\} &:= \{\omega \in \Omega \mid f(\omega) = c\} \\ \{f \ne c\} &:= \{\omega \in \Omega \mid f(\omega) \ne c\} \\ \{f \ge c\} &:= \{\omega \in \Omega \mid f(\omega) \ge c\} \\ \{f > c\} &:= \{\omega \in \Omega \mid f(\omega) > c\}\end{aligned}$$

Die Messbarkeit einer numerischen Funktion lässt sich wie folgt charakterisieren:

7.1.2 Satz. *Für eine numerische Funktion f sind folgende Aussagen äquivalent*:
(a) *f ist messbar.*
(b) *Für alle $c \in \mathbb{R}$ gilt $\{f \le c\} \in \mathcal{F}$.*
(c) *Für alle $c \in \mathbb{R}$ gilt $\{f < c\} \in \mathcal{F}$.*
(d) *Für alle $c \in \mathbb{R}$ gilt $\{f \ge c\} \in \mathcal{F}$.*
(e) *Für alle $c \in \mathbb{R}$ gilt $\{f > c\} \in \mathcal{F}$.*
In diesem Fall gilt $\{f = c\} \in \mathcal{F}$ und $\{f \ne c\} \in \mathcal{F}$ für alle $c \in \mathbb{R}$.

Beweis. Wir nehmen zunächst an, dass (a) gilt. Für alle $c \in \mathbb{R}$ gilt dann

$$\{f \le c\} = f^{-1}([-\infty, c]) \in f^{-1}(\mathcal{B}(\bar{\mathbb{R}})) \subseteq \mathcal{F}$$

Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Für das Mengensystem

$$\bar{\mathcal{E}} = \Big\{[-\infty, c] \;\Big|\; c \in \mathbb{R}\Big\}$$

gilt $\sigma(\bar{\mathcal{E}}) = \mathcal{B}(\bar{\mathbb{R}})$; vgl. Aufgabe 1.5.N. Für alle $c \in \mathbb{R}$ gilt nach Voraussetzung $f^{-1}([-\infty, c]) = \{f \le c\} \in \mathcal{F}$. Daraus folgt zunächst $f^{-1}(\bar{\mathcal{E}}) \subseteq \mathcal{F}$ und sodann

$$f^{-1}(\mathcal{B}(\bar{\mathbb{R}})) = f^{-1}(\sigma(\bar{\mathcal{E}})) = \sigma(f^{-1}(\bar{\mathcal{E}})) \subseteq \mathcal{F}$$

Daher folgt (a) aus (b).

Die Äquivalenz von (b), (c), (d), (e) ergibt sich nun aus den Gleichungen

$$\begin{aligned}
\{f < c\} &= \bigcup_{n\in\mathbb{N}} \left\{ f \leq c - \frac{1}{n} \right\} \\
\{f \geq c\} &= \Omega \setminus \{f < c\} \\
\{f > c\} &= \bigcup_{n\in\mathbb{N}} \left\{ f \geq c + \frac{1}{n} \right\} \\
\{f \leq c\} &= \Omega \setminus \{f > c\}
\end{aligned}$$

und die abschließende Behauptung ergibt sich aus den Gleichungen

$$\begin{aligned}
\{f = c\} &= \{f \leq c\} \setminus \{f < c\} \\
\{f \neq c\} &= \Omega \setminus \{f = c\}
\end{aligned}$$

Damit ist der Satz gezeigt. □

Wir ziehen nun erste Folgerungen aus Satz 7.1.2:

7.1.3 Folgerung. *Seien f und g messbare numerische Funktionen und sei $a \in \bar{\mathbb{R}}$. Dann sind die Funktionen af sowie $f \vee g$ und $f \wedge g$ messbar. Insbesondere sind die Funktionen $f^+, f^-, |f|$ messbar.*

Beweis. Im Fall $a = 0$ gilt $af = 0$ und die Messbarkeit von af ist klar; analog argumentiert man in den Fällen $a = \infty$ und $a = -\infty$. Im Fall $a \in (0, \infty)$ erhält man für alle $c \in \mathbb{R}$ aus der Messbarkeit von f und Satz 7.1.2 zunächst

$$\{af \leq c\} = \{f \leq c/a\} \in \mathcal{F}$$

und durch nochmalige Anwendung von Satz 7.1.2 die Messbarkeit von af; analog argumentiert man im Fall $a \in (-\infty, 0)$. Für alle $c \in \mathbb{R}$ gilt ferner

$$\begin{aligned}
\{f \vee g \leq c\} &= \{f \leq c\} \cap \{g \leq c\} \\
\{f \wedge g \geq c\} &= \{f \geq c\} \cap \{g \geq c\}
\end{aligned}$$

und wie vorher folgt aus Satz 7.1.2 die Messbarkeit von $f \vee g$ und $f \wedge g$. Aus dem bereits gezeigten folgt nun die Messbarkeit von $f^+, f^-, |f|$. □

Die Summe von zwei numerischen Funktionen ist unter Umständen nicht auf der gesamten Grundmenge Ω definiert. Sind jedoch f, g, h numerische Funktionen mit $h = f + g$, so lässt sich zeigen, dass aus der Messbarkeit von f und g die Messbarkeit von h folgt; vgl. Folgerung 7.1.15. Wir betrachten nun Folgen von messbaren numerischen Funktionen:

7.1.4 Folgerung. *Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren numerischen Funktionen. Dann gilt:*
(1) *Die Funktionen $\sup_{n\in\mathbb{N}} f_n$ und $\inf_{n\in\mathbb{N}} f_n$ sind messbar.*
(2) *Die Funktionen $\liminf_{n\to\infty} f_n$ und $\limsup_{n\to\infty} f_n$ sind messbar.*
(3) *Ist f eine numerische Funktion und ist die Folge $\{f_n\}_{n\in\mathbb{N}}$ konvergent mit $f = \lim_{n\to\infty} f_n$, so ist f messbar.*

Beweis. Die Messbarkeit von $\sup_{n\in\mathbb{N}} f_n$ und $\inf_{n\in\mathbb{N}} f_n$ zeigt man genau wie im Beweis von Folgerung 7.1.3. Wegen

$$\begin{aligned}\liminf_{n\to\infty} f_n &= \sup_{n\in\mathbb{N}} \inf_{k\in\mathbb{N}(n)} f_k \\ \limsup_{n\to\infty} f_n &= \inf_{n\in\mathbb{N}} \sup_{k\in\mathbb{N}(n)} f_k\end{aligned}$$

sind auch $\liminf_{n\to\infty} f_n$ und $\limsup_{n\to\infty} f_n$ messbar. Ist schließlich f eine numerische Funktion und die Folge $\{f_n\}_{n\in\mathbb{N}}$ konvergent mit $f = \lim_{n\to\infty} f_n$, so gilt $f = \liminf_{n\to\infty} f_n = \limsup_{n\to\infty} f_n$ und damit ist f messbar. □

Für numerische Funktionen f und g setzen wir

$$\begin{aligned}\{f < g\} &:= \{\omega \in \Omega \mid f(\omega) < g(\omega)\} \\ \{f \le g\} &:= \{\omega \in \Omega \mid f(\omega) \le g(\omega)\} \\ \{f = g\} &:= \{\omega \in \Omega \mid f(\omega) = g(\omega)\} \\ \{f \ne g\} &:= \{\omega \in \Omega \mid f(\omega) \ne g(\omega)\} \\ \{f \ge g\} &:= \{\omega \in \Omega \mid f(\omega) \ge g(\omega)\} \\ \{f > g\} &:= \{\omega \in \Omega \mid f(\omega) > g(\omega)\}\end{aligned}$$

Auch diese Mengen sind messbar, wenn f und g messbar sind:

7.1.5 Folgerung. *Seien f und g messbare numerische Funktionen. Dann gilt $\{f < g\}, \{f \le g\}, \{f = g\}, \{f \ne g\}, \{f \ge g\}, \{f > g\} \in \mathcal{F}$.*

Beweis. Nach Satz 7.1.2 gilt $\{f < g\} = \bigcup_{q\in\mathbb{Q}} \{f < q\} \cap \{q < g\} \in \mathcal{F}$ und damit $\{f \ge g\} = \Omega \setminus \{f < g\} \in \mathcal{F}$, und aus Symmetriegründen gilt dann auch $\{f > g\} \in \mathcal{F}$ und $\{f \le g\} \in \mathcal{F}$, und damit $\{f = g\} = \{f \le g\} \cap \{f \ge g\} \in \mathcal{F}$ und $\{f \ne g\} = \Omega \setminus \{f = g\} \in \mathcal{F}$. □

Wir betrachten nun reelle Funktionen. Eine reelle Funktion heißt *$\mathcal{F}$–messbar* oder kurz *messbar*, wenn sie $\mathcal{F}$–$\mathcal{B}(\mathbb{R})$–messbar ist. Aus der in Beispiel 1.2.6 angegebenen Beziehung zwischen den Borelschen σ–Algebren $\mathcal{B}(\mathbb{R})$ und $\mathcal{B}(\bar{\mathbb{R}})$ erkennt man, dass eine reelle Funktion genau dann $\mathcal{F}$–$\mathcal{B}(\mathbb{R})$–messbar ist, wenn sie, aufgefasst als eine numerische Funktion, die die Werte $+\infty$ und $-\infty$ nicht annimmt, $\mathcal{F}$–$\mathcal{B}(\bar{\mathbb{R}})$–messbar ist.

7.1.6 Lemma. *Seien f und g messbare reelle Funktionen. Dann sind auch die Funktionen $f+g$ und fg messbar.*

Beweis. Nach Beispiel 3.3.1 gilt $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. Daher folgt aus der Messbarkeit von f und g die Messbarkeit der Funktion $H : \Omega \to \mathbb{R}^2$ mit

$$H(\omega) := \begin{pmatrix} f(\omega) \\ g(\omega) \end{pmatrix}$$

Da sowohl die Addition $+ : \mathbb{R}^2 \to \mathbb{R}$ als auch die Multiplikation $\cdot : \mathbb{R}^2 \to \mathbb{R}$ stetig und damit messbar ist und da nach Lemma 2.3.4 die Komposition von messbaren Abbildungen messbar ist, sind auch die Funktionen $f+g$ und fg messbar. □

Sei nun

$$\mathcal{L}^0(\mathcal{F}) := \Big\{ f : \Omega \to \mathbb{R} \,\Big|\, f \text{ ist } \mathcal{F}\text{–messbar} \Big\}$$

Dann lassen sich die wichtigsten Stabilitätseigenschaften der Familie der messbaren reellen Funktionen wie folgt zusammenfassen:

7.1.7 Satz. *$\mathcal{L}^0(\mathcal{F})$ ist ein Vektorverband.*

Beweis. Aus Lemma 7.1.6 und Folgerung 7.1.3 ergibt sich zunächst, dass die messbaren reellen Funktionen einen Vektorraum bilden.
Für alle messbaren reellen Funktionen f, g, h mit $f \le g$ und für alle $a \in \mathbb{R}_+$ gilt $f+h \le g+h$ und $af \le ag$. Daher bilden die messbaren reellen Funktionen einen geordneten Vektorraum.
Für alle messbaren reellen Funktionen f, g sind nach Folgerung 7.1.3 auch $f \vee g$ und $f \wedge g$ messbare reelle Funktionen. Für jede messbare reelle Funktion h mit $f \le h$ und $g \le h$ gilt offenbar $f \vee g \le h$; damit ist $f \vee g$ das Supremum von f und g in der Familie aller messbaren reellen Funktionen, und analog zeigt man, dass $f \wedge g$ das Infimum von f und g in der Familie aller messbaren reellen Funktionen ist. Daher bilden die messbaren reellen Funktionen einen Vektorverband. □

Sei ferner

$$\mathcal{L}^\infty(\mathcal{F}) := \left\{ f \in \mathcal{L}^0(\mathcal{F}) \,\middle|\, \sup_{\omega \in \Omega} |f(\omega)| \le c \text{ für ein } c \in \mathbb{R} \right\}$$

die Familie der beschränkten messbaren reellen Funktionen.

7.1.8 Satz. *$\mathcal{L}^\infty(\mathcal{F})$ ist ein Vektorverband und ein Ideal in $\mathcal{L}^0(\mathcal{F})$.*

Wir kommen auf den Raum $\mathcal{L}^\infty(\mathcal{F})$ erst im nächsten Abschnitt zurück.

Wir betrachten abschließend eine Familie messbarer reeller Funktionen, die für die Integrationstheorie von zentraler Bedeutung ist. Eine reelle Funktion heißt *einfach*, wenn sie messbar ist und nur endlich viele Werte annimmt. Die wichtigsten Stabilitätseigenschaften der Familie der einfachen Funktionen lassen sich wie folgt zusammenfassen:

7.1.9 Satz. *Die einfachen Funktionen bilden einen Vektorverband.*

Beweis. Die Behauptung ergibt sich unmittelbar aus Satz 7.1.7, da sowohl Linearkombinationen als auch das Maximum und das Minimum von zwei einfachen Funktionen nur endlich viele Werte annehmen. □

Sei f eine einfache Funktion. Dann ist das Bild $f(\Omega)$ von f eine endliche Menge und f besitzt die Darstellung

$$f = \sum_{a \in f(\Omega)} a\chi_{\{f=a\}}$$

Diese Darstellung heißt *Standarddarstellung* von f. Da f messbar ist, gilt für alle $a \in f(\Omega)$ nach Folgerung 7.1.5 $\{f = a\} \in \mathcal{F}$. Daher ist jede einfache Funktion eine Linearkombination von Indikatorfunktionen von messbaren Mengen; andererseits ist nach Beispiel 7.1.1 und Satz 7.1.9 jede Linearkombination von Indikatorfunktionen von messbaren Mengen eine einfache Funktion. Damit erhalten wir das folgenden Lemma:

7.1.10 Lemma. *Für eine reelle Funktion f sind folgende Aussagen äquivalent*:
(a) *f ist einfach.*
(b) *Es gibt ein $n \in \mathbb{N}$ sowie eine Familie $\{C_i\}_{i\in\{1,\dots,n\}} \subseteq \mathcal{F}$ und eine Familie $\{c_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}$ mit $f = \sum_{i=1}^n c_i\chi_{C_i}$.*

Dieses Lemma besitzt ein Analogon für positive reelle Funktionen:

7.1.11 Lemma. *Für eine positive reelle Funktion f sind folgende Aussagen äquivalent*:
(a) *f ist einfach.*
(b) *Es gibt ein $n \in \mathbb{N}$ sowie eine Familie $\{C_i\}_{i\in\{1,\dots,n\}} \subseteq \mathcal{F}$ und eine Familie $\{c_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}_+$ mit $f = \sum_{i=1}^n c_i\chi_{C_i}$.*

Die Darstellung einer einfachen Funktion durch eine Linearkombination von Indikatorfunktionen von messbaren Mengen ist im allgemeinen nicht eindeutig; dies gilt auch für positive einfache Funktionen:

7.1.12 Beispiel. Seien $A, B \in \mathcal{F}$ und $a, b \in \mathbb{R}_+$. Dann gilt für alle $\omega \in \Omega$

$$a\,\chi_A(\omega) + b\,\chi_B(\omega) = a\,\chi_{A\setminus B}(\omega) + (a+b)\,\chi_{A\cap B}(\omega) + b\,\chi_{B\setminus A}(\omega)$$

Damit liegen zwei Darstellungen derselben positiven einfachen Funktion vor.

Nach Folgerung 7.1.4 ist das Supremum einer monoton wachsenden Folge von positiven einfachen Funktionen eine positive messbare numerische Funktion. Der folgende Satz zeigt, dass andererseits jede positive messbare numerische Funktion als Supremum einer monoton wachsenden Folge von positiven einfachen Funktionen dargestellt werden kann:

7.1.13 Satz (Approximationssatz; positiver Fall). *Für eine positive numerische Funktion f sind folgende Aussagen äquivalent:*
(a) *f ist messbar.*
(b) *Es gibt eine monoton wachsende Folge $\{f_n\}_{n\in\mathbb{N}}$ von positiven einfachen Funktionen mit $f = \sup_{n\in\mathbb{N}} f_n$.*

Beweis. Es genügt zu zeigen, dass (b) aus (a) folgt, denn die umgekehrte Implikation ist wegen Folgerung 7.1.4 klar.
Sei also f messbar. Für $n \in \mathbb{N}$ und alle $k \in \{0, 1, \dots, 4^n\}$ setzen wir

$$C_{n,k} := \begin{cases} \{k2^{-n} \le f < (k+1)2^{-n}\} & \text{falls } k \in \{0,1,\dots,4^n - 1\} \\ \{2^n \le f\} & \text{falls } k = 4^n \end{cases}$$

und für $n \in \mathbb{N}$ definieren wir

$$f_n := \sum_{k=0}^{4^n} k2^{-n}\, \chi_{C_{n,k}}$$

Aus der Messbarkeit von f folgt zunächst $\{C_{n,k}\}_{k\in\{0,1,\dots,4^n\}} \subseteq \mathcal{F}$ und sodann die Messbarkeit von f_n. Daher ist f_n eine einfache Funktion mit $0 \le f_n \le f$, und aus der Gleichung

$$C_{n,k} = \begin{cases} C_{n+1,2k} + C_{n+1,2k+1} & \text{falls } k \in \{0,1,\dots,4^n - 1\} \\ \sum_{j=2^{2n+1}}^{4^{n+1}} C_{n+1,j} & \text{falls } k = 4^n \end{cases}$$

ergibt sich ferner $f_n \le f_{n+1}$. Daher ist $\{f_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge von positiven einfachen Funktionen mit

$$\sup_{n\in\mathbb{N}} f_n \le f$$

Sei nun $\omega \in \Omega$.
– Im Fall $f(\omega) < \infty$ gilt für alle $n \in \mathbb{N}$, die hinreichend groß sind, $f(\omega) < 2^n$ und damit $f(\omega) < f_n(\omega) + 2^{-n}$, also $f(\omega) \le \sup_{m\in\mathbb{N}} f_m(\omega) + 2^{-n}$, und damit $f(\omega) \le \sup_{m\in\mathbb{N}} f_m(\omega)$.
– Im Fall $f(\omega) = \infty$ gilt $f_n(\omega) = 2^n$ für alle $n \in \mathbb{N}$, und damit $f(\omega) = \infty = \sup_{n\in\mathbb{N}} f_n(\omega)$.

Daher gilt auch

$$f \le \sup_{n\in\mathbb{N}} f_n$$

und damit $f = \sup_{n\in\mathbb{N}} f_n$. Daher folgt (b) aus (a). □

Aus dem letzten Satz erhält man mit Hilfe der Zerlegung $f = f^+ - f^-$ ein Analogon für beliebige numerische Funktionen:

7.1.14 Satz (Approximationssatz; allgemeiner Fall). *Für eine numerische Funktion f sind folgende Aussagen äquivalent*:
(a) *f ist messbar.*
(b) *Es gibt eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von einfachen Funktionen mit $f = \lim_{n\to\infty} f_n$.*
(c) *Es gibt eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von einfachen Funktionen mit $f = \lim_{n\to\infty} f_n$ und $|f| = \sup_{n\in\mathbb{N}} |f_n|$.*

Aus dem Approximationssatz folgt, dass die Familie der messbaren numerischen Funktionen unter der Bildung von Summen und Produkten stabil ist:

7.1.15 Folgerung. *Seien f, g, h numerische Funktionen mit $h = f + g$ oder $h = fg$. Sind f und g messbar, so ist auch h messbar.*

Als eine weitere Anwendung des Approximationssatzes ergibt sich der folgende Faktorisierungssatz, den wir in Kapitel 18 benötigen:

7.1.16 Satz (Faktorisierungssatz). *Sei $(\Omega', \mathcal{F}')$ ein Messraum und sei $g : \Omega \to \Omega'$ eine Abbildung und $\mathcal{G} := \sigma(g)$. Für eine numerische Funktion f sind dann folgende Aussagen äquivalent*:
(a) *f ist $\mathcal{G}$–messbar.*
(b) *Es gibt eine messbare numerische Funktion $h : \Omega' \to \bar{\mathbb{R}}$ mit $f = h \circ g$.*
Ist f $\mathcal{G}$–messbar und positiv bzw. reell, so kann auch h positiv bzw. reell gewählt werden.

Beweis. Wir nehmen zunächst an, dass f eine positive $\mathcal{G}$–messbare einfache Funktion ist. Dann gibt es eine disjunkte Familie $\{G_i\}_{i\in\{1,\dots,m\}} \subseteq \mathcal{G}$ und eine Familie $\{a_i\}_{i\in\{1,\dots,m\}} \subseteq \mathbb{R}_+$ mit

$$f = \sum_{i=1}^{m} a_i \, \chi_{G_i}$$

Nach Definition von $\mathcal{G}$ gibt es für alle $i \in \{1, \dots, m\}$ eine Menge $H_i \in \mathcal{F}'$ mit $G_i = g^{-1}(H_i)$. Sei

$$h := \sum_{i=1}^{m} a_i \, \chi_{H_i}$$

Dann ist h eine positive $\mathcal{F}'$–messbare einfache Funktion und für alle $\omega \in \Omega$ gilt

$$f(\omega) = \sum_{i=1}^{m} a_i \, \chi_{G_i}(\omega) = \sum_{i=1}^{m} a_i \, \chi_{g^{-1}(H_i)}(\omega) = \sum_{i=1}^{m} a_i \, \chi_{H_i}(g(\omega)) = h(g(\omega))$$

und damit $f = h \circ g$.

Wir nehmen nun an, dass f eine positive $\mathcal{G}$–messbare numerische Funktion ist. Nach dem Approximationssatz 7.1.13 gibt es eine monoton wachsende Folge von positiven $\mathcal{G}$–messbaren einfachen Funktionen $\{f_n\}_{n\in\mathbb{N}}$ mit $f = \sup_{n\in\mathbb{N}} f_n$ und nach dem bereits gezeigten gibt es für alle $n \in \mathbb{N}$ eine positive $\mathcal{F}'$–messbare einfache Funktion h_n mit $f_n = h_n \circ g$. Sei $h := \sup_{n\in\mathbb{N}} h_n$. Dann ist h eine positive $\mathcal{F}'$–messbare numerische Funktion und es gilt

$$f = \sup_{n\in\mathbb{N}} f_n = \sup_{n\in\mathbb{N}} (h_n \circ g) = \left(\sup_{n\in\mathbb{N}} h_n\right) \circ g = h \circ g$$

Wir nehmen schließlich an, dass f eine beliebige $\mathcal{G}$–messbare numerische Funktion ist. Dann besitzt f die Zerlegung $f = f^+ - f^-$ und es gibt positive $\mathcal{F}'$–messbare numerische Funktionen h_1 und h_2 mit $f^+ = h_1 \circ g$ und $f^- = h_2 \circ g$. Für die Menge $H := \{h_1 < \infty\} \cup \{h_2 < \infty\}$ gilt $g(\Omega) \subseteq H$ und damit $f^+ = (h_1\chi_H) \circ g$ und $f^- = (h_2\chi_H) \circ g$. Wegen $H \in \mathcal{F}'$ sind die Funktionen $h_1\chi_H$ und $h_2\chi_H$ $\mathcal{F}'$–messbar Damit ist auch die Funktion $h := h_1\chi_H - h_2\chi_H$ $\mathcal{F}'$–messbar und es gilt

$$f = f^+ - f^- = (h_1\chi_H) \circ g - (h_2\chi_H) \circ g = (h_1\chi_H - h_2\chi_H) \circ g = h \circ g$$

Ist f reell, so kann man in der Definition der Funktion h die Menge H durch die Menge $H_0 := \{h_1 < \infty\} \cap \{h_2 < \infty\}$ ersetzen und erhält eine $\mathcal{F}'$–messbare reelle Funktion h_0 mit $f = h_0 \circ g$. □

Aufgaben

7.1.A **Indikatorfunktion:**

(1) Für alle $A, B \in 2^\Omega$ gilt

$$\begin{aligned} \chi_{A\cap B} &= \chi_A \chi_B \\ \chi_{A\cup B} &= \chi_A + \chi_B - \chi_A\chi_B \\ \chi_{A\setminus B} &= \chi_A - \chi_A\chi_B \\ \chi_{A\triangle B} &= |\chi_A - \chi_B| \end{aligned}$$

Sind A, B disjunkt, so gilt $\chi_{A+B} = \chi_A + \chi_B$.

(2) Für alle $A, B \in 2^\Omega$ sind folgende Aussagen äquivalent:
   (a) Es gilt $B \subseteq A$.
   (b) Es gilt $\chi_B \leq \chi_A$.
   (c) Es gilt $\chi_{A\setminus B} = \chi_A - \chi_B$.

(3) Für jede Familie $\{A_i\}_{i\in I} \subseteq 2^\Omega$ gilt

$$\begin{aligned} \chi_{\bigcap_{i\in I} A_i} &= \inf_{i\in I} \chi_{A_i} \\ \chi_{\bigcup_{i\in I} A_i} &= \sup_{i\in I} \chi_{A_i} \end{aligned}$$

Ist die Familie $\{A_i\}_{i\in I}$ disjunkt, so gilt $\chi_{\sum_{i\in I} A_i} = \sum_{i\in I} \chi_{A_i}$.

(4) Für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq 2^\Omega$ gilt

$$\begin{aligned} \chi_{\liminf_{n\to\infty} A_n} &= \liminf_{n\to\infty} \chi_{A_n} \\ \chi_{\limsup_{n\to\infty} A_n} &= \limsup_{n\to\infty} \chi_{A_n} \end{aligned}$$

7.1.B Für eine numerische Funktion f sind folgende Aussagen äquivalent:
(a) f ist messbar.
(b) f^+ und f^- sind messbar.
(c) Es gibt messbare numerische Funktionen g und h mit $f = g - h$.
In diesem Fall ist auch $|f|$ messbar. Folgt umgekehrt aus der Messbarkeit von $|f|$ die Messbarkeit von f?

7.1.C Sei $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ disjunkt und sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren numerischen Funktionen. Dann ist auch $f := \sum_{n=1}^{\infty} f_n \chi_{A_n}$ messbar.

7.1.D **Approximationssatz:** Vergleichen Sie die Approximationssätze 7.1.13 und 7.1.14 mit dem Approximationssatz 5.4.1.

7.1.E Zu jeder σ–Algebra $\mathcal{G}$ auf Ω mit $\mathcal{G} \subseteq \mathcal{F}$ gibt es einen Messraum $(\Omega', \mathcal{F}')$ und eine $\mathcal{G}$–messbare Abbildung $g : \Omega \to \Omega'$ mit $\mathcal{G} = \sigma(g)$.

7.1.F **Komplexe Funktionen:** Für eine *komplexe Funktion* $f : \Omega \to \mathbb{C}$ ist
– der *Realteil* $\mathrm{Re} f : \Omega \to \mathbb{R}$ mit $(\mathrm{Re} f)(\omega) := \mathrm{Re} f(\omega)$,
– der *Imaginärteil* $\mathrm{Im} f : \Omega \to \mathbb{R}$ mit $(\mathrm{Im} f)(\omega) := \mathrm{Im} f(\omega)$, und
– der *Betrag* $|f| : \Omega \to \mathbb{R}$ mit $|f|(\omega) := |f(\omega)|$

eine reelle Funktion. Eine komplexe Funktion heißt *messbar*, wenn sie $\mathcal{F}$–$\mathcal{B}(\mathbb{R}^2)$–messbar ist.
(1) Für eine komplexe Funktion f sind folgende Aussagen äquivalent:
  (a) f ist messbar.
  (b) $\mathrm{Re} f$ und $\mathrm{Im} f$ sind messbar.
(2) Die messbaren komplexen Funktionen bilden einen Vektorraum.

7.2 Messbare Funktionen auf einem Maßraum

Ist $(\Omega, \mathcal{F})$ ein Messraum und $\mu : \mathcal{F} \to [0, \infty]$ ein Maß, so heißt das Tripel $(\Omega, \mathcal{F}, \mu)$ *Maßraum*. Ein Maßraum $(\Omega, \mathcal{F}, \mu)$ heißt *endlich*, wenn μ endlich ist, und er heißt *σ–endlich*, wenn μ σ–endlich ist.

Im gesamten Abschnitt sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum.

Eine Menge $N \in \mathcal{F}$ heißt *μ–Nullmenge* oder kurz *Nullmenge*, wenn $\mu[N] = 0$ gilt. Das folgende Lemma fasst die wichtigsten Stabilitätseigenschaften der Familie der Nullmengen zusammen; vgl. Aufgabe 4.1.F.

7.2.1 Lemma.
(1) *Sei $N \in \mathcal{F}$ eine μ–Nullmenge. Dann ist jede Menge $M \in \mathcal{F}$ mit $M \subseteq N$ eine μ–Nullmenge.*
(2) *Sei I abzählbar und sei $\{N_i\}_{i\in I}$ eine Familie von μ–Nullmengen. Dann ist $\bigcup_{i\in I} N_i$ eine μ–Nullmenge.*

Beweis. Jedes Maß ist monoton und σ–subadditiv. □

Nullmengen sind vor allem in der Integrationstheorie von Bedeutung, denn es wird sich zeigen, dass die Eigenschaften einer Funktion auf einer Nullmenge die Eigenschaften ihres Integrals nicht beeinflussen.

Von einer Eigenschaft, die für jedes $\omega \in \Omega$ entweder gilt oder nicht gilt, sagt man, sie gelte *μ–fast überall* oder kurz *fast überall*, wenn es eine μ–Nullmenge N gibt derart, dass die Eigenschaft für alle $\omega \in \Omega \setminus N$ gilt. Dies bedeutet im allgemeinen jedoch nicht, dass die Menge aller $\omega \in \Omega$, für die die Eigenschaft nicht gilt, selbst eine μ–Nullmenge ist. Insbesondere ergeben sich für Funktionen, die auf einer Teilmenge von Ω definiert sind, die folgenden Definitionen:

– Eine Funktion $h : \Omega_h \to \bar{\mathbb{R}}$ heißt *μ–fast überall definiert*, wenn es eine μ–Nullmenge N gibt mit $\Omega \setminus \Omega_h \subseteq N$.
– Eine Funktion $h : \Omega_h \to \bar{\mathbb{R}}$ heißt *μ–fast überall konstant*, wenn es eine μ–Nullmenge N und ein $c \in \bar{\mathbb{R}}$ gibt mit $\Omega \setminus \Omega_h \subseteq N$ und $h(\omega) = c$ für alle $\omega \in \Omega \setminus N$.
– Eine Funktion $h : \Omega_h \to \bar{\mathbb{R}}$ heißt *μ–fast überall reell* oder *μ–fast überall endlich*, wenn es eine μ–Nullmenge N gibt mit $\Omega \setminus \Omega_h \subseteq N$ und $h(\omega) \in \mathbb{R}$ für alle $\omega \in \Omega \setminus N$.
– Zwei Funktionen $g : \Omega_g \to \bar{\mathbb{R}}$ und $h : \Omega_h \to \bar{\mathbb{R}}$ heißen *μ–fast überall gleich*, wenn es eine μ–Nullmenge N gibt mit $(\Omega \setminus \Omega_g) \cup (\Omega \setminus \Omega_h) \subseteq N$ und $g(\omega) = h(\omega)$ für alle $\omega \in \Omega \setminus N$.
– Eine Folge $\{h_n\}_{n\in\mathbb{N}}$ von Funktionen $h_n : \Omega_{h_n} \to \bar{\mathbb{R}}$ heißt *μ–fast überall konvergent*, wenn es eine Funktion $h : \Omega_h \to \bar{\mathbb{R}}$ und eine μ–Nullmenge N gibt mit $(\Omega \setminus \Omega_h) \cup \bigcup_{n\in\mathbb{N}} (\Omega \setminus \Omega_{h_n}) \subseteq N$ und für jedes $\omega \in \Omega \setminus N$ die Folge $\{h_n(\omega)\}_{n\in\mathbb{N}}$ in $\bar{\mathbb{R}}$ gegen $h(\omega)$ konvergiert.
– Eine Folge $\{h_n\}_{n\in\mathbb{N}}$ von Funktionen $h_n : \Omega_{h_n} \to \bar{\mathbb{R}}$ heißt *μ–fast überall Cauchy*, wenn es eine μ–Nullmenge N gibt mit $\bigcup_{n\in\mathbb{N}} (\Omega \setminus \Omega_{h_n}) \subseteq N$ und für jedes $\omega \in \Omega \setminus N$ die Folge $\{h_n(\omega)\}_{n\in\mathbb{N}}$ eine Cauchy–Folge in $\mathbb{R}$ ist.

Die Konvergenz μ–fast überall wird im folgenden ein wichtige Rolle spielen. Das folgende Ergebnis ist offensichtlich und gilt auch dann, wenn alle Funktionen nur μ–fast überall definiert sind:

7.2.2 Lemma. *Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von Funktionen.*
(1) *Ist $\{f_n\}_{n\in\mathbb{N}}$ μ–fast überall konvergent und sind f und g Funktionen mit $\lim_{n\to\infty} f_n(\omega) = f(\omega)$ μ–fast überall und $\lim_{n\to\infty} f_n(\omega) = g(\omega)$ μ–fast überall, so gilt $f(\omega) = g(\omega)$ μ–fast überall.*
(2) *Sind alle f_n μ–fast überall reell, so ist die Folge $\{f_n\}_{n\in\mathbb{N}}$ genau dann μ–fast überall konvergent gegen eine μ–fast überall reelle Funktion, wenn sie μ–fast überall Cauchy ist.*

Das folgende Beispiel zeigt, dass beim Umgang mit μ–fast überall bestehenden Eigenschaften Vorsicht geboten ist:

7.2.3 Beispiel (Fast überall stetige Funktionen). Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$. Eine Funktion $h : \Omega_h \to \mathbb{R}$ mit $\Omega_h \subseteq \Omega$ heißt *μ–fast überall stetig*, wenn es eine μ–Nullmenge N gibt mit $\Omega \setminus \Omega_h \subseteq N$ und h für alle $\omega \in \Omega \setminus N$ stetig in ω ist.
(1) Die Funktion $h : \Omega \to \mathbb{R}$ mit $h(\omega) := \chi_{\mathbb{Q}}(\omega)$ ist messbar und es gibt eine stetige Funktion $f : \Omega \to \mathbb{R}$ mit $\mu[\{h \neq f\}] = 0$. Dennoch ist h nicht μ–fast überall stetig.

(2) Die Funktion $h : \Omega \to \mathbb{R}$ mit $h(\omega) := \chi_{\mathbb{R}_+}(\omega)$ ist messbar und μ–fast überall stetig. Dennoch gibt es keine stetige Funktion $f : \Omega \to \mathbb{R}$ mit $\mu[\{h \neq f\}] = 0$.

Für $g : \Omega_g \to \bar{\mathbb{R}}$ und $h : \Omega_h \to \bar{\mathbb{R}}$ mit $\Omega_g, \Omega_h \subseteq \Omega$ schreiben wir

$$g =_\mu h$$

wenn es eine Funktion $f \in \mathcal{L}^0(\mathcal{F})$ und eine μ–Nullmenge N gibt mit

$$(\Omega \setminus \Omega_g) \cup \{\omega \in \Omega_g \mid g(\omega) \neq f(\omega)\} \subseteq N$$
$$(\Omega \setminus \Omega_h) \cup \{\omega \in \Omega_h \mid h(\omega) \neq f(\omega)\} \subseteq N$$

Dann gilt $g =_\mu f$ und $h =_\mu f$.

7.2.4 Beispiele. Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$.

(1) Für die Funktionen $h : \Omega_h \to \mathbb{R}$ mit $\Omega_h := \Omega \setminus \{0\}$ und

$$h(\omega) := 1/\omega$$

und $f : \Omega \to \mathbb{R}$ mit

$$f(\omega) := \begin{cases} 0 & \text{falls } \omega = 0 \\ 1/\omega & \text{sonst} \end{cases}$$

gilt $h =_\mu f$. Die Funktion f ist reell und messbar.

(2) **Dirichlet–Funktion:** Für die Dirichlet–Funktion $h : \Omega \to \mathbb{R}$ mit

$$h(\omega) := \begin{cases} \infty & \text{falls } \omega \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

und die Funktion $f : \Omega \to \mathbb{R}$ mit

$$f(\omega) := 0$$

gilt $h =_\mu f$. Die Funktion f ist reell und messbar.

Wir betrachten nun die Familie

$$\mathcal{L}^0(\mathcal{F}, \mu) := \left\{ h : \Omega_h \to \bar{\mathbb{R}} \;\middle|\; \Omega_h \subseteq \Omega \text{ und es gibt ein } f \in \mathcal{L}^0(\mathcal{F}) \text{ mit } h =_\mu f \right\}$$

Nach Lemma 7.2.1 ist die Vereinigung von zwei μ–Nullmengen eine μ–Nullmenge. Daher ist $=_\mu$ eine Äquivalenzrelation auf $\mathcal{L}^0(\mathcal{F}, \mu)$.

Für $h \in \mathcal{L}^0(\mathcal{F}, \mu)$ bezeichnen wir die Äquivalenzklasse, die die Funktion h enthält, mit $[h]_\mu$; es gilt also $[h]_\mu \subseteq \mathcal{L}^0(\mathcal{F}, \mu)$. Nach Definition von $=_\mu$ enthält jede Äquivalenzklasse $[h]_\mu \subseteq \mathcal{L}^0(\mathcal{F}, \mu)$ einen Repräsentanten $f \in \mathcal{L}^0(\mathcal{F})$, also eine messbare Funktion $f : \Omega \to \mathbb{R}$ mit $[h]_\mu = [f]_\mu$. Da nach Lemma 7.2.1 sogar die Vereinigung einer abzählbaren Familie von μ–Nullmengen eine μ–Nullmenge ist, gibt es zu jeder abzählbaren Familie $\{h_i\}_{i \in I} \subseteq \mathcal{L}^0(\mathcal{F}, \mu)$ mit $h_i : \Omega_{h_i} \to \bar{\mathbb{R}}$ eine gemeinsame μ–Nullmenge $N \in \mathcal{F}$ und eine Familie $\{f_i\}_{i \in I} \subseteq \mathcal{L}^0(\mathcal{F})$

derart, dass für alle $i \in I$ sowohl $\Omega \setminus \Omega_{h_i} \subseteq N$ als auch $h_i(\omega) = f_i(\omega)$ für alle $\omega \in \Omega \setminus N$ gilt. Bei der Betrachtung einer abzählbaren Familie $\{[f_i]_\mu\}_{i\in I}$ von Äquivalenzklassen in $\mathcal{L}^0(\mathcal{F},\mu)$ können und werden wir daher annehmen, dass $\{f_i\}_{i\in I} \subseteq \mathcal{L}^0(\mathcal{F})$ gilt.

Sei nun

$$L^0(\mathcal{F},\mu) := \mathcal{L}^0(\mathcal{F},\mu)/_{=_\mu}$$

Dann ist $L^0(\mathcal{F},\mu)$ die Familie aller Äquivalenzklassen von $\mathcal{L}^0(\mathcal{F},\mu)$ unter $=_\mu$. Die Addition, die Skalarmultiplikation und die Ordnungsrelation auf $\mathcal{L}^0(\mathcal{F})$ induzieren in kanonischer Weise eine Addition, eine Skalarmultiplikation und eine Ordnungsrelation auf $L^0(\mathcal{F},\mu)$. Mit Satz 7.1.7 erhalten wir daher das folgende Ergebnis:

7.2.5 Satz. *$L^0(\mathcal{F},\mu)$ ist ein Vektorverband.*

Sofern keine Missverständnisse zu befürchten sind, verzichten wir darauf, die Elemente von $L^0(\mathcal{F},\mu)$ als Äquivalenzklassen von $\mathcal{L}^0(\mathcal{F},\mu)$ zu kennzeichnen. Wir schreiben also $f \in L^0(\mathcal{F},\mu)$ anstelle von $[f] \in L^0(\mathcal{F},\mu)$ und nehmen dabei an, dass der Repräsentant f der Äquivalenzklasse $[f]$ auf ganz Ω definiert und reell ist.

Da der Limes einer μ–fast überall konvergenten Folge reeller Funktionen μ–fast überall eindeutig bestimmt ist, lässt sich diese Art der Konvergenz auch als Konvergenzbegriff in $L^0(\mathcal{F},\mu)$ verstehen. Wir betrachten nun eine weitere Art der Konvergenz:

Eine Folge $\{f_n\}_{n\in\mathbb{N}}$ messbarer reeller Funktionen *konvergiert im Maß* μ gegen eine messbare reelle Funktion f, wenn für alle $\varepsilon \in (0,\infty)$

$$\lim_{n\to\infty} \mu[\{|f_n - f| \geq \varepsilon\}] = 0$$

gilt. Das folgende Ergebnis zeigt, dass auch der Limes einer im Maß μ konvergenten Folge μ–fast überall eindeutig bestimmt ist:

7.2.6 Lemma. *Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge messbarer reeller Funktionen, die im Maß μ gegen eine messbare reelle Funktion f und gegen eine messbare reelle Funktion g konvergiert. Dann gilt $f(\omega) = g(\omega)$ μ–fast überall.*

Beweis. Wegen $|f-g| \leq |f-f_n| + |f_n - g|$ gilt für alle $\varepsilon \in (0,\infty)$

$$\{|f-g| \geq \varepsilon\} \subseteq \{|f-f_n| \geq \varepsilon/2\} \cup \{|f_n - g| \geq \varepsilon/2\}$$

und damit

$$\mu[\{|f-g| \geq \varepsilon\}] \leq \mu[\{|f-f_n| \geq \varepsilon/2\}] + \mu[\{|f_n - g| \geq \varepsilon/2\}]$$

und aus der Voraussetzung folgt nun

$$\mu[\{|f-g| \geq \varepsilon\}] = 0$$

Wegen $\{f \neq g\} = \{|f-g| > 0\} = \bigcup_{k\in\mathbb{N}}\{|f-g| \geq 1/k\}$ ergibt sich daraus

$$\mu[\{f \neq g\}] = \mu\left[\bigcup_{k\in\mathbb{N}}\{|f-g| \geq 1/k\}\right] \leq \sum_{k=1}^{\infty} \mu[\{|f-g| \geq 1/k\}] = 0$$

und damit $f(\omega) = g(\omega)$ fast überall. □

Aufgrund des letzten Lemmas lässt sich auch die Konvergenz im Maß μ als Konvergenzbegriff in $L^0(\mathcal{F},\mu)$ verstehen. Das folgende Lemma zeigt, dass die Konvergenz im Maß μ mit der linearen Struktur von $L^0(\mathcal{F},\mu)$ verträglich ist:

7.2.7 Lemma. *Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren reellen Funktionen, die im Maß μ gegen eine messbare reelle Funktion f konvergiert, und sei $\{g_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren reellen Funktionen, die im Maß μ gegen eine messbare reelle Funktion g konvergiert. Dann konvergiert für jede Wahl von $a, b \in \mathbb{R}$ die Folge $\{af_n+bg_n\}_{n\in\mathbb{N}}$ im Maß μ gegen $af+bg$.*

Beweis. Im Fall $a = 0$ oder $b = 0$ ist nichts zu zeigen. Sei nun $a \neq 0 \neq b$. Wegen $|(af_n+bg_n) - (af+bg)| \leq |a||f_n-f| + |b||g_n-g|$ gilt für alle $\varepsilon \in (0,\infty)$

$$\{|(af_n+bg_n) - (af+bg)| \geq \varepsilon\} \subseteq \{|a||f_n-f| \geq \varepsilon/2\} \cup \{|b||g_n-g| \geq \varepsilon/2\}$$

und damit

$$\begin{aligned}
&\mu[\{|(af_n+bg_n) - (af+bg)| \geq \varepsilon\}] \\
&\leq \mu[\{|f_n-f| \geq \varepsilon/|2a|\}] + \mu[\{|g_n-g| \geq \varepsilon/|2b|\}]
\end{aligned}$$

Daraus folgt die Behauptung. □

Wir vergleichen nun die Konvergenz im Maß mit der Konvergenz fast überall:

7.2.8 Satz. *Sei μ endlich. Dann konvergiert jede Folge von messbaren reellen Funktionen, die μ–fast überall gegen eine messbare reelle Funktion f konvergiert, auch im Maß μ gegen f.*

Beweis. Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren reellen Funktionen, die fast überall gegen eine messbare reelle Funktion f konvergiert. Sei $\varepsilon \in (0,\infty)$. Für alle $m \in \mathbb{N}$ sei

$$A_m := \bigcap_{n\in\mathbb{N}(m)} \{|f_n-f| < \varepsilon\}$$

Dann ist $\{A_m\}_{m\in\mathbb{N}}$ eine monoton wachsende Folge in $\mathcal{F}$ und nach Voraussetzung gibt es eine Nullmenge $N \in \mathcal{F}$ mit $\Omega = N \cup \bigcup_{m\in\mathbb{N}} A_m = \bigcup_{m\in\mathbb{N}}(N\cup A_m)$. Daher gilt

$$\sup_{m\in\mathbb{N}} \mu[A_m] = \sup_{m\in\mathbb{N}} \mu[N \cup A_m] = \mu[\Omega]$$

Da μ endlich ist, folgt daraus

$$\inf_{m\in\mathbb{N}} \mu[\Omega \setminus A_m] = 0$$

Wegen $\{|f_m - f| \geq \varepsilon\} \subseteq \Omega \setminus A_m$ erhalten wir nun

$$\lim_{m\to\infty} \mu[\{|f_m - f| \geq \varepsilon\}] = 0$$

und damit die Behauptung. □

Die folgenden Beispiele zeigen, dass man im letzten Satz auf die Forderung der Endlichkeit des Maßes nicht verzichten kann und dass auch im Fall eines endlichen Maßes aus der Konvergenz im Maß nicht die Konvergenz fast überall folgt:

7.2.9 Beispiele.

(1) Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$. Dann ist μ σ–endlich, aber nicht endlich. Für $n \in \mathbb{N}$ sei $f_n : \Omega \to \mathbb{R}$ gegeben durch

$$f_n(\omega) := \frac{1}{n}\, e^{\omega}$$

Dann konvergiert die Folge $\{f_n\}_{n\in\mathbb{N}}$ fast überall gegen 0. Andererseits gilt für alle $\varepsilon \in (0,\infty)$

$$\mu[\{f_n \geq \varepsilon\}] = \mu[\,[\log(n\varepsilon), \infty)\,] = \infty$$

Daher konvergiert die Folge $\{f_n\}_{n\in\mathbb{N}}$ nicht im Maß gegen 0.

(2) Sei $\Omega := (0,1]$, sei $\mathcal{F}$ die kleinste σ–Algebra auf $(0,1]$, die alle Intervalle der Form $((k-1)2^{-m}, k2^{-m}]$ mit $m \in \mathbb{N}_0$ und $k \in \{1,\ldots,2^m\}$ enthält, und sei $\mu : \mathcal{F} \to [0,\infty]$ gegeben durch $\mu[A] := \boldsymbol{\lambda}[A]$. Dann ist μ endlich. Für $m \in \mathbb{N}_0$ und $k \in \{1,\ldots,2^m\}$ sei f_{2^m+k-1} gegeben durch

$$f_{2^m+k-1}(\omega) := \chi_{((k-1)2^{-m}, k2^{-m}]}(\omega)$$

Dann gilt für alle $\varepsilon \in (0,\infty)$

$$\mu[\{f_{2^m+k-1} \geq \varepsilon\}] \leq \mu[\,((k-1)2^{-m}, k2^{-m}]\,] = 2^{-m}$$

Daher konvergiert die Folge $\{f_n\}_{n\in\mathbb{N}}$ im Maß gegen 0, aber sie ist nicht fast überall konvergent.

Sei nun

$$L^\infty(\mathcal{F},\mu) := \Big\{[h]_\mu \in L^0(\mathcal{F},\mu) \;\Big|\; \text{es gibt ein } f \in [h]_\mu \cap \mathcal{L}^\infty(\mathcal{F})\Big\}$$

Da $\mathcal{L}^\infty(\mathcal{F})$ ein Vektorverband und ein Ideal in $\mathcal{L}^0(\mathcal{F})$ ist, ist auch $L^\infty(\mathcal{F},\mu)$ ein Vektorverband und ein Ideal in $L^0(\mathcal{F},\mu)$.

Des weiteren gilt für alle $f, g \in \mathcal{L}^\infty(\mathcal{F})$ mit $f =_\mu g$

$$\inf\{c \in \mathbb{R}_+ \mid |f| \leq_\mu c\} = \inf\{c \in \mathbb{R}_+ \mid |g| \leq_\mu c\}$$

Daher ist die Abbildung $\|\,.\,\|_\infty : L^\infty(\mathcal{F}, \mu) \to \mathbb{R}_+$ mit

$$\|[f]_\mu\|_\infty := \inf\{c \in \mathbb{R}_+ \mid |f| \leq_\mu c\}$$

wohldefiniert; im folgenden schreiben wir $\|f\|_\infty$ anstelle von $\|[f]_\mu\|_\infty$. Es ist klar, dass $\|\,.\,\|_\infty$ eine Norm ist. Das folgende Lemma zeigt, dass der normierte Raum $(L^\infty(\mathcal{F}, \mu), \|\,.\,\|_\infty)$ vollständig ist:

7.2.10 Lemma (Riesz/Fischer). $(L^\infty(\mathcal{F}, \mu), \|\,.\,\|_\infty)$ *ist vollständig.*

Beweis. Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Cauchy–Folge in $(L^\infty(\mathcal{F}, \mu), \|\,.\,\|_\infty)$. Wir können annehmen, dass $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^\infty(\mathcal{F})$ gilt.
Da die Folge $\{f_n\}_{n\in\mathbb{N}}$ eine Cauchy–Folge in $(L^\infty(\mathcal{F}, \mu), \|\,.\,\|_\infty)$ ist, ist für jedes $\omega \in \Omega$ die Folge $\{f_n(\omega)\}_{n\in\mathbb{N}}$ eine Cauchy–Folge und damit in $\mathbb{R}$ konvergent. Da die Funktion $f : \Omega \to \mathbb{R}$ mit

$$f(\omega) := \lim_{n\to\infty} f_n(\omega)$$

messbar ist, erhalten wir zunächst $f \in \mathcal{L}^0(\mathcal{F})$. Außerdem gilt für alle $\omega \in \Omega$

$$\begin{aligned} |f(\omega)| &\leq |f(\omega) - f_n(\omega)| + |f_n(\omega) - f_m(\omega)| + |f_m(\omega)| \\ &\leq |f(\omega) - f_n(\omega)| + \|f_n - f_m\|_\infty + \|f_m\|_\infty \end{aligned}$$

Für $m \in \mathbb{N}$ hinreichend groß gilt $\|f_n - f_m\|_\infty \leq 1$ für alle $n \in \mathbb{N}(m)$, und für $n \in \mathbb{N}(m)$ hinreichend groß gilt $|f(\omega) - f_n(\omega)| \leq 1$. Daher gilt $f \in \mathcal{L}^\infty(\mathcal{F})$.
Sei nun $\varepsilon \in (0, \infty)$. Dann gibt es ein $m \in \mathbb{N}$ mit $\|f_n - f_m\|_\infty \leq \varepsilon$ für alle $n \in \mathbb{N}(m)$, und für alle $\omega \in \Omega$ gilt

$$|f(\omega) - f_m(\omega)| = \lim_{n\to\infty} |f_n(\omega) - f_m(\omega)| \leq \lim_{n\to\infty} \|f_n - f_m\|_\infty \leq \varepsilon$$

Daraus ergibt sich

$$\|f - f_m\|_\infty \leq \sup_{\omega\in\Omega} |f(\omega) - f_m(\omega)| \leq \varepsilon$$

Da $\varepsilon \in (0, \infty)$ beliebig war, erhalten wir $\lim_{m\to\infty} \|f - f_m\|_\infty = 0$. □

Nach Lemma 7.2.10 ist $(L^\infty(\mathcal{F}, \mu), \|\,.\,\|_\infty)$ ein Banach–Raum. Außerdem ist die Norm $\|\,.\,\|_\infty$ mit der Verbandsstruktur von $L^\infty(\mathcal{F}, \mu)$ in dem Sinne verträglich, dass für alle $f, g \in L^\infty(\mathcal{F}, \mu)$ mit $|f| \leq_\mu |g|$

$$\|\,|f|\,\|_\infty \leq \|\,|g|\,\|_\infty$$

gilt. Daher ist $(L^\infty(\mathcal{F}, \mu), \|\,.\,\|_\infty)$ ein normierter Vektorverband und aufgrund der Vollständigkeit sogar ein Banach–Verband. Der folgende Satz fasst die Eigenschaften von $(L^\infty(\mathcal{F}, \mu), \|\,.\,\|_\infty)$ zusammen:

7.2.11 Satz. *$(L^\infty(\mathcal{F},\mu), \|\,.\,\|_\infty)$ ist ein Banach–Verband und ein Ideal in $L^0(\mathcal{F},\mu)$.*

Im folgenden werden wir bei den in diesem Abschnitt eingeführten Begriffen und Bezeichnungen oft auf die Angabe der σ–Algebra $\mathcal{F}$ oder des Maßes μ verzichten.

Aufgaben

7.2.A Für $A, B \in \mathcal{F}$ sind äquivalent:
(a) Es gilt $A =_\mu B$.
(b) Es gilt $\chi_A =_\mu \chi_B$.

7.2.B Sei μ endlich.
(1) Die Abbildung $d_\mu : L^0(\mathcal{F},\mu) \times L^0(\mathcal{F},\mu) \to \mathbb{R}_+$ mit

$$d_\mu(f,g) := \mu[\{f \neq g\}]$$

ist (wohldefiniert und) eine Metrik und diese Metrik ist translationsinvariant, aber nicht absolut homogen.
(2) Für alle $A, B \in \mathcal{F}$ gilt $d_\mu(\chi_A, \chi_B) = \mu[A \triangle B]$.

7.2.C **Konvergenz im Maß:** Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren reellen Funktionen, die im Maß μ gegen eine messbare reelle Funktion f konvergiert, und sei $\{g_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren reellen Funktionen, die im Maß μ gegen eine messbare reelle Funktion g konvergiert. Dann konvergiert die Folge $\{f_n \vee g_n\}_{n\in\mathbb{N}}$ im Maß μ gegen $f \vee g$.

7.2.D **Lokale Konvergenz im Maß:** Eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von messbaren reellen Funktionen *konvergiert lokal im Maß* μ gegen eine messbare reelle Funktion f, wenn für jede Menge $C \in \mathcal{F}$ mit $\mu[C] < \infty$ und für alle $\varepsilon \in (0,\infty)$

$$\lim_{n\to\infty} \mu[C \cap \{|f_n - f| \geq \varepsilon\}] = 0$$

gilt.
(1) Sei μ σ–endlich. Dann ist der Limes einer lokal im Maß μ konvergenten Folge von messbaren reellen Funktionen μ–fast überall eindeutig bestimmt. Kann man auf die Forderung der σ–Endlichkeit von μ verzichten?
(2) Jede Folge von messbaren reellen Funktionen, die μ–fast überall gegen eine messbare reelle Funktion f konvergiert, konvergiert auch lokal im Maß μ gegen f.
(3) Jede Folge von messbaren reellen Funktionen, die im Maß μ gegen eine messbare reelle Funktion f konvergiert, konvergiert auch lokal im Maß μ gegen f. Ist μ endlich, so gilt auch die Umkehrung dieser Implikation.

7.2.E Sei $\{f_n\}_{n\in\mathbb{N}} \subseteq L^\infty(\mathcal{F},\mu)$ eine Folge, die in $L^\infty(\mathcal{F},\mu)$ gegen eine Funktion $f \in L^\infty(\mathcal{F},\mu)$ konvergiert. Dann konvergiert die Folge $\{f_n\}_{n\in\mathbb{N}}$ auch im Maß gegen f.

8

Lebesgue–Integral

In diesem Kapitel konstruieren wir ein Integral für messbare (numerische) Funktionen, die auf einem Maßraum $(\Omega, \mathcal{F}, \mu)$ definiert sind. Dieses Integral wird als *Lebesgue–Integral* bezeichnet.

Ausgangspunkt für die Konstruktion des Integrals ist die Interpretation des Maßes $\mu[A]$ einer Menge $A \in \mathcal{F}$ als Integral ihrer Indikatorfunktion χ_A; wegen $A \in \mathcal{F}$ ist χ_A messbar und wir setzen

$$\int_\Omega \chi_A \, d\mu := \mu[A]$$

Die Grundidee ist nun, dieses Integral zunächst auf Linearkombinationen von messbaren Indikatorfunktionen und sodann auf deren Limites unter der punktweisen Konvergenz fortzusetzen. Die Ausführung dieser Grundidee wird allerdings dadurch erschwert, dass das Maß μ im allgemeinen nicht endlich ist; man betrachte etwa den Maßraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$. Es zeigt sich jedoch, dass das Integral für messbare Indikatorfunktionen in natürlicher Weise zunächst zu einem Integral für alle positiven einfachen Funktionen (Abschnitt 8.1) und sodann zu einem Integral für alle positiven messbaren Funktionen (Abschnitt 8.2) fortgesetzt werden kann; da jede messbare Funktion als Differenz von zwei positiven messbaren Funktionen dargestellt werden kann, lässt sich das Integral schließlich auch für bestimmte messbare Funktionen definieren, die nicht positiv sein müssen (Abschnitt 8.3). Diese drei Schritte der Konstruktion des Integrals werden als *algebraische Induktion* bezeichnet; wir haben sie bereits im Beweis des Faktorisierungssatzes verwendet, und es wird sich zeigen, dass die algebraische Induktion auch eine grundlegende Technik zum Beweis von Aussagen über Integrale ist.

Abschließend betrachten wir Räume integrierbarer Funktionen (Abschnitt 8.4) und damit erste funktionalanalytische Aspekte der Integrationstheorie.

Im gesamten Kapitel sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum. Im folgenden bezeichnen wir eine numerische Funktion kurz als *Funktion*.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_8,

8.1 Positive einfache Funktionen

In diesem Abschnitt definieren wir das Lebesgue–Integral für positive einfache Funktionen.

Ausgangspunkt für die Konstruktion des Integrals ist die Interpretation des Maßes $\mu[A]$ einer Menge $A \in \mathcal{F}$ als Integral ihrer Indikatorfunktion χ_A: Für eine messbare Menge A setzen wir

$$\int_\Omega \chi_A \, d\mu := \mu[A]$$

und nennen $\int_\Omega \chi_A \, d\mu$ das *Lebesgue–Integral* bezüglich μ oder das μ*–Integral* oder kurz das *Integral* von χ_A.

Die Indikatorfunktion einer messbaren Menge ist eine positive einfache Funktion. Wir erweitern die Definition des Integrals nun auf beliebige positive einfache Funktionen:

Für eine positive einfache Funktion f mit der Standarddarstellung

$$f = \sum_{a \in f(\Omega)} a \, \chi_{\{f=a\}}$$

setzen wir

$$\int_\Omega f \, d\mu := \sum_{a \in f(\Omega)} a \, \mu[\{f=a\}]$$

und nennen $\int_\Omega f \, d\mu$ das *Lebesgue–Integral* bezüglich μ oder das μ*–Integral* oder kurz das *Integral* von f; vgl. Aufgabe 8.1.B. Die Definition des μ–Integrals für positive einfache Funktionen ist offensichtlich mit der Definition des μ–Integrals für Indikatorfunktionen von messbaren Mengen verträglich.

Das folgende Lemma liefert eine allgemeine Darstellung des Integrals einer positiven einfachen Funktion, die sich bei der Herleitung der Eigenschaften des Integrals als nützlich erweist und auch seine Berechnung erleichtern kann:

8.1.1 Lemma. *Sei f eine positive einfache Funktion und sei*

$$f = \sum_{i=1}^n c_i \, \chi_{C_i}$$

eine Darstellung von f mit $\{C_i\}_{i \in \{1,\dots,n\}} \subseteq \mathcal{F}$ und $\{c_i\}_{i \in \{1,\dots,n\}} \subseteq \mathbb{R}_+$. Dann gilt

$$\int_\Omega f \, d\mu = \sum_{i=1}^n c_i \, \mu[C_i]$$

Beweis. Wir betrachten die Standarddarstellung

$$f = \sum_{a \in f(\Omega)} a\, \chi_{\{f=a\}}$$

von f. Dann gilt

$$\sum_{a \in f(\Omega)} a\, \chi_{\{f=a\}} = \sum_{i=1}^{n} c_i\, \chi_{C_i}$$

und daraus ergibt sich für alle $a \in f(\Omega)$

$$a\, \chi_{\{f=a\}} = \sum_{i=1}^{n} c_i\, \chi_{C_i \cap \{f=a\}}$$

Es genügt zu zeigen, dass für alle $a \in f(\Omega)$ die Gleichung

$$a\, \mu[\{f=a\}] = \sum_{i=1}^{n} c_i\, \mu[C_i \cap \{f=a\}] \qquad (*)$$

gilt, denn dann ergibt sich aus $\sum_{a \in f(\Omega)} \{f=a\} = \Omega$ und der Additivität von μ sofort

$$\begin{aligned}
\sum_{a \in f(\Omega)} a\, \mu[\{f=a\}] &= \sum_{a \in f(\Omega)} \sum_{i=1}^{n} c_i\, \mu[C_i \cap \{f=a\}] \\
&= \sum_{i=1}^{n} c_i \sum_{a \in f(\Omega)} \mu[C_i \cap \{f=a\}] \\
&= \sum_{i=1}^{n} c_i\, \mu\left[\sum_{a \in f(\Omega)} C_i \cap \{f=a\}\right] \\
&= \sum_{i=1}^{n} c_i\, \mu[C_i]
\end{aligned}$$

und damit die Behauptung des Lemmas.
Zum Beweis der Gleichung $(*)$ zeigen wir, dass für alle $n \in \mathbb{N}$ und für jede Wahl von $B, D_1, \ldots, D_n \in \mathcal{F}$ und $b, d_1, \ldots, d_n \in \mathbb{R}_+$ mit

$$b\, \chi_B = \sum_{i=1}^{n} d_i\, \chi_{D_i}$$

die Gleichung

$$b\, \mu[B] = \sum_{i=1}^{n} d_i\, \mu[D_i]$$

gilt. Wir führen den Beweis durch vollständige Induktion:

- $n = 1$: Wir betrachten $B, D_1 \in \mathcal{F}$ und $b, d_1 \in \mathbb{R}_+$ mit
$$b\, \chi_B = d_1\, \chi_{D_1}$$
Im Fall $d_1 = 0$ oder $D_1 = \emptyset$ gilt $b = 0$ oder $B = \emptyset$, und damit
$$b\, \mu[B] = d_1\, \mu[D_1]$$
Im Fall $d_1 \neq 0$ und $D_1 \neq \emptyset$ gilt $b = d_1$ und $B = D_1$, und damit ebenfalls
$$b\, \mu[B] = d_1\, \mu[D_1]$$
Damit ist die Behauptung für $n = 1$ gezeigt.
- $n \to n+1$: Wir nehmen an, die Behauptung sei für n bereits bewiesen, und betrachten $B, D_1, \dots, D_n, D_{n+1} \in \mathcal{F}$ und $b, d_1, \dots, d_n, d_{n+1} \in \mathbb{R}_+$ mit
$$b\, \chi_B = \sum_{i=1}^{n+1} d_i\, \chi_{D_i}$$
Im Fall $d_{n+1} = 0$ oder $D_{n+1} = \emptyset$ gilt $b\, \chi_B = \sum_{i=1}^{n} d_i\, \chi_{D_i}$, und damit
$$\begin{aligned} b\, \mu[B] &= \sum_{i=1}^{n} d_i\, \mu[D_i] \\ &= \sum_{i=1}^{n+1} d_i\, \mu[D_i] \end{aligned}$$
Im Fall $d_{n+1} \neq 0$ und $D_{n+1} \neq \emptyset$ gilt $D_{n+1} \subseteq B$ und $0 < d_{n+1} \leq b$. Aus $D_{n+1} \subseteq B$ folgt zunächst $b\, \chi_{D_{n+1}} = \sum_{i=1}^{n} d_i\, \chi_{D_i \cap D_{n+1}} + d_{n+1}\, \chi_{D_{n+1}}$ und damit $(b - d_{n+1})\, \chi_{D_{n+1}} = \sum_{i=1}^{n} d_i\, \chi_{D_i \cap D_{n+1}}$, und aus $b - d_{n+1} \in \mathbb{R}_+$ folgt nun
$$(b - d_{n+1})\, \mu[D_{n+1}] = \sum_{i=1}^{n} d_i\, \mu[D_i \cap D_{n+1}]$$
Im Fall $\mu[D_{n+1}] < \infty$ erhalten wir aus der letzten Gleichung
$$b\, \mu[D_{n+1}] = \sum_{i=1}^{n} d_i\, \mu[D_i \cap D_{n+1}] + d_{n+1}\, \mu[D_{n+1}]$$
und wegen $0 < d_{n+1} \leq b$ gilt diese Gleichung auch im Fall $\mu[D_{n+1}] = \infty$. Außerdem gilt $b\, \chi_{B \setminus D_{n+1}} = \sum_{i=1}^{n} d_i\, \chi_{D_i \setminus D_{n+1}}$, und daraus folgt
$$b\, \mu[B \setminus D_{n+1}] = \sum_{i=1}^{n} d_i\, \mu[D_i \setminus D_{n+1}]$$

Durch Summation ergibt sich nun

$$\begin{aligned} b\,\mu[B] &= b\,\mu[D_{n+1}] + b\,\mu[B \setminus D_{n+1}] \\ &= \sum_{i=1}^{n} d_i\,\mu[D_i \cap D_{n+1}] + d_{n+1}\,\mu[D_{n+1}] + \sum_{i=1}^{n} d_i\,\mu[D_i \setminus D_{n+1}] \\ &= \sum_{i=1}^{n+1} d_i\,\mu[D_i] \end{aligned}$$

Damit ist die Behauptung bewiesen. □

Mit Hilfe des letzten Lemmas können wir nun leicht die elementaren Eigenschaften des Integrals für positive einfache Funktionen beweisen:

8.1.2 Lemma. *Seien f und g positive einfache Funktionen und sei $a \in \mathbb{R}_+$. Dann gilt*

$$\int_\Omega (f+g)\,d\mu = \int_\Omega f\,d\mu + \int_\Omega g\,d\mu$$

und

$$\int_\Omega af\,d\mu = a\int_\Omega f\,d\mu$$

Im Fall $f \le g$ gilt

$$\int_\Omega f\,d\mu \le \int_\Omega g\,d\mu$$

Insbesondere gilt $\int_\Omega f\,d\mu \ge 0$.

Beweis. Mit f und g sind auch $f+g$ und af positiv und einfach. Ausgehend von den Standarddarstellungen von f und g gewinnen wir Darstellungen der Form

$$\begin{aligned} f &= \sum_{i=1}^{n} c_i\,\chi_{B_i} \\ g &= \sum_{i=1}^{n} d_i\,\chi_{B_i} \end{aligned}$$

mit einer disjunkten Familie $\{B_i\}_{i\in\{1,\dots,n\}} \subseteq \mathcal{F}$ sowie $\{c_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}_+$ und $\{d_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}_+$. Dann besitzen $f+g$ und af die Darstellungen

$$\begin{aligned} f+g &= \sum_{i=1}^{n} (c_i+d_i)\,\chi_{B_i} \\ af &= \sum_{i=1}^{n} a\,c_i\,\chi_{B_i} \end{aligned}$$

und aus Lemma 8.1.1 ergibt sich nun

$$\begin{aligned}\int_\Omega (f+g)\,d\mu &= \sum_{i=1}^n (c_i+d_i)\,\mu[B_i] \\ &= \sum_{i=1}^n c_i\,\mu[B_i] + \sum_{i=1}^n d_i\,\mu[B_i] \\ &= \int_\Omega f\,d\mu + \int_\Omega g\,d\mu\end{aligned}$$

und

$$\begin{aligned}\int_\Omega af\,d\mu &= \sum_{i=1}^n a\,c_i\,\mu[B_i] \\ &= a\sum_{i=1}^n c_i\,\mu[B_i] \\ &= a\int_\Omega f\,d\mu\end{aligned}$$

Im Fall $f \le g$ gilt $c_i \le d_i$ für alle $i \in \{1,\dots,n\}$ mit $B_i \ne \emptyset$, und daraus folgt

$$\begin{aligned}\int_\Omega f\,d\mu &= \sum_{i=1}^n c_i\,\mu[B_i] \\ &\le \sum_{i=1}^n d_i\,\mu[B_i] \\ &= \int_\Omega g\,d\mu\end{aligned}$$

Die letzte Behauptung des Lemmas ist ohnehin klar. □

Lemma 8.1.2 besagt, dass das Integral als Abbildung von der Familie der positiven einfachen Funktionen in die Menge der erweiterten reellen Zahlen positiv linear und monoton ist.

Aufgaben

8.1.A Sei f eine positive einfache Funktion. Dann gilt

$$\int_\Omega f\,d\mu = \sum_{a\in f(\Omega)} \int_\Omega a\,\chi_{\{f=a\}}\,d\mu$$

8.1.B Ist μ endlich, so lässt sich die hier für positive einfache Funktionen gegebene Definition des Integrals auf beliebige einfache Funktionen erweitern.

8.1.C Was ergibt sich aus Lemma 8.1.2 für $f := \chi_A$ und $g := \chi_B$ mit $A, B \in \mathcal{F}$, wenn $A \cap B = \emptyset$ bzw. $A \subseteq B$ gilt?

8.2 Positive messbare Funktionen

In diesem Abschnitt erweitern wir das Lebesgue–Integral auf beliebige positive messbare Funktionen.

Für eine positive messbare Funktion f setzen wir

$$\int_\Omega f\,d\mu := \sup\left\{\int_\Omega h\,d\mu \;\middle|\; h \text{ ist einfach mit } 0 \le h \le f\right\}$$

und nennen $\int_\Omega f\,d\mu$ das *Lebesgue–Integral* bezüglich μ oder das *μ–Integral* oder kurz das *Integral* von f. Die Definition des μ–Integrals für positive messbare Funktionen ist offensichtlich mit der Definition des μ–Integrals für positive einfache Funktionen verträglich.

Der folgende Satz über die monotone Konvergenz ist das zentrale Ergebnis der Integrationstheorie:

8.2.1 Satz (Monotone Konvergenz; Levi). *Sei $\{f_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge von positiven messbaren Funktionen. Dann gilt*

$$\int_\Omega \sup_{n\in\mathbb{N}} f_n\,d\mu = \sup_{n\in\mathbb{N}} \int_\Omega f_n\,d\mu$$

Beweis. Wir bemerken zunächst, dass die Funktion $\sup_{n\in\mathbb{N}} f_n$ messbar ist. Für alle $k \in \mathbb{N}$ gilt $f_k \le \sup_{n\in\mathbb{N}} f_n$ und aus der Definition des Integrals für positive messbare Funktionen ergibt sich sofort

$$\int_\Omega f_k\,d\mu \le \int_\Omega \sup_{n\in\mathbb{N}} f_n\,d\mu$$

Daher gilt

$$\sup_{k\in\mathbb{N}} \int_\Omega f_k\,d\mu \le \int_\Omega \sup_{n\in\mathbb{N}} f_n\,d\mu$$

Zum Beweis der umgekehrten Ungleichung betrachten wir eine einfache Funktion h mit $0 \le h \le \sup_{n\in\mathbb{N}} f_n$ sowie ein $\alpha \in (0,1)$. Für $k \in \mathbb{N}$ sei

$$D_k := \{\alpha h \le f_k\}$$

Dann gilt $D_k \in \mathcal{F}$ und mit h sind auch die Funktionen $h\chi_{D_k}$ und $\alpha h\chi_{D_k}$ positiv und einfach. Insbesondere ergibt sich aus der Darstellung

$$h = \sum_{i=1}^m c_i\,\chi_{C_i}$$

von h mit $\{C_i\}_{i\in\{1,\dots,n\}} \subseteq \mathcal{F}$ und $\{c_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}_+$ die Darstellung

$$h\chi_{D_k} = \sum_{i=1}^m c_i \, \chi_{C_i \cap D_k}$$

von $h\chi_{D_k}$. Da die Folge $\{D_k\}_{k\in\mathbb{N}}$ monoton wachsend ist mit $\bigcup_{k\in\mathbb{N}} D_k = \Omega$ und da jedes Maß stetig von unten ist, erhalten wir zunächst

$$\begin{aligned}\int_\Omega h \, d\mu &= \sum_{i=1}^m c_i \, \mu[C_i] \\ &= \sum_{i=1}^m c_i \sup_{k\in\mathbb{N}} \mu[C_i \cap D_k] \\ &= \sup_{k\in\mathbb{N}} \sum_{i=1}^m c_i \, \mu[C_i \cap D_k] \\ &= \sup_{k\in\mathbb{N}} \int_\Omega h\chi_{D_k} \, d\mu\end{aligned}$$

Wegen $0 \leq \alpha h \chi_{D_k} \leq f_k$ erhalten wir außerdem aufgrund der Definition des Integrals von f_k

$$\int_\Omega \alpha h \chi_{D_k} \, d\mu \leq \int_\Omega f_k \, d\mu$$

Unter Verwendung von Lemma 8.1.2 erhalten wir daher

$$\begin{aligned}\alpha \int_\Omega h \, d\mu &= \alpha \sup_{k\in\mathbb{N}} \int_\Omega h\chi_{D_k} \, d\mu \\ &= \sup_{k\in\mathbb{N}} \int_\Omega \alpha h\chi_{D_k} \, d\mu \\ &\leq \sup_{k\in\mathbb{N}} \int_\Omega f_k \, d\mu\end{aligned}$$

Durch Bildung des Supremums über $\alpha \in (0,1)$ ergibt sich nun

$$\int_\Omega h \, d\mu \leq \sup_{k\in\mathbb{N}} \int_\Omega f_k \, d\mu$$

und durch Bildung des Supremums über alle einfachen Funktionen h mit $0 \leq h \leq \sup_{n\in\mathbb{N}} f_n$ ergibt sich sodann aufgrund der Definition des Integrals von $\sup_{n\in\mathbb{N}} f_n$

$$\int_\Omega \sup_{n\in\mathbb{N}} f_n \, d\mu \leq \sup_{k\in\mathbb{N}} \int_\Omega f_k \, d\mu$$

Damit ist der Satz bewiesen. $\square$

Nach dem Approximationssatz ist eine positive Funktion f genau dann messbar, wenn es eine monoton wachsende Folge $\{f_n\}_{n\in\mathbb{N}}$ von positiven einfachen Funktionen gibt mit $f = \sup_{n\in\mathbb{N}} f_n$; das folgende Ergebnis zeigt, dass für jede Wahl der approximierenden Folge das Integral von f mit dem Supremum der Integrale der approximierenden Funktionen f_n übereinstimmt. Dies ist eine unmittelbare Folgerung aus dem Satz über die monotone Konvergenz.

8.2.2 Folgerung. *Sei f eine positive messbare Funktion und sei $\{f_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge von positiven einfachen Funktionen mit $f = \sup_{n\in\mathbb{N}} f_n$. Dann gilt*

$$\int_\Omega f\,d\mu = \sup_{n\in\mathbb{N}} \int_\Omega f_n\,d\mu$$

Das letzte Ergebnis erweist sich bei der Herleitung weiterer Eigenschaften des Integrals als nützlich und kann auch seine Berechnung erleichtern. Insbesondere können wir nun zeigen, dass das Integral einer positiven messbaren Funktion dieselben elementaren Eigenschaften besitzt wie das Integral einer positiven einfachen Funktion:

8.2.3 Lemma. *Seien f und g positive messbare Funktionen und sei $a \in \mathbb{R}_+$. Dann gilt*

$$\int_\Omega (f+g)\,d\mu = \int_\Omega f\,d\mu + \int_\Omega g\,d\mu$$

und

$$\int_\Omega af\,d\mu = a\int_\Omega f\,d\mu$$

Im Fall $f \leq g$ gilt

$$\int_\Omega f\,d\mu \leq \int_\Omega g\,d\mu$$

Insbesondere gilt $\int_\Omega f\,d\mu \geq 0$.

Beweis. Seien $\{f_n\}_{n\in\mathbb{N}}$ und $\{g_n\}_{n\in\mathbb{N}}$ monoton wachsende Folgen von positiven einfachen Funktionen mit

$$\begin{aligned} f &= \sup_{n\in\mathbb{N}} f_n \\ g &= \sup_{n\in\mathbb{N}} g_n \end{aligned}$$

Dann sind auch $\{f_n+g_n\}_{n\in\mathbb{N}}$ und $\{af_n\}_{n\in\mathbb{N}}$ monoton wachsende Folgen von positiven einfachen Funktionen und es gilt

$$\begin{aligned} f+g &= \sup_{n\in\mathbb{N}} (f_n+g_n) \\ af &= \sup_{n\in\mathbb{N}} af_n \end{aligned}$$

Aus Folgerung 8.2.2 und Lemma 8.1.2 folgt nun

$$\begin{aligned}\int_\Omega (f+g)\,d\mu &= \sup_{n\in\mathbb{N}} \int_\Omega (f_n+g_n)\,d\mu \\ &= \sup_{n\in\mathbb{N}} \biggl(\int_\Omega f_n\,d\mu + \int_\Omega g_n\,d\mu \biggr) \\ &= \sup_{n\in\mathbb{N}} \int_\Omega f_n\,d\mu + \sup_{n\in\mathbb{N}} \int_\Omega g_n\,d\mu \\ &= \int_\Omega f\,d\mu + \int_\Omega g\,d\mu\end{aligned}$$

und

$$\begin{aligned}\int_\Omega af\,d\mu &= \sup_{n\in\mathbb{N}} \int_\Omega af_n\,d\mu \\ &= a \sup_{n\in\mathbb{N}} \int_\Omega f_n\,d\mu \\ &= a \int_\Omega f\,d\mu\end{aligned}$$

Die übrigen Behauptungen sind klar. □

Lemma 8.2.3 besagt, dass das Integral als Abbildung von der Familie der positiven messbaren Funktionen in die Menge der erweiterten reellen Zahlen positiv linear und monoton ist.

Der Satz über die monotone Konvergenz lässt sich nun auch als Aussage über das Integral einer Reihe von positiven messbaren Funktionen formulieren:

8.2.4 Folgerung (Monotone Konvergenz; Levi). *Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von positiven messbaren Funktionen. Dann gilt*

$$\int_\Omega \biggl(\sum_{n=1}^\infty f_n \biggr) d\mu = \sum_{n=1}^\infty \int_\Omega f_n\,d\mu$$

Beweis. Die Folge der Partialsummen $\sum_{k=1}^n f_k$ ist eine monoton wachsende Folge von positiven messbaren Funktionen. Aus dem Satz über die monotone Konvergenz und der Additivität des Integrals folgt nun

$$\begin{aligned}\int_\Omega \biggl(\sum_{k=1}^\infty f_k \biggr) d\mu &= \int_\Omega \biggl(\sup_{n\in\mathbb{N}} \sum_{k=1}^n f_k \biggr) d\mu \\ &= \sup_{n\in\mathbb{N}} \int_\Omega \biggl(\sum_{k=1}^n f_k \biggr) d\mu\end{aligned}$$

$$
\begin{aligned}
&= \sup_{n\in\mathbb{N}} \sum_{k=1}^{n} \int_\Omega f_k \, d\mu \\
&= \sum_{k=1}^{\infty} \int_\Omega f_k \, d\mu
\end{aligned}
$$

Damit ist die Folgerung bewiesen. □

Wir beweisen nun einige weitere Ergebnisse über das Integral einer positiven messbaren Funktion. Ein gleichzeitig elementares und zentrales Ergebnis, das vielfältige Anwendungen besitzt, ist die folgende Ungleichung von Markov, die wir im Hinblick auf ihre Anwendungen für den Betrag einer beliebigen messbaren Funktion formulieren:

8.2.5 Lemma (Ungleichung von Markov). *Sei f eine messbare Funktion. Dann gilt für alle $c \in (0,\infty)$*

$$\mu[\{|f| \geq c\}] \leq \frac{1}{c} \int_\Omega |f| \, d\mu$$

Beweis. Unter Verwendung von Lemma 8.2.3 ergibt sich

$$\mu[\{|f| \geq c\}] = \int_\Omega \chi_{\{|f|\geq c\}} \, d\mu \leq \int_\Omega \frac{1}{c} \, |f| \, \chi_{\{|f|\geq c\}} \, d\mu \leq \frac{1}{c} \int_\Omega |f| \, d\mu$$

Damit ist das Lemma bewiesen. □

Lemma 8.2.5 ist die Basisversion der Ungleichung von Markov. Aus ihr lässt sich eine ganze Schar von Ungleichungen ableiten:

8.2.6 Folgerung (Ungleichung von Markov). *Sei f eine messbare Funktion. Dann gilt für jede monoton wachsende Funktion $h : \mathbb{R}_+ \to \mathbb{R}_+$ und für alle $c \in \mathbb{R}_+$ mit $h(c) > 0$*

$$\mu[\{|f| \geq c\}] \leq \frac{1}{h(c)} \int_\Omega h \circ |f| \, d\mu$$

Beweis. Aus Lemma 8.2.5 ergibt sich

$$\mu[\{|f| \geq c\}] \leq \mu[\{h \circ |f| \geq h(c)\}] \leq \frac{1}{h(c)} \int_\Omega h \circ |f| \, d\mu$$

Damit ist die Folgerung bewiesen. □

Aus der Ungleichung von Markov folgt zunächst, dass jede messbare Funktion, für die das Integral ihres Betrages endlich ist, fast überall endlich ist:

8.2.7 Lemma. *Sei f eine messbare Funktion mit*

$$\int_\Omega |f| \, d\mu < \infty$$

Dann ist f μ–fast überall endlich.

Beweis. Aus der Ungleichung von Markov folgt für alle $n \in \mathbb{N}$

$$\mu[\{|f| = \infty\}] \leq \mu[\{|f| \geq n\}] \leq \frac{1}{n} \int_\Omega |f| \, d\mu$$

und aus der Voraussetzung $\int_\Omega |f| \, d\mu < \infty$ folgt nun $\mu[\{|f|{=}\infty\}] = 0$. □

Aus der Ungleichung von Markov folgt ferner, dass eine messbare Funktion genau dann fast überall verschwindet, wenn das Integral ihres Betrages verschwindet; dieses Ergebnis ist gleichzeitig eine Anwendung des Satzes über die monotone Konvergenz:

8.2.8 Lemma. *Sei f eine messbare Funktion. Dann sind folgende Aussagen äquivalent*:
(a) *Es gilt $f(\omega) = 0$ μ–fast überall.*
(b) *Es gilt $\int_\Omega |f| \, d\mu = 0$.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Für $n \in \mathbb{N}$ sei $g_n : \Omega \to \mathbb{R}$ gegeben durch

$$g_n(\omega) := n \, \chi_{\{f \neq 0\}}(\omega)$$

Dann ist $\{g_n\}_{n \in \mathbb{N}}$ eine monoton wachsende Folge von positiven messbaren Funktionen mit $|f| \leq \sup_{n \in \mathbb{N}} g_n$ und für alle $n \in \mathbb{N}$ gilt

$$\int_\Omega g_n \, d\mu = n \, \mu[\{f \neq 0\}] = 0$$

Aus dem Satz über die monotone Konvergenz folgt nun

$$\int_\Omega |f| \, d\mu \leq \int_\Omega \sup_{n \in \mathbb{N}} g_n \, d\mu = \sup_{n \in \mathbb{N}} \int_\Omega g_n \, d\mu = 0$$

und damit $\int_\Omega |f| \, d\mu = 0$. Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Aus der Ungleichung von Markov folgt für alle $n \in \mathbb{N}$

$$\mu[\{|f| \geq 1/n\}] \leq n \int_\Omega |f| \, d\mu = 0$$

Daraus ergibt sich

$$\mu[\{|f| > 0\}] = \sup_{n \in \mathbb{N}} \mu[\{|f| \geq 1/n\}] = 0$$

und damit $\mu[\{|f| \neq 0\}] = 0$. Daher folgt (a) aus (b). □

Die letzten beiden Ergebnisse sind typische Beispiele dafür, wie sich eine fast überall bestehende Eigenschaft einer positiven messbaren Funktion aus einer Eigenschaft ihres Integrals ergibt. Außerdem zeigt das folgende Beispiel, dass die Ausnahmemenge nicht leer sein muss. Dies ist letztlich der Grund dafür, überhaupt fast überall bestehende Eigenschaften zu betrachten.

8.2.9 Beispiel (Dirichlet–Funktion). Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$. Für die Dirichlet–Funktion $f : \Omega \to \bar{\mathbb{R}}$ mit

$$f(\omega) := \begin{cases} \infty & \text{falls } \omega \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

gilt $\mu[\{f \neq 0\}] = \mu[\mathbb{Q}] = 0$ und damit $f(\omega) = 0$ μ–fast überall. Aus Lemma 8.2.8 folgt $\int_\Omega f \, d\mu = 0$ und damit $\int_\Omega f \, d\mu < \infty$, aber es gilt weder $f = 0$ noch $f < \infty$.

Andererseits zeigen die folgenden Ergebnisse, dass bei der Integration positiver messbaren Funktionen Nullmengen vernachlässigt werden können:

8.2.10 Lemma. *Seien f und g positive messbare Funktionen mit $f(\omega) \leq g(\omega)$ μ–fast überall. Dann gilt*

$$\int_\Omega f \, d\mu \leq \int_\Omega g \, d\mu$$

Beweis. Nach Voraussetzung gilt $(f \chi_{\{f>g\}})(\omega) = 0$ fast überall und aus Lemma 8.2.8 folgt $\int_\Omega f \chi_{\{f>g\}} \, d\mu = 0$. Außerdem gilt $f \chi_{\{f \leq g\}} \leq g$. Unter Verwendung von Lemma 8.2.3 erhalten wir

$$\int_\Omega f \, d\mu = \int_\Omega f \chi_{\{f \leq g\}} \, d\mu + \int_\Omega f \chi_{\{f>g\}} \, d\mu = \int_\Omega f \chi_{\{f \leq g\}} \, d\mu \leq \int_\Omega g \, d\mu$$

Damit ist das Lemma bewiesen. □

Insbesondere sind die Integrale von positiven messbaren Funktionen, die sich nur auf einer Nullmenge unterscheiden, identisch:

8.2.11 Folgerung. *Seien f und g positive messbare Funktionen mit $f(\omega) = g(\omega)$ μ–fast überall. Dann gilt*

$$\int_\Omega f \, d\mu = \int_\Omega g \, d\mu$$

Dieses Ergebnis legt es nahe, die Definition des Integrals wie folgt zu erweitern: Für eine Funktion h, die μ–fast überall definiert ist und μ–fast überall mit einer positiven messbaren Funktion f übereinstimmt, setzen wir

$$\int_\Omega h \, d\mu := \int_\Omega f \, d\mu$$

und nennen $\int_\Omega h\,d\mu$ das *Lebesgue–Integral* bezüglich μ oder das μ*–Integral* oder kurz das *Integral* von h. Nach Folgerung 8.2.11 ist das μ–Integral von h unabhängig von der Wahl der positiven messbaren Funktion f mit $h =_\mu f$. Alle Ergebnisse dieses Abschnitts gelten auch für diese Verallgemeinerung des μ–Integrals. Insbesondere bleibt der Satz über die monotone Konvergenz auch dann gültig, wenn man annimmt, dass die Folge $\{f_n\}_{n\in\mathbb{N}}$ nur μ–fast überall monoton wachsend ist.

Wir beschließen diesen Abschnitt mit einer Ungleichung, die auf dem Satz über die monotone Konvergenz beruht und die Grundlage für einen weiteren Konvergenzsatz bildet, den wir im nächsten Abschnitt beweisen:

8.2.12 Lemma (Fatou). *Sei* $\{f_n\}_{n\in\mathbb{N}}$ *eine Folge von positiven messbaren Funktionen. Dann gilt*

$$\int_\Omega \liminf_{n\to\infty} f_n\,d\mu \le \liminf_{n\to\infty} \int_\Omega f_n\,d\mu$$

Beweis. Für $n \in \mathbb{N}$ sei

$$g_n := \inf_{k\in\mathbb{N}(n)} f_k$$

Dann ist $\{g_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge von positiven messbaren Funktionen mit $g_n \le f_n$ für alle $n \in \mathbb{N}$ und aus dem Satz über die monotone Konvergenz ergibt sich

$$\begin{aligned}
\int_\Omega \liminf_{n\to\infty} f_n\,d\mu &= \int_\Omega \sup_{n\in\mathbb{N}} \inf_{k\in\mathbb{N}(n)} f_k\,d\mu \\
&= \int_\Omega \sup_{n\in\mathbb{N}} g_n\,d\mu \\
&= \sup_{n\in\mathbb{N}} \int_\Omega g_n\,d\mu \\
&= \lim_{n\to\infty} \int_\Omega g_n\,d\mu \\
&= \liminf_{n\to\infty} \int_\Omega g_n\,d\mu \\
&\le \liminf_{n\to\infty} \int_\Omega f_n\,d\mu
\end{aligned}$$

Damit ist das Lemma bewiesen. □

Das folgende Beispiel zeigt, dass man im Lemma von Fatou die Ungleichung im allgemeinen nicht durch eine Gleichung ersetzen kann:

8.2.13 Beispiel. Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$. Für alle $n \in \mathbb{N}$ sei f_n gegeben durch

$$f_n(\omega) := n\, \chi_{(0,1/n]}(\omega)$$

Dann gilt $\lim_{n\to\infty} f_n = 0$ und damit $\int_\Omega \liminf_{n\to\infty} f_n \, d\mu = 0$. Andererseits gilt $\liminf_{n\to\infty} \int_\Omega f_n \, d\mu = 1$.

Aufgaben

8.2.A Sei $(\Omega, \mathcal{F})$ ein Messraum und $\{\mu_n\}_{n\in\mathbb{N}}$ eine Folge von Maßen $\mathcal{F} \to [0, \infty]$. Sei ferner $\mu : \mathcal{F} \to [0, \infty]$ gegeben durch

$$\mu[A] := \sum_{n=1}^{\infty} \mu_n[A]$$

Dann ist μ ein Maß und für jede positive messbare Funktion f gilt

$$\int_\Omega f \, d\mu = \sum_{n=1}^{\infty} \int_\Omega f \, d\mu_n$$

8.2.B **Monotone Konvergenz:** Was ergibt sich aus Folgerung 8.2.4 im Fall $f_n := \chi_{A_n}$ mit einer disjunkten Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$?

8.2.C **Ungleichung von Markov:** Zu jedem $c \in (0, \infty)$ gibt es eine messbare Funktion f mit

$$\mu[\{|f| \geq c\}] = \frac{1}{c} \int_\Omega |f| \, d\mu$$

8.2.D **Lemma von Fatou:** Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von positiven messbaren Funktionen mit $\int_\Omega \sup_{n\in\mathbb{N}} f_n \, d\mu < \infty$. Dann gilt

$$\limsup_{n\to\infty} \int_\Omega f_n \, d\mu \leq \int_\Omega \limsup_{n\to\infty} f_n \, d\mu$$

8.2.E Sei $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$. Dann gilt

$$\mu\left[\liminf_{n\to\infty} A_n\right] \leq \liminf_{n\to\infty} \mu[A_n]$$

Im Fall $\mu[\bigcup_{n\in\mathbb{N}} A_n] < \infty$ gilt auch

$$\limsup_{n\to\infty} \mu[A_n] \leq \mu\left[\limsup_{n\to\infty} A_n\right]$$

8.2.F Konstruieren Sie einen endlichen Maßraum $(\Omega, \mathcal{F}, \mu)$ und eine positive messbare Funktion $f : \Omega \to \bar{\mathbb{R}}$ derart, dass f μ–fast überall endlich ist und $\int_\Omega f \, d\mu = \infty$ gilt.

8.3 Integrierbare Funktionen

In diesem Abschnitt erweitern wir die Definition des Lebesgue–Integrals auf bestimmte messbare Funktionen, die nicht positiv sein müssen.

Eine messbare Funktion f heißt *μ–quasiintegrierbar* oder kurz *quasiintegrierbar*, wenn

$$\min\left\{\int_\Omega f^+ \, d\mu, \int_\Omega f^- \, d\mu\right\} < \infty$$

gilt; in diesem Fall setzen wir

$$\int_\Omega f \, d\mu := \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu$$

und nennen $\int_\Omega f \, d\mu$ das *Lebesgue–Integral* bezüglich μ oder das *μ–Integral* oder kurz das *Integral* von f. Die Definition des μ–Integrals für μ–quasiintegrierbare Funktionen ist mit der Definition des μ–Integrals für positive messbare Funktionen verträglich, denn für jede positive messbare Funktion f gilt $f^+ = f$ und $f^- = 0$. Ist f μ–quasiintegrierbar, so ist nach Lemma 8.2.7 f^+ oder f^- μ–fast überall endlich.

Eine messbare Funktion f heißt *Lebesgue–integrierbar* bezüglich μ oder *μ–integrierbar* oder kurz *integrierbar*, wenn

$$\max\left\{\int_\Omega f^+ \, d\mu, \int_\Omega f^- \, d\mu\right\} < \infty$$

gilt. Das folgende Lemma charakterisiert die μ–Integrierbarkeit einer messbaren Funktion:

8.3.1 Lemma. *Sei f eine messbare Funktion. Dann sind folgende Aussagen äquivalent*:
(a) *f ist μ–integrierbar.*
(b) *Es gilt*

$$\int_\Omega |f| \, d\mu < \infty$$

(c) *Es gibt positive messbare Funktionen g und h mit $f = g - h$ und*

$$\max\left\{\int_\Omega g \, d\mu, \int_\Omega h \, d\mu\right\} < \infty$$

In diesem Fall ist f μ–fast überall endlich.

Beweis. Es gilt $|f| = f^+ + f^-$ und aus Lemma 8.2.3 folgt

$$\int_\Omega |f|\,d\mu = \int_\Omega f^+\,d\mu + \int_\Omega f^-\,d\mu$$

Daher sind (a) und (b) äquivalent, und es ist klar, dass (c) aus (a) folgt. Wir nehmen nun an, dass (c) gilt, und betrachten positive messbare Funktionen g und h mit $f = g - h$ und

$$\max\left\{\int_\Omega g\,d\mu, \int_\Omega h\,d\mu\right\} < \infty$$

Wegen $f = g - h \leq g$ gilt $f^+ \leq g$ und wegen $-f = h - g \leq h$ gilt $f^- \leq h$. Aus Lemma 8.2.3 folgt nun

$$\int_\Omega f^+\,d\mu \leq \int_\Omega g\,d\mu$$
$$\int_\Omega f^-\,d\mu \leq \int_\Omega h\,d\mu$$

und damit

$$\max\left\{\int_\Omega f^+\,d\mu, \int_\Omega f^-\,d\mu\right\} < \infty$$

Damit ist gezeigt, dass (a) aus (c) folgt. Die letzte Aussage folgt aus Lemma 8.2.7. □

Aus dem letzten Lemma folgt, dass eine positive messbare Funktion nicht notwendigerweise integrierbar ist; dies gilt sogar für positive einfache Funktionen:

8.3.2 Beispiel. Sei f gegeben durch

$$f(\omega) := 1$$

Dann ist f messbar und positiv, und folgende Aussagen sind äquivalent:
(a) μ ist endlich.
(b) f ist μ–integrierbar.

Offenbar ist eine positive messbare Funktion genau dann integrierbar, wenn ihr Integral endlich ist. Daher lässt sich das letzte Lemma auch wie folgt formulieren:

8.3.3 Folgerung. *Sei f eine messbare Funktion. Dann sind folgende Aussagen äquivalent:*
(a) *f ist μ–integrierbar.*
(b) *$|f|$ ist μ–integrierbar.*
(c) *f^+ und f^- sind μ–integrierbar.*
(d) *Es gibt positive μ–integrierbare Funktionen g und h mit $f = g - h$.*

Das folgende Lemma liefert eine allgemeine Darstellung des Integrals einer integrierbaren Funktion, die sich bei der Herleitung der Eigenschaften des Integrals als nützlich erweist und auch seine Berechnung erleichtern kann:

8.3.4 Lemma. *Sei f eine μ–integrierbare Funktion. Dann gilt für jede Wahl von positiven μ–integrierbaren Funktionen g und h mit $f = g - h$*

$$\int_\Omega f \, d\mu = \int_\Omega g \, d\mu - \int_\Omega h \, d\mu$$

Beweis. Seien g und h positive integrierbare Funktionen mit $f = g - h$. Dann gilt $f^+ - f^- = f = g - h$ und damit $f^+ + h = g + f^-$, und aus Lemma 8.2.3 folgt nun

$$\int_\Omega f^+ \, d\mu + \int_\Omega h \, d\mu = \int_\Omega g \, d\mu + \int_\Omega f^- \, d\mu$$

Da alle Integrale endlich sind, folgt daraus

$$\int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu = \int_\Omega g \, d\mu - \int_\Omega h \, d\mu$$

Damit ist das Lemma bewiesen. □

Das nächste Lemma liefert eine hinreichende Bedingung für die Integrierbarkeit einer messbaren Funktion:

8.3.5 Lemma. *Seien f und g messbare Funktionen mit $|f(\omega)| \le |g(\omega)|$ μ–fast überall. Ist g μ–integrierbar, so ist auch f μ–integrierbar.*

Beweis. Nach Lemma 8.2.10 gilt $\int_\Omega |f| \, d\mu \le \int_\Omega |g| \, d\mu$. □

Wie im Fall positiver messbarer Funktionen sind auch die Integrale integrierbarer Funktionen, die sich nur auf einer Nullmenge unterscheiden, identisch:

8.3.6 Lemma. *Seien f und g messbare Funktionen mit $f(\omega) = g(\omega)$ μ–fast überall. Dann sind folgende Aussagen äquivalent:*
(a) *f ist μ–integrierbar.*
(b) *g ist μ–integrierbar.*
In diesem Fall gilt

$$\int_\Omega f \, d\mu = \int_\Omega g \, d\mu$$

Beweis. Es gilt $f^+(\omega) = g^+(\omega)$ fast überall und $f^-(\omega) = g^-(\omega)$ fast überall, und aus Folgerung 8.2.11 ergibt sich nun $\int_\Omega f^+ \, d\mu = \int_\Omega g^+ \, d\mu$ und $\int_\Omega f^- \, d\mu = \int_\Omega g^- \, d\mu$. □

Wir beweisen nun ein Analogon zu Lemma 8.2.3:

8.3.7 Lemma. *Seien f und g μ–integrierbare reelle Funktionen und sei $a \in \mathbb{R}$.*
(1) *Die Funktion $f+g$ ist μ–integrierbar und es gilt*

$$\int_\Omega (f+g)\,d\mu = \int_\Omega f\,d\mu + \int_\Omega g\,d\mu$$

(2) *Die Funktion af ist μ–integrierbar und es gilt*

$$\int_\Omega af\,d\mu = a \int_\Omega f\,d\mu$$

(3) *Im Fall $f \le g$ gilt*

$$\int_\Omega f\,d\mu \le \int_\Omega g\,d\mu$$

(4) *Die Funktionen $f \vee g$ und $f \wedge g$ sind μ–integrierbar.*

Beweis. Wegen $|f+g| \le |f| + |g|$ ist $f+g$ nach Lemma 8.3.5 integrierbar und wegen $f = f^+ - f^-$ und $g = g^+ - g^-$ gilt $f+g = (f^+ + g^+) - (f^- + g^-)$. Aus Lemma 8.3.4 folgt nun

$$\begin{aligned}\int_\Omega (f+g)\,d\mu &= \int_\Omega (f^+ + g^+)\,d\mu - \int_\Omega (f^- + g^-)\,d\mu \\ &= \left(\int_\Omega f^+\,d\mu + \int_\Omega g^+\,d\mu\right) - \left(\int_\Omega f^-\,d\mu + \int_\Omega g^-\,d\mu\right) \\ &= \left(\int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu\right) + \left(\int_\Omega g^+\,d\mu - \int_\Omega g^-\,d\mu\right) \\ &= \int_\Omega f\,d\mu + \int_\Omega g\,d\mu\end{aligned}$$

Damit ist (1) gezeigt. Nach Lemma 8.2.3 gilt

$$\begin{aligned}\int_\Omega |af|\,d\mu &= \int_\Omega |a||f|\,d\mu \\ &= |a| \int_\Omega |f|\,d\mu\end{aligned}$$

Daher ist af integrierbar. Wegen $f = f^+ - f^-$ gilt $-f = f^- - f^+$ und damit

$$\begin{aligned}\int_\Omega (-f)\,d\mu &= \int_\Omega f^-\,d\mu - \int_\Omega f^+\,d\mu \\ &= -\left(\int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu\right) \\ &= -\int_\Omega f\,d\mu\end{aligned}$$

Im Fall $a \in \mathbb{R}_+$ gilt $af = (af)^+ - (af)^- = af^+ - af^-$ und aus Lemma 8.2.3 folgt nun

$$\begin{aligned}\int_\Omega af\,d\mu &= \int_\Omega af^+\,d\mu - \int_\Omega af^-\,d\mu \\ &= a\int_\Omega f^+\,d\mu - a\int_\Omega f^-\,d\mu \\ &= a\biggl(\int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu\biggr) \\ &= a\int_\Omega f\,d\mu\end{aligned}$$

Im Fall $a \in (-\infty, 0)$ gilt $a = -|a|$ und aus dem bisher gezeigten folgt

$$\begin{aligned}\int_\Omega af\,d\mu &= \int_\Omega (-|a|)f\,d\mu \\ &= -\int_\Omega |a|f\,d\mu \\ &= -|a|\int_\Omega f\,d\mu \\ &= a\int_\Omega f\,d\mu\end{aligned}$$

Damit ist (2) gezeigt. Im Fall $f \leq g$ gilt $f^+ \leq g^+$ und $g^- \leq f^-$, und damit

$$\begin{aligned}\int_\Omega f^+\,d\mu &\leq \int_\Omega g^+\,d\mu \\ \int_\Omega g^-\,d\mu &\leq \int_\Omega f^-\,d\mu\end{aligned}$$

Daraus folgt

$$\begin{aligned}\int_\Omega f\,d\mu &= \int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu \\ &\leq \int_\Omega g^+\,d\mu - \int_\Omega g^-\,d\mu \\ &= \int_\Omega g\,d\mu\end{aligned}$$

Damit ist (3) gezeigt. Wegen $|f \vee g| \leq |f| + |g|$ und $|f \wedge g| \leq |f| + |g|$ sind $|f \vee g|$ und $|f \wedge g|$ nach Lemma 8.3.5 integrierbar. Damit ist (4) gezeigt. □

Aus Lemma 8.3.7 ergibt sich ein Spezialfall der Ungleichung von Jensen:

8.3.8 Folgerung (Ungleichung von Jensen). *Sei f μ–integrierbar. Dann gilt*

$$\left|\int_\Omega f\,d\mu\right| \leq \int_\Omega |f|\,d\mu$$

Beweis. Nach Lemma 8.3.7 gilt $\int_\Omega f\,d\mu \leq \int_\Omega |f|\,d\mu$ und $-\int_\Omega f\,d\mu = \int_\Omega (-f)\,d\mu \leq \int_\Omega |f|\,d\mu$. Daraus folgt die Behauptung. □

Wir beweisen nun einen Konvergenzsatz für integrierbare Funktionen:

8.3.9 Satz (Majorisierte Konvergenz; Lebesgue). *Sei* $\{f_n\}_{n\in\mathbb{N}}$ *eine Folge von messbaren reellen Funktionen, die gegen eine reelle Funktion* f *konvergiert. Wenn* $\sup_{n\in\mathbb{N}} |f_n|$ μ*–integrierbar ist, dann ist auch* f μ*–integrierbar und es gilt*

$$\lim_{n\to\infty} \int_\Omega |f_n - f| \, d\mu = 0$$

Insbesondere ist jede der Funktionen f_n μ*–integrierbar und es gilt*

$$\int_\Omega \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_\Omega f_n \, d\mu$$

Beweis. Als Limes einer konvergenten Folge von messbaren Funktionen ist f messbar. Sei nun

$$g := \sup_{n\in\mathbb{N}} |f_n|$$

Nach Voraussetzung ist g integrierbar. Damit ist auch jede der Funktionen f_n integrierbar, und wegen $|f| \le g$ ist auch f integrierbar. Aus der Dreiecksungleichung ergibt sich nun $|f_n - f| \le |f_n| + |f| \le 2g$ und damit

$$2g - |f_n - f| \ge 0$$

Nach Voraussetzung gilt

$$\lim_{n\to\infty} \Big(2g - |f_n - f|\Big) = 2g$$

und aus dem Lemma von Fatou folgt nun

$$\begin{aligned}
\int_\Omega 2g \, d\mu &= \int_\Omega \lim_{n\to\infty} \Big(2g - |f_n - f|\Big) \, d\mu \\
&= \int_\Omega \liminf_{n\to\infty} \Big(2g - |f_n - f|\Big) \, d\mu \\
&\le \liminf_{n\to\infty} \int_\Omega \Big(2g - |f_n - f|\Big) \, d\mu \\
&= \int_\Omega 2g \, d\mu + \liminf_{n\to\infty} \int_\Omega \Big(-|f_n - f|\Big) \, d\mu \\
&= \int_\Omega 2g \, d\mu - \limsup_{n\to\infty} \int_\Omega |f_n - f| \, d\mu
\end{aligned}$$

Da g integrierbar ist, erhalten wir daraus

$$\limsup_{n\to\infty} \int_\Omega |f_n - f| \, d\mu \le 0$$

und damit $\lim_{n\to\infty} \int_\Omega |f_n - f| \, d\mu = 0$. Die abschließende Behauptung folgt nun aus der Ungleichung von Jensen. □

Die Voraussetzungen des Satzes über die majorisierte Konvergenz sind insbesondere dann erfüllt, wenn eine monoton wachsende Folge von positiven messbaren reellen Funktionen vorliegt, deren Supremum integrierbar ist. Das folgende Beispiel zeigt, dass man im Satz über die majorisierte Konvergenz auf die Forderung der Integrierbarkeit von $\sup_{n\in\mathbb{N}} |f_n|$ nicht verzichten kann:

8.3.10 Beispiel. Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$. Für $n \in \mathbb{N}$ sei f_n gegeben durch

$$f_n(\omega) := n\, \chi_{(0,1/n]}(\omega)$$

Dann gilt $\lim_{n\to\infty} f_n = 0$. Andererseits gilt $\int_\Omega f_n \, d\mu = 1$ für alle $n \in \mathbb{N}$, und damit

$$\lim_{n\to\infty} \int_\Omega f_n \, d\mu \neq \int_\Omega \lim_{n\to\infty} f_n \, d\mu$$

Daraus folgt, dass $\sup_{n\in\mathbb{N}} f_n$ nicht μ–integrierbar ist.

Wie der Satz über die monotone Konvergenz liefert auch der Satz über die majorisierte Konvergenz einen Konvergenzsatz für unendliche Reihen:

8.3.11 Folgerung (Majorisierte Konvergenz; Lebesgue). *Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren reellen Funktionen derart, dass die Reihe $\sum_{n=1}^\infty f_n$ gegen eine reelle Funktion konvergiert. Wenn $\sup_{n\in\mathbb{N}} |\sum_{k=1}^n f_k|$ μ–integrierbar ist, dann ist auch die Reihe $\sum_{n=1}^\infty f_n$ sowie jede der Funktionen f_n μ–integrierbar und es gilt*

$$\int_\Omega \left(\sum_{n=1}^\infty f_n \right) d\mu = \sum_{n=1}^\infty \int_\Omega f_n \, d\mu$$

Beweis. Aus dem Satz über die majorisierte Konvergenz folgt zunächst die Integrierbarkeit der Reihe $\sum_{n=1}^\infty f_n$ und

$$\int_\Omega \left(\sum_{k=1}^\infty f_k \right) d\mu = \int_\Omega \left(\lim_{n\to\infty} \sum_{k=1}^n f_k \right) d\mu = \lim_{n\to\infty} \int_\Omega \left(\sum_{k=1}^n f_k \right) d\mu$$

Wegen $f_n = \sum_{k=1}^n f_k - \sum_{k=1}^{n-1} f_k$ sind auch alle f_n integrierbar, und aus der Linearität des Integrals folgt nun

$$\int_\Omega \left(\sum_{k=1}^\infty f_k \right) d\mu = \lim_{n\to\infty} \int_\Omega \left(\sum_{k=1}^n f_k \right) d\mu = \lim_{n\to\infty} \sum_{k=1}^n \int_\Omega f_k \, d\mu = \sum_{k=1}^\infty \int_\Omega f_k \, d\mu$$

Damit ist die Folgerung bewiesen. □

Wir untersuchen die im Satz über die majorisierte Konvergenz auftretende Art der Konvergenz etwas näher. Eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von μ–integrierbaren reellen Funktionen *konvergiert im Mittel* (bezüglich μ) gegen eine messbare reelle Funktion f, wenn

$$\lim_{n\to\infty} \int_\Omega |f_n - f| \, d\mu = 0$$

gilt; wegen

$$\int_\Omega |f| \, d\mu \le \int_\Omega |f - f_n| \, d\mu + \int_\Omega |f_n| \, d\mu$$

ist dann auch f μ–integrierbar.

Der Satz über die majorisierte Konvergenz liefert also eine hinreichende Bedingung dafür, dass eine Folge von integrierbaren reellen Funktionen, die gegen eine reelle Funktion f konvergiert, auch im Mittel gegen f konvergiert. Wir vergleichen nun die Konvergenz im Mittel mit der Konvergenz im Maß:

8.3.12 Satz. *Jede Folge von μ–integrierbaren reellen Funktionen, die im Mittel gegen eine μ–integrierbare reelle Funktion f konvergiert, konvergiert auch im Maß μ gegen f.*

Beweis. Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von integrierbaren reellen Funktionen, die im Mittel gegen eine reelle Funktion f konvergiert. Für alle $\varepsilon \in (0,\infty)$ folgt aus der Ungleichung von Markov

$$\mu[\{|f_n - f| \ge \varepsilon\}] \le \frac{1}{\varepsilon} \int_\Omega |f_n - f| \, d\mu$$

und daraus folgt die Behauptung. □

Lemma 8.3.6 legt es nahe, den Begriff der μ–Integrierbarkeit und damit auch die Definition des μ–Integrals wie folgt zu erweitern: Eine Funktion h, die μ–fast überall definiert ist und μ–fast überall mit einer μ–integrierbaren Funktion f übereinstimmt, heißt *Lebesgue–integrierbar* oder *μ–integrierbar* oder kurz *integrierbar*, und in diesem Fall setzen wir

$$\int_\Omega h \, d\mu := \int_\Omega f \, d\mu$$

und nennen $\int_\Omega h \, d\mu$ das *Lebesgue–Integral* oder das *μ–Integral* oder kurz das *Integral* von h. Nach Lemma 8.3.6 ist das μ–Integral von h unabhängig von der Wahl der μ–integrierbaren Funktion f mit $h =_\mu f$, und nach Lemma 8.3.1 kann f reell gewählt werden. Mit dieser Erweiterung wird die μ–Integrierbarkeit zu einer Eigenschaft bestimmter Äquivalenzklassen von $\mathcal{L}^0(\mathcal{F},\mu)$ unter der Äquivalenzrelation $=_\mu$.

Alle Ergebnisse dieses Abschnitts gelten auch für diese Verallgemeinerung der μ–Integrierbarkeit und des μ–Integrals. Insbesondere kann die an einigen Stellen getroffene Annahme, dass bestimmte μ–integrierbare Funktionen reell sind, entfallen, und der Satz über die majorisierte Konvergenz bleibt auch dann gültig, wenn nur die Konvergenz μ–fast überall vorausgesetzt wird. Wir können daher bei der Betrachtung einer abzählbaren Familie von μ–integrierbaren Funktionen stets annehmen, dass alle Funktionen reell sind.

Im Hinblick auf die verschiedenen Arten der Konvergenz lässt sich zusammenfassend sagen, dass

– aus der Konvergenz fast überall unter Zusatzbedingungen die Konvergenz im Mittel und
– aus der Konvergenz im Mittel ohne Zusatzbedingungen die Konvergenz im Maß

folgt; andererseits zeigen die folgenden Beispiele, dass weder aus der Konvergenz im Maß die Konvergenz im Mittel noch aus der Konvergenz im Mittel die Konvergenz fast überall folgt:

8.3.13 Beispiele. Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$.

(1) Für $n \in \mathbb{N}$ sei f_n gegeben durch

$$f_n(\omega) := n^2 \chi_{(0,1/n]}(\omega)$$

Dann gilt für alle $\varepsilon \in (0, \infty)$

$$\mu[\{f_n > \varepsilon\}] \leq 1/n$$

Daher konvergiert die Folge $\{f_n\}_{n \in \mathbb{N}}$ im Maß gegen 0. Andererseits gilt

$$\int_\Omega f_n \, d\mu = n$$

Daher konvergiert die Folge $\{f_n\}_{n \in \mathbb{N}}$ nicht im Mittel.

(2) Für $m \in \mathbb{N}_0$ und $k \in \{1, \ldots, 2^m\}$ sei f_{2^m+k-1} gegeben durch

$$f_{2^m+k-1}(\omega) := \chi_{((k-1)2^{-m}, k2^{-m}]}(\omega)$$

Dann gilt

$$\int_\Omega f_{2^m+k-1} \, d\mu = \mu\Big[((k-1)2^{-m}, k2^{-m}]\Big] = 2^{-m}$$

Daher konvergiert die Folge $\{f_n\}_{n \in \mathbb{N}}$ im Mittel gegen 0, aber sie ist nicht fast überall konvergent.

Für eine μ–quasiintegrierbare Funktion f setzen wir

$$\int_\Omega f(\omega) \, d\mu(\omega) := \int_\Omega f \, d\mu$$

Die Angabe der Integrationsvariablen ist in vielen Fällen hilfreich. Dies gilt insbesondere für die folgenden Anwendungen des Satzes über die majorisierte Konvergenz:

8.3.14 Lemma (Stetigkeitslemma). *Sei $(a, b) \subseteq \mathbb{R}$ ein nichtleeres Intervall, sei $x_0 \in (a, b)$ und sei $f : (a, b) \times \Omega \to \mathbb{R}$ eine Funktion mit den folgenden Eigenschaften:*
– *Für alle $x \in (a, b)$ ist die Funktion $\omega \mapsto f(x, \omega)$ messbar.*
– *Für alle $\omega \in \Omega$ ist die Funktion $x \mapsto f(x, \omega)$ stetig in x_0.*
– *Es gibt eine μ–integrierbare Funktion $h : \Omega \to \mathbb{R}$ mit*

$$\sup_{x \in (a,b)} |f(x, \omega)| \leq h(\omega)$$

für alle $\omega \in \Omega$.

Dann ist die Funktion $F : (a, b) \to \mathbb{R}$ mit

$$F(x) := \int_\Omega f(x, \omega)\, d\mu(\omega)$$

stetig in x_0.

Beweis. Sei $\{x_n\}_{n \in \mathbb{N}} \subseteq (a, b)$ eine Folge mit $\lim_{n \to \infty} x_n = x_0$. Dann ist für alle $n \in \mathbb{N}_0$ die Funktion $\omega \to f(x_n, \omega)$ messbar und für alle $\omega \in \Omega$ gilt $\lim_{n \to \infty} f(x_n, \omega) = f(x_0, \omega)$ und $\sup_{n \in \mathbb{N}} |f(x_n, \omega)| \leq h(\omega)$. Aus dem Satz über die majorisierte Konvergenz folgt nun

$$\lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} \int_\Omega f(x_n, \omega)\, d\mu(\omega) = \int_\Omega f(x_0, \omega)\, d\mu(\omega) = F(x_0)$$

Damit ist das Lemma bewiesen. □

Ein vergleichbares Ergebnis erhält man für die Vertauschung von Differentiation und Integration:

8.3.15 Lemma (Differentiationslemma). *Sei $(a, b) \subseteq \mathbb{R}$ ein nichtleeres Intervall und sei $f : (a, b) \times \Omega \to \mathbb{R}$ eine Funktion mit den folgenden Eigenschaften:*
– *Für alle $x \in (a, b)$ ist die Funktion $\omega \mapsto f(x, \omega)$ μ–integrierbar.*
– *Für alle $\omega \in \Omega$ ist die Funktion $x \mapsto f(x, \omega)$ differenzierbar.*
– *Es gibt eine μ–integrierbare Funktion $h : \Omega \to \mathbb{R}$ mit*

$$\sup_{x \in (a,b)} \left| \frac{df}{dx}(x, \omega) \right| \leq h(\omega)$$

für alle $\omega \in \Omega$.

Dann ist die Funktion $F : (a, b) \to \mathbb{R}$ mit

$$F(x) := \int_\Omega f(x, \omega)\, d\mu(\omega)$$

differenzierbar und es gilt

$$\frac{dF}{dx}(x) = \int_\Omega \frac{df}{dx}(x, \omega)\, d\mu(\omega)$$

Beweis. Sei $x_0 \in (a,b)$ und sei $\{x_n\}_{n\in\mathbb{N}} \subseteq (a,b)$ eine Folge mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$. Dann ist für alle $n \in \mathbb{N}$ die Funktion $g_n : \Omega \to \mathbb{R}$ mit

$$g_n(\omega) := \frac{f(x_n,\omega) - f(x_0,\omega)}{x_n - x_0}$$

messbar und für alle $\omega \in \Omega$ gilt

$$\lim_{n\to\infty} g_n(\omega) = \frac{df}{dx}(x_0,\omega)$$

Nach dem Mittelwertsatz gibt es für alle $n \in \mathbb{N}$ und $\omega \in \Omega$ ein $z_n(\omega) \in (a,b)$ mit

$$g_n(\omega) = \frac{df}{dx}(z_n(\omega),\omega)$$

Daher gilt

$$\sup_{n\in\mathbb{N}} |g_n(\omega)| \leq h(\omega)$$

Aus dem Satz über die majorisierte Konvergenz folgt nun

$$\begin{aligned}
\int_\Omega \frac{df}{dx}(x_0,\omega)\, d\mu(\omega) &= \int_\Omega \lim_{n\to\infty} g_n(\omega)\, d\mu(\omega) \\
&= \lim_{n\to\infty} \int_\Omega g_n(\omega)\, d\mu(\omega) \\
&= \lim_{n\to\infty} \int_\Omega \frac{f(x_n,\omega) - f(x_0,\omega)}{x_n - x_0}\, d\mu(\omega) \\
&= \lim_{n\to\infty} \frac{F(x_n) - F(x_0)}{x_n - x_0} \\
&= \frac{dF}{dx}(x_0)
\end{aligned}$$

Damit ist das Lemma bewiesen. □

Aufgaben

8.3.A Quasiintegrierbare Funktionen:

(1) Sei f eine messbare Funktion. Dann sind folgende Aussagen äquivalent:
  - (a) f ist μ–quasiintegrierbar.
  - (b) f^+ oder f^- ist μ–integrierbar.
  - (c) Es gibt positive messbare Funktionen g und h mit $f = g - h$ derart, dass g oder h μ–integrierbar ist.

(2) Sei f eine μ–quasiintegrierbare Funktion. Dann gilt für jede Wahl von positiven messbaren Funktionen g und h mit $f = g - h$ und $\min\{\int_\Omega g\, d\mu, \int_\Omega h\, d\mu\} < \infty$

$$\int_\Omega f\, d\mu = \int_\Omega g\, d\mu - \int_\Omega h\, d\mu$$

(3) Seien f und g μ–quasiintegrierbare Funktionen, von denen mindestens eine μ–integrierbar ist. Dann ist $f+g$ μ–quasiintegrierbar und es gilt

$$\int_\Omega (f+g)\, d\mu = \int_\Omega f\, d\mu + \int_\Omega g\, d\mu$$

8.3.B **Satz von Pratt:** Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren Funktionen, die μ–fast überall gegen eine Funktion f konvergiert, und seien $\{g_n\}_{n\in\mathbb{N}}$ und $\{h_n\}_{n\in\mathbb{N}}$ Folgen von μ–integrierbaren Funktionen mit $g_n \le f_n \le h_n$ für alle $n \in \mathbb{N}$, die μ–fast überall gegen eine μ–integrierbare Funktion g bzw. h konvergieren. Gilt $\lim_{n\to\infty} \int_\Omega g_n\, d\mu = \int_\Omega g\, d\mu$ und $\lim_{n\to\infty} \int_\Omega h_n\, d\mu = \int_\Omega h\, d\mu$, so ist f μ–integrierbar und es gilt

$$\lim_{n\to\infty} \int_\Omega f_n\, d\mu = \int_\Omega f\, d\mu$$

Gilt außerdem $g_n \le 0 \le h_n$ für alle $n \in \mathbb{N}$, so gilt auch

$$\lim_{n\to\infty} \int_\Omega |f_n - f|\, d\mu = 0$$

8.3.C **Integrierbare komplexe Funktionen:** Eine Funktion $f : \Omega \to \mathbb{C}$ heißt *Lebesgue–integrierbar* oder *μ–integrierbar* oder kurz *integrierbar*, wenn die reellen Funktionen $\mathrm{Re}f$ und $\mathrm{Im}f$ μ–integrierbar sind; in diesem Fall setzen wir

$$\int_\Omega f\, d\mu := \int_\Omega \mathrm{Re}f\, d\mu + i \int_\Omega \mathrm{Im}f\, d\mu$$

und nennen $\int_\Omega f\, d\mu$ das *Lebesgue–Integral* oder das *μ–Integral* oder kurz das *Integral* von f. Die Definition des μ–Integrals für μ–integrierbare komplexe Funktionen ist mit der Definition des μ–Integrals für μ–integrierbare reelle Funktionen verträglich.

(1) Jede μ–integrierbare komplexe Funktion ist messbar.

(2) Sei f eine messbare komplexe Funktion. Dann sind folgende Aussagen äquivalent:
  (a) f ist μ–integrierbar.
  (b) $|f|$ ist μ–integrierbar.

(3) Die μ–integrierbaren komplexen Funktionen bilden einen Vektorraum.

(4) Sei f eine μ–integrierbare komplexe Funktion. Dann gilt

$$\left| \int_\Omega f\, d\mu \right| \le \int_\Omega |f|\, d\mu$$

Hinweis: Beweisen Sie die Ungleichung zunächst für den Fall, dass der Realteil und der Imaginärteil von f einfach ist, und sodann für den allgemeinen Fall unter Verwendung des Approximationssatzes.

8.4 L^p–Räume

Im gesamten Abschnitt sei $p \in [1, \infty)$. Eine messbare Funktion f heißt *p–fach μ–integrierbar* oder kurz *p–fach integrierbar*, wenn $|f|^p$ μ–integrierbar ist.

Wir betrachten die Familie

$$\mathcal{L}^p_\mu(\mathcal{F}) := \left\{ f \in \mathcal{L}^0(\mathcal{F}) \,\middle|\, \int_\Omega |f|^p \, d\mu < \infty \right\}$$

aller p–fach μ–integrierbaren reellen Funktionen und untersuchen zunächst die Eigenschaften von $\mathcal{L}^p_\mu(\mathcal{F})$ als Teilmenge des Vektorverbandes $\mathcal{L}^0(\mathcal{F})$. Wir benötigen das folgende Lemma:

8.4.1 Lemma. *Für alle $a, b \in \mathbb{R}$ gilt $|a+b| \le |a|+|b|$ und $|a \vee b| \le |a|+|b|$ sowie*

$$\Big(|a|+|b|\Big)^p \le 2^{p-1}\Big(|a|^p+|b|^p\Big)$$

Beweis. Die ersten Ungleichungen sind klar. Da die Funktion $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ mit $\varphi(x) := x^p$ konvex ist, gilt für alle $a, b \in \mathbb{R}_+$

$$\left(\frac{a+b}{2}\right)^p \le \frac{a^p+b^p}{2}$$

und daraus folgt für alle $a, b \in \mathbb{R}$

$$\Big(|a|+|b|\Big)^p \le 2^{p-1}\Big(|a|^p+|b|^p\Big)$$

Damit ist auch die letzte Ungleichung gezeigt. □

Aus Lemma 8.4.1 folgt zunächst, dass $\mathcal{L}^p_\mu(\mathcal{F})$ ein Vektorverband ist, und es ist dann klar, dass $\mathcal{L}^p_\mu(\mathcal{F})$ sogar ein Ideal in $\mathcal{L}^0(\mathcal{F})$ ist.

Sei nun

$$L^p(\mathcal{F},\mu) := \left\{ [h]_\mu \in L^0(\mathcal{F},\mu) \,\middle|\, \text{es gibt ein } f \in [h]_\mu \cap \mathcal{L}^p_\mu(\mathcal{F}) \right\}$$

Da $\mathcal{L}^p_\mu(\mathcal{F})$ ein Vektorverband und ein Ideal in $\mathcal{L}^0(\mathcal{F})$ ist, ist auch $L^p(\mathcal{F},\mu)$ ein Vektorverband und ein Ideal in $L^0(\mathcal{F},\mu)$.

Nach Lemma 8.3.6 gilt für alle $f, g \in \mathcal{L}^p_\mu(\mathcal{F})$ mit $f =_\mu g$

$$\int_\Omega |f|^p \, d\mu = \int_\Omega |g|^p \, d\mu$$

Daher ist die Abbildung $\|\,.\,\|_p : L^p(\mathcal{F},\mu) \to \mathbb{R}_+$ mit

$$\|[f]_\mu\|_p := \left(\int_\Omega |f|^p \, d\mu\right)^{1/p}$$

wohldefiniert; im folgenden schreiben wir $\|f\|_p$ anstelle von $\|[f]_\mu\|_p$.

Unser erstes Ziel ist es zu zeigen, dass die Abbildung $\|\,.\,\|_p$ eine Norm ist. Offenbar gilt $\|f\|_p = 0$ genau dann, wenn $f = 0$ gilt, und für alle $f \in L^p(\mathcal{F}, \mu)$ und $c \in \mathbb{R}$ gilt $\|cf\|_p = |c|\,\|f\|_p$; zu zeigen bleibt also noch, dass für alle $f, g \in L^p(\mathcal{F}, \mu)$ die Dreiecksungleichung

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

gilt. Dies ist die Ungleichung von Minkowski, die wir in Lemma 8.4.3 beweisen. Für $p = 1$ ist die Ungleichung von Minkowski trivial, und für $p = 2$ geben wir einen einfachen Beweis am Ende dieses Abschnitts. Für beliebige $p \in (1, \infty)$ ergibt sich die Ungleichung von Minkowski aus der folgenden Ungleichung von Hölder:

8.4.2 Lemma (Ungleichung von Hölder). *Sei $p \in (1, \infty)$ und $q \in (1, \infty)$ derart, dass $1/p + 1/q = 1$. Für alle $f \in L^p(\mathcal{F}, \mu)$ und $g \in L^q(\mathcal{F}, \mu)$ gilt dann $fg \in L^1(\mathcal{F}, \mu)$ und*

$$\|fg\|_1 \leq \|f\|_p\,\|g\|_q$$

Beweis. Wir betrachten zunächst für $\alpha \in (0, 1)$ die Funktion $\varphi_\alpha : (0, \infty) \to \mathbb{R}$ mit

$$\varphi_\alpha(x) := \alpha\, x + (1-\alpha) - x^\alpha$$

Für alle $x \in (0, \infty)$ gilt $\varphi_\alpha(x) \geq \varphi_\alpha(1) = 0$. Für alle $a, b \in (0, \infty)$ folgt daraus $0 \leq \varphi_\alpha(a/b) = \alpha\,(a/b) + (1-\alpha) - (a/b)^\alpha$ und damit

$$a^\alpha\, b^{1-\alpha} \leq \alpha\, a + (1-\alpha)\, b$$

Für $\alpha := 1/p$ gilt $1 - \alpha = 1/q$. Daher gilt für alle $a, b \in (0, \infty)$

$$a^{1/p}\, b^{1/q} \leq \frac{1}{p}\, a + \frac{1}{q}\, b$$

und diese Ungleichung gilt sogar für alle $a, b \in \mathbb{R}_+$.
Im Fall $\|f\|_p = 1 = \|g\|_p$ folgt aus der letzten Ungleichung

$$|fg| = (|f|^p)^{1/p}(|g|^q)^{1/q} \leq \frac{1}{p}\,|f|^p + \frac{1}{q}\,|g|^q$$

und damit

$$\int_\Omega |fg|\, d\mu \leq \frac{1}{p}\int_\Omega |f|^p\, d\mu + \frac{1}{q}\int_\Omega |g|^q\, d\mu = \frac{1}{p}\,\|f\|_p^p + \frac{1}{q}\,\|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1$$

Im Fall $\|f\|_p > 0$ und $\|g\|_p > 0$ gilt $\|f/\|f\|_p\|_p = 1 = \|g/\|g\|_q\|_q$, also

$$\int_\Omega \left| \frac{f}{\|f\|_p}\, \frac{g}{\|g\|_q} \right| d\mu \leq 1$$

und damit

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q$$

Im Fall $\|f\|_p = 0$ oder $\|g\|_p = 0$ gilt $f =_\mu 0$ oder $g =_\mu 0$, also auch $fg =_\mu 0$, und damit $\|fg\|_1 = 0$; daher gilt die Ungleichung auch in diesem Fall. □

Wir können nun die Ungleichung von Minkowski beweisen:

8.4.3 Lemma (Ungleichung von Minkowski). *Für alle $f, g \in L^p(\mathcal{F}, \mu)$ gilt*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

Insbesondere ist die Abbildung $\|\,.\,\|_p$ eine Norm auf $L^p(\mathcal{F}, \mu)$.

Beweis. Im Fall $p = 1$ ist die Ungleichung klar.
Sei nun $p \in (1, \infty)$ und sei $q \in (1, \infty)$ derart, dass $1/p + 1/q = 1$. Im Fall $\|f+g\|_p = 0$ ist nichts zu beweisen.
Sei nun $\|f+g\|_p > 0$. Wegen $1/p + 1/q = 1$ gilt $q + p = pq$ und damit $p = (p-1)\, q$. Daraus folgt zunächst

$$\|\, |f+g|^{p-1} \,\|_q^q = \int_\Omega (|f+g|^{p-1})^q \, d\mu = \int_\Omega |f+g|^p = \|f+g\|_p^p$$

und aus der Ungleichung von Hölder erhalten wir nun

$$\begin{aligned}
\|f+g\|_p^p &= \|\, |f+g|^p \,\|_1 \\
&\leq \|\, (|f| + |g|)\, |f+g|^{p-1} \,\|_1 \\
&\leq \|\, |f|\, |f+g|^{p-1} \,\|_1 + \|\, |g|\, |f+g|^{p-1} \,\|_1 \\
&\leq \|f\|_p \, \|\, |f+g|^{p-1} \,\|_q + \|g\|_p \, \|\, |f+g|^{p-1} \,\|_q \\
&= \Big(\|f\|_p + \|g\|_p \Big) \, \|\, |f+g|^{p-1} \,\|_q \\
&= \Big(\|f\|_p + \|g\|_p \Big) \, \|f+g\|_p^{p/q} \\
&= \Big(\|f\|_p + \|g\|_p \Big) \, \|f+g\|_p^{p-1}
\end{aligned}$$

und damit

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

Damit ist das Lemma bewiesen. □

Nach Lemma 8.4.3 ist $(L^p(\mathcal{F}, \mu), \|\,.\,\|_p)$ ein normierter Raum. Unser nächstes Ziel ist es zu zeigen, dass der normierte Raum $(L^p(\mathcal{F}, \mu), \|\,.\,\|_p)$ vollständig ist. Dazu benötigen wir eine Variante des Satzes über die majorisierte Konvergenz und die folgende Eigenschaft der Norm $\|\,.\,\|_p$:

8.4.4 Folgerung. *Für jede Folge* $\{f_n\}_{n\in\mathbb{N}} \subseteq L^p(\mathcal{F},\mu)$ *gilt*

$$\left(\int_\Omega \left(\sum_{n=1}^\infty |f_n|\right)^p d\mu\right)^{1/p} \le \sum_{n=1}^\infty \|f_n\|_p$$

Beweis. Nach dem Satz über die monotone Konvergenz gilt

$$\begin{aligned}\int_\Omega \left(\sum_{k=1}^\infty |f_k|\right)^p d\mu &= \int_\Omega \left(\sup_{n\in\mathbb{N}} \sum_{k=1}^n |f_k|\right)^p d\mu \\ &= \int_\Omega \sup_{n\in\mathbb{N}} \left(\sum_{k=1}^n |f_k|\right)^p d\mu \\ &= \sup_{n\in\mathbb{N}} \int_\Omega \left(\sum_{k=1}^n |f_k|\right)^p d\mu\end{aligned}$$

und aus der Ungleichung von Minkowski folgt nun

$$\begin{aligned}\left(\int_\Omega \left(\sum_{k=1}^\infty |f_k|\right)^p d\mu\right)^{1/p} &= \sup_{n\in\mathbb{N}} \left(\int_\Omega \left(\sum_{k=1}^n |f_k|\right)^p d\mu\right)^{1/p} \\ &= \sup_{n\in\mathbb{N}} \left\|\sum_{k=1}^n |f_k|\right\|_p \\ &\le \sup_{n\in\mathbb{N}} \sum_{k=1}^n \|f_k\|_p \\ &= \sum_{k=1}^\infty \|f_k\|_p\end{aligned}$$

Damit ist die Behauptung bewiesen. □

Wir beweisen nun eine Variante des Satzes über die majorisierte Konvergenz:

8.4.5 Satz (Majorisierte Konvergenz; Lebesgue). *Sei* $\{f_n\}_{n\in\mathbb{N}}$ *eine Folge von messbaren Funktionen, die* μ*–fast überall gegen eine Funktion* f *konvergiert. Gilt* $\sup_{n\in\mathbb{N}} |f_n| \in L^p(\mathcal{F},\mu)$, *so gilt auch* $f \in L^p(\mathcal{F},\mu)$ *und*

$$\lim_{n\to\infty} \|f_n - f\|_p = 0$$

Insbesondere gilt $f_n \in L^p(\mathcal{F},\mu)$ *für alle* $n \in \mathbb{N}$ *und*

$$\lim_{n\to\infty} \|f_n\|_p = \|f\|_p$$

Beweis. Nach Voraussetzung ist die Funktion

$$g := \sup_{n\in\mathbb{N}} |f_n|$$

p–fach integrierbar. Damit ist auch jede der Funktionen f_n p–fach integrierbar, und wegen $|f| \leq g$ ist auch f p–fach integrierbar. Aus Lemma 8.4.1 erhalten wir $|f_n - f|^p \leq (|f_n| + |f|)^p \leq 2^{p-1}(|f_n|^p + |f|^p) \leq 2^{p-1}(g^p + g^p) = (2g)^p$ und damit

$$(2g)^p - |f_n - f|^p \geq 0$$

Nach Voraussetzung gilt

$$\lim_{n\to\infty} \Big((2g)^p - |f_n - f|^p \Big) = (2g)^p$$

und aus dem Lemma von Fatou folgt nun

$$\begin{aligned}
\int_\Omega (2g)^p \, d\mu &= \int_\Omega \lim_{n\to\infty} \Big((2g)^p - |f_n - f|^p \Big) \, d\mu \\
&= \int_\Omega \liminf_{n\to\infty} \Big((2g)^p - |f_n - f|^p \Big) \, d\mu \\
&\leq \liminf_{n\to\infty} \int_\Omega \Big((2g)^p - |f_n - f|^p \Big) \, d\mu \\
&= \int_\Omega (2g)^p \, d\mu + \liminf_{n\to\infty} \int_\Omega \Big(-|f_n - f|^p \Big) \, d\mu \\
&= \int_\Omega (2g)^p \, d\mu - \limsup_{n\to\infty} \int_\Omega |f_n - f|^p \, d\mu
\end{aligned}$$

Da g p–fach integrierbar ist, erhalten wir daraus

$$\limsup_{n\to\infty} \int_\Omega |f_n - f|^p \, d\mu \leq 0$$

und damit $\lim_{n\to\infty} \int_\Omega |f_n - f|^p \, d\mu = 0$. Die abschließende Behauptung folgt nun aus der Dreiecksungleichung. □

Wir können nun zeigen, dass der normierte Raum $(L^p(\mathcal{F}, \mu), \|\,.\,\|_p)$ vollständig ist.

8.4.6 Lemma (Riesz/Fischer). $(L^p(\mathcal{F}, \mu), \|\,.\,\|_p)$ *ist vollständig.*

Beweis. Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Cauchy–Folge in $(L^p(\mathcal{F}, \mu), \|\,.\,\|_p)$. Dann gibt es eine streng monoton wachsende Folge $\{n_k\}_{k\in\mathbb{N}} \subseteq \mathbb{N}$ mit

$$\|f_{n_{k+1}} - f_{n_k}\|_p \leq 2^{-k}$$

für alle $k \in \mathbb{N}$. Aus Folgerung 8.4.4 ergibt sich nun

$$\left(\int_\Omega \left(|f_{n_1}| + \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}|\right)^p d\mu\right)^{1/p} \le \|f_{n_1}\|_p + \sum_{k=1}^{\infty} \|f_{n_{k+1}} - f_{n_k}\|_p$$
$$\le \|f_{n_1}\|_p + 1$$

Daher ist die Funktion

$$g := |f_{n_1}| + \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}|$$

p–fach integrierbar und wir können annehmen, dass g reell ist. Daher gibt es eine messbare reelle Funktion f mit

$$f = f_{n_1} + \lim_{r\to\infty} \sum_{k=1}^{r} \bigl(f_{n_{k+1}} - f_{n_k}\bigr)$$

Wir betrachten nun $m \in \mathbb{N}$. Dann gilt

$$f - f_{n_m} = \lim_{r\to\infty} \sum_{k=m}^{r} \bigl(f_{n_{k+1}} - f_{n_k}\bigr)$$

und für alle $r \in \mathbb{N}(m)$ gilt

$$\left|\sum_{k=m}^{r} \bigl(f_{n_{k+1}} - f_{n_k}\bigr)\right| \le \sum_{k=m}^{r} |f_{n_{k+1}} - f_{n_k}| \le |f_{n_1}| + \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| = g$$

Da g p–fach integrierbar ist, ergibt sich nun aus dem Satz über die majorisierte Konvergenz $f - f_{n_m} \in L^p(\mathcal{F}, \mu)$ und

$$\|f - f_{n_m}\|_p = \left\|\lim_{r\to\infty} \sum_{k=m}^{r} \bigl(f_{n_{k+1}} - f_{n_k}\bigr)\right\|_p = \lim_{r\to\infty} \left\|\sum_{k=m}^{r} \bigl(f_{n_{k+1}} - f_{n_k}\bigr)\right\|_p$$

Wegen $f = (f - f_{n_m}) + f_{n_m}$ gilt dann $f \in L^p(\mathcal{F}, \mu)$ und aus der Abschätzung

$$\left\|\sum_{k=m}^{r} \bigl(f_{n_{k+1}} - f_{n_k}\bigr)\right\|_p \le \sum_{k=m}^{r} \|f_{n_{k+1}} - f_{n_k}\|_p = \sum_{k=m}^{r} 2^{-k} \le 2^{-m+1}$$

ergibt sich nun

$$\|f - f_{n_m}\|_p = \lim_{r\to\infty} \left\|\sum_{k=m}^{r} \bigl(f_{n_{k+1}} - f_{n_k}\bigr)\right\|_p \le 2^{-m+1}$$

Daher gilt

$$\lim_{m\to\infty} \|f - f_{n_m}\|_p = 0$$

Da $\{f_n\}_{n\in\mathbb{N}}$ eine Cauchy–Folge ist, folgt daraus $\lim_{n\to\infty} \|f - f_n\|_p = 0$. $\square$

Nach Lemma 8.4.6 ist $(L^p(\mathcal{F},\mu), \|\,.\,\|_p)$ ein Banach–Raum. Der folgende Satz fasst die Eigenschaften von $(L^p(\mathcal{F},\mu), \|\,.\,\|_p)$ zusammen:

8.4.7 Satz. *$(L^p(\mathcal{F},\mu), \|\,.\,\|_p)$ ist ein Banach–Verband und ein Ideal in $L^0(\mathcal{F},\mu)$.*

Eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von p–fach μ–integrierbaren Funktionen *konvergiert im p–ten Mittel* (bezüglich μ) gegen eine messbare Funktion f, wenn

$$\lim_{n\to\infty} \|f_n - f\|_p = 0$$

gilt; in diesem Fall ist auch f p–fach μ–integrierbar. Im Fall $p = 1$ ist die Konvergenz im p–ten Mittel gerade die Konvergenz im Mittel und im Fall $p = 2$ bezeichnet man die Konvergenz im p–ten Mittel auch als *Konvergenz im quadratischen Mittel*. Aus der Ungleichung von Markov ergibt sich der folgende Zusammenhang zwischen der Konvergenz im p–ten Mittel und der Konvergenz im Maß:

8.4.8 Satz. *Jede Folge von p–fach μ–integrierbaren Funktionen, die im p–ten Mittel gegen eine Funktion f konvergiert, konvergiert auch im Maß μ gegen f.*

Beweis. Sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von p–fach μ–integrierbaren Funktionen, die im p–ten Mittel gegen f konvergiert. Da die Funktion $h : \mathbb{R}_+ \to \mathbb{R}_+$ mit $h(x) := x^p$ monoton wachsend ist, gilt nach der Ungleichung von Markov in der Form von Folgerung 8.2.6 für alle $\varepsilon \in (0,\infty)$

$$\mu[\{|f_n - f| \ge \varepsilon\}] \le \frac{1}{\varepsilon^p}\, \|f_n - f\|_p^p$$

Daraus folgt die Behauptung. □

Wir beziehen nun den Banach–Verband $L^\infty(\mathcal{F},\mu)$ in die Betrachtung mit ein. Unter zusätzlichen Annahmen an das Maß μ lassen sich die Räume $L^p(\mathcal{F},\mu)$ mit $p \in [1,\infty]$ untereinander vergleichen:

8.4.9 Beispiele.

(1) **Folgenräume:** Sei $(\Omega,\mathcal{F},\mu) := (\mathbb{N}_0, 2^{\mathbb{N}_0}, \boldsymbol{\zeta})$ mit $\boldsymbol{\zeta}[A] := |A|$. Dann ist jede Funktion $f : \Omega \to \bar{\mathbb{R}}$ messbar und für alle $p \in [1,\infty)$ gilt

$$\int_\Omega |f|^p \, d\mu = \sum_{n=0}^{\infty} |f(n)|^p$$

Wir identifizieren nun jede Funktion $\Omega \to \bar{\mathbb{R}}$ mit einer Folge $\{a_n\}_{n\in\mathbb{N}_0} \subseteq \bar{\mathbb{R}}$. Für $p \in \{0\} \cup [1,\infty]$ sei

$$\ell^p := L^p(\mathcal{F},\mu)$$

Dann ist ℓ^0 der Raum aller Folgen in $\mathbb{R}$ und ℓ^∞ ist der Raum aller beschränkten Folgen, und für $p \in [1,\infty)$ ist ℓ^p der Raum aller Folgen $\{a_n\}_{n\in\mathbb{N}_0}$ mit

$$\sum_{n=0}^{\infty} |a_n|^p < \infty$$

Insbesondere ist jede Folge in ℓ^p mit $p \in [1,\infty)$ eine Nullfolge. Die Folgen in ℓ^p heißen *p–fach summierbar*.
Sei $p, r \in [1,\infty)$ mit $p < r$. Dann gilt $r/p > 1$. Für jede Folge $\{a_n\}_{n\in\mathbb{N}_0} \in \ell^p$ und alle hinreichend großen $n \in \mathbb{N}_0$ gilt daher

$$|a_n|^r = (|a_n|^p)^{r/p} \leq |a_n|^p$$

und damit $\{a_n\}_{n\in\mathbb{N}_0} \in \ell^r$. Daher gilt $\ell^p \subseteq \ell^r$. Da jede r–fach summierbare Folge eine Nullfolge und damit beschränkt ist, gilt außerdem $\ell^r \subseteq \ell^\infty$.
Insgesamt erhält man für alle $p, r \in (1,\infty)$ mit $p < r$

$$\ell^1 \subseteq \ell^p \subseteq \ell^r \subseteq \ell^\infty$$

(2) **Endliche Maßräume:** Sei $(\Omega, \mathcal{F}, \mu)$ ein endlicher Maßraum. Für $p, r \in [1,\infty)$ mit $p < r$ gilt $r/p > 1$. Aufgrund der Ungleichung von Hölder gilt für jede Funktion $f \in L^r(\mathcal{F},\mu)$

$$\begin{aligned}
\int_\Omega |f|^p \, d\mu &= \|\, |f|^p \,\|_1 \\
&\leq \|\, |f|^p \,\|_{r/p} \, \|\, 1 \,\|_{r/(r-p)} \\
&= \left(\int_\Omega \Bigl(|f|^p \Bigr)^{r/p} d\mu \right)^{p/r} \left(\int_\Omega 1^{r/(r-p)} \, d\mu \right)^{(r-p)/r} \\
&= \|\, f \,\|_r^p \, (\mu[\Omega])^{(r-p)/r}
\end{aligned}$$

und aus der Endlichkeit des Maßes folgt nun $f \in L^p(\mathcal{F},\mu)$. Daher gilt $L^r(\mathcal{F},\mu) \subseteq L^p(\mathcal{F},\mu)$. Aus der Endlichkeit des Maßes folgt außerdem $L^\infty(\mathcal{F},\mu) \subseteq L^r(\mathcal{F},\mu)$.
Insgesamt erhält man für alle $p, r \in (1,\infty)$ mit $p < r$

$$L^\infty(\mathcal{F},\mu) \subseteq L^r(\mathcal{F},\mu) \subseteq L^p(\mathcal{F},\mu) \subseteq L^1(\mathcal{F},\mu)$$

Wir betrachten abschließend den Fall $p = 2$, der Besonderheiten aufweist und in dem der Beweis der Ungleichungen von Hölder und Minkowski vereinfacht werden kann. Eine Funktion $f \in \mathcal{L}^0(\mathcal{F})$ heißt *quadratisch μ–integrierbar*, wenn $\int_\Omega f^2 \, d\mu < \infty$ und damit $f \in \mathcal{L}^2_\mu(\mathcal{F})$ gilt.

Sind f und g quadratisch μ–integrierbare Funktionen, so sind f^2 und g^2 μ–integrierbar und wegen $|fg| \leq f^2 \chi_{\{|g|\leq|f|\}} + g^2 \chi_{\{|g|>|f|\}} \leq f^2 + g^2$ ist auch das Produkt fg μ–integrierbar.

Daher ist die Abbildung $\langle \,.\,, \,.\, \rangle_2 : L^2(\mathcal{F},\mu) \times L^2(\mathcal{F},\mu) \to \mathbb{R}$ mit

$$\langle f, g \rangle_2 := \int_\Omega f g \, d\mu$$

wohldefiniert, und aus den Eigenschaften des Integrals erhält man sofort das folgende Ergebnis:

8.4.10 Lemma. *Die Abbildung* $\langle\, .\,,\, .\,\rangle_2$ *ist ein Skalarprodukt auf* $L^2(\mathcal{F}, \mu)$.

Für jedes Skalarprodukt gilt die Ungleichung von Cauchy/Schwarz:

8.4.11 Lemma (Ungleichung von Cauchy/Schwarz). *Für alle* $f, g \in L^2(\mathcal{F}, \mu)$ *gilt*

$$\langle f, g\rangle_2^2 \leq \langle f, f\rangle_2 \, \langle g, g\rangle_2$$

Beweis. Im Fall $g = 0$ ist nichts zu zeigen. Im Fall $g \neq 0$ gilt für alle $c \in \mathbb{R}$

$$\begin{aligned} 0 &\leq \langle f+cg, f+cg\rangle_2 \\ &= \langle f, f\rangle_2 + 2c\, \langle f, g\rangle_2 + c^2\, \langle g, g\rangle_2 \end{aligned}$$

und mit $c := -\langle f, g\rangle_2 / \langle g, g\rangle_2$ ergibt sich daraus die Behauptung. □

Für alle $f \in L^2(\mathcal{F}, \mu)$ gilt

$$\|f\|_2 = \langle f, f\rangle_2^{1/2}$$

Aus dieser Gleichung und der Ungleichung von Cauchy/Schwarz erhält man das folgende Ergebnis, das aus Lemma 8.4.3 bereits bekannt ist:

8.4.12 Lemma. *Die Abbildung* $\|\, .\, \|_2$ *ist eine Norm auf* $L^2(\mathcal{F}, \mu)$.

Beweis. Für alle $f \in L^2(\mathcal{F}, \mu)$ folgt wegen $\|f\|_2 = \langle f, f\rangle_2^{1/2}$ aus den Eigenschaften des Skalarproduktes, dass $\|f\| = 0$ genau dann gilt, wenn $f = 0$ gilt, und dass für alle $c \in \mathbb{R}$ die Gleichung $\|cf\| = |c|\, \|f\|$ gilt. Aus der Ungleichung von Cauchy/Schwarz ergibt sich ferner für alle $f, g \in L^2(\mathcal{F}, \mu)$

$$\begin{aligned} \|f+g\|_2^2 &= \langle f+g, f+g\rangle_2 \\ &= \langle f, f\rangle_2 + 2\, \langle f, g\rangle_2 + \langle g, g\rangle_2 \\ &\leq \langle f, f\rangle_2 + 2\, \langle f, f\rangle_2^{1/2}\, \langle g, g\rangle_2^{1/2} + \langle g, g\rangle_2 \\ &= \Bigl(\langle f, f\rangle_2^{1/2} + \langle g, g\rangle_2^{1/2}\Bigr)^2 \\ &= \Bigl(\|f\|_2 + \|g\|_2\Bigr)^2 \end{aligned}$$

und damit die Dreiecksungleichung. □

Unter Verwendung der Normen $\|\, .\, \|_2$ und $\|\, .\, \|_1$ lässt sich die Ungleichung von Cauchy/Schwarz auch in der Form

$$\|fg\|_1 \leq \|f\|_2\, \|g\|_2$$

schreiben. Sie ist damit ein Spezialfall der Ungleichung von Hölder, und es ist klar, dass die Dreiecksungleichung für die Norm $\|\, .\, \|_2$ ein Spezialfall der Ungleichung von Minkowski ist.

Nach Lemma 8.4.6 ist der normierte Raum $(L^2(\mathcal{F},\mu), \|\,.\,\|_2)$ vollständig. Daher ist $(L^2(\mathcal{F},\mu), \langle\,.\,,\,.\,\rangle_2)$ ein Hilbert–Raum. Der folgende Satz fasst die Eigenschaften von $(L^2(\mathcal{F},\mu), \langle\,.\,,\,.\,\rangle_2)$ zusammen:

8.4.13 Satz. *$(L^2(\mathcal{F},\mu), \langle\,.\,,\,.\,\rangle_2)$ ist ein Hilbert–Verband und ein Ideal in $L^0(\mathcal{F},\mu)$.*

Aufgaben

8.4.A **Ungleichung von Hölder:** Für alle $f \in L^1(\mathcal{F},\mu)$ und $g \in L^\infty(\mathcal{F},\mu)$ gilt $fg \in L^1(\mathcal{F},\mu)$ und

$$\|fg\|_1 \le \|f\|_1 \|g\|_\infty$$

8.4.B Für alle $p \in [1,\infty)$ besitzt jede Folge von p–fach μ–integrierbaren Funktionen, die im p–ten Mittel gegen eine p–fach μ–integrierbare Funktion f konvergiert, eine Teilfolge, die μ–fast überall gegen f konvergiert.
Hinweis: Beweis von Lemma 8.4.6.

8.4.C Für $p \in (0,1)$ heißt eine messbare Funktion f *p–fach μ–integrierbar*, wenn $|f|^p$ μ–integrierbar ist. Wir setzen

$$\mathcal{L}^p_\mu(\mathcal{F}) := \left\{ f \in \mathcal{L}^0(\mathcal{F}) \;\middle|\; \int_\Omega |f|^p \, d\mu < \infty \right\}$$

und

$$L^p(\mathcal{F},\mu) := \left\{ [h]_\mu \in L^0(\mathcal{F},\mu) \;\middle|\; \text{es gibt ein } f \in [h]_\mu \cap \mathcal{L}^p_\mu(\mathcal{F}) \right\}$$

(1) $L^p(\mathcal{F},\mu)$ ist ein Vektorraum.
(2) Die Abbildung $d_p : L^p(\mathcal{F},\mu) \times L^p(\mathcal{F},\mu)$ mit

$$d_p(f,g) := \int_\Omega |f-g|^p \, d\mu$$

ist (wohldefiniert und) eine Metrik und diese Metrik ist translationsinvariant, aber nicht absolut homogen.
(3) Der metrische Raum $(L^p(\mathcal{F},\mu), d_p)$ ist vollständig.

9

Berechnung des Lebesgue–Integrals

Nachdem das Lebesgue–Integral auf der Grundlage eines Maßes konstruiert ist, betrachten wir nun auch Maße, die durch Integrale dargestellt werden können. Ausgangspunkt ist die Beobachtung, dass man zu jeder positiven messbaren Funktion durch Integration ein Maß erhält, das als unbestimmtes Integral bezeichnet wird (Abschnitt 9.1) und auf den Begriff eines Maßes mit Dichte führt. Maße mit Dichte spielen in der Wahrscheinlichkeitstheorie eine zentrale Rolle. Sie sind aber auch in der allgemeinen Integrationstheorie von Interesse, da die Berechnung eines Integrals nach einem Maß mit Dichte sich in vielen Fällen durch die Kettenregel vereinfachen lässt (Abschnitt 9.2). Schließlich liefert der Satz von Radon/Nikodym eine Bedingung dafür, dass ein Maß eine Dichte bezüglich einem anderen Maß besitzt (Abschnitt 9.3).

Mit der Anwendung der Kettenregel wird das integrierende Maß gewechselt. Dies ist auch das Wesen der Substitutionsregel, bei der ein gegebenes Integral durch ein Integral nach dem Bildmaß des integrierenden Maßes unter einer messbaren Transformation dargestellt wird (Abschnitt 9.4). Um die Vorteile der Substitutionsregel nutzen zu können, ist es gelegentlich erforderlich, die Integration nach einem eingeschränkten Maß zu betrachten (Abschnitt 9.5).

Das enge Wechselspiel zwischen Maßen und Integralen zeigt sich auch bei der Konstruktion von Produktmaßen (Abschnitt 9.6). Das Hauptergebnis über die Integration nach einem Produktmaß ist der Satz von Fubini, der es ermöglicht, Integrale von positiven oder integrierbaren Funktionen in mehreren Variablen durch die sukzessive Integration nach einer einzigen Variablen zu berechnen (Abschnitt 9.7).

Für die Berechnung des Integrals einer Funktion in einer reellen Variablen ist schließlich der Zusammenhang zwischen dem Lebesgue–Integral und dem Riemann–Integral von Nutzen (Abschnitt 9.8).

Im gesamten Kapitel sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_9,

9.1 Integralinduzierte Maße und signierte Maße

Sei f eine messbare Funktion, die positiv oder μ–integrierbar ist. Ist f positiv, so ist für alle $A \in \mathcal{F}$ auch $f\chi_A$ positiv, und ist f μ–integrierbar, so ist für alle $A \in \mathcal{F}$ wegen $|f\chi_A| \le |f|$ auch $f\chi_A$ μ–integrierbar. Daher ist in beiden Fällen die Abbildung $\int f\,d\mu : \mathcal{F} \to \bar{\mathbb{R}}$ mit

$$\left(\int f\,d\mu\right)[A] := \int_\Omega f\chi_A\,d\mu$$

erklärt; zur Vereinfachung der Notation setzen wir

$$\int_A f\,d\mu := \int_\Omega f\chi_A\,d\mu$$

Die Abbildung $\int f\,d\mu$ heißt das *unbestimmte μ–Integral* von f. Der folgende Satz klärt die Eigenschaften des unbestimmten μ–Integrals:

9.1.1 Satz. *Sei f messbar.*
(1) *Ist f positiv, so ist das unbestimmte μ–Integral von f ein Maß.*
(2) *Ist f μ–integrierbar, so ist das unbestimmte μ–Integral von f ein endliches signiertes Maß.*

Beweis. (1) Sei f positiv. Dann gilt

$$\int_\emptyset f\,d\mu = \int_\Omega f\chi_\emptyset\,d\mu = 0$$

Außerdem gilt für jede disjunkte Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ nach dem Satz über die monotone Konvergenz

$$\begin{aligned}
\int_{\sum_{n=1}^\infty A_n} f\,d\mu &= \int_\Omega f\,\chi_{\sum_{n=1}^\infty A_n} d\mu \\
&= \int_\Omega \left(f\sum_{n=1}^\infty \chi_{A_n}\right) d\mu \\
&= \int_\Omega \left(\sum_{n=1}^\infty f\,\chi_{A_n}\right) d\mu \\
&= \sum_{n=1}^\infty \int_\Omega f\,\chi_{A_n} d\mu \\
&= \sum_{n=1}^\infty \int_{A_n} f\,d\mu
\end{aligned}$$

Daher ist das unbestimmte μ–Integral von f ein Maß.

(2) Sei f μ–integrierbar. Dann sind auch f^+ und f^- μ–integrierbar und nach (1) sind die unbestimmten μ–Integrale von f^+ und f^- Maße, die aufgrund der μ–Integrierbarkeit von f^+ und f^- sogar endlich sind. Für alle $A \in \mathcal{F}$ gilt $f\chi_A = f^+\chi_A - f^-\chi_A$ und damit

$$\int_A f\,d\mu = \int_\Omega f\chi_A\,d\mu = \int_\Omega f^+\chi_A\,d\mu - \int_\Omega f^-\chi_A\,d\mu = \int_A f^+\,d\mu - \int_A f^-\,d\mu$$

Daher ist das unbestimmte μ–Integral von f ein endliches signiertes Maß. □

Aufgaben

9.1.A **Unbestimmtes Integral:** Es gilt $\int 1\,d\mu = \mu$.

9.1.B **Unbestimmtes Integral:** Erweitern Sie den Begriff des unbestimmten Integrals auf μ–quasiintegrierbare Funktionen und zeigen Sie, dass das unbestimmte Integral einer μ–quasiintegrierbaren Funktion ein signiertes Maß ist.

9.2 Integration nach einem Maß mit Dichte

Ein Maß $\nu : \mathcal{F} \to [0,\infty]$ heißt *Maß mit μ–Dichte* oder kurz *Maß mit Dichte*, wenn es eine positive messbare Funktion f gibt mit

$$\nu = \int f\,d\mu$$

Ist ν ein Maß mit μ–Dichte, so heißt jede positive messbare Funktion f mit

$$\nu = \int f\,d\mu$$

μ–Dichte von ν. Offenbar ist ein Maß genau dann ein Maß mit μ–Dichte, wenn es das unbestimmte μ–Integral einer positiven messbaren Funktion ist.

9.2.1 Lemma. *Sei $\nu : \mathcal{F} \to [0,\infty]$ ein Maß mit μ–Dichte. Dann ist jede μ–Nullmenge eine ν–Nullmenge.*

Beweis. Sei f eine μ–Dichte von ν. Für jede Menge $A \in \mathcal{F}$ mit $\mu[A] = 0$ gilt dann $f\chi_A = 0$ μ–fast überall, und aus Lemma 8.2.8 folgt nun

$$\nu[A] = \int_A f\,d\mu = \int_\Omega f\chi_A\,d\mu = 0$$

Damit ist die Behauptung gezeigt. □

Der folgende Satz zeigt, dass jedes Integral nach einem Maß mit μ–Dichte als Integral nach μ dargestellt werden kann:

9.2.2 Satz (Kettenregel). *Sei $\nu : \mathcal{F} \to [0,\infty]$ ein Maß mit μ–Dichte und sei f eine μ–Dichte von ν. Sei ferner h eine messbare Funktion.*
(1) *Ist h positiv, so gilt*
$$\int h \, d\nu = \int hf \, d\mu$$
(2) *h ist genau dann ν–integrierbar, wenn hf μ–integrierbar ist, und in diesem Fall gilt*
$$\int h \, d\nu = \int hf \, d\mu$$

Beweis. Sei $A \in \mathcal{F}$ beliebig.
(1) Für alle $C \in \mathcal{F}$ gilt
$$\int_A \chi_C \, d\nu = \int_\Omega \chi_{C\cap A} \, d\nu = \nu[C \cap A] = \int_{C\cap A} f \, d\mu = \int_A \chi_C \, f \, d\mu$$
Damit ist die Behauptung für $h = \chi_C$ mit $C \in \mathcal{F}$ gezeigt. Aus der positiven Linearität des Integrals folgt dann die Behauptung für den Fall, dass h eine positive einfache Funktion ist, und aus dem Satz über die monotone Konvergenz folgt sodann die Behauptung für den Fall, dass h eine beliebige positive messbare Funktion ist.
(2) Wegen (1) gilt
$$\int_\Omega |h| \, d\nu = \int_\Omega |h| f \, d\mu = \int_\Omega |hf| \, d\mu$$
Daraus folgt die Behauptung über die Integrierbarkeit von h und hf. Sei nun h ν–integrierbar und damit hf μ–integrierbar. Wegen $h = h^+ - h^-$ gilt $hf = h^+ f - h^- f = (hf)^+ - (hf)^-$ und aus (1) folgt
$$\begin{aligned}
\int_A h \, d\nu &= \int_A h^+ \, d\nu - \int_A h^- \, d\nu \\
&= \int_A h^+ f \, d\mu - \int_A h^- f \, d\mu \\
&= \int_A (hf)^+ \, d\mu - \int_A (hf)^- \, d\mu \\
&= \int_A hf \, d\mu
\end{aligned}$$
Damit ist die Behauptung bewiesen. □

Es stellt sich die Frage, ob die μ–Dichte eines Maßes mit μ–Dichte eindeutig bestimmt ist. Das folgende Beispiel zeigt, dass dies im allgemeinen nicht der Fall ist:

9.2.3 Beispiel. Sei $\mathcal{F} := \{\emptyset, \Omega\}$ und $\mu[\Omega] = \infty$. Sei

$$\nu := \mu$$

Dann ist für alle $n \in \mathbb{N}$ die Funktion $f_n : \Omega \to \mathbb{R}$ mit

$$f_n(\omega) := n$$

eine μ–Dichte von ν. Die Maße μ und ν sind weder endlich noch σ–endlich.

Unser Ziel ist es nun zu zeigen, dass die μ–Dichte eines Maßes ν mit μ–Dichte eindeutig bestimmt ist, wenn mindestens eines der Maße μ oder ν σ–endlich ist. Wir betrachten zunächst unbestimmte Integrale von μ–integrierbaren Funktionen:

9.2.4 Lemma. *Seien f, g μ–integrierbare Funktionen mit*

$$\int f \, d\mu \le \int g \, d\mu$$

Dann gilt $f \le g$ μ–fast überall.

Beweis. Für alle $n \in \mathbb{N}$ sei $A_n := \{f \ge g + 1/n\}$. Dann ist die Folge $\{A_n\}_{n\in\mathbb{N}}$ monoton wachsend mit

$$\{f > g\} = \bigcup_{n\in\mathbb{N}} A_n$$

Für alle $n \in \mathbb{N}$ gilt

$$\int_{A_n} f \, d\mu + \frac{1}{n} \, \mu[A_n] \le \int_{A_n} g \, d\mu + \frac{1}{n} \, \mu[A_n] = \int_{A_n} \left(g + \frac{1}{n} \right) d\mu \le \int_{A_n} f \, d\mu$$

und damit $\mu[A_n] = 0$, da f integrierbar ist. Da die Vereinigung einer Folge von Nullmengen wieder eine Nullmenge ist, ergibt sich $\mu[\{f > g\}] = 0$. □

9.2.5 Folgerung. *Seien f, g μ–integrierbare Funktionen mit*

$$\int f \, d\mu = \int g \, d\mu$$

Dann gilt $f = g$ μ–fast überall.

Der folgende Satz liefert eine Charakterisierung σ–endlicher Maße. Er zeigt insbesondere, dass jedes σ–endliche Maß als unbestimmtes Integral bezüglich einem endlichen Maß dargestellt werden kann, und liefert damit in Verbindung mit der Kettenregel eine Möglichkeit, Eigenschaften von unbestimmten Integralen bezüglich einem endlichen Maß auf unbestimmte Integrale bezüglich einem σ–endlichen Maß zu übertragen:

9.2.6 Satz. *Folgende Aussagen sind äquivalent:*
(a) *μ ist σ–endlich.*
(b) *Es gibt eine μ–integrierbare Funktion h mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$.*
(c) *Es gibt ein endliches Maß φ und eine messbare Funktion h mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$ und*

$$\varphi = \int h \, d\mu$$

In diesem Fall gilt

$$\mu = \int \frac{1}{h} \, d\varphi$$

und μ und φ besitzen dieselben Nullmengen.

Beweis. Wir nehmen zunächst an, dass (a) gilt. Dann gibt es eine Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ mit $\Omega = \bigcup_{n\in\mathbb{N}} A_n$ und $\mu[A_n] < \infty$ für alle $n \in \mathbb{N}$. Für alle $n \in \mathbb{N}$ wählen wir ein $a_n \in (0, 2^{-n}]$ mit $a_n \mu[A_n] \leq 2^{-n}$. Sei

$$h := \sum_{n=1}^{\infty} a_n \chi_{A_n}$$

Dann gilt $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$, und nach dem Satz über die monotone Konvergenz ist h μ–integrierbar. Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Für alle $n \in \mathbb{N}$ sei $A_n := \{h \geq 1/n\}$. Dann gilt

$$\Omega = \{h > 0\} = \bigcup_{n\in\mathbb{N}} A_n$$

und für alle $n \in \mathbb{N}$ gilt

$$\mu[A_n] = \int_{A_n} 1 \, d\mu \leq \int_{A_n} nh \, d\mu = n \int_{A_n} h \, d\mu = n \, \varphi[A_n]$$

und damit $\mu[A_n] < \infty$. Daher ist μ σ–endlich und (a) folgt aus (b).
Die Äquivalenz von (b) und (c) ist klar.
Wir nehmen schließlich an, dass (c) gilt, Aus der Kettenregel 9.2.2 ergibt sich dann

$$\mu = \int 1 \, d\mu = \int \frac{1}{h} h \, d\mu = \int \frac{1}{h} \, d\varphi$$

Die letzte Behauptung ergibt sich aus Lemma 9.2.1. □

9.2.7 Lemma. *Sei μ σ–endlich und seien f, g positive messbare Funktionen mit*

$$\int f \, d\mu \leq \int g \, d\mu$$

Dann gilt $f \leq g$ μ–fast überall.

Beweis. Wir nehmen zunächst an, dass μ endlich ist. Für alle $n \in \mathbb{N}$ sei $B_n := \{g \leq n\}$. Dann gilt $\bigcup_{n\in\mathbb{N}} B_n = \{g < \infty\}$. Für alle $A \in \mathcal{F}$ gilt

$$\int_A f\chi_{B_n}\, d\mu = \int_{A\cap B_n} f\, d\mu \leq \int_{A\cap B_n} g\, d\mu = \int_A g\chi_{B_n}\, d\mu$$

Insbesondere gilt

$$\int_\Omega f\chi_{B_n}\, d\mu \leq \int_\Omega g\chi_{B_n}\, d\mu \leq \int_\Omega n\, d\mu = n\,\mu[\Omega]$$

Da μ endlich ist, folgt hieraus die Integrierbarkeit von $f\chi_{B_n}$ und $g\chi_{B_n}$, und aus Lemma 9.2.4 erhalten wir nun $f\chi_{B_n} \leq g\chi_{B_n}$ μ–fast überall. Da die Vereinigung einer Folge von Nullmengen wieder eine Nullmenge ist, ergibt sich daraus $f\chi_{\{g<\infty\}} \leq g\chi_{\{g<\infty\}}$ μ–fast überall, und damit $f \leq g$ μ–fast überall. Wir nehmen nun an, dass μ σ–endlich ist. Nach Satz 9.2.6 gibt es ein endliches Maß φ und eine messbare Funktion h mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$ und

$$\mu = \int \frac{1}{h}\, d\varphi$$

Aus der Kettenregel 9.2.2 ergibt sich dann

$$\int f\,\frac{1}{h}\, d\varphi = \int f\, d\mu \leq \int g\, d\mu = \int g\,\frac{1}{h}\, d\varphi$$

Da φ ein endliches Maß ist, erhalten wir aus dem ersten Teil des Beweises $f/h \leq g/h$ φ–fast überall und damit $f \leq g$ φ–fast überall. Da jede φ–Nullmenge auch eine μ–Nullmenge ist, erhalten wir $f \leq g$ μ–fast überall. □

9.2.8 Folgerung. *Sei μ σ–endlich und seien f, g positive messbare Funktionen mit*

$$\int f\, d\mu = \int g\, d\mu$$

Dann gilt $f = g$ μ–fast überall.

Wir kommen nun zu dem angekündigten Eindeutigkeitssatz für Maße mit Dichten:

9.2.9 Satz. *Sei ν ein Maß mit μ–Dichte. Ist μ oder ν σ–endlich, so ist die μ–Dichte von ν μ–fast überall eindeutig bestimmt.*

Beweis. Seien f, g μ–Dichten von ν. Dann gilt

$$\int f\, d\mu = \nu = \int g\, d\mu$$

(1) Ist μ σ–endlich, so ergibt sich aus Folgerung 9.2.8 $f = g$ μ–fast überall.

(2) Ist μ beliebig und ν endlich, so ergibt sich aus Folgerung 9.2.5 $f = g$ μ–fast überall.
(3) Sei nun μ beliebig und ν σ–endlich. Nach Satz 9.2.6 gibt es ein endliches Maß φ und eine messbare Funktion h mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$ und

$$\varphi = \int h \, d\nu$$

Aus der Kettenregel 9.2.2 folgt nun

$$\int hf \, d\mu = \int h \, d\nu = \varphi = \int h \, d\nu = \int hg \, d\mu$$

Daher ist φ ein endliches Maß mit μ–Dichte und jede der Funktionen hf und hg ist eine μ–Dichte von φ. Aus (2) ergibt sich nun $hf = hg$ μ–fast überall und damit $f = g$ μ–fast überall. □

Abschließend charakterisieren wir endliche und σ–endliche Maße mit μ–Dichte durch Eigenschaften ihrer μ–Dichten:

9.2.10 Satz. *Sei μ σ–endlich und sei ν ein Maß mit μ–Dichte f. Dann gilt:*
(1) *ν ist genau dann endlich, wenn f μ–integrierbar ist.*
(2) *ν ist genau dann σ–endlich, wenn f μ–fast überall endlich ist.*

Beweis. Aussage (1) ist klar.
Wir nehmen nun an, dass ν σ–endlich ist. Nach Satz 9.2.6 gibt es ein endliches Maß φ und eine messbare Funktion h mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$ und

$$\varphi = \int h \, d\nu$$

und aus der Kettenregel 9.2.2 folgt nun

$$\varphi = \int h \, d\nu = \int hf \, d\mu$$

Nach (1) ist die Funktion hf μ–integrierbar und aus Lemma 8.2.7 folgt nun, dass hf μ–fast überall endlich ist. Dann ist aber auch f μ–fast überall endlich.
Wir nehmen abschließend an, dass f μ–fast überall endlich ist.

– Sei zunächst μ endlich. Für alle $n \in \mathbb{N}$ sei $A_n := \{f \leq n\}$ und $B_n := A_n \cup \{f = \infty\}$. Dann gilt $\bigcup_{n\in\mathbb{N}} B_n = \Omega$ und für alle $n \in \mathbb{N}$ gilt wegen $f\chi_{B_n} = f\chi_{A_n}$ μ–fast überall

$$\nu[B_n] = \int_{B_n} f \, d\mu = \int_\Omega f\chi_{B_n} \, d\mu = \int_\Omega f\chi_{A_n} \, d\mu \leq \int_\Omega n \, d\mu = n \, \mu[\Omega]$$

und damit $\nu[B_n] < \infty$. Daher ist ν σ–endlich.

– Sei nun μ σ–endlich. Nach Satz 9.2.6 gibt es ein endliches Maß φ und eine messbare Funktion h mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$ sowie

$$\varphi = \int h \, d\mu$$

und

$$\mu = \int \frac{1}{h} \, d\varphi$$

Aus der Kettenregel 9.2.2 folgt nun

$$\nu = \int f \, d\mu = \int f \frac{1}{h} \, d\varphi$$

Da f μ–fast überall endlich ist, ist auch f/h μ–fast überall endlich und damit φ–fast überall endlich. Da φ endlich ist, folgt aus dem bereits gezeigten, dass ν σ–endlich ist.

Damit ist der Satz bewiesen. □

Aufgaben

9.2.A **Lokales Zählmaß:** Sei $C \in \mathcal{F}$ abzählbar. Dann besitzt das lokale Zählmaß $\boldsymbol{\zeta}_C$ bezüglich C die $\boldsymbol{\zeta}$–Dichte χ_C.

9.2.B **Kettenregel:** Sei $\nu : \mathcal{F} \to [0, \infty]$ ein Maß mit μ–Dichte und sei f eine μ–Dichte von ν. Sei ferner h eine messbare Funktion. Dann gilt: h ist genau dann ν–quasiintegrierbar, wenn hf μ–quasiintegrierbar ist, und in diesem Fall gilt

$$\int h \, d\nu = \int hf \, d\mu$$

Hinweis: Es gilt $h^+ = h\chi_{\{h \geq 0\}}$ und $h^- = -h\chi_{\{h \leq 0\}}$.

9.2.C Sei $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), \boldsymbol{\lambda})$. Geben Sie eine messbare Funktion h an mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$, für die das unbestimmte Integral $\int h \, d\mu$ ein endliches Maß ist.

9.3 Absolutstetige und singuläre Maße

Ein Maß $\nu : \mathcal{F} \to [0, \infty]$ heißt *absolutstetig bezüglich* μ oder kurz μ*–stetig*, wenn für alle $A \in \mathcal{F}$ mit $\mu[A] = 0$ auch

$$\nu[A] = 0$$

gilt; in diesem Fall schreiben wir

$$\nu \ll \mu$$

Nach Lemma 9.2.1 ist jedes Maß mit μ–Dichte μ–stetig. Wir zeigen nun, dass unter der Voraussetzung, dass μ σ-endlich ist, auch die Umkehrung dieser Implikation gilt:

9.3.1 Satz (Radon/Nikodym). *Sei μ σ–endlich und sei ν ein Maß mit $\nu \ll \mu$. Dann gibt es eine positive messbare Funktion f mit*

$$\nu = \int f \, d\mu$$

Die Funktion f ist μ–fast überall eindeutig bestimmt.

Beweis. Die Aussage über die Eindeutigkeit der Funktion f ergibt sich aus Satz 9.2.9. Wir führen den Beweis ihrer Existenz in drei Schritten:
(1) Wir nehmen zunächst an, dass μ und ν endlich sind. Sei

$$\mathcal{H} := \left\{ h : \Omega \to [0, \infty] \text{ messbar} \,\middle|\, \int_A h \, d\mu \leq \nu[A] \text{ für alle } A \in \mathcal{F} \right\}$$

Die Menge $\mathcal{H}$ ist nichtleer und sie enthält mit je zwei Funktionen auch ihr Supremum, denn für $g, h \in \mathcal{H}$ gilt für alle $A \in \mathcal{F}$

$$\begin{aligned} \int_A (g \vee h) \, d\mu &= \int_{A \cap \{g \geq h\}} g \, d\mu + \int_{A \cap \{g < h\}} h \, d\mu \\ &\leq \nu[A \cap \{g \geq h\}] + \nu[A \cap \{g < h\}] \\ &= \nu[A] \end{aligned}$$

und damit $g \vee h \in \mathcal{H}$. Des weiteren gibt es eine Folge $\{h_n\}_{n \in \mathbb{N}} \subseteq \mathcal{H}$ mit

$$\sup_{n \in \mathbb{N}} \int_\Omega h_n \, d\mu = \sup_{h \in \mathcal{H}} \int_\Omega h \, d\mu$$

Dann ist die Folge $\{f_n\}_{n \in \mathbb{N}}$ mit

$$f_n := \sup_{k \in \{1, \ldots, n\}} h_k$$

positiv und monoton wachsend mit $h_n \leq f_n \in \mathcal{H}$ und aus der Monotonie des Integrals folgt

$$\sup_{n \in \mathbb{N}} \int_\Omega f_n \, d\mu = \sup_{h \in \mathcal{H}} \int_\Omega h \, d\mu$$

Wir setzen

$$f := \sup_{n \in \mathbb{N}} f_n$$

Aus dem Satz über die monotone Konvergenz ergibt sich für alle $A \in \mathcal{F}$

$$\begin{aligned}\int_A f\,d\mu &= \int_\Omega f\chi_A\,d\mu \\ &= \int_\Omega \sup_{n\in\mathbb{N}} f_n\chi_A\,d\mu \\ &= \sup_{n\in\mathbb{N}} \int_\Omega f_n\chi_A\,d\mu \\ &= \sup_{n\in\mathbb{N}} \int_A f_n\,d\mu \\ &\le \nu[A]\end{aligned}$$

und damit $f \in \mathcal{H}$. Sei nun

$$\varphi := \int f\,d\mu$$

Wegen $f \in \mathcal{H}$ gilt dann

$$\varphi \le \nu$$

Unser Ziel ist es nun zu zeigen, dass $\varphi = \nu$ gilt. Sei $\varepsilon \in (0,\infty)$. Mit ν ist auch φ endlich, und da auch μ endlich ist, ist die Mengenfunktion

$$\tau_\varepsilon := \nu - \varphi - \varepsilon\mu$$

ein endliches signiertes Maß. Aus der Hahn–Zerlegung ergibt sich nun die Existenz von disjunkten Mengen $P, N \in \mathcal{F}$ mit $\Omega = P + N$ und

$$\begin{aligned}\tau_\varepsilon^+[A] &= \tau_\varepsilon[A \cap P] \\ \tau_\varepsilon^-[A] &= -\,\tau_\varepsilon[A \cap N]\end{aligned}$$

für alle $A \in \mathcal{F}$. Für alle $A \in \mathcal{F}$ gilt dann wegen $\varphi \le \nu$

$$\begin{aligned}\int_A \Bigl(f + \varepsilon\chi_P\Bigr)\,d\mu &= \int_A f\,d\mu + \varepsilon\,\mu[A \cap P] \\ &= \varphi[A] + \varepsilon\,\mu[A \cap P] \\ &= \nu[A] - (\nu-\varphi)[A] + \varepsilon\,\mu[A \cap P] \\ &\le \nu[A] - (\nu-\varphi)[A \cap P] + \varepsilon\,\mu[A \cap P] \\ &= \nu[A] - \tau_\varepsilon[A \cap P] \\ &= \nu[A] - \tau_\varepsilon^+[A] \\ &\le \nu[A]\end{aligned}$$

Daher gilt $f + \varepsilon\chi_P \in \mathcal{H}$ und damit

$$\int_\Omega f\,d\mu + \varepsilon\,\mu[P] = \int_\Omega \Bigl(f + \varepsilon\chi_P\Bigr)\,d\mu \le \sup_{h\in\mathcal{H}} \int_\Omega h\,d\mu = \int_\Omega f\,d\mu$$

Da φ endlich ist, gilt $\int_\Omega f\,d\mu = \varphi[\Omega] < \infty$ und wir erhalten $\mu[P] = 0$. Wegen $\varphi \le \nu \ll \mu$ ergibt sich sodann $\tau_\varepsilon^+[P] = \tau_\varepsilon[P] = \nu[P] - \varphi[P] - \varepsilon\,\mu[P] = 0$. Daher gilt für alle $A \in \mathcal{F}$

$$\begin{aligned}\nu[A] &= \varphi[A] + \varepsilon\,\mu[A] + \tau_\varepsilon[A]\\ &= \varphi[A] + \varepsilon\,\mu[A] + \tau_\varepsilon^+[A] - \tau_\varepsilon^-[A]\\ &= \varphi[A] + \varepsilon\,\mu[A] - \tau_\varepsilon^-[A]\\ &\le \varphi[A] + \varepsilon\,\mu[A]\end{aligned}$$

Da μ endlich ist und $\varepsilon \in (0,\infty)$ beliebig war, erhalten wir

$$\nu \le \varphi$$

Zusammen mit der Ungleichung $\varphi \le \nu$ ergibt sich daher

$$\nu = \varphi$$

und aus der Definition von φ folgt nun

$$\nu = \int f\,d\mu$$

(2) Wir nehmen nun an, dass μ endlich und ν beliebig ist. Sei

$$\mathcal{C} := \Bigl\{ C \in \mathcal{F} \Bigm| \nu[C] < \infty \Bigr\}$$

Dann ist das Mengensystem $\mathcal{C}$ nichtleer und $\cup$–stabil und es gibt eine monoton wachsende Folge $\{C_n\}_{n\in\mathbb{N}_0} \subseteq \mathcal{C}$ mit $C_0 := \emptyset$ und

$$\sup_{n\in\mathbb{N}} \mu[C_n] = \sup_{C\in\mathcal{C}} \mu[C]$$

Für alle $n \in \mathbb{N}$ sei $B_n := C_n \setminus C_{n-1}$. Dann ist die Folge $\{B_n\}_{n\in\mathbb{N}}$ disjunkt mit $B_n \in \mathcal{C}$ für alle $n \in \mathbb{N}$ und

$$\sum_{n=1}^\infty B_n = \bigcup_{n\in\mathbb{N}} C_n$$

Sei ferner

$$B_0 := \Omega \setminus \sum_{n=1}^\infty B_n$$

Für alle $n \in \mathbb{N}_0$ seien $\mu_n, \nu_n : \mathcal{F} \to [0,\infty]$ gegeben durch

$$\begin{aligned}\mu_n[A] &:= \mu[A \cap B_n]\\ \nu_n[A] &:= \nu[A \cap B_n]\end{aligned}$$

Dann ist μ_n ein endliches Maß mit

$$\mu_n = \int \chi_{B_n} d\mu$$

und ν_n ist ein Maß mit $\nu_n \ll \mu_n$.
Für alle $n \in \mathbb{N}$ ist auch ν_n endlich und nach (1) gibt es eine positive messbare Funktion f_n mit

$$\nu_n = \int f_n \, d\mu_n$$

Es bleibt zu zeigen, dass es eine positive messbare Funktion f_0 gibt mit

$$\nu_0 = \int f_0 \, d\mu_0$$

denn dann ergibt sich für alle $A \in \mathcal{F}$ aus der Kettenregel

$$\begin{aligned}
\nu[A] &= \sum_{n=0}^{\infty} \nu[A \cap B_n] \\
&= \sum_{n=0}^{\infty} \nu_n[A] \\
&= \sum_{n=0}^{\infty} \int_A f_n \, d\mu_n \\
&= \sum_{n=0}^{\infty} \int_A f_n \, \chi_{B_n} d\mu \\
&= \int_A \left(\sum_{n=0}^{\infty} f_n \chi_{B_n} \right) d\mu
\end{aligned}$$

und für die positive messbare Funktion $f := \sum_{n=0}^{\infty} f_n \chi_{B_n}$ erhalten wir

$$\nu = \int f \, d\mu$$

Wir zeigen nun, dass ν_0 tatsächlich eine μ_0–Dichte besitzt. Sei $A \in \mathcal{F}$.

– Im Fall $\nu_0[A] < \infty$ gilt $\nu[A \cap B_0] < \infty$ und damit $A \cap B_0 \in \mathcal{C}$. Für alle $n \in \mathbb{N}$ gilt daher $C_n + (A \cap B_0) \in \mathcal{C}$ und daraus folgt

$$\begin{aligned}
\sup_{C \in \mathcal{C}} \mu[C] + \mu_0[A] &= \sup_{n \in \mathbb{N}} \mu[C_n] + \mu[A \cap B_0] \\
&= \sup_{n \in \mathbb{N}} \Big(\mu[C_n] + \mu[A \cap B_0] \Big) \\
&= \sup_{n \in \mathbb{N}} \mu \Big[C_n + (A \cap B_0) \Big] \\
&\le \sup_{C \in \mathcal{C}} \mu[C]
\end{aligned}$$

Da μ endlich ist, erhalten wir $\mu_0[A] = 0$ und aus $\nu \ll \mu$ folgt $\nu_0[A] = 0$.

– Im Fall $\nu_0[A] = \infty$ gilt, wiederum wegen $\nu \ll \mu$, $\mu_0[A] > 0$.
Es gilt also entweder $\nu_0[A] = 0 = \mu_0[A]$ oder $\nu_0[A] = \infty$ und $\mu_0[A] > 0$. Daher folgt in beiden Fällen aus dem Satz über die monotone Konvergenz

$$\nu_0[A] = \sup_{n \in \mathbb{N}} \Big(n\, \mu_0[A] \Big) = \sup_{n \in \mathbb{N}} \int_A n\, d\mu_0 = \int_A \sup_{n \in \mathbb{N}} n\, d\mu_0$$

Für die positive messbare Funktion f_0 mit $f_0(\omega) := \infty$ erhalten wir daher

$$\nu_0 = \int f_0\, d\mu_0$$

Daher ist f_0 eine μ_0–Dichte von ν_0.
(3) Wir nehmen schließlich an, dass μ σ–endlich und ν beliebig ist. Nach Satz 9.2.6 gibt es ein endliches Maß φ und eine messbare Funktion h mit $0 < h(\omega) < \infty$ für alle $\omega \in \Omega$ und

$$\varphi = \int h\, d\mu$$

und μ und φ besitzen dieselben Nullmengen. Wegen $\nu \ll \mu$ gilt daher $\nu \ll \varphi$. Da φ endlich ist, ergibt sich aus (2) die Existenz einer positiven messbaren Funktion g mit

$$\nu = \int g\, d\varphi$$

und aus der Kettenregel 9.2.2 folgt nun

$$\nu = \int g\, d\varphi = \int gh\, d\mu$$

Daher ist $f := gh$ eine μ–Dichte von ν. □

Ein Maß $\nu : \mathcal{F} \to [0, \infty]$ heißt *singulär bezüglich* μ oder kurz μ*–singulär*, wenn es eine Menge $N \in \mathcal{F}$ gibt mit

$$\nu[\Omega \setminus N] + \mu[N] = 0$$

in diesem Fall schreiben wir

$$\nu \perp \mu$$

Die Singularität von Maßen ist eine symmetrische Relation, denn ν ist genau dann μ–singulär, wenn μ ν–singulär ist.

9.3.2 Satz (Lebesgue–Zerlegung). *Sei μ σ–endlich. Dann gibt es zu jedem σ–endlichen Maß ν zwei Maße φ, ψ mit $\varphi \ll \mu$ und $\psi \perp \mu$ sowie*

$$\nu = \varphi + \psi$$

Die Zerlegung ist eindeutig und es gilt $\varphi \perp \psi$.

Beweis. Mit μ und ν ist auch das Maß

$$\tau := \mu + \nu$$

σ–endlich. Wegen $\mu \leq \tau$ gilt $\mu \ll \tau$, und aus dem Satz von Radon/Nikodym folgt nun die Existenz einer positiven messbaren Funktion g mit

$$\mu = \int g \, d\tau$$

Sei $N := \{g = 0\}$. Dann gilt $g(\omega)\chi_N(\omega) = 0$ für alle $\omega \in \Omega$ und damit

$$\mu[N] = \int_N g \, d\tau = 0$$

Wir definieren nun zwei Maße $\varphi, \psi : \mathcal{F} \to [0, \infty]$ durch

$$\begin{aligned} \varphi[A] &:= \nu[A \setminus N] \\ \psi[A] &:= \nu[A \cap N] \end{aligned}$$

Dann gilt

$$\nu = \varphi + \psi$$

und wegen $\mu[N] = 0 = \varphi[N]$ und $\psi[\Omega \setminus N] = 0$ gilt

$$\begin{aligned} \mu &\perp \psi \\ \varphi &\perp \psi \end{aligned}$$

Wir zeigen nun, dass auch $\varphi \ll \mu$ gilt. Wir betrachten dazu eine Menge $A \in \mathcal{F}$ mit $\mu[A] = 0$. Dann gilt

$$\int_{A \setminus N} g \, d\tau = \mu[A \setminus N] = 0$$

und damit $g(\omega)\chi_{A \setminus N}(\omega) = 0$ τ–fast überall. Wegen $A \setminus N \subseteq \Omega \setminus N = \{g > 0\}$ folgt daraus $\chi_{A \setminus N}(\omega) = 0$ τ–fast überall, und wir erhalten

$$\varphi[A] = \nu[A \setminus N] \leq \mu[A \setminus N] + \nu[A \setminus N] = \tau[A \setminus N] = 0$$

und damit $\varphi[A] = 0$. Daher gilt

$$\varphi \ll \mu$$

Damit ist die Existenz der Zerlegung gezeigt.
Wir betrachten nun für $i \in \{1, 2\}$ Maße φ_i und ψ_i mit $\varphi_i \ll \mu$ und $\psi_i \perp \mu$ sowie

$$\nu = \varphi_i + \psi_i$$

Dann gibt es Mengen $N_i \in \mathcal{F}$ mit $\mu[N_i] = 0 = \psi_i[\Omega \setminus N_i]$ sowie $\varphi_i[N_i] = 0$. Daher gilt für alle $A \in \mathcal{F}$

$$\begin{aligned} \varphi_i[A] &= \varphi_i[A \setminus N_i] \\ \psi_i[A] &= \psi_i[A \cap N_i] \end{aligned}$$

und damit

$$\begin{aligned} \varphi_i[A] &= \varphi_i[A \setminus N_i] + \psi_i[A \setminus N_i] \\ &= \nu[A \setminus N_i] \end{aligned}$$

und

$$\begin{aligned} \psi_i[A] &= \varphi_i[A \cap N_i] + \psi_i[A \cap N_i] \\ &= \nu[A \cap N_i] \end{aligned}$$

Sei $N := N_1 \cup N_2$. Dann gilt $\mu[N] = 0$. Für alle $A \in \mathcal{F}$ erhalten wir wegen $\varphi_i \ll \mu$

$$\varphi_i[A \cap N] = 0$$

und wegen $N_i \subseteq N$ gilt

$$\psi_i[A \setminus N] = 0$$

Daraus ergibt sich für alle $A \in \mathcal{F}$

$$\begin{aligned} \varphi_i[A] &= \varphi_i[A \setminus N] + \varphi_i[A \cap N] \\ &= \varphi_i[A \setminus N] \\ &= \varphi_i[A \setminus N] + \psi_i[A \setminus N] \\ &= \nu[A \setminus N] \end{aligned}$$

und

$$\begin{aligned} \psi_i[A] &= \psi_i[A \cap N] + \psi_i[A \setminus N] \\ &= \psi_i[A \cap N] \\ &= \psi_i[A \cap N] + \varphi_i[A \cap N] \\ &= \nu[A \cap N] \end{aligned}$$

und damit $\varphi_1 = \varphi_2$ und $\psi_1 = \psi_2$. □

Aufgaben

9.3.A **Äquivalente Maße:** Zwei Maße $\varphi, \psi : \mathcal{F} \to [0, \infty]$ heißen *äquivalent*, wenn $\varphi \ll \psi$ und $\psi \ll \varphi$ gilt; in diesem Fall schreiben wir

$$\varphi \approx \psi$$

Dann ist $\approx$ eine Äquivalenzrelation.

9.3.B **Äquivalente Maße:** Zu jedem σ–endlichen Maß gibt es ein äquivalentes endliches Maß.

9.3.C **Hahn–Zerlegung:** Sei $\nu : \mathcal{F} \to \bar{\mathbb{R}}$ ein signiertes Maß. Dann gilt $\nu^+ \perp \nu^-$.

9.4 Integration nach einem Bildmaß

Wie für das Riemann–Integral gibt es auch für das Lebesgue–Integral eine Substitutionsregel, die die Berechnung eines Integrals in vielen Fällen vereinfacht. Bei der Formulierung der Substitutionsregel für das Lebesgue–Integral ist es wiederum hilfreich, die jeweiligen Integrationsvariablen anzugeben. Die folgende Substitutionsregel wird auch als *Transformationssatz* bezeichnet:

9.4.1 Satz (Substitutionsregel). *Sei $(\Omega', \mathcal{F}')$ ein Messraum und sei $T : \Omega \to \Omega'$ messbar. Sei ferner $h : \Omega' \to \bar{\mathbb{R}}$ eine messbare Funktion.*
(1) *Ist h positiv, so gilt für alle $A' \in \mathcal{F}'$*

$$\int_{T^{-1}(A')} (h \circ T)(\omega)\, d\mu(\omega) = \int_{A'} h(\omega')\, d\mu_T(\omega')$$

(2) *$h \circ T$ ist genau dann μ–integrierbar, wenn h μ_T–integrierbar ist, und in diesem Fall gilt für alle $A' \in \mathcal{F}'$*

$$\int_{T^{-1}(A')} (h \circ T)(\omega)\, d\mu(\omega) = \int_{A'} h(\omega')\, d\mu_T(\omega')$$

Im Fall $A' = \Omega'$ gilt $T^{-1}(A') = \Omega$.

Beweis. (1) Für alle $C' \in \mathcal{F}'$ gilt $(\chi_{C'} \circ T)(\omega) = \chi_{C'}(T(\omega)) = \chi_{T^{-1}(C')}(\omega)$ und damit

$$\begin{aligned}
\int_{T^{-1}(A')} (\chi_{C'} \circ T)(\omega)\, d\mu(\omega) &= \int_{\Omega} \chi_{T^{-1}(A')}(\omega)\, \chi_{T^{-1}(C')}(\omega)\, d\mu(\omega) \\
&= \int_{\Omega} \chi_{T^{-1}(A' \cap C')}(\omega)\, d\mu(\omega) \\
&= \mu[T^{-1}(A' \cap C')] \\
&= \mu_T[A' \cap C'] \\
&= \int_{\Omega'} \chi_{A' \cap C'}(\omega')\, d\mu_T(\omega') \\
&= \int_{A'} \chi_{C'}(\omega')\, d\mu_T(\omega')
\end{aligned}$$

Damit ist die Behauptung für $h = \chi_{C'}$ mit $C' \in \mathcal{F}'$ gezeigt. Aus der positiven Linearität des Integrals folgt dann die Behauptung für den Fall, dass h eine positive einfache Funktion ist, und aus dem Satz über die monotone Konvergenz folgt sodann die Behauptung für den Fall, dass h eine beliebige positive messbare Funktion ist.
(2) Wegen (1) gilt

$$\int_{\Omega} |h \circ T|(\omega)\, d\mu(\omega) = \int_{\Omega} (|h| \circ T)(\omega)\, d\mu(\omega) = \int_{\Omega'} |h|(\omega')\, d\mu_T(\omega')$$

Daraus folgt die Behauptung über die Integrierbarkeit von h und $h \circ T$. Im Fall der Integrierbarkeit von h und $h \circ T$ folgt die Gleichung für die Integrale mit Hilfe der Zerlegung $h = h^+ - h^-$ wegen $(h \circ T)^+ = h^+ \circ T$ und $(h \circ T)^- = h^- \circ T$ aus (1). □

Im Fall $h = \chi_{\Omega'}$ gilt $h \circ T = \chi_\Omega$, und in diesem Fall reduziert sich die Gleichung

$$\int_{T^{-1}(A')} h(T(\omega))\, d\mu(\omega) = \int_{A'} h(\omega')\, d\mu_T(\omega')$$

aus Satz 9.4.1 auf die Gleichung

$$\mu[T^{-1}(A')] = \mu_T[A']$$

Dies ist gerade die Definition des Bildmaßes μ_T von μ unter T.

Die Substitutionsregel lässt sich unter anderem verwenden, um für ein Maß mit Dichte und bestimmte Transformationen das Bildmaß zu bestimmen:

9.4.2 Folgerung. *Sei $(\Omega', \mathcal{F}')$ ein Messraum und sei $T : \Omega \to \Omega'$ messbar. Sei ferner $h : \Omega \to \mathbb{R}_+$ messbar und*

$$\nu := \int h(\omega)\, d\mu(\omega)$$

Ist T bijektiv und T^{-1} messbar, so gilt

$$\nu_T = \int (h \circ T^{-1})(\omega')\, d\mu_T(\omega')$$

Beweis. Für alle $A' \in \mathcal{F}'$ gilt

$$\begin{aligned}\nu_T[A'] &= \nu[T^{-1}(A')] \\ &= \int_{T^{-1}(A')} h(\omega)\, d\mu(\omega) \\ &= \int_{T^{-1}(A')} (h \circ T^{-1} \circ T)(\omega))\, d\mu(\omega) \\ &= \int_{A'} (h \circ T^{-1})(\omega')\, d\mu_T(\omega')\end{aligned}$$

Damit ist die Folgerung bewiesen. □

Aufgabe

9.4.A **Substitutionsregel:** Sei $(\Omega', \mathcal{F}')$ ein Messraum und sei $T : \Omega \to \Omega'$ messbar. Sei ferner $h : \Omega' \to \bar{\mathbb{R}}$ eine messbare Funktion. Dann ist $h \circ T$ genau dann μ–quasiintegrierbar, wenn h μ_T–quasiintegrierbar ist, und in diesem Fall gilt für alle $A' \in \mathcal{F}'$

$$\int_{T^{-1}(A')} (h \circ T)(\omega)\, d\mu(\omega) = \int_{A'} h(\omega')\, d\mu_T(\omega')$$

9.5 Integration nach einem eingeschränkten Maß

Das μ–Integral wurde bisher nur für solche Funktionen definiert, die auf der gesamten Grundmenge Ω oder auf dem Komplement $\Omega \setminus N$ einer μ–Nullmenge N definiert sind. Gelegentlich treten allerdings auch Funktionen auf, die nur auf einer messbaren Menge $C \in \mathcal{F}$ definiert sind, deren Komplement keine μ–Nullmenge ist.

Im gesamten Abschnitt sei $C \in \mathcal{F}$. Dann ist das Mengensystem

$$\mathcal{F}(C) := \Big\{ D \in \mathcal{F} \;\Big|\; D = A \cap C \text{ für eine Menge } A \in \mathcal{F} \Big\}$$

eine σ–Algebra auf C und die Abbildung $S_C : C \to \Omega$ mit

$$S_C(c) := c$$

ist messbar, denn für alle $A \in \mathcal{F}$ gilt $S_C^{-1}(A) = A \cap C \in \mathcal{F}(C)$.

– Ist $\nu : \mathcal{F} \to [0, \infty]$ ein Maß, so heißt das Maß $\nu|_{\mathcal{F}(C)} : \mathcal{F}(C) \to [0, \infty]$ mit

$$\nu|_{\mathcal{F}(C)}[D] := \nu[D]$$

die *Restriktion* von ν auf $\mathcal{F}(C)$.

– Ist $(\Omega', \mathcal{F}')$ ein Messraum und $h : \Omega \to \Omega'$ eine messbare Abbildung, so heißt die Abbildung $h|_C : C \to \Omega'$ mit

$$(h|_C)(c) := h(c)$$

die *Restriktion* von h auf C; wegen $h|_C = h \circ S_C$ ist $h|_C$ messbar.

Wir benötigen die folgende Darstellung des Bildmaßes von $\mu|_{\mathcal{F}(C)}$ unter S_C:

9.5.1 Lemma. *Es gilt*

$$(\mu|_{\mathcal{F}(C)})_{S_C} = \int \chi_C \, d\mu$$

Beweis. Für alle $A \in \mathcal{F}$ gilt

$$\begin{aligned}
(\mu|_{\mathcal{F}(C)})_{S_C}[A] &= (\mu|_{\mathcal{F}(C)})[S_C^{-1}(A)] \\
&= (\mu|_{\mathcal{F}(C)})[A \cap C] \\
&= \mu[A \cap C] \\
&= \int_{A \cap C} d\mu \\
&= \int_A \chi_C \, d\mu
\end{aligned}$$

Damit ist die Behauptung gezeigt. □

Aus Lemma 9.5.1 ergibt sich eine Darstellung des Integrals der Restriktion einer messbaren Funktion bezüglich der Restriktion des Maßes:

9.5.2 Lemma. *Sei h eine messbare Funktion.*
(1) *Ist h positiv, so gilt für alle $D \in \mathcal{F}(C)$*

$$\int_D (h|_C)(c)\, d\mu|_{\mathcal{F}(C)}(c) = \int_D h(\omega)\, d\mu(\omega)$$

(2) *$h|_C$ ist genau dann $\mu|_{\mathcal{F}(C)}$–integrierbar, wenn $h\chi_C$ μ–integrierbar ist, und in diesem Fall gilt für alle $D \in \mathcal{F}(C)$*

$$\int_D (h|_C)(c)\, d\mu|_{\mathcal{F}(C)}(c) = \int_D h(\omega)\, d\mu(\omega)$$

Beweis. Ist h positiv, so ist auch $h|_C$ positiv und aus der Substitutionsregel und der Kettenregel folgt für alle $A \in \mathcal{F}$

$$\begin{aligned}\int_{A\cap C} (h|_C)(c)\, d(\mu|_{\mathcal{F}(C)})(c) &= \int_{S_C^{-1}(A)} (h \circ S_C)(c)\, d(\mu|_{\mathcal{F}(C)})(c)\\ &= \int_A h(\omega)\, d(\mu|_{\mathcal{F}(C)})_{S_C}(\omega)\\ &= \int_A h(\omega)\, \chi_C(\omega)\, d\mu(\omega)\\ &= \int_{A\cap C} h(\omega)\, d\mu(\omega)\end{aligned}$$

Ist h beliebig, so folgt aus der Substitutionsregel und der Kettenregel, dass $h|_C$ genau dann $\mu|_{\mathcal{F}(C)}$–integrierbar ist, wenn $h\chi_C$ μ–integrierbar ist, und auch in diesem Fall gilt für alle $A \in \mathcal{F}$

$$\int_{A\cap C} (h|_C)(c)\, d(\mu|_{\mathcal{F}(C)})(c) = \int_{A\cap C} h(\omega)\, d\mu(\omega)$$

Die Behauptung folgt nun aus der Definition von $\mathcal{F}(C)$. □

Das Lemma legt es nahe, für eine $\mathcal{F}(C)$–$\mathcal{B}(\bar{\mathbb{R}})$–messbare positive oder $\mu|_{\mathcal{F}(C)}$–integrierbare Funktion $g : C \to \bar{\mathbb{R}}$ und $D \in \mathcal{F}(C)$ die Notation zu vereinfachen und

$$\int_D g\, d\mu := \int_D g\, d(\mu|_{\mathcal{F}(C)})$$

zu setzen und die Funktion g als *Lebesgue–integrierbar* oder als *μ–integrierbar* oder kurz als *integrierbar* zu bezeichnen, wenn sie $\mu|_{\mathcal{F}(C)}$–integrierbar ist.

Mit Hilfe geeigneter Restriktion von Maßen und messbaren Funktionen ist es insbesondere möglich, Folgerung 9.4.2 auf Transformationen zu erweitern, die nur lokal injektiv sind:

9.5.3 Lemma. *Sei h eine positive messbare Funktion und sei*

$$\nu := \int h(\omega)\, d\mu(\omega)$$

Sei ferner $(\Omega', \mathcal{F}')$ ein Messraum und $T : \Omega \to \Omega'$ messbar. Ist die Restriktion $T|_C$ injektiv mit $T(C) \in \mathcal{F}'$ und ihre Umkehrfunktion $(T|_C)^{-1} : T(C) \to C$ messbar, so gilt für alle $B' \in \mathcal{F}'$

$$(\nu|_{\mathcal{F}(C)})_{(T|_C)}[B'] = \int_{B'} h\Bigl((T|_C)^{-1}(\omega')\Bigr)\, d(\mu|_{\mathcal{F}(C)})_{(T|_C)}(\omega')$$

Beweis. Nach Lemma 9.5.2 gilt für alle $D \in \mathcal{F}(C)$

$$(\nu|_{\mathcal{F}(C)})[D] = \nu[D] = \int_D h(\omega)\, d\mu(\omega) = \int_D (h|_C)(c)\, d(\mu|_{\mathcal{F}(C)})(c)$$

Es gilt also

$$\nu|_{\mathcal{F}(C)} = \int (h|_C)(c)\, d(\mu|_{\mathcal{F}(C)})(c)$$

Die Behauptung ergibt sich nun aus Folgerung 9.4.2. □

Wir gelangen damit zu dem folgenden allgemeinen Ergebnis:

9.5.4 Satz. *Sei h eine positive messbare Funktion und sei*

$$\nu := \int h(\omega)\, d\mu(\omega)$$

Sei ferner $(\Omega', \mathcal{F}')$ ein Messraum und $T : \Omega \to \Omega'$ messbar. Ist $\{C_i\}_{i \in I} \subseteq \mathcal{F}$ eine abzählbare disjunkte Familie nichtleerer Mengen mit $\nu[\Omega \setminus \sum_{i \in I} C_i] = 0$ derart, dass für alle $i \in I$ die Restriktion $T|_{C_i}$ injektiv mit $T(C_i) \in \mathcal{F}'$ und ihre Umkehrfunktion $(T|_{C_i})^{-1} : T(C_i) \to C_i$ messbar ist, so gilt für alle $B' \in \mathcal{F}'$

$$\nu_T[B'] = \sum_{i \in I} \int_{B' \cap T(C_i)} h\Bigl((T|_{C_i})^{-1}(\omega')\Bigr)\, d(\mu|_{\mathcal{F}(C_i)})_{(T|_{C_i})}(\omega')$$

Beweis. Sei $B' \in \mathcal{F}'$. Dann gilt für alle $i \in I$

$$T^{-1}(B') \cap C_i = (T|_{C_i})^{-1}(B' \cap T(C_i))$$

und aus Lemma 9.5.3 ergibt sich nun

$$\begin{aligned}
\nu_T[B'] &= \nu[T^{-1}(B')] \\
&= \nu\left[T^{-1}(B') \cap \sum_{i\in I} C_i\right] \\
&= \sum_{i\in I} \nu[T^{-1}(B') \cap C_i] \\
&= \sum_{i\in I} (\nu|_{\mathcal{F}(C_i)})[T^{-1}(B') \cap C_i] \\
&= \sum_{i\in I} (\nu|_{\mathcal{F}(C_i)})[(T|_{C_i})^{-1}(B' \cap T(C_i))] \\
&= \sum_{i\in I} (\nu|_{\mathcal{F}(C_i)})_{(T|_{C_i})}[B' \cap T(C_i)] \\
&= \sum_{i\in I} \int_{B'\cap T(C_i)} h\Big((T|_{C_i})^{-1}(\omega')\Big)\, d(\mu|_{\mathcal{F}(C_i)})_{(T|_{C_i})}(\omega')
\end{aligned}$$

Damit ist die Behauptung gezeigt. □

9.6 Produktmaße

In diesem Abschnitt betrachten wir zwei Maßräume $(M, \mathcal{M}, \mu)$ und $(N, \mathcal{N}, \nu)$ sowie das Produkt $(M, \mathcal{M}) \otimes (N, \mathcal{N}) = (M \times N, \mathcal{M} \otimes \mathcal{N})$ der Messräume $(M, \mathcal{M})$ und $(N, \mathcal{N})$. Nach Satz 3.3.4 gilt $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{M} \times \mathcal{N})$, und nach Aufgabe 3.1.A ist $\mathcal{M} \times \mathcal{N}$ ein Halbring.

Ein Maß $\varrho : \mathcal{M} \otimes \mathcal{N} \to [0, \infty]$ heißt *Produktmaß* von μ und ν, wenn für alle $A \times B \in \mathcal{M} \times \mathcal{N}$

$$\varrho[A \times B] = \mu[A]\, \nu[B]$$

gilt und es kein weiteres Maß mit dieser Eigenschaft gibt. Im Fall der Existenz bezeichnen wir das Produktmaß von μ und ν mit

$$\mu \otimes \nu$$

und nennen den Maßraum

$$(M, \mathcal{M}, \mu) \otimes (N, \mathcal{N}, \nu) := (M \times N, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$$

das *Produkt* der Maßräume $(M, \mathcal{M}, \mu)$ und $(N, \mathcal{N}, \nu)$.

Wir betrachten daher die Abbildung $\varrho : \mathcal{M} \times \mathcal{N} \to [0, \infty]$ mit

$$\varrho[A \times B] := \mu[A]\, \nu[B]$$

und zeigen zunächst, dass ϱ ein Maß ist. Dazu benötigen wir eine Darstellung der Werte von ϱ durch Integrale.

Für $C \in 2^{M \times N}$ und $x \in M$ heißt die Menge

$$C_x := \Big\{ y \in N \Big| (x,y) \in C \Big\}$$

x–Schnitt von C. Das folgende Lemma liefert die angekündigte Darstellung der Werte von ϱ durch Integrale:

9.6.1 Lemma. *Sei $A \times B \in \mathcal{M} \times \mathcal{N}$. Für alle $x \in M$ gilt dann $(A \times B)_x \in \mathcal{N}$ und*

$$\nu[(A \times B)_x] = \nu[B]\, \chi_A(x)$$

Insbesondere ist die Abbildung $M \to \bar{\mathbb{R}} : x \mapsto \nu[(A \times B)_x]$ messbar und positiv, und es gilt

$$\varrho[A \times B] = \int_M \nu[(A \times B)_x]\, d\mu(x)$$

Beweis. Für alle $x \in M$ gilt

$$(A \times B)_x = \begin{cases} B & \text{falls } x \in A \\ \emptyset & \text{sonst} \end{cases}$$

und damit $(A \times B)_x \in \mathcal{N}$ und

$$\nu[(A \times B)_x] = \nu[B]\, \chi_A(x)$$

Daher ist die Abbildung $M \to \bar{\mathbb{R}} : x \mapsto \nu[(A \times B)_x]$ messbar und positiv, und es gilt

$$\begin{aligned} \varrho\big[A \times B\big] &= \mu[A]\, \nu[B] \\ &= \int_A \nu[B]\, d\mu(x) \\ &= \int_M \nu[B]\, \chi_A(x)\, d\mu(x) \\ &= \int_M \nu[(A \times B)_x]\, d\mu(x) \end{aligned}$$

Damit ist das Lemma bewiesen. □

Mit Hilfe von Lemma 9.6.1 können wir nun zeigen, dass ϱ ein Maß ist:

9.6.2 Lemma. *Die Abbildung $\varrho : \mathcal{M} \times \mathcal{N} \to [0, \infty]$ ist ein Maß.*

Beweis. Sei $\emptyset = A \times B$ eine Darstellung der leeren Menge $\emptyset \in 2^{M \times N}$. Dann gilt $A = \emptyset$ oder $B = \emptyset$, und damit $\mu[A] = 0$ oder $\nu[B] = 0$. Daher gilt

$$\varrho[\emptyset] = \varrho[A \times B] = \mu[A]\, \nu[B] = 0$$

Sei nun $\{C_k\}_{k\in\mathbb{N}} \subseteq \mathcal{M}\times\mathcal{N}$ eine disjunkte Folge mit $\sum_{k=1}^\infty C_k \in \mathcal{M}\times\mathcal{N}$. Für alle $x \in M$ gilt dann $(\sum_{k=1}^\infty C_k)_x = \sum_{k=1}^\infty (C_k)_x$, und aus Lemma 9.6.1 und dem Satz über die monotone Konvergenz folgt nun

$$\begin{aligned}
\varrho\left[\sum_{k=1}^\infty C_k\right] &= \int_M \nu\left[\left(\sum_{k=1}^\infty C_k\right)_x\right] d\mu(x)\\
&= \int_M \nu\left[\sum_{k=1}^\infty (C_k)_x\right] d\mu(x)\\
&= \int_M \left(\sum_{k=1}^\infty \nu[(C_k)_x]\right) d\mu(x)\\
&= \sum_{k=1}^\infty \int_M \nu[(C_k)_x]\, d\mu(x)\\
&= \sum_{k=1}^\infty \varrho[C_k]
\end{aligned}$$

Daher ist ϱ σ–additiv. Damit ist gezeigt, dass ϱ ein Maß ist. □

Da das Mengensystem $\mathcal{M}\times\mathcal{N}$ ein Halbring ist, folgt aus Lemma 9.6.2 und dem Existenzsatz die Existenz einer Fortsetzung von ϱ auf $\mathcal{M}\otimes\mathcal{N}$. Da das Mengensystem $\mathcal{M}\times\mathcal{N}$ insbesondere $\cap$–stabil ist, erhält man unter der zusätzlichen Voraussetzung, dass ϱ σ–endlich ist, aus dem Eindeutigkeitssatz außerdem die Eindeutigkeit der Fortsetzung. Die σ–Endlichkeit von ϱ ist gewährleistet, wenn μ und ν σ–endlich sind. Aus dem Satz von Caratheodory erhält man daher das folgende Ergebnis:

9.6.3 Satz (Existenzsatz). *Seien μ und ν σ–endlich. Dann besitzen μ und ν ein Produktmaß, und das Produktmaß ist σ–endlich.*

Unser nächstes Ziel ist es, die in Lemma 9.6.1 angegebene Darstellung der Werte von ϱ durch Integrale zu einer Darstellung der Werte des Produktmaßes $\mu\otimes\nu$ durch Integrale zu verallgemeinern.

9.6.4 Lemma. *Sei $C \in \mathcal{M}\otimes\mathcal{N}$. Für alle $x \in M$ gilt dann $C_x \in \mathcal{N}$.*

Beweis. Sei $x \in M$ und

$$\mathcal{D}_x := \left\{ D \in 2^{M\times N} \;\middle|\; D_x \in \mathcal{N}\right\}$$

Dann ist $\mathcal{D}_x$ ein Dynkin–System:
(i) Es gilt $(M\times N)_x = N \in \mathcal{N}$ und damit $M\times N \in \mathcal{D}_x$.
(ii) Sei $D \in \mathcal{D}_x$. Dann gilt $D_x \in \mathcal{N}$ und $D_x + ((M\times N)\setminus D)_x = (M\times N)_x = N \in \mathcal{N}$, also $((M\times N)\setminus D)_x = N\setminus D_x \in \mathcal{N}$, und damit $(M\times N)\setminus D \in \mathcal{D}_x$.
(iii) Sei $\{D_k\}_{k\in\mathbb{N}} \subseteq \mathcal{D}_x$ disjunkt. Dann gilt $\left(\sum_{k=1}^\infty D_k\right)_x = \sum_{k=1}^\infty (D_k)_x \in \mathcal{N}$ und damit $\sum_{k=1}^\infty D_k \in \mathcal{D}_x$.

Außerdem ist das Mengensystem $\mathcal{M} \times \mathcal{N}$ $\cap$–stabil und nach Lemma 9.6.1 gilt $\mathcal{M} \times \mathcal{N} \subseteq \mathcal{D}_x$. Daraus folgt $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{M} \times \mathcal{N}) = \delta(\mathcal{M} \times \mathcal{N}) \subseteq \bigcap_{x \in M} \mathcal{D}_x$. $\square$

9.6.5 Lemma. *Sei ν σ–endlich. Dann ist für alle $C \in \mathcal{M} \otimes \mathcal{N}$ die Abbildung $M \to \bar{\mathbb{R}} : x \mapsto \nu[C_x]$ messbar und positiv.*

Beweis. Für $C \in \mathcal{M} \otimes \mathcal{N}$ sei die Funktion $f_C : M \to \bar{\mathbb{R}}$ gegeben durch

$$f_C(x) := \nu[C_x]$$

Dann ist f_C positiv und es bleibt zu zeigen, dass f_C messbar ist.
(1) Wir nehmen zunächst an, dass ν endlich ist. Sei

$$\mathcal{D} := \left\{ D \in 2^{M \times N} \;\middle|\; f_D \text{ ist messbar} \right\}$$

Dann ist $\mathcal{D}$ ein Dynkin–System:
(i) Für alle $x \in M$ gilt

$$f_{M \times N}(x) = \nu[(M \times N)_x] = \nu[N]$$

Daher ist $f_{M \times N}$ messbar, und daraus folgt $M \times N \in \mathcal{D}$.
(ii) Sei $D \in \mathcal{D}$. Für alle $x \in M$ gilt $((M \times N) \setminus D)_x = N \setminus D_x$ und damit

$$\begin{aligned} f_{(M \times N) \setminus D}(x) &= \nu[((M \times N) \setminus D)_x] \\ &= \nu[N \setminus D_x] \\ &= \nu[N] - \nu[D_x] \\ &= f_{M \times N}(x) - f_D(x) \end{aligned}$$

Daher ist $f_{(M \times N) \setminus D}$ messbar, und daraus folgt $(M \times N) \setminus D \in \mathcal{D}$. .
(iii) Sei $\{D_k\}_{k \in \mathbb{N}} \subseteq \mathcal{D}$ disjunkt. Für alle $x \in M$ gilt $(\sum_{k=1}^\infty D_k)_x = \sum_{k=1}^\infty (D_k)_x$ und damit

$$\begin{aligned} f_{\sum_{k=1}^\infty D_k}(x) &= \nu\left[\left(\sum_{k=1}^\infty D_k\right)_x\right] \\ &= \nu\left[\sum_{k=1}^\infty (D_k)_x\right] \\ &= \sum_{k=1}^\infty \nu[(D_k)_x] \\ &= \sum_{k=1}^\infty f_{D_k}(x) \end{aligned}$$

Daher ist $f_{\sum_{k=1}^\infty D_k}(x)$ messbar, und daraus folgt $\sum_{k=1}^\infty D_k \in \mathcal{D}$.
Damit ist gezeigt, dass $\mathcal{D}$ ein Dynkin–System ist. Außerdem ist das Mengensystem $\mathcal{M} \times \mathcal{N}$ $\cap$–stabil und nach Lemma 9.6.1 gilt $\mathcal{M} \times \mathcal{N} \subseteq \mathcal{D}$. Daraus folgt $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{M} \times \mathcal{N}) = \delta(\mathcal{M} \times \mathcal{N}) \subseteq \mathcal{D}$. Daher ist die Abbildung f_C für alle $C \in \mathcal{M} \otimes \mathcal{N}$ messbar.

(2) Wir nehmen nun an, dass ν σ–endlich ist. Dann gibt es eine monoton wachsende Folge $\{B_n\}_{n\in\mathbb{N}} \subseteq \mathcal{N}$ mit $\bigcup_{n\in\mathbb{N}} B_n = N$ und $\nu[B_n] < \infty$ für alle $n \in \mathbb{N}$, und für alle $n \in \mathbb{N}$ ist die Mengenfunktion $\nu_n : \mathcal{N} \to [0,\infty]$ mit

$$\nu_n[B] := \nu[B \cap B_n]$$

ein endliches Maß. Für $C \in \mathcal{M} \otimes \mathcal{N}$ ist daher nach (1) für alle $n \in \mathbb{N}$ die Funktion $f_{n,C} : M \to \bar{\mathbb{R}}$ mit

$$f_{n,C}(x) := \nu_n[C_x]$$

messbar und positiv, und wegen

$$f_C(x) = \nu[C_x] = \sup_{n\in\mathbb{N}} \nu[C_x \cap B_n] = \sup_{n\in\mathbb{N}} \nu_n[C_x] = \sup_{n\in\mathbb{N}} f_{n,C}(x)$$

ist auch f_C messbar und positiv. $\square$

Mit diesen Vorbereitungen können wir die angekündigte Verallgemeinerung der Integraldarstellung von Lemma 9.6.1 beweisen:

9.6.6 Satz. *Seien μ und ν σ–endlich. Dann gilt für alle $C \in \mathcal{M} \otimes \mathcal{N}$*

$$\big(\mu \otimes \nu\big)[C] = \int_M \nu[C_x]\, d\mu(x)$$

Beweis. Nach Lemma 9.6.5 ist die Funktion $M \to \bar{\mathbb{R}} : x \mapsto \nu[C_x]$ messbar und positiv. Wir betrachten die Mengenfunktion $\varrho_M : \mathcal{M} \otimes \mathcal{N} \to [0,\infty]$ mit

$$\varrho_M[C] := \int_M \nu[C_x]\, d\mu(x)$$

Dann gilt $\varrho_M[\emptyset] = 0$ und für jede disjunkte Folge $\{C_k\}_{k\in\mathbb{N}} \subseteq \mathcal{M} \times \mathcal{N}$ mit $\sum_{k=1}^\infty C_k \in \mathcal{M} \times \mathcal{N}$ gilt nach dem Satz über die monotone Konvergenz

$$\begin{aligned}
\varrho_M\left[\sum_{k=1}^\infty C_k\right] &= \int_M \nu\left[\left(\sum_{k=1}^\infty C_k\right)_x\right] d\mu(x) \\
&= \int_M \nu\left[\sum_{k=1}^\infty (C_k)_x\right] d\mu(x) \\
&= \int_M \left(\sum_{l=1}^\infty \nu[(C_k)_x]\right) d\mu(x) \\
&= \sum_{k=1}^\infty \int_M \nu[(C_k)_x]\, d\mu(x) \\
&= \sum_{l=1}^\infty \varrho_M[C_k]
\end{aligned}$$

Daher ist ϱ_M ein Maß. Außerdem gilt für alle $A \times B \in \mathcal{M} \times \mathcal{N}$

$$\begin{aligned}\varrho_M[A \times B] &= \int_M \nu[(A \times B)_x]\, d\mu(x) \\ &= \int_M \nu[B]\, \chi_A(x)\, d\mu(x) \\ &= \mu[A]\, \nu[B] \\ &= \varrho[A \times B]\end{aligned}$$

Daher gilt $\varrho_M|_{\mathcal{M} \times \mathcal{N}} = \varrho$. Da μ und ν σ–endlich sind, ist auch ϱ σ–endlich, und aus dem Eindeutigkeitssatz folgt nun $\varrho_M = \mu \otimes \nu$. □

Wir dehnen die Konstruktion des Produktmaßes nun auf endlich viele σ–endliche Maßräume aus:

Sei $\{(\Omega_i, \mathcal{F}_i, \mu_i)\}_{i \in \{1,\dots,n\}}$ eine endliche Familie von Maßräumen. Ein Maß $\varrho : \bigotimes_{i=1}^n \mathcal{F}_i \to [0, \infty]$ heißt *Produktmaß* der Familie $\{\mu_i\}_{i \in \{1,\dots,n\}}$, wenn für alle $\prod_{i=1}^n A_i \in \prod_{i=1}^n \mathcal{F}_i$

$$\varrho\left[\prod_{i=1}^n A_i\right] = \prod_{i=1}^n \mu_i[A_i]$$

gilt und es kein weiteres Maß mit dieser Eigenschaft gibt. Im Fall der Existenz bezeichnen wir das Produktmaß der Familie $\{\mu_i\}_{i \in \{1,\dots,n\}}$ mit

$$\bigotimes_{i=1}^n \mu_i$$

und nennen den Maßraum

$$\bigotimes_{i=1}^n (\Omega_i, \mathcal{F}_i, \mu_i) := \left(\prod_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathcal{F}_i, \bigotimes_{i=1}^n \mu_i\right)$$

das *Produkt* der Familie der Maßräume $\{(\Omega_i, \mathcal{F}_i, \mu_i)\}_{i \in \{1,\dots,n\}}$.

9.6.7 Satz. *Sei $\{(\Omega_i, \mathcal{F}_i, \mu_i)\}_{i \in \{1,\dots,n\}}$ eine Familie von σ–endlichen Maßräumen. Dann gibt es genau ein Maß $\varrho : \bigotimes_{i=1}^n \mathcal{F}_i \to [0, \infty]$ mit*

$$\varrho\left[\prod_{i=1}^n A_i\right] = \prod_{i=1}^n \mu_i[A_i]$$

für alle $\prod_{i=1}^n A_i \in \prod_{i=1}^n \mathcal{F}_i$. Das Maß ϱ ist σ–endlich.

Beweis. Die Behauptung ergibt sich unter Verwendung von Aufgabe 3.3.A und Aufgabe 9.6.A durch vollständige Induktion aus Satz 9.6.3. □

Aus dem letzten Satz erhalten wir eine Darstellung des n–dimensionalen Lebesgue–Maßes als Produktmaß:

9.6.8 Folgerung (Lebesgue–Maß). *Es gilt*

$$\boldsymbol{\lambda}^n = \bigotimes_{i=1}^n \boldsymbol{\lambda}$$

Beweis. Da das Lebesgue–Maß $\boldsymbol{\lambda}$ σ–endlich ist, existiert das Produktmaß $\bigotimes_{i=1}^n \boldsymbol{\lambda}$, und nach Beispiel 3.3.1 gilt $\mathcal{B}(\mathbb{R}^n) = \bigotimes_{i=1}^n \mathcal{B}(\mathbb{R})$. Daher ist das n–dimensionale Lebesgue–Maß $\boldsymbol{\lambda}^n$ auf der selben σ–Algebra definiert wie das Produktmaß $\bigotimes_{i=1}^n \boldsymbol{\lambda}$.
Für alle $(\mathbf{a}, \mathbf{b}] \in \mathcal{J}(\mathbb{R}^n)$ mit $\mathbf{a} < \mathbf{b}$ gilt

$$\begin{aligned}
\boldsymbol{\lambda}^n\big[(\mathbf{a}, \mathbf{b}]\big] &= \prod_{i=1}^n (b_i - a_i) \\
&= \prod_{i=1}^n \boldsymbol{\lambda}\big[(a_i, b_i]\big] \\
&= \left(\bigotimes_{i=1}^n \boldsymbol{\lambda}\right)\left[\prod_{i=1}^n (a_i, b_i]\right] \\
&= \left(\bigotimes_{i=1}^n \boldsymbol{\lambda}\right)\big[(\mathbf{a}, \mathbf{b}]\big]
\end{aligned}$$

Daher stimmen die Maße $\boldsymbol{\lambda}^n$ und $\bigotimes_{i=1}^n \boldsymbol{\lambda}$ auf dem Halbring $\mathcal{J}(\mathbb{R}^n)$ überein. Da $\mathcal{J}(\mathbb{R}^n)$ als Halbring $\cap$–stabil ist und die Restriktion von $\boldsymbol{\lambda}^n$ auf $\mathcal{J}(\mathbb{R}^n)$ σ–endlich ist, folgt aus dem Eindeutigkeitssatz, dass die Maße $\boldsymbol{\lambda}^n$ und $\bigotimes_{i=1}^n \boldsymbol{\lambda}$ sogar auf $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{J}(\mathbb{R}^n))$ übereinstimmen. □

Im Fall $(\Omega_i, \mathcal{F}_i, \mu_i) = (\Omega, \mathcal{F}, \mu)$ für alle $i \in \{1, \ldots, n\}$ setzen wir, im Fall der Existenz des Produktmaßes,

$$\mu^n := \bigotimes_{i=1}^n \mu$$

Dies steht im Einklang mit dem letzten Ergebnis.

Aufgabe

9.6.A Für $i \in \{1, 2, 3\}$ sei $(\Omega_i, \mathcal{F}_i, \mu_i)$ ein σ–endlicher Maßraum. Dann gilt

$$(\mu_1 \otimes \mu_2) \otimes \mu_3 = \mu_1 \otimes (\mu_2 \otimes \mu_3)$$

9.7 Integration nach einem Produktmaß

Im gesamten Abschnitt seien $(M, \mathcal{M}, \mu)$ und $(N, \mathcal{N}, \nu)$ Maßräume.

Sind μ und ν σ–endlich, so existiert nach Satz 9.6.3 das Produktmaß $\mu \otimes \nu$, und wir zeigen nun, dass das Integral

$$\int_{M \times N} h(x,y)\, d(\mu \otimes \nu)(x,y)$$

einer messbaren Funktion $h : M \times N \to \bar{\mathbb{R}}$, die positiv oder $\mu \otimes \nu$–integrierbar ist, iterativ berechnet werden kann.

Für eine Funktion $h : M \times N \to \bar{\mathbb{R}}$ und $x \in M$ heißt die Funktion $h_x : N \to \bar{\mathbb{R}}$ mit

$$h_x(y) := h(x,y)$$

der *x–Schnitt* von h.

9.7.1 Lemma. *Sei $h : M \times N \to \bar{\mathbb{R}}$ eine Funktion und sei $x \in M$.*
(1) *Für alle $D \in \mathcal{B}(\bar{\mathbb{R}})$ gilt*

$$(h_x)^{-1}(D) = (h^{-1}(D))_x$$

(2) *Ist h messbar, so ist auch h_x messbar.*

Beweis. Es gilt

$$\begin{aligned}
(h_x)^{-1}(D) &= \Bigl\{ y \in N \Bigm| h_x(y) \in D \Bigr\} \\
&= \Bigl\{ y \in N \Bigm| h(x,y) \in D \Bigr\} \\
&= \Bigl\{ y \in N \Bigm| (x,y) \in h^{-1}(D) \Bigr\} \\
&= (h^{-1}(D))_x
\end{aligned}$$

Ist h messbar, so folgt aus dieser Gleichung und Lemma 9.6.4, dass auch h_x messbar ist. □

Wir betrachten zunächst positive messbare Funktionen. Ist $h : M \times N \to \bar{\mathbb{R}}$ eine positive messbare Funktion, so ist für jedes $x \in M$ auch die Funktion h_x positiv und nach Lemma 9.7.1 messbar. Daher ist die Funktion $f_h : M \to \bar{\mathbb{R}}$ mit

$$f_h(x) := \int_N h_x(y)\, d\nu(y)$$

wohldefiniert und positiv.

9.7.2 Satz (Fubini; positiver Fall). *Seien μ und ν σ–endlich und sei $h : M \times N \to \bar{\mathbb{R}}$ messbar und positiv. Dann ist auch f_h messbar und positiv, und es gilt*

$$\int_{M\times N} h(x,y)\, d(\mu\otimes\nu)(x,y) = \int_M f_h(x)\, d\mu(x)$$

Insbesondere ist h genau dann $\mu\otimes\nu$–integrierbar, wenn f_h μ–integrierbar ist.

Beweis. Wir führen den Beweis der Messbarkeit von f_h und der Gleichheit der Integrale durch algebraische Induktion:
(1) Sei $C \in \mathcal{M}\otimes\mathcal{N}$. Da χ_C messbar und positiv ist, ist für alle $x \in M$ auch $(\chi_C)_x = \chi_{C_x}$ messbar und positiv, und aus Lemma 9.7.1 folgt dann

$$f_{\chi_C}(x) = \int_N (\chi_C)_x(y)\, d\nu(y) = \int_N \chi_{C_x}(y)\, d\nu(y) = \nu[C_x]$$

Nach Lemma 9.6.5 ist f_{χ_C} messbar und aus Satz 9.6.6 folgt nun

$$\begin{aligned}
\int_M f_{\chi_C}(x)\, d\mu(x) &= \int_M \nu[C_x]\, d\mu(x) \\
&= (\mu\otimes\nu)[C] \\
&= \int_{M\times N} \chi_C(x,y)\, d(\mu\otimes\nu)(x,y)
\end{aligned}$$

Damit ist die Behauptung für $h = \chi_C$ mit $C \in \mathcal{M}\otimes\mathcal{N}$ gezeigt.
(2) Sei h eine positive einfache Funktion mit

$$h = \sum_{i=1}^n c_i\, \chi_{C_i}$$

mit $\{C_i\}_{i\in\{1,\dots,n\}} \subseteq \mathcal{M}\otimes\mathcal{N}$ und $\{c_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}_+$. Für alle $x \in M$ gilt $h_x = (\sum_{i=1}^n c_i\, \chi_{C_i})_x = \sum_{i=1}^n c_i\, (\chi_{C_i})_x$, und aus der positiven Linearität des Integrals folgt dann

$$\begin{aligned}
f_h(x) &= \int_N h_x(y)\, d\nu(y) \\
&= \int_N \left(\sum_{i=1}^n c_i\, (\chi_{C_i})_x(y) \right) d\nu(y) \\
&= \sum_{i=1}^n c_i \int_N (\chi_{C_i})_x(y)\, d\nu(y) \\
&= \sum_{i=1}^n c_i\, f_{\chi_{C_i}}(x)
\end{aligned}$$

Aus (1) ergibt sich nun zunächst die Messbarkeit von f_h, und aus der positiven Linearität des Integrals und (1) ergibt sich des weiteren

$$
\begin{aligned}
\int_M f_h(x)\,d\mu(x) &= \int_M \left(\sum_{i=1}^n c_i\, f_{\chi_{C_i}}(x)\right) d\mu(x) \\
&= \sum_{i=1}^n c_i \int_M f_{\chi_{C_i}}(x)\,d\mu(x) \\
&= \sum_{i=1}^n c_i \int_{M\times N} \chi_{C_i}(x,y)\,d(\mu\otimes\nu)(x,y) \\
&= \int_{M\times N} \left(\sum_{i=1}^n c_i\,\chi_{C_i}(x,y)\right) d(\mu\otimes\nu)(x,y) \\
&= \int_{M\times N} h(x,y)\,d(\mu\otimes\nu)(x,y)
\end{aligned}
$$

Damit ist die Behauptung für positive einfache Funktionen gezeigt.
(3) Sei h eine positive messbare Funktion. Dann gibt es eine monoton wachsende Folge $\{h_n\}_{n\in\mathbb{N}}$ von positiven einfachen Funktionen mit $h = \sup_{n\in\mathbb{N}} h_n$. Für alle $x \in M$ gilt $h_x = (\sup_{n\in\mathbb{N}} h_n)_x = \sup_{n\in\mathbb{N}}(h_n)_x$, und aus dem Satz über die monotone Konvergenz folgt dann

$$
\begin{aligned}
f_h(x) &= \int_N h_x(y)\,d\nu(y) \\
&= \int_N \sup_{n\in\mathbb{N}}(h_n)_x(y)\,d\nu(y) \\
&= \sup_{n\in\mathbb{N}} \int_N (h_n)_x(y)\,d\nu(y) \\
&= \sup_{n\in\mathbb{N}} f_{h_n}(x)
\end{aligned}
$$

Aus (2) ergibt sich nun zunächst die Messbarkeit von f_h, und aus dem Satz über die monotone Konvergenz und (2) ergibt sich des weiteren

$$
\begin{aligned}
\int_M f_h(x)\,d\mu(x) &= \int_M \sup_{n\in\mathbb{N}} f_{h_n}(x)\,d\mu(x) \\
&= \sup_{n\in\mathbb{N}} \int_M f_{h_n}(x)\,d\mu(x) \\
&= \sup_{n\in\mathbb{N}} \int_{M\times N} h_n(x,y)\,d(\mu\otimes\nu)(x,y) \\
&= \int_{M\times N} \sup_{n\in\mathbb{N}} h_n(x,y)\,d(\mu\otimes\nu)(x,y) \\
&= \int_{M\times N} h(x,y)\,d(\mu\otimes\nu)(x,y)
\end{aligned}
$$

Damit ist der Satz bewiesen. □

Wir betrachten nun integrierbare Funktionen. Hierfür ist eine Vorüberlegung erforderlich:

Seien μ und ν σ–endlich und sei $h : M \times N \to \bar{\mathbb{R}}$ $\mu \otimes \nu$–integrierbar. Dann sind auch die Funktionen h^+ und h^- $\mu \otimes \nu$–integrierbar, und aus dem Satz von Fubini für positive messbare Funktionen folgt nun die μ–Integrierbarkeit von f_{h^+} und f_{h^-}. Dann sind aber f_{h^+} und f_{h^-} μ–fast überall endlich, und damit gibt es eine μ–Nullmenge $M_0 \in \mathcal{M}$ derart, dass für alle $x \in M \setminus M_0$ die Funktionen $(h^+)_x$ und $(h^-)_x$ ν–integrierbar sind; wegen $h_x = (h_x)^+ - (h_x)^-$ sowie $(h_x)^+ = (h^+)_x$ und $(h_x)^- = (h^-)_x$ ist dann für alle $x \in M \setminus M_0$ auch die Funktion h_x ν–integrierbar und es gilt

$$\begin{aligned}\int_N h_x(y)\, d\nu(y) &= \int_N (h^+)_x(y)\, d\nu(y) - \int_N (h^-)_x(y)\, d\nu(y) \\ &= f_{h^+}(x) - f_{h^-}(x)\end{aligned}$$

Sei nun $f_h : M \to \bar{\mathbb{R}}$ gegeben durch

$$f_h(x) := \begin{cases} \displaystyle\int_N h_x(y)\, d\nu(y) & \text{falls } x \in M \setminus M_0 \\ 0 & \text{sonst} \end{cases}$$

Dann gilt $f_h(x) = f_{h^+}(x)\chi_{M\setminus M_0}(x) - f_{h^-}(x)\chi_{M\setminus M_0}(x)$ und damit ist auch f_h μ–integrierbar.

Der folgende Satz von Fubini für integrierbare Funktionen ergibt sich nun unmittelbar aus dem Satz von Fubini für positive messbare Funktionen:

9.7.3 Satz (Fubini; integrierbarer Fall). *Seien μ und ν σ–endlich und sei $h : M \times N \to \bar{\mathbb{R}}$ $\mu \otimes \nu$–integrierbar. Dann ist f_h μ–integrierbar und es gilt*

$$\int_{M\times N} h(x,y)\, d(\mu\otimes\nu)(x,y) = \int_M f_h(x)\, d\mu(x)$$

Im folgenden verzichten wir auf die Kennzeichnung der x–Schnitte und schreiben

$$f_h(x) = \int_N h(x,y)\, d\nu(y)$$

Entsprechend schreiben wir die Gleichung des Satzes von Fubini in der Form

$$\int_{M\times N} h(x,y)\, d(\mu\otimes\nu)(x,y) = \int_M \left(\int_N h(x,y)\, d\nu(y) \right) d\mu(x)$$

oder, unter Weglassung der Klammern, in der Form

$$\int_{M\times N} h(x,y)\, d(\mu\otimes\nu)(x,y) = \int_M \int_N h(x,y)\, d\nu(y)\, d\mu(x)$$

Vertauscht man schließlich die Rollen der Maße μ und ν, so erhält man unter den Voraussetzungen des Satzes von Fubini die zusätzliche Gleichung

$$\int_{M\times N} h(x,y)\, d(\mu\otimes\nu)(x,y) = \int_N \int_M h(x,y)\, d\mu(x)\, d\nu(y)$$

Beide Gleichungen lassen sich zu der Gleichung

$$\begin{aligned}\int_M \int_N h(x,y)\, d\nu(y)\, d\mu(x) &= \int_{M\times N} h(x,y)\, d(\mu\otimes\nu)(x,y)\\ &= \int_N \int_M h(x,y)\, d\mu(x)\, d\nu(y)\end{aligned}$$

zusammenfassen. Unter der Voraussetzung des Satzes von Fubini kann das Integral von h also in beliebiger Reihenfolge iterativ ausgewertet werden.

Neben der iterativen Berechnung des Integrals einer messbaren positiven oder integrierbaren Funktion auf dem Produkt zweier σ–endlicher Maßräume kann der Satz von Fubini auch für die Berechnung des Integrals einer messbaren positiven oder integrierbaren Funktion f auf einem σ–endlichen Maßraum $(M,\mathcal{M},\mu)$ von Nutzen sein; zu diesem Zweck muss ein σ–endlicher Maßraum $(N,\mathcal{N},\nu)$ und eine messbare positive oder integrierbare Funktion h auf $(M,\mathcal{M},\mu)\otimes(N,\mathcal{N},\nu)$ mit $f(x)=\int_N h(x,y)\, d\nu(y)$ gefunden werden, um den Satz von Fubini anwenden zu können.

9.7.4 Beispiel. Sei $(\Omega,\mathcal{F},\mu)$ ein σ–endlicher Maßraum und sei f messbar und positiv. Dann gilt

$$\int_\Omega f(\omega)\, d\mu(\omega) = \int_{\mathbb{R}_+} \mu[\{f\geq x\}]\, d\boldsymbol{\lambda}(x)$$

In der Tat: Wir betrachten den σ–endlichen Maßraum $(\mathbb{R},\mathcal{B}(\mathbb{R}),\boldsymbol{\lambda})$ und die Indikatorfunktion χ_C der Menge

$$C := \Big\{(\omega,x)\in\Omega\times\mathbb{R} \,\Big|\, 0\leq x\leq f(\omega)\Big\}$$

Die Indikatorfunktion χ_C ist positiv; außerdem ist sie genau dann messbar, wenn $C\in\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$ gilt. Wir zeigen daher zunächst, dass $C\in\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$ gilt. Dazu betrachten wir die Menge

$$D := \left\{\begin{pmatrix} x\\ y\end{pmatrix}\in\mathbb{R}^2 \,\middle|\, 0\leq x\leq y\right\}$$

und die Funktion $T:\Omega\times\mathbb{R}\to\mathbb{R}^2$ mit

$$T(\omega,x) := \begin{pmatrix} x\\ f(\omega)\end{pmatrix}$$

Dann gilt $C = T^{-1}(D)$. Da die Menge D abgeschlossen ist, gilt außerdem $D \in \mathcal{B}(\mathbb{R}^2)$. Die Koordinaten $T_1, T_2 : \Omega \times \mathbb{R} \to \mathbb{R}$ von T mit

$$\begin{aligned} T_1(\omega, x) &:= x \\ T_2(\omega, x) &:= f(\omega) \end{aligned}$$

sind wegen

$$\begin{aligned} T_1^{-1}(B) &:= \Omega \times B \\ T_2^{-1}(B) &:= f^{-1}(B) \times \mathbb{R} \end{aligned}$$

messbar; daher ist auch T messbar und es folgt $C = T^{-1}(D) \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R})$. Aus dem Satz von Fubini ergibt sich nun einerseits

$$\begin{aligned} \int_{\Omega \times \mathbb{R}} \chi_C(\omega, x)\, d(\mu \otimes \boldsymbol{\lambda})(\omega, x) &= \int_\Omega \int_\mathbb{R} \chi_C(\omega, x)\, d\boldsymbol{\lambda}(x)\, d\mu(\omega) \\ &= \int_\Omega \int_\mathbb{R} \chi_{[0, f(\omega)]}(x)\, d\boldsymbol{\lambda}(x)\, d\mu(\omega) \\ &= \int_\Omega \boldsymbol{\lambda}[[0, f(\omega)]]\, d\mu(\omega) \\ &= \int_\Omega f(\omega)\, d\mu(\omega) \end{aligned}$$

und andererseits

$$\begin{aligned} \int_{\Omega \times \mathbb{R}} \chi_C(\omega, x)\, d(\mu \otimes \boldsymbol{\lambda})(\omega, x) &= \int_\mathbb{R} \int_\Omega \chi_C(\omega, x)\, d\mu(\omega)\, d\boldsymbol{\lambda}(x) \\ &= \int_\mathbb{R} \int_\Omega \chi_{\mathbb{R}_+}(x)\, \chi_{\{f \geq x\}}(\omega)\, d\mu(\omega)\, d\boldsymbol{\lambda}(x) \\ &= \int_\mathbb{R} \chi_{\mathbb{R}_+}(x) \int_\Omega \chi_{\{f \geq x\}}(\omega)\, d\mu(\omega)\, d\boldsymbol{\lambda}(x) \\ &= \int_{\mathbb{R}_+} \mu[\{f \geq x\}]\, d\boldsymbol{\lambda}(x) \end{aligned}$$

Damit ist die Behauptung gezeigt.

Aufgabe

9.7.A Sei $(\Omega, \mathcal{F}, \mu)$ ein σ–endlicher Maßraum und sei f positiv und messbar. Dann gilt

$$\int_\Omega f(\omega)\, d\mu(\omega) = \int_{\mathbb{R}_+} \mu[\{f > x\}]\, d\boldsymbol{\lambda}(x)$$

9.8 Lebesgue–Integral und Riemann–Integral

In diesem Abschnitt stellen wir Beziehungen zwischen dem Lebesgue–Integral nach dem Lebesgue–Maß und dem Riemann–Integral bzw. dem uneigentlichen Riemann–Integral her.

Wir betrachten zunächst Funktionen auf einem abgeschlossenen Intervall $[a, b] \subseteq \mathbb{R}$ und bezeichnen mit $\mathcal{B}[a, b]$ die Spur–σ–Algebra von $\mathcal{B}(\mathbb{R})$ auf $[a, b]$.

Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt *Treppenfunktion*, wenn es ein $n \in \mathbb{N}$ und eine streng monoton wachsende endliche Folge $\{x_i\}_{i\in\{0,1,\dots,n\}} \subseteq \mathbb{R}$ mit $a = x_0$ und $x_n = b$ sowie eine Familie $\{c_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}$ gibt mit

$$f(x) = c_i$$

für alle $i \in \{1,\dots,n\}$ und $x \in (x_{i-1}, x_i)$; in diesem Fall ist die reelle Zahl

$$\int_a^b f(x)\,dx := \sum_{i=1}^n c_i(x_i - x_{i-1})$$

unabhängig von der Darstellung von f und heißt das *Riemann–Integral* von f.

9.8.1 Lemma. *Sei $f : [a,b] \to \mathbb{R}$ eine Treppenfunktion. Dann ist f $\boldsymbol{\lambda}$–integrierbar und es gilt*

$$\int_a^b f(x)\,dx = \int_{[a,b]} f(x)\,d\boldsymbol{\lambda}(x)$$

Beweis. Da f eine Treppenfunktion ist, gibt es eine streng monoton wachsende endliche Folge $\{x_i\}_{i\in\{0,1,\dots,n\}} \subseteq \mathbb{R}$ mit $a = x_0$ und $x_n = b$ sowie eine Familie $\{c_i\}_{i\in\{1,\dots,n\}} \subseteq \mathbb{R}$ mit

$$f(x) = c_i$$

für alle $i \in \{1,\dots,n\}$ und $x \in (x_{i-1}, x_i)$. Da jedes offene Intervall und jede einelementige Teilmenge des Intervalls $[a,b]$ messbar ist, ist f $\mathcal{B}[a,b]$–messbar. Es gilt

$$f = \sum_{i=1}^n c_i\,\chi_{(x_{i-1},x_i)}$$

$\boldsymbol{\lambda}$–fast überall und damit

$$|f| = \sum_{i=1}^n |c_i|\,\chi_{(x_{i-1},x_i)}$$

$\boldsymbol{\lambda}$–fast überall. Wegen

$$\begin{aligned}
\int_{[a,b]} |f|(x)\,d\boldsymbol{\lambda}(x) &= \int_{[a,b]} \left(\sum_{i=1}^n |c_i|\,\chi_{(x_{i-1},x_i)}(x)\right) d\boldsymbol{\lambda}(x) \\
&= \sum_{i=1}^n |c_i|\,\boldsymbol{\lambda}[(x_{i-1},x_i)] \\
&= \sum_{i=1}^n |c_i|\,(x_i - x_{i-1})
\end{aligned}$$

ist $|f|$ und damit auch f $\boldsymbol{\lambda}$–integrierbar und es gilt

$$\begin{aligned}\int_{[a,b]} f(x)\,d\boldsymbol{\lambda}(x) &= \int_{[a,b]} \left(\sum_{i=1}^n c_i\,\chi_{(x_{i-1},x_i)}(x)\right) d\boldsymbol{\lambda}(x)\\ &= \sum_{i=1}^n c_i\,\boldsymbol{\lambda}[(x_{i-1},x_i)]\\ &= \sum_{i=1}^n c_i\,(x_i - x_{i-1})\\ &= \int_a^b f(x)\,dx\end{aligned}$$

Damit ist das Lemma bewiesen. □

Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt *Riemann–integrierbar*, wenn sie beschränkt ist und die reellen Zahlen

$$U\!-\!\int_a^b f(x)\,dx := \sup\left\{\int_a^b g(x)\,dx \;\middle|\; g \text{ ist Treppenfunktion mit } g \le f\right\}$$

und

$$O\!-\!\int_a^b f(x)\,dx := \inf\left\{\int_a^b h(x)\,dx \;\middle|\; h \text{ ist Treppenfunktion mit } f \le h\right\}$$

übereinstimmen; in diesem Fall heißt die reelle Zahl

$$\int_a^b f(x)\,dx := U\!-\!\int_a^b f(x)\,dx = O\!-\!\int_a^b f(x)\,dx$$

das *Riemann–Integral* von f. Jede Treppenfunktion ist Riemann–integrierbar und die Definition des Riemann–Integrals für Riemann–integrierbare Funktionen ist mit der für Treppenfunktionen verträglich.

9.8.2 Satz. *Sei $f : [a,b] \to \mathbb{R}$ Riemann–integrierbar. Dann gibt es eine $\boldsymbol{\lambda}$–integrierbare Funktion $g : [a,b] \to \mathbb{R}$ mit $f = g$ $\boldsymbol{\lambda}$–fast überall und*

$$\int_a^b f(x)\,dx = \int_{[a,b]} g(x)\,d\boldsymbol{\lambda}(x)$$

Beweis. Da f Riemann–integrierbar ist, gibt es eine monoton wachsende Folge $\{g_n\}_{n\in\mathbb{N}}$ von Treppenfunktionen mit $\sup_{n\in\mathbb{N}} g_n \le f$ und

$$\sup_{n\in\mathbb{N}} \int_a^b g_n(x)\,dx = \int_a^b f(x)\,dx$$

sowie eine monoton fallende Folge $\{h_n\}_{n\in\mathbb{N}}$ von Treppenfunktionen mit $f \leq \inf_{n\in\mathbb{N}} h_n$ und

$$\int_a^b f(x)\,dx = \inf_{n\in\mathbb{N}} \int_a^b h_n(x)\,dx$$

Aufgrund der Messbarkeit von Treppenfunktionen sind auch die Funktionen

$$\begin{aligned} g &:= \sup_{n\in\mathbb{N}} g_n \\ h &:= \inf_{n\in\mathbb{N}} h_n \end{aligned}$$

messbar, und aus der Monotonie der Folgen $\{g_n\}_{n\in\mathbb{N}}$ und $\{h_n\}_{n\in\mathbb{N}}$ ergibt sich

$$g = \lim_{n\to\infty} g_n \leq f \leq \lim_{n\to\infty} h_n = h$$

und

$$\lim_{n\to\infty} \int_a^b g_n(x)\,dx = \int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b h_n(x)\,dx$$

Wir zeigen nun, dass $h-g=0$ $\boldsymbol{\lambda}$–fast überall gilt.
Wegen $g \leq f \leq h$ gilt

$$0 \leq h-g = \lim_{n\to\infty} h_n - \lim_{n\to\infty} g_n = \lim_{n\to\infty} (h_n - g_n)$$

und für alle $n \in \mathbb{N}$ ist $h_n - g_n$ eine Treppenfunktion mit

$$|h_n - g_n| = h_n - g_n \leq h_1 - g_1$$

Aus dem Satz über die majorisierte Konvergenz und Lemma 9.8.1 folgt nun

$$\begin{aligned} 0 &\leq \int_{[a,b]} (h-g)(x)\,d\boldsymbol{\lambda}(x) \\ &= \int_{[a,b]} \lim_{n\to\infty} (h_n - g_n)(x)\,d\boldsymbol{\lambda}(x) \\ &= \lim_{n\to\infty} \int_{[a,b]} (h_n - g_n)(x)\,d\boldsymbol{\lambda}(x) \\ &= \lim_{n\to\infty} \int_a^b (h_n - g_n)(x)\,dx \\ &= \lim_{n\to\infty} \int_a^b h_n(x)\,dx - \lim_{n\to\infty} \int_a^b g_n(x)\,dx \\ &= \int_a^b f(x)\,dx - \int_a^b f(x)\,dx \\ &= 0 \end{aligned}$$

Daher gilt $h-g=0$ $\boldsymbol{\lambda}$–fast überall, und wegen $g \leq f \leq h$ folgt daraus $f=g$ $\boldsymbol{\lambda}$–fast überall.

Für alle $n \in \mathbb{N}$ gilt $g_1 \leq g_n \leq g \leq h \leq h_1$, und daraus folgt $\sup_{n\in\mathbb{N}} |g_n| \leq (-g_1) \vee h_1$ und $|g| \leq (-g_1) \vee h_1$. Nach Lemma 9.8.1 ist $(-g_1) \vee h_1$ und damit auch $\sup_{n\in\mathbb{N}} |g_n|$ und g $\boldsymbol{\lambda}$–integrierbar. Aus dem Satz über die majorisierte Konvergenz folgt nun mit Lemma 9.8.1

$$\begin{aligned}\int_a^b f(x)\,dx &= \lim_{n\to\infty} \int_a^b g_n(x)\,dx \\ &= \lim_{n\to\infty} \int_{[a,b]} g_n(x)\,d\boldsymbol{\lambda}(x) \\ &= \int_{[a,b]} \lim_{n\to\infty} g_n(x)\,d\boldsymbol{\lambda}(x) \\ &= \int_{[a,b]} g(x)\,d\boldsymbol{\lambda}(x)\end{aligned}$$

Damit ist der Satz bewiesen. □

Für messbare Riemann–integrierbare Funktionen erhält man eine stärkere Aussage:

9.8.3 Folgerung. *Sei $f : [a,b] \to \mathbb{R}$ Riemann–integrierbar und messbar. Dann ist f $\boldsymbol{\lambda}$–integrierbar und es gilt*

$$\int_a^b f(x)\,dx = \int_{[a,b]} f(x)\,d\boldsymbol{\lambda}(x)$$

Beweis. Nach Satz 9.8.2 gibt es eine $\boldsymbol{\lambda}$–integrierbare Funktion $g : [a,b] \to \mathbb{R}$ mit $f = g$ $\boldsymbol{\lambda}$–fast überall und

$$\int_a^b f(x)\,dx = \int_{[a,b]} g(x)\,d\boldsymbol{\lambda}(x)$$

Da f messbar ist, ist nach Lemma 8.3.6 auch f $\boldsymbol{\lambda}$–integrierbar und es gilt

$$\int_{[a,b]} f(x)\,d\boldsymbol{\lambda}(x) = \int_{[a,b]} g(x)\,d\boldsymbol{\lambda}(x)$$

Daraus folgt die Behauptung. □

Das folgende Beispiel zeigt, dass eine beschränkte $\boldsymbol{\lambda}$–integrierbare Funktion nicht Riemann–integrierbar zu sein braucht:

9.8.4 Beispiel (Dirichlet–Funktion). Die Dirichlet–Funktion $f : [a,b] \to \mathbb{R}$ mit

$$f(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \cap [a,b] \\ 0 & \text{sonst} \end{cases}$$

ist $\boldsymbol{\lambda}$–integrierbar und beschränkt, aber sie ist nicht Riemann–integrierbar.

Wir betrachten nun Funktionen auf einem offenen Intervall $(a,b) \subseteq \bar{\mathbb{R}}$.

Eine Funktion $f : (a, b) \to \mathbb{R}$ heißt *uneigentlich Riemann–integrierbar*, wenn für jedes Intervall $[c, d] \subseteq (a, b)$ die Funktion $f|_{[c,d]}$ Riemann–integrierbar ist und für ein $t \in (a, b)$ die Grenzwerte

$$\lim_{s \downarrow a} \int_s^t f(x)\, dx \qquad \text{und} \qquad \lim_{u \uparrow b} \int_t^u f(x)\, dx$$

in $\mathbb{R}$ existieren. In diesem Fall ist die Summe dieser Grenzwerte unabhängig von der Wahl von t und die reelle Zahl

$$\int_a^b f(x)\, dx := \lim_{s \downarrow a} \int_s^t f(x)\, dx + \lim_{u \uparrow b} \int_t^u f(x)\, dx$$

heißt das *uneigentliche Riemann–Integral* von f.

Für positive messbare Funktionen erhalten wir das folgende Ergebnis:

9.8.5 Satz. *Sei $f : (a, b) \to \mathbb{R}$ eine positive messbare Funktion derart, dass für jedes Intervall $[c, d] \subseteq (a, b)$ die Funktion $f|_{[c,d]}$ Riemann–integrierbar ist. Dann sind folgende Aussagen äquivalent:*
(a) *f ist uneigentlich Riemann–integrierbar.*
(b) *f ist $\boldsymbol{\lambda}$–integrierbar.*
In diesem Fall gilt

$$\int_a^b f(x)\, dx = \int_{(a,b)} f(x)\, d\boldsymbol{\lambda}(x)$$

Beweis. Wir betrachten ein $t \in (a, b)$ sowie eine monoton fallende Folge $\{a_n\}_{n \in \mathbb{N}} \subseteq (a, t]$ mit $\lim_{n \to \infty} a_n = a$ und eine monoton wachsende Folge $\{b_n\}_{n \in \mathbb{N}} \subseteq [t, b)$ mit $\lim_{n \to \infty} b_n = b$.
Aus dem Satz über die monotone Konvergenz ergibt sich mit Folgerung 9.8.3

$$\int_{(a,t]} f(x)\, d\boldsymbol{\lambda}(x) = \lim_{n \to \infty} \int_{[a_n,t]} f(x)\, d\boldsymbol{\lambda}(x) = \lim_{n \to \infty} \int_{a_n}^t f(x)\, dx$$

sowie

$$\int_{[t,b)} f(x)\, d\boldsymbol{\lambda}(x) = \lim_{n \to \infty} \int_{[t,b_n]} f(x)\, d\boldsymbol{\lambda}(x) = \lim_{n \to \infty} \int_t^{b_n} f(x)\, dx$$

Wegen $\boldsymbol{\lambda}[\{t\}] = 0$ ergibt sich aus diesen Gleichungen

$$\begin{aligned} \int_{(a,b)} f(x)\, d\boldsymbol{\lambda}(x) &= \int_{(a,t]} f(x)\, d\boldsymbol{\lambda}(x) + \int_{[t,b)} f(x)\, d\boldsymbol{\lambda}(x) \\ &= \lim_{n \to \infty} \int_{a_n}^t f(x)\, dx + \lim_{n \to \infty} \int_t^{b_n} f(x)\, dx \\ &= \int_a^b f(x)\, dx \end{aligned}$$

Daraus folgt die Behauptung. $\square$

Die folgenden Beispiele zeigen, dass die in der Definition der Gamma–Funktion und der Beta–Funktion auftretenden uneigentlichen Riemann–Integrale als Lebesgue–Integrale nach dem Lebesgue–Maß dargestellt werden können:

9.8.6 Beispiele.
(1) **Gamma–Funktion:** Für alle $\gamma \in (0, \infty)$ gilt

$$\Gamma(\gamma) = \int_0^\infty e^{-z} z^{\gamma-1}\, dz = \int_{(0,\infty)} e^{-z} z^{\gamma-1}\, d\boldsymbol{\lambda}(z)$$

(2) **Beta–Funktion:** Für alle $\alpha, \beta \in (0, \infty)$ gilt

$$\mathrm{B}(\alpha, \beta) = \int_0^1 z^{\alpha-1}(1-z)^{\beta-1}\, dz = \int_{(0,1)} z^{\alpha-1}(1-z)^{\beta-1}\, d\boldsymbol{\lambda}(z)$$

Das folgende Beispiel zeigt, dass im letzten Satz die Forderung der Positivität der Funktion wesentlich ist:

9.8.7 Beispiel. Die Funktion $f : (0, \infty) \to \mathbb{R}$ mit

$$f(x) := \frac{\sin(x)}{x}$$

ist bekanntlich uneigentlich Riemann–integrierbar mit

$$\int_0^\infty f(x)\, dx = \frac{\pi}{2}$$

Wegen

$$\begin{aligned}
\int_{(0,\infty)} |f|(x)\, d\boldsymbol{\lambda}(x) &= \int_{(0,\infty)} \frac{|\sin(x)|}{x}\, d\boldsymbol{\lambda}(x) \\
&\geq \sum_{n=1}^\infty \int_{[(n-1)\pi, n\pi]} \frac{|\sin(x)|}{n\pi}\, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{\pi} \sum_{n=1}^\infty \frac{1}{n} \int_{[(n-1)\pi, n\pi]} |\sin(x)|\, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{\pi} \sum_{n=1}^\infty \frac{1}{n} \int_{(n-1)\pi}^{n\pi} |\sin(x)|\, dx \\
&= \frac{2}{\pi} \sum_{n=1}^\infty \frac{1}{n}
\end{aligned}$$

ist f nicht $\boldsymbol{\lambda}$–integrierbar.

Wir geben einige Beispiele, die in der Wahrscheinlichkeitstheorie von Interesse sind:

9.8.8 Beispiele.
(1) Für alle $a \in (0, \infty)$ gilt

$$\int_{(0,\infty)} x\, e^{-ax^2}\, d\boldsymbol{\lambda}(x) = \int_0^\infty x\, e^{-ax^2}\, dx = \frac{1}{2\,a}$$

(2) Es gilt

$$\int_{(0,\infty)} \frac{1}{1+x^2}\, d\boldsymbol{\lambda}(x) = \int_0^\infty \frac{1}{1+x^2}\, dx = \frac{\pi}{2}$$

und

$$\int_{\mathbb{R}} \frac{1}{1+x^2}\, d\boldsymbol{\lambda}(x) = \int_{-\infty}^\infty \frac{1}{1+x^2}\, dx = \pi$$

(3) Es gilt

$$\int_{(0,\infty)} e^{-x^2}\, d\boldsymbol{\lambda}(x) = \frac{1}{2}\sqrt{\pi}$$

und

$$\int_{\mathbb{R}} e^{-x^2} d\boldsymbol{\lambda}(x) = \sqrt{\pi}$$

In der Tat: Die Funktion $h : \mathbb{R}^2 \to \mathbb{R}$ mit

$$h(x,y) := y\, e^{-(1+x^2)y^2}\, \chi_{(0,\infty)^2}(x,y)$$

ist messbar und positiv. Die erste Gleichung ergibt sich mit Hilfe des Satzes von Fubini durch die beiden iterativen Auswertungen des Integrals

$$\int_{\mathbb{R}^2} h(x,y)\, d\boldsymbol{\lambda}^2(x,y)$$

Aus dem Satz von Fubini ergibt sich unter Verwendung von (1) und (2)

$$\begin{aligned}
\int_{\mathbb{R}^2} h(x,y)\, d\boldsymbol{\lambda}^2(x,y) &= \int_{\mathbb{R}} \int_{\mathbb{R}} h(x,y)\, d\boldsymbol{\lambda}(y)\, d\boldsymbol{\lambda}(x) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} y\, e^{-(1+x^2)y^2}\, \chi_{(0,\infty)^2}(x,y)\, d\boldsymbol{\lambda}(y)\, d\boldsymbol{\lambda}(x) \\
&= \int_{(0,\infty)} \int_{(0,\infty)} y\, e^{-(1+x^2)y^2} d\boldsymbol{\lambda}(y)\, d\boldsymbol{\lambda}(x) \\
&= \int_{(0,\infty)} \frac{1}{2}\, \frac{1}{1+x^2}\, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{2} \cdot \frac{\pi}{2} \\
&= \frac{\pi}{4}
\end{aligned}$$

Wir betrachten nun für $y \in (0,\infty)$ die lineare Abbildung $T_y : \mathbb{R} \to \mathbb{R}$ mit $T_y(x) := xy$. Dann ist T_y bijektiv mit $T_y^{-1}(0,\infty) = (0,\infty)$ und $|\det(T_y)| = y$, und aus Satz 6.3.1 ergibt sich $\boldsymbol{\lambda}_{T_y} = y^{-1}\boldsymbol{\lambda}$. Aus dem Satz von Fubini und der Substitutionsregel ergibt sich

$$\int_{\mathbb{R}^2} h(x,y)\, d\boldsymbol{\lambda}^2(x,y) = \int_{\mathbb{R}} \int_{\mathbb{R}} h(x,y)\, d\boldsymbol{\lambda}(x)\, d\boldsymbol{\lambda}(y)$$

$$
\begin{aligned}
&= \int_{\mathbb{R}} \int_{\mathbb{R}} y\, e^{-(1+x^2)y^2} \chi_{(0,\infty)^2}(x,y)\, d\boldsymbol{\lambda}(x)\, d\boldsymbol{\lambda}(y) \\
&= \int_{(0,\infty)} \int_{(0,\infty)} y\, e^{-x^2y^2} d\boldsymbol{\lambda}(x)\, e^{-y^2} d\boldsymbol{\lambda}(y) \\
&= \int_{(0,\infty)} \int_{(T_y)^{-1}(0,\infty)} y\, e^{-(T_y(x))^2} d\boldsymbol{\lambda}(x)\, e^{-y^2} d\boldsymbol{\lambda}(y) \\
&= \int_{(0,\infty)} \int_{(0,\infty)} y\, e^{-z^2} d\boldsymbol{\lambda}_{T_y}(z)\, e^{-y^2} d\boldsymbol{\lambda}(y) \\
&= \int_{(0,\infty)} \int_{(0,\infty)} y\, e^{-z^2}\, y^{-1} d\boldsymbol{\lambda}(z)\, e^{-y^2} d\boldsymbol{\lambda}(y) \\
&= \int_{(0,\infty)} \int_{(0,\infty)} e^{-z^2} d\boldsymbol{\lambda}(z)\, e^{-y^2} d\boldsymbol{\lambda}(y) \\
&= \int_{(0,\infty)} e^{-z^2}\, d\boldsymbol{\lambda}(z) \int_{(0,\infty)} e^{-y^2} d\boldsymbol{\lambda}(y) \\
&= \left(\int_{(0,\infty)} e^{-z^2} d\boldsymbol{\lambda}(z) \right)^2
\end{aligned}
$$

Daher gilt

$$\left(\int_{(0,\infty)} e^{-z^2} d\boldsymbol{\lambda}(z) \right)^2 = \int_{\mathbb{R}^2} h(x,y)\, d\boldsymbol{\lambda}^2(x,y) = \frac{\pi}{4}$$

und damit

$$\int_{(0,\infty)} e^{-x^2}\, d\boldsymbol{\lambda}(x) = \frac{1}{2}\sqrt{\pi}$$

Damit ist die erste Gleichung gezeigt.
Des weiteren ist die Abbildung $S : \mathbb{R} \to \mathbb{R}$ mit $S(x) := -x$ bijektiv mit $S^{-1}(0,\infty) = (-\infty, 0)$ und $|\det(S)| = 1$, und aus Satz 6.3.1 ergibt sich $\boldsymbol{\lambda}_S = \boldsymbol{\lambda}$. Daher gilt

$$
\begin{aligned}
\int_{(-\infty,0)} e^{-x^2} d\boldsymbol{\lambda}(x) &= \int_{S^{-1}(0,\infty)} e^{-(S(x))^2} d\boldsymbol{\lambda}(x) \\
&= \int_{(0,\infty)} e^{-z^2} d\boldsymbol{\lambda}_S(z) \\
&= \int_{(0,\infty)} e^{-z^2} d\boldsymbol{\lambda}(z) \\
&= \frac{1}{2}\sqrt{\pi}
\end{aligned}
$$

und wegen $\boldsymbol{\lambda}[\{0\}] = 0$ ergibt sich nun

$$\int_{\mathbb{R}} e^{-x^2} d\boldsymbol{\lambda}(x) = \int_{(-\infty,0)} e^{-x^2} d\boldsymbol{\lambda}(x) + \int_{(0,\infty)} e^{-x^2} d\boldsymbol{\lambda}(x) = \frac{1}{2}\sqrt{\pi} + \frac{1}{2}\sqrt{\pi} = \sqrt{\pi}$$

Damit ist auch die zweite Gleichung gezeigt.

Aus dem Zusammenhang zwischen dem Lebesgue–Integral und dem Riemann–Integral erhalten wir auch eine weitere Eigenschaft von Transformationen des Lebesgue–Maßes:

9.8.9 Lemma. *Sei $C \subseteq \mathbb{R}$ ein nichtleeres offenes Intervall und sei $T : C \to \mathbb{R}$ stetig differenzierbar mit $T'(x) \neq 0$ für alle $x \in C$. Dann gilt*

$$(\boldsymbol{\lambda}|_{\mathcal{B}(C)})_T|_{\mathcal{B}(T(C))} = \int \left| \frac{dT^{-1}}{dz}(z) \right| d\boldsymbol{\lambda}|_{\mathcal{B}(T(C))}(z)$$

Beweis. Nach Voraussetzung ist T stetig und entweder streng monoton wachsend oder streng monoton fallend. Daher ist $T(C)$ ein offenes Intervall und T besitzt eine Umkehrfunktion $T^{-1} : T(C) \to C$, die stetig differenzierbar ist und dieselben Monotonieeigenschaften besitzt wie T.
Wir betrachten nun die Messräume $(C, \mathcal{B}(C))$ und $(T(C), \mathcal{B}(T(C)))$ und die Restriktionen $\boldsymbol{\lambda}|_{\mathcal{B}(C)}$ und $\boldsymbol{\lambda}|_{\mathcal{B}(T(C))}$ des Lebesgue–Maßes $\boldsymbol{\lambda}$. Da die Abbildung $S := T^{-1}$ stetig differenzierbar ist, ist ihre Ableitung $S' : T(C) \to C$ messbar. Wir betrachten das Maß $\nu : \mathcal{B}(T(C)) \to [0, \infty]$ mit

$$\nu := \int |S'(z)| \, d\boldsymbol{\lambda}|_{\mathcal{B}(T(C))}(z)$$

und berechnen das Bildmaß $\nu_S : \mathcal{B}(C) \to [0, \infty]$:

– Ist T streng monoton wachsend, so ist auch S streng monoton wachsend und für alle $[a, b] \subseteq C$ gilt

$$\begin{aligned}
\nu_S[[a,b]] &= \nu[S^{-1}([a,b])] \\
&= \nu[[S^{-1}(a), S^{-1}(b)]] \\
&= \int_{[S^{-1}(a),S^{-1}(b)]} |S'(z)| \, d\boldsymbol{\lambda}|_{\mathcal{B}(T(C))}(z) \\
&= \int_{[S^{-1}(a),S^{-1}(b)]} S'(z) \, d\boldsymbol{\lambda}|_{\mathcal{B}(T(C))}(z) \\
&= \int_{S^{-1}(a)}^{S^{-1}(b)} S'(z) \, dz \\
&= S(S^{-1}(b)) - S(S^{-1}(a)) \\
&= b - a \\
&= \boldsymbol{\lambda}|_{\mathcal{B}(C)}[[a,b]]
\end{aligned}$$

– Ist T streng monoton fallend, so ist auch S streng monoton fallend und man erhält wiederum für alle $[a, b] \subseteq C$

$$\nu_S[[a,b]] = \boldsymbol{\lambda}|_{\mathcal{B}(C)}[[a,b]]$$

Daher gilt in jedem Fall für alle $[a,b] \subseteq C$

$$\nu_S[[a,b]] = \boldsymbol{\lambda}|_{\mathcal{B}(C)}[[a,b]]$$

Da das Mengensystem $\mathcal{E} := \{[a,b] \mid [a,b] \subseteq C\}$ ein $\cap$–stabiler Erzeuger von $\mathcal{B}(C)$ ist und $(\boldsymbol{\lambda}|_{\mathcal{B}(C)})|_{\mathcal{E}}$ σ–endlich ist, folgt aus dem Eindeutigkeitssatz

$$\nu_S = \boldsymbol{\lambda}|_{\mathcal{B}(C)}$$

und damit

$$(\boldsymbol{\lambda}|_{\mathcal{B}(C)})_T|_{\mathcal{B}(T(C))} = (\nu_S)_T|_{\mathcal{B}(T(C))} = \nu_{T \circ S}|_{\mathcal{B}(T(C))} = \nu$$

Die Behauptung des Lemmas folgt nun aus der Definition von ν. $\square$

Teil IV

Wahrscheinlichkeitstheorie

10

Wahrscheinlichkeitsräume

Ist $(\Omega, \mathcal{F})$ ein Messraum und ist $P : \mathcal{F} \to [0,1]$ ein Maß mit $P[\Omega] = 1$, so heißt das Maß P Wahrscheinlichkeitsmaß und der Maßraum $(\Omega, \mathcal{F}, P)$ heißt Wahrscheinlichkeitsraum. Diese Definitionen gehen auf Kolmogorov [1933] zurück, der so die Wahrscheinlichkeitstheorie in die Maß– und Integrationstheorie eingebettet und die Grundlage für die moderne Wahrscheinlichkeitstheorie gelegt hat.

Der Wahrscheinlichkeitstheorie stehen damit alle Begriffe und Ergebnisse der Maß– und Integrationstheorie zur Verfügung; darüber hinaus gewinnt sie aus der Konzentration auf Wahrscheinlichkeitsmaße eine eigene Kraft, die neue Begriffe hervorbringt, die in der allgemeinen Maß– und Integrationstheorie ohne Bedeutung sind und die zu einer reichen Theorie führen, die unzählige Anwendungen besitzt und die wir in diesem und den folgenden Kapiteln nur in ihren Grundzügen darstellen können.

In diesem Kapitel betrachten wir zunächst die elementaren Begriffe der Wahrscheinlichkeitstheorie (Abschnitt 10.1) und behandeln sodann die wesentlichen Prinzipien zur Konstruktion von Wahrscheinlichkeitsräumen. Die einfachsten Wahrscheinlichkeitsräume sind die diskreten (Abschnitt 10.2) oder sogar symmetrischen Wahrscheinlichkeitsräume (Abschnitt 10.3), deren Konstruktion durch elementare Überlegungen erfolgen kann. Theoretisch anspruchsvoller ist hingegen die Konstruktion des Produktes einer Familie von Wahrscheinlichkeitsräumen, die sich im Fall einer endlichen Familie aus der Maß– und Integrationstheorie ergibt (Abschnitt 10.4), im Fall einer unendlichen Familie aber spezifisch für die Wahrscheinlichkeitstheorie ist (Abschnitt 10.6) und auf dem Begriff einer projektiven Familie von Wahrscheinlichkeitsmaßen beruht (Abschnitt 10.5).

Ein weiteres Prinzip zur Konstruktion von Wahrscheinlichkeitsräumen, das in der Theorie der stochastischen Prozesse von Bedeutung ist und ebenfalls auf dem Begriff einer projektiven Familie von Wahrscheinlichkeitsmaßen beruht, betrachten wir erst in Kapitel 20.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_10,

10.1 Wahrscheinlichkeitsräume und Zufallsgrößen

Sei Ω eine nichtleere Menge und sei $\mathcal{C} \subseteq 2^\Omega$ ein Mengensystem mit $\emptyset, \Omega \in \mathcal{C}$.

Eine Mengenfunktion $P : \mathcal{C} \to [0,1]$ heißt
– *Wahrscheinlichkeitsinhalt*, wenn P ein Inhalt mit $P[\Omega] = 1$ ist.
– *Wahrscheinlichkeitsmaß*, wenn P ein Maß mit $P[\Omega] = 1$ ist.
Jedes Wahrscheinlichkeitsmaß ist ein Wahrscheinlichkeitsinhalt und jeder Wahrscheinlichkeitsinhalt ist endlich. Das folgende Ergebnis ist elementar, wird aber häufig verwendet:

10.1.1 Lemma. *Sei $\mathcal{C}$ eine Algebra und $P : \mathcal{C} \to [0,1]$ ein Wahrscheinlichkeitsinhalt. Dann gilt für alle $A \in \mathcal{C}$*

$$P[\Omega \setminus A] = 1 - P[A]$$

Da jeder Wahrscheinlichkeitsinhalt endlich ist, lässt sich die Charakterisierung der σ–Additivität gegenüber beliebigen Inhalten geringfügig vereinfachen. Das folgende Lemma ergibt sich unmittelbar aus Folgerung 4.2.10:

10.1.2 Lemma. *Sei $\mathcal{C}$ eine Algebra und $P : \mathcal{C} \to [0,1]$ ein Wahrscheinlichkeitsinhalt. Dann sind äquivalent:*
(a) *P ist ein Wahrscheinlichkeitsmaß.*
(b) *P ist σ–additiv.*
(c) *P ist stetig von unten.*
(d) *P ist stetig von oben.*
(e) *P ist $\emptyset$–stetig.*

Ist $(\Omega, \mathcal{F})$ ein Messraum und ist $P : \mathcal{F} \to [0,1]$ ein Wahrscheinlichkeitsmaß, so heißt $(\Omega, \mathcal{F}, P)$ *Wahrscheinlichkeitsraum*. Jeder Wahrscheinlichkeitsraum ist also ein endlicher Maßraum.

Ist $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $(\Omega', \mathcal{F}')$ ein Messraum, so heißt jede messbare Abbildung $X : \Omega \to \Omega'$ *Zufallsgröße* und das Bildmaß P_X von P unter X heißt *Verteilung* von X; insbesondere heißt X
– *Zufallsvariable*, wenn $(\Omega', \mathcal{F}') = (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$ gilt.
– *reelle Zufallsvariable*, wenn $(\Omega', \mathcal{F}') = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ gilt.
– *Zufallsvektor*, wenn $(\Omega', \mathcal{F}') = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ gilt.
Das folgende Lemma ergibt sich unmittelbar aus der Definition des Bildmaßes:

10.1.3 Lemma. *Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $(\Omega', \mathcal{F}')$ ein Messraum und sei $X : \Omega \to \Omega'$ eine Zufallsgröße. Dann ist P_X ein Wahrscheinlichkeitsmaß.*

Ist $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $(\Omega', \mathcal{F}')$ ein Messraum und ist $X : \Omega \to \Omega'$ eine Zufallsgröße, so ist insbesondere $(\Omega', \mathcal{F}', P_X)$ ein Wahrscheinlichkeitsraum.

Die vorher gegebenen Definitionen zeigen bereits, dass für bestimmte mathematische Objekte in der Wahrscheinlichkeitstheorie andere Begriffe verwendet werden als in der allgemeinen Maß– und Integrationstheorie. Wir geben einige weitere Definitionen, die für die Wahrscheinlichkeitstheorie typisch sind:

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum. Dann wird
- jedes Element $\omega \in \Omega$ als *Ergebnis*,
- die Grundmenge Ω als *Ergebnismenge*,
- jede Menge $A \in \mathcal{F}$ als *Ereignis*,
- jede Menge $A \in \mathcal{F}$ mit $A = \{\omega\}$ für ein $\omega \in \Omega$ als *Elementarereignis*, und
- jedes Mengensystem $\mathcal{E} \subseteq \mathcal{F}$ als *Ereignissystem*

bezeichnet. Neben dem Unterschied zwischen Ergebnissen und Elementarereignissen ist zu beachten, dass nicht jede Teilmenge der Ergebnismenge ein Ereignis ist und dass nicht jedes Mengensystem ein Ereignissystem ist. Für ein festes Ergebnis $\omega \in \Omega$ sagt man auch, dass ein Ereignis $A \in \mathcal{F}$ *eintritt*, wenn $\omega \in A$ gilt; auf diese Sprechweise ist es zurückzuführen, dass
- die Grundmenge Ω als *sicheres Ereignis*, und
- die leere Menge $\emptyset$ als *unmögliches Ereignis*

bezeichnet wird.

Diesen Bezeichnungen liegt die Vorstellung zugrunde, dass die Ergebnismenge die möglichen Ergebnisse eines *Zufallsexperimentes* beschreibt und dass die Ereignisse diejenigen Teilmengen der Ergebnismenge sind, deren Wahrscheinlichkeit gemessen werden soll.

10.1.4 Beispiel (Wurf eine Würfels). Als Zufallsexperiment betrachten wir den Wurf eines Würfels. Sieht man davon ab, dass der Würfel möglicherweise vom Tisch fällt oder gegen einen Gegenstand rollt und auf einer seiner Kanten stehenbleibt, so ist die Menge

$$\Omega := \{1,2,3,4,5,6\}$$

eine geeignete Wahl der Ergebnismenge. Interessiert man sich nur dafür, ob beim Wurf des Würfels eine gerade oder eine ungerade Augenzahl auftritt, so genügt es, die σ–Algebra

$$\mathcal{F} := \Big\{\emptyset, \{1,3,5\}, \{2,4,6\}, \Omega\Big\}$$

zu betrachten. Ist schließlich der Würfel unverfälscht, so liegt es nahe, mittels

$$P[A] := \begin{cases} 0 & \text{falls } A = \emptyset \\ 1/2 & \text{falls } A = \{1,3,5\} \text{ oder } A = \{2,4,6\} \\ 1 & \text{falls } A = \Omega \end{cases}$$

ein Wahrscheinlichkeitsmaß $P : \mathcal{F} \to [0,1]$ zu definieren.

In den folgenden Abschnitten betrachten wir weitere Zufallsexperimente und ihre Modellierung durch einen Wahrscheinlichkeitsraum; auch diese Zufallsexperimente sind wieder extrem einfach gewählt, weil sie nur als Beispiele für die Anwendung bestimmter Prinzipien zur Konstruktion von Wahrscheinlichkeitsräumen dienen sollen.

Aufgaben

10.1.A Führen Sie die fehlenden Beweise aus.

10.1.B Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum. Dann gilt für alle $A, B \in \mathcal{F}$

$$|P[A] - P[B]| \leq P[A \triangle B]$$

10.1.C Sei $(\Omega, \mathcal{F})$ ein Messraum und sei $\{P_n\}_{n\in\mathbb{N}}$ eine Folge von Wahrscheinlichkeitsmaßen $\mathcal{F} \to [0,1]$ und $\{a_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$ eine Folge mit $\sum_{n=1}^{\infty} a_n = 1$. Dann ist die Mengenfunktion $P : \mathcal{F} \to [0,\infty]$ mit

$$P[A] := \sum_{n=1}^{\infty} a_n \, P_n[A]$$

ein Wahrscheinlichkeitsmaß.

10.2 Diskrete Wahrscheinlichkeitsräume

Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ heißt *diskret*, wenn er die folgenden Eigenschaften besitzt:
(i) Ω ist abzählbar.
(ii) Es gilt $\mathcal{F} = 2^{\Omega}$.
Wir zeigen, dass diskrete Wahrscheinlichkeitsräume auf besonders einfache Weise konstruiert werden können.

Sei Ω eine abzählbare Menge. Dann heißt jede Funktion $p : \Omega \to [0,1]$ mit

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

Wahrscheinlichkeitsfunktion auf Ω. Jede Wahrscheinlichkeitsfunktion erzeugt ein eindeutig bestimmtes Wahrscheinlichkeitsmaß auf der Potenzmenge:

10.2.1 Lemma. *Sei Ω abzählbar und $p : \Omega \to [0,1]$ eine Wahrscheinlichkeitsfunktion. Dann gibt es genau ein Wahrscheinlichkeitsmaß $P : 2^{\Omega} \to [0,1]$ mit*

$$P[\{\omega\}] = p(\omega)$$

für alle $\omega \in \Omega$ und es gilt

$$P[A] = \sum_{\omega\in A} p(\omega)$$

für alle $A \in 2^{\Omega}$.

Beweis. Sei zunächst $P : 2^{\Omega} \to [0,1]$ gegeben durch

$$P[A] := \sum_{\omega\in A} p(\omega)$$

Dann gilt $P[\emptyset] = 0$ und $P[\Omega] = \sum_{\omega\in\Omega} p(\omega) = 1$, und für jede disjunkte Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ gilt

$$P\left[\sum_{n=1}^{\infty} A_n\right] = \sum_{\omega\in\sum_{n=1}^{\infty} A_n} p(\omega) = \sum_{n=1}^{\infty} \sum_{\omega\in A_n} p(\omega) = \sum_{n=1}^{\infty} P[A_n]$$

Daher ist P ein Wahrscheinlichkeitsmaß mit

$$P[\{\omega\}] = p(\omega)$$

für alle $\omega \in \Omega$.
Sei nun $Q : 2^\Omega \to [0,1]$ ein beliebiges Wahrscheinlichkeitsmaß mit

$$Q[\{\omega\}] = p(\omega)$$

für alle $\omega \in \Omega$. Dann gilt für alle $A \in \mathcal{F}$

$$Q[A] = \sum_{\omega\in A} Q[\{\omega\}] = \sum_{\omega\in A} p(\omega) = P[A]$$

Daher gilt $Q = P$. □

Von besonderem Interesse ist der Fall $\Omega = \mathbb{N}_0$: Eine Folge $\{p_n\}_{n\in\mathbb{N}_0} \subseteq \mathbb{R}_+$ mit

$$\sum_{n=0}^{\infty} p_n = 1$$

heißt *stochastische Folge*. Ist $\{p_n\}_{n\in\mathbb{N}_0}$ eine stochastische Folge, so ist die Funktion $p : \mathbb{N}_0 \to [0,1]$ mit

$$p(n) := p_n$$

eine Wahrscheinlichkeitsfunktion.

10.2.2 Folgerung. *Sei $\{p_n\}_{n\in\mathbb{N}_0}$ eine stochastische Folge. Dann gibt es genau ein Wahrscheinlichkeitsmaß $P : 2^{\mathbb{N}_0} \to [0,1]$ mit*

$$P[\{n\}] = p_n$$

für alle $n \in \mathbb{N}_0$.

Dieses Ergebnis ist bemerkenswert: Obwohl die Potenzmenge von $\mathbb{N}_0$ überabzählbar ist, ist jedes Wahrscheinlichkeitsmaß auf der Potenzmenge durch eine stochastische Folge und damit durch abzählbar viele Werte bestimmt.

Aufgaben

10.2.A Jede Abbildung von einem diskreten Wahrscheinlichkeitsraum in einen Messraum ist eine Zufallsgröße.

10.2.B Beweisen Sie Lemma 10.2.1 mit Hilfe des Satzes von Caratheodory.

10.3 Symmetrische Wahrscheinlichkeitsräume

Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ heißt *symmetrisch*, wenn er die folgenden Eigenschaften besitzt:
(i) Ω ist endlich.
(ii) Es gilt $\mathcal{F} = 2^\Omega$.
(iii) Es gibt ein $p \in [0, 1]$ mit $P[\{\omega\}] = p$ für alle $\omega \in \Omega$.
Jeder symmetrische Wahrscheinlichkeitsraum ist ein diskreter Wahrscheinlichkeitsraum.

In einem symmetrischen Wahrscheinlichkeitsraum wird jedem Elementarereignis dieselbe Wahrscheinlichkeit zugeordnet. Die Wahl eines symmetrischen Wahrscheinlichkeitsraumes als Modell für ein Zufallsexperiment kann mit dem *Prinzip des unzureichenden Grundes* begründet werden: Wenn ein Zufallsexperiment im wesentlichen nur endlich viele Ergebnisse hervorbringen kann und kein Grund dafür erkennbar ist, dass eines der Ergebnisse eine höhere Chance der Realisierung besitzt als die anderen Ergebnisse, dann ist der zugehörige symmetrische Wahrscheinlichkeitsraum ein geeignetes Modell. Symmetrische Wahrscheinlichkeitsräume werden auch als *Laplace–Experimente* bezeichnet, obwohl sie natürlich nur Modelle für Zufallsexperimente sind.

Im Prinzip ist die Berechnung der Wahrscheinlichkeiten der Ereignisse eines symmetrischen Wahrscheinlichkeitsraumes einfach:

10.3.1 Lemma. *Sei* $(\Omega, \mathcal{F}, P)$ *ein symmetrischer Wahrscheinlichkeitsraum. Dann gilt für alle* $A \in \mathcal{F}$

$$P[A] = \frac{|A|}{|\Omega|}$$

Beweis. Nach Voraussetzung gibt es ein $p \in [0, 1]$ mit

$$P[\{\omega\}] = p$$

für alle $\omega \in \Omega$. Wegen $P[\Omega] = 1$ folgt aus der Additivität von P zunächst

$$P[\{\omega\}] = 1/|\Omega|$$

für alle $\omega \in \Omega$ und sodann

$$P[A] = |A|/|\Omega|$$

für alle $A \in \mathcal{F}$. □

In einem symmetrischen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ ist daher die Wahrscheinlichkeit $P[A]$ eines Ereignisses $A \in \mathcal{F}$ gerade das Verhältnis $|A|/|\Omega|$ zwischen der Anzahl $|A|$ der Ergebnisse, für die A eintritt, und der Anzahl $|\Omega|$ aller Ergebnisse. Dabei bezeichnet man
– $|A|$ als die *Anzahl der* (für das Eintreten von A) *günstigen Fälle* und
– $|\Omega|$ als die *Anzahl der möglichen Fälle*.

Aufgrund des Lemmas ist es erstrebenswert, für ein Zufallsexperiment, das im wesentlichen nur endlich viele Ergebnisse hervorbringen kann, die Ergebnismenge so zu wählen, dass die Wahl des zugehörigen symmetrischen Wahrscheinlichkeitsraumes als Modell plausibel ist.

10.3.2 Beispiel (n–maliger Wurf einer Münze). Wir betrachten den n–maligen Wurf einer Münze. Wir nehmen an, dass
- zwischen den verschiedenen Würfen der Münze keine gegenseitige Beeinflussung besteht,
- bei jedem Wurf nur *Kopf* oder *Zahl* auftreten kann und
- die Chance für das Auftreten von *Kopf* beim einmaligen Wurf gleich 1/2 ist.

Wir interessieren uns für die Anzahl der Würfe, bei denen *Kopf* auftritt.
Als Ergebnismenge für dieses Zufallsexperiment bietet sich natürlich die Menge $\Omega := \{0, 1, \ldots, n\}$ an; da aber diese Ergebnismenge offenbar nicht dem Prinzip des unzureichenden Grundes genügt, ist die Wahl des zugehörigen symmetrischen Wahrscheinlichkeitsraumes als Modell nicht plausibel.
Aufgrund der Annahme, dass bei jedem einzelnen Wurf *Kopf* und *Zahl* dieselbe Chance haben und dass zwischen den verschiedenen Würfen keine gegenseitige Beeinflussung besteht, ist es jedoch plausibel, anzunehmen, dass die 2^n Folgen von *Kopf* und *Zahl*, die beim n–maligen Wurf auftreten können, ebenfalls alle dieselbe Chance haben. Wir setzen daher $\Omega := \{K, Z\}^n$ und $\mathcal{F} := 2^\Omega$, und für alle $A \in \mathcal{F}$ setzen wir

$$P[A] := \frac{|A|}{|\Omega|}$$

Dann ist $(\Omega, \mathcal{F}, P)$ ein symmetrischer Wahrscheinlichkeitsraum mit

$$|\Omega| = 2^n$$

Für $k \in \{0, 1, \ldots, n\}$ bezeichne $E_k \in \mathcal{F}$ das Ereignis, dass beim n–maligen Wurf der Münze genau k–mal *Kopf* auftritt. Dann erhält man

$$P[E_k] = \frac{|E_k|}{|\Omega|} = \binom{n}{k} \frac{1}{2^n}$$

Die Zufallsvariable $X : \Omega \to \mathbb{R}$ mit $X(\omega) := k$ für alle $k \in \{0, 1, \ldots, n\}$ und $\omega \in E_k$ gibt die Anzahl der Würfe an, bei denen *Kopf* auftritt, und für alle $k \in \{0, 1, \ldots, n\}$ gilt

$$P[\{X = k\}] = \binom{n}{k} \frac{1}{2^n}$$

Plausibilitätsprüfung: Es gilt $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

Das Beispiel zeigt:
- Wählt man die Ergebnismenge nur so groß wie nötig, so ist Symmetrie oft nicht plausibel und ein Wahrscheinlichkeitsmaß nur schwer zu bestimmen.
- Wünscht man Symmetrie, so muss man oft eine größere Ergebnismenge wählen mit der Konsequenz, dass Ereignisse sehr viele Ergebnisse enthalten können und ihre Mächtigkeit nicht leicht zu bestimmen ist.

Hier hilft die Kombinatorik:

10.3.3 Beispiele (Urnenmodelle). Wir betrachten eine Urne mit $N \geq 2$ Kugeln, von denen $K \in \{1, \dots, N{-}1\}$ Kugeln rot sind und $N{-}K$ Kugeln eine beliebige andere Farbe besitzen. Wir nehmen an, dass alle Kugeln bis auf die Farbe gleichartig sind. Wir interessieren uns für die Anzahl der roten Kugeln bei der zufälligen Auswahl von n Kugeln aus der Urne. Zu diesem Zweck nehmen wir zusätzlich an, die Kugeln seien numeriert derart, dass die Kugeln $1, \dots, K$ rot sind und die Kugeln $K+1, \dots, N$ eine andere Farbe besitzen.

(1) **Ziehen ohne Zurücklegen:** Sei $n \leq N$. Wir setzen

$$\Omega := \text{Menge aller } n\text{–Tupel ohne Wiederholung aus } \{1, ..., N\}$$

und $\mathcal{F} := 2^\Omega$. Aufgrund der Annahme der Gleichartigkeit der Kugeln setzen wir für alle $A \in \mathcal{F}$

$$P[A] := \frac{|A|}{|\Omega|}$$

Dann ist $(\Omega, \mathcal{F}, P)$ ein symmetrischer Wahrscheinlichkeitsraum mit

$$|\Omega| = n! \binom{N}{n}$$

Für $k \in \{0, 1, \dots, n\}$ betrachten wir das Ereignis

$$E_k := \{\omega \in \Omega \mid \text{genau } k \text{ Kugeln sind rot}\}$$

Dann ist die Anzahl der günstigen Fälle das Produkt

– der Anzahl der Möglichkeiten, aus den n Ziehungen k Ziehungen (mit einer roten Kugel als Ergebnis) auszuwählen,
– der Anzahl der Möglichkeiten, aus den K numerierten roten Kugeln k Kugeln ohne Zurücklegen auszuwählen, und
– der Anzahl der Möglichkeiten, aus den $N{-}K$ numerierten Kugeln mit einer anderen Farbe $n{-}k$ Kugeln ohne Zurücklegen auszuwählen.

Es gilt also

$$|E_k| = \binom{n}{k} \cdot k! \binom{K}{k} \cdot (n{-}k)! \binom{N{-}K}{n{-}k} = n! \binom{K}{k} \binom{N{-}K}{n{-}k}$$

und damit

$$P[E_k] = \frac{|E_k|}{|\Omega|} = \frac{n! \binom{K}{k} \binom{N{-}K}{n{-}k}}{n! \binom{N}{n}} = \frac{\binom{K}{k} \binom{N{-}K}{n{-}k}}{\binom{N}{n}}$$

Die Zufallsvariable $X : \Omega \to \mathbb{R}$ mit $X(\omega) := k$ für alle $k \in \{0, 1, \dots, n\}$ und $\omega \in E_k$ gibt die Anzahl der roten Kugeln bei der zufälligen Auswahl von n Kugeln an und für alle $k \in \{0, 1, \dots, n\}$ gilt

$$P[\{X = k\}] = \frac{\binom{K}{k} \binom{N{-}K}{n{-}k}}{\binom{N}{n}}$$

Plausibilitätsprüfung: Es gilt $\sum_{k=0}^n \binom{K}{k}\binom{N{-}K}{n{-}k} = \binom{N}{n}$.

(2) **Ziehen mit Zurücklegen:** Wir setzen

$$\Omega := \text{Menge aller } n\text{–Tupel mit möglicher Wiederholung aus } \{1, ..., N\}$$

und $\mathcal{F} := 2^\Omega$. Aufgrund der Annahme der Gleichartigkeit der Kugeln setzen wir für alle $A \in \mathcal{F}$

$$P[A] := \frac{|A|}{|\Omega|}$$

Dann ist $(\Omega, \mathcal{F}, P)$ ein symmetrischer Wahrscheinlichkeitsraum mit

$$|\Omega| = N^n$$

Für $k \in \{0, 1, \ldots, n\}$ betrachten wir das Ereignis

$$E_k := \{\omega \in \Omega \mid \text{genau } k \text{ Kugeln sind rot}\}$$

Dann ist die Anzahl der günstigen Fälle das Produkt
– der Anzahl der Möglichkeiten, aus den n Ziehungen k Ziehungen (mit einer roten Kugel als Ergebnis) auszuwählen,
– der Anzahl der Möglichkeiten, aus den K numerierten roten Kugeln k Kugeln mit Zurücklegen auszuwählen, und
– der Anzahl der Möglichkeiten, aus den $N-K$ numerierten Kugeln mit einer anderen Farbe $n-k$ Kugeln mit Zurücklegen auszuwählen.

Es gilt also

$$|E_k| = \binom{n}{k} \cdot K^k \cdot (N-K)^{n-k}$$

und damit

$$P[E_k] = \frac{|E_k|}{|\Omega|} = \frac{\binom{n}{k} K^k \, (N-K)^{n-k}}{N^n} = \binom{n}{k} \left(\frac{K}{N}\right)^k \left(\frac{N-K}{N}\right)^{n-k}$$

Mit $\vartheta := K/N$ gilt daher

$$P[E_k] = \binom{n}{k} \vartheta^k \, (1-\vartheta)^{n-k}$$

Die Zufallsvariable $X : \Omega \to \mathbb{R}$ mit $X(\omega) := k$ für alle $k \in \{0, 1, \ldots, n\}$ und $\omega \in E_k$ gibt die Anzahl der roten Kugeln bei der zufälligen Auswahl von n Kugeln an und für alle $k \in \{0, 1, \ldots, n\}$ gilt

$$P[\{X = k\}] = \binom{n}{k} \vartheta^k \, (1-\vartheta)^{n-k}$$

Plausibilitätsprüfung: Es gilt $\sum_{k=0}^{n} \binom{n}{k} \vartheta^k \, (1-\vartheta)^{n-k} = 1$.

Urnenmodelle besitzen vielfältige Anwendungen. Als Beispiel betrachten wir nochmals den n–maligen Wurf einer Münze:

10.3.4 Beispiel (n–maliger Wurf einer Münze). Wir betrachten den n–maligen Wurf einer Münze. Wir nehmen an, dass

- zwischen den verschiedenen Würfen der Münze keine gegenseitige Beeinflussung besteht,
- bei jedem Wurf nur *Kopf* oder *Zahl* auftreten kann und
- die Chance für das Auftreten von *Kopf* beim einmaligen Wurf gleich einer rationalen Zahl $\vartheta \in \mathbb{Q} \cap (0,1)$ ist.

Aufgrund der Annahme $\vartheta \in \mathbb{Q} \cap (0,1)$ gibt es $N, K \in \mathbb{N}$ mit $K \in \{1, \dots, N-1\}$ und $\vartheta = K/N$. Der n–malige Wurf einer Münze entspricht daher dem n–maligen Ziehen mit Zurücklegen aus einer Urne mit N Kugeln, von denen K rot sind. Sei also $(\Omega, \mathcal{F}, P)$ der beim Ziehen mit Zurücklegen betrachtete Wahrscheinlichkeitsraum. Für $k \in \{0, 1, \dots, n\}$ bezeichne $E_k \in \mathcal{F}$ das Ereignis, dass beim n–maligen Ziehen mit Zurücklegen, und damit beim n–maligen Wurf der Münze, genau k–mal *Kopf* auftritt. Dann erhält man

$$P[E_k] = \binom{n}{k} \vartheta^k \, (1-\vartheta)^{n-k}$$

Im Fall $\vartheta = 1/2$ ist dies das bekannte Ergebnis.

Aufgaben

10.3.A **Problem der Doppelsechs:** Eine Spielbank bietet zwei Glücksspiele an:
- Beim ersten Spiel wirft der Spieler vier Mal einen Würfel und gewinnt, wenn keine *Sechs* auftritt.
- Beim zweiten Spiel wirft der Spieler 24 Mal zwei Würfel und gewinnt, wenn keine *Doppelsechs* auftritt.

Welches Spiel ist für den Spieler günstiger?

10.3.B **Problème des parties:** Spieler C und Spieler W werfen abwechselnd eine Münze. Vor jedem Wurf zahlen beide Spieler einen Taler in die Kasse; Spieler C erhält einen Punkt, wenn *Kopf* auftritt, und Spieler W erhält einen Punkt, wenn *Zahl* auftritt. Das Spiel endet, sobald einer der Spieler sieben Punkte erreicht hat; dieser Spieler gewinnt das Spiel und erhält die gesamte Kasse. Wie ist die Kasse aufzuteilen, wenn das Spiel abgebrochen wird und Spieler C vier und Spieler W drei Punkte erreicht hat?

10.3.C Verallgemeinern Sie Beispiel 10.3.3 auf Urnenmodelle mit mehr als zwei Sorten Kugeln.

10.4 Endliche Produkte von Wahrscheinlichkeitsräumen

In Abschnitt 9.6 haben wir gezeigt, dass ausgehend von einer endlichen Familie von Maßräumen ein neuer Maßraum konstruiert werden kann, der als Produkt der ursprünglichen Maßräume bezeichnet wird. Der folgende Satz ergibt sich unmittelbar aus der Definition des Produktmaßes:

10.4.1 Satz. *Sei $\{(\Omega_i, \mathcal{F}_i, P_i)\}_{i \in \{1,\dots,n\}}$ eine endliche Familie von Wahrscheinlichkeitsräumen. Dann ist $\bigotimes_{i \in \{1,\dots,n\}} (\Omega_i, \mathcal{F}_i, P_i)$ ein Wahrscheinlichkeitsraum.*

Für das Produkt einer endlichen Familie von diskreten oder symmetrischen Wahrscheinlichkeitsräumen ergibt sich das folgende Ergebnis:

10.4.2 Folgerung. *Sei* $\{(\Omega_i, \mathcal{F}_i, P_i)\}_{i\in\{1,\dots,n\}}$ *eine endliche Familie von Wahrscheinlichkeitsräumen.*
(1) *Sind alle* $(\Omega_i, \mathcal{F}_i, P_i)$ *diskret, so ist auch das Produkt diskret.*
(2) *Sind alle* $(\Omega_i, \mathcal{F}_i, P_i)$ *symmetrisch, so ist auch das Produkt symmetrisch.*

Als Beispiel betrachten wir wieder den n–maligen Wurf einer Münze:

10.4.3 Beispiel (n–maliger Wurf einer Münze). Wir betrachten den n–maligen Wurf einer Münze. Wir nehmen an, dass
– zwischen den verschiedenen Würfen der Münze keine gegenseitige Beeinflussung besteht,
– bei jedem Wurf nur *Kopf* oder *Zahl* auftreten kann und
– die Chance für das Auftreten von *Kopf* beim einmaligen Wurf gleich einer reellen Zahl $\vartheta \in (0,1)$ ist.

Wir wählen folgende Modelle:
– Als Modell für den i–ten Wurf wählen wir den diskreten Wahrscheinlichkeitsraum $(\Omega_i, \mathcal{F}_i, P_i)$ mit $\Omega_i := \{K, Z\}$ sowie $\mathcal{F}_i = 2^{\Omega_i}$ und dem durch $P_i[\{K\}] := \vartheta$ festgelegten Wahrscheinlichkeitsmaß.
– Als Modell für den n–fachen Wurf wählen wir das Produkt
$$(\Omega, \mathcal{F}, P) := \bigotimes_{i\in\{1,\dots,n\}} (\Omega_i, \mathcal{F}_i, P_i)$$

Dann ist auch $(\Omega, \mathcal{F}, P)$ ein diskreter Wahrscheinlichkeitsraum. Es gilt $\Omega = \{K, Z\}^n$ und $\mathcal{F} = 2^\Omega$. Sei $k \in \{0, 1, \dots, n\}$. Bezeichnet $\omega = (\omega_1, \dots, \omega_n)$ das Ergebnis, dass bei den ersten k Würfen *Kopf* und bei den letzten $n-k$ Würfen *Zahl* auftritt, so gilt aufgrund der Definition des Produktmaßes
$$P[\{\omega\}] = \prod_{i=1}^{n} P_i[\{\omega_i\}] = \prod_{i=1}^{k} P_i[\{K\}] \prod_{j=k+1}^{n} P_j[\{Z\}] = \vartheta^k \, (1-\vartheta)^{n-k}$$

Allgemeiner gilt für jedes Ergebnis $\omega \in \Omega$, bei dem genau k–mal *Kopf* (und $n-k$–mal *Zahl*) auftritt,
$$P[\{\omega\}] = \vartheta^k \, (1-\vartheta)^{n-k}$$

Bezeichnet $E_k \in \mathcal{F}$ das Ereignis, dass genau k–mal *Kopf* auftritt, so ergibt sich aus der Additivität von P
$$P[E_k] = \binom{n}{k} \vartheta^k \, (1-\vartheta)^{n-k}$$

Im Fall $\vartheta \in \mathbb{Q} \cap (0,1)$ ist dies das bekannte Ergebnis. Im Fall $\vartheta = 1/2$ ist außerdem jeder der Wahrscheinlichkeitsräume $(\Omega_i, \mathcal{F}_i, P_i)$ und damit auch das Produkt $(\Omega, \mathcal{F}, P)$ symmetrisch.

10.5 Projektive Familien von Wahrscheinlichkeitsräumen

In diesem Abschnitt bereiten wir zwei weitere Prinzipien zur Konstruktion von Wahrscheinlichkeitsräumen vor.

Sei I eine nichtleere Indexmenge und sei $\mathcal{H}(I)$ die Familie der endlichen nichtleeren Teilmengen von I.

Wir betrachten eine Familie von Messräumen $\{(\Omega_i, \mathcal{F}_i)\}_{i \in I}$ und ihr Produkt

$$(\Omega, \mathcal{F}) := \bigotimes_{i \in I} (\Omega_i, \mathcal{F}_i)$$

Für $j \in I$ sei $\pi_j : \Omega \to \Omega_j$ mit

$$\pi_j(\{\omega_i\}_{i \in I}) := \omega_j$$

die Projektion von Ω auf Ω_j. Für $J \in \mathcal{H}(I)$ sei

$$(\Omega_J, \mathcal{F}_J) := \bigotimes_{i \in J} (\Omega_i, \mathcal{F}_i)$$

das Produkt der Familie $\{(\Omega_i, \mathcal{F}_i)\}_{i \in J}$ und $\pi_J : \Omega \to \Omega_J$ mit

$$\pi_J(\omega) := \{\omega_i\}_{i \in J}$$

die Projektion von Ω auf Ω_J. Für $J \in \mathcal{H}(I)$ und $j \in J$ sei $\pi_{j,J} : \Omega_J \to \Omega_j$ mit

$$\pi_{j,J}(\{\omega_i\}_{i \in J}) := \omega_j$$

die Projektion von Ω_J auf Ω_j. Es gilt $\pi_j = \pi_{j,J} \circ \pi_J$.

10.5.1 Lemma. *Für alle $J \in \mathcal{H}(I)$ ist die Projektion π_J messbar.*

Beweis. Für alle $j \in J$ gilt $\pi_{j,J} \circ \pi_J = \pi_j$ und nach Definition von $\mathcal{F}_J$ und $\mathcal{F}$ sind $\pi_{j,J}$ und π_j messbar. Die Messbarkeit von π_J folgt nun aus Satz 3.3.2. □

Für $K, J \in \mathcal{H}(I)$ mit $K \subseteq J$ sei $\pi_{K,J} : \Omega_J \to \Omega_K$ mit

$$\pi_{K,J}(\{\omega_i\}_{i \in J}) := \{\omega_i\}_{i \in K}$$

die Projektion von Ω_J auf Ω_K. Es gilt $\pi_K = \pi_{K,J} \circ \pi_J$.

10.5.2 Lemma. *Für alle $K, J \in \mathcal{H}(I)$ mit $K \subseteq J$ ist die Projektion $\pi_{K,J}$ messbar.*

Beweis. Für alle $j \in K$ gilt $\pi_{j,K} \circ \pi_{K,J} = \pi_{j,J}$ und nach Definition von $\mathcal{F}_K$ und $\mathcal{F}_J$ sind $\pi_{j,K}$ und $\pi_{j,J}$ messbar. Die Messbarkeit von $\pi_{K,J}$ folgt nun aus Satz 3.3.2. □

Wir wenden uns nun der Frage zu, unter welchen Bedingungen es zu einer Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ von Wahrscheinlichkeitsmaßen mit $Q_J : \mathcal{F}_J \to [0,1]$ für alle $J \in \mathcal{H}(I)$ ein Wahrscheinlichkeitsmaß $Q : \mathcal{F} \to [0,1]$ gibt derart, dass für alle $J \in \mathcal{H}(I)$

$$Q_{\pi_J} = Q_J$$

gilt. Das folgende Lemma liefert eine notwendige Bedingung:

10.5.3 Lemma. *Sei $Q : \mathcal{F} \to [0,1]$ ein Wahrscheinlichkeitsmaß. Dann gilt für alle $K, J \in \mathcal{H}(I)$ mit $K \subseteq J$*

$$Q_{\pi_K} = (Q_{\pi_J})_{\pi_{K,J}}$$

Beweis. Es gilt $Q_{\pi_K} = Q_{\pi_{K,J}\circ\pi_J} = (Q_{\pi_J})_{\pi_{K,J}}$. □

Lemma 10.5.3 legt die folgende Definition nahe:

Eine Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ von Wahrscheinlichkeitsmaßen mit $Q_J : \mathcal{F}_J \to [0,1]$ für alle $J \in \mathcal{H}(I)$ heißt *projektiv*, wenn für alle $K, J \in \mathcal{H}(I)$ mit $K \subseteq J$

$$Q_K = (Q_J)_{\pi_{K,J}}$$

gilt; in diesem Fall heißt auch die Familie $\{(\Omega_J, \mathcal{F}_J, Q_J)\}_{J\in\mathcal{H}(I)}$ *projektiv*.

Nach Lemma 10.5.3 ist für jedes Wahrscheinlichkeitsmaß $Q : \mathcal{F} \to [0,1]$ die Familie $\{Q_{\pi_J}\}_{J\in\mathcal{H}(I)}$ projektiv. Daher kann es nur für eine projektive Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ ein Wahrscheinlichkeitsmaß $Q : \mathcal{F} \to [0,1]$ geben derart, dass für alle $J \in \mathcal{H}(I)$

$$Q_{\pi_J} = Q_J$$

gilt.

Als erstes machen wir die σ–Algebren $\mathcal{F}_J$, die auf unterschiedlichen Grundmengen definiert sind, miteinander vergleichbar, indem wir ihnen σ–Algebren auf Ω zuordnen: Für $J \in \mathcal{H}(I)$ sei

$$\mathcal{Z}_J := \pi_J^{-1}(\mathcal{F}_J)$$

Nach Folgerung 2.1.2 ist $\mathcal{Z}_J$ eine σ–Algebra auf Ω. Die σ–Algebra $\mathcal{Z}_J$ heißt *System der J–Zylinder* auf Ω.

10.5.4 Lemma. *Die Familie $\{\mathcal{Z}_J\}_{J\in\mathcal{H}(I)}$ ist unter Inklusion gerichtet.*

Beweis. Für $K, L \in \mathcal{H}(I)$ sei $J := K \cup L$. Dann gilt $J \in \mathcal{H}(I)$. Wegen $\pi_K = \pi_{K,J} \circ \pi_J$ gilt $\pi_K^{-1} = \pi_J^{-1} \circ \pi_{K,J}^{-1}$ und damit

$$\begin{aligned}\mathcal{Z}_K &= \pi_K^{-1}(\mathcal{F}_K)\\ &= (\pi_J^{-1} \circ \pi_{K,J}^{-1})(\mathcal{F}_K)\\ &= \pi_J^{-1}(\pi_{K,J}^{-1}(\mathcal{F}_K))\\ &\subseteq \pi_J^{-1}(\mathcal{F}_J)\\ &= \mathcal{Z}_J\end{aligned}$$

Analog erhält man

$$\mathcal{Z}_L \subseteq \mathcal{Z}_J$$

Daher gilt $\mathcal{Z}_K \cup \mathcal{Z}_L \subseteq \mathcal{Z}_J$. □

Sei nun

$$\mathcal{Z} := \bigcup_{J \in \mathcal{H}(I)} \mathcal{Z}_J$$

Dann gilt $\mathcal{Z} \subseteq 2^\Omega$. Jede Menge in $\mathcal{Z}$ heißt *Zylindermenge* auf Ω und das Mengensystem $\mathcal{Z}$ heißt *System der Zylindermengen* auf Ω.

10.5.5 Satz. *$\mathcal{Z}$ ist eine Algebra und es gilt $\sigma(\mathcal{Z}) = \mathcal{F}$.*

Beweis. Wir zeigen zunächst, dass $\mathcal{Z}$ eine Algebra ist:
(i) Für alle $J \in \mathcal{H}(I)$ gilt $\Omega \in \mathcal{Z}_J$ und aus der Definition von $\mathcal{Z}$ ergibt sich nun $\Omega \in \mathcal{Z}$.
(ii) Sei $A \in \mathcal{Z}$. Dann gibt es ein $J \in \mathcal{H}(I)$ mit $A \in \mathcal{Z}_J$. Daraus folgt zunächst $\Omega \setminus A \in \mathcal{Z}_J$ und sodann $\Omega \setminus A \in \mathcal{Z}$.
(iii) Seien $A, B \in \mathcal{Z}$. Dann gibt es ein $K \in \mathcal{H}(I)$ mit $A \in \mathcal{Z}_K$ und ein $L \in \mathcal{H}(I)$ mit $B \in \mathcal{Z}_L$. Nach Lemma 10.5.4 gibt es ein $J \in \mathcal{H}(I)$ mit $K, L \subseteq J$ und $\mathcal{Z}_K \cup \mathcal{Z}_L \subseteq \mathcal{Z}_J$. Dann gilt aber $A, B \in \mathcal{Z}_J$. Daraus folgt zunächst $A \cup B \in \mathcal{Z}_J$ und sodann $A \cup B \in \mathcal{Z}$.

Damit ist gezeigt, dass $\mathcal{Z}$ eine Algebra ist.
Für alle $i \in I$ gilt $\pi_i = \pi_{i,\{i\}} \circ \pi_{\{i\}}$ und damit

$$\pi_i^{-1}(\mathcal{F}_i) = \pi_{\{i\}}^{-1}(\pi_{i,\{i\}}^{-1}(\mathcal{F}_i)) = \pi_{\{i\}}^{-1}(\mathcal{F}_{\{i\}}) = \mathcal{Z}_{\{i\}} \subseteq \mathcal{Z} \subseteq \sigma(\mathcal{Z})$$

Aus der Definition von $\mathcal{F}$ folgt nun

$$\mathcal{F} \subseteq \sigma(\mathcal{Z})$$

Andererseits gilt für alle $J \in \mathcal{H}(I)$ nach Lemma 10.5.1

$$\mathcal{Z}_J = \pi_J^{-1}(\mathcal{F}_J) \subseteq \mathcal{F}$$

Daraus folgt zunächst $\mathcal{Z} = \bigcup_{J \in \mathcal{H}(I)} \mathcal{Z}_J \subseteq \mathcal{F}$ und sodann

$$\sigma(\mathcal{Z}) \subseteq \mathcal{F}$$

Damit ist auch die Identität $\sigma(\mathcal{Z}) = \mathcal{F}$ gezeigt. □

Wir zeigen nun, dass es zu jeder projektiven Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ einen eindeutig bestimmten Wahrscheinlichkeitsinhalt $Q : \mathcal{Z} \to [0,1]$ gibt derart, dass für alle $J \in \mathcal{H}(I)$ und alle $C \in \mathcal{F}_J$

$$Q[\pi_J^{-1}(C)] = Q_J[C]$$

gilt:

10.5.6 Satz. *Sei $\{Q_J\}_{J\in\mathcal{H}(I)}$ eine projektive Familie von Wahrscheinlichkeitsmaßen mit $Q_J : \mathcal{F}_J \to [0,1]$ für alle $J \in \mathcal{H}(I)$. Dann gibt es genau einen Wahrscheinlichkeitsinhalt $Q : \mathcal{Z} \to [0,1]$ derart, dass für alle $J \in \mathcal{H}(I)$ und alle $C \in \mathcal{F}_J$*

$$Q[\pi_J^{-1}(C)] = Q_J[C]$$

gilt.

Beweis. Die Definition von Q erfordert eine Vorüberlegung: Sei $A \in \mathcal{Z}$. Für $K, L \in \mathcal{H}(I)$ mit $A \in \mathcal{Z}_K \cap \mathcal{Z}_L$ sei $M := K \cup L$. Dann gilt $K, L \subseteq M$ und damit $A \in \mathcal{Z}_M$. Daher gibt es $C_K \in \mathcal{F}_K$, $C_L \in \mathcal{F}_L$ und $C_M \in \mathcal{F}_M$ mit

$$\begin{aligned} A &= \pi_K^{-1}(C_K) \\ A &= \pi_L^{-1}(C_L) \\ A &= \pi_M^{-1}(C_M) \end{aligned}$$

Wegen $\pi_M^{-1}(\pi_{K,M}^{-1}(C_K)) = (\pi_{K,M} \circ \pi_M)^{-1}(C_K) = \pi_K^{-1}(C_K) = A = \pi_M^{-1}(C_M)$ gilt

$$\pi_{K,M}^{-1}(C_K) = C_M$$

und aus der Projektivität der Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ folgt nun

$$\begin{aligned} Q_K[C_K] &= (Q_M)_{\pi_{K,M}}[C_K] \\ &= Q_M[\pi_{K,M}^{-1}(C_K)] \\ &= Q_M[C_M] \end{aligned}$$

Aus Symmetriegründen gilt auch

$$Q_L[C_L] = Q_M[C_M]$$

und damit

$$Q_K[C_K] = Q_L[C_L]$$

Aufgrund dieser Vorüberlegung ist die Abbildung $Q : \mathcal{Z} \to [0,1]$ mit

$$Q[A] := Q_J[C]$$

für eine beliebige Menge $J \in \mathcal{H}(I)$ mit $A \in \mathcal{Z}_J$ und eine Menge $C \in \mathcal{F}_J$ mit $A = \pi_J^{-1}(C)$ wohldefiniert.

Wir zeigen nun, dass Q ein Wahrscheinlichkeitsinhalt ist:

(i) Für alle $A, B \in \mathcal{Z}$ mit $A \cap B = \emptyset$ gibt es ein $J \in \mathcal{H}(I)$ mit $A, B \in \mathcal{Z}_J$. Daher gibt es $C, D \in \mathcal{F}_J$ mit $C \cap D = \emptyset$ und

$$\begin{aligned} A &= \pi_J^{-1}(C) \\ B &= \pi_J^{-1}(D) \end{aligned}$$

und es gilt

$$\begin{aligned} Q[A+B] &= Q[\pi_J^{-1}(C) + \pi_J^{-1}(D)] \\ &= Q[\pi_J^{-1}(C+D)] \\ &= Q_J[C+D] \\ &= Q_J[C] + Q_J[D] \\ &= Q[\pi_J^{-1}(C)] + Q[\pi_J^{-1}(D)] \\ &= Q[A] + Q[B] \end{aligned}$$

(ii) Für alle $J \in \mathcal{H}(I)$ gilt $\Omega = \pi_J^{-1}(\Omega_J)$ und damit

$$Q[\Omega] = Q[\pi_J^{-1}(\Omega_J)] = Q_J[\Omega_J] = 1$$

(iii) Aus (i) und (ii) folgt $Q[\emptyset] = 0$.

Daher ist Q ein Wahrscheinlichkeitsinhalt. Die Aussage über die Eindeutigkeit ist dann klar. □

Für eine projektive Familie von Wahrscheinlichkeitsmaßen $\{Q_J\}_{J \in \mathcal{H}(I)}$ mit $Q_J : \mathcal{F}_J \to [0,1]$ für alle $J \in \mathcal{H}(I)$ stellt sich nun die Frage, unter welchen Bedingungen der nach Satz 10.5.6 eindeutig bestimmte Wahrscheinlichkeitsinhalt $Q : \mathcal{Z} \to [0,1]$ mit

$$Q[\pi_J^{-1}(C)] = Q_J[C]$$

für alle $J \in \mathcal{H}(I)$ und $C \in \mathcal{F}_J$ sogar ein Wahrscheinlichkeitsmaß ist; in diesem Fall besitzt Q nach dem Satz von Caratheodory eine eindeutige Fortsetzung zu einem Wahrscheinlichkeitsmaß $Q : \mathcal{F} \to [0,1]$, denn nach Satz 10.5.5 ist $\mathcal{Z}$ eine Algebra mit $\sigma(\mathcal{Z}) = \mathcal{F}$, und damit ein Halbring mit $\Omega \in \mathcal{Z}$ und $\sigma(\mathcal{Z}) = \mathcal{F}$. Für das Wahrscheinlichkeitsmaß Q gilt dann für alle $J \in \mathcal{H}(I)$

$$Q_{\pi_J} = Q_J$$

Es gibt zwei Bedingungen, die die σ–Additivität von Q gewährleisten:

– Die eine dieser Bedingungen betrifft die Wahrscheinlichkeitsmaße Q_J und führt auf den Satz von Andersen/Jessen, den wir im nächsten Abschnitt beweisen.
– Die andere dieser Bedingungen betrifft die Messräume $(\Omega_i, \mathcal{F}_i)$ und führt auf den Satz von Kolmogorov, den wir in Abschnitt 20.2 beweisen.

10.6 Satz von Andersen/Jessen

Sei I eine nichtleere Indexmenge und sei $\mathcal{H}(I)$ die Familie der endlichen nichtleeren Teilmengen von I.

Wir betrachten eine Familie von Wahrscheinlichkeitsräumen $\{(\Omega_i, \mathcal{F}_i, Q_i)\}_{i\in I}$ und das Produkt

$$(\Omega, \mathcal{F}) := \bigotimes_{i\in I} (\Omega_i, \mathcal{F}_i)$$

der Familie $\{(\Omega_i, \mathcal{F}_i)\}_{i\in I}$ von Messräumen. Unser Ziel ist es zu zeigen, dass es genau ein Wahrscheinlichkeitsmaß $Q : \mathcal{F} \to [0,1]$ gibt derart, dass für alle $J \in \mathcal{H}(I)$

$$Q_{\pi_J} = \bigotimes_{i\in J} Q_i$$

gilt. Nach Satz 10.4.1 ist für alle $J \in \mathcal{H}(I)$ das endliche Produkt

$$(\Omega_J, \mathcal{F}_J, Q_J) := \bigotimes_{i\in J} (\Omega_i, \mathcal{F}_i, Q_i)$$

der Familie $\{(\Omega_i, \mathcal{F}_i, Q_i)\}_{i\in J}$ ein Wahrscheinlichkeitsraum mit

$$Q_J = \bigotimes_{i\in J} Q_i$$

Wir untersuchen zunächst die Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$:

10.6.1 Lemma. *Die Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ ist projektiv.*

Beweis. Sei $K, J \in \mathcal{H}(I)$ mit $K \subseteq J$. Wir betrachten eine Familie $\{A_i\}_{i\in K}$ mit $A_i \in \mathcal{F}_i$ für alle $i \in K$. Für alle $i \in J$ sei

$$B_i := \begin{cases} A_i & \text{falls} \quad i \in K \\ \Omega_i & \text{falls} \quad i \in J \setminus K \end{cases}$$

Dann gilt $B_i \in \mathcal{F}_i$ für alle $i \in J$ sowie $\prod_{i\in J} B_i = \pi_{K,J}^{-1}(\prod_{i\in K} A_i)$. Daher gilt

$$\begin{aligned} Q_K\left[\prod_{i\in K} A_i\right] &= \prod_{i\in K} Q_i[A_i] \\ &= \prod_{i\in J} Q_i[B_i] \\ &= Q_J\left[\prod_{i\in J} B_i\right] \end{aligned}$$

$$= Q_J\left[\pi_{K,J}^{-1}\left(\prod_{i\in K} A_i\right)\right]$$

$$= (Q_J)_{\pi_{K,J}}\left[\prod_{i\in K} A_i\right]$$

Da $\prod_{i\in K}\mathcal{F}_i$ ein $\cap$–stabiler Erzeuger von $\mathcal{F}_K$ ist, folgt hieraus

$$Q_K = (Q_J)_{\pi_{K,J}}$$

Daher ist die Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ projektiv. □

Wir können nun den angekündigten Satz von Andersen/Jessen beweisen:

10.6.2 Satz (Andersen/Jessen). *Es gibt genau ein Wahrscheinlichkeitsmaß $Q : \mathcal{F} \to [0,1]$ derart, dass für alle $J \in \mathcal{H}(I)$*

$$Q_{\pi_J} = \bigotimes_{i\in J} Q_i$$

gilt.

Beweis. Im Fall einer endlichen Indexmenge I ist nichts zu zeigen. Sei daher I unendlich. Sei ferner $\mathcal{Z}$ die Algebra der Zylindermengen auf $\mathbb{R}^I$ und für $J \in \mathcal{H}(I)$ sei $\mathcal{Z}_J = \pi_J^{-1}(\mathcal{B}(\mathbb{R}^J))$ die σ–Algebra der J–Zylinder.
Nach Lemma 10.6.1 und Satz 10.5.6 existiert genau ein Wahrscheinlichkeitsinhalt $Q : \mathcal{Z} \to [0,1]$ mit

$$Q[\pi_J^{-1}(C)] = Q_J[C]$$

für alle $J \in \mathcal{H}(I)$ und alle $C \in \mathcal{F}_J$. Wir zeigen im folgenden, dass Q $\emptyset$–stetig ist. Dann ist Q nach Lemma 10.1.2 σ–additiv und besitzt nach dem Satz von Caratheodory eine eindeutige Fortsetzung zu einem Wahrscheinlichkeitsmaß $\mathcal{F} \to [0,1]$, das wir wieder mit Q bezeichnen; außerdem gilt für alle $J \in \mathcal{H}(I)$ und alle $C \in \mathcal{F}_J$

$$Q[\pi_J^{-1}(C)] = Q_J[C]$$

und damit für alle $J \in \mathcal{H}(I)$

$$Q_{\pi_J} = Q_J$$

(1) Wir betrachten zunächst $K, J \in \mathcal{H}(I)$ mit $K \subseteq J$ und $K \neq J$. Aufgrund der Definition der Wahrscheinlichkeitsmaße $Q_K, Q_{J\setminus K}, Q_J$ ist der Wahrscheinlichkeitsraum $(\Omega_J, \mathcal{F}_J, Q_J)$ das Produkt der Wahrscheinlichkeitsräume $(\Omega_K, \mathcal{F}_K, Q_K)$ und $(\Omega_{J\setminus K}, \mathcal{F}_{J\setminus K}, Q_{J\setminus K})$; es gilt also

$$(\Omega_J, \mathcal{F}_J, Q_J) = (\Omega_K, \mathcal{F}_K, Q_K) \otimes (\Omega_{J\setminus K}, \mathcal{F}_{J\setminus K}, Q_{J\setminus K})$$

Für $C \in \mathcal{F}_J$ und $\omega_K \in \Omega_K$ sei

$$C(\omega_K) := \Bigl\{\omega \in \Omega_{J\setminus K} \Bigm| \pi_{K,J}^{-1}(\{\omega_K\}) \cap \pi_{J\setminus K,J}^{-1}(\{\omega\}) \subseteq C\Bigr\}$$

Dann ist $C(\omega_K)$ der ω_K–Schnitt von C in $\Omega_{J\setminus K}$. Aus Lemma 9.6.4 folgt nun $C(\omega_K) \in \mathcal{F}_{J\setminus K}$, nach Lemma 9.6.5 ist die Abbildung $\Omega_K \to [0,1] : \omega_K \mapsto Q_{J\setminus K}[C(\omega_K)]$ messbar, und aus Satz 9.6.6 folgt

$$\begin{aligned} Q_J[C] &= (Q_K \otimes Q_{J\setminus K})[C] \\ &= \int_{\Omega_K} Q_{J\setminus K}[C(\omega_K)]\, dQ_K(\omega_K) \end{aligned}$$

(2) Wir betrachten nun eine Zylindermenge $A \in \mathcal{Z}$ sowie $K \in \mathcal{H}(I)$. Dann gibt es ein $J \in \mathcal{H}(I)$ mit $A \in \mathcal{Z}_J$ und $K \subseteq J$ sowie $J \setminus K \neq \emptyset$, und es gibt eine Menge $C \in \mathcal{F}_J$ mit $A = \pi_J^{-1}(C)$. Für $\omega_K \in \Omega_K$ sei

$$A(\omega_K) := \Bigl\{\omega \in \Omega \Bigm| \pi_K^{-1}(\{\omega_K\}) \cap \pi_{I\setminus K}^{-1}(\{\pi_{I\setminus K}(\omega)\}) \subseteq A\Bigr\}$$

Wegen $A \in \mathcal{Z}_J$ und $C \in \mathcal{F}_J$ gilt dann

$$\begin{aligned} A(\omega_K) &= \Bigl\{\omega \in \Omega \Bigm| \pi_K^{-1}(\{\omega_K\}) \cap \pi_{I\setminus K}^{-1}(\{\pi_{I\setminus K}(\omega)\}) \subseteq A\Bigr\} \\ &= \Bigl\{\omega \in \Omega \Bigm| \pi_K^{-1}(\{\omega_K\}) \cap \pi_{J\setminus K}^{-1}(\{\pi_{J\setminus K}(\omega)\}) \subseteq A\Bigr\} \\ &= \Bigl\{\omega \in \Omega \Bigm| \pi_J^{-1}(\pi_{K,J}^{-1}(\{\omega_K\})) \cap \pi_J^{-1}(\pi_{J\setminus K,J}^{-1}(\{\pi_{J\setminus K}(\omega)\})) \subseteq \pi_J^{-1}(C)\Bigr\} \\ &= \Bigl\{\omega \in \Omega \Bigm| \pi_{K,J}^{-1}(\{\omega_K\}) \cap \pi_{J\setminus K,J}^{-1}(\{\pi_{J\setminus K}(\omega)\}) \subseteq C\Bigr\} \\ &= \Bigl\{\omega \in \Omega \Bigm| \pi_{J\setminus K}(\omega) \in C(\omega_K)\Bigr\} \\ &= \pi_{J\setminus K}^{-1}(C(\omega_K)) \end{aligned}$$

Aus (1) ergibt sich nun $A(\omega_K) = \pi_{J\setminus K}^{-1}(C(\omega_K)) \in \pi_{J\setminus K}^{-1}(\mathcal{F}_{J\setminus K}) = \mathcal{Z}_{J\setminus K}$ und wegen

$$\begin{aligned} Q[A(\omega_K)] &= Q[\pi_{J\setminus K}^{-1}(C(\omega_K))] \\ &= Q_{J\setminus K}[C(\omega_K)] \end{aligned}$$

die Messbarkeit der Abbildung $\Omega_K \to [0,1] : \omega_K \mapsto Q[A(\omega_K)]$ sowie

$$\begin{aligned} Q[A] &= Q[\pi_J^{-1}(C)] \\ &= Q_J[C] \\ &= \int_{\Omega_K} Q_{J\setminus K}[C(\omega_K)]\, dQ_K(\omega_K) \\ &= \int_{\Omega_K} Q[A(\omega_K)]\, dQ_K(\omega_K) \end{aligned}$$

(3) Wir betrachten nun eine monoton fallende Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Z}$ mit

$$\inf_{n\in\mathbb{N}} Q[A_n] > 0$$

und zeigen, dass

$$\bigcap_{n\in\mathbb{N}} A_n \neq \emptyset$$

gilt.
Sei $\{K_n\}_{n\in\mathbb{N}} \subseteq \mathcal{H}(I)$ eine streng monoton wachsende Folge mit $A_n \in \mathcal{Z}_{K_n}$ für alle $n \in \mathbb{N}$ und sei

$$\alpha := \inf_{n\in\mathbb{N}} Q[A_n]$$

Wir konstruieren weiter unten eine Folge $\{\omega_{K_m}\}_{m\in\mathbb{N}}$ mit

$$\begin{aligned} \omega_{K_m} &\in \Omega_{K_m} \\ \pi^{-1}_{K_m}(\{\omega_{K_m}\}) &\subseteq A_m \\ Q[A_n(\omega_{K_m})] &\geq \alpha/2^m \end{aligned}$$

sowie

$$\pi_{K_m,K_{m+1}}(\omega_{K_{m+1}}) = \omega_{K_m}$$

für alle $m, n \in \mathbb{N}$. Dann gilt $\omega_{K_{m+1}} \subseteq \pi^{-1}_{K_m,K_{m+1}}(\{\omega_{K_m}\})$ und damit

$$\begin{aligned} \pi^{-1}_{K_{m+1}}(\{\omega_{K_{m+1}}\}) &\subseteq \pi^{-1}_{K_{m+1}}(\pi^{-1}_{K_m,K_{m+1}}(\{\omega_{K_m}\})) \\ &= \pi^{-1}_{K_m}(\{\omega_{K_m}\}) \end{aligned}$$

Daher ist die Folge $\{\pi^{-1}_{K_m}(\{\omega_{K_m}\})\}_{m\in\mathbb{N}}$ monoton fallend mit

$$\bigcap_{m\in\mathbb{N}} \pi^{-1}_{K_m}(\{\omega_{K_m}\}) \neq \emptyset$$

und für alle $n \in \mathbb{N}$ gilt $\bigcap_{m\in\mathbb{N}} \pi^{-1}_{K_m}(\{\omega_{K_m}\}) \subseteq \pi^{-1}_{K_n}(\{\omega_{K_n}\}) \subseteq A_n$. Daher gilt

$$\bigcap_{m\in\mathbb{N}} \pi^{-1}_{K_m}(\{\omega_{K_m}\}) \subseteq \bigcap_{n\in\mathbb{N}} A_n$$

und damit

$$\bigcap_{n\in\mathbb{N}} A_n \neq \emptyset$$

(4) Nach (3) ist Q $\emptyset$–stetig.

(5) Wir tragen nun die Konstruktion der Folge $\{\omega_{K_m}\}_{m\in\mathbb{N}}$ nach:

– $m = 1$: Mit der Folge $\{A_n\}_{n\in\mathbb{N}}$ ist für jedes $\omega' \in \Omega_{K_1}$ auch die Folge $\{A_n(\omega')\}_{n\in\mathbb{N}}$ monoton fallend; damit ist auch die Folge $\{B_n\}_{n\in\mathbb{N}}$ mit

$$B_n := \left\{\omega' \in \Omega_{K_1} \;\middle|\; Q[A_n(\omega')] \geq \alpha/2\right\}$$

monoton fallend. Für alle $n \in \mathbb{N}$ gilt nach (2) $B_n \in \mathcal{F}_{K_1}$ und

$$\begin{aligned}\alpha &\leq Q[A_n] \\ &= \int_{\Omega_{K_1}} Q[A_n(\omega')]\, dQ_{K_1}(\omega') \\ &= \int_{B_n} Q[A_n(\omega')]\, dQ_{K_1}(\omega') + \int_{\Omega_{K_1}\setminus B_n} Q[A_n(\omega')]\, dQ_{K_1}(\omega') \\ &\leq \int_{B_n} 1\, dQ_{K_1}(\omega') + \int_{\Omega_{K_1}\setminus B_n} \frac{\alpha}{2}\, dQ_{K_1}(\omega') \\ &\leq Q_{K_1}[B_n] + \frac{\alpha}{2}\end{aligned}$$

und damit

$$\frac{\alpha}{2} \leq \inf_{n\in\mathbb{N}} Q_{K_1}[B_n] = Q_{K_1}\left[\bigcap_{n\in\mathbb{N}} B_n\right]$$

Da Q_{K_1} ein Wahrscheinlichkeitsmaß ist, gilt $\bigcap_{n\in\mathbb{N}} B_n \neq \emptyset$. Daher gibt es ein $\omega_{K_1} \in \bigcap_{n=1}^{\infty} B_n$. Aus der Definition der Folge $\{B_n\}_{n\in\mathbb{N}}$ ergibt sich

$$\omega_{K_1} \in \Omega_{K_1}$$

sowie

$$\frac{\alpha}{2} \leq Q[A_n(\omega_{K_1})]$$

für alle $n \in \mathbb{N}$. Insbesondere ist die Menge

$$A_1(\omega_{K_1}) = \left\{\omega \in \Omega \;\middle|\; \pi_{K_1}^{-1}(\{\omega_{K_1}\}) \cap \pi_{I\setminus K_1}^{-1}(\{\pi_{I\setminus K_1}(\omega)\}) \subseteq A_1\right\}$$

nichtleer und wegen $A_1 \in \mathcal{Z}_{K_1}$ gilt

$$\pi_{K_1}^{-1}(\{\omega_{K_1}\}) \subseteq A_1$$

– $m \to m+1$: Sei $\omega_{K_m} \in \Omega_{K_m}$ bereits konstruiert. Mit der Folge $\{A_n\}_{n\in\mathbb{N}}$ ist auch die Folge $\{A(\omega_{K_m})\}_{n\in\mathbb{N}}$ und damit für jedes $\omega' \in \Omega_{K_{m+1}\setminus K_m}$ auch die Folge $\{(A_n(\omega_{K_m}))(\omega')\}_{n\in\mathbb{N}}$ monoton fallend; damit ist auch die Folge $\{B_n\}_{n\in\mathbb{N}}$ mit

$$B_n := \left\{\omega' \in \Omega_{K_{m+1}\setminus K_m} \;\middle|\; Q[(A_n(\omega_{K_m}))(\omega')] \geq \alpha/2^{m+1}\right\}$$

monoton fallend. Für alle $n \in \mathbb{N}$ gilt nach (2) $B_n \in \mathcal{F}_{K_{m+1}\setminus K_m}$ und

$$\begin{aligned}
\frac{\alpha}{2^m} &\leq Q[(A_n(\omega_{K_m}))] \\
&= \int_{\Omega_{K_{m+1}\setminus K_m}} Q[(A_n(\omega_{K_m}))(\omega')]\, dQ_{K_{m+1}\setminus K_m}(\omega') \\
&= \int_{B_n} Q[(A_n(\omega_{K_m}))(\omega')]\, dQ_{K_{m+1}\setminus K_m}(\omega') \\
&\quad + \int_{(\Omega_{K_{m+1}\setminus K_m})\setminus B_n} Q[(A_n(\omega_{K_m}))(\omega')]\, dQ_{K_{m+1}\setminus K_m}(\omega') \\
&\leq \int_{B_n} 1\, dQ_{K_{m+1}\setminus K_m}(\omega') + \int_{(\Omega_{K_{m+1}\setminus K_m})\setminus B_n} \frac{\alpha}{2^{m+1}}\, dQ_{K_{m+1}\setminus K_m}(\omega') \\
&\leq Q_{K_{m+1}\setminus K_m}[B_n] + \frac{\alpha}{2^{m+1}}
\end{aligned}$$

und damit

$$\frac{\alpha}{2^{m+1}} \leq \inf_{n\in\mathbb{N}} Q_{K_{m+1}\setminus K_m}[B_n] = Q_{K_{m+1}\setminus K_m}\left[\bigcap_{n\in\mathbb{N}} B_n\right]$$

Da $Q_{K_{m+1}\setminus K_m}$ ein Wahrscheinlichkeitsmaß ist, gilt $\bigcap_{n\in\mathbb{N}} B_n \neq \emptyset$. Daher gibt es ein $\omega_{K_{m+1}\setminus K_m} \in \bigcap_{n\in\mathbb{N}} B_n$. Aus der Definition der Folge $\{B_n\}_{n\in\mathbb{N}}$ ergibt sich

$$\omega_{K_{m+1}\setminus K_m} \in \Omega_{K_{m+1}\setminus K_m}$$

sowie

$$\frac{\alpha}{2^{m+1}} \leq Q[(A_n(\omega_{K_m}))(\omega_{K_{m+1}\setminus K_m})]$$

für alle $n \in \mathbb{N}$. Insbesondere ist die Menge $(A_{m+1}(\omega_{K_m}))(\omega_{K_{m+1}\setminus K_m})$ nichtleer und wegen $A_{m+1}(\omega_{K_m}) \in \mathcal{Z}_{K_{m+1}\setminus K_m}$ gilt

$$\pi^{-1}_{K_{m+1}\setminus K_m}(\{\omega_{K_{m+1}\setminus K_m}\}) \subseteq A_{m+1}(\omega_{K_m})$$

Sei nun

$$\omega_{K_{m+1}} := \pi_{K_{m+1}}(\pi^{-1}_{K_m}(\{\omega_{K_m}\}) \cap \pi^{-1}_{K_{m+1}\setminus K_m}(\{\omega_{K_{m+1}\setminus K_m}\}))$$

Dann gilt

$$\omega_{K_{m+1}} \in \Omega_{K_{m+1}}$$

und

$$\pi_{K_m,K_{m+1}}(\omega_{K_{m+1}}) = \omega_{K_m}$$

und

$$\pi^{-1}_{K_{m+1}}(\{\omega_{K_{m+1}}\}) = \pi^{-1}_{K_m}(\{\omega_{K_m}\}) \cap \pi^{-1}_{K_{m+1}\setminus K_m}(\{\omega_{K_{m+1}\setminus K_m}\})$$

Daher gilt für alle $\omega \in \Omega$

$$\begin{aligned}&\pi^{-1}_{K_{m+1}}(\{\omega_{K_{m+1}}\}) \cap \pi^{-1}_{I\setminus K_{m+1}}(\{\pi_{I\setminus K_{m+1}}(\omega)\})\\ &= \pi^{-1}_{K_m}(\{\omega_{K_m}\}) \cap \pi^{-1}_{K_{m+1}\setminus K_m}(\{\omega_{K_{m+1}\setminus K_m}\}) \cap \pi^{-1}_{I\setminus K_{m+1}}(\{\pi_{I\setminus K_{m+1}}(\omega)\})\end{aligned}$$

Wegen

$$\begin{aligned}&\pi^{-1}_{K_{m+1}\setminus K_m}(\{\omega_{K_{m+1}\setminus K_m}\}) \cap \pi^{-1}_{I\setminus K_{m+1}}(\{\pi_{I\setminus K_{m+1}}(\omega)\})\\ &= \pi^{-1}_{I\setminus K_m}(\{\pi_{I\setminus K_m}(\pi^{-1}_{K_{m+1}\setminus K_m}(\{\omega_{K_{m+1}\setminus K_m}\}) \cap \pi^{-1}_{I\setminus K_{m+1}}(\{\pi_{I\setminus K_{m+1}}(\omega)\}))\})\end{aligned}$$

gilt daher

$$\omega \in A_n(\omega_{K_{m+1}})$$

genau dann, wenn

$$\pi^{-1}_{K_{m+1}\setminus K_m}(\{\omega_{K_{m+1}\setminus K_m}\}) \cap \pi^{-1}_{I\setminus K_{m+1}}(\{\pi_{I\setminus K_{m+1}}(\omega)\}) \in A_n(\omega_{K_m})$$

gilt, und diese Bedingung ist genau dann erfüllt, wenn

$$\omega \in (A_n(\omega_{K_m}))(\omega_{K_{m+1}\setminus K_m})$$

gilt. Daher gilt für alle $n \in \mathbb{N}$

$$A_n(\omega_{K_{m+1}}) = (A_n(\omega_{K_m}))(\omega_{K_{m+1}\setminus K_m})$$

und damit

$$\frac{\alpha}{2^{m+1}} \leq Q[A_n(\omega_{K_{m+1}})]$$

Insbesondere gilt $A_{m+1}(\omega_{K_{m+1}}) \neq \emptyset$. Daher gibt es ein $\omega \in \Omega$ mit

$$\pi^{-1}_{K_{m+1}}(\{\omega_{K_{m+1}}\}) \cap \pi^{-1}_{I\setminus K_{m+1}}(\{\pi_{I\setminus K_{m+1}})(\omega)\} \subseteq A_{m+1}$$

Wegen $A_{m+1} \in \mathcal{Z}_{K_{m+1}}$ gilt dann aber

$$\pi^{-1}_{K_{m+1}}(\{\omega_{K_{m+1}}\}) \subseteq A_{m+1}$$

Damit ist die Folge $\{\omega_{K_m}\}_{m\in\mathbb{N}}$ konstruiert. □

Der durch den Satz von Andersen/Jessen gegebene Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, Q)$ heißt *Produkt* der Familie $\{(\Omega_i, \mathcal{F}_i, Q_i)\}_{i\in I}$ und das Wahrscheinlichkeitsmaß Q heißt *Produkt* der Familie $\{Q_i\}_{i\in I}$. Es gilt

$$\Omega = \prod_{i \in I} \Omega_i$$
$$\mathcal{F} = \bigotimes_{i \in I} \mathcal{F}_i$$

und wir setzen

$$\bigotimes_{i \in I} Q_i := Q$$

und

$$\bigotimes_{i \in I} (\Omega_i, \mathcal{F}_i, Q_i) := \left(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{F}_i, \bigotimes_{i \in I} Q_i \right)$$

Die Notation ist damit analog zum Fall eines endlichen Produktes.

Wir geben abschließend ein einfaches Beispiel für die Anwendung des Satzes von Andersen/Jessen:

10.6.3 Beispiel (Wurf einer Münze bis zum ersten Kopf). Wir betrachten den wiederholten Wurf einer Münze. Wir nehmen an, dass
- zwischen den verschiedenen Würfen der Münze keine gegenseitige Beeinflussung besteht,
- bei jedem Wurf nur *Kopf* oder *Zahl* auftreten kann,
- die Chance für das Auftreten von *Kopf* beim einmaligen Wurf gleich einer reellen Zahl $\vartheta \in (0, 1)$ ist, und
- die Münze solange geworfen wird, bis zum erstenmal *Kopf* auftritt.

Wir interessieren uns für die Anzahl der Würfe, bis zum erstenmal *Kopf* auftritt. Wir wählen folgende Modelle:
- Als Modell für den i–ten Wurf wählen wir den diskreten Wahrscheinlichkeitsraum $(\Omega_i, \mathcal{F}_i, P_i)$ mit $\Omega_i := \{K, Z\}$ sowie $\mathcal{F}_i = 2^{\Omega_i}$ und dem durch $P_i[\{K\}] := \vartheta$ festgelegten Wahrscheinlichkeitsmaß.
- Als Modell für den wiederholten Wurf wählen wir den Wahrscheinlichkeitsraum

$$(\Omega, \mathcal{F}, P) := \bigotimes_{i \in \mathbb{N}} (\Omega_i, \mathcal{F}_i, P_i)$$

Dann gilt $\Omega = \{K, Z\}^{\mathbb{N}}$ und damit ist Ω überabzählbar. Sei $n \in \mathbb{N}$ und sei

$$A_n := \prod_{i=1}^{n-1} \{Z\} \times \{K\} \times \prod_{j=n+1}^{\infty} \Omega_j$$

Dann gilt $A_n \in \mathcal{Z} \subseteq \mathcal{F}$ und A_n ist gerade das Ereignis, dass *Kopf* zum erstenmal beim n–ten Wurf auftritt. Sei ferner

$$(\Omega_{\{1,\dots,n\}}, \mathcal{F}_{\{1,\dots,n\}}, P_{\{1,\dots,n\}}) := \bigotimes_{i \in \{1,\dots,n\}} (\Omega_i, \mathcal{F}_i, P_i)$$

und

$$C_n := \prod_{i=1}^{n-1} \{Z\} \times \{K\}$$

Dann gilt $C_n \in \mathcal{F}_{\{1,\dots,n\}}$ und $A_n = \pi^{-1}_{\{1,\dots,n\}}(C_n)$, und nach Konstruktion der Wahrscheinlichkeitsmaße P und $P_{\{1,\dots,n\}}$ gilt

$$P_{\pi_{\{1,\dots,n\}}} = P_{\{1,\dots,n\}} = \bigotimes_{i\in\{1,\dots,n\}} P_i$$

Wir erhalten daher

$$\begin{aligned} P[A_n] &= P[\pi^{-1}_{\{1,\dots,n\}}(C_n)] \\ &= P_{\pi_{\{1,\dots,n\}}}[C_n] \\ &= \left(\bigotimes_{i\in\{1,\dots,n\}} P_i\right)\left[\prod_{i=1}^{n-1}\{Z\}\times\{K\}\right] \\ &= \left(\prod_{i=1}^{n-1} P_i[\{Z\}]\right) P_n[\{K\}] \\ &= (1-\vartheta)^{n-1}\,\vartheta \end{aligned}$$

Die Folge $\{A_n\}_{n\in\mathbb{N}}$ ist disjunkt mit $P[\sum_{n=1}^\infty A_n] = 1$. Die Abbildung $X : \Omega \to \bar{\mathbb{R}}$ mit

$$X(\omega) := \begin{cases} n & \text{falls } \omega \in A_n \\ \infty & \text{falls } \omega \in \Omega \setminus \sum_{n=1}^\infty A_n \end{cases}$$

gibt die Anzahl der Würfe bis zum ersten Auftreten von *Kopf* an. Dann ist X eine Zufallsvariable und für alle $n \in \mathbb{N}$ gilt

$$P[\{X = n\}] = (1-\vartheta)^{n-1}\,\vartheta$$

Daher gilt $P[\{X \in \mathbb{N}\}] = 1$ und die Münze wird mit Wahrscheinlichkeit Eins nur endlich viele Male geworfen.

Die in diesem Beispiel und auch in den vorher betrachteten Beispielen zum wiederholten Münzwurf getroffene Annahme, dass zwischen den verschiedenen Würfen der Münze keine gegenseitige Beeinflussung besteht, bildet den heuristischen Hintergrund für einen zentralen wahrscheinlichkeitstheoretischen Begriff, den wir im nächsten Kapitel einführen und untersuchen.

Aufgabe

10.6.A Wurf einer Münze bis zum ersten Kopf: Verwenden Sie die Modelle aus Beispiel 10.6.3.

(1) Beschreiben Sie alle Ereignisse der Algebra $\mathcal{Z}$.
(2) Beschreiben Sie alle Ereignisse der σ–Algebra $\mathcal{F}$.
(3) Sei $B := \prod_{i=1}^\infty \{Z\}$. Dann gilt $B \in \mathcal{F} \setminus \mathcal{Z}$ und $P[B] = 0$.

11

Unabhängigkeit

In diesem Kapitel und in den weiteren Kapiteln sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, auf dem alle Zufallsgrößen definiert sind.

In diesem Kapitel betrachten wir den für die Wahrscheinlichkeitstheorie grundlegenden Begriff der Unabhängigkeit, der in drei Formen auftreten kann:
- Unabhängigkeit einer Familie von Ereignissen (Abschnitt 11.1)
- Unabhängigkeit einer Familie von Ereignissystemen (Abschnitt 11.2)
- Unabhängigkeit einer Familie von Zufallsgrößen (Abschnitt 11.3)

Diese drei Formen der Unabhängigkeit hängen eng miteinander zusammen und treten in natürlicher Weise in Produkten von Wahrscheinlichkeitsräumen auf (Abschnitt 11.4).

Sei I eine nichtleere Indexmenge und sei $\mathcal{H}(I)$ die Familie der endlichen nichtleeren Teilmengen von I.

11.1 Unabhängige Familien von Ereignissen

Ziel dieses Abschnitts ist es, den Begriff der Unabhängigkeit für eine Familie von Ereignissen zu definieren und zu untersuchen.

Wir beginnen mit einer Vorbetrachtung.

Sei $C \in \mathcal{F}$ ein Ereignis mit $P[C] \in (0,1)$. Dann ist das unbestimmte Integral

$$P[\,.\,|C] := \int \frac{\chi_C}{P[C]}\, dP$$

ein Wahrscheinlichkeitsmaß mit

$$P[A|C] = \int_A \frac{\chi_C}{P[C]}\, dP = \frac{1}{P[C]} \int_\Omega \chi_{A\cap C}\, dP = \frac{P[A\cap C]}{P[C]}$$

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_11, © Springer-Verlag Berlin Heidelberg 2011

Das Wahrscheinlichkeitsmaß $P[\,.\,|C]$ heißt *bedingtes Wahrscheinlichkeitsmaß* unter C, und für $A \in \mathcal{F}$ heißt $P[A|C]$ die *bedingte Wahrscheinlichkeit* von A unter C. Wegen $P[C|C] = 1$ und $P[\overline{C}|C] = 0$ lebt das Wahrscheinlichkeitsmaß $P[\,.\,|C]$ auf dem Ereignis C und verschwindet auf seinem Komplement. Der Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P[\,.\,|C])$ kann als Modell für das nach dem Eintritt des Ereignisses C verbleibende Zufallsgeschehen interpretiert werden.

Ist $C \in \mathcal{F}$ ein Ereignis mit $P[C] \in (0,1)$, so ist auch $\overline{C} \in \mathcal{F}$ ein Ereignis mit $P[\overline{C}] \in (0,1)$, und in diesem Fall heißt ein Ereignis $A \in \mathcal{F}$ *unabhängig von C*, wenn

$$P[A|C] = P[A|\overline{C}]$$

gilt.

11.1.1 Lemma. *Sei $C \in \mathcal{F}$ ein Ereignis mit $P[C] \in (0,1)$. Dann sind für jedes Ereignis $A \in \mathcal{F}$ folgende Aussagen äquivalent:*
(a) *A ist unabhängig von C.*
(b) *Es gilt $P[A|C] = P[A]$.*
(c) *Es gilt $P[A \cap C] = P[A]\,P[C]$.*

Beweis. Aus der Definition der bedingten Wahrscheinlichkeit ergibt sich

$$\begin{aligned} P[A] &= P[A \cap C] + P[A \cap \overline{C}] \\ &= P[A|C]\,P[C] + P[A|\overline{C}]\,P[\overline{C}] \end{aligned}$$

und wegen $P[C] + P[\overline{C}] = 1$ folgt daraus die Äquivalenz von (a) und (b). Die Äquivalenz von (b) und (c) ist klar. □

Die Äquivalenz der Aussagen (a) und (c) von Lemma 11.1.1 deutet bereits eine gewisse Symmetrie zwischen den bei der Unabhängigkeit betrachteten Ereignissen an; diese Symmetrie lässt sich wie folgt präzisieren:

11.1.2 Folgerung. *Seien $A, C \in \mathcal{F}$ Ereignisse mit $P[A], P[C] \in (0,1)$. Dann sind äquivalent:*
(a) *A ist unabhängig von C.*
(b) *C ist unabhängig von A.*

Folgerung 11.1.2 legt es nahe, die Unabhängigkeit von zwei Ereignissen in symmetrischer Weise und ohne Einschränkung an die Wahrscheinlichkeiten dieser Ereignisse wie folgt zu definieren. Aufgrund der Äquivalenz von (a) und (c) im Lemma 11.1.1 ist die folgende Definition der Unabhängigkeit mit der vorher gegebenen Definition verträglich:

Zwei Ereignisse $A, B \in \mathcal{F}$ heißen *unabhängig* (voneinander), wenn

$$P[A \cap B] = P[A]\,P[B]$$

gilt. In diesem Fall sagen wir auch:

– *A und B sind unabhängig.*
– *A ist unabhängig von B.*
– *B ist unabhängig von A.*
– *$\{A, B\}$ ist unabhängig.*

Im folgenden verstehen wir die Unabhängigkeit von Ereignissen $A, B \in \mathcal{F}$ als eine Eigenschaft der Familie $\{A, B\}$ dieser Ereignisse. Zwei Ereignisse heißen *abhängig* (voneinander), wenn sie nicht unabhängig sind.

11.1.3 Beispiele (Urnenmodelle). Wir betrachten eine Urne mit $N \geq 2$ Kugeln, von denen $K \in \{1, \ldots, N{-}1\}$ Kugeln rot sind und $N{-}K$ Kugeln eine beliebige andere Farbe besitzen. Wir nehmen an, dass alle Kugeln bis auf die Farbe gleichartig sind. Wir ziehen $n = 2$ Kugeln aus der Urne.

(1) **Ziehen ohne Zurücklegen:** Wir wählen den vorher eingeführten symmetrischen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und betrachten die Ereignisse

$$\begin{aligned} A &\mathrel{\widehat{=}} \text{bei der 1. Ziehung tritt eine rote Kugel auf} \\ B &\mathrel{\widehat{=}} \text{bei der 2. Ziehung tritt eine rote Kugel auf} \end{aligned}$$

Dann gilt

$$\begin{aligned} |\Omega| &= N(N-1) \\ |A \cap B| &= K(K-1) \\ |A \cap \overline{B}| &= K(N-K) \\ |\overline{A} \cap B| &= (N-K)K \end{aligned}$$

und damit

$$\begin{aligned} |A| &= K(N-1) \\ |B| &= K(N-1) \end{aligned}$$

Es gilt also

$$\begin{aligned} P[A \cap B] &= K(K-1)/(N(N-1)) \\ P[A] &= K/N \\ P[B] &= K/N \end{aligned}$$

und damit $P[A \cap B] \neq P[A]\, P[B]$. Daher ist $\{A, B\}$ nicht unabhängig.

(2) **Ziehen mit Zurücklegen:** Wir wählen den vorher eingeführten symmetrischen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und betrachten die Ereignisse

$$\begin{aligned} A &\mathrel{\widehat{=}} \text{bei der 1. Ziehung tritt eine rote Kugel auf} \\ B &\mathrel{\widehat{=}} \text{bei der 2. Ziehung tritt eine rote Kugel auf} \end{aligned}$$

Dann gilt

$$\begin{aligned} |\Omega| &= N^2 \\ |A \cap B| &= K^2 \\ |A \cap \overline{B}| &= K(N-K) \\ |\overline{A} \cap B| &= (N-K)K \end{aligned}$$

und damit

$$\begin{aligned} |A| &= KN \\ |B| &= KN \end{aligned}$$

Es gilt also

$$\begin{aligned} P[A \cap B] &= K^2/N^2 \\ P[A] &= K/N \\ P[B] &= K/N \end{aligned}$$

und damit $P[A \cap B] = P[A]\, P[B]$. Daher ist $\{A, B\}$ unabhängig.

Der Begriff der Unabhängigkeit von Ereignissen ist ein wahrscheinlichkeitstheoretischer Begriff, der in der allgemeinen Maßtheorie keinen Sinn macht: Sei $\mu : \mathcal{F} \to [0, \infty]$ ein Maß und seien $A, B \in \mathcal{F}$ Ereignisse mit

$$\mu[A \cap B] = \mu[A]\, \mu[B]$$

Dann gilt wegen $\mu[A \cap B] \leq \mu[A]$ und $\mu[A \cap B] \leq \mu[B]$

$$\begin{aligned} (\mu[A \cap B])^2 &\leq \mu[A]\, \mu[B] \\ &= \mu[A \cap B] \end{aligned}$$

und damit entweder $\mu[A \cap B] = \infty$ oder $\mu[A \cap B] \leq 1$. Im Fall $\mu[A \cap B] = \infty$ ist die Gleichung $\mu[A \cap B] = \mu[A]\, \mu[B]$ aber trivial, und im Fall $\mu[A \cap B] \leq 1$ folgt aus ihr $\min\{\mu[A], \mu[B]\} \leq 1$.

Wir untersuchen zunächst die Frage, welche Ereignisse von einem gegebenen Ereignis unabhängig sind.

11.1.4 Lemma. *Sei $A \in \mathcal{F}$ ein Ereignis und sei $B \in \mathcal{F}$ ein Ereignis mit $P[B] \in \{0, 1\}$. Dann ist $\{A, B\}$ unabhängig.*

Beweis. Im Fall $P[B] = 0$ gilt $P[A \cap B] = 0$ und damit $P[A \cap B] = P[A]\, P[B]$. Im Fall $P[B] = 1$ gilt $P[A \cup B] = 1$ und aus der Gleichung

$$P[A \cup B] + P[A \cap B] = P[A] + P[B]$$

folgt $P[A \cap B] = P[A] = P[A]\, P[B]$. □

Insbesondere ist das sichere Ereignis und das unmögliche Ereignis von jedem Ereignis unabhängig. Andererseits sind ein Ereignis und sein Komplement im allgemeinen abhängig; vgl. Aufgabe 11.1.F.

11.1.5 Lemma. *Sei $A \in \mathcal{F}$ ein Ereignis und sei $B \in \mathcal{F}$ ein Ereignis derart, dass $\{A, B\}$ unabhängig ist. Dann ist $\{A, \overline{B}\}$ unabhängig.*

Beweis. Es gilt

$$\begin{aligned} P[A \cap B] + P[A \cap \overline{B}] &= P[A] \\ &= P[A]\,(P[B] + P[\overline{B}]) \\ &= P[A]\,P[B] + P[A]\,P[\overline{B}] \end{aligned}$$

Daraus folgt die Behauptung. □

11.1.6 Lemma. *Sei $A \in \mathcal{F}$ ein Ereignis und sei $\{B_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ eine Folge von Ereignissen derart, dass $\{A, B_n\}$ für alle $n \in \mathbb{N}$ unabhängig ist. Dann gilt:*
(1) *Ist $\{B_n\}_{n\in\mathbb{N}}$ disjunkt, so ist $\{A, \sum_{n=1}^\infty B_n\}$ unabhängig.*
(2) *Ist $\{B_n\}_{n\in\mathbb{N}}$ monoton wachsend, so ist $\{A, \bigcup_{n\in\mathbb{N}} B_n\}$ unabhängig.*
(3) *Ist $\{B_n\}_{n\in\mathbb{N}}$ monoton fallend, so ist $\{A, \bigcap_{n\in\mathbb{N}} B_n\}$ unabhängig.*

Beweis. Ist die Folge $\{B_n\}_{n\in\mathbb{N}}$ disjunkt, so folgt aus der σ–Additivität

$$\begin{aligned} P\left[A \cap \sum_{n=1}^\infty B_n\right] &= P\left[\sum_{n=1}^\infty A \cap B_n\right] \\ &= \sum_{n=1}^\infty P[A \cap B_n] \\ &= \sum_{n=1}^\infty P[A]\,P[B_n] \\ &= P[A] \sum_{n=1}^\infty P[B_n] \\ &= P[A]\,P\left[\sum_{n=1}^\infty B_n\right] \end{aligned}$$

Daher gilt (1). Die Beweise von (2) und (3) verlaufen analog zu dem von (1) und verwenden anstelle der σ–Additivität die Stetigkeit von unten bzw. die Stetigkeit von oben. □

Das folgende Beispiel zeigt, dass die Voraussetzung der Monotonie in den Aussagen (2) und (3) von Lemma 11.1.6 wesentlich ist:

11.1.7 Beispiel. Sei $(\Omega, \mathcal{F}, P)$ der symmetrische Wahrscheinlichkeitsraum mit $\Omega := \{1, 2, 3, 4\}$. Sei ferner

$$\begin{aligned} A &:= \{1, 2\} \\ B &:= \{2, 3\} \\ C &:= \{1, 3\} \end{aligned}$$

Dann ist sowohl $\{A, B\}$ als auch $\{A, C\}$ unabhängig, aber es ist weder $\{A, B \cup C\}$ noch $\{A, B \cap C\}$ unabhängig.

Der folgende Satz fasst die wichtigsten Aussagen der letzten drei Lemmata zusammen:

11.1.8 Satz. *Sei $A \in \mathcal{F}$ ein Ereignis und sei*

$$\mathcal{D}_A := \Big\{ B \in \mathcal{F} \,\Big|\, \{A, B\} \text{ ist unabhängig} \Big\}$$

Dann ist $\mathcal{D}_A$ ein Dynkin–System.

Beweis. Der Beweis ergibt sich aus den bereits bekannten Ergebnissen:
(i) Wegen $P[\Omega] = 1$ und Lemma 11.1.4 gilt $\Omega \in \mathcal{D}_A$.
(ii) Sei $B \in \mathcal{D}_A$. Nach Lemma 11.1.5 gilt dann $\overline{B} \in \mathcal{D}_A$.
(iii) Sei $\{B_n\}_{n \in \mathbb{N}} \subseteq \mathcal{D}_A$ disjunkt. Nach Lemma 11.1.6 gilt dann $\sum_{n=1}^{\infty} B_n \in \mathcal{D}_A$.
Daher ist $\mathcal{D}_A$ ein Dynkin–System. □

Das Dynkin–System $\mathcal{D}_A := \{B \in \mathcal{F} \mid \{A, B\} \text{ ist unabhängig}\}$ der von $A \in \mathcal{F}$ unabhängigen Ereignisse ist im allgemeinen nicht $\cap$–stabil und damit im allgemeinen keine σ–Algebra; vgl. Beispiel 11.1.7.

Wir verallgemeinern die Definition der Unabhängigkeit von Ereignissen nun auf beliebige Familien von Ereignissen:

Eine Familie $\{A_i\}_{i \in I}$ von Ereignissen $A_i \in \mathcal{F}$ heißt *unabhängig*, wenn für alle $J \in \mathcal{H}(I)$

$$P\left[\bigcap_{i \in J} A_i\right] = \prod_{i \in J} P[A_i]$$

gilt. Wir veranschaulichen diese Definition an zwei Beispielen:

11.1.9 Beispiele.
(1) Sei $(\Omega, \mathcal{F}, P)$ der symmetrische Wahrscheinlichkeitsraum mit $\Omega := \{0, 1\}^3$. Sei ferner

$$\begin{aligned} A &:= \{000, 001, 010, 011\} \\ B &:= \{000, 001, 010, 100\} \\ C &:= \{000, 011, 100, 111\} \end{aligned}$$

Dann gilt

$$P[A \cap B \cap C] = P[A]\, P[B]\, P[C]$$

und

$$\begin{aligned} P[A \cap B] &\neq P[A]\, P[B] \\ P[A \cap C] &= P[A]\, P[C] \\ P[B \cap C] &= P[B]\, P[C] \end{aligned}$$

Daher ist die Familie $\{A, B, C\}$ nicht unabhängig.

(2) Sei $(\Omega, \mathcal{F}, P)$ der symmetrische Wahrscheinlichkeitsraum mit $\Omega := \{0,1\}^2$. Sei ferner

$$\begin{aligned} A &:= \{00, 01\} \\ B &:= \{00, 10\} \\ C &:= \{00, 11\} \end{aligned}$$

Dann gilt

$$\begin{aligned} P[A \cap B] &= P[A]\, P[B] \\ P[A \cap C] &= P[A]\, P[C] \\ P[B \cap C] &= P[B]\, P[C] \end{aligned}$$

und

$$P[A \cap B \cap C] \neq P[A]\, P[B]\, P[C]$$

Daher ist die Familie $\{A, B, C\}$ nicht unabhängig.

Der folgende Satz zeigt, dass jede Teilfamilie einer unabhängigen Familie von Ereignissen wieder unabhängig ist und dass umgekehrt die Unabhängigkeit jeder endlichen Teilfamilie einer Familie von Ereignissen die Unabhängigkeit der gesamten Familie impliziert:

11.1.10 Satz. *Sei $\{A_i\}_{i \in I}$ eine Familie von Ereignissen. Dann sind äquivalent*:
(a) *Die Familie $\{A_i\}_{i \in I}$ ist unabhängig.*
(b) *Für jede nichtleere Menge $K \subseteq I$ ist $\{A_i\}_{i \in K}$ unabhängig.*
(c) *Für jede endliche nichtleere Menge $K \subseteq I$ ist $\{A_i\}_{i \in K}$ unabhängig.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Sei $K \subseteq I$ nichtleer. Für alle $J \in \mathcal{H}(K)$ gilt $J \in \mathcal{H}(I)$ und damit

$$P\left[\bigcap_{i \in J} A_i\right] = \prod_{i \in J} P[A_i]$$

Daher folgt (b) aus (a).
Offensichtlich folgt (c) aus (b).
Wir nehmen nun an, dass (c) gilt. Für alle $J \in \mathcal{H}(I)$ gilt $J \in \mathcal{H}(J)$ und damit

$$P\left[\bigcap_{i \in J} A_i\right] = \prod_{i \in J} P[A_i]$$

Daher folgt (a) aus (c). □

Der folgende Satz verallgemeinert Lemma 11.1.5; er zeigt, dass für eine Familie von Ereignissen die Unabhängigkeit erhalten bleibt, wenn man beliebig viele der Ereignisse durch ihr Komplement ersetzt:

11.1.11 Satz. *Sei* $\{A_i\}_{i \in I}$ *eine Familie von Ereignissen und sei* $\{B_i\}_{i \in I}$ *eine Familie von Ereignissen mit* $B_i \in \{A_i, \overline{A_i}\}$ *für alle* $i \in I$*. Dann sind äquivalent*:
(a) *Die Familie* $\{A_i\}_{i \in I}$ *ist unabhängig.*
(b) *Die Familie* $\{B_i\}_{i \in I}$ *ist unabhängig.*

Beweis. Offenbar genügt es, eine der beiden Implikationen zu beweisen, und wegen Satz 11.1.10 genügt es ferner, den Beweis für den Fall $I = \{1, \dots, n\}$ zu führen.
Sei also die Familie $\{A_i\}_{i \in \{1,\dots,n\}}$ unabhängig. Wir zeigen durch vollständige Induktion, dass für alle $k \in \{0, 1, \dots, n\}$ die Aussage

(k) *Für alle* $J \subseteq \{1, \dots, n\}$ *mit* $|\{i \in J \mid B_i = \overline{A_i}\}| = k$ *gilt*

$$P\left[\bigcap_{i \in J} B_i\right] = \prod_{i \in J} P[B_i]$$

gilt:

- $k = 0$: In diesem Fall ist nichts zu zeigen.
- $k \to k+1$: Wir nehmen an, die Aussage sei für ein $k \in \{0, 1, \dots, n-1\}$ bereits bewiesen, und betrachten $J \subseteq \{1, \dots, n\}$ mit $|\{i \in J \mid B_i = \overline{A_i}\}| = k + 1$. Für alle $j \in J$ mit $B_j = \overline{A_j}$ gilt aufgrund der Induktionsannahme

$$P\left[\bigcap_{i \in J \setminus \{j\}} B_i\right] = \prod_{i \in J \setminus \{j\}} P[B_i]$$

und

$$P\left[A_j \cap \bigcap_{i \in J \setminus \{j\}} B_i\right] = P[A_j] \prod_{i \in J \setminus \{j\}} P[B_i]$$

und damit

$$P\left[A_j \cap \bigcap_{i \in J \setminus \{j\}} B_i\right] = P[A_j]\, P\left[\bigcap_{i \in J \setminus \{j\}} B_i\right]$$

Also ist $\{A_j, \bigcap_{i \in J \setminus \{j\}} B_i\}$ unabhängig. Nach Lemma 11.1.5 ist dann auch $\{\overline{A_j}, \bigcap_{i \in J \setminus \{j\}} B_i\}$ unabhängig. Wegen $B_j = \overline{A_j}$ gilt also

$$\begin{aligned} P\left[\bigcap_{i \in J} B_i\right] &= P\left[B_j \cap \bigcap_{i \in J \setminus \{j\}} B_i\right] \\ &= P\left[\overline{A_j} \cap \bigcap_{i \in J \setminus \{j\}} B_i\right] \end{aligned}$$

$$\begin{aligned}
&= P[\overline{A_j}]\, P\left[\bigcap_{i \in J \setminus \{j\}} B_i\right] \\
&= P[B_j]\, P\left[\bigcap_{i \in J \setminus \{j\}} B_i\right] \\
&= P[B_j] \prod_{i \in J \setminus \{j\}} P[B_i] \\
&= \prod_{i \in J} P[B_i]
\end{aligned}$$

Daher gilt die Aussage (k) für alle $k \in \{0, 1, \dots, n\}$, und damit ist die Familie $\{B_i\}_{i \in \{1,\dots,n\}}$ unabhängig. □

Wir untersuchen abschließend für eine Folge von Ereignissen die Bedeutung der Unabhängigkeit der Folge für die Wahrscheinlichkeit ihres Limes superior:

11.1.12 Lemma (1. Lemma von Borel/Cantelli). *Sei $\{A_n\}_{n \in \mathbb{N}}$ eine Folge von Ereignissen mit $\sum_{n=1}^\infty P[A_n] < \infty$. Dann gilt*

$$P\left[\limsup_{n \to \infty} A_n\right] = 0$$

Beweis. Da P monoton und σ–subadditiv ist, gilt für alle $m \in \mathbb{N}$

$$P\left[\limsup_{n \to \infty} A_n\right] = P\left[\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k\right] \le P\left[\bigcup_{k=m}^\infty A_k\right] \le \sum_{k=m}^\infty P[A_k]$$

Aus der Voraussetzung $\sum_{k=1}^\infty P[A_k] < \infty$ folgt nun die Behauptung. □

Für eine Folge von Ereignissen $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ mit $\sum_{n=1}^\infty P[A_n] = \infty$ erhält man im Fall der Unabhängigkeit der Folge $\{A_n\}_{n \in \mathbb{N}}$ ein ähnliches Ergebnis:

11.1.13 Lemma (2. Lemma von Borel/Cantelli). *Sei $\{A_n\}_{n \in \mathbb{N}}$ eine unabhängige Folge von Ereignissen mit $\sum_{n=1}^\infty P[A_n] = \infty$. Dann gilt*

$$P\left[\limsup_{n \to \infty} A_n\right] = 1$$

Beweis. Es gilt

$$1 - P\left[\limsup_{n \to \infty} A_n\right] = P\left[\Omega \setminus \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k\right] = P\left[\bigcup_{n=1}^\infty \bigcap_{k=n}^\infty \overline{A_k}\right] \le \sum_{n=1}^\infty P\left[\bigcap_{k=n}^\infty \overline{A_k}\right]$$

Für alle $n \in \mathbb{N}$ und für alle $m \in \mathbb{N}(n)$ gilt aufgrund der Unabhängigkeit der Folge $\{A_n\}_{n\in\mathbb{N}}$ und wegen Satz 11.1.11

$$\begin{aligned} P\left[\bigcap_{k=n}^{\infty} \overline{A_k}\right] &\le P\left[\bigcap_{k=n}^{m} \overline{A_k}\right] \\ &= \prod_{k=n}^{m} P[\overline{A_k}] \\ &= \prod_{k=n}^{m} \Bigl(1 - P[A_k]\Bigr) \\ &\le \prod_{k=n}^{m} \exp\Bigl(-P[A_k]\Bigr) \\ &= \exp\left(-\sum_{k=n}^{m} P[A_k]\right) \end{aligned}$$

Aus der Voraussetzung $\sum_{k=1}^{\infty} P[A_k] = \infty$ folgt nun die Behauptung. □

Aus dem 1. und 2. Lemma von Borel/Cantelli ergibt sich für unabhängige Folgen von Ereignissen das folgende *Null–Eins–Gesetz von Borel*:

11.1.14 Satz (Null–Eins–Gesetz; Borel). *Sei $\{A_n\}_{n\in\mathbb{N}}$ eine unabhängige Folge von Ereignissen. Dann gilt*

$$P\left[\limsup_{n\to\infty} A_n\right] = \begin{cases} 0 & \textit{falls } \sum_{n=1}^{\infty} P[A_n] < \infty \\ 1 & \textit{falls } \sum_{n=1}^{\infty} P[A_n] = \infty \end{cases}$$

Allgemein ist ein *Null–Eins–Gesetz* ein Satz, der besagt, dass unter gewissen Voraussetzungen für bestimmte Ereignisse $A \in \mathcal{F}$ entweder $P[A] = 0$ oder $P[A] = 1$ gilt.

Aufgaben

11.1.A Urnenmodelle: Betrachten Sie die Urnenmodelle aus Beispiel 11.1.3.
- Bestimmen Sie für jedes der Urnenmodelle die Wahrscheinlichkeiten $P[A|B]$ und $P[A|\overline{B}]$ und vergleichen Sie die Ergebnisse.
- Bestimmen Sie für jedes der Urnenmodelle die Wahrscheinlichkeiten $P[B|A]$ und $P[B|\overline{A}]$ und vergleichen Sie die Ergebnisse.

Warum ist die Wahrscheinlichkeit $P[B]$ nicht davon abhängig, ob mit oder ohne Zurücklegen gezogen wird?

11.1.B Formel von der totalen Wahrscheinlichkeit: Sei $C \in \mathcal{F}$ ein Ereignis mit $P[C] \in (0,1)$. Dann gilt für alle $A \in \mathcal{F}$

$$P[A] = P[A|C]\,P[C] + P[A|\overline{C}]\,P[\overline{C}]$$

Interpretieren Sie diese Gleichung.

11.1.C **Formel von Bayes:** Sei $C \in \mathcal{F}$ ein Ereignis mit $P[C] \in (0,1)$. Dann gilt für alle $A \in \mathcal{F}$ mit $P[A] > 0$

$$P[C|A] = \frac{P[A|C]\,P[C]}{P[A|C]\,P[C] + P[A|\overline{C}]\,P[\overline{C}]}$$

11.1.D Sei $C \in \mathcal{F}$ ein Ereignis mit $P[C] \in (0,1)$. Dann sind die Wahrscheinlichkeitsmaße $P[\,.\,|C]$ und $P[\,.\,|\overline{C}]$ P–stetig und singulär zueinander.

11.1.E **Lebensversicherungsmathematik:** Die Lebensdauer einer Person wird durch eine Zufallsvariable T mit $P[\{T \in \mathbb{R}_+\}] = 1$ und $P[\{T > x\}] > 0$ für alle $x \in \mathbb{N}_0$ beschrieben. Für $x, k \in \mathbb{N}_0$ sei

$${}_k p_x := P[\{T > x+k\}|\{T > x\}]$$

Dann gilt für alle $x, k, l \in \mathbb{N}_0$

$${}_{k+l}p_x = {}_l p_{x+k}\; {}_k p_x$$

und

$${}_k p_x = \prod_{j=1}^{k} {}_1 p_{x+j-1}$$

Interpretieren Sie die bedingten Wahrscheinlichkeiten ${}_k p_x$ und die beiden Gleichungen.

11.1.F Geben Sie eine hinreichende und notwendige Bedingung dafür an, dass zwei disjunkte Ereignisse unabhängig sind.

11.1.G Seien $A, B \in \mathcal{F}$ Ereignisse mit $P[A \cap B] = 1$. Dann ist $\{A, B\}$ unabhängig.

11.1.H Sei $\{A, B\} \subseteq \mathcal{F}$ unabhängig und sei $C \in \mathcal{F}$ ein Ereignis mit $C \subseteq B$. Dann sind äquivalent:
(a) $\{A, C\}$ ist unabhängig.
(b) $\{A, B \setminus C\}$ ist unabhängig.

11.1.I Führen Sie die Beweise der Aussagen (2) und (3) von Lemma 11.1.6 aus.

11.1.J Sei $\{A_n\}_{n\in\mathbb{N}}$ eine unabhängige Folge von Ereignissen mit $P[A_n] = c$ für alle $n \in \mathbb{N}$ und ein $c \in (0,1]$. Dann gilt $P[\limsup_{n\to\infty} A_n] = 1$.

11.1.K Konstruieren Sie einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und eine Folge von Ereignissen $\{A_n\}_{n\in\mathbb{N}}$ mit $\sum_{n=1}^{\infty} P[A_n] = \infty$ und $P[\limsup_{n\to\infty} A_n] < 1$.

11.1.L **Paarweise unabhängige Familien von Ereignissen:** Eine Familie von Ereignissen $\{A_i\}_{i\in I}$ heißt *paarweise unabhängig*, wenn für alle $i, j \in I$ mit $i \neq j$ die Familie $\{A_i, A_j\}$ unabhängig ist. Jede unabhängige Familie von Ereignissen ist paarweise unabhängig, aber eine paarweise unabhängige Familie von Ereignissen ist nicht notwendigerweise unabhängig.

11.2 Unabhängige Familien von Ereignissystemen

Wir übertragen nun den Begriff der Unabhängigkeit auf Familien von Ereignissystemen.

Wir beginnen wieder mit einer Vorbetrachtung: Sind $A, B \in \mathcal{F}$ Ereignisse (die beide von $\emptyset$ und Ω verschieden sind), so ist nach Lemma 11.1.4 und Lemma 11.1.5 die Familie $\{A, B\}$ genau dann unabhängig, wenn für jede Wahl von $C \in \{\emptyset, A, \overline{A}, \Omega\}$ und $D \in \{\emptyset, B, \overline{B}, \Omega\}$ die Familie $\{C, D\}$ unabhängig ist. Offensichtlich sind die Ereignissysteme $\{\emptyset, A, \overline{A}, \Omega\}$ und $\{\emptyset, B, \overline{B}, \Omega\}$ gerade die von den Ereignissen A bzw. B erzeugten σ–Algebren $\sigma(A)$ bzw. $\sigma(B)$.

Diese Beobachtung legt es nahe, den Begriff der Unabhängigkeit auf Familien von Ereignissystemen zu übertragen.

Eine Familie $\{\mathcal{E}_i\}_{i\in I}$ von Ereignissystemen $\mathcal{E}_i \subseteq \mathcal{F}$ heißt *unabhängig*, wenn jede Familie von Ereignissen $\{A_i\}_{i\in I}$ mit $A_i \in \mathcal{E}_i$ für alle $i \in I$ unabhängig ist.

Der folgende Satz ergibt sich unmittelbar aus der Definition der Unabhängigkeit einer Familie von Ereignissystemen und Satz 11.1.10:

11.2.1 Satz. *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine Familie von Ereignissystemen. Dann sind äquivalent:*
(a) *Die Familie $\{\mathcal{E}_i\}_{i\in I}$ ist unabhängig.*
(b) *Für jede nichtleere Menge $K \subseteq I$ ist $\{\mathcal{E}_i\}_{i\in K}$ unabhängig.*
(c) *Für jede endliche nichtleere Menge $K \subseteq I$ ist $\{\mathcal{E}_i\}_{i\in K}$ unabhängig.*

Auch das folgende Lemma ergibt sich unmittelbar aus der Definition der Unabhängigkeit einer Familie von Ereignissystemen:

11.2.2 Lemma. *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine Familie von Ereignissystemen und sei $\{\mathcal{C}_i\}_{i\in I}$ eine Familie von Ereignissystemen mit $\mathcal{C}_i \subseteq \mathcal{E}_i$ für alle $i \in I$. Ist $\{\mathcal{E}_i\}_{i\in I}$ unabhängig, so ist auch $\{\mathcal{C}_i\}_{i\in I}$ unabhängig.*

Es stellt sich nun die Frage, unter welchen Umständen man umgekehrt von der Unabhängigkeit einer Familie von Ereignissystemen auf die Unabhängigkeit einer Familie von größeren Ereignissystemen schließen kann; von besonderem Interesse ist natürlich die Familie der von den einzelnen Ereignissystemen erzeugten σ–Algebren. Wir benötigen die folgenden Lemmata:

11.2.3 Lemma. *Seien $\mathcal{E}_1$ und $\mathcal{E}_2$ Ereignissysteme derart, dass $\mathcal{E}_1$ $\cap$–stabil und $\{\mathcal{E}_1, \mathcal{E}_2\}$ unabhängig ist. Dann ist $\{\sigma(\mathcal{E}_1), \mathcal{E}_2\}$ unabhängig.*

Beweis. Für jedes Ereignis $A \in \mathcal{E}_2$ ist nach Satz 11.1.8 das Mengensystem

$$\mathcal{D}_A := \Big\{ B \in \mathcal{F} \,\Big|\, \{A, B\} \text{ ist unabhängig} \Big\}$$

ein Dynkin–System mit $\mathcal{E}_1 \subseteq \mathcal{D}_A$. Dann ist auch

$$\mathcal{D} := \bigcap_{A\in\mathcal{E}_2} \mathcal{D}_A$$

ein Dynkin–System mit $\mathcal{E}_1 \subseteq \mathcal{D}$. Da $\mathcal{E}_1$ $\cap$–stabil ist, gilt $\sigma(\mathcal{E}_1) = \delta(\mathcal{E}_1) \subseteq \mathcal{D}$. Nach Definition von $\mathcal{D}$ ist $\{\mathcal{D}, \mathcal{E}_2\}$ unabhängig und aus Lemma 11.2.2 folgt nun, dass auch $\{\sigma(\mathcal{E}_1), \mathcal{E}_2\}$ unabhängig ist. □

11.2.4 Lemma. *Sei $\{\mathcal{C}_i\}_{i \in I}$ eine endliche Familie von Ereignissystemen. Sei ferner $k \in I$ und*

$$\mathcal{D}_k := \left\{ D \in \mathcal{F} \,\middle|\, D = \bigcap_{i \in J} C_i \text{ mit } J \in \mathcal{H}(I \setminus \{k\}) \text{ und } C_i \in \mathcal{C}_i \text{ für alle } i \in J \right\}$$

Dann sind äquivalent:
(a) $\{\mathcal{C}_i\}_{i \in I}$ *ist unabhängig.*
(b) $\{\mathcal{C}_i\}_{i \in I \setminus \{k\}}$ *ist unabhängig und* $\{\mathcal{C}_k, \mathcal{D}_k\}$ *ist unabhängig.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Mit $\{\mathcal{C}_i\}_{i \in I}$ ist nach Satz 11.2.1 auch $\{\mathcal{C}_i\}_{i \in I \setminus \{k\}}$ unabhängig. Für $C_k \in \mathcal{C}_k$ und $D \in \mathcal{D}_k$ gibt es eine Darstellung $D = \bigcap_{i \in J} C_i$ mit $J \in \mathcal{H}(I \setminus \{k\})$ und $C_i \in \mathcal{C}_i$ für alle $i \in J$ und es gilt

$$\begin{aligned}
P[C_k \cap D] &= P\left[C_k \cap \bigcap_{i \in J} C_i \right] \\
&= P\left[\bigcap_{i \in J \cup \{k\}} C_i \right] \\
&= \prod_{i \in J \cup \{k\}} P[C_i] \\
&= P[C_k] \prod_{i \in J} P[C_i] \\
&= P[C_k]\, P\left[\bigcap_{i \in J} C_i \right] \\
&= P[C_k]\, P[D]
\end{aligned}$$

Damit ist auch $\{\mathcal{C}_k, \mathcal{D}_k\}$ unabhängig. Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Sei $\{C_i\}_{i \in I}$ eine Familie von Ereignissen mit $C_i \in \mathcal{C}_i$ für alle $i \in I$ und sei $J \in \mathcal{H}(I)$. Im Fall $k \in J$ gilt

$$\begin{aligned}
P\left[\bigcap_{i \in J} C_i \right] &= P\left[C_k \cap \bigcap_{i \in J \setminus \{k\}} C_i \right] \\
&= P[C_k]\, P\left[\bigcap_{i \in J \setminus \{k\}} C_i \right] \\
&= P[C_k] \prod_{i \in J \setminus \{k\}} P[C_i] \\
&= \prod_{i \in J} P[C_i]
\end{aligned}$$

und im Fall $k \notin J$ gilt ebenfalls

$$P\left[\bigcap_{i\in J} C_i\right] = \prod_{i\in J} P[C_i]$$

Damit gilt für alle $J \in \mathcal{H}(I)$

$$P\left[\bigcap_{i\in J} C_i\right] = \prod_{i\in J} P[C_i]$$

Daher folgt (a) aus (b). □

Der folgende Satz klärt den Zusammenhang zwischen der Unabhängigkeit von σ–Algebren und der Unabhängigkeit ihrer Erzeuger:

11.2.5 Satz. *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine Familie von $\cap$–stabilen Ereignissystemen. Dann sind äquivalent*:
(a) *Die Familie $\{\mathcal{E}_i\}_{i\in I}$ ist unabhängig.*
(b) *Die Familie $\{\sigma(\mathcal{E}_i)\}_{i\in I}$ ist unabhängig.*

Beweis. Wegen Lemma 11.2.2 folgt (a) aus (b), und wegen Satz 11.2.1 genügt es, den Beweis der umgekehrten Implikation für $I = \{1, \ldots, n\}$ zu führen.
Sei also die Familie $\{\mathcal{E}_i\}_{i\in\{1,\ldots,n\}}$ unabhängig. Für $k \in \{1, \ldots, n\}$ sei

$$\mathcal{D}_k := \left\{ D\in\mathcal{F} \,\middle|\, D = \bigcap_{i\in J} C_i,\, J\in\mathcal{H}(\{1,\ldots,n\}\setminus\{k\}),\, \begin{array}{ll} C_i\in\sigma(\mathcal{E}_i) & \text{für } i\le k-1 \\ C_i\in\mathcal{E}_i & \text{für } i\ge k+1 \end{array} \right\}$$

Wir zeigen durch vollständige Induktion, dass für alle $k \in \{0, 1, \ldots, n\}$ die Aussage

(k) $\quad \{\sigma(\mathcal{E}_1), \ldots, \sigma(\mathcal{E}_k), \mathcal{E}_{k+1}, \ldots, \mathcal{E}_n\}$ *ist unabhängig.*

gilt:
- $k = 0$: In diesem Fall ist nichts zu zeigen.
- $k \to k + 1$: Wir nehmen an, die Aussage sei für ein $k \in \{0, 1, \ldots, n-1\}$ bereits bewiesen. Dann ist die Familie

  $$\{\sigma(\mathcal{E}_1), \ldots, \sigma(\mathcal{E}_k), \mathcal{E}_{k+1}, \mathcal{E}_{k+2}, \ldots, \mathcal{E}_n\}$$

  unabhängig. Nach Lemma 11.2.4 ist jede der Familien

  $$\{\sigma(\mathcal{E}_1), \ldots, \sigma(\mathcal{E}_k), \mathcal{E}_{k+2}, \ldots, \mathcal{E}_n\} \quad \text{und} \quad \{\mathcal{E}_{k+1}, \mathcal{D}_{k+1}\}$$

  unabhängig, nach Lemma 11.2.3 ist jede der Familien

  $$\{\sigma(\mathcal{E}_1), \ldots, \sigma(\mathcal{E}_k), \mathcal{E}_{k+2}, \ldots, \mathcal{E}_n\} \quad \text{und} \quad \{\sigma(\mathcal{E}_{k+1}), \mathcal{D}_{k+1}\}$$

  unabhängig, und nach Lemma 11.2.4 ist die Familie

  $$\{\sigma(\mathcal{E}_1), \ldots, \sigma(\mathcal{E}_k), \sigma(\mathcal{E}_{k+1}), \mathcal{E}_{k+2}, \ldots, \mathcal{E}_n\}$$

  unabhängig.

Daher gilt die Aussage (n), und damit ist die Familie $\{\sigma(\mathcal{E}_i)\}_{i\in I}$ unabhängig. □

Insbesondere ist die Unabhängigkeit einer Familie von Ereignissen äquivalent mit der Unabhängigkeit der Familie der von den einzelnen Ereignissen erzeugten σ–Algebren; dies ist im wesentlichen nur eine andere Formulierung von Satz 11.1.11 und schließt den Kreis zur Vorbetrachtung am Anfang dieses Abschnittes:

11.2.6 Folgerung. *Sei $\{A_i\}_{i\in I}$ eine Familie von Ereignissen. Dann sind äquivalent:*
(a) *Die Familie $\{A_i\}_{i\in I}$ ist unabhängig.*
(b) *Die Familie $\{\sigma(A_i)\}_{i\in I}$ ist unabhängig.*

Wir betrachten nun eine weitere Stabilitätseigenschaft der Unabhängigkeit einer Familie von $\cap$–stabilen Ereignissystemen. Das folgende Blocklemma wird sich als äußerst nützlich erweisen und lässt sich auf eine beliebige disjunkte Familie von Teilmengen der Indexmenge verallgemeinern; vgl. Aufgabe 11.2.C.

11.2.7 Lemma (Blocklemma). *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine unabhängige Familie von $\cap$–stabilen Ereignissystemen. Sind $M, N \subseteq I$ nichtleer und disjunkt, so ist auch die Familie*

$$\left\{\sigma\left(\bigcup_{i\in M}\mathcal{E}_i\right), \sigma\left(\bigcup_{i\in N}\mathcal{E}_i\right)\right\}$$

unabhängig.

Beweis. Für jede nichtleere Menge $L \subseteq I$ sei

$$\mathcal{E}_L := \left\{A\in\mathcal{F}\;\middle|\; A = \bigcap_{i\in J} A_i \text{ mit } J\in\mathcal{H}(L) \text{ und } A_i\in\mathcal{E}_i \text{ für alle } i\in J\right\}$$

Da jedes der Mengensysteme $\mathcal{E}_i$ $\cap$–stabil ist, ist auch $\mathcal{E}_L$ $\cap$–stabil. Außerdem gilt

$$\sigma(\mathcal{E}_L) = \sigma\left(\bigcup_{i\in L}\mathcal{E}_i\right)$$

Es genügt daher zu zeigen, dass $\{\mathcal{E}_M, \mathcal{E}_N\}$ unabhängig ist, denn dann folgt aus Satz 11.2.5, dass auch $\{\sigma(\mathcal{E}_M), \sigma(\mathcal{E}_N)\}$ unabhängig ist.
Wir betrachten daher $A \in \mathcal{E}_M$ und $B \in \mathcal{E}_N$ sowie Darstellungen

$$A = \bigcap_{i\in J} A_i$$
$$B = \bigcap_{i\in K} A_i$$

von A und B mit $J \in \mathcal{H}(M)$ und $K \in \mathcal{H}(N)$ sowie $A_i \in \mathcal{E}_i$ für alle $i \in M+N$. Dann gilt $J, K, J+K \in \mathcal{H}(I)$ und damit

$$\begin{aligned}
P[A \cap B] &= P\left[\left(\bigcap_{i\in J} A_i\right) \cap \left(\bigcap_{i\in K} A_i\right)\right]\\
&= P\left[\bigcap_{i\in J+K} A_i\right]\\
&= \prod_{i\in J+K} P[A_i]\\
&= \prod_{i\in J} P[A_i] \cdot \prod_{i\in K} P[A_i]\\
&= P\left[\bigcap_{i\in J} A_i\right] P\left[\bigcap_{i\in K} A_i\right]\\
&= P[A]\,P[B]
\end{aligned}$$

Daher ist $\{\mathcal{E}_M, \mathcal{E}_N\}$ unabhängig. □

Sei $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ eine Folge von Ereignissystemen. Dann heißt die σ–Algebra

$$\mathcal{E}_\infty := \bigcap_{m=1}^{\infty} \sigma\left(\bigcup_{n=m}^{\infty} \mathcal{E}_n\right)$$

die zu der Folge $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ gehörige *terminale σ–Algebra* und jedes Ereignis $A \in \mathcal{E}_\infty$ heißt *terminales Ereignis* der Folge $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$.

Es wird sich zeigen, dass die terminalen Ereignisse einer Folge von Ereignissystemen bei Konvergenzbetrachtungen von Interesse sind:

11.2.8 Beispiel. Sei $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ eine Folge von Ereignissystemen und sei $\mathcal{E}_\infty$ die zugehörige terminale σ–Algebra. Für jede Folge $\{A_n\}_{n\in\mathbb{N}}$ mit $A_n \in \mathcal{E}_n$ für alle $n \in \mathbb{N}$ gilt dann $\limsup_{n\to\infty} A_n \in \mathcal{E}_\infty$.
In der Tat: Für alle $k \in \mathbb{N}$ sei $B_k := \bigcup_{n=k}^{\infty} A_n$ und $\mathcal{D}_k := \sigma(\bigcup_{n=k}^{\infty} \mathcal{E}_n)$. Für alle $m \in \mathbb{N}$ und $k \in \mathbb{N}(m)$ gilt dann $B_k \in \mathcal{D}_k \subseteq \mathcal{D}_m$. Da die Folge $\{B_k\}_{k\in\mathbb{N}}$ monoton fallend ist, erhalten wir für alle $m \in \mathbb{N}$

$$\limsup_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n = \bigcap_{k=1}^{\infty} B_k = \bigcap_{k=m}^{\infty} B_k \in \mathcal{D}_m$$

und damit

$$\limsup_{n\to\infty} A_n \in \bigcap_{m=1}^{\infty} \mathcal{D}_m = \bigcap_{m=1}^{\infty} \sigma\left(\bigcup_{n=m}^{\infty} \mathcal{E}_n\right) = \mathcal{E}_\infty$$

Damit ist die Behauptung gezeigt.

Für die terminalen Ereignisse einer unabhängigen Folge von $\cap$–stabilen Ereignissystemen gilt das *Null–Eins–Gesetz von Kolmogorov*:

11.2.9 Satz (Null–Eins–Gesetz; Kolmogorov). *Sei $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ eine unabhängige Folge von $\cap$–stabilen Ereignissystemen und sei $\mathcal{E}_\infty$ die zugehörige terminale σ–Algebra. Dann gilt für alle $A, B \in \mathcal{E}_\infty$*

$$P[A\cap B] = P[A]\,P[B]$$

Insbesondere gilt für alle $A \in \mathcal{E}_\infty$ entweder $P[A] = 0$ oder $P[A] = 1$.

Beweis. Für $m \in \mathbb{N}$ betrachten wir die σ–Algebren

$$\mathcal{C}_m := \sigma\left(\bigcup_{n=1}^{m} \mathcal{E}_n\right)$$

$$\mathcal{D}_m := \sigma\left(\bigcup_{n=m+1}^{\infty} \mathcal{E}_n\right)$$

Nach dem Blocklemma ist

$$\{\mathcal{C}_m, \mathcal{D}_m\}$$

unabhängig, und wegen $\mathcal{E}_\infty \subseteq \mathcal{D}_m$ ist nach Lemma 11.2.2 auch

$$\{\mathcal{C}_m, \mathcal{E}_\infty\}$$

unabhängig.
Daraus folgt zunächst, dass sogar

$$\left\{\bigcup_{m=1}^{\infty} \mathcal{C}_m, \mathcal{E}_\infty\right\}$$

unabhängig ist. Da die Folge $\{\mathcal{C}_m\}_{m\in\mathbb{N}}$ monoton wachsend ist, ist das Mengensystem $\bigcup_{m=1}^{\infty} \mathcal{C}_m$ eine Algebra und insbesondere $\cap$–stabil, und nach Lemma 11.2.3 ist dann auch

$$\left\{\sigma\left(\bigcup_{m=1}^{\infty} \mathcal{C}_m\right), \mathcal{E}_\infty\right\}$$

unabhängig. Wegen $\mathcal{E}_m \subseteq \mathcal{C}_m$ gilt $\mathcal{E}_\infty \subseteq \sigma(\bigcup_{m=1}^{\infty} \mathcal{C}_m)$, und nach Lemma 11.2.2 ist dann auch

$$\{\mathcal{E}_\infty, \mathcal{E}_\infty\}$$

unabhängig. Daher gilt für alle $A, B \in \mathcal{E}_\infty$

$$P[A\cap B] = P[A]\,P[B]$$

und insbesondere $P[A] \in \{0, 1\}$. □

Das Null–Eins–Gesetz von Kolmogorov ist allgemeiner als das Null–Eins–Gesetz von Borel, denn es stellt für jedes terminale Ereignis $A \in \mathcal{E}_\infty$ einer unabhängigen Folge von $\cap$–stabilen Ereignissystemen $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ sicher, dass $P[A] \in \{0, 1\}$ gilt; dagegen betrifft das Null–Eins–Gesetz von Borel nur das spezielle terminale Ereignis $\limsup_{n\to\infty} A_n \in \mathcal{E}_\infty$ einer unabhängigen Folge von Ereignissen $\{A_n\}_{n\in\mathbb{N}}$, gibt aber zusätzlich hinreichende und notwendige Bedingungen für $P[\limsup_{n\to\infty} A_n] = 1$ und $P[\limsup_{n\to\infty} A_n] = 0$ an.

Aufgaben

11.2.A Sei $\{\mathcal{E}_k\}_{k\in\mathbb{N}}$ eine Folge von Ereignissystemen. Dann sind äquivalent:
(a) Die Folge $\{\mathcal{F}_k\}_{k\in\mathbb{N}}$ ist unabhängig.
(b) Für alle $n \in \mathbb{N}$ ist die Familie $\{\mathcal{F}_k\}_{k\in\{1,\ldots,n\}}$ unabhängig.

11.2.B Sei $\{\mathcal{F}_k\}_{k\in\{1,\ldots,n\}}$ eine endliche Familie von σ–Algebren mit $\mathcal{F}_k \subseteq \mathcal{F}$ für alle $k \in \{1,\ldots,n\}$. Dann sind äquivalent:
(a) Die Familie $\{\mathcal{F}_k\}_{k\in\{1,\ldots,n\}}$ ist unabhängig.
(b) Für jede Familie $\{A_k\}_{k\in\{1,\ldots,n\}}$ mit $A_k \in \mathcal{F}_k$ für alle $k \in \{1,\ldots,n\}$ gilt

$$P\left[\bigcap_{k=1}^n A_k\right] = \prod_{k=1}^n P[A_k]$$

11.2.C **Blocklemma:** Verallgemeinern Sie das Blocklemma auf eine beliebige disjunkte Familie von Teilmengen der Indexmenge.

11.2.D **Terminale σ–Algebra:** Sei $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ eine Folge von Ereignissystemen und sei $\mathcal{E}_\infty$ die zugehörige terminale σ–Algebra. Dann gilt

$$\mathcal{E}_\infty = \bigcap_{m=1}^\infty \sigma\left(\bigcup_{n=m}^\infty \sigma(\mathcal{E}_n)\right)$$

11.2.E Sei $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ eine Folge von Ereignissystemen und sei $\mathcal{E}_\infty$ die zugehörige terminale σ–Algebra. Für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ mit $A_n \in \mathcal{E}_n$ für alle $n \in \mathbb{N}$ gilt dann $\liminf_{n\to\infty} A_n \in \mathcal{E}_\infty$.

11.2.F **Null–Eins–Gesetz (Kolmogorov):** Verallgemeinern Sie die Definition der terminalen σ–Algebra und das Null–Eins–Gesetz von Kolmogorov in geeigneter Weise auf beliebige unendliche Familien von Ereignissystemen.

11.2.G Seien $\mathcal{D}$ und $\mathcal{E}$ Ereignissysteme mit $\mathcal{D} \subseteq \mathcal{E}$. Dann sind äquivalent:
(a) $\{\mathcal{D}, \mathcal{E}\}$ ist unabhängig.
(b) Für alle $D \in \mathcal{D}$ gilt $P[D] \in \{0, 1\}$.

11.2.H **Paarweise unabhängige Familien von Ereignissystemen:** Eine Familie von Ereignissystemen $\{\mathcal{E}_i\}_{i\in I}$ heißt *paarweise unabhängig*, wenn für alle $i, j \in I$ mit $i \neq j$ die Familie $\{\mathcal{E}_i, \mathcal{E}_j\}$ unabhängig ist. Jede unabhängige Familie von Ereignissystemen ist paarweise unabhängig, aber eine paarweise unabhängige Familie von Ereignissystemen ist nicht notwendigerweise unabhängig.

11.3 Unabhängige Familien von Zufallsgrößen

Schließlich übertragen wir den Begriff der Unabhängigkeit auf Familien von Zufallsgrößen.

Auch hier beginnen wir mit einer Vorbetrachtung: Ist $(\Omega', \mathcal{F}')$ ein Messraum und $X : \Omega \to \Omega'$ eine Zufallsgröße, so stimmt die von X erzeugte σ–Algebra $\sigma(X)$ mit dem Mengensystem $X^{-1}(\mathcal{F}')$ überein und aufgrund der Messbarkeit von X gilt $X^{-1}(\mathcal{F}') \subseteq \mathcal{F}$. Insbesondere ist $\sigma(X)$ ein Ereignissystem. Damit können die von Zufallsgrößen erzeugten σ–Algebren verwendet werden, um den Begriff der Unabhängigkeit von Familien von Ereignissystemen auf Familien von Zufallsgrößen zu übertragen.

Wir betrachten eine Familie $\{(\Omega'_i, \mathcal{F}'_i)\}_{i\in I}$ von Messräumen und eine Familie $\{X_i\}_{i\in I}$ von Zufallsgrößen $X_i : \Omega \to \Omega'_i$. Die Familie $\{X_i\}_{i\in I}$ heißt *unabhängig*, wenn die Familie $\{\sigma(X_i)\}_{i\in I}$ unabhängig ist.

11.3.1 Satz. *Sei $\{\mathcal{E}'_i\}_{i\in I}$ eine Familie von $\cap$–stabilen Mengensystemen mit $\sigma(\mathcal{E}'_i) = \mathcal{F}'_i$ für alle $i \in I$. Dann sind äquivalent*:
(a) *Die Familie $\{X_i\}_{i\in I}$ ist unabhängig.*
(b) *Die Familie $\{X_i^{-1}(\mathcal{F}'_i)\}_{i\in I}$ ist unabhängig.*
(c) *Die Familie $\{X_i^{-1}(\mathcal{E}'_i)\}_{i\in I}$ ist unabhängig.*
(d) *Für jede Familie $\{B'_i\}_{i\in I}$ mit $B'_i \in \mathcal{E}'_i$ für alle $i \in I$ ist die Familie $\{X_i^{-1}(B'_i)\}_{i\in I}$ unabhängig.*

Beweis. Die Äquivalenz von (a) und (b) ist klar. Mit $\mathcal{E}'_i$ ist auch $X_i^{-1}(\mathcal{E}'_i)$ $\cap$–stabil und nach Lemma 2.3.2 gilt $X_i^{-1}(\mathcal{F}'_i) = X_i^{-1}(\sigma(\mathcal{E}'_i)) = \sigma(X_i^{-1}(\mathcal{E}'_i))$. Die Äquivalenz von (b) und (c) ergibt sich nun aus Satz 11.2.5. Die Äquivalenz von (c) und (d) ist klar. □

Der folgende Satz ergibt sich unmittelbar aus der Definition der Unabhängigkeit einer Familie von Zufallsgrößen und Satz 11.2.1:

11.3.2 Satz. *Folgende Aussagen sind äquivalent*:
(a) *Die Familie $\{X_i\}_{i\in I}$ ist unabhängig.*
(b) *Für jede nichtleere Menge $K \subseteq I$ ist $\{X_i\}_{i\in K}$ unabhängig.*
(c) *Für jede endliche nichtleere Menge $K \subseteq I$ ist $\{X_i\}_{i\in K}$ unabhängig.*

Der folgende wichtige Satz zeigt, dass für eine Familie von Zufallsgrößen die Unabhängigkeit unter messbaren Transformationen erhalten bleibt:

11.3.3 Satz. *Sei $\{X_i\}_{i\in I}$ unabhängig. Ist $\{(\Omega''_i, \mathcal{F}''_i)\}_{i\in I}$ eine Familie von Messräumen und ist $\{T_i\}_{i\in I}$ eine Familie von messbaren Abbildungen $T_i : \Omega'_i \to \Omega''_i$, so ist auch die Familie $\{T_i \circ X_i\}_{i\in I}$ unabhängig.*

Beweis. Für alle $i \in I$ gilt

$$\sigma(T_i \circ X_i) = (T_i \circ X_i)^{-1}(\mathcal{F}''_i) = X_i^{-1}(T_i^{-1}(\mathcal{F}''_i)) \subseteq X_i^{-1}(\mathcal{F}'_i) = \sigma(X_i)$$

Aus Lemma 11.2.2 folgt nun, dass mit der Familie $\{X_i\}_{i\in I}$ auch die Familie $\{T_i \circ X_i\}_{i\in I}$ unabhängig ist. □

Wir betrachten nun das Produkt

$$(\Omega', \mathcal{F}') := \bigotimes_{i \in I} (\Omega'_i, \mathcal{F}'_i)$$

der Familie $\{(\Omega'_i, \mathcal{F}'_i)\}_{i \in I}$ und die Abbildung $X : \Omega \to \Omega'$ mit

$$\pi_i \circ X = X_i$$

für alle $i \in I$. Wir bezeichnen X_i als i–te Koordinate von X. Nach Satz 3.3.2 ist aufgrund der Messbarkeit aller Koordinaten auch X messbar.

Wir geben zunächst eine Variante des Blocklemmas an, die sich natürlich ebenfalls auf eine beliebige disjunkte Familie von Teilmengen der Indexmenge verallgemeinern lässt; vgl. Aufgabe 11.3.D.

11.3.4 Lemma (Blocklemma). *Sei $\{X_i\}_{i \in I}$ unabhängig. Sind $M, N \subseteq I$ nichtleer und disjunkt, so ist auch die Familie $\{\pi_M \circ X, \pi_N \circ X\}$ unabhängig.*

Beweis. Für jede nichtleere Menge $L \subseteq I$ gilt

$$\sigma(\pi_L \circ X) = \sigma(\{\pi_i \circ X\}_{i \in L}) = \sigma(\{X_i\}_{i \in L}) = \sigma(\{\sigma(X_i)\}_{i \in L})$$

Die Behauptung folgt nun aus Lemma 11.2.7. □

Der folgende Satz liefert eine wichtige Charakterisierung der Unabhängigkeit der Familie $\{X_i\}_{i \in I}$ im Bildbereich:

11.3.5 Satz. *Folgende Aussagen sind äquivalent:*
(a) *Die Familie $\{X_i\}_{i \in I}$ ist unabhängig.*
(b) *Es gilt $P_X = \bigotimes_{i \in I} P_{X_i}$.*

Beweis. Sei zunächst $J \in \mathcal{H}(I)$ und sei $\{B'_i\}_{i \in J}$ eine Familie von Ereignissen mit $B'_i \in \mathcal{F}'_i$ für alle $i \in J$. Dann gilt einerseits

$$\begin{aligned}
P\left[\bigcap_{i \in J} \{X_i \in B'_i\}\right] &= P\left[\bigcap_{i \in J} \{\pi_i \circ X \in B'_i\}\right] \\
&= P\left[\bigcap_{i \in J} X^{-1}(\pi_i^{-1}(B'_i))\right] \\
&= P\left[X^{-1}\left(\bigcap_{i \in J} \pi_i^{-1}(B'_i)\right)\right] \\
&= P_X\left[\pi_J^{-1}\left(\prod_{i \in J} B'_i\right)\right] \\
&= (P_X)_{\pi_J}\left[\prod_{i \in J} B'_i\right]
\end{aligned}$$

und andererseits

$$\begin{aligned}\prod_{i\in J} P[\{X_i \in B'_i\}] &= \prod_{i\in J} P_{X_i}[B'_i] \\ &= \left(\bigotimes_{i\in J} P_{X_i}\right)\left(\prod_{i\in J} B'_i\right)\end{aligned}$$

Die Familie $\{X_i\}_{i\in I}$ ist genau dann unabhängig, wenn für alle $J \in \mathcal{H}(I)$ und jede Familie $\{B'_i\}_{i\in J}$ mit $B'_i \in \mathcal{F}'_i$ für alle $i \in J$ die Gleichung

$$P\left[\bigcap_{i\in J}\{X_i \in B'_i\}\right] = \prod_{i\in J} P[\{X_i \in B'_i\}]$$

gilt. Nach dem vorher gezeigten ist diese Bedingung genau dann erfüllt, wenn für alle $J \in \mathcal{H}(I)$ und jede Familie $\{B'_i\}_{i\in J}$ mit $B'_i \in \mathcal{F}'_i$ für alle $i \in J$ die Gleichung

$$(P_X)_{\pi_J}\left[\prod_{i\in J} B'_i\right] = \left(\bigotimes_{i\in J} P_{X_i}\right)\left(\prod_{i\in J} B'_i\right)$$

gilt. Nach dem Eindeutigkeitssatz ist diese letzte Bedingung genau dann erfüllt, wenn für alle $J \in \mathcal{H}(I)$

$$(P_X)_{\pi_J} = \bigotimes_{i\in J} P_{X_i}$$

gilt, und nach dem Satz von Andersen/Jessen ist dies genau dann der Fall, wenn

$$P_X = \bigotimes_{i\in I} P_{X_i}$$

gilt. □

Wir illustrieren dieses Ergebnis an einem Beispiel:

11.3.6 Beispiele (Urnenmodelle). Wir betrachten eine Urne mit $N \geq 2$ Kugeln, von denen $K \in \{1, \ldots, N–1\}$ Kugeln rot sind und $N–K$ Kugeln eine beliebige andere Farbe besitzen. Wir nehmen an, dass alle Kugeln bis auf die Farbe gleichartig sind. Wir ziehen $n = 2$ Kugeln aus der Urne.

(1) **Ziehen ohne Zurücklegen:** Wir bezeichnen für $i \in \{1, 2\}$ mit X_i die Anzahl der roten Kugeln bei der i–ten Ziehung und setzen

$$X := \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

Dann gilt

$$\begin{aligned}
P\left[\left\{X = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\}\right] &= \frac{(N-K)\,(N-K-1)}{N\,(N-1)} \\
P\left[\left\{X = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}\right] &= \frac{(N-K)\,K}{N\,(N-1)} \\
P\left[\left\{X = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}\right] &= \frac{K\,(N-K)}{N\,(N-1)} \\
P\left[\left\{X = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}\right] &= \frac{K\,(K-1)}{N\,(N-1)}
\end{aligned}$$

Für die Koordinaten von X ergibt sich daraus

$$\begin{aligned}
P[\{X_1 = 0\}] &= (N-K)/N \\
P[\{X_1 = 1\}] &= K/N
\end{aligned}$$

und

$$\begin{aligned}
P[\{X_2 = 0\}] &= (N-K)/N \\
P[\{X_2 = 1\}] &= K/N
\end{aligned}$$

Insbesondere gilt $P_X \neq P_{X_1} \otimes P_{X_2}$. Daher sind X_1 und X_2 nicht unabhängig.

(2) **Ziehen mit Zurücklegen:** Wir bezeichnen für $i \in \{1,2\}$ mit Y_i die Anzahl der roten Kugeln bei der i–ten Ziehung und setzen

$$Y := \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

Dann gilt

$$\begin{aligned}
P\left[\left\{Y = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\}\right] &= \frac{(N-K)^2}{N^2} \\
P\left[\left\{Y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}\right] &= \frac{(N-K)\,K}{N^2} \\
P\left[\left\{Y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}\right] &= \frac{K\,(N-K)}{N^2} \\
P\left[\left\{Y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}\right] &= \frac{K^2}{N^2}
\end{aligned}$$

Für die Koordinaten von Y folgt daraus

$$\begin{aligned}
P[\{Y_1 = 0\}] &= (N-K)/N \\
P[\{Y_1 = 1\}] &= K/N
\end{aligned}$$

und

$$\begin{aligned}
P[\{Y_2 = 0\}] &= (N-K)/N \\
P[\{Y_2 = 1\}] &= K/N
\end{aligned}$$

Insbesondere gilt $P_Y = P_{Y_1} \otimes P_{Y_2}$. Daher sind Y_1 und Y_2 unabhängig.

Es gilt also $P_{X_1} = P_{Y_1}$ und $P_{X_2} = P_{Y_2}$, aber $P_X \neq P_Y$.

Wir beschließen diesen Abschnitt mit einem Null–Eins–Gesetz über die Konvergenz einer unendlichen Reihe von unabhängigen reellen Zufallsvariablen:

11.3.7 Satz (Null–Eins–Gesetz). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge von reellen Zufallsvariablen und sei*

$$A := \left\{ \liminf_{n\to\infty} \sum_{k=1}^{n} X_k = \limsup_{n\to\infty} \sum_{k=1}^{n} X_k \right\}$$

Dann gilt $P[A] \in \{0,1\}$.

Beweis. Sei $\mathcal{E}_\infty$ die terminale σ–Algebra der Folge $\{\sigma(X_k)\}_{k\in\mathbb{N}}$. Für alle $m \in \mathbb{N}$ gilt

$$A = \left\{ \liminf_{n\to\infty} \sum_{k=m}^{n} X_k = \limsup_{n\to\infty} \sum_{k=m}^{n} X_k \right\}$$

und damit $A \in \sigma(\bigcup_{k=m}^{\infty} \sigma(X_k))$. Daher gilt $A \in \bigcap_{m\in\mathbb{N}} \sigma(\bigcup_{k=m}^{\infty} \sigma(X_k)) = \mathcal{E}_\infty$ und aus dem Null–Eins–Gesetz von Kolmogorov folgt nun $P[A] \in \{0,1\}$. □

Aufgaben

11.3.A Sei $\{X_k\}_{k\in\mathbb{N}}$ eine Folge von Zufallsgrößen. Dann sind äquivalent:
(a) Die Folge $\{X_k\}_{k\in\mathbb{N}}$ ist unabhängig.
(b) Für alle $n \in \mathbb{N}$ ist die Familie $\{X_k\}_{k\in\{1,\dots,n\}}$ unabhängig.

11.3.B Sei $\{(\Omega'_k, \mathcal{F}'_k)\}_{k\in\{1,\dots,n\}}$ eine endliche Familie von Messräumen und sei $\{X_k\}_{k\in\{1,\dots,n\}}$ eine Familie von Zufallsgrößen $X_k : \Omega \to \Omega'_k$. Dann sind äquivalent:
(a) Die Familie $\{X_k\}_{k\in\{1,\dots,n\}}$ ist unabhängig.
(b) Für jede Familie $\{B'_k\}_{k\in\{1,\dots,n\}}$ mit $B'_k \in \mathcal{F}'_k$ für alle $k \in \{1,\dots,n\}$ gilt

$$P\left[\bigcap_{k\in\{1,\dots,n\}} \{X_k \in B'_k\} \right] = \prod_{k\in\{1,\dots,n\}} P[\{X_k \in B'_k\}]$$

11.3.C Sei $\{(\Omega'_i, \mathcal{F}'_i)\}_{i\in I}$ eine Familie von Messräumen und sei

$$(\Omega', \mathcal{F}') := \bigotimes_{i\in I} (\Omega'_i, \mathcal{F}'_i)$$

Sei ferner $\{X_i\}_{i\in I}$ eine Familie von Zufallsgrößen $X_i : \Omega \to \Omega'_i$ und sei $X : \Omega \to \Omega'$ die Zufallsgröße mit

$$\pi_i \circ X = X_i$$

für alle $i \in I$. Ist $\{X_i\}_{i\in I}$ unabhängig (bezüglich P), so ist auch $\{\pi_i\}_{i\in I}$ unabhängig (bezüglich P_X).

11.3.D **Blocklemma:** Verallgemeinern Sie das Blocklemma auf eine beliebige disjunkte Familie von Teilmengen der Indexmenge.

11.3.E **Paarweise unabhängige Familien von Zufallsgrößen:** Eine Familie von Zufallsgrößen $\{X_i\}_{i\in I}$ heißt *paarweise unabhängig*, wenn für alle $i,j\in I$ mit $i \neq j$ die Familie $\{X_i, X_j\}$ unabhängig ist. Jede unabhängige Familie von Zufallsgrößen ist paarweise unabhängig, aber eine paarweise unabhängige Familie von Zufallsgrößen ist nicht notwendigerweise unabhängig.

11.4 Produkte von Wahrscheinlichkeitsräumen

Wir klären abschließend das Problem der Existenz eines Wahrscheinlichkeitsraumes mit einer möglicherweise überabzählbaren Familie von unabhängigen Zufallsgrößen.

Zu diesem Zweck betrachten wir eine Familie von Wahrscheinlichkeitsräumen $\{(\Omega_i, \mathcal{F}_i, Q_i)\}_{i\in I}$ und ihr Produkt

$$(\Omega, \mathcal{F}, Q) := \bigotimes_{i\in I} (\Omega_i, \mathcal{F}_i, Q_i)$$

Der folgende Satz zeigt, dass der Produktraum reich an unabhängigen Familien von Ereignissen ist:

11.4.1 Satz. *Die Familie $\{\pi_i^{-1}(\mathcal{F}_i)\}_{i\in I}$ ist unabhängig in $(\Omega, \mathcal{F}, Q)$.*

Beweis. Wir betrachten eine Familie $\{A_i\}_{i\in I} \subseteq \mathcal{F}$ mit $A_i \in \pi_i^{-1}(\mathcal{F}_i)$ für alle $i \in I$ und zeigen, dass die Familie $\{A_i\}_{i\in I}$ unabhängig ist. Dazu wählen wir eine Familie $\{B_i\}_{i\in I}$ mit $B_i \in \mathcal{F}_i$ und $A_i = \pi_i^{-1}(B_i)$ für alle $i \in I$. Für alle $J \subseteq \mathcal{H}(I)$ gilt dann

$$\begin{aligned}
Q\left[\bigcap_{i\in J} A_i\right] &= Q\left[\bigcap_{i\in J} \pi_i^{-1}(B_i)\right] \\
&= Q\left[\pi_J^{-1}\left(\prod_{i\in J} B_i\right)\right] \\
&= Q_{\pi_J}\left[\prod_{i\in J} B_i\right] \\
&= Q_J\left[\prod_{i\in J} B_i\right] \\
&= \prod_{i\in J} Q_i[B_i]
\end{aligned}$$

$$= \prod_{i \in J} Q_{\pi_i}[B_i]$$

$$= \prod_{i \in J} Q[\pi_i^{-1}(B_i)]$$

$$= \prod_{i \in J} Q[A_i]$$

Daher ist die Familie $\{A_i\}_{i \in I}$ unabhängig. □

Aus dem Satz ergibt sich nun sofort die Existenz eines Wahrscheinlichkeitsraumes mit einer unabhängigen Familie von reellen Zufallsvariablen, deren Verteilungen beliebig gewählt werden können:

11.4.2 Folgerung. *Sei $\{Q_i\}_{i \in I}$ eine Familie von Wahrscheinlichkeitsmaßen auf $\mathcal{B}(\mathbb{R})$. Dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und eine unabhängige Familie $\{X_i\}_{i \in I}$ von Zufallsvariablen $\Omega \to \mathbb{R}$ mit*

$$P_{X_i} = Q_i$$

für alle $i \in I$.

Beweis. Sei

$$(\Omega, \mathcal{F}, P) := \bigotimes_{i \in I} (\mathbb{R}, \mathcal{B}(\mathbb{R}), Q_i)$$

und für alle $i \in I$ sei

$$X_i := \pi_i$$

Dann ist $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $\{X_i\}_{i \in I}$ ist eine Familie von Zufallsvariablen $\Omega \to \mathbb{R}$ mit $\sigma(X_i) = X_i^{-1}(\mathcal{B}(\mathbb{R})) = \pi_i^{-1}(\mathcal{B}(\mathbb{R}))$ für alle $i \in I$. Nach Satz 11.4.1 ist die Familie $\{X_i\}_{i \in I}$ unabhängig, und aus der Definition von $(\Omega, \mathcal{F}, P)$ folgt $P_{X_i} = P_{\pi_i} = Q_i$ für alle $i \in I$. □

Im letzten Ergebnis ist die Annahme, dass alle Wahrscheinlichkeitsmaße auf der Borelschen σ–Algebra $\mathcal{B}(\mathbb{R})$ definiert sind, unwesentlich; vgl. Aufgabe 11.4.B.

Aufgaben

11.4.A Vereinfachen Sie den Beweis von Satz 11.4.1 für den Fall von zwei Wahrscheinlichkeitsräumen.

11.4.B Sei $\{(\Omega_i, \mathcal{F}_i, Q_i)\}_{i \in I}$ eine Familie von Wahrscheinlichkeitsräumen. Dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und eine unabhängige Familie $\{X_i\}_{i \in I}$ von Zufallsgrößen $X_i : \Omega \to \Omega_i$ mit $P_{X_i} = Q_i$ für alle $i \in I$.

12

Univariate Verteilungen

Univariate Verteilungen sind Wahrscheinlichkeitsmaße auf $\mathcal{B}(\mathbb{R})$. Sie sind von Interesse, weil für jede reelle Zufallsvariable $X : \Omega \to \mathbb{R}$ das Bildmaß P_X von P unter X eine univariate Verteilung ist.

In diesem Kapitel untersuchen wir univariate Verteilungen. Wir beginnen mit der Charakterisierung univariater Verteilungen durch Verteilungsfunktionen und einer Reihe von Beispielen für diskrete, absolutstetige oder stetigsinguläre univariate Verteilungen (Abschnitt 12.1). Sodann untersuchen wir Transformationen univariater Verteilungen (Abschnitt 12.2) sowie die Eigenschaften ihres Erwartungswertes und ihrer höheren Momente (Abschnitt 12.3) und die ihrer Varianz und ihrer höheren zentralen Momente (Abschnitt 12.4). Dabei werden viele der im ersten Abschnitt eingeführten univariaten Verteilungen auch im Hinblick auf die Berechnung von Transformationen oder Momenten oder zentralen Momenten diskutiert.

12.1 Verteilungen und Verteilungsfunktionen

Ein Wahrscheinlichkeitsmaß $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ heißt *Verteilung auf* $\mathcal{B}(\mathbb{R})$ oder *univariate Verteilung*.

Wir haben bereits gesehen, dass univariate Verteilungen insbesondere als Bildmaße von reellen Zufallsvariablen auftreten: Ist $X : \Omega \to \mathbb{R}$ eine reelle Zufallsvariable, so ist das Bildmaß P_X von P unter X eine univariate Verteilung. Andererseits kann jede univariate Verteilung als Verteilung einer reellen Zufallsvariablen auf einem geeigneten Wahrscheinlichkeitsraum dargestellt werden: Ist $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung, so ist $(\Omega, \mathcal{F}, P) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), Q)$ ein Wahrscheinlichkeitsraum und die Abbildung $X : \Omega \to \mathbb{R}$ mit $X(\omega) := \omega$ ist eine reelle Zufallsvariable mit $P_X = P$ und damit $Q = P_X$.

In der Tat ist es sogar so, dass zahlreiche Eigenschaften von Zufallsvariablen, die in der Wahrscheinlichkeitstheorie studiert werden, in Wirklichkeit nicht

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_12, © Springer-Verlag Berlin Heidelberg 2011

Eigenschaften der Zufallsvariablen, sondern Eigenschaften ihrer Verteilungen sind.

Vor der Begründung der Wahrscheinlichkeitstheorie durch die Maß– und Integrationstheorie stand der Begriff der Verteilung naturgemäß nicht zur Verfügung; statt dessen stand der von der Analysis geprägte Begriff der Verteilungsfunktion im Vordergrund des Interesses:

Eine Funktion $F : \mathbb{R} \to [0,1]$ heißt *Verteilungsfunktion auf* $\mathbb{R}$ oder *univariate Verteilungsfunktion*, wenn sie die folgenden Eigenschaften besitzt:
(i) F ist monoton wachsend.
(ii) F ist rechtsseitig stetig.
(iii) Es gilt $\lim_{x\to-\infty} F(x) = 0$ und $\lim_{x\to\infty} F(x) = 1$.
(In Teilen der Literatur wird in der Definition einer Verteilungsfunktion anstelle der rechtsseitigen Stetigkeit die linksseitige Stetigkeit verlangt.)

Der folgende Satz beschreibt den Zusammenhang zwischen Verteilungen und Verteilungsfunktionen:

12.1.1 Satz (Korrespondenzsatz).
(1) *Zu jeder Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ gibt es genau eine Verteilungsfunktion $F_Q : \mathbb{R} \to [0,1]$ mit $F_Q(x) = Q[(-\infty, x]]$ für alle $x \in \mathbb{R}$.*
(2) *Zu jeder Verteilungsfunktion $F : \mathbb{R} \to [0,1]$ gibt es genau eine Verteilung $Q_F : \mathcal{B}(\mathbb{R}) \to [0,1]$ mit $Q_F[(-\infty, x]] = F(x)$ für alle $x \in \mathbb{R}$.*
(3) *Es gilt $Q_{(F_Q)} = Q$ und $F_{(Q_F)} = F$.*

Beweis. Wir betrachten zunächst eine Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ und zeigen, dass die Funktion $F_Q : \mathbb{R} \to \mathbb{R}$ mit

$$F_Q(x) := Q[(-\infty, x]]$$

eine Verteilungsfunktion ist:
(i) Da jede Verteilung monoton ist, gilt für alle $x, y \in \mathbb{R}$ mit $x \le y$

$$F_Q(x) = Q[(-\infty, x]] \le Q[(-\infty, y]] = F_Q(y)$$

(ii) Da jede Verteilung stetig von oben ist, gilt für alle $x \in \mathbb{R}$

$$\begin{aligned} \inf_{n\in\mathbb{N}} F_Q(x+\tfrac{1}{n}) &= \inf_{n\in\mathbb{N}} Q[(-\infty, x+\tfrac{1}{n}]] \\ &= Q\left[\bigcap_{n\in\mathbb{N}} (-\infty, x+\tfrac{1}{n}]\right] \\ &= Q[(-\infty, x]] \\ &= F_Q(x) \end{aligned}$$

und aus (i) folgt nun, dass F_Q rechtsseitig stetig ist.

(iii) Da jede Verteilung stetig von oben und stetig von unten ist, gilt

$$\inf_{n\in\mathbb{N}} F_Q(-n) = \inf_{n\in\mathbb{N}} Q[(-\infty,-n]] = Q\left[\bigcap_{n\in\mathbb{N}}(-\infty,-n]\right] = Q[\emptyset] = 0$$

und

$$\sup_{n\in\mathbb{N}} F_Q(n) = \sup_{n\in\mathbb{N}} Q[(-\infty,n]] = Q\left[\bigcup_{n\in\mathbb{N}}(-\infty,n]\right] = Q[\mathbb{R}] = 1$$

und aus (i) folgt nun $\lim_{x\to-\infty} F_Q(x) = 0$ und $\lim_{x\to\infty} F_Q(x) = 1$.

Daher ist F_Q eine Verteilungsfunktion mit $F_Q(x) = Q[(-\infty,x]]$ für alle $x \in \mathbb{R}$. Ist andererseits F eine Verteilungsfunktion mit $F(x) = Q[(-\infty,x]]$ für alle $x \in \mathbb{R}$, so gilt offenbar $F = F_Q$. Damit ist (1) gezeigt.

Wir betrachten nun eine Verteilungsfunktion $F : \mathbb{R} \to [0,1]$ und zeigen, dass die Mengenfunktion $Q_F : \mathcal{J}(\mathbb{R}) \to [0,1]$ mit

$$Q_F[(a,b]] := F(b) - F(a)$$

ein σ–endliches Maß ist. Wir führen den Beweis in mehreren Schritten:

Wir zeigen zunächst, dass Q_F ein Inhalt ist.

(i) Für alle $c \in \mathbb{R}$ gilt $(c,c] = \emptyset$ und damit

$$Q_F[\emptyset] = Q_F[(c,c]] = F(c) - F(c) = 0$$

(ii) Sei $(a,b] \in \mathcal{J}(\mathbb{R})$ und sei $\{(a_k,b_k]\}_{k\in\{1,\dots,n\}} \subseteq \mathcal{J}(\mathbb{R})$ eine endliche Familie mit

$$(a,b] = \sum_{k=1}^n (a_k,b_k]$$

Dabei können wir ohne Beschränkung der Allgemeinheit annehmen, dass $a = a_1 \le b_1 = a_2 \le \dots \le b_n = b$ gilt, und wir erhalten

$$\begin{aligned}
Q_F\left[\sum_{k=1}^n (a_k,b_k]\right] &= Q_F[(a,b]] \\
&= F(b) - F(a) \\
&= F(b_n) - F(a_1) \\
&= \sum_{k=1}^n \Big(F(b_k) - F(a_k)\Big) \\
&= \sum_{k=1}^n Q_F[(a_k,b_k]]
\end{aligned}$$

Daher ist Q_F endlich additiv.

Damit ist gezeigt, dass Q_F ein Inhalt ist.

Wir zeigen nun, dass Q_F sogar ein Maß ist. Da $\mathcal{J}(\mathbb{R})$ ein Halbring und Q_F ein Inhalt ist, genügt es nach Lemma 4.2.12 zu zeigen, dass Q_F σ–subadditiv ist. Sei $(a,b] \in \mathcal{J}(\mathbb{R})$ und sei $\{(a_k,b_k]\}_{k\in\mathbb{N}} \subseteq \mathcal{J}(\mathbb{R})$ eine Folge mit

$$(a,b] = \bigcup_{k\in\mathbb{N}} (a_k,b_k]$$

(Hier können wir nicht annehmen, dass die Folgen $\{a_k\}_{k\in\mathbb{N}}$ und $\{b_k\}_{k\in\mathbb{N}}$ monoton wachsend sind.) Da F rechtsseitig stetig ist, gibt es zu jedem $\varepsilon \in (0,\infty)$ ein $\delta \in (0,b-a)$ und eine Folge $\{\delta_k\}_{k\in\mathbb{N}} \subseteq (0,\infty)$ derart, dass

$$F(a+\delta) \leq F(a) + \varepsilon$$

und für alle $k \in \mathbb{N}$

$$F(b_k+\delta_k) \leq F(b_k) + \varepsilon/2^k$$

gilt. Dann gilt aber

$$[a+\delta,b] \subseteq (a,b] = \bigcup_{k\in\mathbb{N}} (a_k,b_k] \subseteq \bigcup_{k\in\mathbb{N}} (a_k,b_k+\delta_k)$$

Da einerseits das Intervall $[a+\delta,b]$ abgeschlossen und beschränkt und damit nach dem Satz von Heine/Borel kompakt ist und andererseits jedes der Intervalle $(a_k,b_k+\delta_k)$ offen ist, gibt es eine endliche Menge $K \subseteq \mathbb{N}$ mit

$$[a+\delta,b] \subseteq \bigcup_{k\in K} (a_k,b_k+\delta_k)$$

Daher gilt $(a+\delta,b] \subseteq [a+\delta,b] \subseteq \bigcup_{k\in K}(a_k,b_k+\delta_k) \subseteq \bigcup_{k\in K}(a_k,b_k+\delta_k]$ und damit

$$(a+\delta,b] = \bigcup_{k\in K} (a_k,b_k+\delta_k] \cap (a+\delta,b]$$

Da Q_F ein Inhalt und damit endlich subadditiv und monoton ist, ergibt sich nun

$$\begin{aligned}
Q_F[(a,b]] - \varepsilon &= F(b) - F(a) - \varepsilon \\
&\leq F(b) - F(a+\delta) \\
&= Q_F[(a+\delta,b]] \\
&= Q_F\left[\bigcup_{k\in K} (a_k,b_k+\delta_k] \cap (a+\delta,b]\right] \\
&\leq \sum_{k\in K} Q_F\Big[(a_k,b_k+\delta_k] \cap (a+\delta,b]\Big]
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{k \in K} Q_F[(a_k, b_k{+}\delta_k]] \\
&= \sum_{k \in K} \Bigl(F(b_k{+}\delta_k) - F(a_k) \Bigr) \\
&\leq \sum_{k \in K} \Bigl(F(b_k) + \varepsilon/2^k - F(a_k) \Bigr) \\
&\leq \sum_{k=1}^{\infty} \Bigl(F(b_k) - F(a_k) \Bigr) + \varepsilon \\
&= \sum_{k=1}^{\infty} Q_F[(a_k, b_k]] + \varepsilon
\end{aligned}
$$

Da $\varepsilon \in (0, \infty)$ beliebig war, erhalten wir

$$
Q_F[(a,b]] \leq \sum_{k=1}^{\infty} Q_F[(a_k, b_k]]
$$

Daher ist der Inhalt Q_F σ–subadditiv und damit ein Maß.
Wegen $\mathbb{R} = \bigcup_{n \in \mathbb{N}} (-n, n]$ und $Q_F[(-n,n]] = F(n) - F(-n) < \infty$ ist das Maß Q_F σ–endlich.
Das σ–endliche Maß Q_F besitzt nach dem Satz von Caratheodory eine eindeutige Fortsetzung zu einem Maß $\mathcal{B}(\mathbb{R}) \to [0, \infty]$, das wir wieder mit Q_F bezeichnen. Außerdem gilt für alle $x \in \mathbb{R}$

$$
\begin{aligned}
Q_F[(-\infty, x]] &= \lim_{n \to \infty} Q_F[(-n, x]] \\
&= \lim_{n \to \infty} \Bigl(F(x) - F(-n) \Bigr) \\
&= F(x) - \lim_{n \to \infty} F(-n) \\
&= F(x)
\end{aligned}
$$

und damit

$$
Q_F[\mathbb{R}] = \lim_{n \to \infty} Q_F[(-\infty, n]] = \lim_{n \to \infty} F(n) = 1
$$

Daher ist Q_F eine Verteilung mit $Q_F[(-\infty, x]] = F(x)$ für alle $x \in \mathbb{R}$.
Ist andererseits $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung mit $Q[(-\infty, x]] = F(x)$ für alle $x \in \mathbb{R}$, so erhält man für alle $(a,b] \in \mathcal{J}(\mathbb{R})$

$$
\begin{aligned}
Q[(a,b]] &= Q[(-\infty, b]] - Q[(-\infty, a]] \\
&= F(b) - F(a) \\
&= Q_F[(-\infty, b]] - Q_F[(-\infty, a]] \\
&= Q_F[(a,b]]
\end{aligned}
$$

Aus dem Eindeutigkeitssatz folgt nun $Q = Q_F$.
Damit ist (2) gezeigt, und (3) ist dann klar. □

Wir stellen nun einige Eigenschaften von Verteilungsfunktionen zusammen:

12.1.2 Lemma. *Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung und sei $F : \mathbb{R} \to [0,1]$ die zugehörige Verteilungsfunktion. Dann gilt:*
(1) *F besitzt höchstens abzählbar viele Sprungstellen.*
(2) *Für alle $x \in \mathbb{R}$ existiert der Grenzwert $\lim_{0<\varepsilon\to 0} F(x-\varepsilon)$ und es gilt*

$$\lim_{0<\varepsilon\to 0} F(x-\varepsilon) = Q[(-\infty, x)]$$

Beweis. Aus der Monotonie von F folgt (1). Da jede Verteilung stetig von unten ist, gilt

$$\sup_{n\in\mathbb{N}} F(x-\tfrac{1}{n}) = \sup_{n\in\mathbb{N}} Q[(-\infty, x-\tfrac{1}{n}]] = Q\left[\bigcup_{n\in\mathbb{N}} (-\infty, x-\tfrac{1}{n}]\right] = Q[(-\infty, x)]$$

und aus der Monotonie von F folgt nun

$$\lim_{0<\varepsilon\to 0} F(x-\varepsilon) = Q[(-\infty, x)]$$

Daher gilt (2). □

Sei $F : \mathbb{R} \to [0,1]$ eine Verteilungsfunktion. Dann setzen wir für alle $x \in \mathbb{R}$

$$F(x-) := \lim_{0<\varepsilon\to 0} F(x-\varepsilon)$$

Wir erhalten das folgende Lemma:

12.1.3 Lemma. *Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung und sei $F : \mathbb{R} \to [0,1]$ die zugehörige Verteilungsfunktion. Dann gilt für alle $x \in \mathbb{R}$*

$$F(x) - F(x-) = Q[\{x\}]$$

Insbesondere ist F genau dann stetig, wenn für alle $x \in \mathbb{R}$

$$Q[\{x\}] = 0$$

gilt.

Beweis. Nach Lemma 12.1.2 gilt für alle $x \in \mathbb{R}$

$$F(x) - F(x-) = Q[(-\infty, x]] - Q[(-\infty, x)] = Q[\{x\}]$$

Damit ist die erste Behauptung gezeigt. Da F rechtsseitig stetig ist, ist F genau dann stetig, wenn für alle $x \in \mathbb{R}$

$$F(x-) = F(x)$$

gilt. Damit ist auch die zweite Behauptung gezeigt. □

Da Verteilungen Mengenfunktionen und Verteilungsfunktionen Punktfunktionen sind, könnte der Eindruck entstehen, dass Verteilungsfunktionen leichter zu handhaben sind als Verteilungen. Dieser Eindruck trügt aus zwei Gründen:

– Verteilungsfunktionen lassen sich nur in seltenen Fällen explizit angeben.
– Selbst dann, wenn eine Verteilungsfunktion explizit angegeben werden kann, ist es oft einfacher, sie durch Summen oder Integrale darzustellen. In diesem Fall lässt sich aber aufgrund des Korrespondenzsatzes auch die Verteilung selbst durch Summen oder Integrale darstellen.

In der maßtheoretischen Wahrscheinlichkeitstheorie sind Verteilungsfunktionen von untergeordneter Bedeutung. Wir werden sie erst wieder gegen Ende dieses Abschnitts zur Konstruktion der Cantor–Verteilung und als Hilfsmittel zur Klassifikation von Verteilungen benutzen.

Von besonderem Interesse sind Verteilungen $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$, die absolutstetig bezüglich einem σ–endlichen Maß $\mu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ sind, denn für jede solche Verteilung gibt es nach dem Satz von Radon/Nikodym eine μ–fast überall eindeutig bestimmte positive messbare Funktion f mit

$$Q = \int f \, d\mu$$

In diesem Fall ist die Berechnung von Q mit Hilfe von f und μ relativ einfach; dies gilt insbesondere dann, wenn μ ein lokales Zählmaß oder das Lebesgue–Maß ist.

Diskrete Verteilungen

Eine Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ heißt *diskret*, wenn es eine abzählbare Menge $C \subseteq \mathbb{R}$ gibt mit $Q[C] = 1$.

12.1.4 Lemma. *Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine diskrete Verteilung. Dann gilt $Q \perp \boldsymbol{\lambda}$.*

Beweis. Für jede abzählbare Menge $C \subseteq \mathbb{R}$ gilt $\boldsymbol{\lambda}[C] = 0$. □

Eine Funktion $f : \mathbb{R} \to \mathbb{R}_+$ heißt *Zähldichte*, wenn es eine abzählbare Menge $C \subseteq \mathbb{R}$ gibt mit $f(x) = 0$ für alle $x \in \mathbb{R} \setminus C$ und

$$\sum_{x \in C} f(x) = 1$$

Jede Zähldichte ist messbar, denn für alle $B \in \mathcal{B}(\mathbb{R})$ gilt

$$f^{-1}(B) = f^{-1}(B \setminus \{0\}) + f^{-1}(B \cap \{0\})$$

und $f^{-1}(B \setminus \{0\})$ ist abzählbar, während $f^{-1}(B \cap \{0\})$ entweder leer ist oder ein abzählbares Komplement hat.

Jede diskrete Verteilung lässt sich durch eine Zähldichte erzeugen, und jede Zähldichte erzeugt eine diskrete Verteilung:

12.1.5 Lemma.

(1) *Sei* $Q : \mathcal{B}(\mathbb{R}) \to [0, 1]$ *eine diskrete Verteilung. Dann ist die Funktion* $f : \mathbb{R} \to \mathbb{R}_+$ *mit*
$$f(x) := Q[\{x\}]$$
eine Zähldichte und für alle $B \in \mathcal{B}(\mathbb{R})$ *gilt*
$$Q[B] = \sum_{x \in B} f(x)$$

(2) *Sei* $f : \mathbb{R} \to \mathbb{R}_+$ *eine Zähldichte. Dann ist die Mengenfunktion* $Q : \mathcal{B}(\mathbb{R}) \to [0, 1]$ *mit*
$$Q[B] := \sum_{x \in B} f(x)$$
eine Verteilung und für alle $x \in \mathbb{R}$ *gilt*
$$Q[\{x\}] = f(x)$$

Der Begriff der Zähldichte weckt Assoziationen im Zusammenhang mit dem Satz von Radon/Nikodym, die es zu klären gilt. Sei $C \subseteq \mathbb{R}$ abzählbar.

– Das lokale Zählmaß $\boldsymbol{\zeta}_C$ bezüglich C ist σ–endlich und es gilt
$$\boldsymbol{\zeta}_C = \int \chi_C \, d\boldsymbol{\zeta}$$

– Ist Q eine diskrete Verteilung mit $Q[C] = 1$ und ist f eine Zähldichte mit
$$Q[B] = \sum_{x \in B} f(x)$$
so gilt
$$Q = \int f \, d\boldsymbol{\zeta}_C = \int f \chi_C \, d\boldsymbol{\zeta}$$
und damit ist f eine $\boldsymbol{\zeta}_C$–Dichte von Q und $f\chi_C$ ist eine $\boldsymbol{\zeta}$–Dichte von Q.

– Ist Q eine diskrete Verteilung mit $Q[C] = 1$ und ist $f : \mathbb{R} \to \mathbb{R}_+$ eine messbare Funktion mit
$$Q = \int f \, d\boldsymbol{\zeta}_C$$
so ist f eine $\boldsymbol{\zeta}_C$–Dichte von Q aber nicht notwendigerweise eine Zähldichte. Der Grund liegt darin, dass die $\boldsymbol{\zeta}_C$–Dichte von Q nur $\boldsymbol{\zeta}_C$–fast überall eindeutig bestimmt ist und die Menge $\mathbb{R} \setminus C$ eine $\boldsymbol{\zeta}_C$–Nullmenge ist, sodass f auf der Menge $\mathbb{R} \setminus C$ beliebige Werte annehmen kann.

Die Darstellung einer diskreten Verteilung als Integral einer Zähldichte bezüglich einem lokalen Zählmaß wird sich gelegentlich als nützlich erweisen.

Das folgende Lemma liefert ein einfaches Prinzip für die Konstruktion von Zähldichten:

12.1.6 Lemma. *Sei $g : \mathbb{R} \to \mathbb{R}_+$ eine Funktion, für die es eine abzählbare Menge $C \subseteq \mathbb{R}$ gibt mit $g(x) = 0$ für alle $x \in \mathbb{R} \setminus C$ und*

$$0 < \sum_{z \in C} g(z) < \infty$$

Dann ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{g(x)}{\sum_{z \in C} g(z)}$$

eine Zähldichte.

Die einfachste diskrete univariate Verteilung ist die Dirac–Verteilung:

12.1.7 Beispiel (Dirac–Verteilung). Für $z \in \mathbb{R}$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} 1 & \text{falls } x = z \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Dirac–Verteilung* und wird mit $\boldsymbol{\delta}_z$ bezeichnet. Für die Verteilung $Q := \boldsymbol{\delta}_z$ gilt

$$Q[B] = \begin{cases} 0 & \text{falls } z \notin B \\ 1 & \text{falls } z \in B \end{cases}$$

und für die zugehörige Verteilungsfunktion F gilt

$$F(x) = \begin{cases} 0 & \text{falls } x < z \\ 1 & \text{falls } x \geq z \end{cases}$$

Die Funktion F wird auch als *Heaviside–Funktion* bezeichnet.

Wir betrachten nun einige parametrische Klassen von diskreten univariaten Verteilungen Q mit $Q[\mathbb{N}_0] = 1$:

12.1.8 Beispiele (Diskrete Verteilungen).

(1) **Hypergeometrische Verteilung:** Für $n, N, K \in \mathbb{N}$ mit $\max\{n, K+1\} \leq N$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \dfrac{\dbinom{K}{x}\dbinom{N-K}{n-x}}{\dbinom{N}{n}} & \text{falls } x \in \{0, 1, \ldots, n\} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *hypergeometrische Verteilung* und wird mit $\mathbf{H}(n, N, K)$ bezeichnet.

Modell: Eine Urne enthalte N Kugeln unterschiedlicher Farben, von denen eine Farbe ausgezeichnet ist und die anderen Farben nicht ausgezeichnet sind. Es bezeichne K die Anzahl der Kugeln der ausgezeichneten Farbe in der Urne und X die zufällige Anzahl der Kugeln dieser Farbe in einer Stichprobe vom Umfang n beim Ziehen ohne Zurücklegen. Dann gilt $P_X = \mathbf{H}(n, N, K)$.

(2) **Binomial–Verteilung:** Für $n \in \mathbb{N}$ und $\vartheta \in (0,1)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \binom{n}{x} \vartheta^x (1-\vartheta)^{n-x} & \text{falls } x \in \{0, 1, \ldots, n\} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Binomial–Verteilung* und wird mit $\mathbf{B}(n, \vartheta)$ bezeichnet.
Modell: Eine Urne enthalte Kugeln unterschiedlicher Farben, von denen eine Farbe ausgezeichnet ist und die anderen Farben nicht ausgezeichnet sind. Es bezeichne ϑ den Anteil der Kugeln der ausgezeichneten Farbe in der Urne und X die zufällige Anzahl der Kugeln dieser Farbe in einer Stichprobe vom Umfang n beim Ziehen mit Zurücklegen. Dann gilt $P_X = \mathbf{B}(n, \vartheta)$.
Spezialfall: Die Binomial–Verteilung $\mathbf{B}(1, \vartheta)$ wird auch als *Bernoulli–Verteilung* oder als *Boole–Verteilung* und mit $\mathbf{B}(\vartheta)$ bezeichnet.

(3) **Poisson–Verteilung:** Für $\alpha \in (0, \infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} e^{-\alpha} \dfrac{\alpha^x}{x!} & \text{falls } x \in \mathbb{N}_0 \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte.
In der Tat: Es gilt

$$\sum_{k=0}^{\infty} \frac{\alpha^k}{k!} = e^\alpha$$

und damit

$$\sum_{k=0}^{\infty} f(k) = 1$$

Die zugehörige Verteilung heißt *Poisson–Verteilung* und wird mit $\mathbf{P}(\alpha)$ bezeichnet.

(4) **Negativbinomial–Verteilung:** Für $\alpha \in (0, \infty)$ und $\vartheta \in (0,1)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \binom{\alpha + x - 1}{x} (1-\vartheta)^\alpha \, \vartheta^x & \text{falls } x \in \mathbb{N}_0 \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Negativbinomial–Verteilung* und wird mit $\mathbf{NB}(\alpha, \vartheta)$ bezeichnet.
Spezialfall: Die Negativbinomial–Verteilung $\mathbf{NB}(n, \vartheta)$ mit $n \in \mathbb{N}$ wird auch als *Pascal–Verteilung* bezeichnet.

(5) **Geometrische Verteilung:** Für $n \in \mathbb{N}$ und $\vartheta \in (0,1)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \binom{x-1}{n-1}(1-\vartheta)^{x-n}\vartheta^n & \text{falls } x \in \mathbb{N}(n) \\ & \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *geometrische Verteilung* und wird mit $\mathbf{Geo}(n,\vartheta)$ bezeichnet.
Modell: Eine Urne enthalte Kugeln unterschiedlicher Farben, von denen eine Farbe ausgezeichnet ist und die anderen Farben nicht ausgezeichnet sind. Es bezeichne ϑ den Anteil der Kugeln der ausgezeichneten Farbe in der Urne und X die zufällige Anzahl der Ziehungen beim Ziehen mit Zurücklegen, bis genau n Kugeln der ausgezeichneten Farbe gezogen sind. Dann gilt $P_X = \mathbf{Geo}(n,\vartheta)$.

Absolutstetige Verteilungen

Eine Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ heißt *absolutstetig*, wenn sie absolutstetig bezüglich dem Lebesgue–Maß $\boldsymbol{\lambda}$ ist, also $Q \ll \boldsymbol{\lambda}$ gilt.

12.1.9 Lemma. *Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine absolutstetige Verteilung und sei $F : \mathbb{R} \to [0,1]$ die zugehörige Verteilungsfunktion. Dann gilt $Q[\{x\}] = 0$ für alle $x \in \mathbb{R}$ und F ist stetig.*

Beweis. Für alle $x \in \mathbb{R}$ gilt $\boldsymbol{\lambda}[\{x\}] = 0$ und aus $Q \ll \boldsymbol{\lambda}$ folgt nun $Q[\{x\}] = 0$. Damit ist die erste Behauptung gezeigt, und die zweite Behauptung folgt aus Lemma 12.1.3. □

Eine messbare Funktion $f : \mathbb{R} \to \mathbb{R}_+$ heißt *Lebesgue–Dichte*, wenn

$$\int_{\mathbb{R}} f(x)\, d\boldsymbol{\lambda}(x) = 1$$

gilt. Jede absolutstetige Verteilung lässt sich durch eine Lebesgue–Dichte erzeugen, und jede Lebesgue-Dichte erzeugt eine absolutstetige Verteilung:

12.1.10 Lemma.
(1) *Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine absolutstetige Verteilung. Dann gibt es eine Lebesgue–Dichte $f : \mathbb{R} \to \mathbb{R}_+$ mit*

$$Q = \int f(x)\, d\boldsymbol{\lambda}(x)$$

und die Lebesgue–Dichte ist $\boldsymbol{\lambda}$–fast überall eindeutig bestimmt.
(2) *Sei $f : \mathbb{R} \to \mathbb{R}_+$ eine Lebesgue–Dichte. Dann ist die Mengenfunktion $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ mit*

$$Q := \int f(x)\, d\boldsymbol{\lambda}(x)$$

eine absolutstetige Verteilung.

Beweis. Nach dem Satz von Radon/Nikodym gibt es zu jeder Verteilung Q mit $Q \ll \boldsymbol{\lambda}$ eine positive messbare Funktion g mit

$$Q = \int g \, d\boldsymbol{\lambda}$$

und g ist $\boldsymbol{\lambda}$–fast überall eindeutig bestimmt. Da Q eine Verteilung ist und damit ein endliches Maß ist, ist g auf dem Komplement einer $\boldsymbol{\lambda}$–Nullmenge endlich. Daher gibt es eine Lebesgue–Dichte f mit $f(x) = g(x)$ $\boldsymbol{\lambda}$–fast überall und $Q = \int f \, d\boldsymbol{\lambda}$. Damit ist (1) gezeigt, und (2) ergibt sich aus Satz 9.1.1. □

Das folgende Lemma liefert ein einfaches Prinzip für die Konstruktion von Lebesgue–Dichten:

12.1.11 Lemma. *Sei $g : \mathbb{R} \to \mathbb{R}_+$ eine messbare Funktion mit*

$$0 < \int_{\mathbb{R}} g(z) \, d\boldsymbol{\lambda}(z) < \infty$$

Dann ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{g(x)}{\int_{\mathbb{R}} g(z) \, d\boldsymbol{\lambda}(z)}$$

eine Lebesgue–Dichte.

Im Gegensatz zu Zähldichten gilt für eine Lebesgue–Dichte f *nicht* notwendigerweise $f(x) \leq 1$ für alle $x \in \mathbb{R}$:

12.1.12 Beispiel. Die Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0, 1]$ mit

$$Q[B] := 2 \, \boldsymbol{\lambda}[B \cap (0, 1/2)]$$

ist absolutstetig und es gilt $Q = \int f \, d\boldsymbol{\lambda}$ mit $f = 2\chi_{(0,1/2)}$.

Das Beispiel zeigt, dass die Werte einer Lebesgue–Dichte einer absolutstetigen Verteilung *nicht* als Wahrscheinlichkeiten interpretiert werden können.

Wir betrachten nun einige parametrische Klassen von absolutstetigen univariaten Verteilungen:

12.1.13 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Für $a, b \in \mathbb{R}$ mit $a < b$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{1}{b - a} \, \chi_{(a,b)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *uniforme Verteilung* auf (a, b) und wird mit $\mathbf{U}(a, b)$ bezeichnet. Für die zugehörige Verteilungsfunktion gilt

$$F(x) = \begin{cases} 0 & \text{falls} \quad x < a \\ \dfrac{x-a}{b-a} & \text{falls } a \leq x < b \\ 1 & \text{falls } b \leq x \end{cases}$$

(2) **Beta–Verteilung:** Für $\alpha, \beta \in (0, \infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{1}{\mathrm{B}(\alpha, \beta)}\, x^{\alpha-1}(1-x)^{\beta-1}\, \chi_{(0,1)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Beta–Verteilung* und wird mit $\mathbf{Be}(\alpha, \beta)$ bezeichnet.
Spezialfall: Es gilt $\mathbf{Be}(1,1) = \mathbf{U}(0,1)$.

(3) **Gamma–Verteilung:** Für $\alpha, \gamma \in (0, \infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{\alpha^\gamma}{\Gamma(\gamma)}\, e^{-\alpha x} x^{\gamma-1}\, \chi_{(0,\infty)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Gamma–Verteilung* und wird mit $\mathbf{Ga}(\alpha, \gamma)$ bezeichnet.
Spezialfälle:

– Die Gamma–Verteilung $\mathbf{Ga}(\alpha, m)$ mit $m \in \mathbb{N}$ wird auch als *Erlang–Verteilung* bezeichnet. Für die zugehörige Verteilungsfunktion gilt

$$F(x) = \begin{cases} 0 & \text{falls } x \leq 0 \\ 1 - e^{-\alpha x} \displaystyle\sum_{k=0}^{m-1} \frac{(\alpha x)^k}{k!} & \text{falls } x > 0 \end{cases}$$

– Die Gamma–Verteilung $\mathbf{Ga}(\alpha, 1)$ wird auch als *Exponential–Verteilung* und mit $\mathbf{Exp}(\alpha)$ bezeichnet. Für die zugehörige Verteilungsfunktion gilt

$$F(x) = \begin{cases} 0 & \text{falls } x \leq 0 \\ 1 - e^{-\alpha x} & \text{falls } x > 0 \end{cases}$$

– Die Gamma–Verteilung $\mathbf{Ga}(1/2, n/2)$ mit $n \in \mathbb{N}$ wird auch als *χ^2–Verteilung* und mit $\boldsymbol{\chi}^2_n$ bezeichnet. Der Parameter n heißt *Anzahl der Freiheitsgrade*.

Zum Nachweis, dass f tatsächlich eine Lebesgue–Dichte ist, betrachten wir die lineare Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit $T(x) := \alpha x$. Nach Satz 6.3.1 gilt $\boldsymbol{\lambda}_T = \alpha^{-1}\boldsymbol{\lambda}$ und aus der Substitutionsregel und der Definition der Gamma–Funktion erhalten wir nun

$$\begin{aligned}
&\int_{\mathbb{R}} e^{-\alpha x} x^{\gamma-1} \chi_{(0,\infty)}(x)\, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{\alpha^{\gamma-1}} \int_{\mathbb{R}} e^{-\alpha x} (\alpha x)^{\gamma-1} \chi_{(0,\infty)}(\alpha x)\, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{\alpha^{\gamma-1}} \int_{T^{-1}(\mathbb{R})} e^{-T(x)} (T(x))^{\gamma-1} \chi_{(0,\infty)}(T(x))\, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{\alpha^{\gamma-1}} \int_{\mathbb{R}} e^{-z} z^{\gamma-1} \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}_T(z)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\alpha^{\gamma-1}} \int_{\mathbb{R}} e^{-z} z^{\gamma-1} \chi_{(0,\infty)}(z)\, \frac{1}{\alpha}\, d\boldsymbol{\lambda}(z) \\
&= \frac{1}{\alpha^{\gamma}} \int_{(0,\infty)} e^{-z} z^{\gamma-1}\, d\boldsymbol{\lambda}(z) \\
&= \frac{\Gamma(\gamma)}{\alpha^{\gamma}}
\end{aligned}
$$

(4) **Normal–Verteilung:** Für $\mu \in \mathbb{R}$ und $\sigma \in (0,\infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$
f(x) := \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}
$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Normal–Verteilung* oder *Gauß–Verteilung* und wird mit $\mathbf{N}(\mu, \sigma^2)$ bezeichnet. (Man beachte, dass hier σ^2 und nicht σ als Parameter verwendet wird; vgl. Beispiel 13.1.7(3).)
Spezialfall: Die Normal–Verteilung $\mathbf{N}(0,1)$ wird als *Standardnormal–Verteilung* bezeichnet.
Zum Nachweis, dass f tatsächlich eine Lebesgue–Dichte ist, betrachten wir die affine Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit

$$
T(x) := \frac{x-\mu}{\sigma\sqrt{2}}
$$

Nach Folgerung 6.3.2 gilt $\boldsymbol{\lambda}_T = (\sigma\sqrt{2})\boldsymbol{\lambda}$ und aus der Substitutionsregel erhalten wir nun

$$
\begin{aligned}
\int_{\mathbb{R}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} d\boldsymbol{\lambda}(x) &= \int_{T^{-1}(\mathbb{R})} e^{-(T(x))^2} d\boldsymbol{\lambda}(x) \\
&= \int_{\mathbb{R}} e^{-z^2} d\boldsymbol{\lambda}_T(z) \\
&= \int_{\mathbb{R}} e^{-z^2} \sigma\sqrt{2}\, d\boldsymbol{\lambda}(z) \\
&= \sigma\sqrt{2} \int_{\mathbb{R}} e^{-z^2} d\boldsymbol{\lambda}(z) \\
&= \sigma\sqrt{2}\,\sqrt{\pi} \\
&= \sqrt{2\pi}\,\sigma
\end{aligned}
$$

(5) t–**Verteilung:** Für $n \in \mathbb{N}$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$
f(x) := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}
$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *t–Verteilung* und wird mit $\mathbf{t}_n$ bezeichnet. Der Parameter n heißt *Anzahl der Freiheitsgrade*.
Dass f tatsächlich eine Lebesgue–Dichte ist, zeigen wir in Beispiel 12.2.5.

Stetigsinguläre Verteilungen

Eine Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ heißt *stetigsingulär*, wenn sie singulär bezüglich dem Lebesgue–Maß $\boldsymbol{\lambda}$ ist, also $Q \perp \boldsymbol{\lambda}$ gilt, und außerdem $Q[\{x\}] = 0$ für alle $x \in \mathbb{R}$ gilt, also die zugehörige Verteilungsfunktion F stetig ist. Das folgende Beispiel zeigt, dass stetigsinguläre univariate Verteilungen existieren:

12.1.14 Beispiel (Cantor–Verteilung). Sei $\mathcal{Z}$ das System aller Teilmengen von $[0,1]$, die als Vereinigung von endlich vielen disjunkten abgeschlossenen (und nichtleeren) Intervallen dargestellt werden können, und sei $\Psi : \mathcal{Z} \to \mathcal{Z}$ gegeben durch

$$\Psi\left(\sum_{i=1}^m [a_i, b_i]\right) := \sum_{i=1}^m \left(\left[a_i, \frac{2a_i + b_i}{3}\right] + \left[\frac{a_i + 2b_i}{3}, b_i\right]\right)$$

Dann gilt $\mathcal{Z} \subseteq \mathcal{B}(\mathbb{R})$ und die Abbildung Ψ ist wohldefiniert. Für alle $A \in \mathcal{Z}$ gilt $\Psi(A) \subseteq A$ und

$$\boldsymbol{\lambda}[\Psi(A)] = \frac{2}{3}\,\boldsymbol{\lambda}[A]$$

Für alle $n \in \mathbb{N}$ sei

$$C_n := \Psi^n([0,1])$$

Dann ist $\{C_n\}_{n\in\mathbb{N}}$ eine monoton fallende Folge in $\mathcal{B}(\mathbb{R}^n)$ und für alle $n \in \mathbb{N}$ gilt

$$\boldsymbol{\lambda}[C_n] = \left(\frac{2}{3}\right)^n$$

Für die Cantor–Menge

$$C := \bigcap_{n\in\mathbb{N}} C_n$$

gilt daher

$$\boldsymbol{\lambda}[C] = 0$$

(vgl. Beispiel 5.5.8).
Wir konstruieren nun eine Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ mit $Q[C] = 1$ und $Q[\{x\}] = 0$ für alle $x \in \mathbb{R}$. Dann gilt $Q[\mathbb{R} \setminus C] + \boldsymbol{\lambda}[C] = 0$ und damit $Q \perp \boldsymbol{\lambda}$, sowie $Q[\{x\}] = 0$ für alle $x \in \mathbb{R}$; daher ist Q stetigsingulär.
(1) Für $n \in \mathbb{N}$ sei $f_n : \mathbb{R} \to \mathbb{R}_+$ gegeben durch

$$f_n(x) := \left(\frac{3}{2}\right)^n \chi_{C_n}(x)$$

Dann ist f_n messbar. Sei ferner $Q_n : \mathcal{B}(\mathbb{R}) \to [0,1]$ gegeben durch

$$Q_n := \int f_n(z)\, d\boldsymbol{\lambda}(z)$$

Dann ist Q_n ein Maß mit $Q_n \ll \boldsymbol{\lambda}$ und für alle $B \in \mathcal{B}(\mathbb{R})$ gilt

$$\begin{aligned}
Q_n[B] &= \int_B f_n(z)\, d\boldsymbol{\lambda}(z) \\
&= \int_{\mathbb{R}} \left(\frac{3}{2}\right)^n \chi_{C_n}(z)\chi_B(z)\, d\boldsymbol{\lambda}(z) \\
&= \left(\frac{3}{2}\right)^n \boldsymbol{\lambda}[C_n \cap B]
\end{aligned}$$

Wegen $\boldsymbol{\lambda}[C_n] = (2/3)^n$ folgt daraus $Q_n[\mathbb{R}] = 1$. Daher ist Q_n eine absolutstetige Verteilung und die zugehörige Verteilungsfunktion F_n ist stetig. Für alle $x \in \mathbb{R}$ gilt

$$F_n(x) = Q_n[(-\infty, x]] = \left(\frac{3}{2}\right)^n \boldsymbol{\lambda}[C_n \cap (-\infty, x]] = \left(\frac{3}{2}\right)^n \boldsymbol{\lambda}[C_n \cap [0, x]]$$

und damit $F_n(0) = 0$ und $F_n(1) = 1$.
(2) Nach Konstruktion besteht C_n aus 2^n abgeschlossenen Intervallen gleicher Länge. Ist $I \subseteq C_n$ ein solches Intervall, so gilt aufgrund der Translationsinvarianz des Lebesgue–Maßes

$$\boldsymbol{\lambda}[I] = \frac{1}{2^n}\, \boldsymbol{\lambda}[C_n] = \frac{1}{2^n}\left(\frac{2}{3}\right)^n = \frac{1}{3^n}$$

und für alle $m \in \mathbb{N}_0$ gilt wegen $C_{n+m} \cap I = \Psi^m(C_n) \cap I = \Psi^m(C_n \cap I) = \Psi^m(I)$

$$\begin{aligned}
\int_I f_{n+m}(z)\, d\boldsymbol{\lambda}(z) &= Q_{n+m}[I] \\
&= \left(\frac{3}{2}\right)^{n+m} \boldsymbol{\lambda}[C_{n+m} \cap I] \\
&= \left(\frac{3}{2}\right)^{n+m} \boldsymbol{\lambda}[\Psi^m(I)] \\
&= \left(\frac{3}{2}\right)^{n+m} \left(\frac{2}{3}\right)^m \boldsymbol{\lambda}[I] \\
&= \left(\frac{3}{2}\right)^{n+m} \left(\frac{2}{3}\right)^m \frac{1}{3^n} \\
&= \frac{1}{2^n}
\end{aligned}$$

(3) Für $n \in \mathbb{N}$ sei

$$F_n^* := F_n|_{[0,1]}$$

Dann ist F_n^* monoton wachsend und stetig mit $F_n^*(0) = 0$ und $F_n^*(1) = 1$, und für alle $m \in \mathbb{N}_0$ und $x \in [0, 1]$ gilt

$$\begin{aligned}
F_{n+m}^*(x) &= F_{n+m}(x) \\
&= Q_{n+m}[(-\infty, x]] \\
&= \int_{(-\infty, x]} f_{n+m}(z)\, d\boldsymbol{\lambda}(z) \\
&= \int_{[0,x]\cap C_n} f_{n+m}(z)\, d\boldsymbol{\lambda}(z)
\end{aligned}$$

– Für $x \in [0,1] \setminus C_n$ folgt aus (2) für alle $m \in \mathbb{N}$

$$\Big| F^*_{n+m}(x) - F^*_n(x) \Big| = 0$$

– Für $x \in C_n$ sei $I_{n,x}$ das eindeutig bestimmte maximale abgeschlossene Intervall mit $x \in I_{n,x} \subseteq C_n$ und sei $a_{n,x}$ seine Untergrenze. Wegen (2) gilt für alle $m \in \mathbb{N}$

$$\begin{aligned}
\Big| F^*_{n+m}(x) - F^*_n(x) \Big| &= \left| \int_{[0,x]\cap C_n} f_{n+m}(z)\, d\boldsymbol{\lambda}(z) - \int_{[0,x]\cap C_n} f_n(z)\, d\boldsymbol{\lambda}(z) \right| \\
&= \left| \int_{[a_{n,x},x]} f_{n+m}(z)\, d\boldsymbol{\lambda}(z) - \int_{[a_{n,x},x]} f_n(z)\, d\boldsymbol{\lambda}(z) \right| \\
&\leq \int_{[a_{n,x},x]} f_{n+m}(z)\, d\boldsymbol{\lambda}(z) + \int_{[a_{n,x},x]} f_n(z)\, d\boldsymbol{\lambda}(z) \\
&\leq \int_{I_{n,x}} f_{n+m}(z)\, d\boldsymbol{\lambda}(z) + \int_{I_{n,x}} f_n(z)\, d\boldsymbol{\lambda}(z) \\
&= 2\,\frac{1}{2^n}
\end{aligned}$$

Daher gilt für alle $x \in [0,1]$ und für alle $n, m \in \mathbb{N}$

$$|F^*_{n+m}(x) - F^*_n(x)| \leq \frac{1}{2^{n-1}}$$

Daher ist die Folge $\{F^*_n\}_{n\in\mathbb{N}}$ eine gleichmäßige Cauchy–Folge und damit gleichmäßig konvergent gegen die Funktion $F^* : [0,1] \to [0,1]$ mit

$$F^*(x) := \lim_{n\to\infty} F^*_n(x)$$

Da alle F^*_n stetig sind, ist auch F^* stetig. Außerdem gilt $F^*(0) = 0$ und $F^*(1) = 1$. Da die Folge $\{F^*_n\}_{n\in\mathbb{N}}$ gleichmäßig gegen F^* konvergiert, gibt es zu jedem $\varepsilon \in (0,\infty)$ ein $n \in \mathbb{N}$ mit

$$|F^*(z) - F^*_n(z)| \leq \varepsilon$$

für alle $z \in [0,1]$, und für $x, y \in [0,1]$ mit $x \leq y$ folgt daraus und aus der Monotonie von F^*_n

$$F^*(x) - \varepsilon \leq F^*_n(x) \leq F^*_n(y) \leq F^*(y) + \varepsilon$$

Da $\varepsilon \in (0,\infty)$ beliebig war, erhalten wir für alle $x, y \in [0,1]$ mit $x \leq y$

$$F^*(x) \leq F^*(y)$$

Daher ist F^* monoton wachsend.
(4) Sei $F : \mathbb{R} \to [0,1]$ gegeben durch

$$F(x) := \begin{cases} 0 & \text{falls} \quad x < 0 \\ F^*(x) & \text{falls} \;\; 0 \leq x < 1 \\ 1 & \text{falls} \;\; 1 \leq x \end{cases}$$

Dann ist F eine stetige Verteilungsfunktion. Die Funktion F heißt *Cantor–Funktion* und die zugehörige Verteilung Q heißt *Cantor–Verteilung*. Es gilt

$$Q[[0,1]] = F(1) - F(0-) = F(1) - F(0) = 1$$

Ist (a,b) eines der maximalen offenen Intervalle aus $[0,1] \setminus C_n$, so gilt

$$\begin{aligned}
Q[(a,b)] &= F(b-) - F(a) \\
&= F(b) - F(a) \\
&= \lim_{n\to\infty} (F_n(b) - F_n(a)) \\
&= \lim_{n\to\infty} (F_n(b-) - F_n(a)) \\
&= \lim_{n\to\infty} Q_n[(a,b)] \\
&= 0
\end{aligned}$$

Daraus folgt $Q[[0,1] \setminus C_n] = 0$ und damit $Q[C_n] = 1$. Dann gilt aber

$$Q[C] = Q\left[\bigcap_{n\in\mathbb{N}} C_n\right] = \lim_{n\to\infty} Q[C_n] = 1$$

und damit $Q[\mathbb{R} \setminus C] = 0$. Wegen $\boldsymbol{\lambda}[C] = 0$ folgt daraus $Q \perp \boldsymbol{\lambda}$. Daher ist Q stetigsingulär.
Da die Cantor–Verteilung stetigsingulär ist, ist die Cantor–Menge nicht abzählbar.

Ein Darstellungssatz

Die Klassen der diskreten, absolutstetigen und stetigsingulären Verteilungen sind offensichtlich disjunkt. Der folgende Satz zeigt, dass jede Verteilung als Konvexkombination einer diskreten, einer absolutstetigen und einer stetigsingulären Verteilung dargestellt werden kann:

12.1.15 Satz (Darstellungssatz). *Zu jeder Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ gibt es $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}_+$ mit $\alpha_1 + \alpha_2 + \alpha_3 = 1$ und Verteilungen Q_1, Q_2, Q_3 mit*

$$Q = \alpha_1 Q_1 + \alpha_2 Q_2 + \alpha_3 Q_3$$

derart, dass Q_1 diskret, Q_2 stetigsingulär, und Q_3 absolutstetig ist.

Beweis. Nach der Lebesgue–Zerlegung gibt es Maße $\mu_0, \mu_3 : \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+$ mit $\mu_0 \perp \boldsymbol{\lambda}$ und $\mu_3 \ll \boldsymbol{\lambda}$ sowie

$$Q = \mu_0 + \mu_3$$

Die Maße μ_0 und μ_3 sind endlich und es gilt $\mu_0[\mathbb{R}] + \mu_3[\mathbb{R}] = 1$.
Wir nehmen zunächst an, dass $\mu_0[\mathbb{R}] = 0$ und damit $Q = \mu_3 \ll \boldsymbol{\lambda}$ gilt. In diesem Fall sei $\alpha_1 := 0$ und Q_1 eine beliebige diskrete Verteilung, $\alpha_2 := 0$ und Q_2 eine beliebige stetigsinguläre Verteilung, sowie $\alpha_3 := 1$ und $Q_3 := Q$. Dann gilt

$$Q = \alpha_1 Q_1 + \alpha_2 Q_2 + \alpha_3 Q_3$$

Wir nehmen für das weitere an, dass $\mu_0[\mathbb{R}] > 0$ gilt. In diesem Fall sind weitere Fallunterscheidungen erforderlich:

– Im Fall $\mu_0[\mathbb{R}] \in (0,1)$ gilt $\mu_3[\mathbb{R}] \in (0,1)$; in diesem Fall sei $\alpha_0 := \mu_0[\mathbb{R}]$ und $Q_0 := \alpha_0^{-1}\mu_0$ sowie $\alpha_3 := \mu_3[\mathbb{R}]$ und $Q_3 := \alpha_3^{-1}\mu_3$. Dann sind Q_0 und Q_3 Verteilungen mit $Q_0 \perp \boldsymbol{\lambda}$ und $Q_3 \ll \boldsymbol{\lambda}$ sowie

$$Q = \alpha_0 Q_0 + \alpha_3 Q_3$$

– Im Fall $\mu_0[\mathbb{R}] = 1$ gilt $Q = \mu_0 \perp \boldsymbol{\lambda}$; in diesem Fall sei $\alpha_0 := 1$ und $Q_0 := Q$ sowie $\alpha_3 := 0$ und Q_3 eine beliebige absolutstetige Verteilung. Dann gilt $Q_0 \perp \boldsymbol{\lambda}$ und $Q_3 \ll \boldsymbol{\lambda}$ sowie

$$Q = \alpha_0 Q_0 + \alpha_3 Q_3$$

Sei nun F_0 die zu der Verteilung Q_0 gehörige Verteilungsfunktion und sei

$$S := \Bigl\{ x \in \mathbb{R} \Bigm| Q_0[\{x\}] \neq 0 \Bigr\}$$

die Menge der Unstetigkeitsstellen von F_0.

– Im Fall $Q_0[S] = 0$ ist F_0 stetig, und wegen $Q_0 \perp \boldsymbol{\lambda}$ ist Q_0 stetigsingulär; in diesem Fall sei $\beta_1 := 0$ und Q_1 eine beliebige diskrete Verteilung sowie $\beta_2 := 1$ und $Q_2 := Q_0$. Dann gilt

$$Q_0 = \beta_1 Q_1 + \beta_2 Q_2$$

– Im Fall $Q_0[S] \in (0,1)$ seien $\mu_1, \mu_2 : \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+$ gegeben durch

$$\begin{aligned} \mu_1[B] &:= Q_0[B \cap S] \\ \mu_2[B] &:= Q_0[B \cap \overline{S}] \end{aligned}$$

Dann gilt $\mu_1[\mathbb{R}], \mu_2[\mathbb{R}] \in (0,1)$ und $Q_0 = \mu_1 + \mu_2$. In diesem Fall sei $\beta_1 := \mu_1[\mathbb{R}]$ und $Q_1 := \beta_1^{-1}\mu_1$ sowie $\beta_2 := \mu_2[\mathbb{R}]$ und $Q_2 := \beta_2^{-1}\mu_2$. Dann ist Q_1 diskret und Q_2 ist stetigsingulär, und es gilt

$$Q_0 = \beta_1 Q_1 + \beta_2 Q_2$$

– Im Fall $Q_0[S] = 1$ ist Q_0 diskret; in diesem Fall sei $\beta_1 := 1$ und $Q_1 := Q_0$ sowie $\beta_2 := 0$ und Q_2 eine beliebige stetigsinguläre Verteilung. Dann gilt

$$Q_0 = \beta_1 Q_1 + \beta_2 Q_2$$

Mit $\alpha_1 := \alpha_0\beta_1$ und $\alpha_2 := \alpha_0\beta_2$ erhalten wir daher auch im Fall $\mu_0[\mathbb{R}] > 0$ eine Darstellung

$$Q = \alpha_1 Q_1 + \alpha_2 Q_2 + \alpha_3 Q_3$$

mit einer diskreten Verteilung Q_1, einer stetigsingulären Verteilung Q_2 und einer absolutstetigen Verteilung Q_3. Damit ist der Satz bewiesen. □

Aufgaben

12.1.A **Symmetrische Verteilungen:** Eine Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ heißt *symmetrisch*, wenn für die Abbildung $S : \mathbb{R} \to \mathbb{R}$ mit $S(x) := -x$

$$Q_S = Q$$

gilt. Für eine Verteilung Q und die zugehörige Verteilungsfunktion F sind äquivalent:
(a) Q ist symmetrisch.
(b) Für alle $x \in \mathbb{R}$ gilt $F(-x) + F(x-) = 1$.

12.1.B **Überlebensfunktion:** Eine Funktion $G : \mathbb{R} \to [0,1]$ heißt *Überlebensfunktion*, wenn sie die folgenden Eigenschaften besitzt:
(i) G ist monoton fallend.
(ii) G ist rechtsseitig stetig.
(iii) Es gilt $\lim_{x\to-\infty} G(x) = 1$ und $\lim_{x\to\infty} G(x) = 0$.
Für Funktionen $F, G : \mathbb{R} \to [0,1]$ mit $F(x) + G(x) = 1$ für alle $x \in \mathbb{R}$ sind folgenden Aussagen äquivalent:
(a) F ist eine Verteilungsfunktion.
(b) G ist eine Überlebensfunktion.

12.1.C Bestimmen Sie die Verteilungsfunktion und die Überlebensfunktion der Exponential–Verteilung.

12.1.D **Stochastische Ordnung:** Sei $\mathcal{Q}^0(\mathbb{R})$ die Familie aller Verteilungen $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$. Für $Q_1, Q_2 \in \mathcal{Q}^0(\mathbb{R})$ schreiben wir

$$Q_1 \leq^0 Q_2$$

wenn für alle $a \in \mathbb{R}$

$$Q_1[(a,\infty)] \leq Q_2[(a,\infty)]$$

gilt. Die Relation $\leq^0$ wird als *stochastische Ordnung* bezeichnet.
(1) Die Relation $\leq^0$ ist eine Ordnungsrelation auf $\mathcal{Q}^0(\mathbb{R})$.
(2) Für $Q_1, Q_2 \in \mathcal{Q}^0(\mathbb{R})$ und die zugehörigen Überlebensfunktionen G_1, G_2 gilt $Q_1 \leq^0 Q_2$ genau dann, wenn für alle $a \in \mathbb{R}$

$$G_1(a) \leq G_2(a)$$

gilt.
(3) Untersuchen Sie die Vergleichbarkeit der Verteilungen $\mathbf{B}(2, \frac{1}{2}), \mathbf{B}(3, \frac{1}{2})$, $\mathbf{B}(3, \frac{1}{3})$ bezüglich der stochastischen Ordnung.

12.1.E **Verteilungsfunktion auf einem abgeschlossenen Intervall:** Sei $[a,b] \subseteq \bar{\mathbb{R}}$ ein Intervall. Eine Funktion $F : [a,b] \to [0,1]$ heißt *Verteilungsfunktion auf* $[a,b]$, wenn sie die folgenden Eigenschaften besitzt:
(i) F ist monoton wachsend.
(ii) F ist rechtsseitig stetig.
(iii) Es gilt $F(a) = 0$ und $\lim_{x\to b} F(x) = 1$.
Formulieren und beweisen Sie ein Analogon des Korrespondenzsatzes und zeigen Sie, dass eine Bijektion zwischen den Verteilungsfunktionen auf $\mathbb{R}$ und den Verteilungsfunktionen auf dem Intervall $[-\infty, \infty]$ besteht.

12.1.F Stochastische Folgen: Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung. Dann gilt $Q[\mathbb{N}_0] = 1$ genau dann, wenn es eine stochastische Folge $\{q_k\}_{k\in\mathbb{N}_0}$ gibt mit

$$Q = \sum_{k=0}^{\infty} q_k \, \boldsymbol{\delta}_k$$

12.1.G Geometrische Verteilung: Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung. Dann sind äquivalent:

(a) Es gilt $Q[\mathbb{N}] = 1$ und für alle $m, n \in \mathbb{N}$ gilt

$$Q[(m+n,\infty)] = Q[(m,\infty)]\, Q[(n,\infty)]$$

(b) Es gilt $Q = \boldsymbol{\delta}_1$ oder es gibt ein $\vartheta \in (0,1)$ mit $Q = \mathbf{Geo}(1,\vartheta)$.

12.1.H Pólya–Verteilung: Für $n \in \mathbb{N}$ und $\alpha, \beta \in (0,\infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \dfrac{\dbinom{\alpha+x-1}{x}\dbinom{\beta+n-x-1}{n-x}}{\dbinom{\alpha+\beta+n-1}{n}} & \text{falls } x \in \{0,1,\dots,n\} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Pólya–Verteilung* und wird mit $\mathbf{Pólya}(n,\alpha,\beta)$ bezeichnet.

12.1.I Gemischte Binomial–Verteilung: Für $n \in \mathbb{N}$ und jede Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ mit $Q[(0,1)] = 1$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \displaystyle\int_{\mathbb{R}} \binom{n}{x} \vartheta^x (1-\vartheta)^{n-x}\, dQ(\vartheta) & \text{falls } x \in \{0,1,\dots,n\} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *gemischte Binomial–Verteilung* bezüglich Q und wird mit $\mathbf{B}(n,Q)$ bezeichnet. Es gilt

$$\mathbf{B}(n, \mathbf{Be}(\alpha,\beta)) = \mathbf{Pólya}(n,\alpha,\beta)$$

Andererseits kann jede Pólya–Verteilung als gemischte Binomial–Verteilung bezüglich einer Beta–Verteilung dargestellt werden.

12.1.J Gemischte Poisson–Verteilung: Für jede Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ mit $Q[(0,\infty)] = 1$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \displaystyle\int_{\mathbb{R}} e^{-\alpha} \frac{\alpha^x}{x!}\, dQ(\alpha) & \text{falls } x \in \mathbb{N}_0 \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *gemischte Poisson–Verteilung* bezüglich Q und wird mit $\mathbf{P}(Q)$ bezeichnet. Es gilt

$$\mathbf{P}(\mathbf{Ga}(\beta,\gamma)) = \mathbf{NB}\biggl(\gamma, \frac{1}{\beta+1}\biggr)$$

Andererseits kann jede Negativbinomial–Verteilung als gemischte Poisson–Verteilung bezüglich einer Gamma–Verteilung dargestellt werden.

12.1.K **Logarithmische Verteilung:** Für $\vartheta \in (0,1)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \begin{cases} \dfrac{1}{|\log(1-\vartheta)|}\, \dfrac{\vartheta^x}{x} & \text{falls } x \in \mathbb{N} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *logarithmische Verteilung* und wird mit $\mathbf{Log}(\vartheta)$ bezeichnet.

12.1.L **Exponential–Verteilung:** Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung. Dann sind äquivalent:

(a) Es gilt $Q[(0,\infty)] = 1$ und für alle $x, y \in (0,\infty)$ gilt

$$Q[(x+y,\infty)] = Q[(x,\infty)]\, Q[(y,\infty)]$$

(b) Es gibt ein $\alpha \in (0,\infty)$ mit $Q = \mathbf{Exp}(\alpha)$.

12.1.M **Normal–Verteilung:** Die Lebesgue–Dichte der Normal–Verteilung $\mathbf{N}(\mu,\sigma^2)$ besitzt ein globales Maximum an der Stelle $x = \mu$ und Wendepunkte an den Stellen $x = \mu-\sigma$ und $x = \mu+\sigma$.

12.1.N F**–Verteilung:** Für $m, n \in \mathbb{N}$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{(\frac{m}{n})^{m/2}}{\mathrm{B}(\frac{m}{2},\frac{n}{2})}\, x^{\frac{m}{2}-1} \biggl(1 + \frac{m}{n}\,x\biggr)^{-\frac{m+n}{2}} \chi_{(0,\infty)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *F–Verteilung.*

12.1.O **Cauchy–Verteilung:** Für $\alpha \in \mathbb{R}$ und $\beta \in (0,\infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{1}{\pi}\, \frac{\beta}{\beta^2 + (x-\alpha)^2}$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Cauchy–Verteilung* und wird mit $\mathbf{Ca}(\alpha,\beta)$ bezeichnet.
Spezialfall: Es gilt $\mathbf{Ca}(0,1) = \mathbf{t}_1$.

12.1.P **Pareto–Verteilung:** Für $\alpha, \beta \in (0,\infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{\beta}{\alpha} \biggl(\frac{\alpha}{x}\biggr)^{\beta+1} \chi_{(\alpha,\infty)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Pareto–Verteilung europäischer Art* und wird mit $\mathbf{Pa}(\alpha,\beta)$ bezeichnet. Bestimmen Sie die Verteilungsfunktion und die Überlebensfunktion.

12.1.Q **Pareto–Verteilung:** Für $\alpha, \beta \in (0, \infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \frac{\beta}{\alpha} \left(\frac{\alpha}{\alpha + x} \right)^{\beta+1} \chi_{(0,\infty)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Pareto–Verteilung amerikanischer Art* und wird mit $\mathbf{Pa}^*(\alpha, \beta)$ bezeichnet. Bestimmen Sie die Verteilungsfunktion und die Überlebensfunktion.

12.1.R **Ausfallrate:** Sei $f : \mathbb{R} \to \mathbb{R}_+$ eine Lebesgue–Dichte mit

$$\int_{(-\infty,x]} f(z)\, d\boldsymbol{\lambda}(z) < 1$$

für alle $x \in \mathbb{R}$. Sei Q die zugehörige Verteilung und sei G die zugehörige Überlebensfunktion. Dann heißt die Funktion $r : \mathbb{R} \to \mathbb{R}_+$ mit

$$r(x) := \frac{f(x)}{G(x)}$$

Ausfallrate von Q.

(1) Ist f stetig, so ist G differenzierbar und es gilt

$$r(x) = -\frac{d(\log \circ G)}{dx}(x)$$

und

$$G(x) = \exp\left(-\int_{(-\infty,x]} r(z)\, d\boldsymbol{\lambda}(z) \right)$$

(2) Für $\alpha \in (0, \infty)$ gilt $r = \alpha\chi_{(0,\infty)}$ genau dann, wenn $Q = \mathbf{Exp}(\alpha)$ gilt.

12.1.S **Weibull–Verteilung:** Für $\alpha, \beta \in (0, \infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \alpha\beta x^{\beta-1} \exp(-\alpha x^{\beta})\, \chi_{(0,\infty)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Weibull–Verteilung*.

(1) Bestimmen Sie die Überlebensfunktion und die Ausfallrate.
(2) Welche Verteilung ergibt sich im Fall $\beta = 1$?

12.1.T **Gompertz–Verteilung:** Für $\alpha \in (0, \infty)$ und $\beta \in (1, \infty)$ ist die Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$f(x) := \alpha\beta^x \exp\left(\frac{\alpha}{\log(\beta)} (1 - \beta^x) \right) \chi_{(0,\infty)}(x)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Gompertz–Verteilung*. Bestimmen Sie die Überlebensfunktion und die Ausfallrate.

12.2 Transformationen von Verteilungen

Ein grundlegendes Problem der Wahrscheinlichkeitstheorie besteht darin, für eine reelle Zufallsvariable X und eine messbare Funktion $T : \mathbb{R} \to \mathbb{R}$ die Verteilung der reellen Zufallsvariablen $T \circ X$ und damit das Bildmaß $P_{T \circ X} = (P_X)_T$ von P_X unter T zu bestimmen.

12.2.1 Beispiel. Sei X eine reelle Zufallsvariable. Dann besitzt die reelle Zufallsvariable X^2 die Darstellung $X^2 = T \circ X$ mit der messbaren Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit $T(x) := x^2$.

Allgemeiner stellt sich das Problem, für eine Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ und eine messbare Funktion $T : \mathbb{R} \to \mathbb{R}$ das Bildmaß Q_T zu bestimmen. Dies ist ein Spezialfall des Problems, für ein Maß $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ und eine messbare Funktion $T : \mathbb{R} \to \mathbb{R}$ das Bildmaß ν_T zu bestimmen.

Für ein Maß $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ mit $\nu \ll \boldsymbol{\zeta}_C$ für eine abzählbare Menge $C \subseteq \mathbb{R}$ ist die Bestimmung des Bildmaßes von ν unter einer beliebigen messbaren Abbildung $T : \mathbb{R} \to \mathbb{R}$ elementar:

12.2.2 Satz. *Sei $h : \mathbb{R} \to \mathbb{R}_+$ messbar und $C \subseteq \mathbb{R}$ abzählbar und sei*

$$\nu := \int h(x)\, d\boldsymbol{\zeta}_C(x)$$

Sei ferner $T : \mathbb{R} \to \mathbb{R}$ messbar. Dann gilt für alle $B \in \mathcal{B}(\mathbb{R})$

$$\nu_T[B] = \sum_{z \in B} \sum_{x \in C \cap T^{-1}(\{z\})} h(x)$$

Insbesondere gilt $\nu_T[\mathbb{R}] = \nu[\mathbb{R}]$.

Beweis. Für alle $B \in \mathcal{B}(\mathbb{R})$ gilt

$$\begin{aligned}\nu_T[B] &= \nu[T^{-1}(B)] \\ &= \int_{T^{-1}(B)} h(x)\, d\boldsymbol{\zeta}_C(x) \\ &= \sum_{x \in C \cap T^{-1}(B)} h(x) \\ &= \sum_{z \in B} \sum_{x \in C \cap T^{-1}(\{z\})} h(x)\end{aligned}$$

Außerdem gilt $\nu_T[\mathbb{R}] = \nu[T^{-1}(\mathbb{R})] = \nu[\mathbb{R}]$. □

Für ein Maß $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ mit $\nu \ll \boldsymbol{\lambda}$ bestimmen wir das Bildmaß von ν zunächst für invertierbare affine Abbildungen $T : \mathbb{R} \to \mathbb{R}$:

12.2.3 Satz. *Sei $h : \mathbb{R} \to \mathbb{R}_+$ messbar und*

$$\nu := \int h(x)\, d\boldsymbol{\lambda}(x)$$

Sei ferner $T : \mathbb{R} \to \mathbb{R}$ gegeben durch $T(x) := a + bx$ mit $a, b \in \mathbb{R}$ und $b \neq 0$. Dann gilt

$$\nu_T = \int h\left(\frac{z-a}{b}\right) \frac{1}{|b|}\, d\boldsymbol{\lambda}(z)$$

Insbesondere gilt $\nu_T[\mathbb{R}] = \nu[\mathbb{R}]$.

Beweis. Die Abbildung T ist invertierbar und T^{-1} ist messbar mit

$$T^{-1}(z) = \frac{z-a}{b}$$

Nach Folgerung 9.4.2 gilt daher

$$\nu_T = \int h\left(\frac{z-a}{b}\right) d\boldsymbol{\lambda}_T(z)$$

und nach Folgerung 6.3.2 gilt

$$\boldsymbol{\lambda}_T = \frac{1}{|b|}\,\boldsymbol{\lambda}$$

Die Behauptung folgt nun aus der Kettenregel. □

Invertierbare affine Abbildungen sind injektiv. Es treten jedoch auch Abbildungen auf, die nur lokal injektiv sind; vgl. Beispiel 12.2.1. Für lokal injektive Abbildungen erhalten wir den folgenden Satz:

12.2.4 Satz. *Sei $h : \mathbb{R} \to \mathbb{R}_+$ messbar und*

$$\nu := \int h(x)\, d\boldsymbol{\lambda}(x)$$

Sei ferner $T : \mathbb{R} \to \mathbb{R}$ messbar. Ist $\{C_i\}_{i\in I}$ eine abzählbare disjunkte Familie nichtleerer offener Intervalle mit $\nu[\mathbb{R}\setminus\sum_{i\in I} C_i] = 0$ derart, dass für alle $i \in I$ die Restriktion $T|_{C_i}$ stetig differenzierbar ist mit $(T|_{C_i})'(x) \neq 0$ für alle $x \in C_i$, so gilt

$$\nu_T = \int \left(\sum_{i\in I} h\Big((T|_{C_i})^{-1}(z)\Big)\left|\frac{d(T|_{C_i})^{-1}}{dz}(z)\right| \chi_{T(C_i)}(z)\right) d\boldsymbol{\lambda}(z)$$

Beweis. Aus Satz 9.5.4 und Lemma 9.8.9 erhält man unter Verwendung des Satzes über die monotone Konvergenz für alle $B' \in \mathcal{B}(\mathbb{R})$

$$\begin{aligned}
\nu_T[B'] &= \sum_{i\in I} \int_{B'\cap T(C_i)} h\Big((T|_{C_i})^{-1}(z)\Big)\, d(\boldsymbol{\lambda}|_{\mathcal{B}(C_i)})_{(T|_{C_i})}(z) \\
&= \sum_{i\in I} \int_{B'\cap T(C_i)} h\Big((T|_{C_i})^{-1}(z)\Big)\left|\frac{d(T|_{C_i})^{-1}}{dz}(z)\right| d\boldsymbol{\lambda}(z) \\
&= \int_{B'} \left(\sum_{i\in I} h\Big((T|_{C_i})^{-1}(z)\Big)\left|\frac{d(T|_{C_i})^{-1}}{dz}(z)\right| \chi_{T(C_i)}(z)\right) d\boldsymbol{\lambda}(z)
\end{aligned}$$

Damit ist der Satz bewiesen. □

Die Voraussetzungen von Satz 12.2.4 sind insbesondere für jede Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit $T(x) = a + bx$ mit $a, b \in \mathbb{R}$ und $b \neq 0$ erfüllt, denn in diesem Fall ist T stetig differenzierbar mit

$$T^{-1}(z) = \frac{z-a}{b}$$

und

$$\left| \frac{dT^{-1}}{dz}(z) \right| = \frac{1}{|b|}$$

Daher ist Satz 12.2.4 eine Verallgemeinerung von Satz 12.2.3.

Als erste Anwendung von Satz 12.2.4 zeigen wir, dass die t–Verteilung wohldefiniert ist:

12.2.5 Beispiel (t–Verteilung). Für alle $n \in \mathbb{N}$ gilt

$$\int_{\mathbb{R}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} d\boldsymbol{\lambda}(x) = 1$$

In der Tat: Wegen

$$\int \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} d\boldsymbol{\lambda}(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} \int \frac{n^{n/2}}{(n+x^2)^{(n+1)/2}}\, d\boldsymbol{\lambda}(x)$$

betrachten wir das Maß

$$\nu := \int \frac{n^{n/2}}{(n+x^2)^{(n+1)/2}}\, d\boldsymbol{\lambda}(x)$$

und die lineare Abbildung $S : \mathbb{R} \to \mathbb{R}$ mit $S(x) := n^{-1/2}\, x$. Aus Satz 12.2.3 ergibt sich

$$\begin{aligned}\nu_S &= \int \frac{n^{n/2}}{(n + ny^2)^{(n+1)/2}}\, n^{1/2}\, d\boldsymbol{\lambda}(y) \\ &= \int \frac{1}{(1 + y^2)^{(n+1)/2}}\, d\boldsymbol{\lambda}(y)\end{aligned}$$

Wir betrachten nun die Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit $T(y) := \arctan y \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Dann ist T injektiv mit $T^{-1}(z) = \tan z$, und aus Satz 12.2.4 ergibt sich nun mit $1 + (\tan z)^2 = (\cos z)^{-2}$

$$\begin{aligned}(\nu_S)_T &= \int \frac{1}{(1 + (\tan z)^2)^{(n+1)/2}} \Bigl(1 + (\tan z)^2\Bigr)\, \chi_{(-\frac{\pi}{2}, \frac{\pi}{2})}(z)\, d\boldsymbol{\lambda}(z) \\ &= \int (\cos z)^{n-1} \chi_{(-\frac{\pi}{2}, \frac{\pi}{2})}(z)\, d\boldsymbol{\lambda}(z)\end{aligned}$$

Wegen $\mathbb{R} = S^{-1}(\mathbb{R})$ und $\mathbb{R} = T^{-1}((-\frac{\pi}{2}, \frac{\pi}{2}))$ erhalten wir daher

$$\begin{aligned}
\nu[\mathbb{R}] &= \nu_S[\mathbb{R}] \\
&= (\nu_S)_T[(-\tfrac{\pi}{2}, \tfrac{\pi}{2})] \\
&= \int_{(-\pi/2,\pi/2)} (\cos z)^{n-1} d\boldsymbol{\lambda}(z) \\
&= \int_{[-\pi/2,\pi/2]} (\cos z)^{n-1} d\boldsymbol{\lambda}(z) \\
&= \int_{-\pi/2}^{\pi/2} (\cos z)^{n-1} dz
\end{aligned}$$

und mit Fallunterscheidung danach, ob n gerade oder ungerade ist, ergibt sich daraus

$$\nu[\mathbb{R}] = \frac{\Gamma(\frac{n}{2})\,\Gamma(\frac{1}{2})}{\Gamma(\frac{n+1}{2})}$$

Daraus folgt die Behauptung.

Des weiteren eignet sich Satz 12.2.4 zur Berechnung der Verteilung einer messbaren Transformation einer Zufallsvariablen. Wir geben zunächst einen weiteren Spezialfall des Satzes an:

12.2.6 Folgerung. *Sei $h : \mathbb{R} \to \mathbb{R}_+$ messbar und*

$$\nu := \int h(x)\, d\boldsymbol{\lambda}(x)$$

Sei ferner $T : \mathbb{R} \to \mathbb{R}$ gegeben durch $T(x) := x^2$. Dann gilt

$$\nu_T = \int \frac{h(-\sqrt{z}) + h(\sqrt{z})}{2\sqrt{z}}\, \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}(z)$$

Beweis. Die Abbildung T ist stetig differenzierbar mit $T'(x) = 2x$. Wegen $T'(0) = 0$ und $T'(x) \neq 0$ für alle $x \neq 0$ betrachten wir die Mengen

$$\begin{aligned}
C_1 &:= (-\infty, 0) \\
C_2 &:= (0, \infty)
\end{aligned}$$

Dann gilt $\boldsymbol{\lambda}[\mathbb{R} \setminus (C_1+C_2)] = \boldsymbol{\lambda}[\{0\}] = 0$ und damit $\nu[\mathbb{R} \setminus (C_1+C_2)] = 0$, sowie

$$\begin{aligned}
T(C_1) &= (0, \infty) \\
T(C_2) &= (0, \infty)
\end{aligned}$$

Insbesondere sind die Restriktionen $T|_{C_1}$ und $T|_{C_2}$ injektiv und es gilt

$$\begin{aligned}
(T|_{C_1})^{-1}(z) &= -\sqrt{z} \\
(T|_{C_2})^{-1}(z) &= \sqrt{z}
\end{aligned}$$

und damit

$$\left|\frac{d(T|_{C_1})^{-1}}{dz}(z)\right| = \frac{1}{2\sqrt{z}}$$
$$\left|\frac{d(T|_{C_2})^{-1}}{dz}(z)\right| = \frac{1}{2\sqrt{z}}$$

Aus Satz 12.2.4 folgt nun

$$\begin{aligned}\nu_T &= \int \left(h(-\sqrt{z})\,\frac{1}{2\sqrt{z}}\,\chi_{(0,\infty)}(z) + h(\sqrt{z})\,\frac{1}{2\sqrt{z}}\,\chi_{(0,\infty)}(z)\right) d\boldsymbol{\lambda}(z)\\ &= \int \frac{h(-\sqrt{z}) + h(\sqrt{z})}{2\sqrt{z}}\,\chi_{(0,\infty)}(z)\,d\boldsymbol{\lambda}(z)\end{aligned}$$

Damit ist die Behauptung gezeigt. □

Wir geben nun einige Beispiele für Transformationen von Verteilungen von reellen Zufallsvariablen:

12.2.7 Beispiele (Absolutstetige Verteilungen).

(1) **Normal–Verteilung:** Sei $P_X = \mathbf{N}(\mu, \sigma^2)$. Dann gilt für alle $a, b \in \mathbb{R}$ mit $b \neq 0$

$$P_{a+bX} = \mathbf{N}(a{+}b\mu, (|b|\sigma)^2)$$

In der Tat: Es gilt

$$P_X = \int \frac{1}{\sqrt{2\pi}\,\sigma}\,\exp\!\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) d\boldsymbol{\lambda}(x)$$

Wir betrachten die affine Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit $T(x) := a + b\,x$. Dann gilt $P_{a+bX} = P_{T\circ X} = (P_X)_T$ und aus Satz 12.2.3 ergibt sich

$$\begin{aligned}P_{a+bX} &= (P_X)_T\\ &= \int \frac{1}{\sqrt{2\pi}\,\sigma}\,\exp\!\left(-\frac{1}{2}\left(\frac{(z{-}a)b^{-1}-\mu}{\sigma}\right)^2\right)\frac{1}{|b|}\,d\boldsymbol{\lambda}(z)\\ &= \int \frac{1}{\sqrt{2\pi}\,|b|\sigma}\,\exp\!\left(-\frac{1}{2}\left(\frac{z-(a+b\mu)}{|b|\sigma}\right)^2\right) d\boldsymbol{\lambda}(z)\\ &= \mathbf{N}(a{+}b\mu, (|b|\sigma)^2)\end{aligned}$$

(2) χ^2–**Verteilung:** Sei $P_X = \mathbf{N}(0,1)$. Dann gilt

$$P_{X^2} = \boldsymbol{\chi}^2_1$$

In der Tat: Es gilt

$$P_X = \int \frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}\,d\boldsymbol{\lambda}(x)$$

Wir betrachten die Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit $T(x) := x^2$. Dann gilt $P_{X^2} = P_{T \circ X} = (P_X)_T$ und aus Folgerung 12.2.6 ergibt sich

$$\begin{aligned} P_{X^2} &= (P_X)_T \\ &= \int \frac{\frac{1}{\sqrt{2\pi}}\, e^{-\frac{z}{2}} + \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z}{2}}}{2\sqrt{z}}\, \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}(z) \\ &= \int \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z}\, z^{-\frac{1}{2}}\, \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}(z) \\ &= \int \frac{\left(\frac{1}{2}\right)^{1/2}}{\Gamma(\frac{1}{2})}\, e^{-\frac{1}{2}z}\, z^{\frac{1}{2}-1}\, \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}(z) \\ &= \mathbf{Ga}\left(\tfrac{1}{2}, \tfrac{1}{2}\right) \end{aligned}$$

und damit $P_{X^2} = \boldsymbol{\chi}^2_1$.

Da das Bildmaß einer Verteilung wieder eine Verteilung ist, kann man die Berechnung einer Verteilung oft vereinfachen: Ist $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung und $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ ein endliches Maß, so schreiben wir

$$Q \sim \nu$$

wenn es ein $c \in (0,\infty)$ gibt derart, dass für alle $B \in \mathcal{B}(\mathbb{R})$

$$Q[B] = c\,\nu[B]$$

gilt. Beispielsweise erhält man in Beispiel 12.2.7(2) unter Vernachlässigung der multiplikativen Konstanten mit

$$\begin{aligned} (P_X)_T &= \int \frac{\frac{1}{\sqrt{2\pi}}\, e^{-\frac{z}{2}} + \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z}{2}}}{2\sqrt{z}}\, \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}(z) \\ &\sim \int e^{-\frac{1}{2}z}\, z^{\frac{1}{2}-1}\, \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}(z) \\ &\sim \mathbf{Ga}\left(\tfrac{1}{2}, \tfrac{1}{2}\right) \end{aligned}$$

dasselbe Ergebnis, da es nur eine Möglichkeit gibt, den letzten Integranden zu einer Lebesgue–Dichte zu normieren.

Aufgaben

12.2.A **Geometrische Verteilung:** Sei $P_X = \mathbf{Geo}(n,\vartheta)$. Dann gilt $P_{X-n} = \mathbf{NB}(n, 1-\vartheta)$.

12.2.B **Uniforme Verteilung:** Sei $P_X = \mathbf{U}(0,1)$ und sei $n \in \mathbb{N}$. Dann gilt $P_{X^n} = \mathbf{Be}(\frac{1}{n}, 1)$.

12.2.C **Uniforme Verteilung:** Sei $P_X = \mathbf{U}(-\frac{\pi}{2}, \frac{\pi}{2})$ und sei $\alpha \in \mathbb{R}$ und $\beta \in (0,\infty)$. Dann gilt $P_{\alpha+\beta \tan X} = \mathbf{Ca}(\alpha, \beta)$.

12.2.D **Cauchy–Verteilung:** Sei $P_X = \mathbf{Ca}(\alpha, \beta)$. Dann gilt $P_{(X-\alpha)/\beta} = \mathbf{t}_1$ sowie $P_{X-\alpha} = \mathbf{Ca}(0, \beta)$ und $P_{\alpha-X} = \mathbf{Ca}(0, \beta)$.

12.2.E **Pareto–Verteilung:** Sei $P_X = \mathbf{Pa}^*(\alpha, \beta)$. Dann gilt $P_{X+\alpha} = \mathbf{Pa}(\alpha, \beta)$.

12.2.F **Symmetrische Verteilungen:** Zeigen Sie, dass
- die Standardnormal–Verteilung,
- jede t–Verteilung, und
- für alle $\beta \in (0, \infty)$ die Cauchy–Verteilung $\mathbf{Ca}(0, \beta)$

symmetrisch ist.

12.2.G Sei $h : \mathbb{R} \to \mathbb{R}_+$ messbar und

$$\nu := \int h(x)\, d\boldsymbol{\lambda}(x)$$

Sei ferner $T : \mathbb{R} \to \mathbb{R}$ gegeben durch $T(x) := |x|$. Dann gilt

$$\nu_T = \int \Big(h(-z) + h(z)\Big)\, \chi_{(0,\infty)}(z)\, d\boldsymbol{\lambda}(z)$$

12.3 Momente

In diesem Abschnitt untersuchen wir die Existenz und die Endlichkeit des Erwartungswertes und der höheren Momente einer Zufallsvariablen.

Wir führen zunächst die in der Wahrscheinlichkeitstheorie übliche Sprechweise ein. Sei $X : \Omega \to \bar{\mathbb{R}}$ eine Zufallsvariable.

– Man sagt, X *besitzt einen Erwartungswert*, wenn X P–quasiintegrierbar ist. Dies ist genau dann der Fall, wenn

$$\min\left\{\int_\Omega X^+(\omega)\, dP(\omega)\,,\, \int_\Omega X^-(\omega)\, dP(\omega)\right\} < \infty$$

gilt, und in diesem Fall heißt

$$E[X] := \int_\Omega X(\omega)\, dP(\omega)$$

der *Erwartungswert* von X. Da jede positive Zufallsvariable einen (möglicherweise unendlichen) Erwartungswert besitzt, besitzt X genau dann einen Erwartungswert, wenn

$$\min\Big\{E[X^+],\, E[X^-]\Big\} < \infty$$

gilt.

– Man sagt, X *besitzt einen endlichen Erwartungswert*, wenn X P–integrierbar ist. Dies ist genau dann der Fall, wenn

$$\max\left\{\int_\Omega X^+(\omega)\, dP(\omega)\,,\, \int_\Omega X^-(\omega)\, dP(\omega)\right\} < \infty$$

gilt. Daher besitzt X genau dann einen endlichen Erwartungswert, wenn

$$\max\Bigl\{E[X^+],\, E[X^-]\Bigr\} < \infty$$

gilt, und dies ist genau dann der Fall, wenn

$$E[|X|] < \infty$$

gilt.

Das folgende Lemma liefert ein weiteres Kriterium für die Existenz eines endlichen Erwartungswertes:

12.3.1 Lemma. *Für jede Zufallsvariable X gilt*

$$E[|X|] = \int_{\mathbb{R}_+} P[\{|X| \geq x\}]\, d\boldsymbol{\lambda}(x)$$

und

$$\sum_{n=1}^{\infty} P[\{|X| \geq n\}] \leq E[|X|] \leq 1 + \sum_{n=1}^{\infty} P[\{|X| \geq n\}]$$

Insbesondere besitzt X genau dann einen endlichen Erwartungswert, wenn die Reihe $\sum_{n=1}^{\infty} P[\{|X| \geq n\}]$ gegen eine reelle Zahl konvergiert.

Beweis. Die Gleichung ergibt sich unmittelbar aus Beispiel 9.7.4. Des weiteren gilt für alle $n \in \mathbb{N}$

$$\begin{aligned} P[\{|X| \geq n\}] &\leq \int_{[n-1,n)} P[\{|X| \geq x\}]\, d\boldsymbol{\lambda}(x) \\ &\leq P[\{|X| \geq n-1\}] \end{aligned}$$

und Summation ergibt

$$\begin{aligned} \sum_{n=1}^{\infty} P[\{|X| \geq n\}] &\leq \int_{\mathbb{R}_+} P[\{|X| \geq x\}]\, d\boldsymbol{\lambda}(x) \\ &\leq \sum_{n=1}^{\infty} P[\{|X| \geq n-1\}] \\ &= \sum_{n=0}^{\infty} P[\{|X| \geq n\}] \end{aligned}$$

Die Behauptung folgt. □

Die Eigenschaften des Erwartungswertes ergeben sich daher unmittelbar aus denen des Lebesgue–Integrals; gegenüber der allgemeinen Integrationstheorie ergibt sich aber die Besonderheit, dass jede konstante Zufallsvariable integrierbar ist:

12.3.2 Lemma. *Sei X eine Zufallsvariable und sei $a, b, c \in \mathbb{R}$.*
(1) *Im Fall $P[\{X \geq 0\}] = 1$ gilt $E[X] \geq 0$.*
(2) *Im Fall $P[\{X = c\}] = 1$ gilt $E[X] = c$.*
(3) *Besitzt X einen Erwartungswert, so besitzt auch $a + bX$ einen Erwartungswert und es gilt $E[a+bX] = a + b\,E[X]$.*

Für die Prüfung der Existenz des Erwartungswertes einer Zufallsvariablen, und darüber hinaus für seine Berechnung im Falle der Existenz, ist die Substitutionsregel von größter Bedeutung. Wir formulieren daher als Spezialfall der allgemeinen Substitutionsregel 9.4.1 die folgende Substitutionsregel für Erwartungswerte:

12.3.3 Satz (Substitutionsregel für Erwartungswerte). *Sei X eine reelle Zufallsvariable und sei $h : \mathbb{R} \to \mathbb{R}$ eine messbare Funktion.*
(1) *Ist h positiv, so gilt*

$$E[h \circ X] = \int_{\mathbb{R}} h(x)\, dP_X(x)$$

(2) *$h \circ X$ ist genau dann P–integrierbar, wenn h P_X–integrierbar ist, und in diesem Fall gilt*

$$E[h \circ X] = \int_{\mathbb{R}} h(x)\, dP_X(x)$$

Aus der Substitutionsregel folgt, dass die Existenz des Erwartungswertes einer Zufallsvariablen und, im Falle der Existenz, auch der Erwartungswert selbst ausschließlich durch ihre Verteilung bestimmt ist:

12.3.4 Folgerung. *Sei X eine Zufallsvariable. Dann gilt:*
(1) *X besitzt genau dann einen Erwartungswert, wenn*

$$\min\left\{ \int_{\mathbb{R}} x^+ \, dP_X(x) \,,\, \int_{\mathbb{R}} x^- \, dP_X(x) \right\} < \infty$$

gilt.
(2) *X besitzt genau dann einen endlichen Erwartungswert, wenn*

$$\max\left\{ \int_{\mathbb{R}} x^+ \, dP_X(x) \,,\, \int_{\mathbb{R}} x^- \, dP_X(x) \right\} < \infty$$

gilt, und diese Bedingung ist genau dann erfüllt, wenn

$$\int_{\mathbb{R}} |x| \, P_X(x) < \infty$$

gilt.
(3) *Besitzt X einen endlichen Erwartungswert, so gilt*

$$E[X] = \int_{\mathbb{R}} x \, dP_X(x)$$

Die Substitutionsregel für Erwartungswerte ist vor allem deshalb von Interesse, weil von einer Zufallsvariablen meist nur ihre Verteilung bekannt ist, während ihre Eigenschaften als Abbildung und oft sogar der Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ unbekannt sind.

Für eine Zufallsvariable, die diskret oder absolutstetig ist und einen endlichen Erwartungswert besitzt, erfolgt die Berechnung des Erwartungswertes meist dadurch, dass man den Erwartungswert als Vielfaches des Integrals einer Zähldichte oder einer Lebesgue–Dichte darstellt; dagegen erfordert der Nachweis dafür, dass eine Zufallsvariable keinen Erwartungswert oder keinen endlichen Erwartungswert besitzt, im allgemeinen eine Abschätzung der Erwartungswerte des Positiv– und/oder Negativteils der Zufallsvariablen.

Der Erwartungswert einer reellen Zufallsvariablen mit einer diskreten Verteilung kann, im Fall seiner Existenz, durch Summation bestimmt werden:

12.3.5 Lemma. *Sei X eine reelle Zufallsvariable mit einer diskreten Verteilung und sei f eine Zähldichte von P_X und $C \subseteq \mathbb{R}$ abzählbar mit $P_X[C] = 1$. Besitzt X einen Erwartungswert, so gilt*

$$E[X] = \sum_{x \in C} x\, f(x)$$

Beweis. Es gilt

$$E[X] = \int_\Omega X(\omega)\, dP(\omega) = \int_\mathbb{R} x\, dP_X(x) = \int_\mathbb{R} x\, f(x)\, d\boldsymbol{\zeta}_C(x) = \sum_{x \in C} x\, f(x)$$

Damit ist die Behauptung gezeigt. □

12.3.6 Beispiele (Diskrete Verteilungen).

(1) **Hypergeometrische Verteilung:** Im Fall $P_X = \mathbf{H}(n, N, K)$ gilt

$$E[X] = n\, \frac{K}{N}$$

(2) **Binomial–Verteilung:** Im Fall $P_X = \mathbf{B}(n, \vartheta)$ gilt

$$E[X] = n\, \vartheta$$

(3) **Poisson–Verteilung:** Im Fall $P_X = \mathbf{P}(\alpha)$ gilt

$$E[X] = \alpha$$

Damit ist die Bedeutung des Parameters der Poisson–Verteilung geklärt.
In der Tat: Es gilt

$$E[X] = \sum_{k=0}^\infty k\, e^{-\alpha}\, \frac{\alpha^k}{k!} = \alpha \sum_{k=1}^\infty e^{-\alpha}\, \frac{\alpha^{k-1}}{(k-1)!} = \alpha \sum_{l=0}^\infty e^{-\alpha}\, \frac{\alpha^l}{l!} = \alpha$$

(4) **Negativbinomial–Verteilung:** Im Fall $P_X = \mathbf{NB}(\alpha, \vartheta)$ gilt

$$E[X] = \alpha \, \frac{\vartheta}{1-\vartheta}$$

(5) **Geometrische Verteilung:** Im Fall $P_X = \mathbf{Geo}(n, \vartheta)$ gilt

$$E[X] = n \, \frac{1}{\vartheta}$$

In den Beispielen 12.3.6 stellt die Existenz des Erwartungswertes kein Problem dar, da die reellen Zufallsvariablen positiv sind; vgl. aber Aufgabe 12.3.B.

Der Erwartungswert einer reellen Zufallsvariablen mit einer absolutstetigen Verteilung kann, im Fall seiner Existenz, durch Integration bestimmt werden:

12.3.7 Lemma. *Sei X eine reelle Zufallsvariable mit einer absolutstetigen Verteilung und sei f eine Lebesgue–Dichte von P_X. Besitzt X einen Erwartungswert, so gilt*

$$E[X] = \int_{\mathbb{R}} x \, f(x) \, d\boldsymbol{\lambda}(x)$$

Beweis. Die Behauptung folgt unmittelbar aus der Kettenregel. □

12.3.8 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Im Fall $P_X = \mathbf{U}(a, b)$ gilt

$$E[X] = \frac{a+b}{2}$$

(2) **Beta–Verteilung:** Im Fall $P_X = \mathbf{Be}(\alpha, \beta)$ gilt

$$E[X] = \frac{\alpha}{\alpha+\beta}$$

(3) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha, \gamma)$ gilt

$$E[X] = \frac{\gamma}{\alpha}$$

Damit ist insbesondere die Bedeutung des Parameters der Exponential–Verteilung $\mathbf{Exp}(\alpha) = \mathbf{Ga}(\alpha, 1)$ geklärt.
In der Tat: Es gilt

$$\begin{aligned}
E[X] &= \int_{\mathbb{R}} x \, dP_X(x) \\
&= \int_{\mathbb{R}} x \, \frac{\alpha^\gamma}{\Gamma(\gamma)} \, e^{-\alpha x} x^{\gamma-1} \, \chi_{(0,\infty)}(x) \, d\boldsymbol{\lambda}(x) \\
&= \frac{\alpha^\gamma}{\Gamma(\gamma)} \int_{\mathbb{R}} e^{-\alpha x} x^{(\gamma+1)-1} \, \chi_{(0,\infty)}(x) \, d\boldsymbol{\lambda}(x)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\alpha^\gamma}{\Gamma(\gamma)} \, \frac{\Gamma(\gamma{+}1)}{\alpha^{\gamma+1}} \int_{\mathbb{R}} \frac{\alpha^{\gamma+1}}{\Gamma(\gamma{+}1)} e^{-\alpha x} x^{(\gamma+1)-1} \, \chi_{(0,\infty)}(x) \, d\boldsymbol{\lambda}(x) \\
&= \frac{\alpha^\gamma}{\Gamma(\gamma)} \, \frac{\Gamma(\gamma{+}1)}{\alpha^{\gamma+1}} \\
&= \frac{\gamma}{\alpha}
\end{aligned}
$$

(4) **Normal–Verteilung:** Im Fall $P_X = \mathbf{N}(\mu, \sigma^2)$ gilt

$$E[X] = \mu$$

Damit ist die Bedeutung des ersten Parameters der Normal–Verteilung geklärt. In der Tat: Wir betrachten die Zufallsvariable $Z := (X - \mu)/\sigma$. Nach Beispiel 12.2.7 gilt $P_Z = \mathbf{N}(0,1)$ und damit

$$
\begin{aligned}
E[Z^+] &= \int_{\mathbb{R}} z^+ \, dP_Z(z) \\
&= \int_{\mathbb{R}} z^+ \frac{1}{\sqrt{2\pi}} \, e^{-\frac{1}{2}z^2} \, d\boldsymbol{\lambda}(z) \\
&= \frac{1}{\sqrt{2\pi}} \int_{(0,\infty)} z \, e^{-\frac{1}{2}z^2} \, d\boldsymbol{\lambda}(z) \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty z \, e^{-\frac{1}{2}z^2} \, dz \\
&= \frac{1}{\sqrt{2\pi}}
\end{aligned}
$$

Nach Aufgabe 12.2.F gilt $P_{-Z} = P_Z$ und damit $E[Z^-] = E[(-Z)^+] = E[Z^+]$. Daher besitzt Z einen endlichen Erwartungswert und es gilt $E[Z] = 0$. Die Behauptung folgt nun aus $X = \mu + \sigma Z$ und der Linearität des Integrals.

(5) t–**Verteilung:** Im Fall $P_X = \mathbf{t}_1$ besitzt X keinen Erwartungswert. In der Tat: Es gilt

$$
\begin{aligned}
E[X^+] &= \int_{\mathbb{R}} x^+ \, dP_X(x) \\
&= \int_{\mathbb{R}} x^+ \, \frac{1}{\pi} \, \frac{1}{1+x^2} \, d\boldsymbol{\lambda}(x) \\
&\ge \int_{(1,\infty)} x \, \frac{1}{\pi} \, \frac{1}{x^2+x^2} \, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{2\pi} \int_{(1,\infty)} \frac{1}{x} \, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{2\pi} \int_1^\infty \frac{1}{x} \, dx \\
&= \infty
\end{aligned}
$$

Nach Aufgabe 12.2.F gilt $P_{-X} = P_X$ und damit $E[X^-] = E[(-X)^+] = E[X^+]$. Daher besitzt X keinen Erwartungswert.

Für jede Zufallsvariable X gilt $E[X^+] \leq E[|X|]$ und $E[X^-] \leq E[|X|]$; besitzt X einen Erwartungswert, so folgt daraus

$$|E[X]| \leq E[|X|]$$

Diese Ungleichung lässt sich wie folgt verallgemeinern:

12.3.9 Satz (Ungleichung von Jensen). *Sei $J \subseteq \mathbb{R}$ ein Intervall und sei X eine Zufallsvariable mit $P[\{X \in J\}] = 1$ und einem endlichen Erwartungswert. Ist $h : J \to \mathbb{R}$ konvex, so besitzt $h \circ X$ einen Erwartungswert und es gilt*

$$h(E[X]) \leq E[h \circ X]$$

Beweis. Wir zeigen zunächst, dass $E[X] \in J$ gilt. Nach Voraussetzung gibt es $a, b \in \bar{\mathbb{R}}$ derart, dass J mit einem der Intervalle $(a, b), (a, b], [a, b), [a, b]$ übereinstimmt, und aus der Monotonie des Erwartungswertes folgt zunächst $E[X] \in [a, b]$.
Im Fall $E[X] = a$ gilt $a \in \mathbb{R}$ und $E[X - a] = 0$, und wegen $P[\{X - a \geq 0\}] = 1$ folgt aus Lemma 8.2.8 $P[\{X - a = 0\}] = 1$ und damit $P[\{X = a\}] = 1$; dann gilt aber $a \in J$ und $P[\{h \circ X = h(a)\}] = 1$, und damit $E[h \circ X] = h(a)$, also

$$h(E[X]) = E[h(X)]$$

Diese Gleichung gilt auch im Fall $E[X] = b$.
Wir nehmen nun an, dass $E[X] \in (a, b)$ gilt. Da h konvex ist, gibt es eine affine Funktion $g : \mathbb{R} \to \mathbb{R}$ mit

$$g(E[X]) = h(E[X])$$

und

$$g(x) \leq h(x)$$

für alle $x \in (a, b)$. Da g affin ist, gibt es $c, d \in \mathbb{R}$ mit

$$g(x) = c + dx$$

Mit X besitzt auch $g \circ X = c + dX$ einen endlichen Erwartungswert, und wegen $-h(x) \leq -g(x)$ gilt

$$E[(h \circ X)^-] \leq E[(g \circ X)^-] \leq E[|g \circ X|]$$

Daher besitzt $h \circ X$ einen Erwartungswert, und wegen $g(x) \leq h(x)$ gilt

$$h(E[X]) = g(E[X]) = c + dE[X] = E[c+dX] = E[g(X)] \leq E[h(X)]$$

Damit ist der Satz bewiesen. □

12.3.10 Folgerung (Ungleichung von Jensen). *Sei $J \subseteq \mathbb{R}$ ein Intervall und sei X eine Zufallsvariable mit $P[\{X \in J\}] = 1$ und einem endlichen Erwartungswert. Ist $h : J \to \mathbb{R}$ konkav, so besitzt $h \circ X$ einen Erwartungswert und es gilt*

$$E[h \circ X] \leq h(E[X])$$

Wir betrachten nun Momente höherer Ordnung. Für eine Zufallsvariable X und $n \in \mathbb{N}$ heißt $E[|X|^n]$ das *n–te absolute Moment* oder das *absolute Moment der Ordnung n* von X und im Fall der Existenz heißt $E[X^n]$ das *n–te Moment* oder das *Moment der Ordnung n* von X. Diese *Momente höherer Ordnung* sind unter anderem für die Abschätzung von Wahrscheinlichkeiten durch die aus der Integrationstheorie bekannte Ungleichung von Markov von Interesse:

12.3.11 Lemma (Ungleichung von Markov). *Sei X eine Zufallsvariable. Dann gilt für alle $c \in (0, \infty)$ und $n \in \mathbb{N}$*

$$P[\{|X| \geq c\}] \leq \frac{E[|X|^n]}{c^n}$$

Des weiteren lässt sich auch der Erwartungswert einer Zufallsvariablen mit einem endlichen Erwartungswert durch höhere Momente abschätzen:

12.3.12 Lemma. *Sei X eine Zufallsvariable mit einem endlichen Erwartungswert und sei $n \in \mathbb{N}$. Dann gilt $|E[X]|^n \leq E[|X|^n]$.*

Beweis. Die Funktion $h : \mathbb{R} \to \mathbb{R}$ mit $h(x) := |x|^n$ ist konvex. Daher folgt die Behauptung aus der Ungleichung von Jensen. □

Da jedes Wahrscheinlichkeitsmaß endlich ist, besitzt jede konstante Zufallsvariable einen endlichen Erwartungswert. Daraus ergibt sich ein allgemeines Ergebnis über die Endlichkeit höherer Momente einer Zufallsvariablen:

12.3.13 Lemma. *Sei X eine Zufallsvariable und seien $m, n \in \mathbb{N}$ mit $m \leq n$. Besitzt X ein endliches Moment der Ordnung n, so besitzt X auch ein endliches Moment der Ordnung m.*

Beweis. Für alle $x \in \mathbb{R}$ gilt $|x|^m \leq 1 + |x|^n$. Daher gilt

$$E[|X|^m] \leq E[1+|X|^n] = 1 + E[|X|^n]$$

und die Behauptung folgt. □

Diese Ungleichungen bleiben gültig, wenn man $m, n \in \mathbb{N}$ durch $p, q \in [1, \infty)$ mit $p \leq q$ ersetzt.

12.3.14 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Im Fall $P_X = \mathbf{U}(a,b)$ gilt für alle $n \in \mathbb{N}$

$$E[X^n] = \frac{1}{n+1}\,\frac{b^{n+1}-a^{n+1}}{b-a}$$

(2) **Beta–Verteilung:** Im Fall $P_X = \mathbf{Be}(\alpha,\beta)$ gilt für alle $n \in \mathbb{N}$

$$E[X^n] = \prod_{k=0}^{n-1} \frac{\alpha+k}{\alpha+\beta+k}$$

(3) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha,\gamma)$ gilt für alle $n \in \mathbb{N}$

$$E[X^n] = \prod_{k=0}^{n-1} \frac{\gamma+k}{\alpha}$$

(4) t–**Verteilung:** Im Fall $P_X = \mathbf{t}_2$ gilt $E[X] = 0$ und $E[|X|^n] = \infty$ für alle $n \geq 2$. In der Tat: Es gilt

$$\begin{aligned} E[X^+] &= \int_{\mathbb{R}} x^+ \, dP_X(x) \\ &= \int_{\mathbb{R}} x^+ \, \frac{1}{(2+x^2)^{3/2}} \, d\boldsymbol{\lambda}(x) \\ &= \int_{\mathbb{R}} \frac{x}{(2+x^2)^{3/2}} \, \chi_{(0,\infty)}(x) \, d\boldsymbol{\lambda}(x) \end{aligned}$$

Wir betrachten das Maß

$$\nu := \int \frac{x}{(2+x^2)^{3/2}} \, \chi_{(0,\infty)}(x) \, d\boldsymbol{\lambda}(x)$$

und die Abbildung $T : \mathbb{R} \to \mathbb{R}$ mit $T(x) := 2 + x^2$. Dann gilt $\nu[\mathbb{R} \setminus (0,\infty)] = 0$, T ist messbar, und die Restriktion $T|_{(0,\infty)} : (0,\infty) \to \mathbb{R}$ ist stetig differenzierbar mit $(T|_{(0,\infty)})'(x) = 2x \neq 0$ für alle $x \in (0,\infty)$. Aus Satz 12.2.4 erhalten wir nun

$$\begin{aligned} \nu_T &= \int \frac{(z-2)^{1/2}}{z^{3/2}} \, \chi_{(0,\infty)}((z-2)^{1/2}) \, \frac{1}{2} \, (z-2)^{-1/2} \, \chi_{(2,\infty)}(z) \, d\boldsymbol{\lambda}(z) \\ &= \int \frac{1}{2} \, z^{-3/2} \, \chi_{(2,\infty)}(z) \, d\boldsymbol{\lambda}(z) \end{aligned}$$

und damit

$$\begin{aligned} E[X^+] &= \nu[\mathbb{R}] \\ &= \nu_T[\mathbb{R}] \\ &= \int_{\mathbb{R}} \frac{1}{2} \, z^{-3/2} \, \chi_{(2,\infty)}(z) \, d\boldsymbol{\lambda}(z) \\ &= \int_{(2,\infty)} \frac{1}{2} \, z^{-3/2} \, d\boldsymbol{\lambda}(z) \end{aligned}$$

$$\begin{aligned}
&= \int_2^\infty \frac{1}{2}\, z^{-3/2}\, dz \\
&= \frac{1}{\sqrt{2}}
\end{aligned}$$

Nach Aufgabe 12.2.F gilt $P_{-X} = P_X$ und damit $E[X^-] = E[(-X)^+] = E[X^+]$, Daher besitzt X einen endlichen Erwartungswert und es gilt $E[X] = 0$. Des weiteren gilt

$$\begin{aligned}
E[2+X^2] &= \int_{\mathbb{R}} (2+x^2)\, dP_X(x) \\
&= \int_{\mathbb{R}} (2+x^2)\, \frac{1}{(2+x^2)^{3/2}}\, d\boldsymbol{\lambda}(x) \\
&= \int_{\mathbb{R}} \frac{1}{(2+x^2)^{1/2}}\, d\boldsymbol{\lambda}(x) \\
&\geq \int_{(1,\infty)} \frac{1}{(3x^2+x^2)^{1/2}}\, d\boldsymbol{\lambda}(x) \\
&= \frac{1}{2} \int_{(1,\infty)} \frac{1}{x}\, d\boldsymbol{\lambda}(x) \\
&= \infty
\end{aligned}$$

und damit $E[X^2] = \infty$. Die Behauptung über $E[|X|^n]$ folgt nun aus Lemma 12.3.13.

Aufgaben

12.3.A Geben Sie einen elementaren Beweis für die Doppelungleichung aus Lemma 12.3.1. Lässt sich die Verwendung des Satzes von Fubini vermeiden?

12.3.B Betrachten Sie den Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $\Omega := \mathbb{N}$ und $\mathcal{F} := 2^\Omega$ sowie $P[\{\omega\}] := 2^{-\omega}$ für alle $\omega \in \Omega$. Konstruieren Sie eine Zufallsvariable $X : \Omega \to \mathbb{R}$ mit $E[X^+] = \infty = E[X^-]$.

12.3.C Sei $(\Omega, \mathcal{F}, P)$ symmetrisch mit $\Omega := \{1, 2, 3, 4, 5, 6\}^3$ und sei $X : \Omega \to \mathbb{R}$ gegeben durch $X((i, j, k)) := i + j + k$. Dann besitzt X einen endlichen Erwartungswert, der sich gemäß

$$E[X] = \int_\Omega X(\omega)\, dP(\omega)$$

oder gemäß

$$E[X] = \int_{\mathbb{R}} x\, dP_X(x)$$

bestimmen lässt. Vergleichen Sie beide Rechnungen.

12.3.D **Pólya–Verteilung:** Sei $P_X = \textbf{Pólya}(n, \alpha, \beta)$. Dann gilt

$$E[X] = n\, \frac{\alpha}{\alpha + \beta}$$

12.3.E **Logarithmische Verteilung:** Sei $P_X = \mathbf{Log}(\vartheta)$. Dann gilt

$$E[X] = \frac{1}{|\log(1-\vartheta)|}\,\frac{\vartheta}{1-\vartheta}$$

12.3.F **Cauchy–Verteilung:** Sei $P_X = \mathbf{Ca}(\alpha,\beta)$. Dann besitzt X keinen Erwartungswert.

12.3.G **Pareto–Verteilung:** Sei $P_X = \mathbf{Pa}(\alpha,\beta)$. Dann gilt für alle $n \in \mathbb{N}$ mit $n < \beta$

$$E[X^n] = \frac{\beta}{\beta - n}\,\alpha^n$$

und für alle $n \in \mathbb{N}$ mit $n \geq \beta$ gilt $E[X^n] = \infty$.

12.3.H **Pareto–Verteilung:** Sei $P_X = \mathbf{Pa}^*(\alpha,\beta)$. Dann gilt für alle $n \in \mathbb{N}$ mit $n < \beta$

$$E[X^n] = \alpha^n \sum_{k=0}^{n} \binom{n}{k} \frac{\beta}{\beta - k}\,(-1)^{n-k}$$

und für alle $n \in \mathbb{N}$ mit $n \geq \beta$ gilt $E[X^n] = \infty$.

12.3.I **Symmetrische Verteilungen:**

(1) Geben Sie ein Beispiel für eine Zufallsvariable mit einer symmetrischen Verteilung, die keinen Erwartungswert besitzt.

(2) Sei X eine Zufallsvariable mit einer symmetrischen Verteilung. Besitzt X ein endliches Moment der Ordnung $2k+1$ für ein $k \in \mathbb{N}_0$, so gilt $E[X^{2k+1}] = 0$.

12.3.J **Integrierte Überlebensfunktion:** Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung mit $\int_{\mathbb{R}} |x|\,dQ(x) < \infty$ und sei G die zugehörige Überlebensfunktion. Dann gilt für alle $a \in \mathbb{R}$

$$\int_{\mathbb{R}} (x-a)^+\,dQ(x) = \int_{(a,\infty)} G(x)\,d\boldsymbol{\lambda}(x)$$

Die Abbildung $\mathbb{R} \to \mathbb{R} : a \mapsto \int_{(a,\infty)} G(x)\,d\boldsymbol{\lambda}(x)$ heißt *integrierte Überlebensfunktion* zu Q.

12.3.K **Stop–Loss Ordnung:** Sei $\mathcal{Q}^1(\mathbb{R})$ die Familie aller Verteilungen $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ mit $\int_{\mathbb{R}} |x|\,dQ(x) < \infty$. Für $Q_1, Q_2 \in \mathcal{Q}^1(\mathbb{R})$ schreiben wir

$$Q_1 \leq^1 Q_2$$

wenn für alle $a \in \mathbb{R}$

$$\int_{\mathbb{R}} (x-a)^+\,dQ_1(x) \leq \int_{\mathbb{R}} (x-a)^+\,dQ_2(x)$$

gilt. Die Relation $\leq^1$ wird als *stop–loss Ordnung* bezeichnet.

(1) Die Relation $\leq^1$ ist eine Ordnungsrelation auf $\mathcal{Q}^1(\mathbb{R})$.
(2) Für $Q_1, Q_2 \in \mathcal{Q}^1(\mathbb{R})$ und die zugehörigen Überlebensfunktionen G_1, G_2 gilt $Q_1 \leq^1 Q_2$ genau dann, wenn für alle $a \in \mathbb{R}$

$$\int_{(a,\infty)} G_1(x)\, d\boldsymbol{\lambda}(x) \leq \int_{(a,\infty)} G_2(x)\, d\boldsymbol{\lambda}(x)$$

gilt.
(3) Für alle $Q_1, Q_2 \in \mathcal{Q}^1(\mathbb{R})$ mit $Q_1 \leq^0 Q_2$ gilt $Q_1 \leq^1 Q_2$.
(4) Verteilungen $Q_1, Q_2 \in \mathcal{Q}^1(\mathbb{R})$ mit

$$\int_{\mathbb{R}} x\, dQ_1(x) = \int_{\mathbb{R}} x\, dQ_2(x)$$

sind bezüglich der stochastischen Ordnung nicht vergleichbar.
(5) Untersuchen Sie die Vergleichbarkeit der Verteilungen $\mathbf{B}(2, \frac{1}{2})$ und $\mathbf{B}(3, \frac{1}{3})$ bezüglich der stop–loss Ordnung.

12.3.L **Bedingter Erwartungswert:** Sei $C \in \mathcal{F}$ ein Ereignis mit $P[C] \in (0, 1)$. Für eine Zufallsvariable X, die positiv oder $P[\,.\,|C]$–integrierbar ist, setzen wir

$$E[X|C] := \int_\Omega X(\omega)\, dP[\,.\,|C](\omega)$$

und nennen $E[X|C]$ den *bedingten Erwartungswert* von X unter C. Es gilt

$$E[X|C] = \frac{1}{P[C]} \int_C X(\omega)\, dP(\omega)$$

Ist X positiv oder P–integrierbar, so gilt

$$E[X] = E[X|C]\, P[C] + E[X|\overline{C}]\, P[\overline{C}]$$

Interpretieren Sie diese Gleichung.

12.3.M Jede Zufallsvariable in $L^\infty(\mathcal{F}, P)$ besitzt endliche Momente beliebiger Ordnung.

12.4 Zentrale Momente

Für eine Zufallsvariable X mit einem endlichen Erwartungswert ist neben dem Erwartungswert vor allem die Abweichung $X - E[X]$ der Zufallsvariablen von ihrem Erwartungswert von Interesse. Die Zufallsvariable $X - E[X]$ heißt die zu X *zentrierte Zufallsvariable* und es gilt

$$E[X - E[X]] = 0$$

Diese Gleichung wird auch als *Schwerpunkteigenschaft* des Erwartungswertes bezeichnet.

Für eine Zufallsvariable X mit einem endlichen Erwartungswert setzen wir

$$\mathrm{var}\,[X] := E[(X - E[X])^2]$$

und nennen $\mathrm{var}\,[X]$ die *Varianz* von X. Die Varianz von X ist genau dann endlich, wenn X ein endliches zweites Moment besitzt. Sie dient als Maß der Abweichung der Zufallsvariablen von ihrem Erwartungswert und wird daher auch als ein *Streuungsmaß* bezeichnet.

12.4.1 Lemma. *Sei X eine Zufallsvariable mit einem endlichen Erwartungswert.*
(1) *Es gilt* $\mathrm{var}\,[X] = E[X^2] - (E[X])^2$.
(2) *Es gilt* $\mathrm{var}\,[X] = 0$ *genau dann, wenn* $P[\{X = E[X]\}] = 1$ *gilt.*
(3) *Es gilt* $\mathrm{var}\,[X] = \inf_{c\in\mathbb{R}} E[(X - c)^2]$.
(4) *Für alle* $a, b \in \mathbb{R}$ *gilt* $\mathrm{var}\,[a+bX] = b^2\,\mathrm{var}\,[X]$.

Beweis. Im Fall $E[X^2] < \infty$ folgt aus der Linearität des Integrals

$$\begin{aligned}\mathrm{var}\,[X] &= E[(X - E[X])^2] \\ &= E[X^2 - 2E[X]\,X + (E[X])^2] \\ &= E[X^2] - 2E[X]\,E[X] + (E[X])^2 \\ &= E[X^2] - (E[X])^2\end{aligned}$$

und diese Gleichung gilt auch im Fall $E[X^2] = \infty$. Damit ist (1) gezeigt. Wegen $(X - E[X])^2 \geq 0$ folgt (2) aus Lemma 8.2.8. Sei nun $E[X^2] < \infty$. Für die Funktion $g : \mathbb{R} \to \mathbb{R}$ mit

$$g(c) := E[(X - c)^2]$$

gilt

$$\begin{aligned}g(c) &= E[X^2] - 2c\,E[X] + c^2 \\ g'(c) &= -\,2\,E[X] + 2c \\ g''(c) &= 2\end{aligned}$$

Daher ist $c^* := E[X]$ der eindeutig bestimmte Minimierer von g. Es gilt also

$$E[(X - E[X])^2] = \inf_{c\in\mathbb{R}} E[(X - c)^2]$$

und diese Gleichung gilt auch im Fall $E[X^2] = \infty$. Damit ist (3) gezeigt. Mit X besitzt auch $a + bX$ einen endlichen Erwartungswert und es gilt

$$\begin{aligned}\mathrm{var}\,[a+bX] &= E\Big[\Big((a+bX) - E[a+bX]\Big)^2\Big] \\ &= E\Big[\Big((a+bX) - (a+b\,E[X])\Big)^2\Big]\end{aligned}$$

$$\begin{aligned}&= E[b^2\,(X-E[X])^2]\\&= b^2\,E[(X-E[X])^2]\\&= b^2\,\mathrm{var}\,[X]\end{aligned}$$

Damit ist (4) gezeigt. □

Für eine Zufallsvariable X mit $P_X[\mathbb{N}_0] = 1$ und einem endlichen Erwartungswert ist es oft vorteilhaft, die Varianz mit Hilfe der Gleichung

$$\mathrm{var}\,[X] = E[X(X-1)] + E[X] - (E[X])^2$$

zu berechnen.

12.4.2 Beispiele (Diskrete Verteilungen).

(1) **Hypergeometrische Verteilung:** Im Fall $P_X = \mathbf{H}(n,N,K)$ gilt

$$\mathrm{var}\,[X] = n\,\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}$$

Der Faktor $(N-n)/(N-1)$ heißt *Korrekturfaktor* (im Vergleich zur Varianz der Binomial–Verteilung $\mathbf{B}(n,K/N)$).

(2) **Binomial–Verteilung:** Im Fall $P_X = \mathbf{B}(n,\vartheta)$ gilt

$$\mathrm{var}\,[X] = n\,\vartheta(1-\vartheta)$$

(3) **Poisson–Verteilung:** Im Fall $P_X = \mathbf{P}(\alpha)$ gilt

$$\mathrm{var}\,[X] = \alpha$$

In der Tat: Es gilt

$$E[X(X-1)] = \sum_{k=0}^{\infty} k(k-1)\,e^{-\alpha}\,\frac{\alpha^k}{k!} = \alpha^2 \sum_{k=2}^{\infty} e^{-\alpha}\,\frac{\alpha^{k-2}}{(k-2)!} = \alpha^2 \sum_{l=0}^{\infty} e^{-\alpha}\,\frac{\alpha^l}{l!} = \alpha^2$$

und damit

$$\mathrm{var}\,[X] = E[X(X-1)] + E[X] - (E[X])^2 = \alpha^2 + \alpha - \alpha^2 = \alpha$$

(4) **Negativbinomial–Verteilung:** Im Fall $P_X = \mathbf{NB}(\alpha,\vartheta)$ gilt

$$\mathrm{var}\,[X] = \alpha\,\frac{\vartheta}{(1-\vartheta)^2}$$

(5) **Geometrische Verteilung:** Im Fall $P_X = \mathbf{Geo}(n,\vartheta)$ gilt

$$\mathrm{var}\,[X] = n\,\frac{1-\vartheta}{\vartheta^2}$$

12.4.3 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Im Fall $P_X = \mathbf{U}(a,b)$ gilt

$$\mathrm{var}\,[X] = \frac{(b-a)^2}{12}$$

(2) **Beta–Verteilung:** Im Fall $P_X = \mathbf{Be}(\alpha,\beta)$ gilt

$$\mathrm{var}\,[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

(3) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha,\gamma)$ gilt

$$\mathrm{var}\,[X] = \frac{\gamma}{\alpha^2}$$

(4) **Normal–Verteilung:** Im Fall $P_X = \mathbf{N}(\mu,\sigma^2)$ gilt

$$\mathrm{var}\,[X] = \sigma^2$$

Damit ist die Bedeutung des zweiten Parameters der Normal–Verteilung geklärt.
In der Tat: Wir betrachten die Zufallsvariable

$$Z := \frac{X-\mu}{\sigma}$$

Nach Beispiel 12.2.7 gilt $P_Z = \mathbf{N}(0,1)$ und $P_{Z^2} = \boldsymbol{\chi}^2_1 = \mathbf{Ga}(\frac{1}{2},\frac{1}{2})$, und aus Beispiel 12.3.8 folgt nun $E[Z]=0$ und $E[Z^2]=1$, und damit $\mathrm{var}\,[Z]=1$. Die Behauptung folgt nun aus $X = \mu+\sigma Z$ und den Eigenschaften der Varianz.

Mit Hilfe der Varianz lassen sich Wahrscheinlichkeiten für Abweichungen einer Zufallsvariablen von ihrem Erwartungswert abschätzen:

12.4.4 Lemma (Ungleichung von Tschebyschev). *Sei X eine Zufallsvariable mit einem endlichen zweiten Moment. Dann gilt für alle $c \in (0,\infty)$*

$$P\Big[\Big\{|X-E[X]| \geq c\Big\}\Big] \leq \frac{\mathrm{var}\,[X]}{c^2}$$

und

$$P\Big[\Big\{|X-E[X]| \leq c\Big\}\Big] \geq 1 - \frac{\mathrm{var}\,[X]}{c^2}$$

Beweis. Die erste Ungleichung ergibt sich unmittelbar aus der Ungleichung von Markov, angewendet auf die zentrierte Zufallsvariable $X - E[X]$. Wegen

$$P\Big[\Big\{|X-E[X]| \leq c\Big\}\Big] \geq P\Big[\Big\{|X-E[X]| < c\Big\}\Big] = 1 - P\Big[\Big\{|X-E[X]| \geq c\Big\}\Big]$$

folgt die zweite Ungleichung aus der ersten. □

Während die Ungleichung von Tschebyschev eine zweiseitige Abschätzung für Abweichungen vom Erwartungswert gibt, liefert die folgende Ungleichung von Cantelli eine einseitige Abschätzung:

12.4.5 Lemma (Ungleichung von Cantelli). *Sei X eine Zufallsvariable mit einem endlichen zweiten Moment. Dann gilt für alle $c \in (0, \infty)$*

$$P\Big[\Big\{X \geq E[X] + c\Big\}\Big] \leq \frac{\text{var}\,[X]}{c^2 + \text{var}\,[X]}$$

Beweis. Sei $Z := X - E[X]$. Dann gilt $E[Z] = 0$ und $\text{var}\,[Z] = \text{var}\,[X]$. Für alle $t \in (-c, \infty)$ gilt $c + t > 0$ und aus der Ungleichung von Markov folgt nun

$$\begin{aligned}
P\Big[\Big\{X - E[X] \geq c\Big\}\Big] &= P\Big[\Big\{Z \geq c\Big\}\Big] \\
&= P\Big[\Big\{Z + t \geq c + t\Big\}\Big] \\
&\leq P\Big[\Big\{|Z + t| \geq c + t\Big\}\Big] \\
&\leq \frac{E[(Z+t)^2]}{(c+t)^2} \\
&= \frac{E[Z^2] + t^2}{(c+t)^2} \\
&= \frac{\text{var}\,[Z] + t^2}{(c+t)^2} \\
&= \frac{\text{var}\,[X] + t^2}{(c+t)^2}
\end{aligned}$$

Wir betrachten nun die Funktion $g : (-c, \infty) \to \mathbb{R}$ mit

$$g(t) := \frac{\text{var}\,[X] + t^2}{(c+t)^2}$$

Dann gilt

$$\begin{aligned}
g'(t) &= 2\,\frac{ct - \text{var}\,[X]}{(c+t)^3} \\
g''(t) &= 2\,\frac{t^2 - 2ct + 3\text{var}\,[X]}{(c+t)^4}
\end{aligned}$$

Daher ist $t^* := \text{var}\,[X]/c$ der eindeutig bestimmte Minimierer von g. Es gilt

$$g(t^*) = \frac{\text{var}\,[X]}{c^2 + \text{var}\,[X]}$$

Daraus folgt die Behauptung. $\square$

Der Beweis der Ungleichung von Cantelli ist ein Beispiel dafür, wie man für bestimmte Wahrscheinlichkeiten zunächst aus der Ungleichung von Markov eine ganze Schar von oberen Schranken gewinnt und sodann das Infimum dieser oberen Schranken bestimmt.

Für eine Zufallsvariable X mit einem endlichen zweiten Moment betrachtet man neben der Varianz auch die *Standardabweichung*

$$\sqrt{\mathrm{var}\,[X]} = \Bigl(E[(X-E[X])^2]\Bigr)^{1/2}$$

Für alle $c \in (0,\infty)$ gilt $\sqrt{\mathrm{var}\,[cX]} = c\sqrt{\mathrm{var}\,[X]}$. Daher ist die Standardabweichung ein *Streuungsmaß*, das dieselbe Dimension wie die Zufallsvariable besitzt.

Die Varianz einer Zufallsvariablen X mit einem endlichen Erwartungswert wird auch als *zweites zentrales Moment* von X bezeichnet. Daneben betrachtet man auch zentrale Momente höherer Ordnung: Für eine Zufallsvariable X mit einem endlichen Erwartungswert und für $n \in \mathbb{N}$ heißt das n–te Moment $E[(X-E[X])^n]$ der zentrierten Zufallsvariablen $X-E[X]$ im Fall der Existenz das *n–te zentrale Moment* oder das *zentrale Moment der Ordnung n* von X; vgl. Aufgaben 12.4.H und 12.4.I. Das n–te zentrale Moment von X existiert für alle $n \in 2\mathbb{N}$.

Aufgaben

12.4.A Sei X eine Zufallsvariable mit einem endlichen Erwartungswert. Dann gilt für alle $c \in \mathbb{R}$

$$E[(X-c)^2] = (E[X]-c)^2 + \mathrm{var}\,[X]$$

12.4.B **Pólya–Verteilung:** Sei $P_X = \mathbf{Pólya}(n,\alpha,\beta)$. Dann gilt

$$\mathrm{var}\,[X] = n\,\frac{\alpha\beta}{(\alpha+\beta)^2}\,\frac{\alpha+\beta+n}{\alpha+\beta+1}$$

12.4.C **Logarithmische Verteilung:** Sei $P_X = \mathbf{Log}(\vartheta)$. Dann gilt

$$\mathrm{var}\,[X] = \frac{|\log(1-\vartheta)|-\vartheta}{|\log(1-\vartheta)|^2}\,\frac{\vartheta}{(1-\vartheta)^2}$$

12.4.D **k–σ–Bereich:** Sei X eine Zufallsvariable mit einem endlichen zweiten Moment und mit Erwartungswert μ und Standardabweichung σ. Für $k \in \mathbb{N}$ wird das Intervall $[\mu-k\sigma, \mu+k\sigma]$ als *k–σ–Bereich* (der Verteilung) von X bezeichnet. Bestimmen Sie eine nur von k abhängige universelle untere Schranke für die Wahrscheinlichkeit

$$P\Bigl[\Bigl\{X \in [\mu-k\sigma, \mu+k\sigma]\Bigr\}\Bigr]$$

und berechnen Sie diese untere Schranke für $k \in \{1,2,3\}$.

12.4.E **Standardabweichung:** Sei X eine Zufallsvariable mit einem endlichen zweiten Moment und $E[X] = 0$. Dann gilt

$$\sqrt{\mathrm{var}\,[X]} = \|X\|_2$$

12.4.F **Variationskoeffizient:** Sei X eine positive Zufallsvariable mit einem endlichen zweiten Moment und $E[X] > 0$. Dann heißt

$$v[X] := \frac{\sqrt{\mathrm{var}\,[X]}}{E[X]}$$

Variationskoeffizient (der Verteilung) von X. Für alle $c \in (0, \infty)$ gilt $v[cX] = v[X]$. Daher ist der Variationskoeffizient ein dimensionsloses Maß für die Streuung von X.

12.4.G **Standardisierung:** Sei X eine Zufallsvariable mit einem endlichen zweiten Moment und mit Erwartungswert μ und Standardabweichung $\sigma > 0$. Dann heißt die Zufallsvariable

$$Z = \frac{X - \mu}{\sigma}$$

die zu X *standardisierte Zufallsvariable*. Die Zufallsvariable Z ist dimensionslos und es gilt $E[Z] = 0$ und $\mathrm{var}\,[Z] = 1$.

12.4.H **Schiefe:** Sei X eine Zufallsvariable mit einem endlichen dritten Moment und mit Erwartungswert μ und Standardabweichung $\sigma > 0$. Dann heißt

$$E\left[\left(\frac{X - \mu}{\sigma}\right)^3\right]$$

Schiefe (der Verteilung) von X. Berechnen Sie die Schiefe für den Fall $P_X = \mathbf{N}(\mu, \sigma^2)$.

12.4.I **Exzess:** Sei X eine Zufallsvariable mit einem endlichen vierten Moment und mit Erwartungswert μ und Standardabweichung $\sigma > 0$. Dann heißt

$$E\left[\left(\frac{X - \mu}{\sigma}\right)^4\right] - 3$$

Exzess (der Verteilung) von X. Berechnen Sie den Exzess für den Fall $P_X = \mathbf{N}(\mu, \sigma^2)$.

12.4.J Sei X eine Zufallsvariable mit einem endlichen Moment der Ordnung n. Dann gilt

$$E[(X - E[X])^n] = \sum_{k=1}^{n} \binom{n}{k} E[X^k]\,(-E[X])^{n-k}$$

und

$$E[X^n] = \sum_{k=1}^{n} \binom{n}{k} E[(X - E[X])^k]\,(E[X])^{n-k}$$

12.4.K Uniforme Verteilung: Im Fall $P_X = \mathbf{U}(a,b)$ gilt

$$E[(X-E[X])^n] = \begin{cases} \dfrac{1}{n+1}\left(\dfrac{b-a}{2}\right)^n & \text{falls } n \in 2\mathbb{N} \\ 0 & \text{sonst} \end{cases}$$

12.4.L Normal–Verteilung: Im Fall $P_X = \mathbf{N}(\mu,\sigma^2)$ gilt

$$E[(X-E[X])^n] = \begin{cases} \sigma^n \displaystyle\prod_{j=1}^{n/2}(2j-1) & \text{falls } n \in 2\mathbb{N} \\ 0 & \text{sonst} \end{cases}$$

13

Multivariate Verteilungen

Multivariate Verteilungen sind Wahrscheinlichkeitsmaße auf $\mathcal{B}(\mathbb{R}^m)$. Sie sind von Interesse, weil für jeden Zufallsvektor $\mathbf{X} : \Omega \to \mathbb{R}^m$ das Bildmaß $P_\mathbf{X}$ von P unter $\mathbf{X}$ eine multivariate Verteilung ist.

In diesem Kapitel untersuchen wir multivariate Verteilungen. Wir beginnen mit der Charakterisierung multivariater Verteilungen durch Verteilungsfunktionen und einer Reihe von Beispielen für diskrete oder absolutstetige Verteilungen (Abschnitt 13.1) und untersuchen dann Transformationen multivariater Verteilungen (Abschnitt 13.2). Für multivariate Verteilungen ergeben sich nun Besonderheiten, für die es im univariaten Fall kein Analogon gibt: Die Berechnung von Randverteilungen (Abschnitt 13.3), die besonderen Eigenschaften der Verteilung eines Zufallsvektors mit unabhängigen Koordinaten (Abschnitt 13.4) und die Berechnung der Verteilung der Summe der Koordinaten eines Zufallsvektors (Abschnitt 13.5). Die letzten Abschnitte sind wieder analog zum univariaten Fall und behandeln die Eigenschaften der Momente (Abschnitt 13.6) und der zentralen Momente (Abschnitt 13.7) einer multivariaten Verteilung. Auch in diesem Kapitel werden viele der im ersten Abschnitt eingeführten Verteilungen auch im Hinblick auf die Berechnung von Transformationen oder Momenten oder zentralen Momenten diskutiert.

13.1 Verteilungen und Verteilungsfunktionen

Ein Wahrscheinlichkeitsmaß

$$Q : \mathcal{B}(\mathbb{R}^m) \to [0, 1]$$

heißt *Verteilung* auf $\mathcal{B}(\mathbb{R}^m)$ (oder kurz *Verteilung* auf $\mathbb{R}^m$). Eine Verteilung auf $\mathcal{B}(\mathbb{R}^m)$ mit $m \geq 2$ wird auch als *multivariate Verteilung* bezeichnet. Wie im univariaten Fall lässt sich auch im multivariaten Fall jede Verteilung als Verteilung eines Zufallsvektors auf einem geeigneten Wahrscheinlichkeitsraum

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_13, © Springer-Verlag Berlin Heidelberg 2011

interpretieren, und wie im univariaten Fall lässt sich auch im multivariaten Fall jede Verteilung durch eine Verteilungsfunktion darstellen. Für die Definition einer multivariaten Verteilungsfunktion benötigen wir das folgende Hilfsmittel:

Sei $F : \mathbb{R}^m \to \mathbb{R}$ eine Funktion und sei $(\mathbf{a}, \mathbf{b}] \in \mathcal{J}(\mathbb{R}^m)$ ein nichtleeres Intervall. Für jede Menge $K \subseteq \{1, \dots, m\}$ definieren wir einen Vektor $\mathbf{c}_{(\mathbf{a},\mathbf{b}],K} \in \mathbb{R}^m$ mit den Koordinaten

$$c_{(\mathbf{a},\mathbf{b}],K,i} := \begin{cases} a_i & \text{falls } i \in K \\ b_i & \text{sonst} \end{cases}$$

und setzen

$$\triangle_F[\mathbf{a}, \mathbf{b}] := \sum_{j=0}^{m} (-1)^j \sum_{K \subseteq \{1,\dots,m\},\, |K|=j} F(\mathbf{c}_{(\mathbf{a},\mathbf{b}],K})$$

Offenbar ist $\{\mathbf{c}_{(\mathbf{a},\mathbf{b}],K}\}_{K \subseteq \{1,\dots,m\}}$ gerade die Menge der 2^m Eckpunkte des Intervalls $(\mathbf{a}, \mathbf{b}]$.

Eine Funktion $F : \mathbb{R}^m \to \mathbb{R}$ heißt *rechtecksmonoton*, wenn für jedes nichtleere Intervall $(\mathbf{a}, \mathbf{b}] \in \mathcal{J}(\mathbb{R}^m)$

$$\triangle_F[\mathbf{a}, \mathbf{b}] \geq 0$$

gilt, und sie heißt *stetig von oben*, wenn für alle $\mathbf{x} \in \mathbb{R}^m$ und jede monoton fallende Folge $\{\mathbf{x}_n\}_{n \in \mathbb{N}} \subseteq \mathbb{R}^m$ mit $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$

$$\lim_{n\to\infty} F(\mathbf{x}_n) = F(\mathbf{x})$$

gilt.

Eine Funktion $F : \mathbb{R}^m \to [0, 1]$ heißt *Verteilungsfunktion*, wenn sie folgende Eigenschaften besitzt:
(i) F ist rechtecksmonoton.
(ii) F ist stetig von oben.
(iii) Es gilt $\lim_{\mathbf{x}\to+\boldsymbol{\infty}} F(\mathbf{x}) = 1$ und für alle $\mathbf{x} \in \mathbb{R}^m$ und $i \in \{1, \dots, m\}$ gilt $\lim_{t\to-\infty} \mathbf{F}(\mathbf{x} + t\mathbf{e}_i) = 0$.

Für $m = 1$ stimmt diese Definition mit der früher gegebenen überein. Für $m \geq 2$ ist sie gerade die richtige Verallgemeinerung, wie der folgende Satz zeigt; vgl. auch Aufgabe 13.1.B.

13.1.1 Satz (Korrespondenzsatz).
(1) *Zu jeder Verteilung* $Q : \mathcal{B}(\mathbb{R}^m) \to [0, 1]$ *gibt es genau eine Verteilungsfunktion* $F_Q : \mathbb{R}^m \to [0, 1]$ *mit* $F_Q(\mathbf{x}) = Q[(-\boldsymbol{\infty}, \mathbf{x}]]$ *für alle* $\mathbf{x} \in \mathbb{R}^m$.
(2) *Zu jeder Verteilungsfunktion* $F : \mathbb{R}^m \to [0, 1]$ *gibt es genau eine Verteilung* $Q_F : \mathcal{B}(\mathbb{R}^m) \to [0, 1]$ *mit* $Q_F[(-\boldsymbol{\infty}, \mathbf{x}]] = F(\mathbf{x})$ *für alle* $\mathbf{x} \in \mathbb{R}^m$.
(3) *Es gilt* $Q_{(F_Q)} = Q$ *und* $F_{(Q_F)} = F$.

Wir verzichten auf den Beweis des Korrespondenzsatzes; der Beweis von (1) ist einfach, der Beweis von (2) verläuft ähnlich wie die Konstruktion des m–dimensionalen Lebesgue–Maßes, und (3) erhält man wie im univariaten Fall.

Diskrete Verteilungen

Eine Verteilung $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ heißt *diskret*, wenn es eine abzählbare Menge $C \subseteq \mathbb{R}^m$ gibt mit $Q[C] = 1$.

13.1.2 Lemma. *Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine diskrete Verteilung. Dann gilt $Q \perp \boldsymbol{\lambda}^m$.*

Eine Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ heißt *Zähldichte*, wenn es eine abzählbare Menge $C \subseteq \mathbb{R}^m$ gibt mit $f(\mathbf{x}) = 0$ für alle $\mathbf{x} \in \mathbb{R}^m \setminus C$ und

$$\sum_{x \in C} f(\mathbf{x}) = 1$$

Jede Zähldichte ist messbar.

Jede diskrete Verteilung lässt sich durch eine Zähldichte erzeugen, und jede Zähldichte erzeugt eine diskrete Verteilung:

13.1.3 Lemma.
(1) *Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine diskrete Verteilung. Dann ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit*

$$f(\mathbf{x}) := Q[\{\mathbf{x}\}]$$

eine Zähldichte und für alle $B \in \mathcal{B}(\mathbb{R}^m)$ gilt

$$Q[B] = \sum_{\mathbf{x} \in B} f(\mathbf{x})$$

(2) *Sei $f : \mathbb{R}^m \to \mathbb{R}_+$ eine Zähldichte. Dann ist die Mengenfunktion $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ mit*

$$Q[B] := \sum_{\mathbf{x} \in B} f(\mathbf{x})$$

eine Verteilung und für alle $\mathbf{x} \in \mathbb{R}^m$ gilt

$$Q[\{\mathbf{x}\}] = f(\mathbf{x})$$

Die in Abschnitt 12.1 für den univariaten Fall geführte Diskussion über Zähldichten und die absolute Stetigkeit einer diskreten Verteilung bezüglich einem lokalen Zählmaß lässt sich unmittelbar auf den multivariaten Fall übertragen.

Die einfachste diskrete multivariate Verteilung ist die Dirac–Verteilung:

13.1.4 Beispiel (Dirac–Verteilung). Für $\mathbf{z} \in \mathbb{R}^m$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \begin{cases} 1 & \text{falls } \mathbf{x} = \mathbf{z} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Dirac–Verteilung* und wird mit $\boldsymbol{\delta}_{\mathbf{z}}$ bezeichnet. Für die Verteilung $Q := \boldsymbol{\delta}_{\mathbf{z}}$ gilt

$$Q[B] = \begin{cases} 0 & \text{falls } \mathbf{z} \notin B \\ 1 & \text{falls } \mathbf{z} \in B \end{cases}$$

und für die zugehörige Verteilungsfunktion F gilt

$$F(\mathbf{x}) = \begin{cases} 0 & \text{falls } \mathbf{x} \not\geq \mathbf{z} \\ 1 & \text{falls } \mathbf{x} \geq \mathbf{z} \end{cases}$$

Die Funktion F wird als *Heaviside–Funktion* bezeichnet.

Wir betrachten nun einige parametrische Klassen von diskreten multivariaten Verteilungen $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ mit $Q[\mathbb{N}_0^m] = 1$:

13.1.5 Beispiele (Diskrete Verteilungen).

(1) **Polyhypergeometrische Verteilung:** Für $n, N \in \mathbb{N}$ und $\mathbf{K} \in \{1, \ldots, N{-}1\}^m$ mit $n \leq N$ und $\mathbf{1}'\mathbf{K} \leq N$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \begin{cases} \dfrac{\dbinom{N - \mathbf{1}'\mathbf{K}}{n - \mathbf{1}'\mathbf{x}} \displaystyle\prod_{i=1}^m \binom{K_i}{x_i}}{\dbinom{N}{n}} & \begin{array}{l}\text{falls } \mathbf{x} \in \mathbb{N}_0^m \\ \text{und } \mathbf{1}'\mathbf{x} \leq n\end{array} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *polyhypergeometrische Verteilung* und wird mit $\mathbf{PH}(n, N, \mathbf{K})$ bezeichnet.
Modell: Eine Urne enthalte N Kugeln unterschiedlicher Farben, von denen m Farben ausgezeichnet sind und die anderen Farben nicht ausgezeichnet sind. Für $i \in \{1, \ldots, m\}$ bezeichne K_i die Anzahl der Kugeln der ausgezeichneten Farbe i in der Urne und X_i die zufällige Anzahl von Kugeln dieser Farbe in einer Stichprobe vom Umfang n beim Ziehen ohne Zurücklegen. Für den Zufallsvektor $\mathbf{X}$ gilt dann $P_{\mathbf{X}} = \mathbf{PH}(n, N, \mathbf{K})$.
Spezialfall: Im Fall $m = 1$ gilt $\mathbf{PH}(n, N, K_1) = \mathbf{H}(n, N, K_1)$.

(2) **Multinomial–Verteilung:** Für $n \in \mathbb{N}$ und $\boldsymbol{\vartheta} \in (\mathbf{0}, \mathbf{1})$ mit $\mathbf{1}'\boldsymbol{\vartheta} \leq 1$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \begin{cases} \dfrac{n!}{(n - \mathbf{1}'\mathbf{x})! \prod_{i=1}^m x_i!} \, (1 - \mathbf{1}'\boldsymbol{\vartheta})^{n - \mathbf{1}'\mathbf{x}} \displaystyle\prod_{i=1}^m \vartheta_i^{x_i} & \begin{array}{l}\text{falls } \mathbf{x} \in \mathbb{N}_0^m \\ \text{und } \mathbf{1}'\mathbf{x} \leq n\end{array} \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Multinomial–Verteilung* und wird mit $\mathbf{M}(n, \boldsymbol{\vartheta})$ bezeichnet.

Modell: Eine Urne enthalte Kugeln unterschiedlicher Farben, von denen m Farben ausgezeichnet sind und die anderen Farben nicht ausgezeichnet sind. Für $i \in \{1, \ldots, m\}$ bezeichne ϑ_i den Anteil der Kugeln der ausgezeichneten Farbe i in der Urne und X_i die zufällige Anzahl von Kugeln dieser Farbe i in einer Stichprobe vom Umfang n beim Ziehen mit Zurücklegen. Für den Zufallsvektor $\mathbf{X}$ gilt dann $P_{\mathbf{X}} = \mathbf{M}(n, \boldsymbol{\vartheta})$.
Spezialfall: Im Fall $m = 1$ gilt $\mathbf{M}(n, \vartheta_1) = \mathbf{B}(n, \vartheta_1)$.

(3) **Poisson–Verteilung:** Für $\boldsymbol{\alpha} \in \mathbb{R}^m$ mit $\mathbf{0} < \boldsymbol{\alpha}$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \begin{cases} \displaystyle\prod_{i=1}^m e^{-\alpha_i} \frac{\alpha_i^{x_i}}{x_i!} & \text{falls } \mathbf{x} \in \mathbb{N}_0^m \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *multivariate Poisson–Verteilung* und wird mit $\mathbf{P}(\boldsymbol{\alpha})$ bezeichnet.

(4) **Negativmultinomial–Verteilung:** Für $\alpha \in (0, \infty)$ und $\boldsymbol{\vartheta} \in (\mathbf{0}, \mathbf{1})$ mit $\mathbf{1}'\boldsymbol{\vartheta} \leq 1$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \begin{cases} \dfrac{\Gamma(\alpha + \mathbf{1}'\mathbf{x})}{\Gamma(\alpha) \prod_{i=1}^m x_i!} (1 - \mathbf{1}'\boldsymbol{\vartheta})^\alpha \displaystyle\prod_{i=1}^m \vartheta_i^{x_i} & \text{falls } \mathbf{x} \in \mathbb{N}_0^m \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Negativmultinomial–Verteilung* und wird mit $\mathbf{NM}(\alpha, \boldsymbol{\vartheta})$ bezeichnet.
Spezialfall: Im Fall $m = 1$ gilt $\mathbf{NM}(\alpha, \vartheta_1) = \mathbf{NB}(\alpha, \vartheta_1)$.

Absolutstetige Verteilungen

Eine Verteilung $Q : \mathcal{B}(\mathbb{R}^m) \to [0, 1]$ heißt *absolutstetig*, wenn sie absolutstetig bezüglich dem m–dimensionalen Lebesgue–Maß $\boldsymbol{\lambda}^m$ ist, also $Q \ll \boldsymbol{\lambda}^m$ gilt.

Eine messbare Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ heißt *Lebesgue–Dichte*, wenn

$$\int_{\mathbb{R}^m} f(\mathbf{x}) \, d\boldsymbol{\lambda}^m(\mathbf{x}) = 1$$

gilt. Jede absolutstetige Verteilung lässt sich durch eine Lebesgue–Dichte erzeugen, und jede Lebesgue-Dichte erzeugt eine absolutstetige Verteilung:

13.1.6 Lemma.

(1) *Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0, 1]$ eine absolutstetige Verteilung. Dann gibt es eine Lebesgue–Dichte $f : \mathbb{R}^m \to \mathbb{R}_+$ mit*

$$Q = \int f(\mathbf{x}) \, d\boldsymbol{\lambda}^m(\mathbf{x})$$

und die Lebesgue–Dichte ist $\boldsymbol{\lambda}^m$–fast überall eindeutig bestimmt.

(2) *Sei* $f : \mathbb{R}^m \to \mathbb{R}_+$ *eine Lebesgue–Dichte. Dann ist die Mengenfunktion* $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ *mit*

$$Q := \int f(\mathbf{x})\, d\boldsymbol{\lambda}^m(\mathbf{x})$$

eine absolutstetige Verteilung.

Wir betrachten nun einige parametrische Klassen von absolutstetigen multivariaten Verteilungen:

13.1.7 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Für jede Menge $C \in \mathcal{B}(\mathbb{R}^m)$ mit $\boldsymbol{\lambda}^m[C] \in (0,\infty)$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \frac{1}{\boldsymbol{\lambda}^m[C]}\, \chi_C(\mathbf{x})$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *uniforme Verteilung* und wird mit $\mathbf{U}(C)$ bezeichnet.

(2) **Dirichlet–Verteilung:** Für $\eta \in (0,\infty)$ und $\boldsymbol{\eta} \in \mathbb{R}^m$ mit $\mathbf{0} < \boldsymbol{\eta}$ und $\mathbf{1}'\boldsymbol{\eta} < \eta$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \frac{\Gamma(\eta)}{\Gamma(\eta-\mathbf{1}'\boldsymbol{\eta}) \prod_{i=1}^m \Gamma(\eta_i)}\, (1-\mathbf{1}'\mathbf{x})^{\eta-\mathbf{1}'\boldsymbol{\eta}-1} \prod_{i=1}^m x_i^{\eta_i-1}\, \chi_{S^m}(\mathbf{x})$$

und

$$S^m := \Big\{ \mathbf{x} \in (0,\infty)^m \,\Big|\, \mathbf{1}'\mathbf{x} < 1 \Big\}$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt *Dirichlet–Verteilung* und wird mit $\mathbf{Dir}(\eta,\boldsymbol{\eta})$ bezeichnet.
Spezialfall: Im Fall $m = 1$ gilt $\mathbf{Dir}(\eta,\eta_1) = \mathbf{Be}(\eta_1, \eta-\eta_1)$.

(3) **Normal–Verteilung:** Für jeden Vektor $\boldsymbol{\mu} \in \mathbb{R}^m$ und jede positiv definite symmetrische Matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{m\times m}$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}}\, \exp\Big(-\frac{1}{2}\,(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big)$$

eine Lebesgue–Dichte. Die zugehörige Verteilung heißt (*multivariate*) *Normal–Verteilung* und wird mit $\mathbf{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ bezeichnet.
Spezialfall: Die Verteilung $\mathbf{N}(\mathbf{0},\mathbf{I})$ wird als (*multivariate*) *Standardnormal–Verteilung* bezeichnet.
In der Tat: Nach Proposition C.1.3 gibt es eine invertierbare Matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ mit $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}'$. Wir betrachten nun die affine Abbildung $T : \mathbb{R}^m \to \mathbb{R}^m$ mit

$$T(\mathbf{x}) := \mathbf{A}^{-1}(\mathbf{x}-\boldsymbol{\mu})$$

Nach Folgerung 6.3.2 gilt $\boldsymbol{\lambda}^m_T = \sqrt{\det(\boldsymbol{\Sigma})}\, \boldsymbol{\lambda}^m$, und aus der Substitutionsregel und dem Satz von Fubini erhalten wir nun

$$
\begin{aligned}
&\int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\Bigl(-\frac{1}{2}\,(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Bigr)\,d\boldsymbol{\lambda}^m(\mathbf{x})\\
&= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\Bigl(-\frac{1}{2}\,(\mathbf{x}-\boldsymbol{\mu})'(\mathbf{A}\mathbf{A}')^{-1}(\mathbf{x}-\boldsymbol{\mu})\Bigr)\,d\boldsymbol{\lambda}^m(\mathbf{x})\\
&= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\Bigl(-\frac{1}{2}\,(\mathbf{A}^{-1}(\mathbf{x}-\boldsymbol{\mu}))'\,\mathbf{A}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Bigr)\,d\boldsymbol{\lambda}^m(\mathbf{x})\\
&= \int_{T^{-1}(\mathbb{R}^m)} \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\Bigl(-\frac{1}{2}\,(T(\mathbf{x}))'\,T(\mathbf{x})\Bigr)\,d\boldsymbol{\lambda}^m(\mathbf{x})\\
&= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\Bigl(-\frac{1}{2}\,\mathbf{z}'\mathbf{z}\Bigr)\,d\boldsymbol{\lambda}^m_T(\mathbf{z})\\
&= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\Bigl(-\frac{1}{2}\,\mathbf{z}'\mathbf{z}\Bigr)\sqrt{\det(\boldsymbol{\Sigma})}\,d\boldsymbol{\lambda}^m(\mathbf{z})\\
&= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m}} \exp\Bigl(-\frac{1}{2}\,\mathbf{z}'\mathbf{z}\Bigr)\,d\boldsymbol{\lambda}^m(\mathbf{z})\\
&= \int_{\mathbb{R}^m} \prod_{i=1}^m \Bigl(\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}\,z_i^2}\Bigr)\,d\Bigl(\bigotimes_{i=1}^m \boldsymbol{\lambda}\Bigr)(\mathbf{z})\\
&= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \prod_{i=1}^m \Bigl(\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}\,z_i^2}\Bigr)\,d\boldsymbol{\lambda}(z_1)\dots d\boldsymbol{\lambda}(z_m)\\
&= \prod_{i=1}^m \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}\,z_i^2}\,d\boldsymbol{\lambda}(z_i)\\
&= 1
\end{aligned}
$$

Aufgaben

13.1.A Beweisen Sie den Korrespondenzsatz 13.1.1.

13.1.B Sei $F : \mathbb{R}^2 \to [0,1]$ gegeben durch

$$
F(\mathbf{x}) := \begin{cases} 1 & \text{falls } \mathbf{1}'\mathbf{x} \geq 1 \\ 0 & \text{sonst} \end{cases}
$$

Dann besitzt F die Eigenschaften (ii) und (iii) einer Verteilungsfunktion, aber F ist nicht rechtecksmonoton. Zeigen Sie ohne Verwendung des Korrespondenzsatzes, dass es keine Verteilung $Q : \mathcal{B}(\mathbb{R}^2) \to [0,1]$ gibt mit $Q[(-\boldsymbol{\infty}, \mathbf{x}]] = F(\mathbf{x})$ für alle $\mathbf{x} \in \mathbb{R}^2$.

13.1.C **Verteilungsfunktion auf einem abgeschlossenen Intervall:** Eine Funktion $F : [\mathbf{a}, \mathbf{b}] \to [0,1]$ heißt *Verteilungsfunktion auf* $[\mathbf{a}, \mathbf{b}]$, wenn sie die folgenden Eigenschaften besitzt:
(i) F ist rechtecksmonoton.
(ii) F ist stetig von oben.
(iii) Für alle $x_1 \in [a_1, b_1]$ und $x_2 \in [a_2, b_2]$ gilt $F(a_1, x_2) = 0 = F(x_1, a_2)$ und es gilt $\lim_{\mathbf{x}\to\mathbf{b}} F(\mathbf{x}) = 1$.
Formulieren und beweisen Sie ein Analogon des Korrespondenzsatzes.

13.1.D Copulas: Eine Funktion $C : [0,1]^2 \to [0,1]$ heißt *Copula*, wenn sie die folgenden Eigenschaften besitzt:

(i) C ist rechtecksmonoton.

(ii) Für alle $u_1, u_2 \in [0,1]$ gilt $C(u_1, 0) = 0 = C(0, u_2)$ sowie $C(u_1, 1) = u_1$ und $C(1, u_2) = u_2$.

Beweisen Sie die folgenden Aussagen:

(1) Ist $C : [0,1]^2 \to [0,1]$ eine Copula, so gilt für alle $\mathbf{u}, \mathbf{v} \in [0,1]^2$

$$|C(\mathbf{v}) - C(\mathbf{u})| \leq |v_1 - u_1| + |v_2 - u_2|$$

(2) Jede Copula ist eine Verteilungsfunktion.

(3) Sind $F_1, F_2 : \mathbb{R} \to [0,1]$ Verteilungsfunktionen und ist $C : [0,1]^2 \to [0,1]$ eine Copula, so ist die Funktion $F : \mathbb{R}^2 \to [0,1]$ mit

$$F(\mathbf{x}) := C(F_1(x_1), F_2(x_2))$$

eine Verteilungsfunktion.

(4) **Fréchet–Schranken:** Jede der Funktionen $M, W : [0,1]^2 \to [0,1]$ mit

$$W(\mathbf{u}) := \max\{u_1 + u_2 - 1, 0\}$$
$$M(\mathbf{u}) := \min\{u_1, u_2\}$$

ist eine Copula und für jede Copula $C : [0,1]^2 \to [0,1]$ gilt $W \leq C \leq M$. Die Copulas W und M heißen *Fréchet–Schranken*.

(5) Die Funktion $\Pi : [0,1]^2 \to [0,1]$ mit

$$\Pi(\mathbf{u}) := u_1 u_2$$

ist eine Copula.

(6) Sei $\mathbf{X} : \Omega \to \mathbb{R}^2$ ein Zufallsvektor, sei $F_{\mathbf{X}}$ die Verteilungsfunktion zu $P_{\mathbf{X}}$ und seien F_{X_1} und F_{X_2} die Verteilungsfunktionen zu P_{X_1} und P_{X_2}. Dann sind äquivalent:

(a) $\mathbf{X}$ besitzt unabhängige Koordinaten.

(b) Für alle $\mathbf{x} \in \mathbb{R}^2$ gilt $F_{\mathbf{X}}(\mathbf{x}) = \Pi(F_{X_1}(x_1), F_{X_2}(x_2))$.

13.1.E Pólya/Eggenberger–Verteilung: Für $n \in \mathbb{N}$ sowie $\eta \in (0, \infty)$ und $\boldsymbol{\eta} \in \mathbb{R}^m$ mit $\mathbf{0} < \boldsymbol{\eta}$ und $\mathbf{1}'\boldsymbol{\eta} \leq \eta$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \begin{cases} \dfrac{\dbinom{\eta - \mathbf{1}'\boldsymbol{\eta} + n - \mathbf{1}'\mathbf{x} - 1}{n - \mathbf{1}'\mathbf{x}} \displaystyle\prod_{i=1}^{m} \binom{\eta_i + x_i - 1}{x_i}}{\dbinom{\eta + n - 1}{n}} & \text{falls } \mathbf{x} \in \mathbb{N}_0^m \\ & \text{und } \mathbf{1}'\mathbf{x} \leq n \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *Pólya/Eggenberger–Verteilung* und wird mit $\mathbf{PE}(n, \eta, \boldsymbol{\eta})$ bezeichnet.

Spezialfall: Im Fall $m = 1$ gilt $\mathbf{PE}(n, \eta, \eta_1) = \textbf{Pólya}(n, \eta_1, \eta - \eta_1)$.

13.1.F Gemischte Multinomial–Verteilung: Für $n \in \mathbb{N}$ und jede Verteilung $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ mit $Q[\{\boldsymbol{\vartheta} \in (0,1)^m \mid \mathbf{1}'\boldsymbol{\vartheta} \leq 1\}] = 1$ ist die Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$f(\mathbf{x}) := \begin{cases} \displaystyle\int_{\mathbb{R}^m} \frac{n!}{(n-\mathbf{1}'\mathbf{x})! \prod_{i=1}^m x_i!} (1-\mathbf{1}'\boldsymbol{\vartheta})^{n-\mathbf{1}'\mathbf{x}} \prod_{i=1}^m \vartheta_i^{x_i} \, dQ(\boldsymbol{\vartheta}) & \text{falls } \mathbf{x} \in \mathbb{N}_0^m \\ & \text{und } \mathbf{1}'\mathbf{x} \leq n \\ 0 & \text{sonst} \end{cases}$$

eine Zähldichte. Die zugehörige Verteilung heißt *gemischte Multinomial–Verteilung*. Untersuchen Sie den Fall $Q = \mathbf{Dir}(\eta, \boldsymbol{\eta})$.

13.1.G Absolutstetige Verteilungen: Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine absolutstetige Verteilung. Sind $f, g : \mathbb{R}^m \to \mathbb{R}_+$ stetige Funktionen mit

$$\int f \, d\boldsymbol{\lambda}^m = Q = \int g \, d\boldsymbol{\lambda}^m$$

so gilt $f = g$.

13.2 Transformationen von Verteilungen

Wie im univariaten Fall sind auch im multivariaten Fall Transformationen von absolutstetigen Maßen $\mathcal{B}(\mathbb{R}^m) \to [0,\infty]$ von größter Bedeutung, und zwar

– für den Nachweis, dass eine messbare Funktion $f : \mathbb{R}^m \to \mathbb{R}_+$ eine Lebesgue–Dichte einer Verteilung ist,
– für den Nachweis der Existenz und Endlichkeit von Momenten einer Verteilung, und
– für die Berechnung von Momenten einer Verteilung.

Der Fall eines Maßes ν mit $\nu \ll \boldsymbol{\zeta}_C$ für eine abzählbare nichtleere Menge $C \in \mathcal{B}(\mathbb{R})^m$ ist elementar zu behandeln; vgl. Satz 12.2.2.

Für ein Maß $\nu : \mathcal{B}(\mathbb{R}^m) \to [0,\infty]$ mit $\nu \ll \boldsymbol{\lambda}^m$ erhält man die folgende Verallgemeinerung von Satz 12.2.3:

13.2.1 Satz. *Sei $f : \mathbb{R}^m \to \mathbb{R}_+$ messbar und*

$$\nu := \int f(\mathbf{x}) \, d\boldsymbol{\lambda}^m(\mathbf{x})$$

Sei ferner $T : \mathbb{R}^m \to \mathbb{R}^m$ gegeben durch $T(\mathbf{x}) := \mathbf{c} + \mathbf{D}\mathbf{x}$ mit $\mathbf{c} \in \mathbb{R}^m$ und $\mathbf{D} \in \mathbb{R}^{m \times m}$ invertierbar. Dann gilt

$$\nu_T = \int f(\mathbf{D}^{-1}(\mathbf{z}-\mathbf{c})) \, \frac{1}{|\det(\mathbf{D})|} \, d\boldsymbol{\lambda}^m(\mathbf{z})$$

Insbesondere gilt $\nu[\mathbb{R}^m] = \nu_T[\mathbb{R}^m]$.

Wir geben nun ein Beispiel für die Transformation der Verteilung eines Zufallsvektors:

13.2.2 Beispiel (Normal–Verteilung). Sei $P_{\mathbf{X}} := \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Dann gilt für alle $\mathbf{c} \in \mathbb{R}^m$ und jede invertierbare Matrix $\mathbf{D} \in \mathbb{R}^{m \times m}$

$$P_{\mathbf{c}+\mathbf{D}\mathbf{X}} = \mathbf{N}(\mathbf{c}+\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')$$

In der Tat: Es gilt

$$P_{\mathbf{X}} = \int \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\Big(-\frac{1}{2}\,(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big)\, d\boldsymbol{\lambda}^m(\mathbf{x})$$

Wir betrachten die affine Abbildung $T : \mathbb{R}^m \to \mathbb{R}^m$ mit $T(\mathbf{x}) := \mathbf{c} + \mathbf{D}\mathbf{x}$. Dann gilt $P_{\mathbf{c}+\mathbf{D}\mathbf{X}} = P_{T\circ\mathbf{X}} = (P_{\mathbf{X}})_T$ und aus Satz 13.2.1 ergibt sich nun

$$\begin{aligned}
P_{\mathbf{c}+\mathbf{D}\mathbf{X}} &= (P_{\mathbf{X}})_T \\
&= \int \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \\
&\qquad \cdot \exp\Big(-\frac{1}{2}\Big(\mathbf{D}^{-1}(\mathbf{z}-\mathbf{c}) - \boldsymbol{\mu}\Big)'\boldsymbol{\Sigma}^{-1}\Big(\mathbf{D}^{-1}(\mathbf{z}-\mathbf{c}) - \boldsymbol{\mu}\Big)\Big)\,\frac{1}{|\det(\mathbf{D})|}\, d\boldsymbol{\lambda}^m(\mathbf{z}) \\
&= \int \frac{1}{\sqrt{(2\pi)^m \det(\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')}} \\
&\qquad \cdot \exp\Big(-\frac{1}{2}\Big(\mathbf{z} - (\mathbf{c}+\mathbf{D}\boldsymbol{\mu})\Big)'(\mathbf{D}^{-1})'\boldsymbol{\Sigma}^{-1}\mathbf{D}^{-1}\Big(\mathbf{z} - (\mathbf{c}+\mathbf{D}\boldsymbol{\mu})\Big)\Big)\, d\boldsymbol{\lambda}^m(\mathbf{z}) \\
&= \int \frac{1}{\sqrt{(2\pi)^m \det(\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')}} \\
&\qquad \cdot \exp\Big(-\frac{1}{2}\Big(\mathbf{z} - (\mathbf{c}+\mathbf{D}\boldsymbol{\mu})\Big)'(\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')^{-1}\Big(\mathbf{z} - (\mathbf{c}+\mathbf{D}\boldsymbol{\mu})\Big)\Big)\, d\boldsymbol{\lambda}^m(\mathbf{z}) \\
&= \mathbf{N}(\mathbf{c}+\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')
\end{aligned}$$

Aufgabe

13.2.A **Ordnungsstatistiken:** Für $k \in \mathbb{N}(2)$ und $\mathbf{x} \in \mathbb{R}^k$ bezeichne $\mathbf{x}_{k-1}$ den um die letzte Koordinate verkürzten Vektor $\mathbf{x}$.

(1) Für alle $k \in \mathbb{N}(2)$ und $j \in \{1, \dots, k-1\}$ ist die Abbildung $h_{k,j} : \mathbb{R}^k \to \mathbb{R}^k$ mit

$$h_{k,j}(\mathbf{x}) := (x_1, \dots, x_{j-1}, x_j \wedge x_{j+1}, x_j \vee x_{j+1}, x_{j+2}, \dots, x_k)'$$

stetig.

(2) Für alle $k \in \mathbb{N}$ ist die Abbildung $T_k : \mathbb{R}^k \to \mathbb{R}^k$ mit

$$T_k(\mathbf{x}) := \begin{cases} x_1 & \text{falls } k = 1 \\ h_{k,1} \circ \cdots \circ h_{k,k-1} \begin{pmatrix} T_{k-1}(\mathbf{x}_{k-1}) \\ x_k \end{pmatrix} & \text{sonst} \end{cases}$$

stetig und für jedes $\mathbf{x} \in \mathbb{R}^k$ ist die Folge der Koordinaten von $T_k(\mathbf{x})$ monoton wachsend.

(3) Für alle $m \in \mathbb{N}$ und für jeden Zufallsvektor $\mathbf{X} : \Omega \to \mathbb{R}^m$ ist $T_m \circ \mathbf{X}$ messbar.

Für einen Zufallsvektor $\mathbf{X} : \Omega \to \mathbb{R}^m$ werden die Koordinaten von $T_m \circ \mathbf{X}$ als *Ordnungsstatistiken* und mit $X_{m:1}, \dots, X_{m:m}$ bezeichnet.

13.3 Randverteilungen

Im gesamten Abschnitt sei $m \geq 2$ und $J \subseteq \{1,\ldots,m\}$ eine Menge mit $1 \leq |J| \leq m-1$. Wir identifizieren wieder den Messraum $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ mit dem Produkt

$$(\mathbb{R}^J, \mathcal{B}(\mathbb{R}^J)) \otimes (\mathbb{R}^{\{1,\ldots,m\}\setminus J}, \mathcal{B}(\mathbb{R}^{\{1,\ldots,m\}\setminus J}))$$

und jede Funktion $\mathbb{R}^m \to \mathbb{R}$ mit einer Funktion $\mathbb{R}^J \times \mathbb{R}^{\{1,\ldots,m\}\setminus J} \to \mathbb{R}$.

Sei zunächst $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor. Dann ist neben der Verteilung $P_\mathbf{X} : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ von $\mathbf{X}$ auch die Verteilung $(P_\mathbf{X})_{\pi_J} : \mathcal{B}(\mathbb{R}^J) \to [0,1]$ des reduzierten Zufallsvektors $\pi_J \circ \mathbf{X}$ von Interesse; die Verteilung $(P_\mathbf{X})_{\pi_J}$ heißt die *Randverteilung* von $\mathbf{X}$ bezüglich J. Für $i \in \{1,\ldots,m\}$ ist die *eindimensionale Randverteilung* $(P_\mathbf{X})_{\pi_i}$ gerade die Verteilung P_{X_i} der i–ten Koordinate $X_i = \pi_i \circ \mathbf{X}$ von $\mathbf{X}$. Die eindimensionalen Randverteilungen von $\mathbf{X}$ sind von besonderem Interesse, denn nach Satz 11.3.5 besitzt $\mathbf{X}$ genau dann unabhängige Koordinaten, wenn $P_\mathbf{X} = \bigotimes_{i=1}^m P_{X_i}$ gilt.

Allgemein ist für eine Verteilung $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ das Bildmaß Q_{π_J} von Q unter der Projektion $\pi_J : \mathbb{R}^m \to \mathbb{R}^J$ eine Verteilung $\mathcal{B}(\mathbb{R}^J) \to [0,1]$; die Verteilung Q_{π_J} heißt die *Randverteilung* von Q bezüglich J. Wir untersuchen die Berechnung von Randverteilungen und zeigen, dass gewisse parametrische Klassen von Verteilungen unter dem Übergang zu beliebigen Randverteilungen stabil sind.

Für eine Verteilung $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$, die absolutstetig bezüglich einem Produktmaß auf $\mathcal{B}(\mathbb{R}^J) \otimes \mathcal{B}(\mathbb{R}^{\{1,\ldots,m\}\setminus J})$ ist, lässt sich die Randverteilung Q_{π_J} mit Hilfe des folgenden Satzes bestimmen:

13.3.1 Satz. *Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine Verteilung mit $Q \ll \mu \otimes \nu$ für σ–endliche Maße $\mu : \mathcal{B}(\mathbb{R}^J) \to [0,\infty]$ und $\nu : \mathcal{B}(\mathbb{R}^{\{1,\ldots,m\}\setminus J}) \to [0,\infty]$ und sei $f : \mathbb{R}^J \times \mathbb{R}^{\{1,\ldots,m\}\setminus J} \to \mathbb{R}_+$ eine messbare Funktion mit*

$$Q = \int f(\mathbf{y},\mathbf{z})\, d(\mu \otimes \nu)(\mathbf{y},\mathbf{z})$$

Dann gilt

$$Q_{\pi_J} = \int \left(\int_{\mathbb{R}^{\{1,\ldots,m\}\setminus J}} f(\mathbf{y},\mathbf{z})\, d\nu(\mathbf{z}) \right) d\mu(\mathbf{y})$$

Insbesondere ist Q_{π_J} μ–stetig und es gibt eine messbare Funktion $g : \mathbb{R}^J \to \mathbb{R}_+$ mit $Q_{\pi_J} = \int g(\mathbf{y})\, d\mu(\mathbf{y})$, und für alle $\mathbf{y} \in \mathbb{R}^J$ mit $g(\mathbf{y}) \neq 0$ gibt es eine messbare Funktion $h_\mathbf{y} : \mathbb{R}^{\{1,\ldots,m\}\setminus J} \to \mathbb{R}_+$ mit

$$f(\mathbf{y},\mathbf{z}) = g(\mathbf{y})\, h_\mathbf{y}(\mathbf{z})$$

und $\int_{\mathbb{R}^{\{1,\ldots,m\}\setminus J}} h_\mathbf{y}(\mathbf{z})\, d\nu(\mathbf{z}) = 1$.

Beweis. Für alle $B \in \mathcal{B}(\mathbb{R}^J)$ gilt nach dem Satz von Fubini

$$\begin{aligned} Q_{\pi_J}[B] &= Q[\pi_J^{-1}(B)] \\ &= Q[B \times \mathbb{R}^{\{1,\dots,m\}\setminus J}] \\ &= \int_{B \times \mathbb{R}^{\{1,\dots,m\}\setminus J}} f(\mathbf{y},\mathbf{z})\, d(\mu \otimes \nu)(\mathbf{y},\mathbf{z}) \\ &= \int_B \left(\int_{\mathbb{R}^{\{1,\dots,m\}\setminus J}} f(\mathbf{y},\mathbf{z})\, d\nu(\mathbf{z}) \right) d\mu(\mathbf{y}) \end{aligned}$$

Damit ist die erste Gleichung gezeigt und es ist dann klar, dass Q_{π_J} μ–stetig ist. Sei nun $g : \mathbb{R}^J \to \mathbb{R}_+$ gegeben durch

$$g(\mathbf{y}) := \int_{\mathbb{R}^{\{1,\dots,m\}\setminus J}} f(\mathbf{y},\mathbf{z})\, d\nu(\mathbf{z})$$

Dann ist g messbar mit $Q_{\pi_J} = \int g(\mathbf{y})\, d\mu(\mathbf{y})$. Für $\mathbf{y} \in \mathbb{R}^J$ mit $g(\mathbf{y}) \neq 0$ sei $h_{\mathbf{y}} : \mathbb{R}^{\{1,\dots,m\}\setminus J} \to \mathbb{R}_+$ gegeben durch

$$h_{\mathbf{y}}(\mathbf{z}) := f(\mathbf{y},\mathbf{z})/g(\mathbf{y})$$

Dann ist $h_{\mathbf{y}}$ messbar und es gilt

$$f(\mathbf{y},\mathbf{z}) = g(\mathbf{y})\, h_{\mathbf{y}}(\mathbf{z})$$

und $\int_{\mathbb{R}^{\{1,\dots,m\}\setminus J}} h_{\mathbf{y}}(\mathbf{z})\, d\nu(\mathbf{z}) = 1$. □

Nach Konstruktion des Produktmaßes sind die Voraussetzungen des Satzes insbesondere für jede Verteilung Q mit $Q \ll \boldsymbol{\zeta}^m_{\mathbb{N}_0}$ oder $Q \ll \boldsymbol{\lambda}^m$ erfüllt, und für jede Verteilung, die den Voraussetzungen des Satzes genügt, berechnet man eine Randverteilung am einfachsten dadurch, dass man durch Umformung die Faktorisierung $f(\mathbf{y},\mathbf{z}) = g(\mathbf{y})\, h_{\mathbf{y}}(\mathbf{z})$ der Dichte f von Q bestimmt.

Eine weitere Vereinfachung bei der Berechnung von Randverteilungen ergibt sich für alle Verteilungen, die einer permutationsinvarianten parametrischen Klasse von Verteilungen angehören: Ist $\pi : \{1,\dots,m\} \to \{1,\dots,m\}$ eine Permutation, so ist die Abbildung $T_\pi : \mathbb{R}^m \to \mathbb{R}^m$ mit

$$T_\pi(\mathbf{e}_i) := \mathbf{e}_{\pi(i)}$$

eine Bijektion und damit messbar. Eine parametrische Klasse $\mathcal{Q}$ von Verteilungen $\mathcal{B}(\mathbb{R}^m) \to [0,1]$ heißt *permutationsinvariant*, wenn sie für jede Verteilung $Q \in \mathcal{Q}$ und jede Permutation $\pi : \{1,\dots,m\} \to \{1,\dots,m\}$ auch die Verteilung Q_{T_π} enthält. Für eine Verteilung $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ in einer permutationsinvarianten parametrischen Klasse von Verteilungen genügt es daher, für alle $k \in \{1,\dots,m-1\}$ die Randverteilung $Q_{\pi_{\{1,\dots,k\}}}$ zu bestimmen.

13.3.2 Beispiele (Diskrete Verteilungen).

(1) **Polyhypergeometrische Verteilung:** Die Klasse der m–dimensionalen polyhypergeometrischen Verteilungen ist permutationsinvariant. Für

$$Q = \mathbf{PH}(n, N, \mathbf{K})$$

gilt für alle $J \in \mathcal{H}(\{1, \ldots, m\})$

$$Q_{\pi_J} = \mathbf{PH}(n, N, \pi_J(\mathbf{K}))$$

Insbesondere gilt für alle $i \in \{1, \ldots, m\}$

$$Q_{\pi_i} = \mathbf{H}(n, N, K_i)$$

(2) **Multinomial–Verteilung:** Die Klasse der m–dimensionalen Multinomial–Verteilungen ist permutationsinvariant. Für

$$Q = \mathbf{M}(n, \boldsymbol{\vartheta})$$

gilt für alle $J \in \mathcal{H}(\{1, \ldots, m\})$

$$Q_{\pi_J} = \mathbf{M}(n, \pi_J(\boldsymbol{\vartheta}))$$

Insbesondere gilt für alle $i \in \{1, \ldots, m\}$

$$Q_{\pi_i} = \mathbf{B}(n, \vartheta_i)$$

In der Tat: Für alle $x_1, \ldots, x_{m-1} \in \mathbb{N}_0$ mit $\sum_{i=1}^{m-1} x_i \leq n$ gilt

$$\begin{aligned}
&Q_{\pi_{\{1,\ldots,m-1\}}}[\{(x_1, \ldots, x_{m-1})\}] \\
&= \sum_{x_m=0}^{n-\sum_{i=1}^{m-1} x_i} Q[\{(x_1, \ldots, x_{m-1}, x_m)\}] \\
&= \sum_{x_m=0}^{n-\sum_{i=1}^{m-1} x_i} \frac{n!}{(n-\sum_{i=1}^{m} x_i)! \prod_{i=1}^{m} x_i!} \left(1 - \sum_{i=1}^{m} \vartheta_i\right)^{n-\sum_{i=1}^{m} x_i} \prod_{i=1}^{m} \vartheta_i^{x_i} \\
&= \frac{n!}{(n-\sum_{i=1}^{m-1} x_i)! \prod_{i=1}^{m-1} x_i!} \left(1 - \sum_{i=1}^{m-1} \vartheta_i\right)^{n-\sum_{i=1}^{m-1} x_i} \prod_{i=1}^{m-1} \vartheta_i^{x_i} \\
&\quad \cdot \sum_{x_m=0}^{n-\sum_{i=1}^{m-1} x_i} \frac{(n-\sum_{i=1}^{m-1} x_i)!}{x_m!\,(n-\sum_{i=1}^{m} x_i)!} \left(\frac{\vartheta_m}{1 - \sum_{i=1}^{m-1} \vartheta_i}\right)^{x_m} \left(\frac{1 - \sum_{i=1}^{m} \vartheta_i}{1 - \sum_{i=1}^{m-1} \vartheta_i}\right)^{n-\sum_{i=1}^{m} x_i} \\
&= \frac{n!}{(n-\sum_{i=1}^{m-1} x_i)! \prod_{i=1}^{m-1} x_i!} \left(1 - \sum_{i=1}^{m-1} \vartheta_i\right)^{n-\sum_{i=1}^{m-1} x_i} \prod_{i=1}^{m-1} \vartheta_i^{x_i}
\end{aligned}$$

Dabei ergibt sich die letzte Gleichung daraus, dass die Summation über die Einzelwahrscheinlichkeiten der Binomial–Verteilung

$$\mathbf{B}\left(n - \sum_{i=1}^{m-1} x_i, \frac{\vartheta_m}{1-\sum_{i=1}^{m-1} \vartheta_i}\right)$$

den Wert 1 ergibt. Damit ist die Behauptung für den Fall $J = \{1, \ldots, m-1\}$ bewiesen.

(3) **Poisson–Verteilung:** Die Klasse der m–dimensionalen Poisson–Verteilungen ist permutationsinvariant. Für

$$Q = \mathbf{P}(\boldsymbol{\alpha})$$

gilt für alle $J \in \mathcal{H}(\{1, \ldots, m\})$

$$Q_{\pi_J} = \mathbf{P}(\pi_J(\boldsymbol{\alpha}))$$

Insbesondere gilt für alle $i \in \{1, \ldots, m\}$

$$Q_{\pi_i} = \mathbf{P}(\alpha_i)$$

(4) **Negativmultinomial–Verteilung:** Die Klasse der m–dimensionalen Negativmultinomial–Verteilungen ist permutationsinvariant. Für

$$Q = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$$

gilt für alle $J \in \mathcal{H}(\{1, \ldots, m\})$

$$Q_{\pi_J} = \mathbf{NM}\left(\alpha, \frac{1}{\sum_{j \in J} \vartheta_j + 1 - \mathbf{1}'\boldsymbol{\vartheta}}\, \pi_J(\boldsymbol{\vartheta})\right)$$

Insbesondere gilt für alle $i \in \{1, \ldots, m\}$

$$Q_{\pi_i} = \mathbf{NB}\left(\alpha, \frac{\vartheta_i}{\vartheta_i + 1 - \mathbf{1}'\boldsymbol{\vartheta}}\right)$$

Im Fall der polyhypergeometrischen Verteilung und im Fall der Multinomial–Verteilung lässt sich das Ergebnis mit Hilfe der Urnenmodelle veranschaulichen: In beiden Fällen sind ursprünglich m Farben ausgezeichnet und der Übergang zu einer k–dimensionalen Randverteilung entspricht einer Vergröberung der Betrachtungsweise, die darin besteht, dass von den ursprünglichen m ausgezeichneten Farben nur noch k Farben als ausgezeichnet angesehen werden.

13.3.3 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Die Klasse der m–dimensionalen uniformen Verteilungen ist permutationsinvariant. Für

$$Q = \mathbf{U}(C)$$

mit

$$C := \left\{ \mathbf{x} \in \mathbb{R}^2 \;\middle|\; \left(x_1 - \tfrac{1}{2}\right)^2 + \left(x_2 - \tfrac{1}{2}\right)^2 < \tfrac{1}{4} \right\}$$

gilt für alle $i \in \{1, 2\}$

$$Q_{\pi_i} = \mathbf{Be}\left(\tfrac{3}{2}, \tfrac{3}{2}\right)$$

In der Tat: Es gilt

$$Q = \int \frac{4}{\pi}\, \chi_C(\mathbf{x})\, d\boldsymbol{\lambda}^2(\mathbf{x})$$

und aus Satz 13.3.1 folgt

$$Q_{\pi_1} = \int \left(\int_{\mathbb{R}} \frac{4}{\pi}\, \chi_C(x_1, x_2)\, d\boldsymbol{\lambda}(x_2) \right) d\boldsymbol{\lambda}(x_1)$$

Um das innere Integral ausrechnen zu können, setzen wir für alle $x_1 \in (0, 1)$

$$C(x_1) := \left(\tfrac{1}{2} - \sqrt{x_1(1-x_1)}, \tfrac{1}{2} + \sqrt{x_1(1-x_1)}\right)$$

Dann gilt $\chi_C(x_1, x_2) = \chi_{(0,1)}(x_1)\, \chi_{C(x_1)}(x_2)$ und man erhält

$$\begin{aligned}
Q_{\pi_1} &\sim \int \left(\int_{\mathbb{R}} \chi_C(x_1, x_2)\, d\boldsymbol{\lambda}(x_2) \right) d\boldsymbol{\lambda}(x_1) \\
&= \int \left(\int_{\mathbb{R}} \chi_{(0,1)}(x_1)\, \chi_{C(x_1)}(x_2)\, d\boldsymbol{\lambda}(x_2) \right) d\boldsymbol{\lambda}(x_1) \\
&= \int \left(\int_{\mathbb{R}} \chi_{C(x_1)}(x_2)\, d\boldsymbol{\lambda}(x_2) \right) \chi_{(0,1)}(x_1)\, d\boldsymbol{\lambda}(x_1) \\
&= \int \boldsymbol{\lambda}[C(x_1)]\, \chi_{(0,1)}(x_1)\, d\boldsymbol{\lambda}(x_1) \\
&= \int 2\sqrt{x_1(1-x_1)}\, \chi_{(0,1)}(x_1)\, d\boldsymbol{\lambda}(x_1) \\
&\sim \int x_1^{\frac{3}{2}-1} (1-x_1)^{\frac{3}{2}-1} \chi_{(0,1)}(x_1)\, d\boldsymbol{\lambda}(x_1)
\end{aligned}$$

und damit $Q_{\pi_1} = \mathbf{Be}(\frac{3}{2}, \frac{3}{2})$.

(2) **Dirichlet–Verteilung:** Die Klasse der m–dimensionalen Dirichlet–Verteilungen ist permutationsinvariant. Für

$$Q = \mathbf{Dir}(\eta, \boldsymbol{\eta})$$

gilt für alle $J \in \mathcal{H}(\{1, \ldots, m\})$

$$Q_{\pi_J} = \mathbf{Dir}(\eta, \pi_J(\boldsymbol{\eta}))$$

Insbesondere gilt für alle $i \in \{1, \ldots, m\}$

$$Q_{\pi_i} = \mathbf{Be}(\eta_i, \eta - \eta_i)$$

(3) **Normal–Verteilung:** Die Klasse der m–dimensionalen Normal–Verteilungen ist permutationsinvariant. Für

$$Q = \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

gilt für alle $J \in \mathcal{H}(\{1, \ldots, m\})$

$$Q_{\pi_J} = \mathbf{N}(\boldsymbol{\mu}_J, \boldsymbol{\Sigma}_J)$$

mit $\boldsymbol{\mu}_J := \pi_J(\boldsymbol{\mu})$ und $\boldsymbol{\Sigma}_J := \{\sigma_{ij}\}_{i,j \in J}$. Insbesondere gilt für alle $i \in \{1, \ldots, m\}$

$$Q_{\pi_i} = \mathbf{N}(\mu_i, \sigma_{ii})$$

In der Tat: Es gilt

$$\begin{aligned}
Q &= \int \frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) d\boldsymbol{\lambda}^m(\mathbf{x}) \\
&\sim \int \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) d\boldsymbol{\lambda}^m(\mathbf{x})
\end{aligned}$$

Wir betrachten $k \in \{1, \ldots, m-1\}$ und schreiben $\boldsymbol{\mu}$ als Blockvektor

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}$$

mit $\boldsymbol{\mu}_1 := \{\mu_i\}_{i\in\{1,\dots,k\}}$ und $\boldsymbol{\Sigma}$ als Blockmatrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

mit $\boldsymbol{\Sigma}_{11} := \{\sigma_{ij}\}_{i,j\in\{1,\dots,k\}}$. Dann sind $\boldsymbol{\Sigma}_{11}$ und $\boldsymbol{\Sigma}_{22}$ symmetrisch und positiv definit, und damit ist auch die Matrix $\mathbf{U} := \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ symmetrisch und positiv definit. Es gilt

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I} \end{pmatrix}' \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{U}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I} \end{pmatrix}$$

Wir schreiben nun auch jeden Vektor $\mathbf{x} \in \mathbb{R}^m$ als Blockvektor

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$$

mit $\mathbf{x}_1 := \{x_i\}_{i\in\{1,\dots,k\}}$ und setzen $\boldsymbol{\nu}(\mathbf{x}_1) := \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)$. Dann gilt

$$\begin{aligned}
&(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \\
&= \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{pmatrix}' \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I} \end{pmatrix}' \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{U}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1) \end{pmatrix}' \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{U}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1) \end{pmatrix} \\
&= (\mathbf{x}_1 - \boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) + (\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))'\mathbf{U}^{-1}(\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))
\end{aligned}$$

und damit

$$\begin{aligned}
&\exp\Bigl(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Bigr) \\
&= \exp\Bigl(-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)\Bigr) \cdot \exp\Bigl(-\frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))'\mathbf{U}^{-1}(\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))\Bigr)
\end{aligned}$$

Aufgrund der Translationsinvarianz des Lebesgue–Maßes ist das Integral

$$\int_{\mathbb{R}^{m-k}} \exp\Bigl(-\frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))'\mathbf{U}^{-1}(\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))\Bigr)\, d\boldsymbol{\lambda}^{m-k}(\mathbf{x}_2)$$

unabhängig von $\boldsymbol{\nu}(\mathbf{x}_1)$ und damit unabhängig von $\mathbf{x}_1$, und wegen

$$\mathbf{N}(\boldsymbol{\nu}(\mathbf{x}_1), \mathbf{U}) \sim \int_{\mathbb{R}^{m-k}} \exp\Bigl(-\frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))'\mathbf{U}^{-1}(\mathbf{x}_2 - \boldsymbol{\nu}(\mathbf{x}_1))\Bigr)\, d\boldsymbol{\lambda}^{m-k}(\mathbf{x}_2)$$

ergibt sich nun aus Satz 13.3.1

$$Q_{\pi_{\{1,\dots,k\}}} \sim \int \exp\Bigl(-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)\Bigr)\, d\boldsymbol{\lambda}^{k}(\mathbf{x}_1)$$

und damit $Q_{\pi_{\{1,\dots,k\}}} = \mathbf{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$.

Aufgaben

13.3.A Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine Verteilung und sei F die zugehörige Verteilungsfunktion. Dann gilt für die Verteilungsfunktion F_{π_J} zu Q_{π_J}

$$F_{\pi_J}(\mathbf{y}) = \lim_{n\to\infty} F\Big(\pi_J^{-1}(\mathbf{y}) \cap \pi^{-1}_{\{1,\dots,m\}\setminus J}(n\mathbf{1}_{\{1,\dots,m\}\setminus J})\Big)$$

13.3.B **Multinomial–Verteilung:** Sei $n \in \mathbb{N}$ und sei $\{\vartheta_i\}_{i\in\mathbb{N}} \subseteq (0,\infty)$ eine Folge mit $\sum_{i=1}^\infty \vartheta_i = 1$. Dann ist die Familie $\{\mathbf{M}(n,\{\vartheta_i\}_{i\in K})\}_{K\in\mathcal{H}(\mathbb{N})}$ projektiv.

13.3.C **Pólya/Eggenberger–Verteilung:** Sei $n \in \mathbb{N}$ und $\eta \in (0,\infty)$ und sei $\{\eta_i\}_{i\in\mathbb{N}} \subseteq (0,\infty)$ eine Folge mit $\sum_{i=1}^\infty \eta_i = \eta$. Dann ist die Familie $\{\mathbf{PE}(n,\eta,\{\eta_i\}_{i\in K})\}_{K\in\mathcal{H}(\mathbb{N})}$ projektiv.

13.3.D **Dirichlet–Verteilung:** Sei $\eta \in (0,\infty)$ und sei $\{\eta_i\}_{i\in\mathbb{N}} \subseteq (0,\infty)$ eine Folge mit $\sum_{i=1}^\infty \eta_i = \eta$. Dann ist die Familie $\{\mathbf{Dir}(\eta,\{\eta_i\}_{i\in K})\}_{K\in\mathcal{H}(\mathbb{N})}$ projektiv.

13.4 Unabhängigkeit

Sei zunächst $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor. Nach Satz 11.3.5 besitzt $\mathbf{X}$ genau dann unabhängige Koordinaten $X_1,\dots,X_m$, wenn

$$P_{\mathbf{X}} = \bigotimes_{i=1}^m P_{X_i}$$

gilt. Wegen $X_i = \pi_i \circ \mathbf{X}$ lässt sich diese Bedingung auch in der Form

$$P_{\mathbf{X}} = \bigotimes_{i=1}^m (P_{\mathbf{X}})_{\pi_i}$$

schreiben. Daher besitzt ein Zufallsvektor genau dann unabhängige Koordinaten, wenn seine Verteilung gleich dem Produkt seiner eindimensionalen Randverteilungen ist.

Wir untersuchen nun die Frage, unter welchen Bedingungen eine Verteilung $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ mit dem Produkt ihrer eindimensionalen Randverteilungen $Q_{\pi_i} : \mathcal{B}(\mathbb{R}) \to [0,1]$ übereinstimmt. Als erstes charakterisieren wir die Gültigkeit der Gleichung

$$Q = \bigotimes_{i=1}^m Q_{\pi_i}$$

durch eine analoge Gleichung für die zugehörigen Verteilungsfunktionen:

13.4.1 Satz. *Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine Verteilung. Sei ferner F die Verteilungsfunktion zu Q und für $i \in \{1,\dots,m\}$ sei F_{π_i} die Verteilungsfunktion zu Q_{π_i}. Dann sind äquivalent:*
(a) *Es gilt $Q = \bigotimes_{i=1}^m Q_{\pi_i}$.*
(b) *Für alle $\mathbf{x} \in \mathbb{R}^m$ gilt $F(\mathbf{x}) = \prod_{i=1}^m F_{\pi_i}(x_i)$.*

Beweis. Für alle $\mathbf{x} \in \mathbb{R}^m$ gilt

$$F(\mathbf{x}) = Q[(-\boldsymbol{\infty}, \mathbf{x}]]$$

und

$$\begin{aligned}
\prod_{i=1}^m F_{\pi_i}(x_i) &= \prod_{i=1}^m Q_{\pi_i}[(-\infty, x_i]] \\
&= \left(\bigotimes_{i=1}^m Q_{\pi_i}\right)\left[\prod_{i=1}^m (-\infty, x_i]\right] \\
&= \left(\bigotimes_{i=1}^m Q_{\pi_i}\right)[(-\boldsymbol{\infty}, \mathbf{x}]]
\end{aligned}$$

Im Fall $Q = \bigotimes_{i=1}^m Q_{\pi_i}$ gilt daher $F(\mathbf{x}) = \prod_{i=1}^m F_{\pi_i}(x_i)$ für alle $\mathbf{x} \in \mathbb{R}^m$, und die umgekehrte Implikation folgt aus dem Eindeutigkeitssatz. □

Da Verteilungsfunktionen nur in seltenen Fällen explizit angegeben werden können, ist der letzte Satz nur von theoretischem Interesse. Für Anwendungen wichtig ist hingegen der folgende Satz über Verteilungen, die absolutstetig bezüglich einem Produktmaß sind; vgl. Aufgabe 13.4.A:

13.4.2 Satz. *Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine Verteilung mit $Q \ll \mu^m$ für ein σ–endliches Maß $\mu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$. Sei ferner f eine μ^m–Dichte von Q und für $i \in \{1,\dots,m\}$ sei f_{π_i} eine μ–Dichte von Q_{π_i}. Dann sind äquivalent:*
(a) *Es gilt $Q = \bigotimes_{i=1}^m Q_{\pi_i}$.*
(b) *Es gilt $f(\mathbf{x}) = \prod_{i=1}^m f_{\pi_i}(x_i)$ μ^m–fast überall.*

Beweis. Für alle $B_1,\dots,B_m \in \mathcal{B}(\mathbb{R})$ gilt

$$Q\left[\prod_{i=1}^m B_i\right] = \int_{\prod_{i=1}^m B_i} f(\mathbf{x})\, d\mu^m(\mathbf{x})$$

und nach dem Satz von Fubini gilt

$$\begin{aligned}
\left(\bigotimes_{i=1}^m Q_{\pi_i}\right)\left[\prod_{i=1}^m B_i\right] &= \prod_{i=1}^m Q_{\pi_i}[B_i] \\
&= \prod_{i=1}^m \int_{B_i} f_{\pi_i}(x_i)\, d\mu(x_i) \\
&= \int_{\prod_{i=1}^m B_i} \left(\prod_{i=1}^m f_{\pi_i}(x_i)\right) d\mu^m(\mathbf{x})
\end{aligned}$$

Nach dem Eindeutigkeitssatz ist die Bedingung $Q = \bigotimes_{i=1}^m Q_{\pi_i}$ gleichwertig mit

$$\int f(\mathbf{x})\, d\mu^m(\mathbf{x}) = \int \left(\prod_{i=1}^m f_{\pi_i}(x_i) \right) d\mu^m(\mathbf{x})$$

und nach dem Satz von Radon/Nikodym ist diese Bedingung gleichwertig mit $f(\mathbf{x}) = \prod_{i=1}^m f_{\pi_i}(x_i)$ fast überall. □

13.4.3 Beispiele (Diskrete Verteilungen).
(1) **Polyhypergeometrische Verteilung:** Für

$$Q = \mathbf{PH}(n, N, \mathbf{K})$$

gilt $Q \neq \bigotimes_{i=1}^m Q_{\pi_i}$.
(2) **Multinomial–Verteilung:** Für

$$Q = \mathbf{M}(n, \boldsymbol{\vartheta})$$

gilt $Q \neq \bigotimes_{i=1}^m Q_{\pi_i}$.
(3) **Poisson–Verteilung:** Für

$$Q = \mathbf{P}(\boldsymbol{\alpha})$$

gilt $Q = \bigotimes_{i=1}^m Q_{\pi_i}$.
(4) **Negativmultinomial–Verteilung:** Für

$$Q = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$$

gilt $Q \neq \bigotimes_{i=1}^m Q_{\pi_i}$.

13.4.4 Beispiele (Absolutstetige Verteilungen).
(1) **Uniforme Verteilung:** Für

$$Q = \mathbf{U}(C)$$

gilt $Q = \bigotimes_{i=1}^m Q_{\pi_i}$ genau dann, wenn es $C_1, \dots, C_m \in \mathcal{B}(\mathbb{R})$ gibt mit

$$\boldsymbol{\lambda}^m \left[C \triangle \prod_{i=1}^m C_i \right] = 0$$

In der Tat: Wir nehmen zunächst an, dass $Q = \bigotimes_{i=1}^m Q_{\pi_i}$ gilt. Für $i \in \{1, \dots, m\}$ sei $f_{\pi_i} : \mathbb{R} \to \mathbb{R}_+$ messbar mit $Q_{\pi_i} = \int f_{\pi_i}\, d\boldsymbol{\lambda}$ und $C_i := \{f_{\pi_i} > 0\}$. Dann gilt nach Satz 13.4.2

$$\frac{1}{\boldsymbol{\lambda}^m[C]}\, \chi_C(\mathbf{x}) = \prod_{i=1}^m f_{\pi_i}(x_i)$$

$\boldsymbol{\lambda}^m$–fast überall, und daraus folgt $\boldsymbol{\lambda}^m[C \triangle \prod_{i=1}^m C_i] = 0$.
Wir nehmen nun an, dass es $C_1, \dots, C_m \in \mathcal{B}(\mathbb{R})$ gibt mit $\boldsymbol{\lambda}^m[C \triangle \prod_{i=1}^m C_i] = 0$. Dann gilt

$$\frac{1}{\boldsymbol{\lambda}^m[C]}\, \chi_C(\mathbf{x}) = \prod_{i=1}^m \frac{1}{\boldsymbol{\lambda}[C_i]}\, \chi_{C_i}(x_i)$$

$\boldsymbol{\lambda}^m$–fast überall, und aus dem Satz von Fubini folgt für alle $B_1, \dots, B_m \in \mathcal{B}(\mathbb{R})$

$$\begin{aligned} Q\left[\prod_{i=1}^m B_i\right] &= \int_{\prod_{i=1}^m B_i} \frac{1}{\boldsymbol{\lambda}^m[C]}\, \chi_C(\mathbf{x})\, d\boldsymbol{\lambda}^m(\mathbf{x}) \\ &= \int_{\prod_{i=1}^m B_i} \left(\prod_{i=1}^m \frac{1}{\boldsymbol{\lambda}[C_i]}\, \chi_{C_i}(x_i)\right) d\boldsymbol{\lambda}^m(\mathbf{x}) \\ &= \prod_{i=1}^m \int_{B_i} \frac{1}{\boldsymbol{\lambda}[C_i]}\, \chi_{C_i}(x_i)\, d\boldsymbol{\lambda}(x_i) \end{aligned}$$

Daraus ergibt sich zunächst

$$Q_{\pi_i}[B_i] = \int_{B_i} \frac{1}{\boldsymbol{\lambda}[C_i]}\, \chi_{C_i}(x_i)\, d\boldsymbol{\lambda}(x_i)$$

für alle $i \in \{1, \dots, m\}$, und sodann

$$\begin{aligned} Q\left[\prod_{i=1}^m B_i\right] &= \prod_{i=1}^m \int_{B_i} \frac{1}{\boldsymbol{\lambda}[C_i]}\, \chi_{C_i}(x_i)\, d\boldsymbol{\lambda}(x_i) \\ &= \prod_{i=1}^m Q_{\pi_i}[B_i] \\ &= \left(\bigotimes_{i=1}^m Q_{\pi_i}\right)\left[\prod_{i=1}^m B_i\right] \end{aligned}$$

Aus dem Eindeutigkeitssatz folgt nun $Q = \bigotimes_{i=1}^m Q_{\pi_i}$.

(2) **Dirichlet–Verteilung:** Für

$$Q = \mathbf{Dir}(\eta, \boldsymbol{\eta})$$

gilt $Q \neq \bigotimes_{i=1}^m Q_{\pi_i}$.

(3) **Normal–Verteilung:** Für

$$Q = \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

gilt $Q = \bigotimes_{i=1}^m Q_{\pi_i}$ genau dann, wenn $\boldsymbol{\Sigma}$ eine Diagonalmatrix ist.
In der Tat: Wir nehmen zunächst an, dass $Q = \bigotimes_{i=1}^m Q_{\pi_i}$ gilt. Nach Satz 13.4.2 gilt dann

$$\frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) = \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma_{ii}}} \exp\left(-\frac{1}{2}\,\frac{(x_i-\mu_i)^2}{\sigma_{ii}}\right)$$

$\boldsymbol{\lambda}^m$–fast überall, und da beide Seiten stetig sind, gilt die Gleichung sogar für alle $\mathbf{x} \in \mathbb{R}^m$. Mit $\mathbf{x} = \boldsymbol{\mu}$ erhalten wir zunächst

$$\frac{1}{\sqrt{(2\pi)^m \det(\boldsymbol{\Sigma})}} = \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma_{ii}}}$$

und sodann für alle $\mathbf{x} \in \mathbb{R}^m$

$$\begin{aligned}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) &= \prod_{i=1}^{m}\exp\left(-\frac{1}{2}\frac{(x_i-\mu_i)^2}{\sigma_{ii}}\right)\\ &= \exp\left(-\frac{1}{2}\sum_{i=1}^{m}\frac{(x_i-\mu_i)^2}{\sigma_{ii}}\right)\end{aligned}$$

Für die Diagonalmatrix $\mathbf{U}$ mit

$$u_{ij} := \begin{cases}\sigma_{ii} & \text{falls } i=j\\ 0 & \text{sonst}\end{cases}$$

gilt daher

$$\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) = \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\mathbf{U}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

und damit

$$(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = (\mathbf{x}-\boldsymbol{\mu})'\mathbf{U}^{-1}(\mathbf{x}-\boldsymbol{\mu})$$

Daraus folgt $\boldsymbol{\Sigma}^{-1} = \mathbf{U}^{-1}$ und damit $\boldsymbol{\Sigma} = \mathbf{U}$.
Die umgekehrte Implikation ist klar.

Aufgaben

13.4.A Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine Verteilung mit $Q \ll \bigotimes_{i=1}^m \mu_i$ für σ–endliche Maße $\mu_1,\dots,\mu_m : \mathcal{B}(\mathbb{R}) \to [0,\infty]$. Sei ferner f eine $(\bigotimes_{i=1}^m \mu_i)$–Dichte von Q und für $i \in \{1,\dots,m\}$ sei f_{π_i} eine μ_i–Dichte von Q_{π_i}. Dann sind äquivalent:
(a) Es gilt $Q = \bigotimes_{i=1}^m Q_{\pi_i}$.
(b) Es gilt $f(\mathbf{x}) = \prod_{i=1}^m f_{\pi_i}(x_i)$ $(\bigotimes_{i=1}^m \mu_i)$–fast überall.

13.4.B Sind X und Y unabhängige Zufallsvariable, so gilt für alle $z \in \mathbb{R}$

$$\begin{aligned}P[\{X \vee Y \le z\}] &= P[\{X \le z\}]\,P[\{Y \le z\}]\\ P[\{X \wedge Y > z\}] &= P[\{X > z\}]\,P[\{Y > z\}]\end{aligned}$$

13.4.C Bestimmen Sie für unabhängige exponential–verteilte Zufallsvariable X, Y die Verteilung von $X \wedge Y$.

13.5 Verteilungen von Summen von Zufallsvariablen

In diesem Abschnitt untersuchen wir für reelle Zufallsvariable $X_1,\dots,X_m$ die Verteilung ihrer Summe $\sum_{i=1}^m X_i$.

Zu diesem Zweck fassen wir die reellen Zufallsvariablen $X_1,\dots,X_m$ als Koordinaten eines Zufallsvektors $\mathbf{X} : \Omega \to \mathbb{R}^m$ auf. Wegen

$$\sum_{i=1}^m X_i = \mathbf{1}'\mathbf{X}$$

geht die Verteilung der Summe über die lineare Abbildung $T : \mathbb{R}^m \to \mathbb{R}$ mit $T(\mathbf{x}) := \mathbf{1}'\mathbf{x}$ aus der Verteilung von $\mathbf{X}$ hervor, denn es gilt

$$P_{\sum_{i=1}^m X_i} = P_{\mathbf{1}'\mathbf{X}} = P_{T\circ\mathbf{X}} = (P_{\mathbf{X}})_T$$

Wir untersuchen die Berechnung der Verteilung der Summe von möglicherweise abhängigen reellen Zufallsvariablen und zeigen, dass gewisse Klassen von univariaten Verteilungen unter der Summation von unabhängigen reellen Zufallsvariablen stabil sind.

Wir betrachten zunächst den Fall, in dem $\mathbf{X}$ eine diskrete Verteilung besitzt. In diesem Fall erhält man die Verteilung von $\mathbf{1}'\mathbf{X}$ sofort aus der multivariaten Version von Satz 12.2.2:

13.5.1 Satz. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor mit einer diskreten Verteilung und sei $C \subseteq \mathbb{R}^m$ abzählbar mit $P_{\mathbf{X}}[C] = 1$. Dann gilt für alle $B \in \mathcal{B}(\mathbb{R})$*

$$P_{\mathbf{1}'\mathbf{X}}[B] = \sum_{x\in B}\left(\sum_{\mathbf{x}\in C,\, \mathbf{1}'\mathbf{x}=x} P[\{\mathbf{X}=\mathbf{x}\}]\right)$$

Insbesondere ist $P_{\mathbf{1}'\mathbf{X}}$ diskret.

Wir geben einige Beispiele für die Anwendung des Satzes:

13.5.2 Beispiele (Diskrete Verteilungen).

(1) **Polyhypergeometrische Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{PH}(n, N, \mathbf{K})$ gilt

$$P_{\mathbf{1}'\mathbf{X}} = \mathbf{H}(n, N, \mathbf{1}'\mathbf{K})$$

(2) **Multinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{M}(n, \boldsymbol{\vartheta})$ gilt

$$P_{\mathbf{1}'\mathbf{X}} = \mathbf{B}(n, \mathbf{1}'\boldsymbol{\vartheta})$$

(3) **Poisson–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{P}(\boldsymbol{\alpha})$ gilt

$$P_{\mathbf{1}'\mathbf{X}} = \mathbf{P}(\mathbf{1}'\boldsymbol{\alpha})$$

(4) **Negativmultinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$ gilt

$$P_{\mathbf{1}'\mathbf{X}} = \mathbf{NB}\big(\alpha, \mathbf{1}'\boldsymbol{\vartheta}\big)$$

Im Fall eines Zufallsvektors mit unabhängigen Koordinaten vereinfacht sich die Darstellung der Verteilung der Summe wie folgt:

13.5.3 Folgerung. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor mit einer diskreten Verteilung und sei $C \subseteq \mathbb{R}^m$ abzählbar mit $P_{\mathbf{X}}[C] = 1$. Besitzt $\mathbf{X}$ unabhängige Koordinaten, so gilt für alle $B \in \mathcal{B}(\mathbb{R})$*

$$P_{\mathbf{1}'\mathbf{X}}[B] = \sum_{x\in B}\left(\sum_{\mathbf{x}\in C,\, \mathbf{1}'\mathbf{x}=x} \prod_{i=1}^m P[\{X_i = x_i\}]\right)$$

Als wichtigen Spezialfall des letzten Ergebnisses notieren wir die folgende Faltungsformel:

13.5.4 Folgerung (Faltungsformel). *Seien X und Y unabhängige reelle Zufallsvariable mit einer diskreten Verteilung. Dann gilt*

$$P[\{X+Y=z\}] = \sum_{x\in\mathbb{R},\, P[\{X=x\}]>0} P[\{X=x\}]\, P[\{Y=z-x\}]$$

Die Faltungsformel bleibt natürlich richtig, wenn man auf der rechten Seite der Gleichung die Rollen von X und Y vertauscht. Wir geben einige Beispiele für die Anwendung der Faltungsformel:

13.5.5 Beispiele (Diskrete Verteilungen). Seien X und Y unabhängige reelle Zufallsvariable.

(1) **Binomial–Verteilung:** Im Fall $P_X = \mathbf{B}(m,\vartheta)$ und $P_Y = \mathbf{B}(n,\vartheta)$ gilt

$$P_{X+Y} = \mathbf{B}(m+n,\vartheta)$$

(2) **Poisson–Verteilung:** Im Fall $P_X = \mathbf{P}(\alpha)$ und $P_Y = \mathbf{P}(\beta)$ gilt

$$P_{X+Y} = \mathbf{P}(\alpha+\beta)$$

In der Tat: Für alle $k \in \mathbb{N}_0$ gilt

$$\begin{aligned} P[\{X+Y=k\}] &= \sum_{l=0}^{k} P[\{X=l\}]\, P[\{Y=k-l\}] \\ &= \sum_{l=0}^{k} e^{-\alpha}\frac{\alpha^l}{l!}\cdot e^{-\beta}\frac{\beta^{k-l}}{(k-l)!} \\ &= e^{-(\alpha+\beta)}\,\frac{(\alpha+\beta)^k}{k!}\sum_{l=0}^{k}\binom{k}{l}\left(\frac{\alpha}{\alpha+\beta}\right)^{l}\left(\frac{\beta}{\alpha+\beta}\right)^{k-l} \\ &= e^{-(\alpha+\beta)}\,\frac{(\alpha+\beta)^k}{k!} \end{aligned}$$

(3) **Negativbinomial–Verteilung:** Im Fall $P_X = \mathbf{NB}(\alpha,\vartheta)$ und $P_Y = \mathbf{NB}(\beta,\vartheta)$ gilt

$$P_{X+Y} = \mathbf{NB}(\alpha+\beta,\vartheta)$$

Wir betrachten nun den Fall, in dem $\mathbf{X}$ eine absolutstetige Verteilung besitzt. In diesem Fall lässt sich die Verteilung der Summe $\mathbf{1}'\mathbf{X}$ zwar nicht direkt, aber mit einem kleinen Umweg aus dem entsprechenden Satz 13.2.1 über lineare Abbildungen gewinnen.

13.5.6 Satz. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor mit $P_{\mathbf{X}} \ll \boldsymbol{\lambda}^m$. Sei ferner f eine $\boldsymbol{\lambda}^m$–Dichte von $P_{\mathbf{X}}$. Dann gilt*

$$P_{\mathbf{1}'\mathbf{X}} = \int\left(\int_{\mathbb{R}^{m-1}} f(y-\mathbf{1}'\mathbf{z},\mathbf{z})\, d\boldsymbol{\lambda}^{m-1}(\mathbf{z})\right) d\boldsymbol{\lambda}(y)$$

Insbesondere gilt $P_{\mathbf{1}'\mathbf{X}} \ll \boldsymbol{\lambda}$.

Beweis. Wir betrachten die lineare Abbildung $T : \mathbb{R}^m \to \mathbb{R}^m$ mit $T(\mathbf{x}) := \mathbf{D}\mathbf{x}$ und der Matrix $\mathbf{D}$ mit den Koordinaten

$$d_{ij} := \begin{cases} 1 & \text{falls } i = 1 \text{ oder } i = j \\ 0 & \text{sonst} \end{cases}$$

Dann gilt $\mathbf{1}'\mathbf{X} = \pi_1 \circ T \circ \mathbf{X}$ und damit

$$P_{\mathbf{1}'\mathbf{X}} = P_{\pi_1 \circ T \circ \mathbf{X}} = ((P_{\mathbf{X}})_T)_{\pi_1}$$

Wegen $\det(\mathbf{D}) = 1$ folgt aus Satz 13.2.1

$$(P_{\mathbf{X}})_T = \int f(\mathbf{D}^{-1}\mathbf{u})\, d\boldsymbol{\lambda}^m(\mathbf{u})$$

Für die Koordinaten c_{ij} der Matrix $\mathbf{D}^{-1}$ gilt

$$c_{ij} = \begin{cases} 1 & \text{falls } i = j \\ -1 & \text{falls } i = 1 \neq j \\ 0 & \text{sonst} \end{cases}$$

Für die Funktion $\tilde{f} : \mathbb{R} \times \mathbb{R}^{m-1} \to \mathbb{R}_+$ mit $\tilde{f}(y, \mathbf{z}) := f(y - \mathbf{1}'\mathbf{z}, \mathbf{z})$ gilt dann

$$\begin{aligned} (P_{\mathbf{X}})_T &= \int f(\mathbf{D}^{-1}\mathbf{u})\, d\boldsymbol{\lambda}^m(\mathbf{u}) \\ &= \int f(y - \mathbf{1}'\mathbf{z}, \mathbf{z})\, d(\boldsymbol{\lambda} \otimes \boldsymbol{\lambda}^{m-1})(y, \mathbf{z}) \\ &= \int \tilde{f}(y, \mathbf{z})\, d(\boldsymbol{\lambda} \otimes \boldsymbol{\lambda}^{m-1})(y, \mathbf{z}) \end{aligned}$$

und aus Satz 13.3.1 folgt nun

$$\begin{aligned} ((P_{\mathbf{X}})_T)_{\pi_1} &= \int \left(\int_{\mathbb{R}^{m-1}} \tilde{f}(y, \mathbf{z})\, d\boldsymbol{\lambda}^{m-1}(\mathbf{z}) \right) d\boldsymbol{\lambda}(y) \\ &= \int \left(\int_{\mathbb{R}^{m-1}} f(y - \mathbf{1}'\mathbf{z}, \mathbf{z})\, d\boldsymbol{\lambda}^{m-1}(\mathbf{z}) \right) d\boldsymbol{\lambda}(y) \end{aligned}$$

Wegen $P_{\mathbf{1}'\mathbf{X}} = ((P_{\mathbf{X}})_T)_{\pi_1}$ folgt daraus die Behauptung. □

Satz 13.5.6 ist vor allem für Zufallsvektoren mit unabhängigen Koordinaten von Interesse; vgl. Folgerung 13.5.8. In anderen Fällen kann es günstiger sein, die im Beweis des Satzes verwendeten Argumente direkt auf die vorliegende Verteilung anzuwenden; dies gilt insbesondere dann, wenn die Eigenschaften von Verteilungen linearer Transformationen und von Randverteilungen bereits bekannt sind:

13.5.7 Beispiel (Normal–Verteilung). Sei $P_{\mathbf{X}} = \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Dann gilt

$$P_{\mathbf{1}'\mathbf{X}} = \mathbf{N}(\mathbf{1}'\boldsymbol{\mu}, \mathbf{1}'\boldsymbol{\Sigma}\mathbf{1})$$

In der Tat: Wir betrachten die Matrix $\mathbf{D} \in \mathbb{R}^{m \times m}$ mit den Koordinaten

$$d_{ij} := \begin{cases} 1 & \text{falls } i = 1 \text{ oder } i = j \\ 0 & \text{sonst} \end{cases}$$

Nach Beispiel 13.2.2 gilt $P_{\mathbf{DX}} = \mathbf{N}(\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')$. Wegen $\mathbf{e}_1'\mathbf{D} = \mathbf{1}'$ gilt $\mathbf{e}_1'\mathbf{D}\boldsymbol{\mu} = \mathbf{1}'\boldsymbol{\mu}$ und $\mathbf{e}_1'\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}'\mathbf{e}_1 = \mathbf{1}'\boldsymbol{\Sigma}\mathbf{1}$, und aus Beispiel 13.3.3 ergibt sich nun

$$P_{\mathbf{1}'\mathbf{X}} = P_{\mathbf{e}_1'\mathbf{DX}} = (P_{\mathbf{DX}})_{\pi_1} = \mathbf{N}(\mathbf{1}'\boldsymbol{\mu}, \mathbf{1}'\boldsymbol{\Sigma}\mathbf{1})$$

Wir betrachten nun wieder den Fall eines Zufallsvektors mit unabhängigen Koordinaten:

13.5.8 Folgerung. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor mit $P_{\mathbf{X}} \ll \boldsymbol{\lambda}^m$ und unabhängigen Koordinaten. Für $i \in \{1, \ldots, m\}$ sei $f_{\pi_i} : \mathbb{R} \to \mathbb{R}_+$ eine $\boldsymbol{\lambda}$–Dichte von P_{X_i}. Dann gilt*

$$P_{\mathbf{1}'\mathbf{X}} = \int \left(\int_{\mathbb{R}^{m-1}} \left(f_{\pi_1}(y - \mathbf{1}'\mathbf{z}) \prod_{i=2}^{m} f_{\pi_i}(z_{i-1}) \right) d\boldsymbol{\lambda}^{m-1}(\mathbf{z}) \right) d\boldsymbol{\lambda}(y)$$

Als wichtigen Spezialfall des letzten Ergebnisses notieren wir die folgende Faltungsformel; vgl. Aufgabe 13.5.E:

13.5.9 Folgerung (Faltungsformel). *Seien X und Y unabhängige Zufallsvariable mit $P_X = \int f(x)\, d\boldsymbol{\lambda}(x)$ und $P_Y = \int g(y)\, d\boldsymbol{\lambda}(y)$. Dann gilt*

$$P_{X+Y} = \int \left(\int_{\mathbb{R}} f(u-t)\, g(t)\, d\boldsymbol{\lambda}(t) \right) d\boldsymbol{\lambda}(u)$$

Die Faltungsformel bleibt natürlich richtig, wenn man auf der rechten Seite der Gleichung die Rollen von f und g vertauscht. Wir geben einige Beispiele für die Anwendung der Faltungsformel:

13.5.10 Beispiele (Absolutstetige Verteilungen). Seien X und Y unabhängige reelle Zufallsvariable.

(1) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha, \gamma)$ und $P_Y = \mathbf{Ga}(\alpha, \delta)$ gilt

$$P_{X+Y} = \mathbf{Ga}(\alpha, \gamma + \delta)$$

In der Tat: Es gilt

$$\begin{aligned} &P_{X+Y} \\ &\sim \int \left(\int_{\mathbb{R}} e^{-\alpha t} t^{\gamma-1} \chi_{(0,\infty)}(t) \cdot e^{-\alpha(u-t)} (u-t)^{\delta-1} \chi_{(0,\infty)}(u-t)\, d\boldsymbol{\lambda}(t) \right) d\boldsymbol{\lambda}(u) \end{aligned}$$

$$= \int e^{-\alpha u} u^{\gamma+\delta-1} \chi_{(0,\infty)}(u) \left(\int_{\mathbb{R}} \left(\frac{t}{u}\right)^{\gamma-1} \left(1-\frac{t}{u}\right)^{\delta-1} \frac{1}{u}\, \chi_{(0,u)}(t)\, d\boldsymbol{\lambda}(t) \right) d\boldsymbol{\lambda}(u)$$

$$= \int e^{-\alpha u} u^{\gamma+\delta-1} \chi_{(0,\infty)}(u) \left(\int_{\mathbb{R}} v^{\gamma-1}\,(1-v)^{\delta-1}\, \chi_{(0,1)}(v)\, d\boldsymbol{\lambda}(v) \right) d\boldsymbol{\lambda}(u)$$

$$\sim \int e^{-\alpha u} u^{\gamma+\delta-1} \chi_{(0,\infty)}(u)\, d\boldsymbol{\lambda}(u)$$

und damit $P_{X+Y} = \mathbf{Ga}(\alpha, \gamma+\delta)$; dabei wurde beim letzten Gleichheitszeichen für $u \in (0,\infty)$ die lineare Abbildung $S_u : \mathbb{R} \to \mathbb{R}$ mit $S_u(t) := t/u$ verwendet.

(2) **Normal–Verteilung:** Im Fall $P_X = \mathbf{N}(\mu, \sigma^2)$ und $P_Y = \mathbf{N}(\nu, \tau^2)$ gilt

$$P_{X+Y} = \mathbf{N}(\mu+\nu, \sigma^2+\tau^2)$$

Dies ergibt sich unmittelbar aus Beispiel 13.5.7.

Aufgaben

13.5.A **Normal–Verteilung:** Sei $P_{\mathbf{X}} = \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Dann gilt für alle $\mathbf{c} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$

$$P_{\mathbf{c}'\mathbf{X}} = \mathbf{N}(\mathbf{c}'\boldsymbol{\mu}, \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c})$$

13.5.B **Binomial–Verteilung:** Seien $X_1, \dots, X_m$ unabhängige reelle Zufallsvariable mit $P_{X_i} = \mathbf{B}(\vartheta)$ für alle $i \in \{1, \dots, m\}$. Dann gilt $P_{\sum_{i=1}^m X_i} = \mathbf{B}(m, \vartheta)$.

13.5.C **Erlang–Verteilung:** Seien $X_1, \dots, X_m$ unabhängige reelle Zufallsvariable mit $P_{X_i} = \mathbf{Exp}(\alpha)$ für alle $i \in \{1, \dots, m\}$. Dann gilt $P_{\sum_{i=1}^m X_i} = \mathbf{Ga}(\alpha, m)$.

13.5.D χ^2**–Verteilung:** Seien $X_1, \dots, X_m$ unabhängige standardnormal–verteilte reelle Zufallsvariable. Dann gilt $P_{\sum_{i=1}^m X_i^2} = \chi^2_m$.

13.5.E **Faltung:** Sei $\mathcal{Q}(\mathbb{R})$ die Familie aller Verteilungen $\mathcal{B}(\mathbb{R}) \to [0,1]$ und sei $T : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $T(\mathbf{x}) := \mathbf{1}'\mathbf{x}$. Die Abbildung $* : \mathcal{Q}(\mathbb{R}) \times \mathcal{Q}(\mathbb{R}) \to \mathcal{Q}(\mathbb{R})$ mit

$$Q * R := (Q \otimes R)_T$$

heißt *Faltung*.

(1) $(\mathcal{Q}(\mathbb{R}), *)$ ist eine kommutative Halbgruppe mit dem neutralen Element $\boldsymbol{\delta}_0$.
(2) Sind $Q, R \in \mathcal{Q}(\mathbb{R})$ absolutstetig, so ist auch $Q * R$ absolutstetig.
(3) Sind X und Y unabhängige Zufallsvariable, so gilt $P_{X+Y} = P_X * P_Y$.

13.6 Momente

Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor und sei $h : \mathbb{R}^m \to \mathbb{R}$ eine messbare Funktion. Ist $h \circ \mathbf{X}$ positiv oder integrierbar, so gilt nach der Substitutionsregel

$$E[h \circ \mathbf{X}] = \int_\Omega (h \circ \mathbf{X})(\omega)\, dP(\omega) = \int_{\mathbb{R}^m} h(\mathbf{x})\, dP_{\mathbf{X}}(\mathbf{x})$$

Insbesondere gilt:

– Besitzt X_i einen endlichen Erwartungswert, so gilt

$$E[X_i] = \int_{\mathbb{R}^m} x_i \, dP_{\mathbf{X}}(\mathbf{x})$$

– Besitzt $X_i X_j$ einen endlichen Erwartungswert, so gilt

$$E[X_i X_j] = \int_{\mathbb{R}^m} x_i x_j \, dP_{\mathbf{X}}(\mathbf{x})$$

Eine hinreichende Bedingung für die Existenz eines endlichen Erwartungswertes des Produktes $X_i X_j$ ist die Endlichkeit der zweiten Momente seiner Faktoren, denn nach der Ungleichung von Cauchy/Schwarz gilt

$$E[|X_i X_j|] \leq (E[X_i^2])^{1/2} (E[X_j^2])^{1/2}$$

Diese Bedingung ist aber nicht notwendig; vgl. Lemma 13.6.8.

– Besitzen alle Koordinaten von $\mathbf{X}$ einen endlichen Erwartungswert, so besitzt für alle $\mathbf{a} \in \mathbb{R}^m$ die Zufallsvariable $\mathbf{a}'\mathbf{X}$ einen endlichen Erwartungswert und es gilt

$$E[\mathbf{a}'\mathbf{X}] = \sum_{i=1}^{m} a_i \, E[X_i]$$

Dies folgt aus der Linearität des Integrals.

Besitzt $\mathbf{X}$ integrierbare Koordinaten $X_1, \ldots, X_m$, so setzen wir

$$E[\mathbf{X}] := (E[X_i])_{i \in \{1,\ldots,m\}}$$

und nennen $E[\mathbf{X}]$ die *Erwartung* von $\mathbf{X}$.

13.6.1 Lemma. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor mit integrierbaren Koordinaten. Dann gilt für alle $\mathbf{c} \in \mathbb{R}^k$ und für jede Matrix $\mathbf{D} \in \mathbb{R}^{k \times m}$*

$$E[\mathbf{c}+\mathbf{D}\mathbf{X}] = \mathbf{c} + \mathbf{D} \, E[\mathbf{X}]$$

Das Lemma ergibt sich unmittelbar aus der Linearität des Integrals.

13.6.2 Beispiele (Diskrete Verteilungen).

(1) **Polyhypergeometrische Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{PH}(n, N, \mathbf{K})$ gilt

$$E[\mathbf{X}] = \frac{n}{N} \, \mathbf{K}$$

(2) **Multinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{M}(n, \boldsymbol{\vartheta})$ gilt

$$E[\mathbf{X}] = n \, \boldsymbol{\vartheta}$$

(3) **Poisson–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{P}(\boldsymbol{\alpha})$ gilt

$$E[\mathbf{X}] = \boldsymbol{\alpha}$$

(4) **Negativmultinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$ gilt

$$E[\mathbf{X}] = \alpha \, \frac{1}{1 - \mathbf{1}'\boldsymbol{\vartheta}} \, \boldsymbol{\vartheta}$$

13.6.3 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{U}(C)$ mit

$$C := \left\{ \mathbf{x} \in \mathbb{R}^2 \,\middle|\, \left(x_1 - \tfrac{1}{2}\right)^2 + \left(x_2 - \tfrac{1}{2}\right)^2 < \tfrac{1}{4} \right\}$$

gilt

$$E[\mathbf{X}] = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

(2) **Dirichlet–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{Dir}(\eta, \boldsymbol{\eta})$ gilt

$$E[\mathbf{X}] = \frac{1}{\eta} \, \boldsymbol{\eta}$$

(3) **Normal–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ gilt

$$E[\mathbf{X}] = \boldsymbol{\mu}$$

Damit ist die Bedeutung des ersten Parameters der Normal–Verteilung geklärt.

Ist $\{U_{ij}\}_{i \in \{1,\dots,m\},\, j \in \{1,\dots,n\}}$ eine Familie von reellen Zufallsvariablen, so heißt die messbare Abbildung $\mathbf{U} : \Omega \to \mathbb{R}^{m \times n}$ mit

$$\mathbf{U} := (U_{ij})_{i \in \{1,\dots,m\},\, j \in \{1,\dots,n\}}$$

Zufallsmatrix mit den *Koordinaten* U_{ij}. Für eine Zufallsmatrix $\mathbf{U} : \Omega \to \mathbb{R}^{m \times n}$ mit integrierbaren Koordinaten setzen wir

$$E[\mathbf{U}] := (E[U_{ij}])_{i \in \{1,\dots,m\},\, j \in \{1,\dots,n\}}$$

und nennen $E[\mathbf{U}]$ die *Erwartung* von $\mathbf{U}$.

13.6.4 Lemma. *Sei* $\mathbf{U}$ *eine Zufallsmatrix mit integrierbaren Koordinaten. Dann gilt für alle Matrizen* $\mathbf{A}, \mathbf{B}$ *passender Dimension*

$$E[\mathbf{AUB}] = \mathbf{A} \, E[\mathbf{U}] \, \mathbf{B}$$

Für einen Zufallsvektor $\mathbf{X}$ mit quadratisch integrierbaren Koordinaten nennen wir $E[\mathbf{XX}']$ die *Matrix der zweiten gemischten Momente* von $\mathbf{X}$. Offenbar ist die Zufallsmatrix $\mathbf{XX}'$ und damit auch die Matrix $E[\mathbf{XX}']$ symmetrisch.

13.6.5 Lemma. *Sei* $\mathbf{X}$ *ein Zufallsvektor mit quadratisch integrierbaren Koordinaten. Dann gilt für jede Matrix* $\mathbf{D}$ *passender Dimension*

$$E[(\mathbf{DX})(\mathbf{DX})'] = \mathbf{D} \, E[\mathbf{XX}'] \, \mathbf{D}'$$

13.6.6 Beispiele (Diskrete Verteilungen).

(1) **Polyhypergeometrische Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{PH}(n, N, \mathbf{K})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$E[X_i X_j] = \frac{n(n-1)}{N(N-1)}\, K_i K_j$$

und damit $E[X_i X_j] \neq E[X_i]\, E[X_j]$.

(2) **Multinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{M}(n, \boldsymbol{\vartheta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$E[X_i X_j] = n(n-1)\, \vartheta_i \vartheta_j$$

und damit $E[X_i X_j] \neq E[X_i]\, E[X_j]$.
In der Tat: Es gilt $P_{X_i, X_j} = \mathbf{M}(n, \vartheta_i, \vartheta_j)$ und aufgrund der Permutationsinvarianz der Klasse der zweidimensionalen Multinomial–Verteilungen genügt es, die Gleichung für $i = 1$ und $j = 2$ zu beweisen. Es gilt

$$\begin{aligned}
E[X_1 X_2] &= \sum_{k=0}^{n} \sum_{l=0}^{n-k} k\, l\, \frac{n!}{(n-k-l)!\, k!\, l!}\, (1-\vartheta_1-\vartheta_2)^{n-k-l} \vartheta_1^k \vartheta_2^l \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n-k} k\, l\, \frac{n!}{(n-k-l)!\, k!\, l!}\, (1-\vartheta_1-\vartheta_2)^{n-k-l} \vartheta_1^k \vartheta_2^l \\
&= \sum_{k=1}^{n-1} \sum_{l=1}^{n-k} k\, l\, \frac{n!}{(n-k-l)!\, k!\, l!}\, (1-\vartheta_1-\vartheta_2)^{n-k-l} \vartheta_1^k \vartheta_2^l \\
&= n(n-1)\, \vartheta_1 \vartheta_2 \\
&\qquad \sum_{k=1}^{n-1} \sum_{l=1}^{n-k} \frac{(n-2)!}{(n-k-l)!\, (k-1)!\, (l-1)!}\, (1-\vartheta_1-\vartheta_2)^{n-k-l} \vartheta_1^{k-1} \vartheta_2^{l-1} \\
&= n(n-1)\, \vartheta_1 \vartheta_2 \\
&\qquad \sum_{p=0}^{n-2} \sum_{q=0}^{(n-2)-p} \frac{(n-2)!}{((n-2)-p-q)!p!\, q!}\, (1-\vartheta_1-\vartheta_2)^{(n-2)-p-q} \vartheta_1^p \vartheta_2^q \\
&= n(n-1)\, \vartheta_1 \vartheta_2
\end{aligned}$$

(3) **Poisson–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{P}(\boldsymbol{\alpha})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$E[X_i X_j] = \alpha_i \alpha_j$$

und damit $E[X_i X_j] = E[X_i]\, E[X_j]$.

(4) **Negativmultinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$E[X_i X_j] = \alpha(\alpha+1)\, \frac{\vartheta_i}{1 - \mathbf{1}'\boldsymbol{\vartheta}}\, \frac{\vartheta_j}{1 - \mathbf{1}'\boldsymbol{\vartheta}}$$

und damit $E[X_i X_j] \neq E[X_i]\, E[X_j]$.

13.6.7 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{U}(C)$ mit

$$C := \left\{ \mathbf{x} \in \mathbb{R}^2 \,\middle|\, \bigl(x_1 - \tfrac{1}{2}\bigr)^2 + \bigl(x_2 - \tfrac{1}{2}\bigr)^2 < \tfrac{1}{4} \right\}$$

gilt

$$E[X_1 X_2] = 1/4$$

und damit $E[X_1 X_2] = E[X_1]\, E[X_2]$.
In der Tat: Für $x_1 \in (0,1)$ sei $C(x_1) := \bigl(\tfrac{1}{2} - \sqrt{x_1(1-x_1)}, \tfrac{1}{2} + \sqrt{x_1(1-x_1)}\bigr)$. Dann gilt

$$\begin{aligned}
E[X_1 X_2] &= \int_{\mathbb{R}^2} x_1 x_2 \, dP_{\mathbf{X}}(\mathbf{x}) \\
&= \int_{\mathbb{R}^2} x_1 x_2 \, \frac{4}{\pi} \, \chi_C(\mathbf{x}) \, d\boldsymbol{\lambda}^2(\mathbf{x}) \\
&= \frac{4}{\pi} \int_{\mathbb{R}^2} x_1 x_2 \, \chi_{(0,1)}(x_1) \, \chi_{C(x_1)}(x_2) \, d\boldsymbol{\lambda}^2(\mathbf{x}) \\
&= \frac{4}{\pi} \int_{\mathbb{R}} x_1 \, \chi_{(0,1)}(x_1) \left(\int_{\mathbb{R}} x_2 \, \chi_{C(x_1)}(x_2) \, d\boldsymbol{\lambda}(x_2) \right) d\boldsymbol{\lambda}(x_1) \\
&= \frac{4}{\pi} \int_{\mathbb{R}} x_1 \, \chi_{(0,1)}(x_1) \, \sqrt{x_1(1-x_1)} \, d\boldsymbol{\lambda}(x_1) \\
&= \frac{4}{\pi} \, \mathrm{B}(\tfrac{5}{2}, \tfrac{3}{2}) \int_{\mathbb{R}} \frac{1}{\mathrm{B}(\frac{5}{2}, \frac{3}{2})} \, x_1^{\frac{5}{2}-1} (1-x_1)^{\frac{3}{2}-1} \, \chi_{(0,1)}(x_1) \, d\boldsymbol{\lambda}(x_1) \\
&= \frac{4}{\pi} \, \mathrm{B}(\tfrac{5}{2}, \tfrac{3}{2}) \\
&= \frac{1}{4}
\end{aligned}$$

(2) **Dirichlet–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{Dir}(\eta, \boldsymbol{\eta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$E[X_i X_j] = \frac{\eta_i \eta_j}{\eta(\eta+1)}$$

und damit $E[X_i X_j] \neq E[X_i]\, E[X_j]$.

Wir betrachten abschließend Zufallsvektoren mit unabhängigen Koordinaten:

13.6.8 Lemma. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor mit unabhängigen Koordinaten. Wenn alle Koordinaten positiv sind oder alle Koordinaten integrierbar sind, dann ist auch ihr Produkt positiv bzw. integrierbar und es gilt*

$$E\left[\prod_{i=1}^m X_i \right] = \prod_{i=1}^m E[X_i]$$

Beweis. Aus der Unabhängigkeit der Koordinaten und dem Satz von Fubini erhält man zunächst

$$\begin{aligned} E\left[\left|\prod_{i=1}^m X_i\right|\right] &= \int_{\mathbb{R}^m} \left|\prod_{i=1}^m x_i\right| dP_{\mathbf{X}}(\mathbf{x}) \\ &= \int_{\mathbb{R}^m} \left(\prod_{i=1}^m |x_i|\right) d\left(\bigotimes_{i=1}^m P_{X_i}\right)(\mathbf{x}) \\ &= \prod_{i=1}^m \int_{\mathbb{R}} |x_i|\, dP_{X_i}(x_i) \\ &= \prod_{i=1}^m E[|X_i|] \end{aligned}$$

Im Fall positiver Koordinaten folgt daraus bereits die Behauptung; im Fall integrierbarer Koordinaten folgt die Integrierbarkeit ihres Produktes, und die Wiederholung des Arguments liefert dann

$$E\left[\prod_{i=1}^m X_i\right] = \prod_{i=1}^m E[X_i]$$

Damit ist das Lemma bewiesen. □

Für integrierbare Zufallsvariable X und Y ist daher die Unabhängigkeit eine hinreichende Bedingung für die Gültigkeit der Gleichung $E[XY] = E[X]\,E[Y]$; andererseits zeigen die Beispiele 13.4.4(1) und 13.6.7(1), dass diese Bedingung nicht notwendig ist.

Aufgabe

13.6.A t–**Verteilung:** Sei

$$P_{\mathbf{X}} = \int \frac{1}{(2+x_1^2)^{3/2}\,(2+x_2^2)^{3/2}}\, d\boldsymbol{\lambda}^2(\mathbf{x})$$

Dann besitzt $X_1 X_2$ einen endlichen Erwartungswert und es gilt

$$E[X_1 X_2] = 0$$

13.7 Zentrale Momente

Für integrierbare Zufallsvariable X und Y, deren Produkt ebenfalls integrierbar ist, setzen wir

$$\mathrm{cov}\,[X,Y] := E[(X-E[X])(Y-E[Y])]$$

und nennen $\mathrm{cov}\,[X,Y]$ die *Kovarianz* von X und Y. Die folgenden Lemmata fassen die Eigenschaften der Kovarianz zusammen:

13.7.1 Lemma. *Seien X und Y integrierbare Zufallsvariable, deren Produkt ebenfalls integrierbar ist, und sei $a, b, c, d \in \mathbb{R}$. Dann gilt:*
(1) $\mathrm{cov}\,[X,X] = \mathrm{var}\,[X]$.
(2) $\mathrm{cov}\,[X,Y] = \mathrm{cov}\,[Y,X]$.
(3) $\mathrm{cov}\,[X,Y] = E[XY] - E[X]\,E[Y]$.
(4) $\mathrm{cov}\,[a{+}bX, c{+}dY] = bd\,\mathrm{cov}\,[X,Y]$.

Beweis. Die Aussagen (1) und (2) ergeben sich unmittelbar aus der Definition der Kovarianz und die Aussagen (3) und (4) folgen aus der Linearität des Erwartungswertes. □

13.7.2 Lemma. *Seien X und Y integrierbare Zufallsvariable. Sind X und Y unabhängig, so ist auch XY integrierbar und es gilt* $\mathrm{cov}\,[X,Y] = 0$.

Beweis. Die Behauptung ergibt sich unmittelbar aus Lemma 13.6.8. □

Die Kovarianz tritt in natürlicher Weise bei der Berechnung der Varianz einer Summe von zwei Zufallsvariablen auf:

13.7.3 Lemma. *Seien X und Y quadratisch integrierbare Zufallsvariable. Dann gilt*

$$\mathrm{var}\,[X{+}Y] = \mathrm{var}\,[X] + 2\,\mathrm{cov}\,[X,Y] + \mathrm{var}\,[Y]$$

Sind X und Y unabhängig, so gilt $\mathrm{var}\,[X{+}Y] = \mathrm{var}\,[X] + \mathrm{var}\,[Y]$.

Beweis. Die erste Aussage folgt aus der Linearität des Erwartungswertes und die zweite Aussage folgt dann aus Lemma 13.7.2. □

Für quadratisch integrierbare Zufallsvariable X und Y ist daher die Unabhängigkeit eine hinreichende Bedingung für die Gültigkeit der Gleichungen $\mathrm{cov}\,[X,Y] = 0$ und $\mathrm{var}\,[X{+}Y] = \mathrm{var}\,[X] + \mathrm{var}\,[Y]$; andererseits zeigt Beispiel 13.7.5(1), dass diese Bedingung nicht notwendig ist.

13.7.4 Beispiele (Diskrete Verteilungen).
(1) **Polyhypergeometrische Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{PH}(n, N, \mathbf{K})$ gilt für alle $i, j \in \{1, \dots, m\}$ mit $i \neq j$

$$\mathrm{cov}\,[X_i, X_j] = -\,n\,\frac{K_i}{N}\,\frac{K_j}{N}\,\frac{N-n}{N-1}$$

(2) **Multinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{M}(n, \boldsymbol{\vartheta})$ gilt für alle $i, j \in \{1, \dots, m\}$ mit $i \neq j$

$$\mathrm{cov}\,[X_i, X_j] = -\,n\,\vartheta_i\vartheta_j$$

In der Tat: Es gilt

$$\begin{aligned} E[X_i] &= n\,\vartheta_i \\ E[X_j] &= n\,\vartheta_j \\ E[X_iX_j] &= n(n{-}1)\,\vartheta_i\vartheta_j \end{aligned}$$

Daraus folgt die Behauptung.

(3) **Poisson–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{P}(\boldsymbol{\alpha})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\mathrm{cov}\,[X_i, X_j] = 0$$

(4) **Negativmultinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\mathrm{cov}\,[X_i, X_j] = \alpha\,\frac{\vartheta_i}{1 - \mathbf{1}'\boldsymbol{\vartheta}}\,\frac{\vartheta_j}{1 - \mathbf{1}'\boldsymbol{\vartheta}}$$

13.7.5 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{U}(C)$ mit

$$C := \left\{ \mathbf{x} \in \mathbb{R}^2 \;\middle|\; \left(x_1 - \tfrac{1}{2}\right)^2 + \left(x_2 - \tfrac{1}{2}\right)^2 < \tfrac{1}{4} \right\}$$

gilt $\mathrm{cov}\,[X_1, X_2] = 0$ (obwohl X_1 und X_2 nicht unabhängig sind).

(2) **Dirichlet–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{Dir}(\eta, \boldsymbol{\eta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\mathrm{cov}\,[X_i, X_j] = -\,\frac{\eta_i \eta_j}{\eta^2(\eta+1)}$$

Für einen Zufallsvektor $\mathbf{X} : \Omega \to \mathbb{R}^m$ mit quadratisch integrierbaren Koordinaten setzen wir

$$\mathrm{var}\,[\mathbf{X}] := (\mathrm{cov}\,[X_i, X_j])_{i,j \in \{1,\ldots,m\}}$$

und nennen $\mathrm{var}\,[\mathbf{X}]$ die *Varianz* von $\mathbf{X}$.

13.7.6 Lemma. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor mit quadratisch integrierbaren Koordinaten.*

(1) *Es gilt* $\mathrm{var}\,[\mathbf{X}] = E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'] = E[\mathbf{X}\mathbf{X}'] - E[\mathbf{X}]\,(E[\mathbf{X}])'$.

(2) *Für jeden Vektor $\mathbf{c} \in \mathbb{R}^k$ und jede Matrix $\mathbf{D} \in \mathbb{R}^{k \times m}$ gilt*

$$\mathrm{var}\,[\mathbf{c} + \mathbf{D}\mathbf{X}] = \mathbf{D}\,\mathrm{var}\,[\mathbf{X}]\,\mathbf{D}'$$

(3) *Die Matrix* $\mathrm{var}\,[\mathbf{X}]$ *ist symmetrisch und positiv semidefinit.*

(4) *Die Matrix* $\mathrm{var}\,[\mathbf{X}]$ *ist genau dann singulär, wenn es ein $\mathbf{d} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ und ein $c \in \mathbb{R}$ gibt mit $P[\{\mathbf{d}'\mathbf{X} = c\}] = 1$.*

(5) *Besitzt $\mathbf{X}$ unabhängige Koordinaten, so ist* $\mathrm{var}\,[\mathbf{X}]$ *eine Diagonalmatrix.*

Beweis. Aus den Eigenschaften der Kovarianz erhält man zunächst

$$\mathrm{var}\,[\mathbf{X}] = E[\mathbf{X}\mathbf{X}'] - E[\mathbf{X}](E[\mathbf{X}])'$$

und sodann

$$\begin{aligned}\mathrm{var}\,[\mathbf{c}+\mathbf{D}\mathbf{X}] &= \mathrm{var}\,[\mathbf{D}\mathbf{X}] \\ &= E[(\mathbf{D}\mathbf{X})(\mathbf{D}\mathbf{X})'] - E[\mathbf{D}\mathbf{X}](E[\mathbf{D}\mathbf{X}])' \\ &= \mathbf{D}E[\mathbf{X}\mathbf{X}']\mathbf{D}' - \mathbf{D}E[\mathbf{X}](E[\mathbf{X}])'\mathbf{D}' \\ &= \mathbf{D}\,\mathrm{var}\,[\mathbf{X}]\,\mathbf{D}'\end{aligned}$$

Damit sind (1) und (2) gezeigt. Die Symmetrie von $\mathrm{var}\,[\mathbf{X}]$ ist klar. Wegen (2) gilt für alle $\mathbf{d} \in \mathbb{R}^m$

$$\mathbf{d}'\,\mathrm{var}\,[\mathbf{X}]\,\mathbf{d} = \mathrm{var}\,[\mathbf{d}'\mathbf{X}] \geq 0$$

Daher ist $\mathrm{var}\,[\mathbf{X}]$ positiv semidefinit. Damit ist (3) gezeigt. Außerdem sind folgende Aussagen äquivalent:
- $\mathrm{var}\,[\mathbf{X}]$ ist singulär.
- Es gibt ein $\mathbf{d} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ mit $\mathbf{d}'\,\mathrm{var}\,[\mathbf{X}]\,\mathbf{d} = 0$.
- Es gibt ein $\mathbf{d} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ mit $\mathrm{var}\,[\mathbf{d}'\mathbf{X}] = 0$.
- Es gibt ein $\mathbf{d} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ und ein $c \in \mathbb{R}$ mit $P[\{\mathbf{d}'\mathbf{X} = c\}] = 1$.

Damit ist (4) gezeigt. Schließlich folgt (5) aus Lemma 13.7.2. □

13.7.7 Beispiele (Diskrete Verteilungen).

(1) **Polyhypergeometrische Verteilung:** Im Fall $P_\mathbf{X} = \mathbf{PH}(n, N, \mathbf{K})$ gilt

$$\mathrm{var}\,[\mathbf{X}] = n\,\frac{N-n}{N-1}\left(\frac{1}{N}\,\mathrm{diag}(\mathbf{K}) - \frac{1}{N^2}\,\mathbf{K}\mathbf{K}'\right)$$

(2) **Multinomial–Verteilung:** Im Fall $P_\mathbf{X} = \mathbf{M}(n, \boldsymbol{\vartheta})$ gilt

$$\mathrm{var}\,[\mathbf{X}] = n\Big(\mathrm{diag}(\boldsymbol{\vartheta}) - \boldsymbol{\vartheta}\boldsymbol{\vartheta}'\Big)$$

(3) **Poisson–Verteilung:** Im Fall $P_\mathbf{X} = \mathbf{P}(\boldsymbol{\alpha})$ gilt

$$\mathrm{var}\,[\mathbf{X}] = \mathrm{diag}(\boldsymbol{\alpha})$$

(4) **Negativmultinomial–Verteilung:** Im Fall $P_\mathbf{X} = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$ gilt

$$\mathrm{var}\,[\mathbf{X}] = \alpha\left(\frac{1}{1-\mathbf{1}'\boldsymbol{\vartheta}}\,\mathrm{diag}(\boldsymbol{\vartheta}) + \frac{1}{(1-\mathbf{1}'\boldsymbol{\vartheta})^2}\,\boldsymbol{\vartheta}\boldsymbol{\vartheta}'\right)$$

13.7.8 Beispiele (Absolutstetige Verteilungen).

(1) **Uniforme Verteilung:** Im Fall $P_\mathbf{X} = \mathbf{U}(C)$ mit

$$C := \left\{\mathbf{x} \in \mathbb{R}^2 \;\middle|\; \left(x_1 - \tfrac{1}{2}\right)^2 + \left(x_2 - \tfrac{1}{2}\right)^2 < \tfrac{1}{4}\right\}$$

gilt

$$\mathrm{var}\,[\mathbf{X}] = \begin{pmatrix} 1/16 & 0 \\ 0 & 1/16 \end{pmatrix}$$

(obwohl die Koordinaten von $\mathbf{X}$ nicht unabhängig sind).

(2) **Dirichlet–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{Dir}(\eta, \boldsymbol{\eta})$ gilt

$$\mathrm{var}\,[\mathbf{X}] = \frac{1}{\eta(\eta+1)}\,\mathrm{diag}(\boldsymbol{\eta}) - \frac{1}{\eta^2(\eta+1)}\,\boldsymbol{\eta}\boldsymbol{\eta}'$$

(3) **Normal–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ gilt

$$\mathrm{var}\,[\mathbf{X}] = \boldsymbol{\Sigma}$$

Damit ist die Bedeutung des zweiten Parameters der Normal–Verteilung geklärt.
In der Tat: Nach Proposition C.1.3 gibt es eine invertierbare Matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ mit $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}'$. Sei

$$\mathbf{Z} := \mathbf{A}^{-1}(\mathbf{X}-\boldsymbol{\mu})$$

Nach Beispiel 13.2.2 gilt $P_{\mathbf{Z}} = \mathbf{N}(\mathbf{0}, \mathbf{I})$, und aus Beispiel 13.3.3(3) sowie Beispiel 13.4.4(3) und Lemma 13.7.2 folgt

$$\mathrm{var}\,[\mathbf{Z}] = \mathbf{I}$$

Wegen $\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}\mathbf{Z}$ folgt nun aus Lemma 13.7.6

$$\mathrm{var}\,[\mathbf{X}] = \mathrm{var}\,[\boldsymbol{\mu}+\mathbf{A}\mathbf{Z}] = \mathbf{A}\,\mathrm{var}\,[\mathbf{Z}]\,\mathbf{A}' = \mathbf{A}\mathbf{I}\mathbf{A}' = \mathbf{A}\mathbf{A}' = \boldsymbol{\Sigma}$$

Insbesondere gilt: $\mathrm{var}\,[\mathbf{X}]$ ist genau dann eine Diagonalmatrix, wenn die Koordinaten von $\mathbf{X}$ unabhängig sind (vgl. Beispiel 13.4.4).

Aus den Eigenschaften der Varianz eines Zufallsvektors ergibt sich eine wahrscheinlichkeitstheoretische Version der aus der Integrationstheorie bekannten Ungleichung von Cauchy/Schwarz:

13.7.9 Folgerung (Cauchy/Schwarz). *Seien X und Y quadratisch integrierbare Zufallsvariable. Dann gilt*

$$(\mathrm{cov}\,[X,Y])^2 \leq \mathrm{var}\,[X]\,\mathrm{var}\,[Y]$$

Außerdem gilt $(\mathrm{cov}\,[X,Y])^2 = \mathrm{var}\,[X]\,\mathrm{var}\,[Y]$ *genau dann, wenn es $a, b, c \in \mathbb{R}$ gibt mit $a \neq 0$ oder $b \neq 0$ und $P[\{aX+bY = c\}] = 1$.*

Beweis. Für den Zufallsvektor $\mathbf{X} : \Omega \to \mathbb{R}^2$ mit $X_1 := X$ und $X_2 := Y$ gilt

$$\det(\mathrm{var}\,[\mathbf{X}]) = \mathrm{var}\,[X]\,\mathrm{var}\,[Y] - (\mathrm{cov}\,[X,Y])^2$$

Nach Lemma 13.7.6 ist $\mathrm{var}\,[\mathbf{X}]$ positiv semidefinit. Daher gilt $\det(\mathrm{var}\,[\mathbf{X}]) \geq 0$, und die Ungleichung folgt.
Des weiteren ist die Bedingung $(\mathrm{cov}\,[X,Y])^2 = \mathrm{var}\,[X]\,\mathrm{var}\,[Y]$ gleichwertig mit $\det(\mathrm{var}\,[\mathbf{X}]) = 0$, also der Singularität von $\mathrm{var}\,[\mathbf{X}]$. Nach Lemma 13.7.6 ist aber $\mathrm{var}\,[\mathbf{X}]$ genau dann singulär, wenn es $a, b, c \in \mathbb{R}$ gibt mit $a \neq 0$ oder $b \neq 0$ und $P[\{aX+bY = c\}] = 1$. $\square$

Wir kommen nochmals auf die Kovarianz zurück: Zwei integrierbare Zufallsvariable X und Y, deren Produkt ebenfalls integrierbar ist, heißen

– *strikt positiv korreliert*, wenn $\text{cov}\,[X, Y] > 0$ gilt.
– *positiv korreliert*, wenn $\text{cov}\,[X, Y] \geq 0$ gilt.
– *unkorreliert*, wenn $\text{cov}\,[X, Y] = 0$ gilt.
– *negativ korreliert*, wenn $\text{cov}\,[X, Y] \leq 0$ gilt.
– *strikt negativ korreliert*, wenn $\text{cov}\,[X, Y] < 0$ gilt.

Nach Lemma 13.7.2 sind unabhängige integrierbare Zufallsvariable unkorreliert; andererseits zeigt Beispiel 13.7.5(1), dass unkorrelierte Zufallsvariable nicht unabhängig sein müssen.

Der in diesen Definitionen verwendete Begriff der Korrelation erklärt sich aus der folgenden Definition:

Für quadratisch integrierbare Zufallsvariable X und Y mit $\text{var}\,[X] \neq 0$ und $\text{var}\,[Y] \neq 0$ setzen wir

$$\varrho_{X,Y} := \frac{\text{cov}\,[X, Y]}{\sqrt{\text{var}\,[X]}\,\sqrt{\text{var}\,[Y]}}$$

und nennen $\varrho_{X,Y}$ den *Korrelationskoeffizienten* von X und Y. Im Gegensatz zur Kovarianz hat der Korrelationskoeffizient den Vorteil, dass sein Wert stets im Intervall $[-1, 1]$ liegt und die beiden extremen Werte eine Interpretation besitzen:

13.7.10 Lemma. *Seien X und Y quadratisch integrierbare Zufallsvariable mit* $\text{var}\,[X] \neq 0 \neq \text{var}\,[Y]$. *Dann gilt*

$$\varrho_{X,Y} = \text{cov}\left[\frac{X - E[X]}{\sqrt{\text{var}\,[X]}}, \frac{Y - E[Y]}{\sqrt{\text{var}\,[Y]}}\right]$$

und

$$|\varrho_{X,Y}| \leq 1$$

Außerdem gilt $|\varrho_{X,Y}| = 1$ genau dann, wenn es $a, b, c \in \mathbb{R}$ gibt mit $a \neq 0$ oder $b \neq 0$ und $P[\{aX + bY = c\}] = 1$.

Beweis. Die erste Gleichung ergibt sich aus Lemma 13.7.1 und der Rest folgt aus der Ungleichung von Cauchy/Schwarz. □

13.7.11 Beispiele (Diskrete Verteilungen).

(1) **Polyhypergeometrische Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{PH}(n, N, \mathbf{K})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\varrho_{X_i, X_j} = -\sqrt{\frac{K_i}{N - K_i}\,\frac{K_j}{N - K_j}}$$

(2) **Multinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{M}(n, \boldsymbol{\vartheta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\varrho_{X_i,X_j} = -\sqrt{\frac{\vartheta_i}{1-\vartheta_i} \frac{\vartheta_j}{1-\vartheta_j}}$$

(3) **Poisson–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{P}(\boldsymbol{\alpha})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\varrho_{X_i,X_j} = 0$$

(4) **Negativmultinomial–Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{NM}(\alpha, \boldsymbol{\vartheta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\varrho_{X_i,X_j} = \sqrt{\frac{\vartheta_i}{\vartheta_i + 1 - \mathbf{1}'\boldsymbol{\vartheta}} \frac{\vartheta_j}{\vartheta_j + 1 - \mathbf{1}'\boldsymbol{\vartheta}}}$$

13.7.12 Beispiele (Absolutstetige Verteilungen).
(1) **Uniforme Verteilung:** Im Fall $P_{\mathbf{X}} = \mathbf{U}(C)$ mit

$$C := \left\{\mathbf{x} \in \mathbb{R}^2 \;\middle|\; \left(x_1 - \tfrac{1}{2}\right)^2 + \left(x_2 - \tfrac{1}{2}\right)^2 < \tfrac{1}{4}\right\}$$

gilt $\varrho_{X_1,X_2} = 0$.
(2) **Dirichlet–Verteilung:** Im Fall $Q = \mathbf{Dir}(\eta, \boldsymbol{\eta})$ gilt für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$\varrho_{X_i,X_j} = -\sqrt{\frac{\eta_i}{\eta - \eta_i} \frac{\eta_j}{\eta - \eta_j}}$$

Der Korrelationskoeffizient ist ein Maß für den Grad des affinen Zusammenhangs zwischen zwei Zufallsvariablen; andere funktionale Zusammenhänge zwischen zwei Zufallsvariablen können jedoch auch dann bestehen, wenn die Zufallsvariablen unkorreliert sind:

13.7.13 Beispiel. Sei X eine reelle Zufallsvariable mit $P_X = \mathbf{N}(0,1)$ und sei $Y := X^2$. Dann gilt $E[X] = 0$ und $E[XY] = E[X^3] = 0$, und damit $\mathrm{cov}\,[X,Y] = 0$ und damit

$$\varrho_{X,Y} = 0$$

obwohl Y eine Funktion von X ist.

Aufgaben

13.7.A **Kovarianz:** Sei $h : \mathbb{R} \to \mathbb{R}$ eine monoton wachsende Funktion. Ist X eine reelle Zufallsvariable derart, dass X und $h(X)$ ein endliches zweites Moment besitzen, so gilt $\mathrm{cov}\,[X, h(X)] \geq 0$.

13.7.B **Kovarianz:** Für Zufallsvektoren $\mathbf{X} : \Omega \to \mathbb{R}^m$ und $\mathbf{Y} : \Omega \to \mathbb{R}^n$, deren Koordinaten alle ein endliches zweites Moment besitzen, heißt die Matrix

$$\mathrm{cov}\,[\mathbf{X}, \mathbf{Y}] := (\mathrm{cov}\,[X_i, Y_j])_{i \in \{1,\dots,m\}, j \in \{1,\dots,n\}}$$

die *Kovarianz* von $\mathbf{X}$ und $\mathbf{Y}$. Für alle Vektoren $\mathbf{a}, \mathbf{c}$ und Matrizen $\mathbf{B}, \mathbf{D}$ passender Dimension gilt

$$\mathrm{cov}\,[\mathbf{a}+\mathbf{B}\mathbf{X}, \mathbf{c}+\mathbf{D}\mathbf{Y}] = \mathbf{B}\,\mathrm{cov}\,[\mathbf{X}, \mathbf{Y}]\,\mathbf{D}'$$

Im Fall $\mathbf{X} = \mathbf{Y}$ gilt $\mathrm{cov}\,[\mathbf{X}, \mathbf{Y}] = \mathrm{var}\,[\mathbf{X}]$.

13.7.C **Korrelationskoeffizient:** Seien X und Y quadratisch integrierbare Zufallsvariable mit $\mathrm{var}\,[X] \neq 0 \neq \mathrm{var}\,[Y]$.

(1) Für alle $a, b, c, d \in \mathbb{R}$ gilt

$$\varrho_{a+bX, c+dY} = \varrho_{X,Y}$$

(2) Es gilt $\varrho_{X,Y} = 1$ genau dann, wenn

$$P\left[\left\{ \frac{X - E[X]}{\sqrt{\mathrm{var}\,[X]}} = \frac{Y - E[Y]}{\sqrt{\mathrm{var}\,[Y]}} \right\}\right] = 1$$

(3) Es gilt $\varrho_{X,Y} = -1$ genau dann, wenn

$$P\left[\left\{ \frac{X - E[X]}{\sqrt{\mathrm{var}\,[X]}} = -\frac{Y - E[Y]}{\sqrt{\mathrm{var}\,[Y]}} \right\}\right] = 1$$

13.7.D Seien X und Y reelle Zufallsvariable mit $P_{X,Y} = \mathbf{M}(n, \vartheta, 1-\vartheta)$. Dann gilt

$$\varrho_{X,Y} = -1$$

Leiten Sie dieses Ergebnis ohne Verwendung von Beispiel 13.7.11 her.

13.7.E Zwei Bernoulli–verteilte reelle Zufallsvariable sind genau dann unkorreliert, wenn sie unabhängig sind.

13.7.F Geben Sie ein Beispiel für drei Bernoulli–verteilte reelle Zufallsvariable, die paarweise unkorreliert, aber in ihrer Gesamtheit nicht unabhängig sind.

13.7.G Seien X und Y reelle Zufallsvariable mit $P_X = \mathbf{P}(\alpha)$ und $P_Y = \mathbf{P}(\beta)$.

(1) Sind X und Y unabhängig, so gilt $P_{X+Y} = \mathbf{P}(\alpha+\beta)$.

(2) Gilt $P_{X+Y} = \mathbf{P}(\alpha+\beta)$, so sind X und Y unkorreliert.

13.7.H **Normal–Verteilung:** Sei $\mathbf{X}$ ein Zufallsvektor mit $P_{\mathbf{X}} = \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Dann sind äquivalent:

(a) $\mathbf{X}$ besitzt unabhängige Koordinaten.

(b) $\mathbf{X}$ besitzt unkorrelierte Koordinaten.

14

Konvergenz von Folgen von Zufallsvariablen

Zur Vorbereitung auf die Gesetze der Großen Zahlen, die wir im nächsten Kapitel behandeln, vergleichen wir in diesem Kapitel die aus der Maß– und Integrationstheorie bekannten drei Arten der Konvergenz für eine Folge von Zufallsvariablen:

– die Konvergenz P–fast überall, die hier als P–fast sichere Konvergenz bezeichnet wird (Abschnitt 14.1),
– die Konvergenz im Maß P, die hier als stochastische Konvergenz bezeichnet wird (Abschnitt 14.2), und
– die Konvergenz im p–ten Mittel bezüglich P für $p \in [1, \infty)$ (Abschnitt 14.3).

Gegenüber der allgemeinen Theorie ergeben sich Besonderheiten daraus, dass P ein Wahrscheinlichkeitsmaß ist. Eine weitere Art der Konvergenz behandeln wir später in Kapitel 17.

14.1 Fast sichere Konvergenz

Von einer Eigenschaft, die für jedes $\omega \in \Omega$ entweder gilt oder nicht gilt, sagt man, sie gelte *P–fast sicher* oder kurz *fast sicher*, wenn sie P–fast überall gilt. Diese Ausdrucksweise entspricht der Bezeichnung der Ergebnismenge Ω als sicheres Ereignis.

Insbesondere heißt eine Folge $\{X_n\}_{n\in\mathbb{N}}$ von Zufallsvariablen *P–fast sicher konvergent* oder kurz *fast sicher konvergent*, wenn sie P–fast überall konvergent ist. Da $\liminf_{n\to\infty} X_n$ und $\limsup_{n\to\infty} X_n$ messbar sind, gilt

$$\left\{ \liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n \right\} \in \mathcal{F}$$

Daher ist die Folge $\{X_n\}_{n\in\mathbb{N}}$ genau dann P–fast sicher konvergent, wenn

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_14, © Springer-Verlag Berlin Heidelberg 2011

$$P\left[\left\{\liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n\right\}\right] = 1$$

gilt, und in diesem Fall sagt man auch, dass die Folge $\{X_n\}_{n\in\mathbb{N}}$ *mit Wahrscheinlichkeit Eins konvergent* ist.

Für eine Folge von reellen Zufallsvariablen lässt sich die fast sichere Konvergenz gegen eine ebenfalls reelle Zufallsvariable wie folgt charakterisieren:

14.1.1 Lemma. *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen. Dann sind äquivalent*:
(a) *Es gilt* $P[\{-\infty < \liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n < \infty\}] = 1$.
(b) *Die Folge $\{X_n\}_{n\in\mathbb{N}}$ konvergiert fast sicher gegen eine reelle Zufallsvariable.*
(c) *Für alle $\varepsilon \in (0,\infty)$ gilt* $\lim_{m\to\infty} P[\{\sup_{n\in\mathbb{N}(m)} |X_n - X_m| \geq \varepsilon\}] = 0$.

Beweis. Die Äquivalenz von (a) und (b) ist klar. Wir zeigen nun die Äquivalenz von (b) und (c).
Für $\omega \in \Omega$ konvergiert die Folge $\{X_n(\omega)\}_{n\in\mathbb{N}}$ genau dann gegen eine reelle Zahl, wenn sie eine Cauchy–Folge ist, und dies ist genau dann der Fall, wenn es zu jedem $k \in \mathbb{N}$ ein $m \in \mathbb{N}$ gibt mit $\sup_{n\in\mathbb{N}(m)} |X_n(\omega) - X_m(\omega)| \leq 1/k$.
Daher konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ genau dann fast sicher gegen eine reelle Zufallsvariable, wenn

$$P\left[\bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \left\{\sup_{n\in\mathbb{N}(m)} |X_n - X_m| \leq 1/k\right\}\right] = 1$$

gilt. Da die Folge $\{A_k\}_{k\in\mathbb{N}}$ mit $A_k := \bigcup_{m\in\mathbb{N}}\{\sup_{n\in\mathbb{N}(m)} |X_n - X_m| \leq 1/k\}$ monoton fallend ist, gilt die letzte Gleichung genau dann, wenn für alle $k \in \mathbb{N}$

$$P\left[\bigcup_{m\in\mathbb{N}} \left\{\sup_{n\in\mathbb{N}(m)} |X_n - X_m| \leq 1/k\right\}\right] = 1$$

gilt, und dies ist, da für jedes $k \in \mathbb{N}$ die Folge $\{C_{k,m}\}_{m\in\mathbb{N}}$ mit $C_{k,m} := \{\sup_{n\in\mathbb{N}(m)} |X_n - X_m| \leq 1/k\}$ monoton wachsend ist, gleichwertig damit, dass für alle $k \in \mathbb{N}$

$$\lim_{m\to\infty} P\left[\left\{\sup_{n\in\mathbb{N}(m)} |X_n - X_m| \leq 1/k\right\}\right] = 1$$

gilt. □

In völlig analoger Weise lässt sich die fast sichere Konvergenz einer Folge von reellen Zufallsvariablen gegen eine gegebene reelle Zufallsvariable charakterisieren:

14.1.2 Lemma. *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen und sei X eine reelle Zufallsvariable. Dann sind äquivalent*:
(a) *Die Folge $\{X_n\}_{n\in\mathbb{N}}$ konvergiert fast sicher gegen X.*
(b) *Für alle $\varepsilon \in (0,\infty)$ gilt* $\lim_{m\to\infty} P[\{\sup_{n\in\mathbb{N}(m)} |X_n - X| \geq \varepsilon\}] = 0$.

Aufgabe

14.1.A Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen und sei X eine reelle Zufallsvariable. Dann sind äquivalent:
(a) Die Folge $\{X_n\}_{n\in\mathbb{N}}$ konvergiert fast sicher gegen X.
(b) Für alle $\varepsilon \in (0,\infty)$ gilt $P[\liminf_{n\to\infty}\{|X_n - X| \leq \varepsilon\}] = 1$.

14.2 Stochastische Konvergenz

Eine Folge $\{X_n\}_{n\in\mathbb{N}}$ von reellen Zufallsvariablen *konvergiert stochastisch* gegen eine reelle Zufallsvariable X, wenn sie im Maß P gegen X konvergiert. Nach Definition der Konvergenz im Maß bedeutet dies, dass für alle $\varepsilon \in (0,\infty)$

$$\lim_{n\to\infty} P[\{|X_n - X| \geq \varepsilon\}] = 0$$

gilt. Der Vergleich dieser Bedingung mit der in Lemma 14.1.2 angegebenen Charakterisierung der fast sicheren Konvergenz liefert erneut das folgende Ergebnis, das bereits aus Satz 7.2.8 bekannt ist:

14.2.1 Lemma. *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen und sei X eine reelle Zufallsvariable. Wenn die Folge $\{X_n\}_{n\in\mathbb{N}}$ fast sicher gegen X konvergiert, dann konvergiert sie auch stochastisch gegen X.*

Das folgende Beispiel zeigt, dass die stochastische Konvergenz schwächer ist als die fast sichere Konvergenz:

14.2.2 Beispiel (Wandernde Türme). Sei $(\Omega,\mathcal{F},P) := ((0,1],\mathcal{B}(0,1],\boldsymbol{\lambda}|_{\mathcal{B}(0,1]})$. Für $m \in \mathbb{N}$ und $k \in \{1,\dots,2^m\}$ sei

$$B_{m,k} := ((k-1)\,2^{-m}, k\,2^{-m}]$$

und

$$X_{2^m+k-2} := \chi_{B_{m,k}}$$

Dann gilt für alle $\varepsilon \in (0,\infty)$

$$\lim_{n\to\infty} P[\{X_n \geq \varepsilon\}] = 0$$

Daher konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ stochastisch gegen 0. Andererseits gilt für alle $\omega \in \Omega$

$$\liminf_{n\to\infty} X_n(\omega) = 0 < 1 = \limsup_{n\to\infty} X_n(\omega)$$

Daher konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ nicht fast sicher gegen 0.

Andererseits besitzt jede Folge von reellen Zufallsvariablen, die stochastisch gegen eine reelle Zufallsvariable konvergiert, immerhin eine Teilfolge, die fast sicher gegen dieselbe reelle Zufallsvariable konvergiert; dies ergibt sich aus dem folgenden Satz, der die stochastische Konvergenz durch die fast sichere Konvergenz von Teilfolgen charakterisiert:

14.2.3 Satz (Teilfolgenprinzip). *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen und sei X eine reelle Zufallsvariable. Dann sind äquivalent:*
(a) *Die Folge $\{X_n\}_{n\in\mathbb{N}}$ konvergiert stochastisch gegen X.*
(b) *Jede Teilfolge von $\{X_n\}_{n\in\mathbb{N}}$ besitzt eine Teilfolge, die fast sicher gegen X konvergiert.*

Beweis. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $X = 0$ gilt.
Wir nehmen zunächst an, dass (a) gilt. Da mit der Folge $\{X_n\}_{n\in\mathbb{N}}$ auch jede Teilfolge stochastisch gegen 0 konvergiert, genügt es zu zeigen, dass die Folge $\{X_n\}_{n\in\mathbb{N}}$ eine Teilfolge besitzt, die fast sicher gegen 0 konvergiert. Dazu setzen wir $n_0 := 0$ und wählen induktiv zu jedem $k \in \mathbb{N}$ ein $n_k \in \mathbb{N}$ mit $n_{k-1} < n_k$ und

$$P[\{|X_{n_k}| > 1/k\}] \le 2^{-(k+1)}$$

Sei nun $\varepsilon \in (0,\infty)$. Dann gilt für alle $m \in \mathbb{N}$ mit $m > 1/\varepsilon$

$$\begin{aligned} P\left[\left\{\sup_{k\in\mathbb{N}(m)} |X_{n_k}| \ge \varepsilon\right\}\right] &\le P\left[\left\{\sup_{k\in\mathbb{N}(m)} |X_{n_k}| > 1/m\right\}\right] \\ &= P\left[\bigcup_{k\in\mathbb{N}(m)} \{|X_{n_k}| > 1/m\}\right] \\ &\le \sum_{k\in\mathbb{N}(m)} P[\{|X_{n_k}| > 1/m\}] \\ &\le \sum_{k\in\mathbb{N}(m)} P[\{|X_{n_k}| > 1/k\}] \\ &\le \sum_{k\in\mathbb{N}(m)} 2^{-(k+1)} \\ &= 2^{-m} \end{aligned}$$

und damit

$$\lim_{m\to\infty} P\left[\left\{\sup_{k\in\mathbb{N}(m)} |X_{n_k}| \ge \varepsilon\right\}\right] = 0$$

Aus Lemma 14.1.2 folgt nun, dass die Teilfolge $\{X_{n_k}\}_{k\in\mathbb{N}}$ fast sicher gegen 0 konvergiert. Damit ist gezeigt, dass (b) aus (a) folgt.

Wir nehmen nun an, dass (b) gilt. Nach dem Teilfolgenprinzip der Analysis konvergiert eine Folge $\{a_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ genau dann gegen 0, wenn jede Teilfolge von $\{a_n\}_{n\in\mathbb{N}}$ eine Teilfolge besitzt, die gegen 0 konvergiert. Sei nun $\varepsilon \in (0,\infty)$. Nach Voraussetzung besitzt jede Teilfolge der Folge $\{X_n\}_{n\in\mathbb{N}}$ eine Teilfolge, die fast sicher und nach Lemma 14.2.1 auch stochastisch gegen 0 konvergiert; daher besitzt jede Teilfolge der Folge $\{a_n\}_{n\in\mathbb{N}}$ mit

$$a_n := P[\{|X_n| \geq \varepsilon\}]$$

eine Teilfolge, die gegen 0 konvergiert, und daraus folgt die Konvergenz der Folge $\{a_n\}_{n\in\mathbb{N}}$ gegen 0, also

$$\lim_{n\to\infty} P[\{|X_n| \geq \varepsilon\}] = \lim_{n\to\infty} a_n = 0$$

Da $\varepsilon \in (0,\infty)$ beliebig war, ist damit gezeigt, dass (a) aus (b) folgt. □

Aufgaben

14.2.A Eine Folge $\{X_n\}_{n\in\mathbb{N}}$ von reellen Zufallsvariablen konvergiert genau dann stochastisch gegen eine reelle Zufallsvariable X, wenn für alle $\varepsilon \in (0,\infty)$

$$\lim_{n\to\infty} P[\{|X_n - X| \leq \varepsilon\}] = 1$$

gilt.

14.2.B Sei $\{X_n\}_{n\in\mathbb{N}}$ eine unabhängige Folge von reellen Zufallsvariablen mit $P_{X_n} = \mathbf{B}(1/n)$ für alle $n \in \mathbb{N}$. Dann konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ stochastisch gegen 0, aber sie ist nicht fast sicher konvergent.

14.2.C Sei $\{\mathbf{X}_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvektoren $\Omega \to \mathbb{R}^m$ und sei $\mathbf{X} : \Omega \to \mathbb{R}^m$ ein Zufallsvektor. Wenn für alle $i \in \{1,\ldots,m\}$ die Folge $\{\pi_i \circ \mathbf{X}_n\}_{n\in\mathbb{N}}$ stochastisch gegen $\pi_i \circ \mathbf{X}$ konvergiert, dann konvergiert für jede stetige Funktion $h : \mathbb{R}^m \to \mathbb{R}$ die Folge $\{h \circ \mathbf{X}_n\}_{n\in\mathbb{N}}$ stochastisch gegen $h \circ \mathbf{X}$.

14.3 Konvergenz im p–ten Mittel

Für $p \in [1,\infty]$ setzen wir

$$L^p := L^p(\mathcal{F}, P)$$

Jede Zufallsvariable in einer Äquivalenzklasse von L^p ist fast sicher reell. Wir erinnern daran, dass für $p \in [1,\infty)$ eine Folge von Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}}$ in L^p im p–ten Mittel gegen eine Zufallsvariable $X \in L^p$ konvergiert, wenn

$$\lim_{n\to\infty} \left(\int_\Omega |X_n - X|^p \, dP\right)^{1/p} = 0$$

gilt; diese Bedingung ist offenbar gleichwertig mit

$$\lim_{n\to\infty} E[|X_n - X|^p] = 0$$

Das folgende Ergebnis ergibt sich unmittelbar aus Satz 8.4.8:

14.3.1 Satz. *Sei $p \in [1, \infty)$. Dann konvergiert jede Folge von reellen Zufallsvariablen in L^p, die im p–ten Mittel gegen eine reelle Zufallsvariable $X \in L^p$ konvergiert, auch stochastisch gegen X.*

Das folgende Beispiel zeigt, dass die stochastische Konvergenz schwächer ist als die Konvergenz im p–ten Mittel:

14.3.2 Beispiel. Sei $(\Omega, \mathcal{F}, P) := ((0,1], \mathcal{B}(0,1], \boldsymbol{\lambda}|_{\mathcal{B}(0,1]})$ und sei $p \in [1, \infty)$. Für $n \in \mathbb{N}$ sei

$$X_n := n^{1/p} \chi_{(0,1/n]}$$

Dann gilt für alle $\varepsilon \in (0, \infty)$

$$\lim_{n \to \infty} P[\{X_n \geq \varepsilon\}] = 0$$

Daher konvergiert die Folge $\{X_n\}_{n \in \mathbb{N}}$ stochastisch gegen 0. Andererseits gilt

$$\lim_{n \to \infty} E[X_n^p] = 1$$

Daher konvergiert die Folge $\{X_n\}_{n \in \mathbb{N}}$ nicht im p–ten Mittel gegen 0.

Die in Beispiel 14.3.2 betrachtete Folge konvergiert sogar fast sicher gegen 0. Sie ist daher gleichzeitig ein Beispiel dafür, dass eine fast sicher konvergente Folge nicht notwendigerweise auch im p–ten Mittel konvergent ist, und sie verletzt offenbar die Bedingungen des Satzes über die monotone Konvergenz und des Satzes über die majorisierte Konvergenz, die beide hinreichend dafür sind, dass eine fast sicher konvergente Folge von reellen Zufallsvariablen auch im Mittel gegen denselben Limes konvergiert.

Da P ein Wahrscheinlichkeitsmaß ist, gilt für alle $p, r \in [0, \infty)$ mit $p \leq r$ die Inklusion $L^r \subseteq L^p$; vgl. Beispiel 8.4.9(2). Darüber hinaus gilt das folgende Ergebnis:

14.3.3 Satz. *Sei $p, r \in [1, \infty)$ mit $p \leq r$. Dann konvergiert jede Folge von Zufallsvariablen in L^r, die im r–ten Mittel gegen eine Zufallsvariable $X \in L^r$ konvergiert, auch im p–ten Mittel gegen X.*

Beweis. Sei $\{X_n\}_{n \in \mathbb{N}} \subseteq L^r$ eine Folge von Zufallsvariablen, die im r–ten Mittel gegen eine Zufallsvariable $X \in L^r$ konvergiert. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $X = 0$ gilt.
Wegen $p \leq r$ ist die Funktion $h : \mathbb{R}_+ \to \mathbb{R}$ mit $h(x) := x^{p/r}$ konkav, und aus der Ungleichung von Jensen folgt nun

$$E[|X_n|^p] = E[(|X_n|^r)^{p/r}] \leq (E[|X_n|^r])^{p/r} = ((E[|X_n|^r])^{1/r})^p$$

Daher gilt

$$\|\, |X_n| \,\|_p = (E[|X_n|^p])^{1/p} \leq (E[|X_n|^r])^{1/r} = \|\, |X_n| \,\|_r$$

Die Behauptung folgt. □

15

Gesetze der Großen Zahlen

Für eine Folge von Zufallsvariablen $\{X_k\}_{k\in\mathbb{N}}$ in L^1 mit $E[X_k] = \mu$ für alle $k \in \mathbb{N}$ gilt aufgrund der Linearität des Erwartungswertes

$$E\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] = \mu$$

Wir untersuchen in diesem Kapitel die Frage, unter welchen Bedingungen und für welchen Konvergenzbegriff für eine derartige Folge

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \mu$$

gilt.

Wir geben zunächst hinreichende Bedingungen dafür an, dass für eine solche Folge die Folge $\{n^{-1}\sum_{k=1}^{n} X_k\}$ stochastisch (Abschnitt 15.1) oder fast sicher (Abschnitt 15.2) gegen μ konvergiert; die entsprechenden Konvergenzsätze werden als schwache bzw. starke Gesetze der Großen Zahlen bezeichnet. Die Gesetze der Großen Zahlen und der aus ihnen abgeleitete Satz von Glivenko/ Cantelli (Abschnitt 15.3) bilden eine wesentliche Grundlage der Statistik. Als eine weitere Anwendung des starken Gesetzes der Großen Zahlen behandeln wir Irrfahrten (Abschnitt 15.4).

15.1 Schwache Gesetze der Großen Zahlen

Eine Folge von Zufallsvariablen $\{X_k\}_{k\in\mathbb{N}}$ in L^1 mit $E[X_k] = \mu$ für alle $k \in \mathbb{N}$ genügt dem *schwachen Gesetz der Großen Zahlen* wenn

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \mu$$

bezüglich der stochastischen Konvergenz gilt.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_15, © Springer-Verlag Berlin Heidelberg 2011

Der folgende Satz enthält eine allgemeine Version des schwachen Gesetzes der Großen Zahlen:

15.1.1 Satz (Schwaches Gesetz der Großen Zahlen). *Sei* $\{X_k\}_{k\in\mathbb{N}}$ *eine Folge in* L^2 *mit* $E[X_k] = \mu$ *für alle* $k \in \mathbb{N}$ *und*

$$\lim_{n\to\infty} \mathrm{var}\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] = 0$$

Dann gilt

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \mu$$

im quadratischen Mittel und stochastisch.

Beweis. Wegen

$$E\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] = \mu$$

gilt

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right\|_2^2 &= E\left[\left(\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right)^2\right] \\
&= \mathrm{var}\left[\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right] \\
&= \mathrm{var}\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right]
\end{aligned}$$

und aus der Voraussetzung folgt nun die Konvergenz im quadratischen Mittel. Aus Satz 14.3.1 ergibt sich dann die stochastische Konvergenz. □

Eine Folge von Zufallsvariablen $\{X_k\}_{k\in\mathbb{N}}$ in L^2 heißt (*paarweise*) *unkorreliert*, wenn für alle $i, k \in \mathbb{N}$ mit $i \neq k$

$$\mathrm{cov}\,[X_i, X_k] = 0$$

gilt. Für unkorrelierte Folgen von Zufallsvariablen in L^2, die alle denselben Erwartungswert und außerdem dieselbe Varianz besitzen, ergibt sich eine besonders einprägsame Version des schwachen Gesetzes der Großen Zahlen:

15.1.2 Folgerung (Schwaches Gesetz der Großen Zahlen). *Sei* $\{X_k\}_{k\in\mathbb{N}}$ *eine unkorrelierte Folge in* L^2 *mit* $E[X_k] = \mu$ *und* $\mathrm{var}\,[X_k] = \sigma^2$ *für alle* $k \in \mathbb{N}$. *Dann gilt*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \mu$$

im quadratischen Mittel und stochastisch.

Beweis. Nach Voraussetzung gilt

$$\mathrm{var}\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] = \frac{1}{n^2}\sum_{k=1}^{n}\mathrm{var}\,[X_k] = \frac{1}{n^2}\,n\sigma^2 = \frac{\sigma^2}{n}$$

Die Behauptung folgt nun aus Satz 15.1.1. □

Das folgende Beispiel zeigt, dass insbesondere die bei der wiederholten Durchführung eines Zufallsexperimentes auftretenden relativen Häufigkeiten für das Eintreten eines bestimmten Ereignisses gegen die Wahrscheinlichkeit des Ereignisses konvergieren; die spezielle Form des Zufallsexperimentes ist dabei unwesentlich.

15.1.3 Beispiel (Wurf einer Münze). Wir betrachten den wiederholten Wurf einer Münze. Wir nehmen an, dass

- zwischen den verschiedenen Würfen der Münze keine gegenseitige Beeinflussung besteht,
- bei jedem Wurf nur *Kopf* oder *Zahl* auftreten kann, und
- die Chance für das Auftreten von *Kopf* beim einmaligen Wurf gleich einer reellen Zahl $\vartheta \in (0,1)$ ist.

Wir interessieren uns für die Konvergenz des Anteils derjenigen Würfe, bei denen *Kopf* auftritt, wenn die Anzahl der Würfe gegen Unendlich strebt.
Wir wählen folgende Modelle:

- Als Modell für den i–ten Wurf wählen wir den diskreten Wahrscheinlichkeitsraum $(\Omega_i, \mathcal{F}_i, P_i)$ mit $\Omega_i := \{K, Z\}$ sowie $\mathcal{F}_i = 2^{\Omega_i}$ und dem durch $P_i[\{K\}] := \vartheta$ festgelegten Wahrscheinlichkeitsmaß.
- Als Modell für den wiederholten Wurf wählen wir den Wahrscheinlichkeitsraum

$$(\Omega, \mathcal{F}, P) := \bigotimes_{i\in\mathbb{N}} (\Omega_i, \mathcal{F}_i, P_i)$$

Für alle $i \in \mathbb{N}$ sei $X_i : \Omega \to \mathbb{R}$ gegeben durch

$$X_i(\omega) := \begin{cases} 1 & \text{falls } \omega_i = K \\ 0 & \text{falls } \omega_i = Z \end{cases}$$

und sei $h_i : \Omega_i \to \mathbb{R}$ gegeben durch

$$h_i(\omega_i) := \begin{cases} 1 & \text{falls } \omega_i = K \\ 0 & \text{falls } \omega_i = Z \end{cases}$$

Dann gilt

$$X_i = h \circ \pi_i$$

Nach Satz 11.4.1 ist die Folge $\{\pi_i\}_{i\in\mathbb{N}}$ unabhängig und aus Satz 11.3.3 folgt nun, dass auch die Folge $\{X_i\}_{i\in\mathbb{N}}$ unabhängig ist. Für alle $i \in \mathbb{N}$ gilt außerdem $P_{X_i} = \mathbf{B}(\vartheta)$.

Für $n \in \mathbb{N}$ gibt die Zufallsvariable

$$\frac{1}{n} \sum_{i=1}^{n} X_i$$

für n Würfe den Anteil derjenigen Würfe an, bei denen *Kopf* auftritt; dieser Anteil wird auch als *relative Häufigkeit* für das Eintreten des Ereignisses $\{K\}$ bei einem einzelnen Wurf bezeichnet. Dann gilt

$$E\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \vartheta$$

und

$$\operatorname{var}\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{\vartheta(1-\vartheta)}{n}$$

Aus dem schwachen Gesetz der Großen Zahlen folgt nun

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \vartheta$$

im quadratischen Mittel und stochastisch, und damit die Konvergenz der relativen Häufigkeiten für das Auftreten von *Kopf* gegen die Wahrscheinlichkeit dafür, dass bei einem einzelnen Wurf das Ereignis $\{K\}$ eintritt.

Aufgaben

15.1.A **Schwaches Gesetz der Großen Zahlen:** Sei $\{X_k\}_{k\in\mathbb{N}}$ eine Folge in L^2 mit

$$\lim_{n\to\infty} \operatorname{var}\left[\frac{1}{n} \sum_{k=1}^{n} X_k\right] = 0$$

Dann gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Bigl(X_k - E[X_k]\Bigr) = 0$$

im quadratischen Mittel und stochastisch.

15.1.B **Schwaches Gesetz der Großen Zahlen:** Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unkorrelierte Folge in L^2 mit $E[X_k] = \mu$ für alle $k \in \mathbb{N}$ und

$$\sup_{k\in\mathbb{N}} \operatorname{var}[X_k] < \infty$$

Dann gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu$$

im quadratischen Mittel und stochastisch. Geben Sie eine geometrische Interpretation der Konvergenz im quadratischen Mittel.

15.2 Starke Gesetze der Großen Zahlen

Eine Folge von Zufallsvariablen $\{X_k\}_{k\in\mathbb{N}}$ in L^1 mit $E[X_k] = \mu$ für alle $k \in \mathbb{N}$ genügt dem *starken Gesetz der Großen Zahlen*, wenn

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu$$

fast sicher gilt.

Das im letzten Abschnitt bewiesene schwache Gesetz der Großen Zahlen gibt für eine solche Folge eine Bedingung an, unter der die Folge $\{n^{-1} \sum_{k=1}^{n} X_k\}_{n\in\mathbb{N}}$ stochastisch gegen μ konvergiert. Nach dem Teilfolgenprinzip konvergiert eine Teilfolge dieser Folge auch fast sicher gegen μ. Wir geben nun Bedingungen an, unter denen sogar die gesamte Folge fast sicher gegen μ konvergiert und damit dem starken Gesetz der Großen Zahlen genügt.

Als erstes zeigen wir, dass für eine unabhängige Folge von reellen Zufallsvariablen $\{X_k\}_{k\in\mathbb{N}}$ die Folge $\{n^{-1} \sum_{k=1}^{n} X_k\}_{n\in\mathbb{N}}$ entweder fast sicher konvergent oder fast sicher divergent ist und dass darüber hinaus im Fall der fast sicheren Konvergenz der Limes eine konstante Zufallsvariable ist:

15.2.1 Satz (Null–Eins–Gesetz). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge von reellen Zufallsvariablen und sei*

$$A := \left\{ \liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \right\}$$

Dann gilt $P[A] \in \{0,1\}$ und im Fall $P[A] = 1$ gibt es ein $c \in [-\infty, \infty]$ mit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = c$$

fast sicher.

Beweis. Sei $\mathcal{E}_\infty$ die terminale σ–Algebra der Folge $\{\sigma(X_k)\}_{k\in\mathbb{N}}$. Für alle $m \in \mathbb{N}$ und alle $n \in \mathbb{N}(m)$ gilt

$$\frac{1}{n} \sum_{k=1}^{n} X_k = \frac{1}{n} \sum_{k=1}^{m-1} X_k + \frac{1}{n} \sum_{k=m}^{n} X_k$$

und wegen $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{m-1} X_k = 0$ folgt daraus

$$A = \left\{ \liminf_{n\to\infty} \frac{1}{n} \sum_{k=m}^{n} X_k = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=m}^{n} X_k \right\}$$

und damit $A \in \sigma(\bigcup_{k=m}^{\infty} \sigma(X_k))$. Daher gilt $A \in \bigcap_{m\in\mathbb{N}} \sigma(\bigcup_{k=m}^{\infty} \sigma(X_k)) = \mathcal{E}_\infty$ und aus dem Null–Eins–Gesetz von Kolmogorov folgt nun $P[A] \in \{0,1\}$.
Im Fall $P[A] = 1$ ist die Folge $\{n^{-1}\sum_{k=1}^{n} X_k\}_{n\in\mathbb{N}}$ fast sicher konvergent und für alle $m \in \mathbb{N}$ gilt

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \lim_{n\to\infty} \frac{1}{n}\sum_{k=m}^{n} X_k$$

fast sicher. Daher ist die Zufallsvariable $\lim_{n\to\infty} n^{-1}\sum_{k=1}^{n} X_k$ $\mathcal{E}_\infty$–messbar und damit fast sicher konstant; es gibt also eine Konstante $c \in [-\infty, \infty]$ mit

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = c$$

fast sicher. □

Es gilt nun, in der Situation des letzten Satzes zusätzliche Bedingungen zu finden, unter denen der Fall der fast sicheren Konvergenz eintritt, und unter diesen Bedingungen auch den fast sicher konstanten Limes zu bestimmen. Wie im Fall des schwachen Gesetzes der Großen Zahlen in der Version von Folgerung 15.1.2 erweist sich auch hier eine Bedingung an die Varianzen als hinreichend.

15.2.2 Lemma (Ungleichung von Kolmogorov). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge in L^2. Dann gilt für alle $\varepsilon \in (0,\infty)$ und $m \in \mathbb{N}$*

$$P\left[\left\{\sup_{n\in\mathbb{N}(m)} \left|\sum_{k=m}^{n}\Big(X_k - E[X_k]\Big)\right| > \varepsilon\right\}\right] \le \frac{1}{\varepsilon^2}\sum_{k=m}^{\infty} \mathrm{var}\,[X_k]$$

Beweis. Es genügt, die Ungleichung für $m = 1$ zu beweisen. Für $n \in \mathbb{N}$ sei

$$S_n := \sum_{k=1}^{n}\Big(X_k - E[X_k]\Big)$$

und

$$A_n := \{|S_n| > \varepsilon\} \cap \bigcap_{k=1}^{n-1} \{|S_k| \le \varepsilon\}$$

Dann gilt $\{\sup_{n\in\mathbb{N}} |S_n| > \varepsilon\} = \sum_{n=1}^{\infty} A_n$ und damit

$$P\left[\left\{\sup_{n\in\mathbb{N}} |S_n| > \varepsilon\right\}\right] = \sum_{n=1}^{\infty} P[A_n]$$

Zur Abschätzung der rechten Seite dieser Gleichung betrachten wir zunächst $r \in \mathbb{N}$ und $n \in \{1, \ldots, r\}$. Da die Familie $\{\sigma(X_k)\}_{k\in\mathbb{N}}$ unabhängig ist, ist nach dem Blocklemma auch $\{\sigma(X_1, \ldots, X_n), \sigma(X_{n+1}, \ldots, X_r)\}$ unabhängig. Daher ist $\{S_n \chi_{A_n}, S_r - S_n\}$ unabhängig, und daraus folgt mit $E[S_r] = 0 = E[S_n]$

$$E[(S_r - S_n)\, S_n \chi_{A_n}] = E[S_r - S_n]\, E[S_n \chi_{A_n}] = 0$$

und damit

$$\begin{aligned}
\varepsilon^2 P[A_n] &= E[\varepsilon^2 \chi_{A_n}] \\
&\leq E[S_n^2 \chi_{A_n}] \\
&\leq E[(S_r - S_n)^2 \chi_{A_n}] + E[2(S_r - S_n) S_n \chi_{A_n}] + E[S_n^2 \chi_{A_n}] \\
&= E[S_r^2 \chi_{A_n}]
\end{aligned}$$

Durch Summation ergibt sich nun aus der Definition der Ereignisse A_n und mit $E[S_r] = 0$ und der Unabhängigkeit der Folge $\{X_k\}_{k\in\mathbb{N}}$

$$\begin{aligned}
\varepsilon^2 \sum_{n=1}^{r} P[A_n] &\leq \sum_{n=1}^{r} E[S_r^2 \chi_{A_n}] \\
&\leq E[S_r^2] \\
&= \mathrm{var}\,[S_r] \\
&= \mathrm{var}\left[\sum_{n=1}^{r} X_n\right] \\
&= \sum_{n=1}^{r} \mathrm{var}\,[X_n] \\
&\leq \sum_{n=1}^{\infty} \mathrm{var}\,[X_n]
\end{aligned}$$

Daher gilt

$$\sum_{n=1}^{\infty} P[A_n] \leq \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \mathrm{var}\,[X_n]$$

und damit

$$P\left[\left\{\sup_{n\in\mathbb{N}} |S_n| > \varepsilon\right\}\right] = \sum_{n=1}^{\infty} P[A_n] \leq \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \mathrm{var}\,[X_n]$$

Damit ist das Lemma bewiesen. □

Aus der Ungleichung von Kolmogorov ergibt sich eine erste Konvergenzaussage:

15.2.3 Lemma. *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge von Zufallsvariablen in L^2 mit*

$$\sum_{k=1}^{\infty} \mathrm{var}\,[X_k] < \infty$$

Dann konvergiert die Reihe

$$\sum_{k=1}^{\infty} \Big(X_k - E[X_k] \Big)$$

fast sicher und in L^2 gegen eine reelle Zufallsvariable.

Beweis. Für $n \in \mathbb{N}$ sei wieder

$$S_n := \sum_{k=1}^{n} \Big(X_k - E[X_k] \Big)$$

Nach der Ungleichung von Kolmogorov gilt für alle $\varepsilon \in (0, \infty)$ und $m \in \mathbb{N}$

$$\begin{aligned} P\left[\left\{ \sup_{n\in\mathbb{N}(m)} |S_n - S_m| > \varepsilon \right\}\right] &= P\left[\left\{ \sup_{n\in\mathbb{N}(m+1)} |S_n - S_m| > \varepsilon \right\}\right] \\ &= P\left[\left\{ \sup_{n\in\mathbb{N}(m+1)} \left| \sum_{k=m+1}^{n} \Big(X_k - E[X_k] \Big) \right| > \varepsilon \right\}\right] \\ &\leq \frac{1}{\varepsilon^2} \sum_{k=m+1}^{\infty} \mathrm{var}\,[X_k] \end{aligned}$$

und aus der Voraussetzung folgt nun für alle $\varepsilon \in (0, \infty)$

$$\lim_{m\to\infty} P\left[\left\{ \sup_{n\in\mathbb{N}(m)} |S_n - S_m| > \varepsilon \right\}\right] = 0$$

Daher folgt aus Lemma 14.1.1, dass die Folge $\{S_n\}_{n\in\mathbb{N}}$ fast sicher gegen eine reelle Zufallsvariable konvergiert.
Des weiteren gilt für alle $m \in \mathbb{N}$ und $n \in \mathbb{N}(m)$

$$\|S_n - S_m\|_2^2 = \left\| \sum_{k=m+1}^{n} \Big(X_k - E[X_k] \Big) \right\|_2^2 = \mathrm{var}\left[\sum_{k=m+1}^{n} X_k \right] = \sum_{k=m+1}^{n} \mathrm{var}\,[X_k]$$

und aus der Voraussetzung folgt nun, dass die Folge $\{S_n\}_{n\in\mathbb{N}}$ eine Cauchy–Folge in L^2 und damit in L^2 konvergent ist.
Da jede im quadratischen Mittel konvergente Folge eine fast sicher konvergente Teilfolge besitzt, stimmen die Limites bezüglich der fast sicheren Konvergenz und der Konvergenz im quadratischen Mittel überein. □

Wir benötigen nun noch eine Eigenschaft unendlicher Reihen:

15.2.4 Lemma (Kronecker). *Sei $\{a_k\}_{k\in\mathbb{N}}$ eine Folge reeller Zahlen mit*

$$\sum_{k=1}^{\infty} \frac{a_k}{k} = a$$

für ein $a \in \mathbb{R}$. Dann gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} a_k = 0$$

Beweis. Für $n \in \mathbb{N}$ sei

$$b_n := \sum_{k=1}^{n} \frac{a_k}{k}$$

(und damit $b_0 = 0$). Dann gilt für alle $m \in \mathbb{N}$ und $n \in \mathbb{N}(m)$

$$\begin{aligned}
\frac{1}{n}\sum_{k=1}^{n} a_k &= \frac{1}{n}\sum_{k=1}^{n} k\,(b_k - b_{k-1}) \\
&= \frac{1}{n}\left(\sum_{k=1}^{n} kb_k - \sum_{k=1}^{n}(k-1)b_{k-1} - \sum_{k=1}^{n} b_{k-1}\right) \\
&= \frac{1}{n}\sum_{k=1}^{n}(b_n - b_{k-1}) \\
&= \frac{1}{n}\sum_{k=1}^{m}(b_m - b_{k-1}) + \frac{1}{n}\left(\sum_{k=1}^{m}(b_n - b_m) + \sum_{k=m+1}^{n}(b_n - b_{k-1})\right)
\end{aligned}$$

Nach Voraussetzung ist die Folge $\{b_n\}_{n\in\mathbb{N}}$ konvergent mit $\lim_{n\to\infty} b_n = a$ und insbesondere eine Cauchy–Folge. Daraus folgt die Behauptung. □

Wir können nun eine allgemeine Version des ersten starken Gesetzes der Großen Zahlen beweisen:

15.2.5 Satz (1. Gesetz der Großen Zahlen; Kolmogorov). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge in L^2 mit $E[X_k] = \mu$ für alle $k \in \mathbb{N}$ und*

$$\sum_{k=1}^{\infty} \frac{\mathrm{var}\,[X_k]}{k^2} < \infty$$

Dann gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu$$

fast sicher.

Beweis. Nach Voraussetzung gilt

$$\sum_{k=1}^{\infty} \operatorname{var}\left[\frac{X_k}{k}\right] < \infty$$

und aus Lemma 15.2.3 folgt, dass die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k}\Big(X_k - E[X_k]\Big) = \sum_{k=1}^{\infty} \left(\frac{X_k}{k} - E\left[\frac{X_k}{k}\right]\right)$$

fast sicher gegen eine reelle Zufallsvariable konvergiert. Aus Kroneckers Lemma folgt nun

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Big(X_k - E[X_k]\Big) = 0$$

fast sicher. □

Für unabhängige Folgen von Zufallsvariablen in L^2, die alle denselben Erwartungswert und außerdem dieselbe Varianz besitzen, ergibt sich eine besonders einprägsame Version des ersten starken Gesetzes der Großen Zahlen:

15.2.6 Folgerung (1. Gesetz der Großen Zahlen; Kolmogorov). *Sei* $\{X_k\}_{k\in\mathbb{N}}$ *eine unabhängige Folge in* L^2 *mit* $E[X_k] = \mu$ *und* $\operatorname{var}[X_k] = \sigma^2$ *für alle* $k \in \mathbb{N}$. *Dann gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu$$

fast sicher und in L^2.

Beweis. Die fast sichere Konvergenz ergibt sich aus Satz 15.2.5 und die Konvergenz in L^2 ergibt sich aus Folgerung 15.1.2, da jede unabhängige Folge in L^2 unkorreliert ist. □

Der Vergleich von Folgerung 15.2.6 mit Folgerung 15.1.2 zeigt, dass eine Folge $\{X_k\}_{k\in\mathbb{N}}$ in L^2 mit $E[X_k] = \mu$ und $\operatorname{var}[X_k] = \sigma^2$ für alle $k \in \mathbb{N}$
- dem schwachen Gesetz der Großen Zahlen genügt, wenn sie unkorreliert ist, und
- dem starken Gesetz der Großen Zahlen genügt, wenn sie unabhängig ist.

Dies entspricht genau der Tatsache, dass
- jede unabhängige Folge unkorreliert ist und
- jede fast sicher konvergente Folge stochastisch konvergent ist,

während die umgekehrten Implikationen im allgemeinen falsch sind.

Nach Folgerung 15.2.6 konvergiert insbesondere die in Beispiel 15.1.3 betrachtete Folge der relativen Häufigkeiten eines Ereignisses nicht nur stochastisch, sondern sogar fast sicher gegen die Wahrscheinlichkeit des Ereignisses.

Im 1. Gesetz der Großen Zahlen wird in der Voraussetzung die quadratische Integrierbarkeit der Zufallsvariablen gefordert, obwohl in der Formulierung der Konvergenzaussage nur die Integrierbarkeit benötigt wird. Damit stellt sich die Frage, ob sich ein starkes Gesetz der Großen Zahlen auch unter Verzicht auf die Varianzbedingung beweisen lässt.

Eine Familie von Zufallsvariablen $\{X_i\}_{i\in I}$ heißt *identisch verteilt*, wenn alle Zufallsvariablen dieselbe Verteilung besitzen; in diesem Fall bezeichnen wir mit X eine beliebige Zufallsvariable mit $P_X = P_{X_i}$ für alle $i \in I$ und nennen X eine *typische Zufallsvariable* der Familie $\{X_i\}_{i\in I}$.

Wir zeigen nun, dass für eine Folge von Zufallsvariablen, die nicht nur unabhängig, sondern auch identisch verteilt sind, die im 1. Gesetz der Großen Zahlen auftretende Varianzbedingung entbehrlich ist; in diesem Fall stellt sich außerdem heraus, dass die Zufallsvariablen genau dann integrierbar sind, wenn die Folge dem starken Gesetz der Großen Zahlen genügt.

15.2.7 Satz (2. Gesetz der Großen Zahlen; Kolmogorov). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge in L^1 mit $E[X] = \mu$. Dann gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu$$

fast sicher.

Beweis. Für alle $k \in \mathbb{N}$ sei $Z_k : \Omega \to \mathbb{R}$ gegeben durch

$$Z_k(\omega) := X_k(\omega)\,\chi_{(-k,k)}(X_k(\omega))$$

und sei

$$J_k := (-k, -k+1] \cup [k-1, k)$$

Dann ist $\{Z_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge in L^2. Wir zeigen:

(1) Es gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Bigl(X_k - Z_k\Bigr) = 0$$

fast sicher.

(2) Es gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Bigl(Z_k - E[Z_k]\Bigr) = 0$$

fast sicher.

(3) Es gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Bigl(E[Z_k] - \mu\Bigr) = 0$$

Dann gilt offenbar

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Bigl(X_k - \mu\Bigr) = 0$$

fast sicher, und daraus folgt die Behauptung.
Wir beweisen nun die genannten Teilergebnisse:
(1) Nach Lemma 12.3.1 gilt

$$\begin{aligned}\sum_{k=1}^{\infty} P[\{X_k \neq Z_k\}] &= \sum_{k=1}^{\infty} P[\{|X_k| \geq k\}] \\ &= \sum_{k=1}^{\infty} P[\{|X| \geq k\}] \\ &\leq E[|X|]\end{aligned}$$

Aus der Endlichkeit des Erwartungswertes von $|X|$ und dem 1. Lemma von Borel/Cantelli folgt nun $P[\limsup_{n\to\infty}\{X_k \neq Z_k\}] = 0$ und damit

$$P\Bigl[\liminf_{n\to\infty}\{X_k = Z_k\}\Bigr] = 1$$

Daher gibt es eine Nullmenge $N \in \mathcal{F}$ derart, dass es zu jedem $\omega \in \Omega \setminus N$ ein $k_\omega \in \mathbb{N}$ gibt derart, dass $X_k(\omega) = Z_k(\omega)$ für alle $k \in \mathbb{N}(k_\omega)$ gilt. Daraus folgt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Bigl(X_k - Z_k\Bigr) = 0$$

fast sicher. Damit ist (1) gezeigt.
(2) Für alle $k \in \mathbb{N}$ gilt

$$\begin{aligned}\mathrm{var}\,[Z_k] &\leq E[Z_k^2] \\ &= E[X_k^2\,\chi_{(-k,k)}(X_k)] \\ &= \int_{(-k,k)} x^2\, dP_X(x) \\ &= \sum_{j=1}^{k} \int_{J_j} x^2\, dP_X(x) \\ &\leq \sum_{j=1}^{k} j \int_{J_j} |x|\, dP_X(x)\end{aligned}$$

Ferner gilt

$$\sum_{k=j}^{\infty} \frac{1}{k^2} \leq \frac{1}{j^2} + \sum_{k=j}^{\infty} \int_{(k,k+1]} \frac{1}{x^2}\, d\boldsymbol{\lambda}(x) = \frac{1}{j^2} + \int_{(j,\infty)} \frac{1}{x^2}\, d\boldsymbol{\lambda}(x) = \frac{1}{j^2} + \frac{1}{j} \leq \frac{2}{j}$$

und damit

$$\sum_{k=j}^{\infty} \frac{j}{k^2} \leq 2$$

Daher gilt

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{\mathrm{var}\,[Z_k]}{k^2} &\leq \sum_{k=1}^{\infty} \frac{1}{k^2} \sum_{j=1}^{k} j \int_{J_j} |x| \, dP_X(x) \\
&= \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} \frac{j}{k^2} \int_{J_j} |x| \, dP_X(x) \\
&\leq 2 \sum_{j=1}^{\infty} \int_{J_j} |x| \, dP_X(x) \\
&= 2 \int_{\mathbb{R}} |x| \, dP_X(x) \\
&= 2\, E[|X|]
\end{aligned}$$

Aus der Endlichkeit des Erwartungswertes von $|X|$ und dem 1. Gesetz der Großen Zahlen folgt nun

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Big(Z_k - E[Z_k] \Big) = 0$$

fast sicher. Damit ist (2) gezeigt.
(3) Für alle $k \in \mathbb{N}$ gilt

$$\begin{aligned}
\Big| E[Z_k] - \mu \Big| &= \Big| E[Z_k] - E[X_k] \Big| \\
&\leq E[|Z_k - X_k|] \\
&= \sum_{j=k+1}^{\infty} \int_{J_j} |x| \, dP_{X_k}(x) \\
&= \sum_{j=k+1}^{\infty} \int_{J_j} |x| \, dP_X(x)
\end{aligned}$$

Aus der Endlichkeit des Erwartungswertes von $|X|$ folgt nun

$$\lim_{n\to\infty} \Big| E[Z_k] - \mu \Big| = 0$$

und damit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Big(E[Z_k] - \mu \Big) = 0$$

Damit ist auch (3) gezeigt. □

Für eine unabhängig und identisch verteilte Folge von Zufallsvariablen in L^2 ist die Varianzbedingung aus dem 1. Gesetz der Großen Zahlen offenbar erfüllt. Der Vorteil des 2. Gesetzes der Großen Zahlen besteht darin, dass es sogar für jede unabhängig und identisch verteilte Folge von Zufallsvariablen in L^1 gilt.

Für eine unabhängig und identisch verteilte Folge $\{X_k\}_{k\in\mathbb{N}}$ von Zufallsvariablen bezeichnen wir die Zufallsvariablen

$$\frac{1}{n}\sum_{k=1}^{n} X_k$$

auch als *Stichprobenmittel* zum *Stichprobenumfang* $n \in \mathbb{N}$.

Als nächstes wenden wir uns der Frage zu, ob sich für eine unabhängig und identisch verteilte Folge $\{X_k\}_{k\in\mathbb{N}}$ von Zufallsvariablen die fast sichere Konvergenz der Folge der Stichprobenmittel gegen eine reelle Zahl auch dann einstellen kann, wenn $E[|X|] = \infty$ gilt:

15.2.8 Satz. *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge von Zufallsvariablen. Dann sind äquivalent*:
(a) *Es gilt $X \in L^1$.*
(b) *Die Folge der Stichprobenmittel konvergiert gegen eine reelle Zahl.*
(c) *Die Folge der Stichprobenmittel konvergiert gegen eine reelle Zufallsvariable.*

Beweis. Nach dem 2. Gesetz der Großen Zahlen folgt (b) aus (a), und nach dem Null–Eins–Gesetz 15.2.1 sind (b) und (c) äquivalent. Wir nehmen nun an, dass (b) gilt. Wegen

$$\frac{1}{n}X_n = \frac{1}{n}\sum_{k=1}^{n} X_k - \frac{n-1}{n}\,\frac{1}{n-1}\sum_{k=1}^{n-1} X_k$$

gilt dann

$$\lim_{n\to\infty}\frac{1}{n}X_n = 0$$

fast sicher, und damit

$$P\left[\limsup_{n\to\infty}\{|X_n| \geq n\}\right] = P\left[\limsup_{n\to\infty}\left\{\left|\frac{1}{n}\,X_n\right| \geq 1\right\}\right] = 0$$

Da die Folge $\{X_n\}_{n\in\mathbb{N}}$ unabhängig und identisch verteilt ist, folgt aus dem Null–Eins–Gesetz von Borel

$$\sum_{n=1}^{\infty} P[\{|X| \geq n\}] = \sum_{n=1}^{\infty} P[\{|X_n| \geq n\}] < \infty$$

Aus Lemma 12.3.1 folgt nun $E[|X|] < \infty$. Damit ist gezeigt, dass (a) aus (b) folgt. □

Wir untersuchen nun die Konvergenz der Folge der Stichprobenmittel einer unabhängig und identisch verteilten Folge von quasiintegrierbaren Zufallsvariablen:

15.2.9 Satz. *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge von Zufallsvariablen.*
(1) *Im Fall $E[X^-] < \infty = E[X^+]$ gilt $\lim_{n\to\infty} n^{-1}\sum_{k=1}^n X_k = +\infty$.*
(2) *Im Fall $E[X^+] < \infty = E[X^-]$ gilt $\lim_{n\to\infty} n^{-1}\sum_{k=1}^n X_k = -\infty$.*

Beweis. Sei $E[X^-] < \infty = E[X^+]$.
Mit $\{X_k\}_{k\in\mathbb{N}}$ ist auch die Folge $\{X_k^-\}_{k\in\mathbb{N}}$ unabhängig und identisch verteilt, und aus dem 2. Gesetz der Großen Zahlen folgt wegen $E[X^-] < \infty$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k^- = E[X^-]$$

Des weiteren ist für $m \in \mathbb{N}$ die Folge $\{(X_k^+ \wedge m)\}_{k\in\mathbb{N}}$ unabhängig und identisch verteilt, und aus dem 2. Gesetz der Großen Zahlen folgt

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k^+ \wedge m = E[X^+ \wedge m]$$

und damit

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k^+ \geq \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k^+ \wedge m = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k^+ \wedge m = E[X^+ \wedge m]$$

Aus dem Satz über die monotone Konvergenz folgt nun

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k^+ \geq \sup_{m\in\mathbb{N}} E[X^+ \wedge m] = E[X^+]$$

und wegen $E[X^+] = \infty$ ergibt sich daraus

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k^+ = E[X^+]$$

Wegen $X_k = X_k^+ - X_k^-$ und $E[X^-] < \infty = E[X^+]$ ergibt sich nun

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k = \infty$$

Damit ist (1) gezeigt, und (2) ist dann klar. □

Unser letztes Ergebnis zeigt, dass für eine unabhängig und identisch verteilte Folge $\{X_k\}_{k\in\mathbb{N}}$ von Zufallsvariablen mit $E[|X|] = \infty$ die Folge der Stichprobenmittel fast sicher nicht gegen eine reelle Zahl konvergiert:

15.2.10 Satz. *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge von Zufallsvariablen mit $E[|X|] = \infty$. Dann gilt*

$$P\left[\left\{\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{k=1}^{n} X_k\right| = \infty\right\}\right] = 1$$

Beweis. Wegen

$$\frac{X_n}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k - \frac{n-1}{n}\,\frac{1}{n-1}\sum_{k=1}^{n-1} X_k$$

gilt

$$\limsup_{n\to\infty}\left|\frac{X_n}{n}\right| \leq 2\,\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{k=1}^{n} X_k\right|$$

und damit

$$P\left[\left\{\limsup_{n\to\infty}\left|\frac{X_n}{n}\right| = \infty\right\}\right] \leq P\left[\left\{\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{k=1}^{n} X_k\right| = \infty\right\}\right]$$

Wir zeigen nun, dass

$$P\left[\left\{\limsup_{n\to\infty}\left|\frac{X_n}{n}\right| = \infty\right\}\right] = 1$$

gilt.
Für alle $m \in \mathbb{N}$ gilt nach Voraussetzung $E[|X|/m] = \infty$ und aus Lemma 12.3.1 folgt

$$\sum_{n=1}^{\infty} P\left[\left\{\left|\frac{X_n}{n}\right| \geq m\right\}\right] = \sum_{n=1}^{\infty} P\left[\left\{\left|\frac{X}{n}\right| \geq m\right\}\right] = \sum_{n=1}^{\infty} P\left[\left\{\left|\frac{X}{m}\right| \geq n\right\}\right] = \infty$$

Da die Folge $\{X_n\}_{n\in\mathbb{N}}$ unabhängig ist, folgt aus dem Null–Eins–Gesetz von Borel

$$P\left[\limsup_{n\to\infty}\left\{\left|\frac{X_n}{n}\right| \geq m\right\}\right] = 1$$

und damit

$$P\left[\left\{\limsup_{n\to\infty}\left|\frac{X_n}{n}\right| \geq m\right\}\right] = 1$$

Da diese Gleichung für alle $m \in \mathbb{N}$ gilt, ergibt sich

$$P\left[\left\{\limsup_{n\to\infty}\left|\frac{X_n}{n}\right| = \infty\right\}\right] = 1$$

Damit ist der Satz bewiesen. □

Aufgaben

15.2.A **1. Gesetz der Großen Zahlen:** Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge in L^2 mit $E[X_k] = \mu$ für alle $k \in \mathbb{N}$ und $\sup_{k\in\mathbb{N}} \mathrm{var}\,[X_k] < \infty$. Dann gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu$$

fast sicher und in L^2.

15.2.B **Poisson–Verteilung:** Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge von reellen Zufallsvariablen mit $P_{X_k} = \mathbf{P}(\alpha_k)$ für alle $k \in \mathbb{N}$ und $\alpha := \sum_{k=1}^{\infty} \alpha_k < \infty$. Dann konvergiert die Reihe

$$\sum_{k=1}^{\infty} X_k$$

fast sicher und in L^2 gegen eine reelle Zufallsvariable X mit $E[X] = \alpha$.

15.3 Satz von Glivenko/Cantelli

Im gesamten Abschnitt sei $F : \mathbb{R} \to [0,1]$ eine Verteilungsfunktion und sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge von reellen Zufallsvariablen mit Verteilungsfunktion F.

Für $n \in \mathbb{N}$ heißt die Abbildung $F_n : \mathbb{R} \times \Omega \to [0,1]$ mit

$$F_n(x,\omega) := \frac{1}{n} \sum_{k=1}^{n} \chi_{(-\infty,x]}(X_k(\omega))$$

die *empirische Verteilungsfunktion* zu F und zum Stichprobenumfang n. Diese Bezeichnung wird durch das folgende Lemma gerechtfertigt:

15.3.1 Lemma. *Für alle* $n \in \mathbb{N}$ *und jedes* $\omega \in \Omega$ *ist die Funktion* $F_n(\,.\,,\omega) : \mathbb{R} \to [0,1]$ *eine Verteilungsfunktion und für alle* $x \in \mathbb{R}$ *gilt*

$$F_n(x-,\omega) = \frac{1}{n} \sum_{k=1}^{n} \chi_{(-\infty,x)}(X_k(\omega))$$

Beweis. Wir zeigen zunächst, dass $F_n(\,.\,,\omega)$ eine Verteilungsfunktion ist:
(i) Für alle $x, y \in \mathbb{R}$ mit $x \le y$ gilt $F_n(x,\omega) \le F_n(y,\omega)$.
(ii) Sei $x \in \mathbb{R}$. Für alle $\varepsilon \in (0,\infty)$ gilt

$$F_n(x+\varepsilon,\omega) = \frac{1}{n} \sum_{k=1}^{n} \chi_{(-\infty,x+\varepsilon]}(X_k(\omega))$$

und damit $\lim_{\varepsilon\to 0} F_n(x+\varepsilon,\omega) = F_n(x,\omega)$.
(iii) Es gilt $\lim_{x\to-\infty} F_n(x,\omega) = 0$ und $\lim_{x\to\infty} F_n(x,\omega) = 1$.

Daher ist $F_n(\,.\, ,\omega)$ eine Verteilungsfunktion. Außerdem gilt für alle $x \in \mathbb{R}$

$$\begin{aligned} F_n(x-,\omega) &= \lim_{\varepsilon\to 0} F_n(x-\varepsilon,\omega) \\ &= \lim_{\varepsilon\to 0} \frac{1}{n}\sum_{k=1}^{n} \chi_{(-\infty,x-\varepsilon]}(X_k(\omega)) \\ &= \frac{1}{n}\sum_{k=1}^{n} \chi_{(-\infty,x)}(X_k(\omega)) \end{aligned}$$

Damit ist das Lemma gezeigt. □

Aus der Definition der empirischen Verteilungsfunktion ist unmittelbar klar, dass für alle $n \in \mathbb{N}$ und $x \in \mathbb{R}$ die Abbildung $F_n(x,\,.\,) : \Omega \to [0,1]$ messbar ist. Wir untersuchen nun die Konvergenz der Folge $\{F_n(x,\,.\,)\}_{n\in\mathbb{N}}$.

15.3.2 Lemma. *Für jedes $x \in \mathbb{R}$ gilt $\lim_{n\to\infty} F_n(x,\omega) = F(x)$ fast sicher.*

Beweis. Mit $\{X_k\}_{k\in\mathbb{N}}$ ist auch die Folge $\{\chi_{(-\infty,x]} \circ X_k\}_{n\in\mathbb{N}}$ unabhängig und identisch verteilt. Außerdem gilt

$$\begin{aligned} E[(\chi_{(-\infty,x]} \circ X_k)] &= E[\chi_{\{X_k \le x\}}] \\ &= P[\{X_k \le x\}] \\ &= F(x) \end{aligned}$$

Aus dem starken Gesetz der Großen Zahlen folgt nun

$$\lim_{n\to\infty} F_n(x,\omega) = \frac{1}{n}\sum_{k=1}^{n}(\chi_{(-\infty,x]} \circ X_k)(\omega) = F(x)$$

Damit ist das Lemma bewiesen. □

Nach Lemma 15.3.2 gibt es zu jedem $x \in \mathbb{R}$ eine Nullmenge $N(x) \in \mathcal{F}$ derart, dass für alle $\omega \in \Omega \setminus N(x)$

$$\lim_{n\to\infty} \Big| F_n(x,\omega) - F(x) \Big| = 0$$

gilt. Das Bemerkenswerte an der Aussage des folgenden Satzes ist, dass man die Nullmengen $N(x)$ durch eine von $x \in \mathbb{R}$ unabhängige Nullmenge $N \in \mathcal{F}$ ersetzen kann und dass die Konvergenz für alle $\omega \in \Omega \setminus N$ gleichmäßig in $x \in \mathbb{R}$ ist:

15.3.3 Satz (Glivenko/Cantelli). *Es gilt*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \Big| F_n(x,\omega) - F(x) \Big| = 0$$

fast sicher.

Beweis. Sei $\varphi : (0,1) \to \mathbb{R}$ gegeben durch

$$\varphi(z) := \inf\Big\{ x \in \mathbb{R} \;\Big|\; z \le F(x) \Big\}$$

Dann gilt für alle $z \in (0,1)$

$$F(\varphi(z)-) \le z \le F(\varphi(z))$$

Für $m \in \mathbb{N}$ und $k \in \{1, \dots, m\}$ sei

$$x_{m,k} := \varphi\left(\frac{k}{m+1}\right)$$

Dann gilt für alle $k \in \{1, \dots, m\}$

$$F(x_{m,k}-) \le \frac{k}{m+1} \le F(x_{m,k})$$

und damit für alle $k \in \{1, \dots, m-1\}$

$$\begin{aligned} F(x_{m,1}-) &\le \frac{1}{m+1} \\ F(x_{m,k+1}-) - F(x_{m,k}) &\le \frac{1}{m+1} \\ 1 - F(x_{m,m}) &\le \frac{1}{m+1} \end{aligned}$$

Für $m, n \in \mathbb{N}$ und $\omega \in \Omega$ sei

$$\begin{aligned} &\Delta_{m,n}(\omega) \\ &:= \max_{1 \le k \le m} \max\Big\{ \Big|F_n(x_{m,k},\omega) - F(x_{m,k})\Big|, \Big|F_n(x_{m,k}-,\omega) - F(x_{m,k}-)\Big| \Big\} \end{aligned}$$

Mit Hilfe dieser Größen erhalten wir für alle $\omega \in \Omega$ eine von $x \in \mathbb{R}$ unabhängige Abschätzung der Differenz $|F_n(x,\omega) - F(x)|$:

– Für $x \in (-\infty, x_{m,1})$ gilt

$$\begin{aligned} F_n(x,\omega) &\le F_n(x_{m,1}-,\omega) \\ &\le F(x_{m,1}-) + \Delta_{m,n}(\omega) \\ &\le \frac{1}{m+1} + \Delta_{m,n}(\omega) \\ &\le F(x) + \frac{1}{m+1} + \Delta_{m,n}(\omega) \end{aligned}$$

und

$$\begin{aligned} F(x) &\le F(x_{m,1}-) \\ &\le \frac{1}{m+1} \\ &\le F_n(x,\omega) + \frac{1}{m+1} + \Delta_{m,n}(\omega) \end{aligned}$$

– Für $x \in [x_{m,k}, x_{m,k+1})$ mit $k \in \{1, \dots, m-1\}$ gilt

$$\begin{aligned} F_n(x,\omega) &\le F_n(x_{m,k+1}-,\omega) \\ &\le F(x_{m,k+1}-) + \Delta_{m,n}(\omega) \\ &\le F(x_{m,k}) + \frac{1}{m+1} + \Delta_{m,n}(\omega) \\ &\le F(x) + \frac{1}{m+1} + \Delta_{m,n}(\omega) \end{aligned}$$

und

$$\begin{aligned} F(x) &\le F(x_{m,k+1}-) \\ &\le F(x_{m,k}) + \frac{1}{m+1} \\ &\le F_n(x_{m,k},\omega) + \Delta_{m,n}(\omega) + \frac{1}{m+1} \\ &\le F_n(x,\omega) + \frac{1}{m+1} + \Delta_{m,n}(\omega) \end{aligned}$$

– Für $x \in [x_{m,m}, +\infty)$ gilt

$$\begin{aligned} F_n(x,\omega) &\le 1 \\ &\le F(x_{m,m}) + \frac{1}{m+1} \\ &\le F(x) + \frac{1}{m+1} \\ &\le F(x) + \frac{1}{m+1} + \Delta_{m,n}(\omega) \end{aligned}$$

und

$$\begin{aligned} F(x) &\le 1 \\ &\le F(x_{m,m}) + \frac{1}{m+1} \\ &\le F_n(x_{m,m},\omega) + \Delta_{m,n}(\omega) + \frac{1}{m+1} \\ &\le F_n(x,\omega) + \frac{1}{m+1} + \Delta_{m,n}(\omega) \end{aligned}$$

Daher gilt für alle $\omega \in \Omega$

$$\sup_{x\in\mathbb{R}} \Big| F_n(x,\omega) - F(x) \Big| \le \frac{1}{m+1} + \Delta_{m,n}(\omega)$$

Für festes $m \in \mathbb{N}$ gilt nach Lemma 15.3.2 $\lim_{n\to\infty} \Delta_{m,n}(\omega) = 0$ fast sicher und damit

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}} \Big|F_n(x,\omega) - F(x)\Big| \leq \frac{1}{m+1}$$

fast sicher. Da $m \in \mathbb{N}$ beliebig war und die Vereinigung einer Folge von Nullmengen wieder eine Nullmenge ist, erhalten wir nun

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}} \Big|F_n(x,\omega) - F(x)\Big| = 0$$

fast sicher, und damit

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \Big|F_n(x,\omega) - F(x)\Big| = 0$$

fast sicher. □

Aufgabe

15.3.A **Empirische Verteilung:** Sei $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung und sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge von Zufallsvariablen mit $P_X = Q$. Für $n \in \mathbb{N}$ heißt die Abbildung $Q_n : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ mit

$$Q_n(B,\omega) := \frac{1}{n}\sum_{k=1}^{n} \chi_B(X_k(\omega))$$

die *empirische Verteilung* zu Q und zum Stichprobenumfang n. Dann ist für alle $\omega \in \Omega$ die Abbildung $Q_n(\,.\,,\omega) : \mathcal{B}(\mathbb{R}) \to [0,1]$ eine Verteilung und für alle $B \in \mathcal{B}(\mathbb{R})$ ist die Abbildung $Q_n(B,\,.\,) : \Omega \to [0,1]$ messbar.

15.4 Irrfahrten

Im gesamten Abschnitt sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge von reellen Zufallsvariablen und sei X eine typische Zufallsvariable dieser Folge.

Wir untersuchen die Konvergenz der Folge $\{S_n\}_{n\in\mathbb{N}_0}$ mit

$$S_n := \sum_{k=1}^{n} X_k$$

und damit die Konvergenz der Reihe $\sum_{k=1}^{\infty} X_k$. Die Folge $\{S_n\}_{n\in\mathbb{N}_0}$ heißt *Irrfahrt* mit Sprüngen X_k.

Es erweist sich als zweckmäßig, das unendliche Produkt

$$(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R})^{\mathbb{N}}) = \bigotimes_{k=1}^{\infty} (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

zu betrachten und die Folge $\{X_k\}_{k\in\mathbb{N}}$ mit der Zufallsgröße $\mathbf{X} : \Omega \to \mathbb{R}^{\mathbb{N}}$ mit den Koordinaten $\pi_k \circ \mathbf{X} = X_k$ zu identifizieren.

Eine bijektive Abbildung $\tau : \mathbb{N} \to \mathbb{N}$ heißt *endliche Permutation*, wenn es ein $m \in \mathbb{N}$ gibt mit $\tau(n) = n$ für alle $n \in \mathbb{N}(m)$. Jede endliche Permutation τ erzeugt eine messbare Bijektion $\mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$, die wir der Einfachheit halber wieder mit τ bezeichnen. Dann ist für jede endliche Permutation τ auch die Abbildung $\tau \circ \mathbf{X} : \Omega \to \mathbb{R}^{\mathbb{N}}$ eine Zufallsgröße.

Eine messbare Funktion $h : \mathbb{R}^{\mathbb{N}} \to [-\infty, \infty]$ heißt $\mathbf{X}$*–permutierbar*, wenn für jede endliche Permutation τ

$$h \circ \tau \circ \mathbf{X} = h \circ \mathbf{X}$$

gilt, und eine Menge $A \in \mathcal{B}(\mathbb{R})^{\mathbb{N}}$ heißt $\mathbf{X}$*–permutierbar*, wenn ihre Indikatorfunktion χ_A $\mathbf{X}$–permutierbar ist.

15.4.1 Lemma. *Sei $\mathcal{C}_{\mathbf{X}}$ das Mengensystem aller $\mathbf{X}$–permutierbaren Mengen. Dann ist $\mathcal{C}_{\mathbf{X}}$ eine σ–Algebra und jede $\mathbf{X}$–permutierbare Funktion ist $\mathcal{C}_{\mathbf{X}}$–messbar.*

Beweis. Aus der Definition ist unmittelbar klar, dass $\mathcal{C}_{\mathbf{X}}$ eine σ–Algebra ist. Sei nun $h : \mathbb{R}^{\mathbb{N}} \to [-\infty, \infty]$ $\mathbf{X}$–permutierbar. Dann gilt für alle $c \in [-\infty, \infty]$ und für jede endliche Permutation $\tau : \mathbb{N} \to \mathbb{N}$

$$\chi_{\{h \le c\}} \circ \tau \circ \mathbf{X} = \chi_{\{h \circ \tau \circ \mathbf{X} \le c\}} = \chi_{\{h \circ \mathbf{X} \le c\}} = \chi_{\{h \le c\}} \circ \mathbf{X}$$

und damit $\{h \le c\} \in \mathcal{C}_{\mathbf{X}}$. Da $c \in [-\infty, \infty]$ beliebig war, ist h $\mathcal{C}_{\mathbf{X}}$–messbar. □

Aus diesem Lemma ergibt sich das Null–Eins–Gesetz von Hewitt/Savage:

15.4.2 Satz (Null–Eins–Gesetz; Hewitt/Savage). *Für jede $\mathbf{X}$–permutierbare Menge $A \in \mathcal{B}(\mathbb{R})^{\mathbb{N}}$ gilt $P_{\mathbf{X}}[A] \in \{0, 1\}$.*

Beweis. Für alle $k \in \mathbb{N}$ sei

$$\mathcal{G}_k := \sigma\Big(\{\pi_i\}_{i \in \{1,\dots,k\}}\Big)$$

und damit

$$\mathcal{G}_k = \sigma(\pi_{\{1,\dots,k\}})$$

Dann ist das Mengensystem

$$\mathcal{G} := \bigcup_{k \in \mathbb{N}} \mathcal{G}_k$$

eine Algebra mit $\sigma(\mathcal{G}) = \mathcal{B}(\mathbb{R})^{\mathbb{N}}$ und nach dem Approximationssatz 5.4.1 gibt es eine Folge $\{G_k\}_{k \in \mathbb{N}} \subseteq \mathcal{G}$ mit

$$\lim_{k \to \infty} P_{\mathbf{X}}[A \triangle G_k] = 0$$

Da die Folge $\{\mathcal{G}_k\}_{k \in \mathbb{N}}$ monoton wachsend ist, können wir annehmen, dass $G_k \in \mathcal{G}_k$ für alle $k \in \mathbb{N}$ gilt.

Für alle $k \in \mathbb{N}$ sei $\tau_k : \mathbb{N} \to \mathbb{N}$ gegeben durch

$$\tau_k(n) := \begin{cases} n+k & \text{falls } n \in \{1, \dots, k\} \\ n-k & \text{falls } n \in \{k+1, \dots, 2k\} \\ n & \text{sonst} \end{cases}$$

Nach Wahl von G_k gibt es eine Menge $B_k \in \mathcal{B}(\mathbb{R}^k)$ mit

$$G_k = \pi^{-1}_{\{1,\dots,k\}}(B_k)$$

und wir setzen

$$H_k := \pi^{-1}_{\{k+1,\dots,2k\}}(B_k)$$

Dann gilt $\pi_{\{1,\dots,k\}} \circ \tau_k \circ \mathbf{X} = \pi_{\{k+1,\dots,2k\}} \circ \mathbf{X}$ und damit

$$\begin{aligned}(\tau_k \circ \mathbf{X})^{-1}(G_k) &= \mathbf{X}^{-1}(\tau_k^{-1}(\pi^{-1}_{\{1,\dots,k\}}(B_k))) \\ &= \mathbf{X}^{-1}(\pi^{-1}_{\{k+1,\dots,2k\}}(B_k)) \\ &= \mathbf{X}^{-1}(H_k)\end{aligned}$$

Da A $\mathbf{X}$–permutierbar ist, gilt außerdem $\chi_A \circ \tau_k \circ \mathbf{X} = \chi_A \circ \mathbf{X}$ und damit

$$(\tau_k \circ \mathbf{X})^{-1}(A) = \mathbf{X}^{-1}(A)$$

Daraus folgt

$$(\tau_k \circ \mathbf{X})^{-1}(A \triangle G_k) = \mathbf{X}^{-1}(A \triangle H_k)$$

Da die Folge $\{X_k\}_{k \in \mathbb{N}}$ unabhängig und identisch verteilt ist, gilt $P_\mathbf{X} = P_{\tau_k \circ \mathbf{X}}$ und damit

$$\begin{aligned}P_\mathbf{X}[G_k] &= P_{\tau_k \circ \mathbf{X}}[G_k] \\ &= P[(\tau_k \circ \mathbf{X})^{-1}(G_k)] \\ &= P[\mathbf{X}^{-1}(H_k)] \\ &= P_\mathbf{X}[H_k]\end{aligned}$$

sowie

$$P_\mathbf{X}[A \triangle G_k] = P_\mathbf{X}[A \triangle H_k]$$

Wegen $A \triangle (G_k \cap H_k) \subseteq (A \triangle G_k) \cup (A \triangle H_k)$ folgt daraus

$$\begin{aligned}P_\mathbf{X}[A \triangle (G_k \cap H_k)] &\leq P_\mathbf{X}[A \triangle G_k] + P_\mathbf{X}[A \triangle H_k] \\ &= 2\, P_\mathbf{X}[A \triangle G_k]\end{aligned}$$

Nach Wahl der Folge $\{G_k\}_{k \in \mathbb{N}}$ gilt $\lim_{k \to \infty} P_\mathbf{X}[A \triangle G_k] = 0$. Wir erhalten daher

$$\begin{aligned}
\lim_{k\to\infty} P_{\mathbf{X}}[A\triangle(G_k\cap H_k)] &= 0\\
\lim_{k\to\infty} P_{\mathbf{X}}[A\triangle G_k] &= 0\\
\lim_{k\to\infty} P_{\mathbf{X}}[A\triangle H_k] &= 0
\end{aligned}$$

und damit wegen $|P_{\mathbf{X}}[A] - P_{\mathbf{X}}[C]| \le P_{\mathbf{X}}[A\triangle C]$

$$\begin{aligned}
\lim_{k\to\infty} P_{\mathbf{X}}[G_k\cap H_k] &= P_{\mathbf{X}}[A]\\
\lim_{k\to\infty} P_{\mathbf{X}}[G_k] &= P_{\mathbf{X}}[A]\\
\lim_{k\to\infty} P_{\mathbf{X}}[H_k] &= P_{\mathbf{X}}[A]
\end{aligned}$$

Aus der Unabhängigkeit der Folge $\{X_k\}_{k\in\mathbb{N}}$ bezüglich P folgt die Unabhängigkeit der Folge $\{\pi_k\}_{k\in\mathbb{N}}$ bezüglich $P_{\mathbf{X}}$. Daher ist für jedes $k \in \mathbb{N}$ das Paar $\{G_k, H_k\}$ unabhängig bezüglich $P_{\mathbf{X}}$ und wir erhalten

$$\begin{aligned}
P_{\mathbf{X}}[A] &= \lim_{k\to\infty} P_{\mathbf{X}}[G_k\cap H_k]\\
&= \lim_{k\to\infty} \left(P_{\mathbf{X}}[G_k]\, P_{\mathbf{X}}[H_k]\right)\\
&= \lim_{k\to\infty} P_{\mathbf{X}}[G_k]\cdot \lim_{k\to\infty} P_{\mathbf{X}}[H_k]\\
&= P_{\mathbf{X}}[A]\, P_{\mathbf{X}}[A]
\end{aligned}$$

und damit $P_{\mathbf{X}}[A] \in \{0,1\}$. □

Als unmittelbare Folgerung aus dem Null–Eins–Gesetz von Hewitt/Savage ergibt sich ein analoges Ergebnis für $\mathbf{X}$–permutierbare Funktionen:

15.4.3 Folgerung. *Sei $h : \mathbb{R}^{\mathbb{N}} \to [-\infty,\infty]$ $\mathbf{X}$–permutierbar. Dann ist h $P_{\mathbf{X}}$–fast sicher konstant und $h\circ\mathbf{X}$ ist P–fast sicher konstant.*

Aus dem Null–Eins–Gesetz von Hewitt/Savage ergibt sich des weiteren die folgende bemerkenswerte Alternative für die Konvergenz einer unendlichen Reihe von unabhängig und identisch verteilten Zufallsvariablen:

15.4.4 Satz. *Sei $P_X \neq \boldsymbol{\delta}_0$. Dann gilt eine der folgenden Alternativen*:
(1) $\lim_{n\to\infty} S_n = \infty$ *fast sicher.*
(2) $\liminf_{n\to\infty} S_n = -\infty$ *und* $\limsup_{n\to\infty} S_n = \infty$ *fast sicher.*
(3) $\lim_{n\to\infty} S_n = -\infty$ *fast sicher.*

Beweis. Sei $g : \mathbb{R}^{\mathbb{N}} \to [-\infty,\infty]$ gegeben durch

$$g(\mathbf{x}) := \limsup_{n\to\infty} \sum_{k=1}^n x_k$$

Für jede endliche Permutation $\tau : \mathbb{N} \to \mathbb{N}$ und alle $n \in \mathbb{N}$ hinreichend groß gilt

$$\sum_{k=1}^{n} x_k = \sum_{k=1}^{n} x_{\tau(k)}$$

Daher ist g $\mathbf{X}$–permutierbar und nach Folgerung 15.4.3 gibt es ein $c \in [-\infty, \infty]$ mit

$$P[\{g \circ \mathbf{X} = c\}] = 1$$

fast sicher.
Wir betrachten nun die Zufallsgröße $\mathbf{Y} : \Omega \to \mathbb{R}^{\mathbb{N}}$ mit $Y_n := X_{n+1}$ für alle $n \in \mathbb{N}$. Dann gilt $P_{\mathbf{Y}} = P_{\mathbf{X}}$ und damit

$$P[\{g \circ \mathbf{X} = c\}] = P_{\mathbf{X}}[\{g = c\}] = P_{\mathbf{Y}}[\{g = c\}] = P[\{g \circ \mathbf{Y} = c\}]$$

Daher gilt

$$\limsup_{n\to\infty} \sum_{k=1}^{n} X_k = c = \limsup_{n\to\infty} \sum_{k=1}^{n} Y_k$$

fast sicher und wegen $\sum_{k=1}^{n+1} X_k = X_1 + \sum_{k=1}^{n} Y_k$ folgt nun

$$c = X_1 + c$$

Wegen $P_X \neq \boldsymbol{\delta}_0$ gilt daher $c \in \{-\infty, \infty\}$. Es gilt also

$$P\left[\left\{\limsup_{n\to\infty} \sum_{k=1}^{n} X_k \in \{-\infty, \infty\}\right\}\right] = 1$$

und aus Symmetriegründen gilt auch

$$P\left[\left\{\liminf_{n\to\infty} \sum_{k=1}^{n} X_k \in \{-\infty, \infty\}\right\}\right] = 1$$

Daraus folgt die Behauptung des Satzes. □

Es bleibt die Aufgabe, die in Satz 15.4.4 angegebenen Alternativen geeignet zu charakterisieren. Dies gelingt in der Tat unter der zusätzlichen Bedingung, dass alle Zufallsvariablen der Folge einen endlichen Erwartungswert besitzen. Wir benötigen das folgende Lemma:

15.4.5 Lemma. *Für alle $m \in \mathbb{N}$ gilt*

$$\sum_{n=0}^{\infty} P[\{|S_n| < m\}] \leq 2\, m \sum_{n=0}^{\infty} P[\{|S_n| < 1\}]$$

Beweis. Sei $J \in \mathcal{J}(\mathbb{R})$ ein halboffenes Intervall der Länge 1 und für $n \in \mathbb{N}$ sei

$$A_n := \{S_n \in J\}$$

und

$$B_n := A_n \cap \bigcap_{k=1}^{n-1} \overline{A_k}$$

Dann gilt

$$\sum_{k=1}^{\infty} \chi_{A_k} = \sum_{k=1}^{\infty} \chi_{A_k} \sum_{n=1}^{k} \chi_{B_n} = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \chi_{B_n} \chi_{A_k}$$

und damit

$$\begin{aligned}
\sum_{k=1}^{\infty} P[\{S_k \in J\}] &= \sum_{k=1}^{\infty} P[A_k] \\
&= \int_\Omega \left(\sum_{k=1}^{\infty} \chi_{A_k} \right) dP \\
&= \int_\Omega \left(\sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \chi_{B_n} \chi_{A_k} \right) dP \\
&= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P[B_n \cap A_k]
\end{aligned}$$

Für alle $n \in \mathbb{N}$ und $k \in \mathbb{N}(n)$ gilt $B_n \cap A_k \subseteq A_n \cap A_k \subseteq \{|S_k - S_n| < 1\}$, also $B_n \cap A_k \subseteq B_n \cap \{|S_k - S_n| < 1\}$, und damit

$$\begin{aligned}
P[B_n \cap A_k] &\leq P[B_n \cap \{|S_k - S_n| < 1\}] \\
&= P[B_n]\, P[\{|S_k - S_n| < 1\}] \\
&= P[B_n]\, P[\{|S_{k-n}| < 1\}]
\end{aligned}$$

Also gilt

$$\begin{aligned}
\sum_{k=1}^{\infty} P[\{S_k \in J\}] &= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P[B_n \cap A_k] \\
&\leq \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P[B_n]\, P[\{|S_{k-n}| < 1\}] \\
&= \sum_{n=1}^{\infty} \sum_{j=0}^{\infty} P[B_n]\, P[\{|S_j| < 1\}] \\
&\leq \sum_{j=0}^{\infty} P[\{|S_j| < 1\}]
\end{aligned}$$

Für alle $n \in \{1,\dots,m\}$ sei $J_n := (-n,-n+1] \cup [n-1,n)$. Dann gilt

$$\sum_{k=0}^{\infty} P[\{S_k \in J_1\}] = \sum_{k=0}^{\infty} P[\{|S_k| < 1\}]$$

und für $n \in \{2,\dots,m\}$ gilt wegen $P[\{S_0 \in J_n\}] = 0$ und der vorher gezeigten Ungleichung

$$\begin{aligned}\sum_{k=0}^{\infty} P[\{S_k \in J_n\}] &= \sum_{k=1}^{\infty} P[\{S_k \in J_n\}] \\ &\le 2 \sum_{k=0}^{\infty} P[\{|S_k| < 1\}]\end{aligned}$$

Also gilt

$$\begin{aligned}\sum_{k=0}^{\infty} P[\{|S_k| < m\}] &= \sum_{k=0}^{\infty} P\left[\left\{S_k \in \sum_{n=1}^{m} J_n\right\}\right] \\ &= \sum_{k=0}^{\infty}\sum_{n=1}^{m} P[\{S_k \in J_n\}] \\ &= \sum_{n=1}^{m}\sum_{k=0}^{\infty} P[\{S_k \in J_n\}] \\ &\le 2m \sum_{k=0}^{\infty} P[\{|S_k| < 1\}]\end{aligned}$$

Damit ist das Lemma bewiesen. □

Der folgende Satz charakterisiert die in Satz 15.4.4 angegebenen Alternativen:

15.4.6 Satz (Chung/Fuchs). *Sei $X \in L^1$ mit $P_X \neq \boldsymbol{\delta}_0$. Dann gilt*
(1) $E[X] > 0$ *genau dann, wenn* $\lim_{n\to\infty} S_n = \infty$ *fast sicher gilt.*
(2) $E[X] = 0$ *genau dann, wenn* $\liminf_{n\to\infty} S_n = -\infty$ *und* $\limsup_{n\to\infty} S_n = \infty$ *fast sicher gilt.*
(3) $E[X] < 0$ *genau dann, wenn* $\lim_{n\to\infty} S_n = -\infty$ *fast sicher gilt.*

Beweis. Wir nehmen zunächst an, dass $E[X] > 0$ gilt. Nach dem 2. Gesetz der Großen Zahlen gilt dann

$$\lim_{n\to\infty} \frac{1}{n} S_n = E[X]$$

fast sicher und damit

$$\liminf_{n\to\infty} S_n \ge \lim_{n\to\infty} n \, \frac{E[X]}{2}$$

fast sicher. Daher gilt in diesem Fall $\lim_{n\to\infty} S_n = \infty$ fast sicher.
Wir nehmen nun an, dass $E[X] < 0$ gilt. Dann gilt aus Symmetriegründen $\lim_{n\to\infty} S_n = -\infty$ fast sicher.

Wir nehmen schließlich an, dass $E[X] = 0$ gilt. Nach dem 2. Gesetz der Großen Zahlen konvergiert die Folge $\{\frac{1}{n}S_n\}_{n\in\mathbb{N}}$ fast sicher und damit auch stochastisch gegen 0. Sei $\varepsilon \in (0,\infty)$. Dann gibt es ein $n(\varepsilon) \in \mathbb{N}$ mit

$$P[\{|S_n| < n\,\varepsilon\}] \geq \frac{1}{2}$$

für alle $n \in \mathbb{N}(n(\varepsilon))$. Für alle $m \in \mathbb{N}$ sei $m(\varepsilon) := \max\{n \in \mathbb{N} \mid n\,\varepsilon \leq m\}$. Dann gilt für alle $n \in \{n(\varepsilon),\dots,m(\varepsilon)\}$

$$P[\{|S_n| < m\}] \geq \frac{1}{2}$$

Für hinreichend große $m \in \mathbb{N}$ gilt $n(\varepsilon) \leq m(\varepsilon)$ und aus Lemma 15.4.5 folgt

$$\begin{aligned}
2\,m\sum_{n=0}^{\infty} P[\{|S_n| < 1\}] &\geq \sum_{n=0}^{\infty} P[\{|S_n| < m\}] \\
&\geq \sum_{n=n(\varepsilon)}^{m(\varepsilon)} P[\{|S_n| < m\}] \\
&\geq \frac{1}{2}\Big(m(\varepsilon)+1-n(\varepsilon)\Big) \\
&\geq \frac{1}{2}\left(\frac{m}{\varepsilon}-n(\varepsilon)\right) \\
&= \frac{m}{2\,\varepsilon}\left(1-\frac{\varepsilon\,n(\varepsilon)}{m}\right)
\end{aligned}$$

und damit

$$\sum_{n=0}^{\infty} P[\{|S_n| < 1\}] \geq \frac{1}{4\,\varepsilon}\left(1-\frac{\varepsilon\,n(\varepsilon)}{m}\right)$$

Daher gilt für alle $\varepsilon \in (0,\infty)$

$$\sum_{n=0}^{\infty} P[\{|S_n| < 1\}] \geq \frac{1}{4\,\varepsilon}$$

und daraus folgt

$$\sum_{n=0}^{\infty} P[\{|S_n| < 1\}] = \infty$$

Sei nun $k \in \mathbb{N}$. Für alle $m \in \mathbb{N}$ setzen wir

$$A_m := \{|S_m| < 1\} \cap \bigcap_{n=m+k}^{\infty} \{|S_n| \geq 1\}$$

Dann ist für jedes $i \in \{1,\dots,k\}$ die Folge $\{A_{i+jk}\}_{j\in\mathbb{N}_0}$ disjunkt, und es gilt

$$\sum_{m=1}^{\infty} P[A_m] = \sum_{i=1}^{k}\sum_{j=0}^{\infty} P[A_{i+jk}] = \sum_{i=1}^{k} P\left[\sum_{j=0}^{\infty} A_{i+jk}\right] \le k$$

Ferner gilt für alle $m \in \mathbb{N}$

$$\begin{aligned}
P[\{|S_m|<1\}]\,P\left[\bigcap_{l=k}^{\infty}\{|S_l|\ge 2\}\right] &= P[\{|S_m|<1\}]\,P\left[\bigcap_{n=m+k}^{\infty}\{|S_n-S_m|\ge 2\}\right]\\
&= P\left[\{|S_m|<1\}\cap\bigcap_{n=m+k}^{\infty}\{|S_n-S_m|\ge 2\}\right]\\
&\le P\left[\{|S_m|<1\}\cap\bigcap_{n=m+k}^{\infty}\{|S_n|\ge 1\}\right]\\
&= P[A_m]
\end{aligned}$$

und damit

$$\sum_{m=1}^{\infty} P[\{|S_m|<1\}]\,P\left[\bigcap_{l=k}^{\infty}\{|S_l|\ge 2\}\right] \le \sum_{m=1}^{\infty} P[A_m] \le k$$

Wegen $\sum_{m=0}^{\infty} P[\{|S_m|<1\}] = \infty$ gilt daher

$$P\left[\bigcap_{l=k}^{\infty}\{|S_l|\ge 2\}\right] = 0$$

Daher gilt

$$P\left[\bigcup_{k=1}^{\infty}\bigcap_{l=k}^{\infty}\{|S_l|\ge 2\}\right] = 0$$

und damit

$$P\left[\limsup_{n\to\infty}\{|S_n|<2\}\right] = P\left[\bigcap_{k=1}^{\infty}\bigcup_{l=k}^{\infty}\{|S_l|<2\}\right] = 1$$

Dann aber gilt weder $\lim_{n\to\infty} S_n = \infty$ noch $\lim_{n\to\infty} S_n = -\infty$, und aus Satz 15.4.4 folgt $\limsup_{n\to\infty} S_n = \infty$ und $\liminf_{n\to\infty} S_n = -\infty$ fast sicher.
Die Behauptung folgt nun aus Satz 15.4.4. □

Wir beenden diesen Abschnitt mit dem denkbar einfachsten Beispiel einer Irrfahrt:

15.4.7 Beispiel (Bernoulli–Irrfahrt). Sei $P_{(X+1)/2} = \mathbf{B}(\vartheta)$. Dann gilt
(1) $\vartheta > 1/2$ genau dann, wenn $\lim_{n\to\infty} S_n = \infty$ fast sicher.
(2) $\vartheta = 1/2$ genau dann, wenn $\liminf_{n\to\infty} S_n = -\infty$ und $\limsup_{n\to\infty} S_n = \infty$ fast sicher.
(3) $\vartheta < 1/2$ genau dann, wenn $\lim_{n\to\infty} S_n = -\infty$ fast sicher.

Teil V

Vertiefung der Wahrscheinlichkeitstheorie

16

Erzeugende Funktionen

In diesem Kapitel untersuchen wir für reelle Zufallsvariable Funktionen auf der Menge der reellen Zahlen, die durch die Erwartungswerte bestimmter Transformationen der Zufallsvariablen definiert sind und als erzeugende Funktionen bezeichnet werden. Offenbar sind die erzeugenden Funktionen einer reellen Zufallsvariablen durch deren Verteilung vollständig bestimmt. Daher könnte man erzeugende Funktionen auch durch die Integrale bestimmter reeller oder komplexer Funktionen bezüglich einer univariaten Verteilung definieren; da aber jede univariate Verteilung als Verteilung einer reellen Zufallsvariablen auf einem geeigneten Wahrscheinlichkeitsraum dargestellt werden kann, führt dies zu keiner echten Verallgemeinerung.

Das Interesse an erzeugenden Funktionen beruht darauf, dass sie in vielen Fällen die Bestimmung von Momenten einer Zufallsvariablen oder sogar die Bestimmung der Verteilung einer Zufallsvariablen erleichtern; diesen Eigenschaften verdanken die erzeugenden Funktionen ihren Namen.

Wir betrachten
- die wahrscheinlichkeitserzeugende Funktion einer Zufallsvariablen mit Werten in $\mathbb{N}_0$ (Abschnitt 16.1),
- die momenterzeugende Funktion (Abschnitt 16.2),
- die kumulantenerzeugende Funktion (Abschnitt 16.3), und
- die charakteristische Funktion (Abschnitt 16.4).

Dabei untersuchen wir vor allem die Eigenschaften der wahrscheinlichkeitserzeugenden Funktion und der charakteristischen Funktion.

Grundsätzlich können diese erzeugenden Funktionen auch für Zufallsvektoren und damit für multivariate Verteilungen definiert werden. Wir machen von dieser Möglichkeit jedoch keinen Gebrauch und beschränken uns auf die Darstellung der grundlegenden Eigenschaften der erzeugenden Funktionen von Zufallsvariablen.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_16, © Springer-Verlag Berlin Heidelberg 2011

16.1 Wahrscheinlichkeitserzeugende Funktion

Sei X eine Zufallsvariable mit $P_X[\mathbb{N}_0] = 1$. Dann besitzt für alle $t \in [0,1]$ die Zufallsvariable t^X einen endlichen Erwartungswert. Die Funktion $m_X : [0,1] \to [0,1]$ mit

$$m_X(t) := E[t^X]$$

heißt die *wahrscheinlichkeitserzeugende Funktion* von X. Wegen

$$m_X(t) = \sum_{n=0}^{\infty} P[\{X=n\}]\, t^n = \sum_{n=0}^{\infty} P_X[\{n\}]\, t^n$$

ist die wahrscheinlichkeitserzeugende Funktion von X durch die Verteilung von X bestimmt.

16.1.1 Beispiele (Diskrete Verteilungen).
(1) **Binomial–Verteilung:** Im Fall $P_X = \mathbf{B}(n,\vartheta)$ gilt

$$m_X(t) = (1-\vartheta+\vartheta t)^n$$

(2) **Poisson–Verteilung:** Im Fall $P_X = \mathbf{P}(\alpha)$ gilt

$$m_X(t) = e^{-\alpha(1-t)}$$

In der Tat: Es gilt

$$m_X(t) = \sum_{n=0}^{\infty} e^{-\alpha}\, \frac{\alpha^n}{n!}\, t^n = e^{-\alpha} \sum_{k=0}^{\infty} \frac{(\alpha t)^n}{n!} = e^{-\alpha}\, e^{\alpha t} = e^{-\alpha(1-t)}$$

(3) **Negativbinomial–Verteilung:** Im Fall $P_X = \mathbf{NB}(\alpha,\vartheta)$ gilt

$$m_X(t) = \left(\frac{1-\vartheta t}{1-\vartheta}\right)^{-\alpha}$$

Wir untersuchen nun die Eigenschaften der wahrscheinlichkeitserzeugenden Funktion.

16.1.2 Satz. *Sei X eine Zufallsvariable mit $P_X[\mathbb{N}_0] = 1$. Dann gilt:*
(1) *m_X ist monoton wachsend und stetig, und für alle $t \in [0,1]$ gilt*

$$0 \le m_X(t) \le m_X(1) = 1$$

(2) *m_X ist auf dem Intervall $[0,1)$ unendlich oft differenzierbar.*
(3) *Für alle $k \in \mathbb{N}_0$ und für alle $t \in [0,1)$ gilt*

$$m_X^{(k)}(t) = \sum_{n=k}^{\infty} P[\{X=n\}] \left(\prod_{i=0}^{k-1} (n-i)\right) t^{n-k}$$

(4) *Für alle $k \in \mathbb{N}_0$ ist $m_X^{(k)}(t)$ auf dem Intervall $[0,1)$ monoton wachsend und es gilt*

$$\sup_{t\in[0,1)} m_X^{(k)}(t) = \sum_{n=k}^{\infty} P[\{X=n\}] \prod_{i=0}^{k-1} (n-i)$$

Beweis. Da die Potenzreihe $\sum_{n=1}^{\infty} P[\{X = n\}]\, t^n$ auf $[-1, 1]$ konvergent ist, folgt aus den Eigenschaften von Potenzreihen, dass die wahrscheinlichkeitserzeugende Funktion auf $[0, 1)$ unendlich oft stetig differenzierbar ist. Daraus folgt (2).
Insbesondere ist die wahrscheinlichkeitserzeugende Funktion stetig auf $[0, 1)$ und es ist klar, dass sie auch positiv und monoton wachsend ist mit $m_X(1) = 1$.
Am Ende des Beweises zeigen wir, dass die wahrscheinlichkeitserzeugende Funktion auch in $t = 1$ stetig ist. Damit ist dann (1) gezeigt.
Der Nachweis von (3) ist elementar.
Sei nun $k \in \mathbb{N}_0$. Nach (3) ist die k–te Ableitung von m_X auf $[0, 1)$ monoton wachsend. Sei

$$c_k := \sup_{t\in[0,1)} m_X^{(k)}(t)$$

Dann gilt für alle $t \in [0, 1)$

$$\begin{aligned} m_X^{(k)}(t) &= \sum_{n=k}^{\infty} P[\{X = n\}] \left(\prod_{i=0}^{k-1} (n-i) \right) t^{n-k} \\ &\le \sum_{n=k}^{\infty} P[\{X = n\}] \prod_{i=0}^{k-1} (n-i) \end{aligned}$$

und damit

$$c_k \le \sum_{n=k}^{\infty} P[\{X = n\}] \prod_{i=0}^{k-1} (n-i)$$

Andererseits gilt für alle $m \in \mathbb{N}$ und für alle $t \in [0, 1)$

$$\begin{aligned} \sum_{n=k}^{m} P[\{X = n\}] \left(\prod_{i=0}^{k-1} (n-i) \right) t^{n-k} &\le \sum_{n=k}^{\infty} P[\{X = n\}] \left(\prod_{i=0}^{k-1} (n-i) \right) t^{n-k} \\ &= m_X^{(k)}(t) \\ &\le c_k \end{aligned}$$

und aus der Stetigkeit von Polynomen folgt nun

$$\sum_{n=k}^{m} P[\{X = n\}] \prod_{i=0}^{k-1} (n-i) \le c_k$$

Daraus folgt

$$\sum_{n=k}^{\infty} P[\{X = n\}] \prod_{i=0}^{k-1} (n-i) \le c_k$$

Zusammen mit der vorher gezeigten Ungleichung erhalten wir

$$\sum_{n=k}^{\infty} P[\{X=n\}] \prod_{i=0}^{k-1} (n-i) = c_k$$

Damit ist auch (4) bewiesen. Für $k = 0$ lautet die Gleichung aus (4)

$$\sup_{t \in [0,1)} m_X(t) = \sum_{n=0}^{\infty} P[\{X=n\}] = 1 = m_X(1)$$

Damit ist auch die Stetigkeit von m_X an der Stelle $t = 1$ gezeigt. □

Wir zeigen nun, dass die wahrscheinlichkeitserzeugende Funktion ihren Namen verdient. Dazu benötigen wir das folgende Lemma, das sich unmittelbar aus Satz 16.1.2 ergibt:

16.1.3 Lemma. *Sei X eine Zufallsvariable mit $P_X[\mathbb{N}_0] = 1$. Dann gilt für alle $k \in \mathbb{N}_0$*

$$P[\{X=k\}] = \frac{1}{k!}\, m_X^{(k)}(0)$$

Der folgende Satz zeigt, dass die Verteilung einer Zufallsvariablen X mit $P_X[\mathbb{N}_0] = 1$ durch ihre wahrscheinlichkeitserzeugende Funktion bestimmt ist:

16.1.4 Satz (Eindeutigkeitssatz). *Seien X und Y Zufallsvariable mit $P_X[\mathbb{N}_0] = 1 = P_Y[\mathbb{N}_0]$. Dann sind äquivalent*:
(a) *Es gilt $P_X = P_Y$.*
(b) *Es gilt $m_X = m_Y$.*

Die Aussage des Satzes ergibt sich unmittelbar aus Lemma 16.1.3.

Für eine Zufallsvariable X mit $P_X[\mathbb{N}_0] = 1$ und für $k \in \mathbb{N}_0$ heißt

$$E\left[\binom{X}{k}\right]$$

das *Binomial–Moment* der Ordnung k von X. Es gilt

$$E\left[\binom{X}{k}\right] = \sum_{n=k}^{\infty} \binom{n}{k} P[\{X=n\}]$$

und aus Satz 16.1.2 ergibt sich

$$E\left[\binom{X}{k}\right] = \sup_{t \in [0,1)} \frac{1}{k!}\, m_X^{(k)}(t)$$

Der folgende Satz gibt notwendige und hinreichende Bedingungen dafür an, dass das Binomial–Moment der Ordnung k endlich ist:

16.1.5 Satz. *Sei X eine Zufallsvariable mit $P_X[\mathbb{N}_0] = 1$. Für $k \in \mathbb{N}$ sind folgende Aussagen äquivalent:*
(a) *Es gilt*

$$E\left[\binom{X}{k}\right] < \infty$$

(b) *Es gilt*

$$E[X^k] < \infty$$

(c) *m_X ist an der Stelle $t = 1$ k–mal stetig differenzierbar.*
In diesem Fall gilt

$$E\left[\binom{X}{k}\right] = \frac{1}{k!}\, m_X^{(k)}(1)$$

Beweis. Wir nehmen zunächst an, dass (a) gilt. Für alle $l \in \mathbb{N}$ gilt

$$\begin{aligned} E\left[\binom{X}{l}\right] &= \sum_{n=l}^{\infty} \binom{n}{l} P[\{X=n\}] \\ &= \sum_{n=l}^{\infty} \binom{n}{l-1} \frac{n-l+1}{l} P[\{X=n\}] \\ &\geq \sum_{n=2l-1}^{\infty} \binom{n}{l-1} P[\{X=n\}] \end{aligned}$$

Da es für die Konvergenz einer Reihe auf ihre ersten Glieder nicht ankommt, folgt für alle $l \in \{1, \ldots, k\}$ mit

$$E\left[\binom{X}{l}\right] < \infty$$

aus der vorigen Ungleichung

$$E\left[\binom{X}{l-1}\right] < \infty$$

Wegen Satz 16.1.2 und aufgrund der vorher gezeigten Implikation gilt für alle $l \in \{0, 1, \ldots, k\}$

$$\sup_{t\in[0,1)} m_X^{(l)}(t) = l!\, E\left[\binom{X}{l}\right] < \infty$$

Aus den Eigenschaften von Potenzreihen folgt nun, dass m_X an der Stelle $t = 1$ k–mal stetig differenzierbar ist. Daher folgt (c) aus (a).

Wir nehmen nun an, dass (c) gilt. Dann gilt

$$E\left[\binom{X}{k}\right] = \sup_{t\in[0,1)} \frac{1}{k!}\, m_X^{(k)}(t) = \frac{1}{k!}\, m_X^{(k)}(1)$$

Daher folgt (a) aus (c).
Wir nehmen nochmals an, dass (a) gilt. Aufgrund der bereits gezeigten Äquivalenz von (a) und (c) gilt für alle $l \in \{0,1,\dots,k\}$

$$E\left[\binom{X}{l}\right] < \infty$$

Durch vollständige Induktion zeigt man, dass sich alle Momente der Ordnung $l \in \{0,1,\dots,k\}$ von X in der Form

$$E[X^l] = \sum_{j=0}^{l} a_{l,j}\, E\left[\binom{X}{j}\right]$$

mit $a_{l,0},\dots,a_{l,l} \in \mathbb{R}$ darstellen lassen und dass für alle $l \in \{0,1,\dots,k\}$

$$E[X^l] < \infty$$

gilt. Daher folgt (b) aus (a).
Wir nehmen abschließend an, dass (b) gilt. Wegen

$$E\left[\binom{X}{k}\right] \le \frac{1}{k!}\, E[X^k]$$

folgt aus $E[X^k] < \infty$

$$E\left[\binom{X}{k}\right] < \infty$$

Daher folgt (a) aus (b). □

Für eine Zufallsvariable X mit $P_X[\mathbb{N}_0] = 1$ lassen sich der Erwartungswert und die Varianz besonders leicht mit Hilfe der wahrscheinlichkeitserzeugenden Funktion bestimmen:

16.1.6 Folgerung. *Sei X eine Zufallsvariable mit $P_X[\mathbb{N}_0] = 1$.*
(1) *Besitzt X einen endlichen Erwartungswert, so gilt*

$$E[X] = m_X'(1)$$

(2) *Besitzt X ein endliches zweites Moment, so gilt*

$$\mathrm{var}\,[X] = m_X''(1) + m_X'(1) - (m_X'(1))^2$$

Beweis. Die Formel für den Erwartungswert ist unmittelbar klar aus Satz 16.1.5. Die Formel für die Varianz ergibt sich ebenfalls aus Satz 16.1.5, denn es gilt

$$\begin{aligned}\operatorname{var}[X] &= E[X^2] - (E[X])^2 \\ &= E[X(X-1)] + E[X] - (E[X])^2 \\ &= m''_X(1) + m'_X(1) - (m'_X(1))^2\end{aligned}$$

Damit ist die Folgerung gezeigt. □

16.1.7 Beispiele (Diskrete Verteilungen).

(1) **Binomial–Verteilung:** Im Fall $P_X = \mathbf{B}(n, \vartheta)$ gilt

$$\begin{aligned}m'_X(t) &= n\vartheta(1-\vartheta+\vartheta t)^{n-1} \\ m''_X(t) &= n(n-1)\vartheta^2(1-\vartheta+\vartheta t)^{n-2}\end{aligned}$$

und damit $E[X] = n\,\vartheta$ und $\operatorname{var}[X] = n\,\vartheta(1-\vartheta)$.

(2) **Poisson–Verteilung:** Im Fall $P_X = \mathbf{P}(\alpha)$ gilt

$$\begin{aligned}m'_X(t) &= \alpha\, e^{-\alpha(1-t)} \\ m''_X(t) &= \alpha^2\, e^{-\alpha(1-t)}\end{aligned}$$

und damit $E[X] = \alpha$ und $\operatorname{var}[X] = \alpha$.

(3) **Negativbinomial–Verteilung:** Im Fall $P_X = \mathbf{NB}(\alpha, \vartheta)$ gilt

$$\begin{aligned}m'_X(t) &= \alpha\,\frac{\vartheta}{1-\vartheta}\left(\frac{1-\vartheta t}{1-\vartheta}\right)^{-(\alpha+1)} \\ m''_X(t) &= \alpha(\alpha+1)\left(\frac{\vartheta}{1-\vartheta}\right)^2\left(\frac{1-\vartheta t}{1-\vartheta}\right)^{-(\alpha+2)}\end{aligned}$$

und damit $E[X] = \alpha\vartheta/(1-\vartheta)$ und $\operatorname{var}[X] = \alpha\vartheta/(1-\vartheta)^2$.

Mit Hilfe der wahrscheinlichkeitserzeugenden Funktion lässt sich auch die Verteilung einer Summe von unabhängigen Zufallsvariablen bestimmen:

16.1.8 Satz. *Sei* $\mathbf{X} : \Omega \to \mathbb{R}^d$ *ein Zufallsvektor mit* $P_{\mathbf{X}}[\mathbb{N}_0^d] = 1$ *und unabhängigen Koordinaten. Dann gilt*

$$m_{\mathbf{1}'\mathbf{X}}(t) = \prod_{i=1}^d m_{X_i}(t)$$

Beweis. Aus der Unabhängigkeit der Familie $\{X_i\}_{i\in\{1,\dots,d\}}$ folgt die Unabhängigkeit der Familie $\{t^{X_i}\}_{i\in\{1,\dots,d\}}$. Daher gilt

$$m_{\mathbf{1}'\mathbf{X}}(t) = E\big[t^{\mathbf{1}'\mathbf{X}}\big] = E\Big[t^{\sum_{i=1}^d X_i}\Big] = E\left[\prod_{i=1}^d t^{X_i}\right] = \prod_{i=1}^d E\big[t^{X_i}\big] = \prod_{i=1}^d m_{X_i}(t)$$

Damit ist der Satz gezeigt. □

16.1.9 Beispiele (Diskrete Verteilungen). Seien X und Y unabhängige reelle Zufallsvariable.

(1) **Binomial–Verteilung:** Im Fall $P_X = \mathbf{B}(m, \vartheta)$ und $P_Y = \mathbf{B}(n, \vartheta)$ gilt

$$P_{X+Y} = \mathbf{B}(m{+}n, \vartheta)$$

(2) **Poisson–Verteilung:** Im Fall $P_X = \mathbf{P}(\alpha)$ und $P_Y = \mathbf{P}(\beta)$ gilt

$$P_{X+Y} = \mathbf{P}(\alpha{+}\beta)$$

In der Tat: Es gilt

$$m_{X+Y}(t) = m_X(t)\, m_Y(t) = e^{-\alpha(1-t)}\, e^{-\beta(1-t)} = e^{-(\alpha+\beta)(1-t)}$$

(3) **Negativbinomial–Verteilung:** Im Fall $P_X = \mathbf{NB}(\alpha, \vartheta)$ und $P_Y = \mathbf{NB}(\beta, \vartheta)$ gilt

$$P_{X+Y} = \mathbf{NB}(\alpha{+}\beta, \vartheta)$$

Aus Satz 16.1.8 ergibt sich insbesondere ein Spezialfall der aus Folgerung 13.5.4 bekannten Faltungsformel:

16.1.10 Folgerung (Faltungsformel). *Seien X und Y unabhängige Zufallsvariable mit $P_X[\mathbb{N}_0] = 1 = P_Y[\mathbb{N}_0]$. Dann gilt für alle $n \in \mathbb{N}_0$*

$$P[\{X + Y = n\}] = \sum_{k=0}^{n} P[\{X = k\}]\, P[\{Y = n{-}k\}]$$

Beweis. Nach Satz 16.1.8 gilt $m_{X+Y}(t) = m_X(t)\, m_Y(t)$. Aus der Produktregel der Differentialrechnung folgt

$$m_{X+Y}^{(n)}(t) = \sum_{k=0}^{n} \binom{n}{k} m_X^{(k)}(t)\, m_Y^{(n-k)}(t)$$

und damit

$$\begin{aligned} P[\{X{+}Y = n\}] &= \frac{m_{X+Y}^{(n)}(0)}{n!} \\ &= \sum_{k=0}^{n} \frac{m_X^{(k)}(0)}{k!}\, \frac{m_Y^{(n-k)}(0)}{(n{-}k)!} \\ &= \sum_{k=0}^{n} P[\{X = k\}]\, P[\{Y = n{-}k\}] \end{aligned}$$

Damit ist die Folgerung bewiesen. □

Die Verwendung der wahrscheinlichkeitserzeugenden Funktion ist vor allem dann von Vorteil, wenn sie in geschlossener Form darstellbar ist.

Aufgaben

16.1.A Jede wahrscheinlichkeitserzeugende Funktion ist konvex.

16.1.B Für eine Zufallsvariable X mit $P_X[\mathbb{N}_0] = 1$ sind folgende Aussagen äquivalent:
(a) Es gilt $P_X = \boldsymbol{\delta}_0$.
(b) Es gilt $m_X(0) = 1$.
(c) Es gibt ein $t \in [0,1)$ mit $m_X(t) = 1$.
(d) Für alle $t \in [0,1]$ gilt $m_X(t) = 1$.
(e) m_X ist an der Stelle $t = 1$ differenzierbar und es gilt $m'_X(1) = 0$.
(f) m_X ist an der Stelle $t = 1$ differenzierbar und für alle $t \in [0,1]$ gilt $m'_X(t) = 0$.

16.1.C **Panjer–Verteilung:** Eine reelle Zufallsvariable X mit $P_X[\mathbb{N}_0] = 1$ besitzt eine *Panjer–Verteilung*, wenn es $a, b \in \mathbb{R}$ gibt mit $a + b > 0$ derart, dass für alle $n \in \mathbb{N}$

$$P[\{X = n\}] = \left(a + \frac{b}{n}\right) P[\{X = n-1\}]$$

gilt.
(1) Für eine Zufallsvariable X mit $P_X[\mathbb{N}_0] = 1$ und für $a, b \in \mathbb{R}$ mit $a + b > 0$ sind folgende Aussagen äquivalent:
(a) Für alle $n \in \mathbb{N}$ gilt

$$P[\{X = n\}] = \left(a + \frac{b}{n}\right) P[\{X = n-1\}]$$

(b) Für alle $t \in [0,1)$ gilt

$$(1-at)m'_X(t) = (a+b)m_X(t)$$

(c) Für alle $n \in \mathbb{N}$ und $t \in [0,1)$ gilt

$$(1-at)m_X^{(n)}(t) = (na+b)m_X^{(n-1)}(t)$$

In diesem Fall gilt $a < 1$.
(2) Ist X eine reelle Zufallsvariable, die eine Panjer–Verteilung besitzt, so gilt

$$E[X] = \frac{a+b}{1-a}$$
$$\mathrm{var}\,[X] = \frac{a+b}{(1-a)^2}$$

(3) Eine reelle Zufallsvariable besitzt genau dann eine Panjer–Verteilung, wenn sie eine Binomial–, Poisson– oder Negativbinomial–Verteilung besitzt.
(4) Charakterisieren Sie die Binomial–, Poisson– und Negativbinomial–Verteilungen nach der Lage des Parameters a und nach dem Verhältnis zwischen der Varianz und dem Erwartungswert.

16.1.D **Geometrische Verteilung:** Sei $P_X = \mathbf{Geo}(n, \vartheta)$. Dann gilt

$$m_X(t) = \left(\frac{\vartheta t}{1 - (1-\vartheta)t}\right)^n$$

Bestimmen Sie den Erwartungswert und die Varianz von X.

16.1.E Logarithmische Verteilung: Sei $P_X = \mathbf{Log}(\vartheta)$. Dann gilt

$$m_X(t) = \frac{\log(1-\vartheta t)}{\log(1-\vartheta)}$$

Bestimmen Sie den Erwartungswert und die Varianz von X.

16.1.F Wahrscheinlichkeitserzeugende Funktion eines Zufallsvektors: Sei $\mathbf{X} : \Omega \to \mathbb{R}^d$ ein Zufallsvektor mit $P_{\mathbf{X}}[\mathbb{N}_0^d] = 1$. Dann heißt die Funktion $m_{\mathbf{X}} : [\mathbf{0}, \mathbf{1}] \to [0,1]$ mit

$$m_{\mathbf{X}}(\mathbf{t}) := E\left[\prod_{i=1}^{d} t_i^{X_i}\right]$$

die *wahrscheinlichkeitserzeugende Funktion* von $\mathbf{X}$. Untersuchen Sie die Eigenschaften von $m_{\mathbf{X}}$.

16.2 Momenterzeugende Funktion

Sei X eine Zufallsvariable. Dann besitzt für alle $t \in \mathbb{R}$ die Zufallsvariable e^{tX} einen Erwartungswert. Die Funktion $M_X : \mathbb{R} \to [0,\infty]$ mit

$$M_X(t) := E[e^{tX}]$$

heißt die *momenterzeugende Funktion* von X. Wegen

$$M_X(t) = \int_{\mathbb{R}} e^{tx}\, dP_X(x)$$

ist die momenterzeugende Funktion von X durch die Verteilung von X bestimmt.

16.2.1 Beispiele (Absolutstetige Verteilungen).

(1) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha, \gamma)$ gilt

$$M_X(t) = \begin{cases} \left(\dfrac{\alpha}{\alpha - t}\right)^{\gamma} & \text{falls } t \in (-\infty, \alpha) \\ +\infty & \text{sonst} \end{cases}$$

In der Tat: Für alle $t \in (-\infty, \alpha)$ gilt

$$\begin{aligned} M_X(t) &= \int_{\mathbb{R}} e^{tx} \cdot \frac{\alpha^\gamma}{\Gamma(\gamma)}\, e^{-\alpha x} x^{\gamma-1} \chi_{(0,\infty)}(x)\, d\boldsymbol{\lambda}(x) \\ &= \frac{\alpha^\gamma}{(\alpha-t)^\gamma} \int_{\mathbb{R}} \frac{(\alpha-t)^\gamma}{\Gamma(\gamma)}\, e^{-(\alpha-t)x} x^{\gamma-1} \chi_{(0,\infty)}(x)\, d\boldsymbol{\lambda}(x) \\ &= \left(\frac{\alpha}{\alpha-t}\right)^{\gamma} \end{aligned}$$

(2) **Standardnormal–Verteilung:** Im Fall $P_X = \mathbf{N}(0,1)$ gilt

$$M_X(t) = \exp(t^2/2)$$

Das folgende Lemma ergibt sich unmittelbar aus der Definition der momenterzeugenden Funktion:

16.2.2 Lemma. *Sei X eine Zufallsvariable. Dann gilt für alle $a, b \in \mathbb{R}$*

$$M_{a+bX}(t) = e^{at} M_X(bt)$$

Auch die folgenden Eigenschaften der momenterzeugenden Funktion einer Zufallsvariablen sind offensichtlich:

16.2.3 Lemma. *Sei X eine Zufallsvariable. Dann gilt $M_X(0) = 1$ und für alle $t \in \mathbb{R}$ gilt $M_X(t) \in (0, \infty]$.*

Die momenterzeugende Funktion einer Zufallsvariablen ist vor allem dann von Interesse, wenn sie in einer Umgebung von 0 endlich ist:

16.2.4 Satz. *Sei X eine Zufallsvariable. Wenn es ein $a \in (0, \infty)$ gibt mit $M_X(t) < \infty$ für alle $t \in (-a, a)$, dann gilt:*
(1) *X besitzt endliche Momente beliebiger Ordnung.*
(2) *Für alle $t \in (-a, a)$ gilt*

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E[X^k]$$

(3) *M_X ist auf dem Intervall $(-a, a)$ unendlich oft differenzierbar und für alle $k \in \mathbb{N}$ gilt*

$$E[X^k] = M_X^{(k)}(0)$$

(4) *Besitzt M_X die Potenzreihendarstellung*

$$M_X(t) = \sum_{k=0}^{\infty} a_k t^k$$

so gilt für alle $k \in \mathbb{N}$

$$E[X^k] = a_k k!$$

Beweis. Sei $t \in (-a, a)$. Nach Voraussetzung sind e^{tX} und e^{-tX} integrierbar, und wegen $e^{|tX|} \le e^{tX} + e^{-tX}$ ist auch $e^{|tX|}$ integrierbar. Nach dem Satz über die monotone Konvergenz gilt

$$\sum_{k=0}^{\infty} E\left[\frac{|tX|^k}{k!}\right] = E\left[\sum_{k=0}^{\infty} \frac{|tX|^k}{k!}\right] = E\big[e^{|tX|}\big]$$

und aus der Integrierbarkeit von $e^{|tX|}$ folgt nun (mit $t \neq 0$), dass X endliche Momente beliebiger Ordnung besitzt. Des weiteren gilt

$$e^{tX} = \sum_{k=0}^{\infty} \frac{t^k X^k}{k!}$$

und

$$\sup_{n\in\mathbb{N}_0} \left|\sum_{k=0}^{n} \frac{t^k X^k}{k!}\right| \leq \sum_{k=0}^{\infty} \frac{|tX|^k}{k!}$$

und aus Folgerung 8.3.11 ergibt sich nun

$$E[e^{tX}] = \sum_{k=0}^{\infty} \frac{t^k}{k!}\, E[X^k]$$

Damit sind (1) und (2) bewiesen, und die übrigen Aussagen ergeben sich aus den Eigenschaften von Potenzreihen. □

Wir geben abschließend ein Beispiel für die Anwendung der letzten Aussage des Satzes:

16.2.5 Beispiel (Exponential–Verteilung). Sei $P_X = \mathbf{Exp}(\alpha)$. Dann gilt für alle $t \in (-\alpha, \alpha)$

$$M_X(t) = \frac{\alpha}{\alpha - t} = \sum_{k=0}^{\infty} \frac{1}{\alpha^k}\, t^k$$

Für alle $k \in \mathbb{N}$ gilt daher $E[X^k] = k!/\alpha^k$.

Aufgaben

16.2.A Jede momenterzeugende Funktion ist konvex.

16.2.B Sei X eine Zufallsvariable mit $P_X[\mathbb{R}_+] = 1$. Dann ist M_X auf dem Intervall $(-\infty, 0]$ endlich.

16.2.C **Normal–Verteilung:** Sei $P_X = \mathbf{N}(\mu, \sigma^2)$. Dann gilt

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2}{2}\, t^2\right)$$

und

$$E[(X-\mu)^n] = \begin{cases} \sigma^n \prod_{j=1}^{n/2} (2j-1) & \text{falls } n \in 2\mathbb{N} \\ 0 & \text{sonst} \end{cases}$$

16.2.D Sei $\mathbf{X} : \Omega \to \mathbb{R}^d$ ein Zufallsvektor mit unabhängigen Koordinaten. Dann gilt

$$M_{\mathbf{1}'\mathbf{X}}(t) = \prod_{i=1}^{d} M_{X_i}(t)$$

16.2.E **Momenterzeugende Funktion eines Zufallsvektors:** Sei $\mathbf{X} : \Omega \to \mathbb{R}^d$ ein Zufallsvektor. Dann heißt die Funktion $M_{\mathbf{X}} : \mathbb{R}^d \to [0, \infty]$ mit

$$M_{\mathbf{X}}(\mathbf{t}) := E[e^{\mathbf{t}'\mathbf{X}}]$$

die *momenterzeugende Funktion* von $\mathbf{X}$. Untersuchen Sie die Eigenschaften von $M_{\mathbf{X}}$.

16.3 Kumulantenerzeugende Funktion

Sei X eine Zufallsvariable. Die Funktion $C_X : \mathbb{R} \to \bar{\mathbb{R}}$ mit

$$C_X(t) := \log(E[e^{tX}])$$

(mit $\log(\infty) := \infty$) heißt die *kumulantenerzeugende Funktion* von X. Es gilt $C_X = \log \circ M_X$. Wegen

$$C_X(t) = \log\left(\int_{\mathbb{R}} e^{tx} \, dP_X(x)\right)$$

ist die kumulantenerzeugende Funktion von X durch die Verteilung von X bestimmt.

16.3.1 Beispiele (Absolutstetige Verteilungen).

(1) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha, \gamma)$ gilt

$$C_X(t) = \begin{cases} \gamma\Big(\log(\alpha) - \log(\alpha - t)\Big) & \text{falls } t \in (-\infty, \alpha) \\ +\infty & \text{sonst} \end{cases}$$

(2) **Standardnormal–Verteilung:** Im Fall $P_X = \mathbf{N}(0, 1)$ gilt

$$C_X(t) = t^2/2$$

Das folgende Lemma ergibt sich unmittelbar aus der Definition der kumulantenerzeugenden Funktion:

16.3.2 Lemma. *Sei X eine Zufallsvariable. Dann gilt für alle $a, b \in \mathbb{R}$*

$$C_{a+bX}(t) = at + C_X(bt)$$

Die kumulantenerzeugende Funktion einer Zufallsvariablen ist vor allem dann von Interesse, wenn ihre momenterzeugende Funktion in einer Umgebung von 0 endlich ist:

16.3.3 Satz. *Sei X eine Zufallsvariable. Wenn es ein $a \in (0, \infty)$ gibt mit $M_X(t) < \infty$ für alle $t \in (-a, a)$, dann ist C_X auf dem Intervall $(-a, a)$ unendlich oft differenzierbar und es gilt*

$$C'(t) = \frac{M'_X(t)}{M_X(t)}$$

und

$$C''(t) = \frac{M''_X(t)M_X(t) - (M'_X(t))^2}{(M_X(t))^2}$$

und damit $C'(0) = E[X]$ und $C''(0) = \text{var}\,[X]$.

Beweis. Die Aussage ergibt sich unmittelbar aus Satz 16.2.4. □

16.3.4 Beispiele (Absolutstetige Verteilungen).

(1) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha, \gamma)$ gilt für alle $t \in (-\infty, \alpha)$

$$C'_X(t) = \frac{\gamma}{\alpha - t}$$
$$C''_X(t) = \frac{\gamma}{(\alpha - t)^2}$$

und damit $E[X] = \gamma/\alpha$ und $\text{var}\,[X] = \gamma/\alpha^2$.

(2) **Standardnormal–Verteilung:** Im Fall $P_X = \mathbf{N}(0, 1)$ gilt

$$C'_X(t) = t$$
$$C''_X(t) = 1$$

und damit $E[X] = 0$ und $\text{var}\,[X] = 1$.

Aufgaben

16.3.A Sei X eine Zufallsvariable mit einem endlichen Erwartungswert. Dann gilt für alle $t \in (0, \infty)$

$$E[X] \leq \frac{1}{t} \log\Big(E[e^{tX}]\Big)$$

16.3.B **Normal–Verteilung:** Sei $P_X = \mathbf{N}(\mu, \sigma^2)$. Dann gilt

$$C_X(t) = \mu t + \frac{\sigma^2}{2}\, t^2$$

und damit $E[X] = \mu$ und $\text{var}\,[X] = \sigma^2$.

16.3.C Sei $\mathbf{X} : \Omega \to \mathbb{R}^d$ ein Zufallsvektor mit unabhängigen Koordinaten. Dann gilt

$$C_{\mathbf{1}'\mathbf{X}}(t) = \sum_{i=1}^{d} C_{X_i}(t)$$

16.3.D **Kumulantenerzeugende Funktion eines Zufallsvektors:** Sei $\mathbf{X} : \Omega \to \mathbb{R}^d$ ein Zufallsvektor. Dann heißt die Funktion $C_{\mathbf{X}} : \mathbb{R}^d \to \bar{\mathbb{R}}$ mit

$$C_{\mathbf{X}}(\mathbf{t}) := \log(E[e^{\mathbf{t}'\mathbf{X}}])$$

die *kumulantenerzeugende Funktion* von $\mathbf{X}$. Untersuchen Sie die Eigenschaften von $C_{\mathbf{X}}$.

16.4 Charakteristische Funktion

Sei X eine Zufallsvariable. Die Funktion $\psi_X : \mathbb{R} \to \mathbb{C}$ mit

$$\psi_X(t) := E[e^{itX}]$$

heißt die *charakteristische Funktion* von X. Wegen

$$\psi_X(t) = \int_{\mathbb{R}} e^{itx}\, dP_X(x)$$

ist die charakteristische Funktion von X durch die Verteilung von X bestimmt.

16.4.1 Beispiele (Diskrete Verteilungen).
(1) **Binomial–Verteilung:** Im Fall $P_X = \mathbf{B}(n, \vartheta)$ gilt

$$\psi_X(t) = \Bigl(1 - \vartheta + \vartheta\, e^{it}\Bigr)^n$$

(2) **Poisson–Verteilung:** Im Fall $P_X = \mathbf{P}(\alpha)$ gilt

$$\psi_X(t) = \exp(-\alpha(1-e^{it}))$$

(3) **Negativbinomial–Verteilung:** Im Fall $P_X = \mathbf{NB}(\alpha, \vartheta)$ gilt

$$\psi_X(t) = \left(\frac{1-\vartheta e^{it}}{1-\vartheta}\right)^{-\alpha}$$

16.4.2 Beispiele (Absolutstetige Verteilungen).
(1) **Uniforme Verteilung:** Im Fall $P_X = \mathbf{U}(a, b)$ gilt

$$\psi_X(t) = \frac{e^{ibt} - e^{iat}}{i(b-a)t}$$

(2) **Gamma–Verteilung:** Im Fall $P_X = \mathbf{Ga}(\alpha, \gamma)$ gilt

$$\psi_X(t) = \left(\frac{\alpha}{\alpha - it}\right)^{\gamma}$$

(3) **Standardnormal–Verteilung:** Im Fall $P_X = \mathbf{N}(0, 1)$ gilt

$$\psi_X(t) = \exp(-t^2/2)$$

Das folgende Lemma ergibt sich unmittelbar aus der Definition der charakteristischen Funktion:

16.4.3 Lemma. *Sei X eine Zufallsvariable. Dann gilt für alle $a, b \in \mathbb{R}$*

$$\psi_{a+bX}(t) = e^{iat}\, \psi_X(bt)$$

Für den Nachweis vieler Eigenschaften der charakteristischen Funktion einer Zufallsvariablen ist das folgende Lemma wesentlich:

16.4.4 Lemma. *Für alle $x \in \mathbb{R}$ und $n \in \mathbb{N}_0$ gilt*

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \min\left\{ \frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right\}$$

Insbesondere gilt für jede Zufallsvariable X mit $E[|X|^n] < \infty$ und für alle $t \in \mathbb{R}$

$$\left| \psi_X(t) - \sum_{k=0}^{n} \frac{(it)^k}{k!} E[X^k] \right| \leq E\left[\min\left\{ \frac{|tX|^{n+1}}{(n+1)!}, \frac{2|tX|^n}{n!} \right\} \right]$$

Beweis. Wir nehmen zunächst an, dass $x \geq 0$ gilt. Durch partielle Integration erhält man für alle $n \in \mathbb{N}_0$

$$\int_0^x (x-s)^n e^{is}\, ds = \frac{x^{n+1}}{n+1} + \frac{i}{n+1} \int_0^x (x-s)^{n+1} e^{is}\, ds$$

und durch vollständige Induktion erhält man nun die Gleichung

$$e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + \frac{i^{n+1}}{n!} \int_0^x (x-s)^n e^{is}\, ds$$

Wegen $|\int_0^x (x-s)^n e^{is}\, ds| \leq |x|^{n+1}/(n+1)$ ergibt sich daraus

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| = \left| \frac{i^{n+1}}{n!} \int_0^x (x-s)^n e^{is}\, ds \right| \leq \frac{|x|^{n+1}}{(n+1)!}$$

und damit die erste Abschätzung. Für $n = 0$ ist die zweite Abschätzung trivial, und für $n \in \mathbb{N}$ ergibt sich aus der Gleichung

$$\begin{aligned} e^{ix} &= \sum_{k=0}^{n-1} \frac{(ix)^k}{k!} + \frac{i^n}{(n-1)!} \int_0^x (x-s)^{n-1} e^{is}\, ds \\ &= \sum_{k=0}^{n} \frac{(ix)^k}{k!} + \frac{i^n}{(n-1)!} \int_0^x (x-s)^{n-1} e^{is}\, ds - \frac{(ix)^n}{n!} \end{aligned}$$

und der vorher gezeigten Ungleichung

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \left| \frac{i^n}{(n-1)!} \int_0^x (x-s)^{n-1} e^{is}\, ds \right| + \frac{|x|^n}{n!} \leq 2\, \frac{|x|^n}{n!}$$

und damit die zweite Abschätzung.
Im Fall $x < 0$ erhält man zunächst

$$\int_x^0 (s-x)^n e^{is}\, ds = \frac{(-x)^{n+1}}{n+1} - \frac{i}{n+1} \int_x^0 (s-x)^{n+1} e^{is}\, ds$$

und sodann

$$e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + \frac{(-i)^{n+1}}{n!} \int_x^0 (s-x)^n e^{is}\, ds$$

Der Beweis der Abschätzungen verläuft dann analog zum Fall $x \geq 0$. □

Aus dem Lemma ergibt sich insbesondere ein Zusammenhang zwischen der Endlichkeit bestimmter Momente einer Zufallsvariablen und dem Grad der Differenzierbarkeit ihrer charakteristischen Funktion:

16.4.5 Lemma. *Sei X eine Zufallsvariable mit einem endlichen Moment der Ordnung $n \in \mathbb{N}_0$. Dann ist ψ_X n–mal differenzierbar und für alle $k \in \{0, 1, \dots, n\}$ gilt*

$$\psi_X^{(k)}(t) = E[(iX)^k e^{itX}]$$

und insbesondere

$$\psi_X^{(k)}(0) = i^k E[X^k]$$

und $\psi_X^{(k)}$ ist gleichmäßig stetig.

Beweis. Wir beweisen zunächst die Aussage über die Differenzierbarkeit und führen den Beweis durch vollständige Induktion.

- $k = 0$: In diesem Fall ist nichts zu zeigen.
- $k \to k + 1$: Wir nehmen an, die Aussage über die Differenzierbarkeit sei für ein $k \in \{0, 1, \dots, n-1\}$ bereits bewiesen. Dann gilt für alle $h \in \mathbb{R}\setminus\{0\}$

$$\begin{aligned}
&\left| \frac{\psi_X^{(k)}(t+h) - \psi_X^{(k)}(t)}{h} - E[(iX)^{k+1} e^{itX}] \right| \\
&= \left| \frac{E[(iX)^k e^{i(t+h)X}] - E[(iX)^k e^{itX}] - E[h(iX)^{k+1} e^{itX}]}{h} \right| \\
&= \left| E\left[(iX)^k e^{itX} \left(\frac{e^{ihX} - 1 - ihX}{h} \right) \right] \right| \\
&\leq E\left[|X|^k \left| \frac{e^{ihX} - 1 - ihX}{h} \right| \right]
\end{aligned}$$

  Nach Lemma 16.4.4 ist der letzte Integrand durch die integrierbare Zufallsvariable $2|X|^{k+1}$ beschränkt und konvergiert für $h \to 0$ gegen 0. Aus dem Satz über die majorisierte Konvergenz folgt nun, dass $\psi_X^{(k)}$ differenzierbar ist mit

$$\psi_X^{(k+1)}(t) = E[(iX)^{k+1} e^{itX}]$$

  und damit

$$\psi_X^{(k+1)}(0) = i^{k+1} E[X^{k+1}]$$

Wir beweisen nun die Aussage über die gleichmäßige Stetigkeit. Es gilt

$$\begin{aligned}\left|\psi_X^{(k)}(t+h) - \psi_X^{(k)}(t)\right| &= \left|E[(iX)^k e^{i(t+h)X}] - E[(iX)^k e^{itX}]\right| \\ &= \left|E\Big[(iX)^k e^{itX}(e^{ihX}-1)\Big]\right| \\ &\leq E\Big[|X|^k |e^{ihX}-1|\Big]\end{aligned}$$

Da der letzte Integrand durch die integrierbare Zufallsvariable $2|X|^k$ beschränkt ist und für $h \to 0$ gegen 0 konvergiert, folgt die gleichmäßige Stetigkeit von $\psi_X^{(k)}$ nun aus dem Satz über die majorisierte Konvergenz. □

Wir zeigen nun, dass die charakteristische Funktion ihren Namen verdient. Dazu benötigen wir das folgende Lemma:

16.4.6 Lemma. *Sei X eine Zufallsvariable und sei $(a,b] \subseteq \mathbb{R}$ ein nichtleeres Intervall mit $P_X[\{a,b\}] = 0$. Dann ist die Funktion $h : \mathbb{R} \to \mathbb{C}$ mit*

$$h(s) := \begin{cases} \dfrac{e^{-ias} - e^{-ibs}}{is} & \text{falls } s \neq 0 \\ b-a & \text{falls } s = 0 \end{cases}$$

stetig und beschränkt und es gilt

$$P_X[(a,b]] = \frac{1}{2\pi} \lim_{n\to\infty} \int_{[-n,n]} h(s)\,\psi_X(s)\,d\boldsymbol{\lambda}(s)$$

Beweis. Für alle $s \neq 0$ gilt nach Lemma 16.4.4

$$\begin{aligned}|h(s) - h(0)| &= \left|\frac{e^{-ias} - e^{-ibs}}{is} - (b-a)\right| \\ &= \left|\frac{e^{-ias} - 1 + ias}{is} - \frac{e^{-ibs} - 1 + ibs}{is}\right| \\ &\leq \frac{a^2+b^2}{2}\,|s|\end{aligned}$$

und

$$|h(s)| = \left|\frac{e^{-ias} - e^{-ibs}}{is}\right| = |e^{-ibs}|\left|\frac{e^{i(b-a)s} - 1}{is}\right| \leq b-a = h(0)$$

Aus der ersten Abschätzung folgt, dass h an der Stelle $s = 0$ und damit auf $\mathbb{R}$ stetig ist, und aus der zweiten Abschätzung folgt, dass h beschränkt ist.

Wir betrachten nun die Funktionen $g : \mathbb{R} \to \mathbb{C}$ und $f : \mathbb{R}^2 \to \mathbb{C}$ mit

$$\begin{aligned} g(s) &:= h(s)\, \psi_X(s) \\ f(x,s) &:= h(s)\, e^{isx} \end{aligned}$$

Dann sind g und f stetig und damit messbar, und es gilt

$$\begin{aligned} g(s) &= h(s)\, \psi_X(s) \\ &= h(s) \int_{\mathbb{R}} e^{isx}\, dP_X(x) \\ &= \int_{\mathbb{R}} h(s)\, e^{isx}\, dP_X(x) \\ &= \int_{\mathbb{R}} f(x,s)\, dP_X(x) \end{aligned}$$

Für die Funktionen $g_n : \mathbb{R} \to \mathbb{C}$ und $f_n : \mathbb{R}^2 \to \mathbb{C}$ mit

$$\begin{aligned} g_n(s) &:= \chi_{[-n,n]}(s)\, h(s)\, \psi_X(s) \\ f_n(x,s) &:= \chi_{[-n,n]}(s)\, h(s)\, e^{isx} \end{aligned}$$

gilt daher

$$g_n(s) = \int_{\mathbb{R}} f_n(x,s)\, dP_X(x)$$

Da f wegen $|f(x,s)| = |h(s)|$ beschränkt ist, ist f_n $P_X \otimes \boldsymbol{\lambda}$–integrierbar, und aus dem Satz von Fubini folgt nun, dass g_n $\boldsymbol{\lambda}$–integrierbar ist mit

$$\begin{aligned} \int_{\mathbb{R}} g_n(s)\, d\boldsymbol{\lambda}(s) &= \int_{\mathbb{R}} \int_{\mathbb{R}} f_n(x,s)\, dP_X(x)\, d\boldsymbol{\lambda}(s) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f_n(x,s)\, d\boldsymbol{\lambda}(s)\, dP_X(x) \end{aligned}$$

Wir untersuchen das innere Integral. Für alle $x \in \mathbb{R}$ gilt

$$\begin{aligned} \int_{\mathbb{R}} f_n(x,s)\, d\boldsymbol{\lambda}(s) &= \int_{[-n,n]\setminus\{0\}} \frac{e^{-ias} - e^{-ibs}}{is}\, e^{isx}\, d\boldsymbol{\lambda}(s) \\ &= \int_{[-n,n]\setminus\{0\}} \frac{e^{is(x-a)} - e^{is(x-b)}}{is}\, d\boldsymbol{\lambda}(s) \end{aligned}$$

Mit Hilfe der Gleichungen $e^{iz} = \cos(z) + i\sin(z)$ sowie $\cos(-z) = \cos(z)$ und $\sin(-z) = -\sin(z)$ und durch Übergang zum uneigentlichen Riemann–Integral über $(0,n]$ ergibt sich aus der letzten Gleichung

$$\begin{aligned} &\int_{\mathbb{R}} f_n(x,s)\, d\boldsymbol{\lambda}(s) \\ &= 2\left(\int_0^n \frac{\sin(s(x-a))}{s}\, ds - \int_0^n \frac{\sin(s(x-b))}{s}\, ds \right) \\ &= 2\left(\operatorname{sign}(x-a) \int_0^{n|x-a|} \frac{\sin(t)}{t}\, dt - \operatorname{sign}(x-b) \int_0^{n|x-b|} \frac{\sin(t)}{t}\, dt \right) \end{aligned}$$

Es gilt $\lim_{y\to\infty} \int_0^y \sin(t)/t\,dt = \pi/2$. Daher gibt es einerseits ein $C \in \mathbb{R}$ mit

$$\sup_{n\in\mathbb{N}} \left| \int_{\mathbb{R}} f_n(x,s)\,d\boldsymbol{\lambda}(s) \right| \leq C$$

für alle $x \in \mathbb{R}$, und zusammen mit der Voraussetzung $P_X[\{a,b\}] = 0$ ergibt sich andererseits

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n(x,s)\,d\boldsymbol{\lambda}(s) = 2\pi\chi_{(a,b]}(x)$$

P_X–fast sicher. Aus dem Satz über die majorisierte Konvergenz folgt nun

$$\begin{aligned}
\lim_{n\to\infty} \int_{[-n,n]} h(s)\,\psi_X(s)\,d\boldsymbol{\lambda}(s) &= \lim_{n\to\infty} \int_{\mathbb{R}} g_n(s)\,d\boldsymbol{\lambda}(s) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \int_{\mathbb{R}} f_n(x,s)\,d\boldsymbol{\lambda}(s)\,dP_X(x) \\
&= \int_{\mathbb{R}} \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x,s)\,d\boldsymbol{\lambda}(s)\,dP_X(x) \\
&= \int_{\mathbb{R}} 2\pi\chi_{(a,b]}(x)\,dP_X(x) \\
&= 2\pi\,P_X[(a,b]]
\end{aligned}$$

Daraus folgt die Behauptung. □

Der folgende Satz zeigt, dass die Verteilung einer Zufallsvariablen durch ihre charakteristische Funktion bestimmt ist:

16.4.7 Satz (Eindeutigkeitssatz). *Seien X und Y Zufallsvariable. Dann sind äquivalent:*
(a) *Es gilt $P_X = P_Y$.*
(b) *Es gilt $\psi_X = \psi_Y$.*

Beweis. Offensichtlich genügt es zu zeigen, dass (a) aus (b) folgt. Wir nehmen daher an, dass (b) gilt. Sei

$$\mathcal{J}_{X,Y} := \Big\{(a,b] \in 2^{\mathbb{R}} \;\Big|\; a \leq b \text{ und } P_X[\{a,b\}] = 0 = P_Y[\{a,b\}]\Big\}$$

Dann ist $\mathcal{J}_{X,Y}$ ein $\cap$–stabiles Mengensystem mit $\mathcal{J}_{X,Y} \subseteq \mathcal{J}(\mathbb{R})$ und nach Lemma 16.4.6 gilt für alle $J \in \mathcal{J}_{X,Y}$

$$P_X[J] = P_Y[J]$$

Da die Menge $\{z \in \mathbb{R} \mid \max\{P_X[\{z\}], P_Y[\{z\}]\} > 0\}$ abzählbar ist, gibt es zu jedem Intervall $(a,b] \subseteq \mathbb{R}$ mit $P_X[\{a\}] = 0 = P_Y[\{a\}]$ eine monoton fallende Folge $\{b_n\}_{n\in\mathbb{N}}$ mit $P_X[\{b_n\}] = 0 = P_Y[\{b_n\}]$ für alle $n \in \mathbb{N}$ und $\inf_{n\in\mathbb{N}} b_n = b$, und damit $(a, b] = \bigcap_{n\in\mathbb{N}}(a, b_n] \in \sigma(\mathcal{J}_{X,Y})$, und zu jedem Intervall $(a, b] \subseteq \mathbb{R}$ gibt es eine monoton fallende Folge $\{a_n\}_{n\in\mathbb{N}}$ mit $P_X[\{a_n\}] = 0 = P_Y[\{a_n\}]$ für alle $n \in \mathbb{N}$ und $\inf_{n\in\mathbb{N}} a_n = a$, und damit $(a, b] = \bigcup_{n\in\mathbb{N}}(a_n, b] \in \sigma(\mathcal{J}_{X,Y})$. Daher gilt $\mathcal{J}(\mathbb{R}) \subseteq \sigma(\mathcal{J}_{X,Y}) \subseteq \mathcal{B}(\mathbb{R})$, und damit $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{J}(\mathbb{R})) = \sigma(\mathcal{J}_{X,Y})$. Aus dem Eindeutigkeitssatz für Maße folgt nun, dass für alle $B \in \mathcal{B}(\mathbb{R})$

$$P_X[B] = P_Y[B]$$

gilt. Daher folgt (a) aus (b). □

Mit Hilfe der charakteristischen Funktion lässt sich auch die Verteilung einer Summe von unabhängigen Zufallsvariablen bestimmen:

16.4.8 Satz. *Sei $\mathbf{X} : \Omega \to \mathbb{R}^d$ ein Zufallsvektor mit unabhängigen Koordinaten. Dann gilt*

$$\psi_{\mathbf{1}'\mathbf{X}}(t) = \prod_{i=1}^{d} \psi_{X_i}(t)$$

Der Beweis von Satz 16.4.8 verläuft analog zum Beweis von Satz 16.1.8.

Aufgaben

16.4.A **Symmetrische Verteilungen:** Sei X eine reelle Zufallsvariable. Dann sind äquivalent:
(a) P_X ist symmetrisch.
(b) Es gilt $\psi_X = \psi_{-X}$
(c) Es gilt $\psi_X(t) = E[\cos(tX)]$.
(d) ψ_X ist reell.
Geben Sie Beispiele für symmetrische Verteilungen.

16.4.B **Normal–Verteilung:** Sei $P_X = \mathbf{N}(\mu, \sigma^2)$. Dann gilt

$$\psi_X(t) = \exp\left(i\mu t - \frac{\sigma^2}{2}\, t^2\right)$$

16.4.C Sei X eine reelle Zufallsvariable mit $M_{|X|}(t) < \infty$ für alle $t \in \mathbb{R}$. Dann besitzt X endliche Momente beliebiger Ordnung und es gilt

$$\psi_X(t) = \sum_{k=0}^{\infty} \frac{(it)^k}{k!}\, E[X^k]$$

16.4.D **Charakteristische Funktion eines Zufallsvektors:** Sei $\mathbf{X} : \Omega \to \mathbb{R}^d$ ein Zufallsvektor. Dann heißt die Funktion $\psi_{\mathbf{X}} : \mathbb{R}^d \to \mathbb{C}$ mit

$$\psi_{\mathbf{X}}(\mathbf{t}) := E[e^{i\mathbf{t}'\mathbf{X}}]$$

die *charakteristische Funktion* von $\mathbf{X}$.

(1) Verallgemeinern Sie den Eindeutigkeitssatz auf charakteristische Funktionen von Zufallsvektoren.
(2) Ein Zufallsvektor $\mathbf{X} : \Omega \to \mathbb{R}^d$ besitzt genau dann unabhängige Koordinaten, wenn für alle $\mathbf{t} \in \mathbb{R}^d$

$$\psi_{\mathbf{X}}(\mathbf{t}) = \prod_{j=1}^{d} \psi_{X_j}(t_j)$$

gilt.

17

Schwache Konvergenz und Zentraler Grenzwertsatz

Gegenstand dieses Kapitels ist ein Konvergenzbegriff für Verteilungen, der als schwache Konvergenz bezeichnet wird. Diese Begriffsbildung lässt sich auf zweifache Weise erklären: Zum einen wird die schwache Konvergenz einer Folge von Verteilungen durch die Konvergenz einer Familie von Integralen und damit durch die Konvergenz einer Familie von Funktionalen definiert; zum anderen stellt sich heraus, dass die schwache Konvergenz der Verteilungen einer Folge von Zufallsvariablen in der Tat schwächer als die stochastische Konvergenz und damit schwächer als jeder der bisher behandelten Konvergenzbegriffe ist.

Für eine Folge von Zufallsvariablen wird die schwache Konvergenz ihrer Verteilungen auch als Verteilungskonvergenz oder als Konvergenz in Verteilung bezeichnet. Aussagen über die Konvergenz in Verteilung bilden die Grundlage für die näherungsweise Berechnung von Verteilungen, die entweder nicht vollständig bestimmt oder nur mit hohem Aufwand berechnet werden können, durch Verteilungen, deren Eigenschaften wohlbekannt sind.

Das bekannteste Ergebnis über die Konvergenz in Verteilung ist der Zentrale Grenzwertsatz, der besagt, dass bei hinreichend großem Stichprobenumfang das Stichprobenmittel einer unabhängig und identisch verteilten Folge von Zufallsvariablen näherungsweise normal–verteilt ist. Dieses Ergebnis ist von zentraler Bedeutung in der Statistik.

Wir untersuchen zunächst die schwache Konvergenz einer Folge univariater Verteilungen, ihre Charakterisierung durch eine geeignete Art der Konvergenz ihrer Verteilungsfunktionen und durch die Konvergenz ihrer charakteristischen Funktionen, sowie den Zusammenhang zwischen der stochastischen Konvergenz einer Folge von Zufallsvariablen und der schwachen Konvergenz ihrer Verteilungen (Abschnitt 17.1). Wir wenden uns dann dem Begriff der Straffheit einer Familie von Verteilungen zu (Abschnitt 17.2) und beweisen schließlich die einfachste Form des Zentralen Grenzwertsatzes (Abschnitt 17.3).

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_17, © Springer-Verlag Berlin Heidelberg 2011

17.1 Schwache Konvergenz

Im gesamten Abschnitt sei (S, d) ein metrischer Raum und $\mathcal{T}_d$ die von der Metrik d erzeugte Topologie. Wir bezeichnen mit $\mathcal{B}(S)$ die zugehörige Borelsche σ–Algebra, mit

$$\mathcal{Q}(S)$$

die Familie aller Wahrscheinlichkeitsmaße $\mathcal{B}(S) \to [0, 1]$ und mit

$$C_b(S)$$

die Familie aller beschränkten stetigen Funktionen $S \to \mathbb{R}$. Dann ist jede Funktion $f \in C_b(S)$ bezüglich jedem Wahrscheinlichkeitsmaß $Q \in \mathcal{Q}(S)$ integrierbar.

17.1.1 Lemma. *Für $Q_1, Q_2 \in \mathcal{Q}(S)$ sind äquivalent:*
(a) *Es gilt $Q_1 = Q_2$.*
(b) *Für alle $f \in C_b(S)$ gilt $\int_S f\, dQ_1 = \int_S f\, dQ_2$.*

Beweis. Wir nehmen an, dass (b) gilt. Da die Borelsche σ–Algebra $\mathcal{B}(S)$ von der Topologie $\mathcal{T}_d$ erzeugt wird und jede Topologie $\cap$–stabil ist, genügt es zu zeigen, dass die Gleichung

$$Q_1[U] = Q_2[U]$$

für alle $U \in \mathcal{T}_d$ gilt. Sei also $U \in \mathcal{T}_d$ und $G := S \setminus U$. Sei ferner $g : S \to \mathbb{R}$ gegeben durch

$$g(x) := \inf\Big\{ d(x, z) \,\Big|\, z \in G \Big\}$$

Aus der Dreiecksungleichung erhält man $|g(x)-g(y)| \le d(x, y)$ und damit die Stetigkeit von g, und aus der Abgeschlossenheit von G ergibt sich

$$g^{-1}(\{0\}) = G$$

Für alle $m \in \mathbb{N}$ sei $f_m : S \to \mathbb{R}$ gegeben durch

$$f_m(x) := \min\{1, m\, g(x)\}$$

Dann ist $\{f_m\}_{m\in\mathbb{N}}$ eine monoton wachsende Folge in $C_b(S)$ und wegen $g^{-1}(\{0\}) = G$ gilt

$$\sup_{m\in\mathbb{N}} f_m = \chi_U$$

Aus dem Satz über die monotone Konvergenz folgt nun für $i \in \{1, 2\}$

$$Q_i[U] = \int_S \chi_U\, dQ_i = \int_S \left(\sup_{m\in\mathbb{N}} f_m \right) dQ_i = \sup_{m\in\mathbb{N}} \int_S f_m\, dQ_i$$

und damit $Q_1[U] = Q_2[U]$. Daher folgt (a) aus (b). □

Eine Folge $\{Q_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Q}(S)$ heißt *schwach konvergent*, wenn es ein $Q \in \mathcal{Q}(S)$ gibt derart, dass für alle $f \in C_b(S)$

$$\lim_{n\to\infty} \int_S f\,dQ_n = \int_S f\,dQ$$

gilt. Nach Lemma 17.1.1 ist der Limes einer schwach konvergenten Folge von Wahrscheinlichkeitsmaßen eindeutig bestimmt.

17.1.2 Beispiel. Sei $\{x_n\}_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen. Dann sind äquivalent:
(a) Die Folge $\{\boldsymbol{\delta}_{x_n}\}_{n\in\mathbb{N}}$ konvergiert schwach gegen $\boldsymbol{\delta}_{x_0}$.
(b) Die Folge $\{x_n\}_{n\in\mathbb{N}}$ konvergiert gegen x_0.
In der Tat: Wir nehmen zunächst an, dass (a) gilt. Für $\varepsilon \in (0,\infty)$ sei $f_\varepsilon : \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f_\varepsilon(x) := \begin{cases} 1 - \dfrac{|x-x_0|}{\varepsilon} & \text{falls } |x-x_0| \le \varepsilon \\ 0 & \text{sonst} \end{cases}$$

Dann gilt $f_\varepsilon \in C_b(\mathbb{R})$ und damit

$$\lim_{n\to\infty} f_\varepsilon(x_n) = \lim_{n\to\infty} \int_\mathbb{R} f_\varepsilon(x)\,d\boldsymbol{\delta}_{x_n}(x) = \int_\mathbb{R} f_\varepsilon(x)\,d\boldsymbol{\delta}_{x_0}(x) = f_\varepsilon(x_0) = 1$$

Da $\varepsilon \in (0,\infty)$ beliebig war, folgt daraus $\lim_{n\to\infty} x_n = x_0$.
Wir nehmen nun an, dass (b) gilt. Dann gilt für alle $f \in C_b(\mathbb{R})$

$$\lim_{n\to\infty} \int_\mathbb{R} f(x)\,d\boldsymbol{\delta}_{x_n}(x) = \lim_{n\to\infty} f(x_n) = f(x_0) = \int_\mathbb{R} f(x)\,d\boldsymbol{\delta}_{x_0}(x)$$

Daher konvergiert die Folge $\{\boldsymbol{\delta}_{x_n}\}_{n\in\mathbb{N}}$ schwach gegen $\boldsymbol{\delta}_{x_0}$.

Für $Q \in \mathcal{Q}(S)$ heißt eine Menge $B \in \mathcal{B}(S)$ *Q–Stetigkeitsmenge*, wenn

$$Q[B^\circ] = Q[B^\bullet]$$

gilt, wobei B° das Innere und $B^\bullet$ den Abschluss von B bezeichnet.

17.1.3 Satz (Portemanteau–Theorem). *Für $\{Q_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Q}(S)$ und $Q \in \mathcal{Q}(S)$ sind folgende Aussagen äquivalent:*
(a) *Die Folge $\{Q_n\}_{n\in\mathbb{N}}$ konvergiert schwach gegen Q.*
(b) *Für jede offene Menge U gilt $\liminf_{n\to\infty} Q_n[U] \ge Q[U]$.*
(c) *Für jede abgeschlossene Menge G gilt $\limsup_{n\to\infty} Q_n[G] \le Q[G]$.*
(d) *Für jede Q–Stetigkeitsmenge B gilt $\lim_{n\to\infty} Q_n[B] = Q[B]$.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Sei U offen und $G := S \setminus U$. Sei ferner $g : S \to \mathbb{R}$ gegeben durch

$$g(x) := \inf\Big\{ d(x,z) \;\Big|\; z \in G \Big\}$$

und für alle $m \in \mathbb{N}$ sei $f_m : S \to \mathbb{R}$ gegeben durch

$$f_m(x) := \min\{1, mg(x)\}$$

Dann ist $\{f_m\}_{m\in\mathbb{N}}$ eine monoton wachsende Folge in $C_b(S)$ und es gilt

$$\sup_{m\in\mathbb{N}} f_m = \chi_U$$

Nach Voraussetzung gilt für alle $m \in \mathbb{N}$

$$\liminf_{n\to\infty} Q_n[U] = \liminf_{n\to\infty} \int_S \chi_U \, dQ_n \geq \lim_{n\to\infty} \int_S f_m \, dQ_n = \int_S f_m \, dQ$$

und aus dem Satz über die monotone Konvergenz ergibt sich nun

$$\liminf_{n\to\infty} Q_n[U] \geq \sup_{m\in\mathbb{N}} \int_S f_m \, dQ = \int_S \sup_{m\in\mathbb{N}} f_m \, dQ = \int_S \chi_U \, dQ = Q[U]$$

Daher folgt (b) aus (a).
Sei nun U offen und G abgeschlossen mit $U + G = S$. Dann gilt

$$Q[U] + Q[G] = 1$$

und wegen $Q_n[U] + Q_n[G] = 1$ gilt

$$\liminf_{n\to\infty} Q_n[U] + \limsup_{n\to\infty} Q_n[G] = 1$$

Aus diesen Gleichungen ergibt sich sofort die Äquivalenz von (b) und (c).
Wir nehmen nun an, dass (b) und (c) gelten. Dann gilt für alle $B \in \mathcal{B}(S)$

$$\begin{aligned} Q[B^\circ] &\leq \liminf_{n\to\infty} Q_n[B^\circ] \\ &\leq \liminf_{n\to\infty} Q_n[B] \\ &\leq \limsup_{n\to\infty} Q_n[B] \\ &\leq \limsup_{n\to\infty} Q_n[B^\bullet] \\ &\leq Q[B^\bullet] \end{aligned}$$

sowie

$$\begin{aligned} Q[B^\circ] &\leq Q[B] \\ &\leq Q[B^\bullet] \end{aligned}$$

Für jede Q–Stetigkeitsmenge $B \in \mathcal{B}(S)$ folgt daraus

$$\lim_{n\to\infty} Q_n[B] = Q[B]$$

Daher folgt (d) aus (b) und (c).

Wir nehmen schließlich an, dass (d) gilt. Sei $f \in C_b(S)$. Da Q ein Wahrscheinlichkeitsmaß ist, ist die Menge $\{c \in \mathbb{R} \mid Q[\{f = c\}] > 0\}$ abzählbar. Sei $\varepsilon \in (0, \infty)$. Für alle $k \in \mathbb{Z}$ setzen wir

$$B_{\varepsilon,k} := \{\varepsilon k < f \leq \varepsilon(k+1)\}$$

Dann gilt $B^{\bullet}_{\varepsilon,k} \setminus B^{\circ}_{\varepsilon,k} \subseteq \{f = \varepsilon k\} \cup \{f = \varepsilon(k+1)\}$. Daher ist für alle außer abzählbar viele $\varepsilon \in (0, \infty)$ jede der Mengen $B_{\varepsilon,k}$ eine Q–Stetigkeitsmenge. Da f beschränkt ist, ist die Menge $\{k \in \mathbb{Z} \mid B_{\varepsilon,k} \neq \emptyset\}$ endlich. Daher gilt für alle außer abzählbar viele $\varepsilon \in (0, \infty)$

$$\begin{aligned}
\int_S f \, dQ - \varepsilon &= \int_S (f - \varepsilon) \, dQ \\
&\leq \int_S \left(\sum_{k \in \mathbb{Z}} \varepsilon k \chi_{B_{\varepsilon,k}} \right) dQ \\
&= \sum_{k \in \mathbb{Z}} \varepsilon k \, Q[B_{\varepsilon,k}] \\
&= \sum_{k \in \mathbb{Z}} \varepsilon k \lim_{n \to \infty} Q_n[B_{\varepsilon,k}] \\
&= \lim_{n \to \infty} \sum_{k \in \mathbb{Z}} \varepsilon k \, Q_n[B_{\varepsilon,k}] \\
&= \lim_{n \to \infty} \int_S \left(\sum_{k \in \mathbb{Z}} \varepsilon k \chi_{B_{\varepsilon,k}} \right) dQ_n \\
&\leq \liminf_{n \to \infty} \int_S f \, Q_n
\end{aligned}$$

und analog erhält man

$$\limsup_{n \to \infty} \int_S f \, Q_n \leq \int_S f \, dQ + \varepsilon$$

Daher gilt

$$\int_S f \, dQ - \varepsilon \leq \liminf_{n \to \infty} \int_S f \, Q_n \leq \limsup_{n \to \infty} \int_S f \, Q_n \leq \int_S f \, dQ + \varepsilon$$

und daraus ergibt sich

$$\lim_{n \to \infty} \int_S f \, dQ_n = \int_S f \, dQ$$

Daher folgt (a) aus (d). □

Wir untersuchen nun die schwache Konvergenz einer Folge von Verteilungen auf $\mathcal{B}(\mathbb{R})$.

Eine Folge $\{F_n\}_{n\in\mathbb{N}}$ von Verteilungsfunktionen $\mathbb{R} \to [0,1]$ heißt *schwach konvergent*, wenn es eine Verteilungsfunktion $F : \mathbb{R} \to [0,1]$ gibt derart, dass für jede Stetigkeitsstelle $x \in \mathbb{R}$ von F

$$\lim_{n\to\infty} F_n(x) = F(x)$$

gilt. Der folgende Satz zeigt, dass die Definitionen der schwachen Konvergenz für Verteilungen und Verteilungsfunktionen miteinander im Einklang stehen:

17.1.4 Satz (Helly/Bray). *Sei $\{Q_n\}_{n\in\mathbb{N}}$ eine Folge von Verteilungen und $\{F_n\}_{n\in\mathbb{N}}$ die Folge der zugehörigen Verteilungsfunktionen und sei Q eine Verteilung und F die zugehörige Verteilungsfunktion. Dann sind äquivalent*:
(a) *Die Folge $\{Q_n\}_{n\in\mathbb{N}}$ konvergiert schwach gegen Q.*
(b) *Die Folge $\{F_n\}_{n\in\mathbb{N}}$ konvergiert schwach gegen F.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Sei $x \in \mathbb{R}$ eine Stetigkeitsstelle von F. Dann gilt $Q[\{x\}] = 0$. Daher ist die Menge $(-\infty, x]$ eine Q–Stetigkeitsmenge und es gilt

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} Q_n[(-\infty, x]] = Q[(-\infty, x]] = F(x)$$

Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Wir betrachten zunächst das Mengensystem

$$\mathcal{J}_Q := \Big\{(a,b] \in \mathcal{J}(\mathbb{R}) \,\Big|\, Q[\{a\}] = 0 = Q[\{b\}]\Big\}$$

Für $(a,b] \in \mathcal{J}_Q$ gilt nach Voraussetzung

$$\begin{aligned} Q[(a,b]] &= F(b) - F(a) \\ &= \lim_{n\to\infty} F_n(b) - \lim_{n\to\infty} F_n(a) \\ &= \lim_{n\to\infty} \Big(F_n(b) - F_n(a)\Big) \\ &= \lim_{n\to\infty} Q_n[(a,b]] \end{aligned}$$

Da $\mathcal{J}_Q$ $\cap$–stabil ist, ergibt sich für jede endliche Familie $\{J_i\}_{i\in H} \subseteq \mathcal{J}_Q$ aus der Einschluss–Ausschluss–Formel und der letzten Gleichung

$$Q\left[\bigcup_{i\in H} J_i\right] = \lim_{n\to\infty} Q_n\left[\bigcup_{i\in H} J_i\right]$$

Sei U offen. Nach Beispiel 1.1.8 ist U die Vereinigung einer abzählbaren Familie von offenen Intervallen, und jedes offene Intervall ist die Vereinigung einer abzählbaren Familie von halboffenen Intervallen. Da F nur abzählbar viele Sprungstellen besitzt, ist jedes offene Intervall sogar die Vereinigung einer abzählbaren Familie von Intervallen in $\mathcal{J}_Q$, und damit ist auch U die Vereinigung einer abzählbaren Familie von Intervallen in $\mathcal{J}_Q$. Daher gibt es zu jedem $\varepsilon \in (0,\infty)$ eine endliche Familie $\{J_i\}_{i\in H(\varepsilon)} \subseteq \mathcal{J}_Q$ mit $\bigcup_{i\in H(\varepsilon)} J_i \subseteq U$ und

$$Q[U] \le Q\left[\bigcup_{i\in H(\varepsilon)} J_i\right] + \varepsilon$$

und wir erhalten

$$Q[U] - \varepsilon \le Q\left[\bigcup_{i\in H(\varepsilon)} J_i\right] = \lim_{n\to\infty} Q_n\left[\bigcup_{i\in H(\varepsilon)} J_i\right] \le \liminf_{n\to\infty} Q_n[U]$$

Daher gilt

$$Q[U] \le \liminf_{n\to\infty} Q_n[U]$$

Aus dem Portemanteau–Theorem folgt nun, dass die Folge $\{Q_n\}_{n\in\mathbb{N}}$ schwach gegen Q konvergiert. Daher folgt (a) aus (b). □

Eine Folge $\{X_n\}_{n\in\mathbb{N}}$ von reellen Zufallsvariablen heißt *verteilungskonvergent* oder *konvergent in Verteilung*, wenn es eine reelle Zufallsvariable X gibt derart, dass die Folge der Verteilungen $\{P_{X_n}\}_{n\in\mathbb{N}}$ schwach gegen die Verteilung P_X konvergiert; nach dem Satz von Helly/Bray ist diese Bedingung gleichwertig damit, dass die Folge der Verteilungsfunktionen $\{F_{X_n}\}_{n\in\mathbb{N}}$ schwach gegen die Verteilungsfunktion F_X konvergiert.

17.1.5 Beispiel (Poisson–Approximation). Sei X eine Zufallsvariable mit $P_X = \mathbf{P}(\alpha)$ und sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen mit $P_{X_n} = \mathbf{B}(n, \alpha/n)$. Dann konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ in Verteilung gegen X.
In der Tat: Für alle $k \in \mathbb{N}_0$ gilt

$$\lim_{n\to\infty} P[\{X_n = k\}] = \lim_{n\to\infty} \binom{n}{k}\left(\frac{\alpha}{n}\right)^k \left(1-\frac{\alpha}{n}\right)^{n-k} = e^{-\alpha}\,\frac{\alpha^k}{k!}$$

Durch Summation erhält man daraus für alle $x \in \mathbb{R}$

$$\lim_{n\to\infty} P[\{X_n \le x\}] = P[\{X \le x\}]$$

Daraus folgt die Behauptung.

Für eine Folge von reellen Zufallsvariablen vergleichen wir nun die Konvergenz in Verteilung mit der stochastischen Konvergenz. Wir benötigen das folgende Lemma:

17.1.6 Lemma. *Seien Y und Z reelle Zufallsvariable. Dann gilt für alle $y, z \in \mathbb{R}$*

$$P[\{Y \le y\}] \le P[\{Z \le z\}] + P[\{|Z-Y| > z-y\}]$$

Beweis. Es gilt

$$\begin{aligned}\{Y \le y\} &= \{Y \le y\} \cap \{Z \le z\} + \{Y \le y\} \cap \{Z > z\} \\ &\subseteq \{Z \le z\} \cup \{Z - Y > z - y\} \\ &\subseteq \{Z \le z\} \cup \{|Z - Y| > z - y\}\end{aligned}$$

Daraus folgt die Behauptung. □

17.1.7 Satz. *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen und sei X eine reelle Zufallsvariable. Wenn die Folge $\{X_n\}_{n\in\mathbb{N}}$ stochastisch gegen X konvergiert, dann konvergiert sie auch in Verteilung gegen X.*

Beweis. Nach Lemma 17.1.6 gilt für alle $x \in \mathbb{R}$ und $\varepsilon \in (0, \infty)$

$$F_X(x - \varepsilon) \le F_{X_n}(x) + P[\{|X_n - X| > \varepsilon\}]$$

und

$$F_{X_n}(x) \le F_X(x + \varepsilon) + P[\{|X - X_n| > \varepsilon\}]$$

und aus der stochastischen Konvergenz der Folge $\{X_n\}_{n\in\mathbb{N}}$ gegen X folgt nun

$$F_X(x - \varepsilon) \le \liminf_{n\to\infty} F_{X_n}(x) \le \limsup_{n\to\infty} F_{X_n}(x) \le F_X(x + \varepsilon)$$

Für jede Stetigkeitsstelle x von F_X gilt daher

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

Daher konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ in Verteilung gegen X. □

Das folgende Beispiel zeigt, dass die Umkehrung der Implikation von Satz 17.1.7 falsch ist:

17.1.8 Beispiel. Sei X eine reelle Zufallsvariable mit $P_{(X+1)/2} = \mathbf{B}(\frac{1}{2})$ und für alle $n \in \mathbb{N}$ sei $X_n := (-1)^n X$. Dann gilt für alle $n \in \mathbb{N}$

$$P_{X_n} = P_X$$

Daher konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ in Verteilung gegen X. Andererseits ist die Folge $\{X_n\}_{n\in\mathbb{N}}$ nicht stochastisch konvergent.

Im Fall der Konvergenz gegen eine konstante und damit Dirac–verteilte Zufallsvariable hingegen ist die Konvergenz in Verteilung mit der stochastischen Konvergenz äquivalent:

17.1.9 Satz. *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen und sei $c \in \mathbb{R}$. Dann sind äquivalent:*
(a) *$\{X_n\}_{n\in\mathbb{N}}$ konvergiert stochastisch gegen c.*
(b) *$\{X_n\}_{n\in\mathbb{N}}$ konvergiert in Verteilung gegen c.*

Beweis. Nach Satz 17.1.7 folgt (b) aus (a). Wir nehmen nun an, dass (b) gilt. Sei $\varepsilon \in (0,\infty)$. Dann gilt

$$\begin{aligned} P[\{|X_n - c| \le \varepsilon\}] &= P[\{c-\varepsilon \le X_n \le c+\varepsilon\}] \\ &\ge P[\{c-\varepsilon < X_n \le c+\varepsilon\}] \\ &= P[\{X_n \le c+\varepsilon\}] - P[\{X_n \le c-\varepsilon\}] \end{aligned}$$

Da $c+\varepsilon$ und $c-\varepsilon$ Stetigkeitsstellen der Verteilungsfunktion F_c sind, erhalten wir

$$\begin{aligned} \liminf_{n\to\infty} P[\{|X_n - c| \le \varepsilon\}] &\ge \lim_{n\to\infty} F_{X_n}(c+\varepsilon) - \lim_{n\to\infty} F_{X_n}(c-\varepsilon) \\ &= F_c(c+\varepsilon) - F_c(c-\varepsilon) \\ &= 1 - 0 \\ &= 1 \end{aligned}$$

und damit

$$\lim_{n\to\infty} P[\{|X_n - c| \le \varepsilon\}] = 1$$

Daher folgt (a) aus (b). □

Das folgende Beispiel illustriert die Aussage des letzten Satzes:

17.1.10 Beispiel (Wandernde Türme). Sei $(\Omega, \mathcal{F}, P) := ((0,1], \mathcal{B}(0,1], \boldsymbol{\lambda}|_{\mathcal{B}(0,1]})$. Für $m \in \mathbb{N}$ und $k \in \{1, \dots, 2^m\}$ sei

$$B_{m,k} := ((k-1)\,2^{-m}, k\,2^{-m}]$$

und

$$X_{2^m+k-2} := \chi_{B_{m,k}}$$

Daher konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ stochastisch gegen 0; vgl. Beispiel 14.2.2. Außerdem gilt für alle $f \in C_b(\mathbb{R})$

$$\int_{\mathbb{R}} f(x)\, dP_{X_{2^m+k-2}}(x) = \int_{\Omega} f(X_{2^m+k-2}(\omega))\, dP(\omega) = f(0)\,(1-2^{-m}) + f(1)\,2^{-m}$$

und damit

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(x)\, dP_{X_n}(x) = f(0) = \int_{\mathbb{R}} f(x)\, d\boldsymbol{\delta}_0(x)$$

Daher konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ in Verteilung gegen 0.

Aufgabe

17.1.A Sei (S,d) ein metrischer Raum und sei $Q \in \mathcal{Q}(S)$. Dann ist das System der Q–Stetigkeitsmengen eine Algebra.

17.2 Straffheit

In diesem Abschnitt zeigen wir, dass die schwache Konvergenz einer Folge von Verteilungen durch die Konvergenz der Folge ihrer charakteristischen Funktionen charakterisiert werden kann; dabei ist die charakteristische Funktion $\psi : \mathbb{R} \to \mathbb{C}$ einer Verteilung $Q : \mathcal{B}(\mathbb{R}) \to [0,1]$ in Übereinstimmung mit der Darstellung der charakteristischen Funktion einer Zufallsvariablen durch

$$\psi(t) := \int_{\mathbb{R}} e^{itx} \, dQ(x)$$

definiert.

Sei (S,d) ein metrischer Raum und $\mathcal{B}(S)$ die zugehörige Borelsche σ–Algebra. Eine Familie $\mathcal{Q} \subseteq \mathcal{Q}(S)$ von Wahrscheinlichkeitsmaßen heißt *straff*, wenn es zu jedem $\varepsilon \in (0,\infty)$ eine kompakte Menge $K \in 2^S$ gibt derart, dass für alle $Q \in \mathcal{Q}$

$$Q[K] \geq 1 - \varepsilon$$

gilt.

17.2.1 Lemma. *Sei $\mathcal{Q}$ eine Familie von Verteilungen. Dann sind äquivalent*:
(a) *$\mathcal{Q}$ ist straff.*
(b) *Zu jedem $\varepsilon \in (0,\infty)$ gibt es ein Intervall $[a,b]$ derart, dass für alle $Q \in \mathcal{Q}$*

$$Q[[a,b]] \geq 1 - \varepsilon$$

gilt.

Beweis. Nach dem Satz von Heine/Borel ist jede kompakte Menge beschränkt und abgeschlossen, und damit eine Teilmenge eines abgeschlossenen Intervalls. Andererseits ist jedes abgeschlossene Intervall kompakt. □

17.2.2 Beispiel. Sei $\{x_n\}_{n\in\mathbb{N}}$ eine Folge reeller Zahlen. Dann sind äquivalent:
(a) Die Folge $\{\boldsymbol{\delta}_{x_n}\}_{n\in\mathbb{N}}$ ist straff.
(b) Die Folge $\{x_n\}_{n\in\mathbb{N}}$ ist beschränkt.

Im Hinblick auf die Beispiele 17.1.2 und 17.2.2 ist das folgende Ergebnis nicht überraschend:

17.2.3 Satz. *Jede schwach konvergente Folge von Verteilungen ist straff.*

Beweis. Sei $\{Q_n\}_{n\in\mathbb{N}}$ eine Folge von Verteilungen, die schwach gegen eine Verteilung Q konvergiert. Sei $\varepsilon \in (0,\infty)$. Dann gibt es ein $c \in (0,\infty)$ mit

$$Q[[-c,c]] \geq 1 - \varepsilon$$

und die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) := \begin{cases} 1 & \text{falls } |x| \le c \\ 2 - |x/c| & \text{falls } c < |x| \le 2c \\ 0 & \text{falls } 2c < |x| \end{cases}$$

ist stetig und beschränkt. Es gilt $\chi_{[-c,c]} \le f \le \chi_{[-2c,2c]}$ und damit

$$\liminf_{n\to\infty} \int_{\mathbb{R}} \chi_{[-2c,2c]}\, dQ_n \ge \lim_{n\to\infty} \int_{\mathbb{R}} f\, dQ_n = \int_{\mathbb{R}} f\, dQ \ge \int_{\mathbb{R}} \chi_{[-c,c]}\, dQ \ge 1 - \varepsilon$$

Daher gilt

$$Q_n[[-2c, 2c]] \ge 1 - \varepsilon$$

für alle außer endlich viele $n \in \mathbb{N}$, und durch Vergrößerung von c kann man erreichen, dass diese Ungleichung sogar für alle $n \in \mathbb{N}$ gilt. Daher ist die Folge $\{Q_n\}_{n\in\mathbb{N}}$ straff. □

Der folgende Satz charakterisiert die Straffheit einer Folge von Verteilungen durch die Existenz schwach konvergenter Teilfolgen (mit möglicherweise unterschiedlichen Limites):

17.2.4 Satz (Helly). *Sei $\{Q_n\}_{n\in\mathbb{N}}$ ein Folge von Verteilungen. Dann sind äquivalent:*
(a) *Die Folge $\{Q_n\}_{n\in\mathbb{N}}$ ist straff.*
(b) *Jede Teilfolge der Folge $\{Q_n\}_{n\in\mathbb{N}}$ besitzt eine schwach konvergente Teilfolge.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Sei $\{F_n\}_{n\in\mathbb{N}}$ die Folge der zugehörigen Verteilungsfunktionen und sei $\{q_i\}_{i\in\mathbb{N}}$ eine Abzählung der Menge $\mathbb{Q}$ der rationalen Zahlen. Für alle $i \in \mathbb{N}$ ist die Folge $\{F_n(q_i)\}_{n\in\mathbb{N}}$ beschränkt und enthält daher eine konvergente Teilfolge. Durch sukzessive Verdünnung und Anwendung des Diagonalprinzips erhält man eine streng monoton wachsende Folge $\{n_k\}_{k\in\mathbb{N}}$ derart, dass für alle $i \in \mathbb{N}$ die Folge $\{F_{n_k}(q_i)\}_{k\in\mathbb{N}}$ konvergent ist. Sei nun $G : \mathbb{Q} \to \mathbb{R}$ gegeben durch

$$G(q) := \lim_{k\to\infty} F_{n_k}(q)$$

und sei $F : \mathbb{R} \to \mathbb{R}$ gegeben durch

$$F(x) := \inf\Big\{G(q) \,\Big|\, q \in \mathbb{Q} \cap (x, \infty)\Big\}$$

Dann ist F monoton wachsend und für alle $x \in \mathbb{R}$ gilt $F(x) \in [0, 1]$.
Sei $x \in \mathbb{R}$ und $\varepsilon \in (0, \infty)$. Dann gibt es ein $q \in \mathbb{Q} \cap (x, \infty)$ mit $G(q) \le F(x) + \varepsilon$ und für alle $y \in (x, q)$ gilt $F(x) \le F(y) \le G(q) \le F(x) + \varepsilon$. Daher ist F rechtsseitig stetig.

Sei nun x eine Stetigkeitsstelle von F und $\varepsilon \in (0,\infty)$. Dann gibt es $p,q \in \mathbb{Q}$ mit $p < x < q$ sowie $F(x) \le G(p) + \varepsilon$ und $G(q) \le F(x) + \varepsilon$, und man erhält

$$F(x) - \varepsilon \le G(p) = \lim_{k\to\infty} F_{n_k}(p) \le \liminf_{k\to\infty} F_{n_k}(x)$$

und

$$\limsup_{k\to\infty} F_{n_k}(x) \le \lim_{k\to\infty} F_{n_k}(q) = G(q) \le F(x) + \varepsilon$$

Für jede Stetigkeitsstelle x von F gilt daher

$$\lim_{k\to\infty} F_{n_k}(x) = F(x)$$

Es bleibt zu zeigen, dass F eine Verteilungsfunktion ist. Sei $\varepsilon \in (0,\infty)$. Da die Folge $\{Q_n\}_{n\in\mathbb{N}}$ straff ist, gibt es ein abgeschlossenes Intervall J derart, dass für alle $n \in \mathbb{N}$

$$Q_n[J] \ge 1 - \varepsilon$$

gilt. Dann gibt es aber auch ein halboffenes Intervall $(a,b]$ mit $Q_n[(a,b]] \ge 1-\varepsilon$ für alle $n \in \mathbb{N}$ und man kann dieses Intervall so wählen, dass seine Endpunkte Stetigkeitsstellen von F sind. Nach dem bisher Gezeigten gilt dann

$$\begin{aligned}
F(b) - F(a) &= \lim_{k\to\infty} F_{n_k}(b) - \lim_{k\to\infty} F_{n_k}(a) \\
&= \lim_{k\to\infty} \bigl(F_{n_k}(b) - F_{n_k}(a)\bigr) \\
&= \lim_{k\to\infty} Q_{n_k}[(a,b]] \\
&\ge 1 - \varepsilon
\end{aligned}$$

und damit $F(a) \le \varepsilon$ und $F(b) \ge 1 - \varepsilon$. Daher gibt es zu jedem $\varepsilon \in (0,\infty)$ ein $a \in \mathbb{R}$ mit $F(a) \le \varepsilon$ und ein $b \in \mathbb{R}$ mit $F(b) \ge 1-\varepsilon$, und aus der Monotonie von F folgt nun, dass $\lim_{x\to-\infty} F(x) = 0$ und $\lim_{x\to\infty} F(x) = 1$ gilt. Daher ist F eine Verteilungsfunktion. Da die Folge $\{F_{n_k}\}_{k\in\mathbb{N}}$ schwach gegen F konvergiert, konvergiert nach dem Satz von Helly/Bray die Folge $\{Q_{n_k}\}_{k\in\mathbb{N}}$ schwach gegen die zu F gehörige Verteilung Q. Daher folgt (b) aus (a).

Wir nehmen nun an, dass (b) gilt. Wir nehmen des weiteren an, dass die Folge $\{Q_n\}_{n\in\mathbb{N}}$ nicht straff ist. Dann gibt es ein $\varepsilon \in (0,\infty)$ derart, dass für alle $a,b \in \mathbb{R}$ und ein $n \in \mathbb{N}$

$$Q_n[[a,b]] < 1 - \varepsilon$$

und damit

$$Q_n[(a,b]] < 1 - \varepsilon$$

gilt. Daher gibt es für alle $k \in \mathbb{N}$ ein $n_k \in \mathbb{N}$ mit

$$Q_{n_k}[(-k,k]] < 1-\varepsilon$$

Da es kein $n \in \mathbb{N}$ gibt mit $Q_n[(-k,k]] < 1-\varepsilon$ für alle $k \in \mathbb{N}$, können wir annehmen, dass die Folge $\{n_k\}_{k\in\mathbb{N}}$ streng monoton wachsend ist. Daher ist $\{Q_{n_k}\}_{k\in\mathbb{N}}$ eine Teilfolge der Folge $\{Q_n\}_{n\in\mathbb{N}}$ und nach Voraussetzung gibt es eine Teilfolge $\{Q_{n_{k_l}}\}_{l\in\mathbb{N}}$ der Folge $\{Q_{n_k}\}_{k\in\mathbb{N}}$, die schwach gegen eine Verteilung Q konvergiert. Sei nun $(a,b]$ ein Intervall mit $Q[\{a\}] = 0 = Q[\{b\}]$ und

$$Q[(a,b]] \geq 1-\varepsilon$$

Für alle hinreichend großen $l \in \mathbb{N}$ gilt $(a,b] \subseteq (-k_l, k_l]$ und damit

$$Q[(a,b]] = \lim_{l\to\infty} Q_{n_{k_l}}[(a,b]] < 1-\varepsilon$$

Dies ist ein Widerspruch. Daher folgt (a) aus (b). □

Für die schwache Konvergenz gilt das Teilfolgenprinzip:

17.2.5 Lemma (Teilfolgenprinzip). *Sei $\{Q_n\}_{n\in\mathbb{N}}$ eine Folge von Verteilungen und sei Q eine Verteilung. Dann sind äquivalent:*
(a) *Die Folge $\{Q_n\}_{n\in\mathbb{N}}$ konvergiert schwach gegen Q.*
(b) *Jede Teilfolge der Folge $\{Q_n\}_{n\in\mathbb{N}}$ besitzt eine Teilfolge, die schwach gegen Q konvergiert.*

Beweis. Wir nehmen an, dass (b) gilt. Wir nehmen des weiteren an, dass die Folge $\{Q_n\}_{n\in\mathbb{N}}$ nicht schwach gegen Q konvergiert. Dann gibt es eine beschränkte stetige Funktion $f : \mathbb{R} \to \mathbb{R}$ derart, dass die Folge $\{\int_{\mathbb{R}} f\, dQ_n\}_{n\in\mathbb{N}}$ nicht gegen $\int_{\mathbb{R}} f\, dQ$ konvergiert, und daraus folgt, dass es zu jedem $\varepsilon \in (0,\infty)$ eine Teilfolge $\{Q_{n_k}\}_{k\in\mathbb{N}}$ gibt mit $\left|\int_{\mathbb{R}} f\, dQ_{n_k} - \int_{\mathbb{R}} f\, dQ\right| > \varepsilon$ für alle $k \in \mathbb{N}$. Dies ist ein Widerspruch zur Voraussetzung. Damit ist gezeigt, dass die Folge $\{Q_n\}_{n\in\mathbb{N}}$ schwach gegen Q konvergiert. Daher folgt (a) aus (b).
Die umgekehrte Implikation ist klar. □

17.2.6 Lemma. *Sei $\{Q_n\}_{n\in\mathbb{N}}$ eine Folge von Verteilungen und $\{\psi_n\}_{n\in\mathbb{N}}$ die Folge der zugehörigen charakteristischen Funktionen. Konvergiert die Folge $\{Q_n\}_{n\in\mathbb{N}}$ schwach gegen eine Verteilung Q mit charakteristischer Funktion ψ, so konvergiert die Folge $\{\psi_n\}_{n\in\mathbb{N}}$ punktweise gegen ψ.*

Beweis. Sowohl der Realteil als auch der Imaginärteil der Exponentialfunktion $\mathbb{R} \to \mathbb{C} : x \mapsto e^{itx}$ mit $t \in \mathbb{R}$ ist stetig und beschränkt. □

17.2.7 Satz. *Sei $\{Q_n\}_{n\in\mathbb{N}}$ eine Folge von Verteilungen und $\{\psi_n\}_{n\in\mathbb{N}}$ die Folge der zugehörigen charakteristischen Funktionen. Ist $\{Q_n\}_{n\in\mathbb{N}}$ straff und gibt es eine Funktion $\psi : \mathbb{R} \to \mathbb{C}$ derart, dass die Folge $\{\psi_n\}_{n\in\mathbb{N}}$ punktweise gegen ψ konvergiert, so ist ψ die charakteristische Funktion einer Verteilung Q und die Folge $\{Q_n\}_{n\in\mathbb{N}}$ konvergiert schwach gegen Q.*

Beweis. Nach dem Satz von Helly besitzt jede Teilfolge von $\{Q_n\}_{n\in\mathbb{N}}$ eine schwach konvergente Teilfolge. Wegen $\lim_{n\to\infty}\psi_n(t) = \psi(t)$ folgt aus Lemma 17.2.6, dass ψ die charakteristische Funktion jeder der so entstehenden Grenzverteilungen ist, und aus dem Eindeutigkeitssatz 16.4.7 folgt nun, dass alle Grenzverteilungen identisch sind. Aus Lemma 17.2.5 folgt schließlich die schwache Konvergenz der Folge $\{Q_n\}_{n\in\mathbb{N}}$ gegen eine Verteilung mit der charakteristischen Funktion ψ. □

Wir können nun die schwache Konvergenz einer Folge von Verteilungen durch die Konvergenz der Folge ihrer charakteristischen Funktionen charakterisieren:

17.2.8 Satz (Stetigkeitssatz). *Sei $\{Q_n\}_{n\in\mathbb{N}}$ eine Folge von Verteilungen und $\{\psi_n\}_{n\in\mathbb{N}}$ die Folge der zugehörigen charakteristischen Funktionen und sei Q eine Verteilung und ψ die zugehörige charakteristische Funktion. Dann sind äquivalent*:
(a) *Die Folge $\{Q_n\}_{n\in\mathbb{N}}$ konvergiert schwach gegen Q.*
(b) *Die Folge $\{\psi_n\}_{n\in\mathbb{N}}$ konvergiert punktweise gegen ψ.*

Beweis. Nach Lemma 17.2.6 folgt (b) aus (a). Wir nehmen nun an, dass (b) gilt. Nach Satz 17.2.7 genügt es zu zeigen, dass die Folge $\{Q_n\}_{n\in\mathbb{N}}$ straff ist. Sei $\varepsilon \in (0,\infty)$. Da ψ stetig ist mit $\psi(0) = 1$ gibt es ein $c \in (0,\infty)$ mit

$$\frac{1}{c}\int_{[-c,c]} |1-\psi(t)|\, d\boldsymbol{\lambda}(t) \le \varepsilon$$

Andererseits ergibt sich für alle $n \in \mathbb{N}$ aus dem Satz von Fubini

$$\begin{aligned}
\frac{1}{c}\int_{[-c,c]} (1-\psi_n(t))\, d\boldsymbol{\lambda}(t) &= \frac{1}{c}\int_{[-c,c]}\int_{\mathbb{R}} (1-e^{itx})\, dQ_n(x)\, d\boldsymbol{\lambda}(t)\\
&= \frac{1}{c}\int_{\mathbb{R}}\int_{[-c,c]} (1-e^{itx})\, d\boldsymbol{\lambda}(t)\, dQ_n(x)\\
&= \int_{\mathbb{R}} 2\left(1-\frac{\sin(cx)}{cx}\right) dQ_n(x)\\
&\ge \int_{\mathbb{R}\setminus[-2/c,2/c]} 2\left(1-\frac{1}{|cx|}\right) dQ_n(x)\\
&\ge Q_n[\mathbb{R}\setminus[-2/c,2/c]]
\end{aligned}$$

Aus der Voraussetzung und dem Satz über die majorisierte Konvergenz folgt nun

$$\limsup_{n\to\infty} Q_n[\mathbb{R}\setminus[-2/c,2/c]] \le \varepsilon$$

und damit

$$Q_n[[-2/c,2/c]] \ge 1-\varepsilon$$

für alle außer endlich viele $n \in \mathbb{N}$. Daher ist die Folge $\{Q_n\}_{n\in\mathbb{N}}$ straff. □

Aufgabe

17.2.A **Poisson–Verteilung:** Sei $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängige Folge von reellen Zufallsvariablen mit $P_{X_k} = \mathbf{P}(\alpha_k)$ für alle $k \in \mathbb{N}$ sowie $\alpha := \sum_{k=1}^{\infty} \alpha_k < \infty$. Dann konvergiert die Reihe

$$\sum_{k=1}^{\infty} X_k$$

in Verteilung gegen eine reelle Zufallsvariable X mit $P_X = \mathbf{P}(\alpha)$.

17.3 Zentraler Grenzwertsatz

Ist $\{X_k\}_{k\in\mathbb{N}}$ eine unkorrelierte Folge in L^2 mit $E[X_k] = \mu$ und $\mathrm{var}\,[X_k] = \sigma^2$ für alle $k \in \mathbb{N}$, so konvergiert die Folge der Stichprobenmittel

$$\frac{1}{n}\sum_{k=1}^{n} X_k$$

nach dem schwachen Gesetz der Großen Zahlen stochastisch gegen μ und nach Satz 17.1.7 konvergiert sie auch in Verteilung gegen μ. Die Verteilung von μ ist aber degeneriert und eignet sich daher im Fall $\sigma^2 > 0$ nicht zur Approximation der Verteilung der Stichprobenmittel.

Unter der stärkeren Voraussetzung, dass $\{X_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge in L^2 mit $E[X] = \mu$ ist, konvergiert die Folge der Stichprobenmittel nach dem starken Gesetz der Großen Zahlen sogar fast sicher gegen μ und in diesem Fall lässt sich unter der Annahme $\mathrm{var}\,[X] = \sigma^2 > 0$ auch das Problem der Approximation der Stichprobenmittel lösen, indem man nicht die Folge der Stichprobenmittel selbst, sondern die Folge der standardisierten Stichprobenmittel

$$\frac{\frac{1}{n}\sum_{k=1}^{n} X_k - E\bigl[\frac{1}{n}\sum_{k=1}^{n} X_k\bigr]}{\sqrt{\mathrm{var}\,\bigl[\frac{1}{n}\sum_{k=1}^{n} X_k\bigr]}}$$

betrachtet. Es gilt

$$\frac{\frac{1}{n}\sum_{k=1}^{n} X_k - E\bigl[\frac{1}{n}\sum_{k=1}^{n} X_k\bigr]}{\sqrt{\mathrm{var}\,\bigl[\frac{1}{n}\sum_{k=1}^{n} X_k\bigr]}} = \frac{\frac{1}{n}\sum_{k=1}^{n} X_k - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \sum_{k=1}^{n} \frac{X_k - \mu}{\sigma\sqrt{n}}$$

Im folgenden sei Φ die Verteilungsfunktion der Standardnormal–Verteilung.

17.3.1 Satz (Zentraler Grenzwertsatz). *Sei* $\{X_n\}_{n\in\mathbb{N}}$ *eine unabhängig und identisch verteilte Folge von quadratisch integrierbaren reellen Zufallsvariablen mit* $E[X] = \mu$ *und* $\mathrm{var}\,[X] = \sigma^2 > 0$. *Für alle* $n \in \mathbb{N}$ *sei*

$$S_n := \sum_{k=1}^{n} \frac{X_k - \mu}{\sigma\sqrt{n}}$$

Dann gilt für alle $x \in \mathbb{R}$

$$\lim_{n\to\infty} P[\{S_n \le x\}] = \Phi(x)$$

Beweis. Sei ψ_S die charakteristische Funktion der Standardnormal–Verteilung. Es genügt zu zeigen, dass für alle $t \in \mathbb{R}$

$$\lim_{n\to\infty} |\psi_{S_n}(t) - \psi_S(t)| = 0$$

gilt, denn dann folgt die Behauptung des Satzes aus dem Stetigkeitssatz für charakteristische Funktionen.
Für alle $k \in \mathbb{N}$ sei

$$Z_k = \frac{X_k - \mu}{\sigma}$$

Dann ist Z_k die zu X_k gehörige standardisierte Zufallsvariable und es gilt

$$S_n = \sum_{k=1}^{n} \frac{Z_k}{\sqrt{n}}$$

Mit $\{X_k\}_{k\in\mathbb{N}}$ ist auch $\{Z_k\}_{k\in\mathbb{N}}$ eine unabhängig und identisch verteilte Folge in L^2. Nach Satz 16.4.8 und Lemma 16.4.3 gilt daher

$$\psi_{S_n}(t) = \Big(\psi_{Z/\sqrt{n}}(t)\Big)^n = \left(\psi_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

und nach Beispiel 16.4.2(3) gilt

$$\psi_S(t) = e^{-t^2/2}$$

Wir erhalten daher die Ungleichung

$$\begin{aligned}\Big|\psi_{S_n}(t) - \psi_S(t)\Big| &= \left|\left(\psi_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n - e^{-t^2/2}\right| \\ &\le \left|\left(\psi_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n - \left(1 - \frac{t^2}{2n}\right)^n\right| + \left|\left(1 - \frac{t^2}{2n}\right)^n - e^{-t^2/2}\right|\end{aligned}$$

Der zweite Summand konvergiert gegen 0, und es bleibt zu zeigen, dass auch der erste Summand gegen 0 konvergiert.

Für alle $v, w \in \mathbb{C}$ mit $|v|, |w| \leq 1$ gilt $|v^n - w^n| \leq n\,|v-w|$. Daraus ergibt sich zunächst für hinreichend große $n \in \mathbb{N}$ die Ungleichung

$$\left|\left(\psi_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n - \left(1 - \frac{t^2}{2n}\right)^n\right| \leq n \left|\psi_Z\left(\frac{t}{\sqrt{n}}\right) - \left(1 - \frac{t^2}{2n}\right)\right|$$

und aus Lemma 16.4.4 ergibt sich für alle $s \in \mathbb{R}$ und $\eta \in (0, \infty)$

$$\begin{aligned}
\left|\psi_Z(s) - \left(1 - \frac{s^2}{2}\right)\right| &\leq \int_\Omega \min\Bigl\{|sZ|^3, |sZ|^2\Bigr\}\, dP \\
&\leq \int_{\{|Z| \leq \eta\}} |sZ|^3\, dP + \int_{\{|Z| > \eta\}} |sZ|^2\, dP \\
&\leq \eta\,|s|^3 \int_{\{|Z| \leq \eta\}} Z^2\, dP + s^2 \int_{\{|Z| > \eta\}} Z^2\, dP \\
&\leq \eta\,|s|^3 + s^2 \int_{\{|Z| > \eta\}} Z^2\, dP
\end{aligned}$$

Für alle $\varepsilon \in (0, \infty)$ ergibt sich aus diesen Ungleichungen mit $s := t/\sqrt{n}$ und $\eta := \varepsilon\,\sqrt{n}$

$$\begin{aligned}
\left|\left(\psi_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n - \left(1 - \frac{t^2}{2n}\right)^n\right| &\leq n \left|\psi_Z\left(\frac{t}{\sqrt{n}}\right) - \left(1 - \frac{t^2}{2n}\right)\right| \\
&\leq n \left(\varepsilon\,\frac{|t|^3}{n} + \frac{t^2}{n} \int_{\{|Z| > \varepsilon\sqrt{n}\}} Z^2\, dP\right) \\
&\leq \varepsilon\,|t|^3 + t^2 \int_{\{|Z| > \varepsilon\sqrt{n}\}} Z^2\, dP
\end{aligned}$$

und damit

$$\limsup_{n \to \infty} \left|\left(\psi_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n - \left(1 - \frac{t^2}{2n}\right)^n\right| \leq \varepsilon\,|t|^3$$

Da $\varepsilon \in (0, \infty)$ beliebig war, ergibt sich nun

$$\lim_{n \to \infty} \left|\left(\psi_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n - \left(1 - \frac{t^2}{2n}\right)^n\right| = 0$$

Damit ist gezeigt, dass auch der erste Summand gegen 0 konvergiert. □

Der Zentrale Grenzwertsatz besagt, dass die standardisierten Stichprobenmittel einer unabhängig und identisch verteilten Folge von quadratisch integrierbaren reellen Zufallsvariablen in Verteilung gegen eine standardnormal–verteilte Zufallsvariable konvergieren.

Ist $\{X_k\}_{k \in \mathbb{N}}$ eine unabhängig und identisch verteilte Folge in L^2 mit $E[X] = \mu$ und $\operatorname{var}[X] = \sigma^2 > 0$, so ergibt sich aus dem Zentralen Grenzwertsatz für die Verteilung des Stichprobenmittels die Approximation

$$P\left[\left\{\frac{1}{n}\sum_{k=1}^{n} X_k \leq x\right\}\right] \approx \Phi\left(\frac{x-\mu}{\sigma}\sqrt{n}\right)$$

Diese Approximation ist in der Statistik von Interesse.

18

Bedingte Erwartung

Bei der Konstruktion wahrscheinlichkeitstheoretischer Modelle spielt neben dem Begriff der Unabhängigkeit auch das Konzept des Konditionierens eine zentrale Rolle. Dies lässt sich bereits am Beispiel der Urnenmodelle erkennen:

– Beim Ziehen mit Zurücklegen wird die bei der ersten Ziehung gezogene Kugel zurückgelegt und damit die ursprüngliche Zusammensetzung der Urne wiederhergestellt. Daher ist das Ergebnis der zweiten Ziehung unabhängig vom Ergebnis der ersten Ziehung.
– Beim Ziehen ohne Zurücklegen wird die bei der ersten Ziehung gezogene Kugel nicht zurückgelegt und damit die Anzahl der Kugeln reduziert und die Zusammensetzung der Urne verändert. Daher ist das Ergebnis der zweiten Ziehung abhängig vom Ergebnis der ersten Ziehung. In diesem Fall ist es von Interesse, die bedingten Wahrscheinlichkeiten für das Auftreten einer roten Kugel bei der zweiten Ziehung unter den Bedingungen, dass bei der ersten Ziehung eine rote Kugel bzw. eine andersfarbige Kugel auftritt, zu bestimmen; es liegt dann nahe, diese bedingten Wahrscheinlichkeiten zusammenfassend durch eine einzige bedingte Wahrscheinlichkeit darzustellen, die über das zufällige Ergebnis der ersten Ziehung vom Zufall abhängt.

Die allgemeine Konstruktion zufälliger bedingter Wahrscheinlichkeiten beruht auf der bedingten Erwartung einer Zufallsvariablen bezüglich einer σ–Algebra; hier ist daran zu erinnern, dass auch die Unabhängigkeit von Zufallsvariablen durch eine Eigenschaft von σ–Algebren definiert ist.

In diesem Kapitel führen wir zunächst die bedingte Erwartung einer positiven Zufallsvariablen ein (Abschnitt 18.1) und definieren die bedingte Erwartung dann, analog zur Konstruktion des Integrals, auch für eine Klasse von Zufallsvariablen, die nicht positiv sein müssen (Abschnitt 18.2). Von besonderem Interesse ist die bedingte Erwartung für quadratisch integrierbare Zufallsvariable (Abschnitt 18.3). Als erste Anwendung bedingter Erwartungen untersuchen wir schließlich die Konvergenz trivialer Martingale (Abschnitt 18.4).

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_18, © Springer-Verlag Berlin Heidelberg 2011

Für $p \in \{0\} \cup [1, \infty]$ setzen wir

$$L^p(\mathcal{F}) := L^p(\mathcal{F}, P)$$

Eine σ–Algebra $\mathcal{G}$ auf Ω heißt *Unter–σ–Algebra* von $\mathcal{F}$, wenn $\mathcal{G} \subseteq \mathcal{F}$ gilt. Für eine Unter–σ–Algebra $\mathcal{G}$ von $\mathcal{F}$ sei

$$L^0(\mathcal{F}, \mathcal{G}) := \Big\{ [X]_P \in L^0(\mathcal{F}) \;\Big|\; X =_P Y \text{ für ein } Y \in \mathcal{L}^0(\mathcal{G}) \Big\}$$

Dann ist $L^0(\mathcal{F}, \mathcal{G})$ die Menge derjenigen Äquivalenzklassen von $L^0(\mathcal{F})$, die einen $\mathcal{G}$–messbaren Repräsentanten besitzen. Offenbar ist $L^0(\mathcal{F}, \mathcal{G})$ ein Vektorverband.

Im gesamten Kapitel sei $\mathcal{G}$ eine Unter–σ–Algebra von $\mathcal{F}$.

18.1 Bedingte Erwartung einer positiven Zufallsvariablen

In diesem Abschnitt definieren wir die $\mathcal{G}$–bedingte Erwartung einer positiven Zufallsvariablen und untersuchen ihre Eigenschaften.

18.1.1 Satz. *Sei X eine positive Zufallsvariable. Dann gibt es eine positive $\mathcal{G}$–messbare Zufallsvariable U mit*

$$\int_G U \, dP = \int_G X \, dP$$

für alle $G \in \mathcal{G}$. Sind U und V positive $\mathcal{G}$–messbare Zufallsvariable mit

$$\int_G U \, dP = \int_G X \, dP = \int_G V \, dP$$

für alle $G \in \mathcal{G}$, so gibt es eine Nullmenge $N \in \mathcal{G}$ mit $U(\omega) = V(\omega)$ für alle $\omega \in \Omega \setminus N$.

Beweis. Wir betrachten das Maß $\nu : \mathcal{F} \to [0, \infty]$ mit

$$\nu[A] := \int_A X \, dP$$

und seine Restriktion $\nu|_{\mathcal{G}} : \mathcal{G} \to [0, \infty]$ auf $\mathcal{G}$. Da ν P–stetig ist, ist $\nu|_{\mathcal{G}}$ $P|_{\mathcal{G}}$–stetig. Da $P|_{\mathcal{G}}$ σ–endlich ist, folgt aus dem Satz von Radon/Nikodym die Existenz einer positiven $\mathcal{G}$–messbaren Zufallsvariablen U mit

$$\nu|_{\mathcal{G}}[G] = \int_G U \, dP|_{\mathcal{G}}$$

und damit

$$\int_G U \, dP = \int_G U \, dP|_{\mathcal{G}} = \nu|_{\mathcal{G}}[G] = \nu[G] = \int_G X \, dP$$

für alle $G \in \mathcal{G}$. Nach dem Satz von Radon/Nikodym ist U $P|_{\mathcal{G}}$–fast sicher eindeutig bestimmt. $\square$

Ist X eine positive Zufallsvariable, so nennen wir jede positive $\mathcal{G}$–messbare Zufallsvariable U mit

$$\int_G U\, dP = \int_G X\, dP$$

für alle $G \in \mathcal{G}$ eine *Version der $\mathcal{G}$–bedingten Erwartung* von X; wir bezeichnen mit

$$E^{\mathcal{G}}(X)$$

eine beliebige Version der $\mathcal{G}$–bedingten Erwartung von X und nennen $E^{\mathcal{G}}(X)$ die *$\mathcal{G}$–bedingte Erwartung* von X. Da die σ–Algebra $\mathcal{G}$ außer der leeren Menge weitere Nullmengen enthalten kann, besitzt die $\mathcal{G}$–bedingte Erwartung einer positiven Zufallsvariablen im allgemeinen mehrere Versionen, von denen aber je zwei auf dem Komplement einer Nullmenge in $\mathcal{G}$ übereinstimmen. Für eine beliebige Zufallsvariable Y gilt daher

$$U(\omega) = Y(\omega) \quad \text{fast sicher}$$

entweder für jede oder für keine Version U der $\mathcal{G}$–bedingten Erwartung von X; gilt sie für jede Version der $\mathcal{G}$–bedingten Erwartung von X, so schreiben wir kurz

$$E^{\mathcal{G}}(X) = Y$$

Diese Konvention wenden wir auch auf Ungleichungen an.

Aus der Definition der $\mathcal{G}$–bedingten Erwartung einer positiven Zufallsvariablen X ergibt sich unmittelbar die Gleichung

$$E[E^{\mathcal{G}}(X)] = E[X]$$

Außerdem ist klar, dass jede positive $\mathcal{G}$–messbare Zufallsvariable mit ihrer $\mathcal{G}$–bedingten Erwartung übereinstimmt. Aus diesen Bemerkungen ergeben sich sofort zwei Spezialfälle bezüglich der σ–Algebra $\mathcal{G}$:
– Im Fall $\mathcal{G} = \{\emptyset, \Omega\}$ ist jede $\mathcal{G}$–messbare Zufallsvariable konstant und für jede positive Zufallsvariable X gilt $E^{\mathcal{G}}(X) = E[X]$.
– Im Fall $\mathcal{G} = \mathcal{F}$ ist jede positive Zufallsvariable X $\mathcal{G}$–messbar und es gilt $E^{\mathcal{G}}(X) = X$.

Das folgende Ergebnis zeigt, dass die Definition der $\mathcal{G}$–bedingten Erwartung einer positiven Zufallsvariablen mit der Definition ihres bedingten Erwartungswertes bezüglich einem Ereignis im Einklang steht:

18.1.2 Lemma (Fourier–Entwicklung). *Sei $\{G_i\}_{i\in\{1,\dots,m\}} \subseteq \mathcal{F}$ eine disjunkte Familie von Ereignissen mit $\sum_{i=1}^m G_i = \Omega$ und $P[G_i] > 0$ für alle $i \in \{1,\dots,m\}$ und sei $\mathcal{G}$ die von der Familie $\{G_i\}_{i\in\{1,\dots,m\}}$ erzeugte σ–Algebra. Dann gilt für jede positive Zufallsvariable X*

$$E^{\mathcal{G}}(X) = \sum_{i=1}^m E[X|G_i]\, \chi_{G_i}$$

Beweis. Da die von der Familie $\{G_i\}_{i\in\{1,\ldots,m\}}$ erzeugte σ–Algebra $\mathcal{G}$ außer der leeren Menge keine Nullmengen enthält, ist die $\mathcal{G}$–bedingte Erwartung einer positiven Zufallsvariablen X eindeutig bestimmt und es gibt eine Familie $\{a_i\}_{i\in\{1,\ldots,m\}} \subseteq [0,\infty]$ mit

$$E^{\mathcal{G}}(X) = \sum_{i=1}^{m} a_i \chi_{G_i}$$

Für alle $i \in \{1,\ldots,m\}$ gilt daher

$$\int_{G_i} X\,dP = \int_{G_i} E^{\mathcal{G}}(X)\,dP = \int_{G_i} \left(\sum_{j=1}^{m} a_j \chi_{G_j}\right) dP = \int_{G_i} a_i\,dP = a_i\,P[G_i]$$

und damit

$$a_i = \frac{1}{P[G_i]} \int_{G_i} X\,dP = E[X|G_i]$$

Damit ist die Behauptung gezeigt. □

Das Lemma deutet darauf hin, dass der Übergang von einer positiven Zufallsvariablen X zu ihrer $\mathcal{G}$–bedingten Erwartung als ein lokales Mitteln verstanden werden kann: Wenn die σ–Algebra durch endlich viele Ereignisse erzeugt wird, dann nimmt die $\mathcal{G}$–bedingte Erwartung $E^{\mathcal{G}}(X)$ nur endlich viele Werte an, und im Extremfall $\mathcal{G} = \{\emptyset, \Omega\}$ ist die $\mathcal{G}$–bedingte Erwartung von X sogar konstant.

Ist Θ eine reelle Zufallsvariable mit $P_\Theta[\{\vartheta_1,\ldots,\vartheta_m\}] = 1$ und $P[\{\Theta = \vartheta_i\}] > 0$ für alle $i \in \{1,\ldots,m\}$, so gilt nach Lemma 18.1.2

$$E^{\sigma(\Theta)}(X) = \sum_{i=1}^{m} E[X|\{\Theta = \vartheta_i\}]\,\chi_{\{\Theta=\vartheta_i\}}$$

In diesem Fall gibt es also eine messbare Funktion $h : \mathbb{R} \to \mathbb{R}$ mit

$$E^{\sigma(\Theta)}(X) = h \circ \Theta$$

Dieses Ergebnis lässt sich wie folgt verallgemeinern:

18.1.3 Satz (Faktorisierungssatz). *Sei $(\Omega', \mathcal{F}')$ ein Messraum und sei $\Theta : \Omega \to \Omega'$ eine Zufallsgröße. Dann gibt es zu jeder positiven Zufallsvariablen X eine messbare Funktion $h : \Omega' \to \mathbb{R}$ mit*

$$E^{\sigma(\Theta)}(X) = h \circ \Theta$$

Der Satz folgt unmittelbar aus dem Faktorisierungssatz 7.1.16.

Wir untersuchen nun die Eigenschaften der bedingten Erwartung einer positiven Zufallsvariablen:

18.1.4 Lemma. *Sei X eine positive Zufallsvariable und sei Y eine positive $\mathcal{G}$–messbare Zufallsvariable. Dann sind äquivalent*:
(a) *Es gilt $Y = E^{\mathcal{G}}(X)$.*
(b) *Für jede positive $\mathcal{G}$–messbare Zufallsvariable U gilt*

$$\int_\Omega UY\,dP = \int_\Omega UX\,dP$$

Beweis. Wir nehmen zunächst an, dass $Y = E^{\mathcal{G}}(X)$ gilt, und führen den Beweis durch algebraische Induktion. Nach dem Approximationssatz ist jede positive $\mathcal{G}$–messbare Zufallsvariable das Supremum einer monoton wachsenden Folge von positiven einfachen $\mathcal{G}$–messbaren Zufallsvariablen, und es ist klar, dass jede einfache $\mathcal{G}$–messbare Zufallsvariable eine Linearkombination von $\mathcal{G}$–messbaren Indikatorfunktionen ist. Aufgrund der Linearität des Integrals und des Satzes über die monotone Konvergenz genügt es also zu zeigen, dass die Gleichung

$$\int_\Omega UY\,dP = \int_\Omega UX\,dP$$

für jede $\mathcal{G}$–messbare Indikatorfunktion U gilt. Nach Voraussetzung gilt für alle $G \in \mathcal{G}$

$$\int_\Omega \chi_G Y\,dP = \int_G Y\,dP = \int_G X\,dP = \int_\Omega \chi_G X\,dP$$

und damit in der Tat

$$\int_\Omega UY\,dP = \int_\Omega UX\,dP$$

für jede $\mathcal{G}$–messbare Indikatorfunktion U. Damit ist gezeigt, dass (b) aus (a) folgt, und die umgekehrte Implikation ist klar. □

Das folgende Lemma enthält die elementaren Eigenschaften der bedingten Erwartung einer positiven Zufallsvariablen:

18.1.5 Lemma. *Seien X und Y positive Zufallsvariable. Dann gilt*:
(1) $E^{\mathcal{G}}(E^{\mathcal{G}}(X)) = E^{\mathcal{G}}(X)$.
(2) $E^{\mathcal{G}}(X+Y) = E^{\mathcal{G}}(X) + E^{\mathcal{G}}(Y)$.
(3) *Für jede positive $\mathcal{G}$–messbare Zufallsvariable U gilt $E^{\mathcal{G}}(UX) = UE^{\mathcal{G}}(X)$.*
(4) *Im Fall $X \le Y$ gilt $E^{\mathcal{G}}(X) \le E^{\mathcal{G}}(Y)$.*

Beweis. Aussage (1) folgt aus der $\mathcal{G}$–Messbarkeit von $E^{\mathcal{G}}(X)$. Aus der Linearität des Integrals folgt für alle $G \in \mathcal{G}$

$$\int_G (X+Y)\,dP = \int_G X\,dP + \int_G Y\,dP$$

$$
\begin{aligned}
&= \int_G E^{\mathcal{G}}(X)\,dP + \int_G E^{\mathcal{G}}(Y)\,dP \\
&= \int_G \Bigl(E^{\mathcal{G}}(X) + E^{\mathcal{G}}(Y)\Bigr)\,dP
\end{aligned}
$$

Da $E^{\mathcal{G}}(X)+E^{\mathcal{G}}(Y)$ $\mathcal{G}$–messbar ist, folgt daraus $E^{\mathcal{G}}(X+Y) = E^{\mathcal{G}}(X)+E^{\mathcal{G}}(Y)$. Damit ist (2) gezeigt. Mit U ist für alle $G \in \mathcal{G}$ auch $\chi_G U$ $\mathcal{G}$–messbar. Nach Lemma 18.1.4 gilt daher für alle $G \in \mathcal{G}$

$$
\int_G UX\,dP = \int_\Omega \chi_G UX\,dP = \int_\Omega \chi_G U E^{\mathcal{G}}(X)\,dP = \int_G U E^{\mathcal{G}}(X)\,dP
$$

Da mit U auch $UE^{\mathcal{G}}(X)$ $\mathcal{G}$–messbar ist, folgt daraus $E^{\mathcal{G}}(UX) = UE^{\mathcal{G}}(X)$. Damit ist (3) gezeigt. Zum Beweis von (4) betrachten wir eine monoton wachsende Folge $\{X_n\}_{n\in\mathbb{N}}$ von positiven einfachen Zufallsvariablen mit $X = \sup_{n\in\mathbb{N}} X_n$. Für alle $n \in \mathbb{N}$ ist X_n endlich und für jede Zufallsvariable Z mit $X_n \leq Z$ gilt $(Z-X_n) + X_n = Z$. Aus (2) folgt nun

$$
\begin{aligned}
E^{\mathcal{G}}(X_n) &\leq E^{\mathcal{G}}(Z-X_n) + E^{\mathcal{G}}(X_n) \\
&= E^{\mathcal{G}}(Z)
\end{aligned}
$$

Mit $Z := X_{n+1}$ erkennt man, dass die Folge $\{E^{\mathcal{G}}(X_n)\}_{n\in\mathbb{N}}$ monoton wachsend ist und mit $Z := Y$ ergibt sich

$$
\sup_{n\in\mathbb{N}} E^{\mathcal{G}}(X_n) \leq E^{\mathcal{G}}(Y)
$$

Aus dem Satz über die monotone Konvergenz folgt nun für alle $G \in \mathcal{G}$

$$
\begin{aligned}
\int_G X\,dP &= \int_G \sup_{n\in\mathbb{N}} X_n\,dP \\
&= \sup_{n\in\mathbb{N}} \int_G X_n\,dP \\
&= \sup_{n\in\mathbb{N}} \int_G E^{\mathcal{G}}(X_n)\,dP \\
&= \int_G \sup_{n\in\mathbb{N}} E^{\mathcal{G}}(X_n)\,dP
\end{aligned}
$$

und damit

$$
E^{\mathcal{G}}(X) = \sup_{n\in\mathbb{N}} E^{\mathcal{G}}(X_n)
$$

Daher gilt $E^{\mathcal{G}}(X) = \sup_{n\in\mathbb{N}} E^{\mathcal{G}}(X_n) \leq E^{\mathcal{G}}(Y)$. Damit ist auch (4) gezeigt. □

Bemerkenswert ist die Aussage (3) von Lemma 18.1.5, denn sie besagt, dass sich die positiven $\mathcal{G}$–messbaren Zufallsvariablen bezüglich der $\mathcal{G}$–bedingten Erwartung wie Skalare verhalten.

Für Folgen von positiven Zufallsvariablen gilt die folgende Verallgemeinerung des Satzes über die monotone Konvergenz:

18.1.6 Satz (Monotone Konvergenz; Levi). *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine monoton wachsende Folge von positiven Zufallsvariablen. Dann gilt*

$$E^{\mathcal{G}}\left(\sup_{n\in\mathbb{N}} X_n\right) = \sup_{n\in\mathbb{N}} E^{\mathcal{G}}(X_n)$$

Beweis. Nach Lemma 18.1.5 ist auch die Folge $\{E^{\mathcal{G}}(X_n)\}_{n\in\mathbb{N}}$ monoton wachsend. Aus dem Satz über die monotone Konvergenz folgt nun für alle $G \in \mathcal{G}$

$$\int_G \sup_{n\in\mathbb{N}} X_n \, dP = \sup_{n\in\mathbb{N}} \int_G X_n \, dP = \sup_{n\in\mathbb{N}} \int_G E^{\mathcal{G}}(X_n) \, dP = \int_G \sup_{n\in\mathbb{N}} E^{\mathcal{G}}(X_n) \, dP$$

Da $\sup_{n\in\mathbb{N}} E^{\mathcal{G}}(X_n)$ $\mathcal{G}$–messbar ist, folgt die Behauptung. □

Genau wie im unbedingten Fall erhält man nun die folgende Variante des bedingten Satzes über die monotone Konvergenz:

18.1.7 Folgerung (Monotone Konvergenz; Levi). *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von positiven Zufallsvariablen. Dann gilt*

$$E^{\mathcal{G}}\left(\sum_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} E^{\mathcal{G}}(X_n)$$

Aus dem bedingten Satz über die monotone Konvergenz erhält man außerdem, wieder genau wie im unbedingten Fall, das bedingte Lemma von Fatou:

18.1.8 Lemma (Fatou). *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von positiven Zufallsvariablen. Dann gilt*

$$E^{\mathcal{G}}\left(\liminf_{n\to\infty} X_n\right) \le \liminf_{n\to\infty} E^{\mathcal{G}}(X_n)$$

Da die bedingte Erwartung einer positiven Zufallsvariablen nur fast sicher eindeutig bestimmt ist, liegt es nahe, den Begriff der bedingten Erwartung auf alle Funktionen zu erweitern, die nur fast sicher definiert sind und nur fast sicher mit einer positiven Zufallsvariablen übereinstimmen. Wir benötigen dazu das folgende Lemma, das sich unmittelbar aus Folgerung 8.2.11 ergibt:

18.1.9 Lemma. *Seien X und Y positive Zufallsvariable mit $X = Y$ fast sicher und sei U eine positive $\mathcal{G}$–messbare Zufallsvariable. Dann sind äquivalent:*
(a) *Für alle $G \in \mathcal{G}$ gilt $\int_G U \, dP = \int_G X \, dP$.*
(b) *Für alle $G \in \mathcal{G}$ gilt $\int_G U \, dP = \int_G Y \, dP$.*

Sei also Z eine Funktion, die fast sicher definiert ist und fast sicher mit einer positiven Zufallsvariablen X übereinstimmt, und sei W eine Funktion, die fast sicher definiert ist und fast sicher mit einer Version U der $\mathcal{G}$–bedingten Erwartung von X übereinstimmt. Mit der erweiterten Definition des Integrals erhält man dann für alle $G \in \mathcal{G}$

$$\int_G W\,dP = \int_G U\,dP = \int_G X\,dP = \int_G Z\,dP$$

Daher nennen wir W eine *Version der $\mathcal{G}$–bedingten Erwartung* von Z und wir bezeichnen mit

$$E^{\mathcal{G}}(Z)$$

eine beliebige Version der $\mathcal{G}$–bedingten Erwartung von Z und nennen $E^{\mathcal{G}}(Z)$ die *$\mathcal{G}$–bedingte Erwartung* von Z. Nach Lemma 18.1.9 ist diese Definition unabhängig von der Wahl von X, und sie ist offenbar auch unabhängig von der Wahl von U.

18.2 Bedingte Erwartung und bedingte Integrierbarkeit

In diesem Abschnitt erweitern wir den Begriff der bedingten Erwartung auf eine Klasse von Zufallsvariablen, die nicht notwendigerweise positiv sind.

Eine reelle Zufallsvariable X *besitzt eine $\mathcal{G}$–bedingte Erwartung*, wenn

$$\min\Bigl\{E^{\mathcal{G}}(X^+), E^{\mathcal{G}}(X^-)\Bigr\} < \infty$$

gilt. In diesem Fall gibt es eine Version U der $\mathcal{G}$–bedingten Erwartung von X^+ und eine Version V der $\mathcal{G}$–bedingten Erwartung von X^- derart, dass für alle $\omega \in \Omega$

$$\min\Bigl\{U(\omega), V(\omega)\Bigr\} < \infty$$

gilt und wir nennen eine $\mathcal{G}$–messbare Zufallsvariable W eine *Version der $\mathcal{G}$–bedingten Erwartung* von X, wenn es eine Version U der $\mathcal{G}$–bedingten Erwartung von X^+ und eine Version V der $\mathcal{G}$–bedingten Erwartung von X^- gibt mit

$$W = U - V$$

Wir bezeichnen mit

$$E^{\mathcal{G}}(X)$$

eine beliebige Version der $\mathcal{G}$–bedingten Erwartung von X und nennen $E^{\mathcal{G}}(X)$ die *$\mathcal{G}$–bedingte Erwartung* von X. Diese Definition ist mit der Definition der $\mathcal{G}$–bedingten Erwartung einer positiven Zufallsvariablen verträglich. Eine reelle Zufallsvariable, die eine bedingte Erwartung besitzt, heißt *$\mathcal{G}$–bedingt quasi-integrierbar*.

Eine reelle Zufallsvariable X *besitzt eine endliche $\mathcal{G}$–bedingte Erwartung*, wenn

$$\max\Bigl\{E^{\mathcal{G}}(X^+), E^{\mathcal{G}}(X^-)\Bigr\} < \infty$$

gilt. In diesem Fall besitzen X^+ und X^- und damit auch X eine endliche Version ihrer $\mathcal{G}$–bedingten Erwartung. Eine reelle Zufallsvariable, die eine endliche bedingte Erwartung besitzt, heißt *$\mathcal{G}$–bedingt integrierbar*.

Das folgende Lemma 18.2.1 legt es nahe, den Begriff der bedingten Integrierbarkeit auf Funktionen auszudehnen, die nur fast sicher definiert sind:

18.2.1 Lemma. *Seien X und Y reelle Zufallsvariable mit $X(\omega) = Y(\omega)$ fast sicher. Dann sind äquivalent*:
(a) *X ist $\mathcal{G}$–bedingt integrierbar.*
(b) *Y ist $\mathcal{G}$–bedingt integrierbar.*
In diesem Fall gilt $E^{\mathcal{G}}(X) = E^{\mathcal{G}}(Y)$.

Sei also Z eine Funktion, die fast sicher definiert ist und fast sicher mit einer $\mathcal{G}$–bedingt integrierbaren Zufallsvariablen X übereinstimmt, und sei W eine Funktion, die fast sicher definiert ist und fast sicher mit einer Version der $\mathcal{G}$–bedingten Erwartung von X übereinstimmt. Dann heißt auch Z *$\mathcal{G}$–bedingt integrierbar* und W heißt *Version der $\mathcal{G}$–bedingten Erwartung* von Z und wir bezeichnen mit

$$E^{\mathcal{G}}(Z)$$

eine beliebige Version der $\mathcal{G}$–bedingten Erwartung von Z und nennen $E^{\mathcal{G}}(Z)$ die *bedingte Erwartung* von Z.

Im Sinne dieser erweiterten Definition ist die $\mathcal{G}$–bedingte Integrierbarkeit eine Eigenschaft von Elementen des Vektorverbandes $L^0(\mathcal{F})$ und damit eine Eigenschaft von Äquivalenzklassen der Menge $\mathcal{L}^0(\mathcal{F})$. Für $X \in L^0(\mathcal{F})$ bildet auch die Menge aller Versionen der $\mathcal{G}$–bedingten Erwartung von X eine Äquivalenzklasse in $\mathcal{L}^0(\mathcal{F})$. Dementsprechend sind im folgenden alle Gleichungen bzw. Ungleichungen für $\mathcal{G}$–bedingte Erwartungen als Gleichungen bzw. Ungleichungen für Äquivalenzklassen und damit als Gleichungen bzw. Ungleichungen für reelle Repräsentanten zu lesen. Sei

$$L^{1,\mathcal{G}}(\mathcal{F}) := \Bigl\{[X]_P \in L^0(\mathcal{F}) \Bigm| X =_P Y \text{ für ein } Y \text{ mit } E^{\mathcal{G}}(|Y|) < \infty\Bigr\}$$

Dann ist $L^{1,\mathcal{G}}(\mathcal{F})$ die Menge aller Äquivalenzklassen, die (im Sinne der ursprünglichen Definition der $\mathcal{G}$–bedingten Integrierbarkeit) einen $\mathcal{G}$–bedingt integrierbaren Repräsentanten besitzen, und die $\mathcal{G}$–bedingte Erwartung lässt sich als Abbildung

$$L^{1,\mathcal{G}}(\mathcal{F}) \to L^0(\mathcal{F},\mathcal{G}) : X \mapsto E^{\mathcal{G}}(X)$$

verstehen.

Das folgende Lemma charakterisiert die bedingte Integrierbarkeit einer Zufallsvariablen:

18.2.2 Lemma. *Sei X eine Zufallsvariable. Dann sind äquivalent*:
(a) *X ist $\mathcal{G}$–bedingt integrierbar.*
(b) *Es gilt $E^{\mathcal{G}}(|X|) < \infty$.*
(c) *Es gibt positive Zufallsvariable U und V mit $X = U - V$ fast sicher und*

$$\max\Big\{E^{\mathcal{G}}(U), E^{\mathcal{G}}(V)\Big\} < \infty$$

In diesem Fall gilt $|E^{\mathcal{G}}(X)| < \infty$.

Offenbar ist eine positive Zufallsvariable genau dann $\mathcal{G}$–bedingt integrierbar, wenn ihre $\mathcal{G}$–bedingte Erwartung endlich ist. Daher lässt sich das letzte Lemma auch wie folgt formulieren:

18.2.3 Folgerung. *Sei X eine Zufallsvariable. Dann sind äquivalent*:
(a) *X ist $\mathcal{G}$–bedingt integrierbar.*
(b) *$|X|$ ist $\mathcal{G}$–bedingt integrierbar.*
(c) *X^+ und X^- sind $\mathcal{G}$–bedingt integrierbar.*
(d) *Es gibt positive $\mathcal{G}$–bedingt integrierbare Zufallsvariable U und V mit $X(\omega) = U(\omega) - V(\omega)$ fast sicher.*

Eine weitere Charakterisierung der bedingten Integrierbarkeit einer Zufallsvariablen geben wir in Lemma 19.1.4.

Das folgende Lemma liefert eine allgemeine Darstellung der $\mathcal{G}$–bedingten Erwartung einer $\mathcal{G}$–bedingt integrierbaren Zufallsvariablen, die sich bei der Herleitung der Eigenschaften der $\mathcal{G}$–bedingten Erwartung als nützlich erweist und auch ihre Berechnung erleichtern kann:

18.2.4 Lemma. *Sei X eine $\mathcal{G}$–bedingt integrierbare Zufallsvariable. Dann gilt für jede Wahl von positiven $\mathcal{G}$–bedingt integrierbaren Zufallsvariablen U und V mit $X(\omega) = U(\omega) - V(\omega)$ fast sicher*

$$E^{\mathcal{G}}(X) = E^{\mathcal{G}}(U) - E^{\mathcal{G}}(V)$$

Das folgende Lemma ist ein Analogon zu Lemma 18.1.5:

18.2.5 Lemma. *Seien X und Y $\mathcal{G}$–bedingt integrierbare Zufallsvariable. Dann gilt*:
(1) *$E^{\mathcal{G}}(X)$ ist $\mathcal{G}$–bedingt integrierbar und es gilt $E^{\mathcal{G}}(E^{\mathcal{G}}(X)) = E^{\mathcal{G}}(X)$.*
(2) *Die Zufallsvariable $X+Y$ ist $\mathcal{G}$–bedingt integrierbar und es gilt*

$$E^{\mathcal{G}}(X+Y) = E^{\mathcal{G}}(X) + E^{\mathcal{G}}(Y)$$

(3) *Für jede reelle $\mathcal{G}$–messbare Zufallsvariable U ist UX $\mathcal{G}$–bedingt integrierbar und es gilt*
$$E^{\mathcal{G}}(UX) = UE^{\mathcal{G}}(X)$$
(4) *Im Fall $X \leq Y$ gilt*
$$E^{\mathcal{G}}(X) \leq E^{\mathcal{G}}(Y)$$
(5) *Die Zufallsvariablen $X \vee Y$ und $X \wedge Y$ sind $\mathcal{G}$–bedingt integrierbar.*
(6) *Jede Zufallsvariable Z mit $|Z| \leq |X|$ ist $\mathcal{G}$–bedingt integrierbar.*

Der folgende Satz fasst die Struktureigenschaften von $L^{1,\mathcal{G}}(\mathcal{F})$ und die Eigenschaften der $\mathcal{G}$–bedingten Erwartung zusammen:

18.2.6 Satz. *$L^{1,\mathcal{G}}(\mathcal{F})$ ist ein Vektorverband und ein Ideal in $L^0(\mathcal{F})$ und die Abbildung $L^{1,\mathcal{G}}(\mathcal{F}) \to L^0(\mathcal{F},\mathcal{G}) : X \mapsto E^{\mathcal{G}}(X)$ ist linear, positiv und idempotent.*

Bemerkenswert ist wieder die Eigenschaft (3) aus Lemma 18.2.5, denn sie besagt, dass sich die reellen $\mathcal{G}$–messbaren Zufallsvariablen bezüglich der $\mathcal{G}$–bedingten Erwartung wie Skalare verhalten.

Für Folgen von bedingt integrierbaren Zufallsvariablen gilt die folgende Verallgemeinerung des Satzes über die majorisierte Konvergenz:

18.2.7 Satz (Majorisierte Konvergenz; Lebesgue). *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen, die fast sicher gegen eine Zufallsvariable X konvergiert. Wenn $\sup_{n\in\mathbb{N}} |X_n|$ $\mathcal{G}$–bedingt integrierbar ist, dann ist auch X $\mathcal{G}$–bedingt integrierbar und es gilt*
$$\lim_{n\to\infty} E^{\mathcal{G}}(X_n) = E^{\mathcal{G}}\Big(\lim_{n\to\infty} X_n\Big)$$
fast sicher.

Wir betrachten nun bedingte Erwartungen bezüglich geschachtelter Unter–σ–Algebren von $\mathcal{F}$:

18.2.8 Satz (David schlägt Goliath). *Sei $\mathcal{D}$ eine Unter–σ–Algebra von $\mathcal{F}$ mit $\mathcal{D} \subseteq \mathcal{G}$. Dann gilt*
$$L^{1,\mathcal{D}}(\mathcal{F}) \subseteq L^{1,\mathcal{G}}(\mathcal{F})$$
und für jede positive oder $\mathcal{D}$–bedingt integrierbare Zufallsvariable X gilt
$$E^{\mathcal{G}}(E^{\mathcal{D}}(X)) = E^{\mathcal{D}}(X) = E^{\mathcal{D}}(E^{\mathcal{G}}(X))$$

Beweis. Wir betrachten zunächst eine positive Zufallsvariable X. Wegen $\mathcal{D} \subseteq \mathcal{G}$ ist die Zufallsvariable $E^{\mathcal{D}}(X)$ $\mathcal{G}$–messbar, und daraus folgt

$$E^{\mathcal{G}}(E^{\mathcal{D}}(X)) = E^{\mathcal{D}}(X)$$

Wegen $\mathcal{D} \subseteq \mathcal{G}$ gilt außerdem für alle $D \in \mathcal{D}$

$$\int_D X\,dP = \int_D E^{\mathcal{G}}(X)\,dP = \int_D E^{\mathcal{D}}(E^{\mathcal{G}}(X))\,dP$$

und daraus folgt

$$E^{\mathcal{D}}(X) = E^{\mathcal{D}}(E^{\mathcal{G}}(X))$$

Damit ist die Gleichung

$$E^{\mathcal{G}}(E^{\mathcal{D}}(X)) = E^{\mathcal{D}}(X) = E^{\mathcal{D}}(E^{\mathcal{G}}(X))$$

für positive Zufallsvariable gezeigt. Wir nehmen nun zusätzlich an, dass die positive Zufallsvariable X auch $\mathcal{D}$–bedingt integrierbar ist, und betrachten das Maß

$$\nu := \int X\,dP$$

sowie seine Restriktionen $\nu|_{\mathcal{G}}$ und $\nu|_{\mathcal{D}}$ auf $\mathcal{G}$ bzw. $\mathcal{D}$. Da X $\mathcal{D}$–bedingt integrierbar ist, ist $E^{\mathcal{D}}(X)$ endlich, und aus Satz 9.2.10 folgt nun, dass $\nu|_{\mathcal{D}}$ σ–endlich ist. Wegen $\mathcal{D} \subseteq \mathcal{G}$ ist dann auch $\nu|_{\mathcal{G}}$ σ–endlich, und aus Satz 9.2.10 folgt nun, dass $E^{\mathcal{G}}(X)$ endlich ist. Daher ist X $\mathcal{G}$–bedingt integrierbar.
Für eine beliebige $\mathcal{D}$–bedingt integrierbare Zufallsvariable X folgt die Behauptung des Satzes nun aus der Zerlegung $X = X^+ - X^-$ und der Linearität des Erwartungswertes. □

Mit $\mathcal{D} := \{\emptyset, \Omega\}$ ergibt sich aus dem letzten Satz, dass insbesondere jede integrierbare Zufallsvariable auch $\mathcal{G}$–bedingt integrierbar ist:

18.2.9 Folgerung (David schlägt Goliath). *Es gilt*

$$L^1(\mathcal{F}) \subseteq L^{1,\mathcal{G}}(\mathcal{F})$$

und für jede positive oder integrierbare Zufallsvariable X gilt

$$E[X] = E[E^{\mathcal{G}}(X)]$$

Da jede Zufallsvariable trivialerweise $\mathcal{F}$–bedingt integrierbar ist, erkennt man andererseits am Beispiel $\mathcal{G} = \mathcal{F}$, dass eine $\mathcal{G}$–bedingt integrierbare Zufallsvariable nicht notwendigerweise integrierbar ist.

Wir untersuchen nun Eigenschaften der bedingten Erwartung unter Unabhängigkeitsannahmen:

18.2.10 Satz (Reduktionssatz). *Sei $\mathcal{H}$ eine Unter–σ–Algebra von $\mathcal{F}$ und sei X eine positive oder $\mathcal{G}$–bedingt integrierbare Zufallsvariable. Ist $\mathcal{H}$ unabhängig von $\sigma(\mathcal{G} \cup \sigma(X))$, so gilt*

$$E^{\sigma(\mathcal{G}\cup\mathcal{H})}(X) = E^{\mathcal{G}}(X)$$

Beweis. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass X positiv und einfach ist. Dann ist das Mengensystem

$$\mathcal{D} := \left\{ D \in \mathcal{F} \;\middle|\; \int_D E^{\mathcal{G}}(X)\, dP = \int_D X\, dP \right\}$$

ein Dynkin–System und das Mengensystem

$$\mathcal{C} := \left\{ C \in \mathcal{F} \;\middle|\; C = G \cap H \text{ mit } G \in \mathcal{G} \text{ und } H \in \mathcal{H} \right\}$$

ist ein $\cap$–stabiler Erzeuger von $\sigma(\mathcal{G} \cup \mathcal{H})$. Für alle $G \in \mathcal{G}$ und $H \in \mathcal{H}$ gilt aufgrund der vorausgesetzten Unabhängigkeit

$$\begin{aligned}
\int_{G\cap H} E^{\mathcal{G}}(X)\, dP &= \int_\Omega \chi_H \chi_G E^{\mathcal{G}}(X)\, dP \\
&= P[H] \int_G E^{\mathcal{G}}(X)\, dP \\
&= P[H] \int_G X\, dP \\
&= \int_\Omega \chi_H \chi_G X\, dP \\
&= \int_{G\cap H} X\, dP
\end{aligned}$$

Daher gilt $\mathcal{C} \subseteq \mathcal{D}$ und damit $\sigma(\mathcal{G} \cup \mathcal{H}) = \sigma(\mathcal{C}) = \delta(\mathcal{C}) \subseteq \mathcal{D}$. Damit ist gezeigt, dass die Gleichung

$$\int_D E^{\mathcal{G}}(X)\, dP = \int_D X\, dP$$

für alle $D \in \sigma(\mathcal{G} \cup \mathcal{H})$ gilt. Da $E^{\mathcal{G}}(X)$ $\mathcal{G}$–messbar und damit auch $\sigma(\mathcal{G} \cup \mathcal{H})$–messbar ist, folgt daraus $E^{\mathcal{G}}(X) = E^{\sigma(\mathcal{G}\cup\mathcal{H})}(X)$. □

Mit $\mathcal{G} := \{\emptyset, \Omega\}$ erhält man sofort einen wichtigen Spezialfall des Reduktionssatzes:

18.2.11 Folgerung. *Sei $\mathcal{H}$ eine Unter–σ–Algebra von $\mathcal{F}$ und sei X eine positive oder integrierbare Zufallsvariable. Ist $\mathcal{H}$ unabhängig von $\sigma(X)$, so gilt*

$$E^{\mathcal{H}}(X) = E[X]$$

Allgemein lässt sich die Unabhängigkeit von Unter–σ–Algebren durch Eigenschaften ihrer bedingten Erwartungen charakterisieren:

18.2.12 Folgerung. *Sei $\mathcal{H}$ eine Unter–σ–Algebra von $\mathcal{F}$. Dann sind äquivalent*:
(a) *$\mathcal{G}$ und $\mathcal{H}$ sind unabhängig.*
(b) *Für jede positive $\mathcal{G}$–messbare Zufallsvariable X gilt $E^{\mathcal{H}}(X) = E[X]$.*
(c) *Für jede positive $\mathcal{H}$–messbare Zufallsvariable X gilt $E^{\mathcal{G}}(X) = E[X]$.*
(d) *Für jede integrierbare $\mathcal{G}$–messbare Zufallsvariable X gilt $E^{\mathcal{H}}(X) = E[X]$.*
(e) *Für jede integrierbare $\mathcal{H}$–messbare Zufallsvariable X gilt $E^{\mathcal{G}}(X) = E[X]$.*

Beweis. Offenbar genügt es, die Äquivalenz von (a) und (b) zu beweisen.
Wir nehmen zunächst an, dass (a) gilt. Für jede positive $\mathcal{G}$–messbare Zufallsvariable X ist dann $\mathcal{H}$ unabhängig von $\sigma(X)$ und aus Folgerung 18.2.11 ergibt sich $E^{\mathcal{H}}(X) = E[X]$. Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Für alle $G \in \mathcal{G}$ und $H \in \mathcal{H}$ gilt

$$\begin{aligned} P[G \cap H] &= \int_H \chi_G \, dP \\ &= \int_H E^{\mathcal{H}}(\chi_G) \, dP \\ &= \int_H E[\chi_G] \, dP \\ &= \int_H P[G] \, dP \\ &= P[G] \, P[H] \end{aligned}$$

Daher folgt (a) aus (b). □

Für eine $\mathcal{G}$–bedingt integrierbare Zufallsvariable X heißt

$$\mathrm{var}^{\mathcal{G}}(X) := E^{\mathcal{G}}\Bigl((X - E^{\mathcal{G}}(X))^2\Bigr)$$

die *$\mathcal{G}$–bedingte Varianz* von X.

18.2.13 Lemma. *Sei X eine $\mathcal{G}$–bedingt integrierbare Zufallsvariable und sei U eine reelle $\mathcal{G}$–messbare Zufallsvariable. Dann gilt*

$$E^{\mathcal{G}}((X+U)^2) = E^{\mathcal{G}}(X^2) + 2U E^{\mathcal{G}}(X) + U^2$$

Insbesondere gilt

$$E^{\mathcal{G}}((X - E^{\mathcal{G}}(X))^2) = E^{\mathcal{G}}(X^2) - (E^{\mathcal{G}}(X))^2$$

Beweis. Sei $G := \{E^{\mathcal{G}}(X^2) < \infty\}$. Dann sind alle Summanden auf der rechten Seite der Gleichung

$$\chi_G(X+U)^2 = \chi_G X^2 + 2\chi_G U E^{\mathcal{G}}(X) + \chi_G U^2$$

$\mathcal{G}$–bedingt integrierbar und es gilt

$$\begin{aligned}\chi_G E^{\mathcal{G}}((X+U)^2) &= E^{\mathcal{G}}(\chi_G(X+U)^2)\\ &= E^{\mathcal{G}}\Big(\chi_G X^2 + 2\chi_G UX + \chi_G U^2\Big)\\ &= \chi_G E^{\mathcal{G}}(X^2) + 2\chi_G U E^{\mathcal{G}}(X) + \chi_G U^2\\ &= \chi_G\Big(E^{\mathcal{G}}(X^2) + 2UE^{\mathcal{G}}(X) + U^2\Big)\end{aligned}$$

Wegen $(a \pm b)^2 \le 2\,(a^2+b^2)$ gilt einerseits

$$E^{\mathcal{G}}(X^2) = E^{\mathcal{G}}(((X+U) - U)^2) \le 2\,\Big(E^{\mathcal{G}}((X+U)^2) + U^2\Big)$$

und andererseits

$$E^{\mathcal{G}}((X+U)^2) \le 2\,\Big(E^{\mathcal{G}}(X^2) + U^2\Big)$$

Da U endlich ist, folgt aus diesen Ungleichungen $G = \{E^{\mathcal{G}}((X+U)^2) < \infty\}$, und damit

$$\chi_{\Omega\setminus G} E^{\mathcal{G}}((X+U)^2) = \chi_{\Omega\setminus G}\Big(E^{\mathcal{G}}(X^2) + 2UE^{\mathcal{G}}(X) + U^2\Big)$$

Die Behauptung des Lemmas ergibt sich nun durch Summation. □

Das folgende Lemma zeigt, dass die Eigenschaften der $\mathcal{G}$–bedingten Varianz denen der unbedingten Varianz entsprechen, wobei wieder die reellen $\mathcal{G}$–messbaren Zufallsvariablen an die Stelle der Konstanten treten:

18.2.14 Lemma. *Sei X eine $\mathcal{G}$–bedingt integrierbare Zufallsvariable.*
(1) *Es gilt* $\mathrm{var}^{\mathcal{G}}(X) = E^{\mathcal{G}}(X^2) - (E^{\mathcal{G}}(X))^2$.
(2) *Es gilt* $\mathrm{var}^{\mathcal{G}}(X) = 0$ *genau dann, wenn* $P[\{X = E^{\mathcal{G}}(X)\}] = 1$ *gilt.*
(3) *Es gilt* $\mathrm{var}^{\mathcal{G}}(X) = \inf\{E^{\mathcal{G}}((X-U)^2) \mid U \textit{ ist } \mathcal{G}\textit{–messbar}\}$.
(4) *Für jede Wahl von reellen $\mathcal{G}$–messbaren Zufallsvariablen U und V ist $U + VX$ $\mathcal{G}$–bedingt integrierbar und es gilt* $\mathrm{var}^{\mathcal{G}}(U+VX) = V^2\,\mathrm{var}^{\mathcal{G}}(X)$.

Beweis. Da $E^{\mathcal{G}}(X)$ $\mathcal{G}$–messbar ist, ergibt sich (1) unmittelbar aus Lemma 18.2.13. Im Fall $\mathrm{var}^{\mathcal{G}}(X) = 0$ gilt

$$\int_\Omega (X - E^{\mathcal{G}}(X))^2\,dP = \int_\Omega E^{\mathcal{G}}((X - E^{\mathcal{G}}(X))^2)\,dP = \int_\Omega \mathrm{var}^{\mathcal{G}}(X)\,dP = 0$$

und damit $P[\{X = E^{\mathcal{G}}(X)\}] = 1$, und die umgekehrte Implikation ist klar. Damit ist (2) gezeigt. Für jede $\mathcal{G}$–messbare Zufallsvariable U folgt aus Lemma 18.2.13 wegen $E^{\mathcal{G}}(X - E^{\mathcal{G}}(X)) = 0$

$$\begin{aligned} E^{\mathcal{G}}((X-U)^2) &= E^{\mathcal{G}}\Big(\big((X-E^{\mathcal{G}}(X)) - (U-E^{\mathcal{G}}(X))\big)^2\Big) \\ &= E^{\mathcal{G}}\Big((X-E^{\mathcal{G}}(X))^2\Big) + (U-E^{\mathcal{G}}(X))^2 \\ &\geq \mathrm{var}^{\mathcal{G}}(X) \end{aligned}$$

Damit ist (3) gezeigt. Schließlich gilt für reelle $\mathcal{G}$–messbare Zufallsvariable U und V

$$\begin{aligned} \mathrm{var}^{\mathcal{G}}(U+VX) &= E^{\mathcal{G}}\Big(\big((U+VX) - E^{\mathcal{G}}(U+VX)\big)^2\Big) \\ &= E^{\mathcal{G}}\Big(V^2(X-E^{\mathcal{G}}(X))^2\Big) \\ &= V^2 E^{\mathcal{G}}\Big((X-E^{\mathcal{G}}(X))^2\Big) \\ &= V^2\,\mathrm{var}^{\mathcal{G}}(X) \end{aligned}$$

Damit ist auch (4) gezeigt. □

Für $\mathcal{G}$–bedingt integrierbare Zufallsvariable X und Y, deren Produkt ebenfalls $\mathcal{G}$–bedingt integrierbar ist, heißt

$$\mathrm{cov}^{\mathcal{G}}(X,Y) := E^{\mathcal{G}}\Big((X-E^{\mathcal{G}}(X))\,(Y-E^{\mathcal{G}}(Y))\Big)$$

die *$\mathcal{G}$–bedingte Kovarianz* von X und Y, und die Zufallsvariablen heißen *$\mathcal{G}$–bedingt unkorreliert*, wenn $\mathrm{cov}^{\mathcal{G}}(X,Y) = 0$ gilt.

18.2.15 Lemma. *Sei X und Y $\mathcal{G}$–bedingt integrierbare Zufallsvariable, deren Produkt ebenfalls $\mathcal{G}$–bedingt integrierbar ist. Dann gilt*

$$\mathrm{cov}^{\mathcal{G}}(X,Y) = E^{\mathcal{G}}(XY) - E^{\mathcal{G}}(X)\,E^{\mathcal{G}}(Y)$$

und

$$\mathrm{var}^{\mathcal{G}}(X+Y) = \mathrm{var}^{\mathcal{G}}(X) + 2\,\mathrm{cov}^{\mathcal{G}}(X,Y) + \mathrm{var}^{\mathcal{G}}(Y)$$

Wir zeigen abschließend, dass die bedingte Erwartung einer integrierbaren Zufallsvariablen durch eine Familie von Gleichungen charakterisiert wird, die an die Definition der bedingten Erwartung einer positiven Zufallsvariablen erinnert:

18.2.16 Satz. *Sei $X \in L^1(\mathcal{F})$ und sei Y $\mathcal{G}$–messbar. Dann sind äquivalent:*
(a) *Es gilt $Y = E^{\mathcal{G}}(X)$.*
(b) *Für alle $G \in \mathcal{G}$ gilt $\int_G Y\,dP = \int_G X\,dP$.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Nach Lemma 18.2.5 gilt dann für alle $G \in \mathcal{G}$

$$\begin{aligned}\int_G Y\,dP &= \int_G E^{\mathcal{G}}(X)\,dP \\ &= \int_G E^{\mathcal{G}}(X^+)\,dP - \int_G E^{\mathcal{G}}(X^-)\,dP \\ &= \int_G X^+\,dP - \int_G X^-\,dP \\ &= \int_G X\,dP\end{aligned}$$

Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Nach Lemma 18.2.5 gilt dann für alle $G \in \mathcal{G}$

$$\int_\Omega E^{\mathcal{G}}(X)\chi_G\,dP = \int_\Omega E^{\mathcal{G}}(X\chi_G)\,dP = \int_\Omega X\chi_G\,dP = \int_\Omega Y\chi_G\,dP$$

und damit

$$\int_\Omega (E^{\mathcal{G}}(X)-Y)\chi_G\,dP = 0$$

Sei $G^+ := \{E^{\mathcal{G}}(X) \geq Y\}$. Da Y $\mathcal{G}$–messbar ist, gilt $G^+ \in \mathcal{G}$ und aus der letzten Gleichung folgt

$$(E^{\mathcal{G}}(X) - Y)^+ = 0$$

Analog erhält man mit $G^- := \{E^{\mathcal{G}}(X) \leq Y\}$

$$(E^{\mathcal{G}}(X) - Y)^- = 0$$

Aus diesen Gleichungen folgt $E^{\mathcal{G}}(X) = Y$. Daher folgt (a) aus (b). □

Aufgaben

18.2.A Führen Sie die fehlenden Beweise aus.

18.2.B **Bedingt quasiintegrierbare Zufallsvariable:** Erweitern Sie den Begriff der $\mathcal{G}$–bedingten Erwartung auf Funktionen, die fast sicher definiert sind und fast sicher mit einer $\mathcal{G}$–bedingt quasiintegrierbaren Zufallsvariablen übereinstimmen.

18.2.C **Fourier–Entwicklung:** Sei $\{G_i\}_{i\in\{1,\dots,m\}} \subseteq \mathcal{F}$ eine disjunkte Familie von Ereignissen mit $\sum_{i=1}^m G_i = \Omega$ und $P[G_i] > 0$ für alle $i \in \{1,\dots,m\}$ und sei $\mathcal{G}$ die von der Familie $\{G_i\}_{i\in\{1,\dots,m\}}$ erzeugte σ–Algebra. Dann gilt für jede $\mathcal{G}$–bedingt integrierbare Zufallsvariable X

$$E^{\mathcal{G}}(X) = \sum_{i=1}^m E[X|G_i]\,\chi_{G_i}$$

18.2.D **Faktorisierungssatz:** Sei $(\Omega', \mathcal{F}')$ ein Messraum und sei $\Theta : \Omega \to \Omega'$ eine Zufallsgröße. Dann gibt es zu jeder $\sigma(\Theta)$–integrierbaren Zufallsvariablen X eine messbare Funktion $h : \Omega' \to \mathbb{R}$ mit

$$E^{\sigma(\Theta)}(X) = h \circ \Theta$$

18.2.E **Ungleichung von Jensen:** Sei X $\mathcal{G}$–bedingt integrierbar. Dann gilt

$$|E^{\mathcal{G}}(X)| \le E^{\mathcal{G}}(|X|)$$

18.2.F **David schlägt Goliath:** Sei $\mathcal{H}$ eine Unter–σ–Algebra von $\mathcal{F}$. Dann sind äquivalent:
(a) $L^0(\mathcal{F}, \mathcal{H}) \subseteq L^0(\mathcal{F}, \mathcal{G})$.
(b) Für alle $H \in \mathcal{H}$ gilt

$$\chi_H = E^{\mathcal{H}}(E^{\mathcal{G}}(\chi_H))$$

(c) Für jede positive Zufallsvariable X gilt

$$E^{\mathcal{H}}(X) = E^{\mathcal{H}}(E^{\mathcal{G}}(X))$$

(d) Für jede positive Zufallsvariable X gilt

$$E^{\mathcal{G}}(E^{\mathcal{H}}(X)) = E^{\mathcal{H}}(X) = E^{\mathcal{H}}(E^{\mathcal{G}}(X))$$

18.2.G **Bedingte Kovarianz:** Seien X und Y $\mathcal{G}$–bedingt integrierbare Zufallsvariable, deren Produkt XY ebenfalls $\mathcal{G}$–bedingt integrierbar ist. Dann sind äquivalent:
(a) X und Y sind $\mathcal{G}$–bedingt unkorreliert.
(b) Es gilt $E^{\mathcal{G}}(XY) = E^{\mathcal{G}}(X)\, E^{\mathcal{G}}(Y)$.
(c) Es gilt $\mathrm{var}^{\mathcal{G}}(X+Y) = \mathrm{var}^{\mathcal{G}}(X) + \mathrm{var}^{\mathcal{G}}(Y)$.
Für jede $\mathcal{G}$–messbare Zufallsvariable U gilt $\mathrm{cov}^{\mathcal{G}}(X, U) = 0$.

18.3 Bedingte Erwartung als Projektion

Wegen $L^2(\mathcal{F}) \subseteq L^1(\mathcal{F}) \subseteq L^{1,\mathcal{G}}(\mathcal{F})$ gelten alle Ergebnisse über Zufallsvariable in $L^{1,\mathcal{G}}(\mathcal{F})$ insbesondere für Zufallsvariable in $L^2(\mathcal{F})$. Das folgende Lemma zeigt, dass für eine quadratisch integrierbare Zufallsvariable auch ihre bedingte Erwartung quadratisch integrierbar ist:

18.3.1 Lemma. *Sei $X \in L^2(\mathcal{F})$. Dann gilt $E^{\mathcal{G}}(X) \in L^2(\mathcal{F})$.*

Beweis. Nach Lemma 18.2.13 gilt $(E^{\mathcal{G}}(X))^2 \le E^{\mathcal{G}}(X^2)$, und Integration ergibt $E[(E^{\mathcal{G}}(X))^2] \le E[E^{\mathcal{G}}(X^2)] = E[X^2]$. □

Der folgende Satz enthält eine fundamentale Eigenschaft der bedingten Erwartung einer quadratisch integrierbaren Zufallsvariablen:

18.3.2 Satz (Pythagoras). *Sei $X \in L^2(\mathcal{F})$ und $U \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$. Dann gilt*

$$E[(X-U)^2] = E[(X-E^{\mathcal{G}}(X))^2] + E[(E^{\mathcal{G}}(X)-U)^2]$$

Insbesondere gilt

$$E[X^2] = E[(X-E^{\mathcal{G}}(X))^2] + E[(E^{\mathcal{G}}(X))^2]$$

Beweis. Es gilt

$$E[(X-E^{\mathcal{G}}(X))(E^{\mathcal{G}}(X)-U)] = E[E^{\mathcal{G}}((X-E^{\mathcal{G}}(X))(E^{\mathcal{G}}(X)-U))] = 0$$

Daraus folgt die behauptete Gleichung. □

Der folgende Satz charakterisiert die bedingte Erwartung einer quadratisch integrierbaren Zufallsvariablen:

18.3.3 Satz. *Sei $X \in L^2(\mathcal{F})$. Dann sind für $Y \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ folgende Aussagen äquivalent:*
(a) *Es gilt $Y = E^{\mathcal{G}}(X)$.*
(b) *Für alle $Z \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ gilt $E[(X-Y)^2] \leq E[(X-Z)^2]$.*
(c) *Für alle $Z \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ gilt $E[XZ] = E[YZ]$.*

Beweis. Wir nehmen zunächst an, dass (a) gilt. Wegen $Y = E^{\mathcal{G}}(X)$ ergibt sich aus dem Satz von Pythagoras für alle $Z \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$

$$E[(X-Z)^2] = E[(X-Y)^2] + E[(Y-Z)^2]$$

Daher folgt (b) aus (a).
Wir nehmen nun an, dass (b) gilt. Für $Z \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ und $c \in (0,\infty)$ gilt dann $Y \pm cZ \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ und damit

$$\begin{aligned} E[(X-Y)^2] &\leq E[(X - (Y \pm cZ))^2] \\ &= E[((X-Y) \mp cZ)^2] \\ &= E[(X-Y)^2] \mp 2c\, E[(X-Y)Z] + c^2 E[Z^2] \end{aligned}$$

Daraus folgt zunächst $|2\, E[(X-Y)Z]| \leq c\, E[Z^2]$ und da $c \in (0,\infty)$ beliebig war ergibt sich sodann $E[(X-Y)Z] = 0$. Daher folgt (c) aus (b).
Wir nehmen schließlich an, dass (c) gilt. Für $G \in \mathcal{G}$ gilt $\chi_G \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ und damit $E[X\chi_G] = E[Y\chi_G]$. Aus Satz 18.2.16 folgt nun $Y = E^{\mathcal{G}}(X)$. Daher folgt (a) aus (c). □

Aufgrund der Eigenschaft (b) aus Satz 18.3.3 wird die $\mathcal{G}$–bedingte Erwartung einer Zufallsvariablen $X \in L^2(\mathcal{F})$ auch als ihre *Projektion* auf $L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ bezeichnet.

Der Satz von Pythagoras besitzt eine weitere wichtige Anwendung, die unter anderem zeigt, dass die Varianz der bedingten Erwartung einer Zufallsvariablen nie größer ist als die Varianz der Zufallsvariablen:

18.3.4 Lemma (Varianz–Zerlegung). *Sei $X \in L^2(\mathcal{F})$. Dann gilt*

$$\text{var}\,[X] = E[\text{var}^{\mathcal{G}}(X)] + \text{var}\,[E^{\mathcal{G}}(X)]$$

Beweis. Nach dem Satz von Pythagoras gilt

$$\begin{aligned}\text{var}\,[X] &= E[(X - E[X])^2]\\ &= E[(X - E^{\mathcal{G}}(X))^2] + E[(E^{\mathcal{G}}(X) - E[X])^2]\\ &= E[E^{\mathcal{G}}((X - E^{\mathcal{G}}(X))^2)] + E[(E^{\mathcal{G}}(X) - E[E^{\mathcal{G}}(X)])^2]\\ &= E[\text{var}^{\mathcal{G}}(X)] + \text{var}\,[E^{\mathcal{G}}(X)]\end{aligned}$$

Damit ist die Gleichung gezeigt. □

Für eine Zufallsvariable $X \in L^2(\mathcal{F})$ bezeichnet man $E[\text{var}^{\mathcal{G}}(X)]$ als *Varianz in den Klassen* und $\text{var}\,[E^{\mathcal{G}}(X)]$ als *Varianz zwischen den Klassen*.

Aufgabe

18.3.A Kovarianz–Zerlegung: Für alle $X, Y \in L^2(\mathcal{F})$ gilt

$$\text{cov}\,[X,Y] = E[\text{cov}^{\mathcal{G}}(X,Y)] + \text{cov}\,[E^{\mathcal{G}}(X), E^{\mathcal{G}}(Y)]$$

und für jede Zufallsvariable $U \in L^2(\mathcal{F}) \cap L^0(\mathcal{F},\mathcal{G})$ gilt

$$\text{cov}\,[X,U] = \text{cov}\,[E^{\mathcal{G}}(X), U]$$

18.4 Martingale

Eine monoton wachsende Folge $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ von Unter–σ–Algebren von $\mathcal{F}$ heißt *Filtration*. Eine Filtration kann insbesondere durch eine Folge von Zufallsvariablen erzeugt werden:

18.4.1 Beispiel (Natürliche Filtration). Sei $\{X_k\}_{k\in\mathbb{N}}$ eine Folge von Zufallsvariablen. Dann ist die Folge $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ mit $\mathcal{F}_n := \sigma(\{X_k\}_{k\in\{1,\dots,n\}})$ für alle $n \in \mathbb{N}$ eine Filtration; diese Filtration wird als *natürliche Filtration* bezüglich der Folge $\{X_k\}_{k\in\mathbb{N}}$ bezeichnet.

Ist $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ eine Filtration, so ist das Mengensystem $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ eine Algebra, aber im allgemeinen keine σ–Algebra; vgl. Aufgabe 18.4.C.

Eine Folge von Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}}$ heißt
– *adaptiert* bezüglich einer Filtration $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, wenn für alle $n \in \mathbb{N}$ die Zufallsvariable X_n $\mathcal{F}_n$–messbar ist, und sie heißt
– *Martingal* bezüglich einer Filtration $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, wenn sie adaptiert ist und für alle $n \in \mathbb{N}$ die Zufallsvariable X_{n+1} $\mathcal{F}_n$–bedingt integrierbar ist mit $E^{\mathcal{F}_n}(X_{n+1}) = X_n$.

Jede Folge von Zufallsvariablen ist bezüglich ihrer natürlichen Filtration adaptiert. Das folgende Beispiel zeigt, dass viele Martingale durch eine einzige Zufallsvariable erzeugt werden können:

18.4.2 Beispiel. Sei $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ eine Filtration und sei X eine positive oder integrierbare Zufallsvariable. Dann ist die Folge $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ ein Martingal.

Eine Folge von Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}}$ heißt
– *positiv*, wenn jede der Zufallsvariablen X_n positiv ist, und sie heißt
– *integrierbar*, wenn jede der Zufallsvariablen X_n integrierbar ist.
Das folgende Lemma charakterisiert diejenigen positiven oder integrierbaren adaptierten Folgen von Zufallsvariablen, die ein Martingal sind:

18.4.3 Lemma. *Sei $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ eine Filtration und sei $\{X_n\}_{n\in\mathbb{N}}$ eine positive oder integrierbare adaptierte Folge von Zufallsvariablen. Dann sind folgende Aussagen äquivalent:*
(a) *$\{X_n\}_{n\in\mathbb{N}}$ ist ein Martingal.*
(b) *Für alle $n \in \mathbb{N}$ und $r \in \mathbb{N}(n)$ gilt $E^{\mathcal{F}_n}(X_r) = X_n$.*

Beweis. Wir nehmen an, dass (a) gilt. Wir betrachten $n \in \mathbb{N}$ und zeigen durch vollständige Induktion, dass für alle $r \in \mathbb{N}(n)$ die Gleichung $E^{\mathcal{F}_n}(X_r) = X_n$ gilt.
- $r = n$: In diesem Fall ist nichts zu zeigen.
- $r \to r+1$: Wir nehmen an, die Gleichung sei für ein $r \in \mathbb{N}(n)$ bereits bewiesen. Dann ergibt sich aus Satz 18.2.8 und der Martingal–Eigenschaft

$$E^{\mathcal{F}_n}(X_{r+1}) = E^{\mathcal{F}_n}(E^{\mathcal{F}_r}(X_{r+1})) = E^{\mathcal{F}_n}(X_r) = X_n$$

Daher folgt (b) aus (a). Die umgekehrte Implikation ist klar. □

Im weiteren Verlauf dieses Abschnitts sei $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ eine Filtration und $\mathcal{F}_\infty$ die von der Algebra $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ erzeugte σ–Algebra.

Wir klären zunächst die Struktur der Familie aller integrierbaren $\mathcal{F}_\infty$–messbaren Zufallsvariablen:

18.4.4 Satz (Approximationssatz). *Sei Y eine integrierbare Zufallsvariable mit $Y \in L^0(\mathcal{F}, \mathcal{F}_\infty)$. Dann gibt es zu jedem $\eta \in (0,\infty)$ eine integrierbare Zufallsvariable Z mit $Z \in L^0(\mathcal{F}, \mathcal{F}_l)$ für ein $l \in \mathbb{N}$ und $E[|Y-Z|] \le \eta$.*

Beweis. Mit Y sind auch Y^+ und Y^- integrierbar und $\mathcal{F}_\infty$–messbar. Wir können daher annehmen, dass Y positiv ist.
Da Y $\mathcal{F}_\infty$–messbar und positiv ist, gibt es nach dem Approximationssatz 7.1.13 eine monoton wachsenden Folge $\{Y_n\}_{n\in\mathbb{N}}$ von einfachen $\mathcal{F}_\infty$–messbaren Zufallsvariablen mit $Y = \sup_{n\in\mathbb{N}} Y_n$ und damit $E[Y] = \sup_{n\in\mathbb{N}} E[Y_n]$. Da Y integrierbar ist, gilt $\lim_{n\to\infty} E[|Y-Y_n|] = 0$. Wir können daher annehmen, dass Y einfach ist.
Da Y $\mathcal{F}_\infty$–messbar und einfach ist und da $\mathcal{F}_\infty$ die von der Algebra $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ erzeugte σ–Algebra ist, ergibt sich aus dem Approximationssatz 5.4.1 die Existenz einer einfachen Zufallsvariablen Z mit $Z \in L^0(\mathcal{F}, \mathcal{F}_l)$ für ein $l \in \mathbb{N}$ und $E[|Y-Z|] \le \eta$. □

Martingale sind unter anderem im Hinblick auf Konvergenzsätze von Interesse. Unser Ziel ist es nun, einen Konvergenzsatz für Martingale zu beweisen, die von einer quadratisch integrierbaren Zufallsvariablen erzeugt werden. Dazu benötigen wir das folgende Lemma, das, unter den gegebenen Bedingungen, als eine Verschärfung der Ungleichung von Markov verstanden werden kann:

18.4.5 Lemma (Maximale Ungleichung). *Sei $\{X_n\}_{n\in\mathbb{N}}$ ein positives Martingal. Dann gilt für alle $m \in \mathbb{N}$ und $\varepsilon \in (0,\infty)$*

$$P\left[\left\{\sup_{n\in\mathbb{N}(m)} X_n > \varepsilon\right\}\right] \le \frac{1}{\varepsilon}\, E[X_m]$$

Beweis. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $m = 1$ gilt. Für alle $n \in \mathbb{N}$ sei

$$A_n := \{X_n > \varepsilon\} \cap \bigcap_{k=1}^{n-1} \{X_k \le \varepsilon\}$$

Dann ist die Folge $\{A_n\}_{n\in\mathbb{N}}$ disjunkt mit $\{\sup_{n\in\mathbb{N}} X_n > \varepsilon\} = \sum_{n=1}^\infty A_n$ sowie $A_n \in \mathcal{F}_n$ für alle $n \in \mathbb{N}$. Für $r \in \mathbb{N}$ ergibt sich aus Lemma 18.4.3

$$\begin{aligned}
\varepsilon \sum_{n=1}^r P[A_n] &= \sum_{n=1}^r \int_{A_n} \varepsilon\, dP \\
&\le \sum_{n=1}^r \int_{A_n} X_n\, dP \\
&= \sum_{n=1}^r \int_{A_n} E^{\mathcal{F}_n}(X_r)\, dP \\
&= \sum_{n=1}^r \int_{A_n} X_r\, dP \\
&\le \int_\Omega X_r\, dP \\
&= \int_\Omega X_1\, dP
\end{aligned}$$

Daraus ergibt sich

$$\begin{aligned}
P\left[\left\{\sup_{n\in\mathbb{N}} X_n > \varepsilon\right\}\right] &= \sum_{n=1}^\infty P[A_n] \\
&\le \frac{1}{\varepsilon} \int_\Omega X_1\, dP
\end{aligned}$$

und damit die gewünschte Ungleichung. □

Der folgende Satz von Lévy ist der einfachste Konvergenzsatz für Martingale:

18.4.6 Satz (Lévy). *Sei X eine Zufallsvariable in $L^2(\mathcal{F})$. Dann gilt*

$$\lim_{n\to\infty} E^{\mathcal{F}_n}(X) = E^{\mathcal{F}_\infty}(X)$$

fast sicher und in $L^2(\mathcal{F})$.

Beweis. Sei $\varepsilon \in (0,\infty)$. Dann ist die Folge

$$\left\{P\left[\left\{\sup_{n\in\mathbb{N}(m)} |E^{\mathcal{F}_n}(X)-E^{\mathcal{F}_\infty}(X)| > \varepsilon\right\}\right]\right\}_{m\in\mathbb{N}}$$

monoton fallend und damit konvergent. Für den Beweis der fast sicheren Konvergenz des Martingals $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ gegen $E^{\mathcal{F}_\infty}(X)$ genügt es zu zeigen, dass die betrachtete Folge gegen 0 konvergiert.
Sei $Z \in L^2(\mathcal{F})$ eine Zufallsvariable mit $Z \in L^0(\mathcal{F},\mathcal{F}_l)$ für ein $l \in \mathbb{N}$. Dann ist die Folge $\{E^{\mathcal{F}_n}(|X-Z|)\}_{n\in\mathbb{N}}$ ein positives Martingal und für alle $n \in \mathbb{N}(l)$ gilt

$$\begin{aligned}
|E^{\mathcal{F}_n}(X)-E^{\mathcal{F}_\infty}(X)| &\le |E^{\mathcal{F}_n}(X)-Z)| + |E^{\mathcal{F}_\infty}(X)-Z| \\
&\le |E^{\mathcal{F}_n}(E^{\mathcal{F}_\infty}(X)-Z)| + |E^{\mathcal{F}_\infty}(X)-Z| \\
&\le E^{\mathcal{F}_n}|E^{\mathcal{F}_\infty}(X)-Z| + |E^{\mathcal{F}_\infty}(X)-Z|
\end{aligned}$$

Für alle $m \in \mathbb{N}(l)$ ergibt sich durch Übergang zum Supremum über $n \in \mathbb{N}(m)$ und unter Verwendung der maximalen Ungleichung und der Ungleichung von Markov

$$\begin{aligned}
&P\left[\left\{\sup_{n\in\mathbb{N}(m)} |E^{\mathcal{F}_n}(X)-E^{\mathcal{F}_\infty}(X)| > \varepsilon\right\}\right] \\
&\le P\left[\left\{\sup_{n\in\mathbb{N}(m)} E^{\mathcal{F}_n}(|E^{\mathcal{F}_\infty}(X)-Z|) > \varepsilon/2\right\}\right] + P\left[\left\{|E^{\mathcal{F}_\infty}(X)-Z| > \varepsilon/2\right\}\right] \\
&\le \frac{2}{\varepsilon}\, E[E^{\mathcal{F}_n}(|E^{\mathcal{F}_\infty}(X)-Z|)] + \frac{2}{\varepsilon}\, E[|E^{\mathcal{F}_\infty}(X)-Z|] \\
&= \frac{4}{\varepsilon}\, E[|E^{\mathcal{F}_\infty}(X)-Z|]
\end{aligned}$$

Daher gilt

$$\lim_{m\to\infty} P\left[\left\{\sup_{n\in\mathbb{N}(m)} |E^{\mathcal{F}_n}(X)-E^{\mathcal{F}_\infty}(X)| > \varepsilon\right\}\right] \le \frac{4}{\varepsilon}\, E[|E^{\mathcal{F}_\infty}(X)-Z|]$$

Nach dem Approximationssatz 18.4.4 gibt es für alle $\eta \in (0,\infty)$ eine integrierbare Zufallsvariable Z mit $Z \in L^0(\mathcal{F},\mathcal{F}_l)$ für ein $l \in \mathbb{N}$ und

$$E[|E^{\mathcal{F}_\infty}(X)-Z|] \le \frac{\eta\,\varepsilon}{4}$$

Daher gilt

$$\lim_{m\to\infty} P\left[\left\{\sup_{n\in\mathbb{N}(m)} |E^{\mathcal{F}_n}(X)-E^{\mathcal{F}_\infty}(X)| > \varepsilon\right\}\right] \le \eta$$

Da $\eta \in (0,\infty)$ beliebig war, gilt

$$\lim_{m\to\infty} P\left[\left\{\sup_{n\in\mathbb{N}(m)} |E^{\mathcal{F}_n}(X)-E^{\mathcal{F}_\infty}(X)| > \varepsilon\right\}\right] = 0$$

Damit ist gezeigt, dass das Martingal $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ fast sicher gegen $E^{\mathcal{F}_\infty}(X)$ konvergiert.
Sei $m \in \mathbb{N}$. Für alle $k \in \mathbb{N}(m)$ erhält man aus dem Satz von Pythagoras wegen $E^{\mathcal{F}_k}(E^{\mathcal{F}_{k+1}}(X)) = E^{\mathcal{F}_k}(X)$

$$\begin{aligned}
&E[(E^{\mathcal{F}_{k+1}}(X)-E^{\mathcal{F}_m}(X))^2]\\
&= E[(E^{\mathcal{F}_{k+1}}(X)-E^{\mathcal{F}_k}(X))^2] + E[(E^{\mathcal{F}_k}(X)-E^{\mathcal{F}_m}(X))^2]
\end{aligned}$$

Für alle $n \in \mathbb{N}(m)$ ergibt sich nun durch vollständige Induktion

$$E[(E^{\mathcal{F}_{n+1}}(X)-E^{\mathcal{F}_m}(X))^2] = \sum_{k=m}^{n} E[(E^{\mathcal{F}_{k+1}}(X)-E^{\mathcal{F}_k}(X))^2]$$

und aus Lemma 18.4.3 und dem Satz von Pythagoras folgt sodann

$$\begin{aligned}
\sum_{k=m}^{n} E[(E^{\mathcal{F}_{k+1}}(X)-E^{\mathcal{F}_k}(X))^2] &= E[(E^{\mathcal{F}_{n+1}}(X)-E^{\mathcal{F}_m}(X))^2]\\
&= E[(E^{\mathcal{F}_{n+1}}(X)-E^{\mathcal{F}_m}(E^{\mathcal{F}_{n+1}}(X)))^2]\\
&\le E[(E^{\mathcal{F}_{n+1}}(X))^2]\\
&\le E[X^2]
\end{aligned}$$

Daher gilt

$$\sum_{k=m}^{\infty} E[(E^{\mathcal{F}_{k+1}}(X)-E^{\mathcal{F}_k}(X))^2] \le E[X^2]$$

Aus der quadratischen Integrierbarkeit von X folgt nun wegen

$$\begin{aligned}
E[(E^{\mathcal{F}_{n+1}}(X)-E^{\mathcal{F}_m}(X))^2] &= \sum_{k=m}^{n} E[(E^{\mathcal{F}_{k+1}}(X)-E^{\mathcal{F}_k}(X))^2]\\
&\le \sum_{k=m}^{\infty} E[(E^{\mathcal{F}_{k+1}}(X)-E^{\mathcal{F}_k}(X))^2]
\end{aligned}$$

dass die Folge $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ eine Cauchy–Folge in $L^2(\mathcal{F})$ ist. Da $L^2(\mathcal{F})$ vollständig ist, konvergiert die Folge $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ in $L^2(\mathcal{F})$ gegen eine Zufallsvariable $Y \in L^2(\mathcal{F})$. Insbesondere besitzt die Folge $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ eine

Teilfolge, die fast sicher gegen Y konvergiert. Da vorher bereits gezeigt wurde, dass die Folge $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ fast sicher gegen $E^{\mathcal{F}_\infty}(X)$ konvergiert, gilt $Y = E^{\mathcal{F}_\infty}(X)$. Daher konvergiert die Folge $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ auch in $L^2(\mathcal{F})$ gegen $E^{\mathcal{F}_\infty}(X)$. □

Der Satz von Lévy besitzt eine interessante Folgerung:

18.4.7 Folgerung. *Sei $\{X_n\}_{n\in\mathbb{N}}$ ein Martingal in $L^2(\mathcal{F})$. Wenn es eine Zufallsvariable $X \in L^2(\mathcal{F})$ gibt mit*

$$\lim_{n\to\infty} E[(X_n - X)^2] = 0$$

dann ist X $\mathcal{F}_\infty$–messbar, für alle $n \in \mathbb{N}$ gilt $X_n = E^{\mathcal{F}_n}(X)$, und es gilt

$$\lim_{n\to\infty} X_n = X$$

fast sicher.

Beweis. Da $\{X_n\}_{n\in\mathbb{N}}$ ein Martingal ist, gilt

$$\begin{aligned} E[(X_n - E^{\mathcal{F}_n}(X))^2] &= E[(E^{\mathcal{F}_n}(X_{n+1}) - E^{\mathcal{F}_n}(E^{\mathcal{F}_{n+1}}(X)))^2] \\ &= E[(E^{\mathcal{F}_n}\big(X_{n+1} - E^{\mathcal{F}_{n+1}}(X)\big))^2] \\ &\leq E[E^{\mathcal{F}_n}(X_{n+1} - E^{\mathcal{F}_{n+1}}(X))^2] \\ &= E[(X_{n+1} - E^{\mathcal{F}_{n+1}}(X))^2] \end{aligned}$$

Daher ist die Folge $\{E[(X_n - E^{\mathcal{F}_n}(X))^2]\}_{n\in\mathbb{N}}$ monoton wachsend.
Da die Folge $\{X_n\}_{n\in\mathbb{N}}$ in $L^2(\mathcal{F})$ gegen X konvergiert, besitzt sie eine Teilfolge, die fast sicher gegen X konvergiert, und aus $\{X_n\}_{n\in\mathbb{N}} \subseteq L^0(\mathcal{F}, \mathcal{F}_\infty)$ ergibt sich nun $X \in L^0(\mathcal{F}, \mathcal{F}_\infty)$, also $X = E^{\mathcal{F}_\infty}(X)$, und damit

$$\begin{aligned} E[(X_n - E^{\mathcal{F}_n}(X))^2] &\leq E[(X_n - X)^2] + E[(X - E^{\mathcal{F}_n}(X))^2] \\ &= E[(X_n - X)^2] + E[(E^{\mathcal{F}_\infty}(X) - E^{\mathcal{F}_n}(X))^2] \end{aligned}$$

Nach Voraussetzung konvergiert die Folge $\{X_n\}_{n\in\mathbb{N}}$ in $L^2(\mathcal{F})$ gegen X, und nach dem Satz von Lévy konvergiert die Folge $\{E^{\mathcal{F}_n}(X)\}_{n\in\mathbb{N}}$ in $L^2(\mathcal{F})$ gegen $E^{\mathcal{F}_\infty}(X)$. Daher konvergiert die Folge $\{E[(X_n - E^{\mathcal{F}_n}(X))^2]\}_{n\in\mathbb{N}}$ gegen 0.
Aus diesen Eigenschaften der Folge $\{E[(X_n - E^{\mathcal{F}_n}(X))^2]\}_{n\in\mathbb{N}}$ ergibt sich daher für alle $n \in \mathbb{N}$

$$X_n = E^{\mathcal{F}_n}(X)$$

und aus dem Satz von Lévy folgt nun, dass die Folge $\{X_n\}_{n\in\mathbb{N}}$ auch fast sicher gegen X konvergiert. □

Aufgaben

18.4.A **Natürliche Filtration:** Jede Folge von Zufallsvariablen ist bezüglich ihrer natürlichen Filtration adaptiert.

18.4.B **Filtration:** Sei $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ eine Filtration. Dann ist das Mengensystem $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ eine Algebra.

18.4.C **Filtration:** Konstruieren Sie eine Filtration $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, für die die Algebra $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ *keine* σ–Algebra ist.
Hinweis: Betrachten Sie die natürliche Filtration einer geeigneten Folge $\{X_n\}_{n\in\mathbb{N}}$ von Zufallsvariablen.

19

Bedingte Wahrscheinlichkeit und bedingte Verteilung

Mit der Verfügbarkeit des Begriffs der bedingten Erwartung lässt sich nun die bedingte Wahrscheinlichkeit eines Ereignisses durch die bedingte Erwartung seiner Indikatorfunktion definieren; man beschreitet damit den in der Integrationstheorie üblichen Weg vom Maß zum Integral in umgekehrter Richtung.

Wir untersuchen zunächst die elementaren Eigenschaften bedingter Wahrscheinlichkeiten bezüglich einer Unter–σ–Algebra und stellen insbesondere den Zusammenhang mit den bereits bekannten bedingten Wahrscheinlichkeiten bezüglich einem Ereignis dar (Abschnitt 19.1). Aus dem Begriff der bedingten Wahrscheinlichkeit ergibt sich zunächst in natürlicher Weise der Begriff der bedingten Unabhängigkeit, der sich in der Konstruktion wahrscheinlichkeitstheoretischer Modelle als überaus nützlich erweist (Abschnitt 19.2); dies gilt auch für die bedingte Verteilung einer Zufallsvariablen (Abschnitt 19.3) und für bedingte Dichten einer bedingten Verteilung (Abschnitt 19.4). Den Abschluss dieses Kapitels bilden bedingte Versionen der Gesetze der Großen Zahlen.

Im gesamten Kapitel sei $\mathcal{G}$ eine Unter–σ–Algebra von $\mathcal{F}$.

19.1 Bedingte Wahrscheinlichkeit

Für ein Ereignis $A \in \mathcal{F}$ heißt die Zufallsvariable

$$P^{\mathcal{G}}(A) := E^{\mathcal{G}}(\chi_A)$$

die *$\mathcal{G}$–bedingte Wahrscheinlichkeit* von A. Aus dieser Definition ergeben sich sofort zwei Spezialfälle bezüglich der σ–Algebra $\mathcal{G}$:

– Im Fall $\mathcal{G} = \{\emptyset, \Omega\}$ gilt $P^{\mathcal{G}}(A) = P[A]$.
– Im Fall $\mathcal{G} = \mathcal{F}$ gilt $P^{\mathcal{G}}(A) = \chi_A$.

Das folgende Lemma fasst die elementaren Eigenschaften der bedingten Wahrscheinlichkeit zusammen:

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_19, © Springer-Verlag Berlin Heidelberg 2011

19.1.1 Lemma.
(1) *Es gilt* $P^{\mathcal{G}}(\Omega) = 1$ *und* $P^{\mathcal{G}}(\emptyset) = 0$.
(2) *Für alle* $A \in \mathcal{F}$ *gilt* $0 \leq P^{\mathcal{G}}(A) \leq 1$.
(3) *Für jede disjunkte Folge* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ *gilt*

$$P^{\mathcal{G}}\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P^{\mathcal{G}}(A_n)$$

(4) *Für alle* $A \in \mathcal{F}$ *und* $G \in \mathcal{G}$ *gilt* $P^{\mathcal{G}}(A \cap G) = P^{\mathcal{G}}(A)\,\chi_G$.
(5) *Für alle* $G \in \mathcal{G}$ *gilt* $P^{\mathcal{G}}(G) = \chi_G$.

Beweis. Wegen $\Omega \in \mathcal{G}$ und $\emptyset \in \mathcal{G}$ gilt $E^{\mathcal{G}}(\chi_\Omega) = \chi_\Omega = 1$ und $E^{\mathcal{G}}(\chi_\emptyset) = \chi_\emptyset = 0$. Für $A \in \mathcal{F}$ gilt $\chi_\emptyset \leq \chi_A \leq \chi_\Omega$ und aus Lemma 18.1.5 folgt $0 \leq E^{\mathcal{G}}(\chi_A) \leq 1$. Ist $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ eine disjunkte Folge, so folgt aus dem bedingten Satz über die monotone Konvergenz in der Form von Folgerung 18.1.7

$$E^{\mathcal{G}}(\chi_{\sum_{n=1}^{\infty} A_n}) = E^{\mathcal{G}}\left(\sum_{n=1}^{\infty} \chi_{A_n}\right) = \sum_{n=1}^{\infty} E^{\mathcal{G}}(\chi_{A_n})$$

Für alle $A \in \mathcal{F}$ und $G \in \mathcal{G}$ gilt $E^{\mathcal{G}}(\chi_{A\cap G}) = E^{\mathcal{G}}(\chi_A \chi_G) = E^{\mathcal{G}}(\chi_A)\,\chi_G$ und damit gilt insbesondere $E^{\mathcal{G}}(\chi_G) = \chi_G$. □

Da $\mathcal{G}$–bedingte Wahrscheinlichkeiten $\mathcal{G}$–bedingte Erwartungen von Indikatorfunktionen sind, gelten Gleichungen und Ungleichungen für $\mathcal{G}$–bedingte Wahrscheinlichkeiten im allgemeinen nur fast sicher. Insbesondere gilt die Gleichung

$$P^{\mathcal{G}}\left(\sum_{n=1}^{\infty} A_n\right)(\omega) = \sum_{n=1}^{\infty} P^{\mathcal{G}}(A_n)(\omega)$$

im allgemeinen nur auf dem Komplement einer Nullmenge, die von der disjunkten Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ abhängt; vgl. Abschnitt 19.3.

Das folgende Ergebnis ist eine unmittelbare Folgerung aus Lemma 18.1.2; es zeigt, dass die Definition der $\mathcal{G}$–bedingten Wahrscheinlichkeit eines Ereignisses mit der Definition seiner bedingten Wahrscheinlichkeit bezüglich einem Ereignis $G \in \mathcal{G}$ im Einklang steht:

19.1.2 Lemma (Fourier–Entwicklung). *Sei* $\{G_i\}_{i\in\{1,\ldots,m\}} \subseteq \mathcal{F}$ *eine disjunkte Familie von Ereignissen mit* $\sum_{i=1}^m G_i = \Omega$ *und* $P[G_i] > 0$ *für alle* $i \in \{1,\ldots,m\}$ *und sei* $\mathcal{G}$ *die von der Familie* $\{G_i\}_{i\in\{1,\ldots,m\}}$ *erzeugte* σ*–Algebra. Dann gilt für jedes Ereignis* $A \in \mathcal{F}$

$$P^{\mathcal{G}}(A) = \sum_{i=1}^{m} P[A|G_i]\,\chi_{G_i}$$

Wir untersuchen nun die Eigenschaften der bedingten Wahrscheinlichkeit. Genau wie im unbedingten Fall zeigt man, dass die bedingte Wahrscheinlichkeit monoton und σ–subadditiv ist:

19.1.3 Lemma. *Für alle $A, B \in \mathcal{F}$ mit $A \subseteq B$ gilt*

$$P^{\mathcal{G}}(A) \leq P^{\mathcal{G}}(B)$$

und für jede Folge von Ereignissen $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ gilt

$$P^{\mathcal{G}}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P^{\mathcal{G}}(A_n)$$

Auch das folgende Lemma lässt sich genau wie im unbedingten Fall elementar beweisen; vgl. Aufgabe 12.3.A:

19.1.4 Lemma. *Sei X eine Zufallsvariable. Dann gilt*

$$\sum_{n=1}^{\infty} P^{\mathcal{G}}(\{|X| \geq n\}) \leq E^{\mathcal{G}}(|X|) \leq 1 + \sum_{n=1}^{\infty} P^{\mathcal{G}}(\{|X| \geq n\})$$

Insbesondere ist X genau dann $\mathcal{G}$–bedingt integrierbar, wenn die Reihe $\sum_{n=1}^{\infty} P^{\mathcal{G}}(\{|X| \geq n\})$ fast sicher gegen eine reelle Zufallsvariable konvergiert.

Aus diesem Lemma erhält man sofort die bedingte Ungleichung von Markov:

19.1.5 Folgerung (Ungleichung von Markov). *Sei X eine Zufallsvariable. Dann gilt für alle $\varepsilon \in (0, \infty)$*

$$P^{\mathcal{G}}(\{|X| \geq \varepsilon\}) \leq \frac{1}{\varepsilon}\, E^{\mathcal{G}}(|X|)$$

Schließlich lässt sich auch das bedingte 1. Lemma von Borel/Cantelli genau wie im unbedingten Fall beweisen:

19.1.6 Lemma (1. Lemma von Borel/Cantelli). *Sei $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ eine Folge von Ereignissen und sei $G := \{\sum_{n=1}^{\infty} P^{\mathcal{G}}(A_n) < \infty\}$. Dann gilt*

$$P^{\mathcal{G}}\left(\limsup_{n\to\infty} A_n\right) \chi_G = 0$$

Wir betrachten abschließend den Übergang von bedingten zu unbedingten Wahrscheinlichkeiten:

19.1.7 Lemma.
(1) *Für jedes Ereignis $A \in \mathcal{F}$ gilt*

$$P[A] = E[P^{\mathcal{G}}(A)]$$

(2) *Sei $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ eine Folge von Ereignissen mit*

$$\lim_{n\to\infty} P^{\mathcal{G}}(A_n) = 0$$

fast sicher. Dann gilt

$$\lim_{n\to\infty} P[A_n] = 0$$

Beweis. Für jedes Ereignis $A \in \mathcal{F}$ gilt

$$P[A] = E[\chi_A] = E[E^{\mathcal{G}}(\chi_A)] = E[P^{\mathcal{G}}(A)]$$

Damit ist (1) gezeigt. Wegen $\sup_{n\in\mathbb{N}} |P^{\mathcal{G}}(A_n)| \le 1 \in L^1(\mathcal{F})$ folgt nun (2) aus dem Satz über die majorisierte Konvergenz. □

19.2 Bedingte Unabhängigkeit

Sei I eine nichtleere Indexmenge und sei $\mathcal{H}(I)$ die Familie der endlichen nichtleeren Teilmengen von I.

Eine Familie von Ereignissen $\{A_i\}_{i\in I} \subseteq \mathcal{F}$ heißt $\mathcal{G}$*–bedingt unabhängig*, wenn für alle $J \in \mathcal{H}(I)$

$$P^{\mathcal{G}}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} P^{\mathcal{G}}(A_i)$$

gilt. Aus dieser Definition ergeben sich sofort zwei Spezialfälle bezüglich der σ–Algebra $\mathcal{G}$:
– Im Fall $\mathcal{G} = \{\emptyset, \Omega\}$ ist eine Familie von Ereignissen genau dann $\mathcal{G}$–bedingt unabhängig, wenn sie unabhängig ist.
– Im Fall $\mathcal{G} = \mathcal{F}$ ist jede Familie von Ereignissen $\mathcal{G}$–bedingt unabhängig.
Obwohl die bedingte Unabhängigkeit analog zur Unabhängigkeit definiert ist, impliziert im allgemeinen keine dieser Eigenschaften die andere:

19.2.1 Beispiele. Sei Θ eine reelle Zufallsvariable mit $P_\Theta = \mathbf{B}(1/2)$.
(1) Sind $A, B \in \mathcal{F}$ Ereignisse mit

$$\begin{aligned} P^{\sigma(\Theta)}(A \cap B) &= (1{+}3\Theta)/9 \\ P^{\sigma(\Theta)}(A \cap \overline{B}) &= 2/9 \\ P^{\sigma(\Theta)}(\overline{A} \cap B) &= 2/9 \\ P^{\sigma(\Theta)}(\overline{A} \cap \overline{B}) &= (4{-}3\Theta)/9 \end{aligned}$$

so gilt $P^{\sigma(\Theta)}(A) = (1{+}\Theta)/3 = P^{\sigma(\Theta)}(B)$ und wegen $\Theta^2 = \Theta$ folgt daraus

$$P^{\sigma(\Theta)}(A \cap B) = (1{+}3\Theta)/9 = P^{\sigma(\Theta)}(A)\, P^{\sigma(\Theta)}(B)$$

Daher ist $\{A, B\}$ $\sigma(\Theta)$–bedingt unabhängig. Andererseits gilt

$$P[A \cap B] = E[P^{\sigma(\Theta)}(A \cap B)] = E[(1{+}3\Theta)/9] = 5/18$$

sowie $P[A] = 1/2 = P[B]$ und damit

$$P[A \cap B] = 5/18 \ne 1/4 = P[A]\, P[B]$$

Daher ist $\{A, B\}$ nicht unabhängig.

(2) Sind $A, B \in \mathcal{F}$ Ereignisse mit

$$\begin{aligned} P^{\sigma(\Theta)}(A \cap B) &= (1-\Theta)/2 \\ P^{\sigma(\Theta)}(A \cap \overline{B}) &= \Theta/2 \\ P^{\sigma(\Theta)}(\overline{A} \cap B) &= \Theta/2 \\ P^{\sigma(\Theta)}(\overline{A} \cap \overline{B}) &= (1-\Theta)/2 \end{aligned}$$

so gilt $P^{\sigma(\Theta)}(A) = 1/2 = P^{\sigma(\Theta)}(B)$ und damit

$$P^{\sigma(\Theta)}(A \cap B) = (1-\Theta)/2 \neq 1/4 = P^{\sigma(\Theta)}(A)\, P^{\sigma(\Theta)}(B)$$

Daher ist $\{A, B\}$ nicht $\sigma(\Theta)$–bedingt unabhängig. Andererseits gilt

$$P[A \cap B] = E[P^{\sigma(\Theta)}(A \cap B)] = E[(1-\Theta)/2] = 1/4$$

sowie $P[A] = 1/2 = P[B]$ und damit

$$P[A \cap B] = 1/4 = P[A]\, P[B]$$

Daher ist $\{A, B\}$ unabhängig.

Dennoch lassen sich die in Abschnitt 11.1 bewiesenen Ergebnisse über die Unabhängigkeit einer Familie von Ereignissen im wesentlichen auf die bedingte Unabhängigkeit übertragen. Als Beispiel geben wir hier nur das bedingte Null–Eins–Gesetz von Borel an:

19.2.2 Satz (Null–Eins–Gesetz; Borel). *Sei $\{A_n\}_{n\in\mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige Folge von Ereignissen und sei $G := \{\sum_{n=1}^{\infty} P^{\mathcal{G}}(A_n) < \infty\}$. Dann gilt*

$$P^{\mathcal{G}}\left(\limsup_{n\to\infty} A_n\right) = \chi_{\Omega\setminus G}$$

Das bedingte Null–Eins–Gesetz von Borel ist ein erstes Beispiel dafür, dass bei der Verallgemeinerung von Ergebnissen über die Unabhängigkeit einer Familie von Ereignissen auf die bedingte Unabhängigkeit unter Umständen Anpassungen in der Formulierung und damit auch in der Beweisführung angebracht sind.

Die wichtigsten Ergebnisse über die bedingte Unabhängigkeit einer Familie von Ereignissen sind in den Ergebnissen über die bedingte Unabhängigkeit einer Familie von Ereignissystemen enthalten.

Eine Familie von Ereignissystemen $\{\mathcal{E}_i\}_{i\in I} \subseteq \mathcal{F}$ heißt *$\mathcal{G}$–bedingt unabhängig*, wenn jede Familie von Ereignissen $\{A_i\}_{i\in I} \subseteq \mathcal{F}$ mit $A_i \in \mathcal{E}_i$ für alle $i \in I$ $\mathcal{G}$–bedingt unabhängig ist. Wir geben nun die wichtigsten Ergebnisse über die $\mathcal{G}$–bedingte Unabhängigkeit einer Familie von Ereignissystemen an, wobei wir auf die Angabe von Lemmata, deren Beweis wie im Fall der Unabhängigkeit geführt werden kann, verzichten.

19.2.3 Satz. *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine Familie von Ereignissystemen. Dann sind äquivalent*:
(a) *Die Familie $\{\mathcal{E}_i\}_{i\in I}$ ist $\mathcal{G}$–bedingt unabhängig.*
(b) *Für jede nichtleere Menge $K \subseteq I$ ist $\{\mathcal{E}_i\}_{i\in K}$ $\mathcal{G}$–bedingt unabhängig.*
(c) *Für jede endliche nichtleere Menge $K \subseteq I$ ist $\{\mathcal{E}_i\}_{i\in K}$ $\mathcal{G}$–bedingt unabhängig.*

Der Beweis des Satzes verläuft analog zum Beweis von Satz 11.2.1.

19.2.4 Lemma. *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine Familie von Ereignissystemen und sei $\{\mathcal{C}_i\}_{i\in I}$ eine Familie von Ereignissystemen mit $\mathcal{C}_i \subseteq \mathcal{E}_i$ für alle $i \in I$. Ist $\{\mathcal{E}_i\}_{i\in I}$ $\mathcal{G}$–bedingt unabhängig, so ist auch $\{\mathcal{C}_i\}_{i\in I}$ $\mathcal{G}$–bedingt unabhängig.*

Der folgende Satz klärt den Zusammenhang zwischen der bedingten Unabhängigkeit von σ–Algebren und der bedingten Unabhängigkeit ihrer Erzeuger; er enthält unter anderem eine äußerst nützliche Äquivalenz, die im Fall der Unabhängigkeit trivial und daher nicht sichtbar ist:

19.2.5 Satz. *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine Familie von $\cap$–stabilen Ereignissystemen. Dann sind äquivalent*:
(a) *Die Familie $\{\mathcal{E}_i\}_{i\in I}$ ist $\mathcal{G}$–bedingt unabhängig.*
(b) *Die Familie $\{\sigma(\mathcal{E}_i)\}_{i\in I}$ ist $\mathcal{G}$–bedingt unabhängig.*
(c) *Die Familie $\{\sigma(\mathcal{E}_i \cup \mathcal{G})\}_{i\in I}$ ist $\mathcal{G}$–bedingt unabhängig.*

Beweis. Der Beweis der Äquivalenz von (a) und (b) verläuft analog zum Beweis von Satz 11.2.5. Aufgrund von Lemma 19.2.4 ist außerdem klar, dass (b) aus (c) folgt.
Wir nehmen nun an, dass (b) gilt. Für alle $i \in I$ ist das Ereignissystem

$$\mathcal{H}_i := \Big\{ B \in \mathcal{F} \,\Big|\, B = A \cap G \text{ mit } A \in \sigma(\mathcal{E}_i) \text{ und } G \in \mathcal{G} \Big\}$$

$\cap$–stabil mit $\sigma(\mathcal{H}_i) = \sigma(\mathcal{E}_i \cup \mathcal{G})$. Für $J \subseteq \mathcal{H}(I)$ und jede Wahl von $\{A_i\}_{i\in J}$ und $\{G_i\}_{i\in J}$ mit $A_i \in \sigma(\mathcal{E}_i)$ und $G_i \in \mathcal{G}$ für alle $i \in J$ gilt

$$\begin{aligned}
P^{\mathcal{G}}\left(\bigcap_{i\in J}(A_i \cap G_i)\right) &= P^{\mathcal{G}}\left(\left(\bigcap_{i\in J} A_i\right) \cap \left(\bigcap_{i\in J} G_i\right)\right) \\
&= P^{\mathcal{G}}\left(\bigcap_{i\in J} A_i\right) \chi_{\bigcap_{i\in J} G_i} \\
&= \prod_{i\in J} P^{\mathcal{G}}(A_i) \cdot \prod_{i\in J} \chi_{G_i} \\
&= \prod_{i\in J} P^{\mathcal{G}}(A_i \cap G_i)
\end{aligned}$$

Daher ist die Familie $\{\mathcal{H}_i\}_{i\in I}$ und damit auch die Familie $\{\sigma(\mathcal{H}_i)\}_{i\in I}$, also die Familie $\{\sigma(\mathcal{E}_i \cup \mathcal{G})\}_{i\in I}$ $\mathcal{G}$–bedingt unabhängig. Daher folgt (c) aus (b). $\square$

Schließlich besitzt auch das Blocklemma eine bedingte Version:

19.2.6 Lemma (Blocklemma). *Sei $\{\mathcal{E}_i\}_{i\in I}$ eine $\mathcal{G}$–bedingt unabhängige Familie von $\cap$–stabilen Ereignissystemen. Sind $M,N \subseteq I$ disjunkt, so ist auch die Familie*

$$\left\{\sigma\left(\bigcup_{i\in M}\mathcal{E}_i\right),\sigma\left(\bigcup_{i\in N}\mathcal{E}_i\right)\right\}$$

$\mathcal{G}$–bedingt unabhängig.

Der Beweis des Blocklemmas verläuft analog zum Beweis von Lemma 11.2.7.

Als letztes Ergebnis zur bedingten Unabhängigkeit von Ereignissystemen beweisen wir eine bedingte Version des Null–Eins–Gesetzes von Kolmogorov:

19.2.7 Satz (Null–Eins–Gesetz; Kolmogorov). *Sei $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige Folge von $\cap$–stabilen Ereignissystemen und sei $\mathcal{E}_\infty$ die zugehörige terminale σ–Algebra. Dann gilt:*
(1) *Für alle $A,B \in \mathcal{E}_\infty$ gilt $P^\mathcal{G}(A\cap B) = P^\mathcal{G}(A)\,P^\mathcal{G}(B)$.*
(2) *Für alle $A \in \mathcal{E}_\infty$ gilt $P^\mathcal{G}(A) = \chi_A$ und es gibt ein Ereignis $G \in \mathcal{G}$ mit $\chi_A = \chi_G$.*
(3) *Zu jeder $\mathcal{E}_\infty$–messbaren numerischen Funktion X gibt es eine $\mathcal{G}$–messbare numerische Funktion Z mit $X = Z$.*

Beweis. Der Beweis von (1) verläuft analog zum Beweis von Satz 11.2.9. Zum Beweis von (2) betrachten wir $A \in \mathcal{E}_\infty$. Nach (1) gilt $P[\{P^\mathcal{G}(A) \in \{0,1\}\}] = 1$, und da $P^\mathcal{G}(A)$ $\mathcal{G}$–messbar ist folgt hieraus die Existenz eines Ereignisses $G \in \mathcal{G}$ mit

$$P^\mathcal{G}(A) = \chi_G$$

Daher gilt $P[A] = P[G]$ und wegen $\chi_G = \chi_G^2 = \chi_G\,P^\mathcal{G}(A) = P^\mathcal{G}(G\cap A)$ gilt auch $P[G] = P[G\cap A]$. Aus diesen Gleichungen ergibt sich $P[G\triangle A] = 0$ und damit

$$\chi_G = \chi_A$$

Damit ist (2) gezeigt, und (3) ergibt sich dann durch algebraische Induktion. □

Schließlich übertragen wir den Begriff der bedingten Unabhängigkeit auf Zufallsgrößen:

Sei $\{(\Omega_i',\mathcal{F}_i')\}_{i\in I}$ eine Familie von Messräumen. Eine Familie $\{X_i\}_{i\in I}$ von Zufallsgrößen $X_i : \Omega \to \Omega_i'$ heißt *$\mathcal{G}$–bedingt unabhängig*, wenn die Familie $\{\sigma(X_i)\}_{i\in I}$ $\mathcal{G}$–bedingt unabhängig ist.

Die in Abschnitt 11.3 bewiesenen Ergebnisse über die Unabhängigkeit einer Familie von Zufallsgrößen lassen sich weitgehend auf die bedingte Unabhängigkeit übertragen. Eine Ausnahme bildet vorerst die in Satz 11.3.5 angegebene Charakterisierung der Unabhängigkeit einer Familie von Zufallsgrößen durch eine Eigenschaft ihrer gemeinsamen Verteilung; vgl. Abschnitt 19.3.

Das folgende Ergebnis ist eine bedingte Version von Lemma 13.6.8:

19.2.8 Lemma. *Sei $\{X, Y\}$ $\mathcal{G}$–bedingt unabhängig. Wenn X und Y $\mathcal{G}$–bedingt integrierbar sind, dann ist auch XY $\mathcal{G}$–bedingt integrierbar und es gilt* $\mathrm{cov}^{\mathcal{G}}(X, Y) = 0$.

Beweis. Für $\mathcal{G}$–bedingt unabhängige Ereignisse $A, B \in \mathcal{F}$ gilt

$$\begin{aligned} E^{\mathcal{G}}\big(\chi_A \chi_B\big) &= E^{\mathcal{G}}\big(\chi_{A\cap B}\big) \\ &= P^{\mathcal{G}}\big(A \cap B\big) \\ &= P^{\mathcal{G}}\big(A\big)\, P^{\mathcal{G}}\big(B\big) \\ &= E^{\mathcal{G}}\big(\chi_A\big)\, E^{\mathcal{G}}\big(\chi_B\big) \end{aligned}$$

Aus der Linearität der bedingten Erwartung und dem bedingten Satz über die monotone Konvergenz ergibt sich nun im Fall der Positivität von X und Y die Gleichung

$$E^{\mathcal{G}}(XY) = E^{\mathcal{G}}(X)\, E^{\mathcal{G}}(Y)$$

und damit $XY \in L^{1,\mathcal{G}}(\mathcal{F})$ und $\mathrm{cov}^{\mathcal{G}}(X, Y) = 0$. Im allgemeinen Fall folgt die Behauptung dann aus den Gleichungen $(XY)^+ = X^+Y^+ + X^-Y^-$ und $(XY)^- = X^+Y^- + X^-Y^+$. □

19.3 Bedingte Verteilung

Sei $(\Omega', \mathcal{F}')$ ein Messraum. Eine Abbildung $K : \mathcal{F}' \times \Omega \to [0, 1]$ heißt *$\mathcal{G}$–Markov–Kern*, wenn
(i) für jedes $\omega \in \Omega$ die Abbildung $K(\,.\,, \omega) : \mathcal{F}' \to [0, 1]$ ein Wahrscheinlichkeitsmaß ist und
(ii) für jedes $A' \in \mathcal{F}'$ die Abbildung $K(A', \,.\,) : \Omega \to [0, 1]$ $\mathcal{G}$–messbar ist.
Im folgenden betrachten wir nur $\mathcal{G}$–Markov–Kerne auf $\mathcal{B}(\mathbb{R}) \times \Omega$.

Eine Abbildung $F : \mathbb{R} \times \Omega \to [0, 1]$ heißt *$\mathcal{G}$–bedingte Verteilungsfunktion*, wenn
(i) für jedes $\omega \in \Omega$ die Abbildung $F(\,.\,, \omega) : \mathbb{R} \to [0, 1]$ eine Verteilungsfunktion ist und
(ii) für jedes $x \in \mathbb{R}$ die Abbildung $F(x, \,.\,) : \Omega \to [0, 1]$ $\mathcal{G}$–messbar ist.
Für $\mathcal{G}$–Markov–Kerne auf $\mathcal{B}(\mathbb{R}) \times \Omega$ und $\mathcal{G}$–bedingte Verteilungsfunktionen gilt der folgende bedingte Korrespondenzsatz:

19.3.1 Satz (Korrespondenzsatz).

(1) *Zu jedem $\mathcal{G}$–Markov–Kern $K : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ gibt es genau eine $\mathcal{G}$–bedingte Verteilungsfunktion $F_K : \mathbb{R} \times \Omega \to [0,1]$ mit*

$$F_K(x,\omega) = K((-\infty, x], \omega)$$

für alle $x \in \mathbb{R}$ und $\omega \in \Omega$.

(2) *Zu jeder $\mathcal{G}$–bedingten Verteilungsfunktion $F_K : \mathbb{R} \times \Omega \to [0,1]$ gibt es genau einen $\mathcal{G}$–Markov–Kern $K_F : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ mit*

$$K_F((-\infty, x], \omega) = F(x,\omega)$$

für alle $x \in \mathbb{R}$ und $\omega \in \Omega$.

(3) *Es gilt $K_{(F_K)} = K$ und $F_{(K_F)} = F$.*

Beweis. Sei zunächst $K : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ ein $\mathcal{G}$–Markov–Kern und sei $F_K : \mathbb{R} \times \Omega \to [0,1]$ gegeben durch

$$F_K(x,\omega) := K((-\infty, x], \omega)$$

Aus dem Korrespondenzsatz 12.1.1 folgt, dass für jedes $\omega \in \Omega$ die Abbildung $F(\,.\,,\omega)$ eine Verteilungsfunktion ist, und aus der Definition ist unmittelbar klar, dass für jedes $x \in \mathbb{R}$ die Abbildung $F(x,\,.\,)$ $\mathcal{G}$–messbar ist. Daher ist F_K eine $\mathcal{G}$–bedingte Verteilungsfunktion.

Sei nun $F_K : \mathbb{R} \times \Omega \to [0,1]$ eine $\mathcal{G}$–bedingte Verteilungsfunktion. Aus dem Korrespondenzsatz 12.1.1 folgt, dass es zu jedem $\omega \in \Omega$ genau eine Verteilung $K_\omega : \mathcal{B}(\mathbb{R}) \to [0,1]$ gibt mit $K_\omega[(-\infty, x]] = F(x,\omega)$ für alle $x \in \mathbb{R}$. Sei nun $K : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ gegeben durch

$$K(B,\omega) := K_\omega[B]$$

Dann ist für jedes $\omega \in \Omega$ die Abbildung $K(\,.\,,\omega)$ ein Wahrscheinlichkeitsmaß. Es bleibt zu zeigen, dass für jedes $B \in \mathcal{B}(\mathbb{R})$ die Abbildung $K(B,\,.\,)$ $\mathcal{G}$–messbar ist. Sei

$$\mathcal{D} := \Big\{ B \in \mathcal{B}(\mathbb{R}) \;\Big|\; K(B,\,.\,) \text{ ist } \mathcal{G}\text{–messbar} \Big\}$$

und

$$\mathcal{E} := \Big\{ (-\infty, x] \;\Big|\; x \in \mathbb{R} \Big\}$$

Dann ist $\mathcal{D}$ ein Dynkin–System und $\mathcal{E}$ ist $\cap$–stabil mit $\mathcal{E} \subseteq \mathcal{D}$ und $\sigma(\mathcal{E}) = \mathcal{B}(\mathbb{R})$. Daher gilt $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{E}) = \delta(\mathcal{E}) \subseteq \mathcal{D}$, und damit ist für jedes $B \in \mathcal{B}(\mathbb{R})$ die Abbildung $K(B,\,.\,)$ $\mathcal{G}$–messbar.

Die übrigen Aussagen des Satzes ergeben sich wieder aus dem Korrespondenzsatz 12.1.1. □

Ist X eine reelle Zufallsvariable, so heißt eine Abbildung $K : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ *$\mathcal{G}$–bedingte Verteilung* von X, wenn

(i) K ein $\mathcal{G}$–Markov–Kern ist und

(ii) für alle $B \in \mathcal{B}(\mathbb{R})$

$$K(B,\omega) = P^{\mathcal{G}}(X^{-1}(B))(\omega)$$

fast sicher gilt.

Der folgende Satz zeigt, dass jede reelle Zufallsvariable eine im wesentlichen eindeutig bestimmte $\mathcal{G}$–bedingte Verteilung besitzt:

19.3.2 Satz. *Sei X eine reelle Zufallsvariable. Dann gilt:*

(1) *X besitzt eine $\mathcal{G}$–bedingte Verteilung*

(2) *Sind K_1 und K_2 $\mathcal{G}$–bedingte Verteilungen von X, so gilt*

$$K_1(\,.\,,\omega) = K_2(\,.\,,\omega)$$

fast sicher.

Beweis. Für alle $r \in \mathbb{Q}$ sei H_r eine Version der $\mathcal{G}$–bedingten Wahrscheinlichkeit $P^{\mathcal{G}}(\{X \leq r\})$ und für alle $\omega \in \Omega$ sei $G_\omega : \mathbb{Q} \to [0,1]$ gegeben durch

$$G_\omega(r) := H_r(\omega)$$

Da $\mathbb{Q}$ abzählbar ist, gibt es eine Nullmenge $N \in \mathcal{G}$ derart, dass für alle $\omega \in \Omega \setminus N$ die Funktion G_ω die folgenden Eigenschaften besitzt:

(i) Für alle $r,s \in \mathbb{Q}$ mit $r \leq s$ gilt $G_\omega(r) \leq G_\omega(s)$.

(ii) Für alle $r \in \mathbb{Q}$ und jede monoton fallende Folge $\{r_n\}_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ mit $r = \inf_{n\in\mathbb{N}} r_n$ gilt $G_\omega(r) = \lim_{n\to\infty} G_\omega(r_n)$.

(iii) Es gilt $\inf_{r\in\mathbb{Q}} G_\omega(r) = 0$ und $\sup_{r\in\mathbb{Q}} G_\omega(r) = 1$.

Sei $F_X : \mathbb{R} \to [0,1]$ die Verteilungsfunktion von X und sei $F : \mathbb{R} \times \Omega \to [0,1]$ gegeben durch

$$F(x,\omega) := \inf_{r\in\mathbb{Q}\cap[x,\infty)} G_\omega(r)\,\chi_{\Omega\setminus N}(\omega) + F_X(x)\,\chi_N(\omega)$$

Dann ist für alle $\omega \in \Omega$ die Funktion $F(\,.\,,\omega)$ eine Verteilungsfunktion, und an der Darstellung

$$F(x,\omega) = \inf_{r\in\mathbb{Q}\cap[x,\infty)} H_r(\omega)\,\chi_{\Omega\setminus N}(\omega) + F_X(x)\,\chi_N(\omega)$$

erkennt man, dass für alle $x \in \mathbb{R}$ die Funktion $F(x,\,.\,)$ $\mathcal{G}$–messbar ist. Daher ist F eine $\mathcal{G}$–bedingte Verteilungsfunktion.

Wir betrachten nun den zu F gehörigen $\mathcal{G}$–Markov–Kern K. Für $G \in \mathcal{G}$ seien $\mu_G, \nu_G : \mathcal{B}(\mathbb{R}) \to [0,1]$ gegeben durch

$$\mu_G[B] := \int_G \chi_{\{X\in B\}}(\omega)\,dP(\omega)$$

$$\nu_G[B] := \int_G K(B,\omega)\,dP(\omega)$$

Für alle $x \in \mathbb{R}$ erhält man unter Verwendung des bedingten Satzes über die monotone Konvergenz

$$\begin{aligned}
\mu_G[(-\infty, x]] &= \int_G \chi_{\{X \leq x\}}(\omega)\, dP(\omega) \\
&= \int_G P^{\mathcal{G}}(\{X \leq x\})(\omega)\, dP(\omega) \\
&= \int_G \inf_{r \in \mathbb{Q} \cap [x,\infty)} P^{\mathcal{G}}(\{X \leq r\})(\omega)\, dP(\omega) \\
&= \int_G \inf_{r \in \mathbb{Q} \cap [x,\infty)} H_r(\omega)\, dP(\omega) \\
&= \int_G F(x, \omega)\, dP(\omega) \\
&= \int_G K((-\infty, x], \omega)\, dP(\omega) \\
&= \nu_G[(-\infty, x]]
\end{aligned}$$

Aus dem Eindeutigkeitssatz folgt nun $\mu_G = \nu_G$. Für alle $B \in \mathcal{G}$ gilt daher für alle $G \in \mathcal{G}$

$$\int_G \chi_{\{X \in B\}}(\omega)\, dP(\omega) = \int_G K(B, \omega)\, dP(\omega)$$

und aus der $\mathcal{G}$–Messbarkeit von $K(B, .)$ folgt nun

$$K(B, \omega) = P^{\mathcal{G}}(X^{-1}(B))(\omega)$$

fast sicher.
Seien schließlich K_1 und K_2 $\mathcal{G}$–bedingte Verteilungen von X. Dann gibt es eine Nullmenge N derart, dass für alle $\omega \in \Omega \setminus N$ und für alle $r \in \mathbb{Q}$

$$K_1((-\infty, r], \omega) = K_2((-\infty, r], \omega)$$

gilt. Sei

$$\mathcal{D} := \Big\{ B \in \mathcal{B}(\mathbb{R}) \;\Big|\; K_1(B, \omega) = K_2(B, \omega) \text{ für alle } \omega \in \Omega \setminus N \Big\}$$

und

$$\mathcal{E} := \Big\{ (-\infty, r] \;\Big|\; r \in \mathbb{Q} \Big\}$$

Dann ist $\mathcal{D}$ ein Dynkin–System und $\mathcal{E}$ ist $\cap$–stabil mit $\mathcal{E} \subseteq \mathcal{D}$ und $\sigma(\mathcal{E}) = \mathcal{B}(\mathbb{R})$. Daher gilt $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{E}) = \delta(\mathcal{E}) \subseteq \mathcal{D}$ und damit $K_1(B, \omega) = K_2(B, \omega)$ für alle $B \in \mathcal{B}(\mathbb{R})$ und $\omega \in \Omega \setminus N$. □

Im folgenden bezeichnen wir die (fast sicher eindeutig bestimmte) $\mathcal{G}$–bedingte Verteilung einer reellen Zufallsvariablen X mit $P_X^{\mathcal{G}}$.

Der folgende Satz vervollständigt die Eigenschaften der bedingten Erwartung:

19.3.3 Satz. *Sei X eine positive oder $\mathcal{G}$–bedingt integrierbare reelle Zufallsvariable und sei $h : \mathbb{R} \to \mathbb{R}$ messbar. Dann gilt*

$$E^{\mathcal{G}}(h \circ X) = \int_{\mathbb{R}} h(x)\, dP_X^{\mathcal{G}}(x, .)$$

Beweis. Für alle $B \in \mathcal{B}(\mathbb{R})$ gilt

$$\begin{aligned} E^{\mathcal{G}}(\chi_B \circ X) &= E^{\mathcal{G}}(\chi_{\{X \in B\}}) \\ &= P^{\mathcal{G}}(\{X \in B\}) \\ &= P_X^{\mathcal{G}}(B, .) \\ &= \int_{\mathbb{R}} \chi_B(x)\, dP_X^{\mathcal{G}}(x, .) \end{aligned}$$

Die Behauptung folgt nun durch algebraische Induktion. □

Wir notieren abschließend zwei Folgerungen aus Satz 19.3.3; vgl. Aufgabe 19.3.D:

19.3.4 Folgerung. *Sei X eine reelle Zufallsvariable und sei $h : \mathbb{R} \to \mathbb{R}$ messbar. Dann gilt für alle $B \in \mathcal{B}(\mathbb{R})$*

$$P^{\mathcal{G}}(\{h \circ X \in B\}) = \int_{\mathbb{R}} \chi_B(h(x))\, dP_X^{\mathcal{G}}(x, .)$$

19.3.5 Folgerung. *Sei X eine reelle Zufallsvariable und sei $h : \mathbb{R} \to \mathbb{R}$ messbar. Dann gilt für alle $B \in \mathcal{B}(\mathbb{R})$*

$$P_{h \circ X}^{\mathcal{G}}(B, .) = P_X^{\mathcal{G}}(h^{-1}(B), .)$$

Aufgaben

19.3.A **Empirische Verteilung:** Jede empirische Verteilung ist ein $\mathcal{F}$–Markov–Kern.

19.3.B **Empirische Verteilungsfunktion:** Jede empirische Verteilungsfunktion ist eine $\mathcal{F}$–bedingte Verteilungsfunktion.

19.3.C **Bedingte Verteilung:** Verallgemeinern Sie den Begriff der bedingten Verteilung auf den multivariaten Fall und verallgemeinern Sie Satz 19.3.2 auf Zufallsvektoren. Beweisen Sie außerdem eine bedingte Version von Satz 11.3.5.

19.3.D Beweisen Sie die Folgerungen 19.3.4 und 19.3.5.

19.4 Bedingte Dichte

Ist $K : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ ein $\mathcal{G}$–Markov–Kern und $\mu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ ein σ–endliches Maß, so heißt eine Abbildung $k : \mathbb{R} \times \Omega \to \mathbb{R}_+$ *$\mathcal{G}$–bedingte μ–Dichte* von K, wenn

(i) für jedes $\omega \in \Omega$ die Abbildung $k(\,.\,,\omega) : \mathbb{R} \to \mathbb{R}_+$ eine μ–Dichte von $K(\,.\,,\omega)$ ist und

(ii) für jedes $x \in \mathbb{R}$ die Abbildung $k(x,\,.\,) : \Omega \to \mathbb{R}_+$ $\mathcal{G}$–messbar ist.

Wir zeigen die Existenz einer $\mathcal{G}$–bedingten μ–Dichte eines $\mathcal{G}$–Markov–Kerns $\mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ in einem Spezialfall:

19.4.1 Satz. *Sei $(\Omega', \mathcal{F}')$ ein Messraum und sei $\Theta : \Omega \to \Omega'$ eine Zufallsgröße. Sei X eine reelle Zufallsvariable. Wenn es σ–endliche Maße $\mu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ und $\nu : \mathcal{F}' \to [0,\infty]$ sowie eine messbare Funktion $h : \mathbb{R} \times \Omega' \to \mathbb{R}_+$ gibt mit*

$$P_{X,\Theta} = \int h(x,\vartheta)\, d(\mu \otimes \nu)(x,\vartheta)$$

dann besitzt die $\sigma(\Theta)$–bedingte Verteilung von X eine $\sigma(\Theta)$–bedingte μ–Dichte und für alle $B \in \mathcal{B}(\mathbb{R})$ gilt

$$P_X^{\sigma(\Theta)}(B,\omega) = \int_B \frac{h(x,\Theta(\omega))}{\int_{\mathbb{R}} h(z,\Theta(\omega))\, d\mu(z)}\, d\mu(x)$$

fast sicher.

Beweis. Nach Voraussetzung ist die Funktion $f : \mathbb{R} \to [0,\infty]$ mit

$$f(x) := \int_{\Omega'} h(x,\vartheta)\, d\nu(\vartheta)$$

eine μ–Dichte von P_X und die Funktion $g : \Omega' \to [0,\infty]$ mit

$$g(\vartheta) := \int_{\mathbb{R}} h(x,\vartheta)\, d\mu(x)$$

ist eine ν–Dichte von P_Θ. Daher gilt $g(\vartheta) < \infty$ ν–fast überall, und wir zeigen nun, dass wir ohne Beschränkung der Allgemeinheit annehmen können, dass auch $g(\vartheta) > 0$ ν–fast überall gilt.
Die Abbildung $\widetilde{\nu} : \mathcal{F}' \to [0,\infty]$ mit

$$\widetilde{\nu}[C] := \int_C \chi_{\{g>0\}}(\vartheta)\, d\nu(\vartheta)$$

ist ein σ–endliches Maß mit $\widetilde{\nu}[\{g=0\}] = 0$ und aus dem Satz von Fubini erhält man für alle $B \in \mathcal{B}(\mathbb{R})$ und $C \in \mathcal{F}'$

$$
\begin{aligned}
P[\{X \in B\} \cap \{\Theta \in C\}] &= \int_{B\times C} h(x,\vartheta)\, d(\mu \otimes \nu)(x,\vartheta) \\
&= \int_C \int_B h(x,\vartheta)\, d\mu(x)\, d\nu(\vartheta) \\
&= \int_C \int_B h(x,\vartheta)\, d\mu(x)\, \chi_{\{g>0\}}(\vartheta)\, d\nu(\vartheta) \\
&= \int_C \int_B h(x,\vartheta)\, d\mu(x)\, d\widetilde{\nu}(\vartheta) \\
&= \int_{B\times C} h(x,\vartheta)\, d(\mu \otimes \widetilde{\nu})(x,\vartheta)
\end{aligned}
$$

und insbesondere

$$
\begin{aligned}
\int_B f(x)\, d\mu(x) &= P[\{X \in B\}] \\
&= P[\{X \in B\} \cap \{\Theta \in \Omega'\}] \\
&= \int_{B\times\Omega'} h(x,\vartheta)\, d(\mu \otimes \widetilde{\nu})(x,\vartheta) \\
&= \int_B \int_{\Omega'} h(x,\vartheta)\, d\widetilde{\nu}(\vartheta)\, d\mu(x)
\end{aligned}
$$

und damit

$$
f(x) = \int_{\Omega'} h(x,\vartheta)\, d\widetilde{\nu}(\vartheta)
$$

μ–fast überall. Wir können daher ohne Beschränkung der Allgemeinheit annehmen, dass $0 < g(\vartheta) < \infty$ ν–fast überall gilt.
Damit ist die Menge $N' := \{g \in \{0,\infty\}\}$ eine ν–Nullmenge und die Menge $N := \{g \circ \Theta \in \{0,\infty\}\}$ ist wegen $N = \Theta^{-1}(N')$ und

$$
P[N] = P[\Theta^{-1}(N')] = P_\Theta[N'] = \int_{N'} g(\vartheta)\, d\nu(\vartheta) = 0
$$

eine P–Nullmenge. Daher sind die Funktionen $k' : \mathbb{R} \times \Omega' \to \mathbb{R}_+$ mit

$$
k'(x,\vartheta) := \frac{h(x,\vartheta)}{\int_{\mathbb{R}} h(z,\vartheta)\, d\mu(z)}\, \chi_{\Omega'\setminus N'}(\vartheta) + f(x)\, \chi_{N'}(\vartheta)
$$

und $k : \mathbb{R} \times \Omega \to \mathbb{R}_+$ mit

$$
k(x,\omega) := k'(x,\Theta(\omega))
$$

wohldefiniert und messbar und es gilt

$$
k(x,\omega) = \frac{h(x,\Theta(\omega))}{\int_{\mathbb{R}} h(z,\Theta(\omega))\, d\mu(z)}\, \chi_{\Omega\setminus N}(\omega) + f(x)\, \chi_N(\omega)
$$

Sei nun $K : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ gegeben durch

$$K(B,\omega) = \int_B k(x,\omega)\, d\mu(x)$$

Dann ist K ein $\sigma(\Theta)$–Markov–Kern und k ist eine $\sigma(\Theta)$–bedingte μ–Dichte von K. Sei $B \in \mathcal{B}(\mathbb{R})$. Dann gilt für alle $C \in \mathcal{F}'$

$$\begin{aligned}
\int_{\Theta^{-1}(C)} K(B,\omega)\, dP(\omega) &= \int_{\Theta^{-1}(C)} \int_B k(x,\omega)\, d\mu(x)\, dP(\omega) \\
&= \int_B \int_{\Theta^{-1}(C)} k'(x,\Theta(\omega))\, dP(\omega)\, d\mu(x) \\
&= \int_B \int_C k'(x,\vartheta)\, dP_\Theta(\vartheta)\, d\mu(x) \\
&= \int_B \int_C k'(x,\vartheta)\left(\int_{\mathbb{R}} h(z,\vartheta)\, d\mu(z)\right) d\nu(\vartheta)\, d\mu(x) \\
&= \int_B \int_C h(x,\vartheta)\, d\nu(\vartheta)\, d\mu(x) \\
&= \int_{B\times C} h(x,\vartheta)\, d(\mu\otimes\nu)(x,\vartheta) \\
&= P[\{X \in B\} \cap \{\Theta \in C\}] \\
&= \int_{\Theta^{-1}(C)} \chi_{X^{-1}(B)}(\omega)\, dP(\omega)
\end{aligned}$$

und damit gilt für alle $A \in \sigma(\Theta)$

$$\int_A K(B,\omega)\, dP(\omega) = \int_A \chi_{X^{-1}(B)}(\omega)\, dP(\omega)$$

Da $K(B,\,.\,)$ $\sigma(\Theta)$–messbar ist, folgt daraus

$$K(B,\omega) = P^{\sigma(\Theta)}(X^{-1}(B))(\omega) = P_X^{\sigma(\Theta)}(B,\omega)$$

fast sicher. Daher ist k eine $\sigma(\Theta)$–bedingte μ–Dichte von $P_X^{\sigma(\Theta)}$. □

Der folgende Satz liefert eine Umkehrung der Aussage des letzten Satzes:

19.4.2 Satz. *Sei $(\Omega',\mathcal{F}')$ ein Messraum und sei $\Theta : \Omega \to \Omega'$ eine Zufallsgröße. Sei X eine reelle Zufallsvariable. Wenn es σ–endliche Maße $\mu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ und $\nu : \mathcal{F}' \to [0,\infty]$ sowie messbare Funktionen $k' : \mathbb{R}_+ \times \Omega' \to \mathbb{R}_+$ und $g : \Omega' \to \mathbb{R}_+$ gibt derart, dass für alle $B \in \mathcal{B}(\mathbb{R})$*

$$P_X^{\sigma(\Theta)}(B,\omega) = \int_B k'(x,\Theta(\omega))\, d\mu(x)$$

fast sicher und

$$P_\Theta = \int g(\vartheta)\, d\nu(\vartheta)$$

gilt, dann gilt

$$P_{X,\Theta} = \int k'(x,\vartheta)\, g(\vartheta)\, d(\mu \otimes \nu)(x,\vartheta)$$

Beweis. Für alle $B \in \mathcal{B}(\mathbb{R})$ und $C \in \mathcal{F}'$ gilt

$$\begin{aligned} P^{\sigma(\Theta)}(\{X \in B\} \cap \{\Theta \in C\}) &= \chi_{\{\Theta \in C\}}\, P^{\sigma(\Theta)}(\{X \in B\}) \\ &= \chi_{\{\Theta \in C\}}\, P_X^{\sigma(\Theta)}(B, \,.\,) \\ &= \chi_{\{\Theta \in C\}} \int_B k'(x, \Theta(\,.\,))\, d\mu(x) \end{aligned}$$

und damit

$$\begin{aligned} P[\{X \in B\} \cap \{\Theta \in C\}] &= E[P^{\sigma(\Theta)}(\{X \in B\} \cap \{\Theta \in C\})] \\ &= \int_\Omega \chi_{\{\Theta \in C\}} \int_B k'(x, \Theta(\omega))\, d\mu(x) dP(\omega) \\ &= \int_{\Theta^{-1}(C)} \int_B k'(x, \Theta(\omega))\, d\mu(x)\, dP(\omega) \\ &= \int_B \int_{\Theta^{-1}(C)} k'(x, \Theta(\omega))\, dP(\omega)\, d\mu(x) \\ &= \int_B \int_C k'(x,\vartheta)\, dP_\Theta(\vartheta)\, d\mu(x) \\ &= \int_B \int_C k'(x,\vartheta)\, g(\vartheta)\, d\nu(\vartheta)\, d\mu(x) \\ &= \int_{B \times C} k'(x,\vartheta)\, g(\vartheta)\, d(\mu \otimes \nu)(x,\vartheta) \end{aligned}$$

Damit ist der Satz bewiesen. □

Wir geben abschließend anhand eines Beispiels eine Anwendung der Sätze über bedingte Dichten:

19.4.3 Beispiel. Sei Θ eine reelle Zufallsvariable mit $P_\Theta = \mathbf{Ga}(\alpha,\gamma)$ und sei X eine reelle Zufallsvariable mit $P_X^{\sigma(\Theta)} = \mathbf{P}(\Theta)$. Dann gilt für alle $B, C \in \mathcal{B}(\mathbb{R})$

$$P[\{X \in B\} \cap \{\Theta \in C\}] = \sum_{k \in B \cap \mathbb{N}_0} \int_{C \cap (0,\infty)} \frac{\alpha^\gamma}{\Gamma(\gamma)\, k!} e^{-(\alpha+1)\vartheta} \vartheta^{\gamma+k-1} d\boldsymbol{\lambda}(\vartheta)$$

Insbesondere gilt $P_X = \mathbf{NB}(\gamma, 1/(\alpha+1))$ und $P_\Theta^{\sigma(X)} = \mathbf{Ga}(\alpha+1, \gamma+X)$.

In der Tat: Wegen

$$P_\Theta[C] = \int_C \frac{\alpha^\gamma}{\Gamma(\gamma)} e^{-\alpha\vartheta}\, \vartheta^{\gamma-1}\, \chi_{(0,\infty)}(\vartheta)\, d\boldsymbol{\lambda}(\vartheta)$$

und

$$P_X^{\sigma(\Theta)}(B,\omega) = \int_B e^{-\Theta(\omega)}\, \frac{(\Theta(\omega))^x}{x!}\, \chi_{\mathbb{N}_0}(x)\, d\boldsymbol{\zeta}(x)$$

gilt nach Satz 19.4.2

$$\begin{aligned}
&P[\{X \in B\} \cap \{\Theta \in C\}] \\
&= \int_{B\times C} e^{-\vartheta}\, \frac{\vartheta^x}{x!}\, \chi_{\mathbb{N}_0}(x)\, \frac{\alpha^\gamma}{\Gamma(\gamma)}\, e^{-\alpha\vartheta}\, \vartheta^{\gamma-1}\, \chi_{(0,\infty)}(\vartheta)\, d(\boldsymbol{\zeta} \otimes \boldsymbol{\lambda})(x,\vartheta) \\
&= \int_{B\times C} e^{-\vartheta}\, \frac{\vartheta^x}{x!}\, \frac{\alpha^\gamma}{\Gamma(\gamma)}\, e^{-\alpha\vartheta}\, \vartheta^{\gamma-1}\, \chi_{(0,\infty)}(\vartheta)\, d(\boldsymbol{\zeta}_{\mathbb{N}_0} \otimes \boldsymbol{\lambda})(x,\vartheta)
\end{aligned}$$

Daraus folgt einerseits für alle $k \in \mathbb{N}_0$

$$\begin{aligned}
P[\{X = k\}] &= \int_{\{k\}\times\mathbb{R}} e^{-\vartheta}\, \frac{\vartheta^x}{x!}\, \frac{\alpha^\gamma}{\Gamma(\gamma)}\, e^{-\alpha\vartheta}\, \vartheta^{\gamma-1}\, \chi_{(0,\infty)}(\vartheta)\, d(\boldsymbol{\zeta}_{\mathbb{N}_0} \otimes \boldsymbol{\lambda})(x,\vartheta) \\
&= \int_{\mathbb{R}} \frac{\alpha^\gamma}{\Gamma(\gamma)\, k!} e^{-(\alpha+1)\vartheta}\, \vartheta^{\gamma+k-1}\, \chi_{(0,\infty)}(\vartheta)\, d\boldsymbol{\lambda}(\vartheta) \\
&= \frac{\alpha^\gamma}{\Gamma(\gamma)\, k!}\, \frac{\Gamma(\gamma+k)}{(\alpha+1)^{\gamma+k}} \int_{\mathbb{R}} \frac{(\alpha+1)^{\gamma+k}}{\Gamma(\gamma+k)} e^{-(\alpha+1)\vartheta}\, \vartheta^{\gamma+k-1}\, \chi_{(0,\infty)}(\vartheta)\, d\boldsymbol{\lambda}(\vartheta) \\
&= \frac{\alpha^\gamma}{\Gamma(\gamma)\, k!}\, \frac{\Gamma(\gamma+k)}{(\alpha+1)^{\gamma+k}} \\
&= \binom{\gamma+k-1}{k} \left(\frac{\alpha}{\alpha+1}\right)^\gamma \left(\frac{1}{\alpha+1}\right)^k
\end{aligned}$$

und damit $P_X = \mathbf{NB}(\gamma, 1/(\alpha+1))$, und andererseits ergibt sich aus Satz 19.4.1 für alle $C \in \mathcal{B}(\mathbb{R})$

$$\begin{aligned}
P_\Theta^{\sigma(X)}(C,\omega) &= \int_C \frac{e^{-\vartheta}\, \dfrac{\vartheta^{X(\omega)}}{X(\omega)!}\, \dfrac{\alpha^\gamma}{\Gamma(\gamma)}\, e^{-\alpha\vartheta}\, \vartheta^{\gamma-1}\, \chi_{(0,\infty)}(\vartheta)}{\displaystyle\int_{\mathbb{R}} e^{-\tau}\, \frac{\tau^{X(\omega)}}{X(\omega)!}\, \frac{\alpha^\gamma}{\Gamma(\gamma)}\, e^{-\alpha\tau}\, \tau^{\gamma-1}\, \chi_{(0,\infty)}(\tau)\, d\boldsymbol{\lambda}(\tau)}\, d\boldsymbol{\lambda}(\vartheta) \\
&= \int_C \frac{\dfrac{(\alpha+1)^\gamma}{\Gamma(\gamma+X(\omega))} e^{-(\alpha+1)\vartheta}\, \vartheta^{\gamma+X(\omega)-1}\, \chi_{(0,\infty)}(\vartheta)}{\displaystyle\int_{\mathbb{R}} \frac{(\alpha+1)^\gamma}{\Gamma(\gamma+X(\omega))} e^{-(\alpha+1)\tau}\, \tau^{\gamma+X(\omega)-1}\, \chi_{(0,\infty)}(\tau)\, d\boldsymbol{\lambda}(\tau)}\, d\boldsymbol{\lambda}(\vartheta) \\
&= \int_C \frac{(\alpha+1)^\gamma}{\Gamma(\gamma+X(\omega))} e^{-(\alpha+1)\vartheta}\, \vartheta^{\gamma+X(\omega)-1}\, \chi_{(0,\infty)}(\vartheta)\, d\boldsymbol{\lambda}(\vartheta)
\end{aligned}$$

und damit $P_\Theta^{\sigma(X)} = \mathbf{Ga}(\alpha+1, \gamma+X)$.

19.5 Bedingte Gesetze der Großen Zahlen

Der folgende Satz ist eine bedingte Version des Null–Eins–Gesetzes 15.2.1:

19.5.1 Satz (Null–Eins–Gesetz). *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige Folge von Zufallsvariablen und sei*

$$A := \left\{ \liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \right\}$$

Dann gibt es eine $\mathcal{G}$–messbare Zufallsvariable U mit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \chi_A = U$$

fast sicher.

Beweis. Sei $\mathcal{E}_\infty$ die terminale σ–Algebra der Folge $\{\sigma(X_k)\}_{k\in\mathbb{N}}$. Dann gilt $A \in \mathcal{E}_\infty$ und die Folge $\{n^{-1}\sum_{k=1}^{n} X_k\chi_A\}$ konvergiert fast sicher gegen eine $\mathcal{E}_\infty$–messbare Zufallsvariable. Aus dem bedingten Null–Eins–Gesetz von Kolmogorov folgt nun die Existenz eines Ereignisses $G \in \mathcal{G}$ mit $\chi_A = \chi_G$ und die Existenz einer $\mathcal{G}$–messbaren Zufallsvariablen U mit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \, \chi_A = U$$

fast sicher. □

Sei

$$L^{2,\mathcal{G}}(\mathcal{F}) := \Big\{ [X]_P \in L^0(\mathcal{F}) \;\Big|\; X =_P Y \text{ für ein } Y \text{ mit } E^{\mathcal{G}}(Y^2) < \infty \Big\}$$

Dann gilt

$$L^{2,\mathcal{G}}(\mathcal{F}) \subseteq L^{1,\mathcal{G}}(\mathcal{F})$$

Wir beweisen nun eine bedingte Version des 1. Gesetzes der Großen Zahlen.

19.5.2 Lemma (Ungleichung von Kolmogorov). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige Folge in $L^{2,\mathcal{G}}(\mathcal{F})$. Dann gilt für alle $\varepsilon \in (0,\infty)$ und $m \in \mathbb{N}$*

$$P^{\mathcal{G}}\left(\left\{ \sup_{n\in\mathbb{N}(m)} \left| \sum_{k=m}^{n} \Big(X_k - E^{\mathcal{G}}(X_k) \Big) \right| > \varepsilon \right\} \right) \le \frac{1}{\varepsilon^2} \sum_{k=m}^{\infty} \mathrm{var}^{\mathcal{G}}(X_k)$$

Beweis. Der Beweis verläuft mit einer kleinen Anpassung wie der Beweis von Lemma 15.2.2: Da im vorliegenden Fall die Zufallsvariablen nicht an ihrem

Erwartungswert, sondern an ihrer $\mathcal{G}$–bedingten Erwartung zentriert werden, benötigt man im Beweis die aus Satz 19.2.5 bekannte Tatsache, dass mit der Folge $\{X_k\}_{k\in\mathbb{N}}$ auch die Folge $\{\sigma(\sigma(X_k)\cup\mathcal{G})\}_{k\in\mathbb{N}}$ $\mathcal{G}$–bedingt unabhängig ist. □

Das im Beweis von Lemma 19.5.2 verwendete Argument wird auch zum Beweis der folgenden bedingten Version von Lemma 15.2.3 benötigt:

19.5.3 Lemma. *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige Folge in $L^{2,\mathcal{G}}(\mathcal{F})$ und sei $G := \{\sum_{k=1}^\infty \mathrm{var}^{\mathcal{G}}(X_k) < \infty\}$. Dann konvergiert die Reihe*

$$\sum_{k=1}^\infty \Big(X_k - E^{\mathcal{G}}(X_k)\Big)\chi_G$$

fast sicher gegen eine reelle Zufallsvariable.

Aus dem Lemma ergibt sich sofort eine bedingte Version des ersten Gesetzes der Großen Zahlen:

19.5.4 Satz (1. Gesetz der Großen Zahlen; Kolmogorov). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige Folge in $L^{2,\mathcal{G}}(\mathcal{F})$ und sei*

$$G := \left\{\sum_{k=1}^\infty \frac{1}{k^2}\,\mathrm{var}^{\mathcal{G}}(X_k) < \infty\right\}$$

Dann gilt

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \Big(X_k - E^{\mathcal{G}}(X_k)\Big)\chi_G = 0$$

fast sicher.

Unter etwas stärkeren Voraussetzungen erhält man neben der fast sicheren Konvergenz auch die Konvergenz in $L^2(\mathcal{F})$:

19.5.5 Folgerung (1. Gesetz der Großen Zahlen; Kolmogorov). *Sei $\{X_k\}_{k\in\mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige Folge in $L^2(\mathcal{F})$ mit*

$$\sup_{k\in\mathbb{N}} E[\mathrm{var}^{\mathcal{G}}(X_k)] < \infty$$

Dann gilt

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \Big(X_k - E^{\mathcal{G}}(X_k)\Big) = 0$$

fast sicher und in $L^2(\mathcal{F})$.

Eine Familie von Zufallsvariablen $\{X_i\}_{i\in I}$ heißt *$\mathcal{G}$–bedingt identisch verteilt*, wenn alle Zufallsvariablen dieselbe $\mathcal{G}$–bedingte Verteilung besitzen; in diesem

Fall bezeichnen wir mit X eine beliebige Zufallsvariable mit $P_X^{\mathcal{G}} = P_{X_i}^{\mathcal{G}}$ für alle $i \in I$ und nennen X eine *typische Zufallsvariable* der Familie $\{X_i\}_{i \in I}$.

Mit geringfügigen Anpassungen des Beweises erhält man auch eine bedingte Version des zweiten Gesetzes der Großen Zahlen:

19.5.6 Satz (2. Gesetz der Großen Zahlen; Kolmogorov). *Sei $\{X_k\}_{k \in \mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige und $\mathcal{G}$–bedingt identisch verteilte Folge in $L^{1,\mathcal{G}}(\mathcal{F})$. Dann gilt*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \Big(X_k - E^{\mathcal{G}}(X_k) \Big) = 0$$

fast sicher.

Schließlich erhält man aus dem 2. bedingten Gesetz der Großen Zahlen eine bedingte Version des Satzes von Glivenko/Cantelli:

19.5.7 Satz (Glivenko/Cantelli). *Sei $\{X_k\}_{k \in \mathbb{N}}$ eine $\mathcal{G}$–bedingt unabhängige und $\mathcal{G}$–bedingt identisch verteilte Folge von Zufallsvariablen und sei $\{F_n\}_{n \in \mathbb{N}}$ die Folge der zugehörigen empirischen Verteilungsfunktionen. Dann gilt*

$$\lim_{n \to \infty} \sup_{x \in \mathbb{R}} \Big| F_n(x, \omega) - F_X^{\mathcal{G}}(x, \omega) \Big| = 0$$

fast sicher.

Aufgaben

19.5.A Führen Sie die fehlenden Beweise aus.

19.5.B Formulieren und beweisen Sie geeignete bedingte Versionen der übrigen Ergebnisse aus Abschnitt 15.2.

19.5.C Jede $\mathcal{G}$–bedingt identisch verteilte Familie von Zufallsvariablen ist identisch verteilt.

19.5.D Je zwei $\mathcal{G}$–bedingt unabhängige und $\mathcal{G}$–bedingt identisch verteilte Zufallsvariablen sind positiv korreliert.

19.5.E **Austauschbare Familien von Zufallsvariablen:** Eine Familie von Zufallsvariablen $\{X_i\}_{i \in I}$ heißt *austauschbar*, wenn für alle $n \in \mathbb{N}$ jede Familie $\{X_i\}_{i \in J}$ mit $|J| = n$ dieselbe Verteilung besitzt.
(1) Jede $\mathcal{G}$–bedingt unabhängige und $\mathcal{G}$–bedingt identisch verteilte Familie von Zufallsvariablen ist austauschbar.
(2) Jede austauschbare Familie von Zufallsvariablen ist identisch verteilt.

20

Regularität und Satz von Kolmogorov

Sei I eine nichtleere Indexmenge und sei $\mathcal{H}(I)$ die Familie der endlichen nichtleeren Teilmengen von I.

Wir betrachten den Messraum $\bigotimes_{i\in I}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und eine projektive Familie von Verteilungen $\{Q_J\}_{J\in\mathcal{H}(I)}$ mit $Q_J : \mathcal{B}(\mathbb{R}^J) \to [0,1]$ für alle $J \in \mathcal{H}(I)$ und untersuchen die Existenz und Eindeutigkeit eines Wahrscheinlichkeitsmaßes $Q : \bigotimes_{i\in I} \mathcal{B}(\mathbb{R}) \to [0,1]$ mit

$$Q_{\pi_J} = Q_J$$

für alle $J \in \mathcal{H}(I)$.

Ist insbesondere $\{Q_i\}_{i\in I}$ eine Familie von Verteilungen auf $\mathcal{B}(\mathbb{R})$, so ist die Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ mit

$$Q_J := \bigotimes_{i\in J} Q_i$$

projektiv, und in diesem Spezialfall folgt die Existenz und Eindeutigkeit eines Wahrscheinlichkeitsmaßes $Q : \bigotimes_{i\in I} \mathcal{B}(\mathbb{R}) \to [0,1]$ mit

$$Q_{\pi_J} = Q_J$$

für alle $J \in \mathcal{H}(I)$ aus dem Satz von Andersen/Jessen. Der Satz von Andersen/Jessen liefert daher insbesondere zu jeder Familie von Verteilungen $\{Q_i\}_{i\in I}$ auf $\mathcal{B}(\mathbb{R})$ die Existenz eines Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$ und einer unabhängigen Familie von Zufallsvariablen $\{X_i\}_{i\in I}$ mit $X_i : \Omega \to \mathbb{R}$ und $P_{X_i} = Q_i$ für alle $i \in I$.

Wir zeigen zunächst, dass jede multivariate Verteilung regulär ist (Abschnitt 20.1), und beweisen dann den Satz von Kolmogorov (Abschnitt 20.2), der das eingangs genannte Problem für jede projektive Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ von endlichdimensionalen Verteilungen $Q_J : \mathcal{B}(\mathbb{R}^J) \to [0,1]$ löst.

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6_20, © Springer-Verlag Berlin Heidelberg 2011

20.1 Regularität

Sei (S,d) ein metrischer Raum und $\mathcal{B}(S)$ die zugehörige Borelsche σ–Algebra. Ein Maß $\mu : \mathcal{B}(S) \to [0,\infty]$ heißt *regulär*, wenn für alle $B \in \mathcal{B}(S)$

$$\mu[B] = \inf\Big\{\mu[U] \,\Big|\, U \in \mathcal{B}(S) \text{ ist offen mit } B \subseteq U\Big\}$$

und

$$\mu[B] = \sup\Big\{\mu[K] \,\Big|\, K \in \mathcal{B}(S) \text{ ist kompakt mit } K \subseteq B\Big\}$$

gilt. Das folgende Lemma liefert die Regularität aller Verteilungen auf $\mathcal{B}(\mathbb{R}^m)$:

20.1.1 Lemma. *Jede Verteilung auf $\mathcal{B}(\mathbb{R}^m)$ ist regulär.*

Beweis. Sei $Q : \mathcal{B}(\mathbb{R}^m) \to [0,1]$ eine Verteilung und sei $\varepsilon \in (0,\infty)$.
Sei zunächst $(\mathbf{a},\mathbf{b}] \in \mathcal{J}(\mathbb{R}^m)$. Dann ist die Folge $\{(\mathbf{a},\mathbf{b}+k^{-1}\mathbf{1})\}_{k\in\mathbb{N}}$ eine monoton fallende Folge offener Mengen mit $(\mathbf{a},\mathbf{b}] = \bigcap_{k\in\mathbb{N}}(\mathbf{a},\mathbf{b}+k^{-1}\mathbf{1})$ und es gilt

$$\begin{aligned} Q[(\mathbf{a},\mathbf{b}]] &\le \inf\Big\{Q[U] \,\Big|\, U \in 2^\Omega \text{ ist offen mit } (\mathbf{a},\mathbf{b}] \subseteq U\Big\} \\ &\le \inf\Big\{Q[(\mathbf{a},\mathbf{b}+k^{-1}\mathbf{1})] \,\Big|\, k \in \mathbb{N}\Big\} \\ &= Q[(\mathbf{a},\mathbf{b}]] \end{aligned}$$

und damit

$$Q[(\mathbf{a},\mathbf{b}]] = \inf\Big\{Q[U] \,\Big|\, U \in 2^\Omega \text{ ist offen mit } (\mathbf{a},\mathbf{b}] \subseteq U\Big\}$$

Sei nun $B \in \mathcal{B}(\mathbb{R}^m)$. Nach dem Satz von Caratheodory stimmt Q mit der Restriktion des von der Restriktion von Q auf $\mathcal{J}(\mathbb{R}^m)$ erzeugten äußeren Maßes $(Q|_{\mathcal{J}(\mathbb{R}^m)})^*$ auf $\mathcal{B}(\mathbb{R}^m)$ überein; es gilt also

$$Q[B] = (Q|_{\mathcal{J}(\mathbb{R}^m)})^*[B]$$

Sei $\varepsilon \in (0,\infty)$. Dann gibt es eine Folge $\{(\mathbf{a}_n,\mathbf{b}_n]\}_{n\in\mathbb{N}} \subseteq \mathcal{J}(\mathbb{R}^m)$ mit $B \subseteq \bigcup_{n\in\mathbb{N}}(\mathbf{a}_n,\mathbf{b}_n]$ und

$$\sum_{n=1}^\infty Q[(\mathbf{a}_n,\mathbf{b}_n]] \le Q[B] + \frac{\varepsilon}{2}$$

und nach dem vorher Gezeigten gibt es für alle $n \in \mathbb{N}$ eine offene Menge U_n mit $(\mathbf{a}_n,\mathbf{b}_n] \subseteq U_n$ und

$$Q[U_n] \le Q[(\mathbf{a}_n,\mathbf{b}_n]) + \frac{\varepsilon}{2^{n+1}}$$

Sei $U := \bigcup_{n\in\mathbb{N}} U_n$. Dann ist U offen mit $B \subseteq U$ und es gilt

$$\begin{aligned} Q[U] &\le \sum_{n=1}^{\infty} Q[U_n] \\ &\le \sum_{n=1}^{\infty} \left(Q[(\mathbf{a}_n, \mathbf{b}_n]) + \frac{\varepsilon}{2^{n+1}} \right) \\ &= \sum_{n=1}^{\infty} Q[(\mathbf{a}_n, \mathbf{b}_n]) + \frac{\varepsilon}{2} \\ &\le Q[B] + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= Q[B] + \varepsilon \end{aligned}$$

Daraus folgt

$$Q[B] = \inf\Big\{ Q[U] \,\Big|\, U \in \mathcal{B}(\mathbb{R}^m) \text{ ist offen mit } B \subseteq U \Big\}$$

Sei nochmals $\varepsilon \in (0,\infty)$. Nach dem bisher Gezeigten gibt es eine offene Menge $V \in \mathcal{B}(\mathbb{R}^m)$ mit $\overline{B} \subseteq V$ und $Q[V] \le Q[\overline{B}] + \varepsilon/2$ und damit

$$Q[V \cap B] + Q[\overline{B}] = Q[V \cap B] + Q[V \setminus B] = Q[V] \le Q[\overline{B}] + \frac{\varepsilon}{2}$$

Da Q endlich ist, ergibt sich nun

$$Q[B \cap V] \le \frac{\varepsilon}{2}$$

und wegen $\inf_{n\in\mathbb{N}} Q[B \setminus [-n,n]] = 0$ gibt es eine kompakte Menge $L \in \mathcal{B}(\mathbb{R}^m)$ mit

$$Q[B \setminus L] \le \frac{\varepsilon}{2}$$

Sei $K := L \setminus V$. Dann ist K kompakt mit $K \subseteq B$ und $B \setminus K \subseteq (B \setminus L) \cup (B \cap V)$ und es gilt

$$Q[B \setminus K] \le Q[B \setminus L] + Q[B \cap V] \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

und damit

$$Q[B] = Q[B \cap K] + Q[B \setminus K] \le Q[K] + \varepsilon$$

Daraus folgt

$$Q[B] = \sup\Big\{ Q[K] \,\Big|\, K \in \mathcal{B}(\mathbb{R}^m) \text{ ist kompakt mit } K \subseteq B \Big\}$$

Daher ist Q regulär. □

20.2 Satz von Kolmogorov

Wir beweisen nun den angekündigten Satz von Kolmogorov:

20.2.1 Satz (Kolmogorov). *Sei $\{Q_J\}_{J\in\mathcal{H}(I)}$ eine projektive Familie von Wahrscheinlichkeitsmaßen mit $Q_J : \mathcal{B}(\mathbb{R}^J) \to [0,1]$ für alle $J \in \mathcal{H}(I)$. Dann gibt es genau ein Wahrscheinlichkeitsmaß $Q : \bigotimes_{i\in I} \mathcal{B}(\mathbb{R}) \to [0,1]$ derart, dass für alle $J \in \mathcal{H}(I)$*

$$Q_{\pi_J} = Q_J$$

gilt.

Beweis. Im Fall einer endlichen Indexmenge I ist nichts zu zeigen. Sei daher I unendlich. Sei ferner $\mathcal{Z}$ die Algebra der Zylindermengen auf $\mathbb{R}^I$ und für $J \in \mathcal{H}(I)$ sei $\mathcal{Z}_J = \pi_J^{-1}(\mathcal{B}(\mathbb{R}^J))$ die σ–Algebra der J–Zylinder.
Nach Satz 10.5.6 existiert genau ein Wahrscheinlichkeitsinhalt $Q : \mathcal{Z} \to [0,1]$ mit

$$Q[\pi_J^{-1}(C)] = Q_J[C]$$

für alle $J \in \mathcal{H}(I)$ und alle $C \in \mathcal{F}_J$. Wir zeigen im folgenden, dass Q $\emptyset$–stetig ist. Dann ist Q nach Lemma 10.1.2 σ–additiv und besitzt nach dem Satz von Caratheodory eine eindeutige Fortsetzung zu einem Wahrscheinlichkeitsmaß $\mathcal{F} \to [0,1]$, das wir wieder mit Q bezeichnen.
Der Wahrscheinlichkeitsinhalt Q ist nach Definition genau dann $\emptyset$–stetig, wenn für jede monoton fallende Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Z}$ mit $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$

$$\inf_{n\in\mathbb{N}} Q[A_n] = 0$$

gilt. Diese Bedingung ist gleichwertig damit, dass für jede monoton fallende Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Z}$ mit $\inf_{n\in\mathbb{N}} Q[A_n] > 0$

$$\bigcap_{n\in\mathbb{N}} A_n \neq \emptyset$$

gilt.
Sei also $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Z}$ eine monoton fallende Folge mit $\inf_{n\in\mathbb{N}} Q[A_n] > 0$ und sei

$$\alpha := \inf_{n\in\mathbb{N}} Q[A_n]$$

Dann gibt es eine streng monoton wachsende Folge $\{J_n\}_{n\in\mathbb{N}} \subseteq \mathcal{H}(I)$ mit

$$A_n \in \mathcal{Z}_{J_n}$$

für alle $n \in \mathbb{N}$.

(1) Für alle $n \in \mathbb{N}$ gibt es ein $C_n \in \mathcal{F}_{J_n}$ mit

$$A_n = \pi_{J_n}^{-1}(C_n)$$

und nach Lemma 20.1.1 gibt es eine kompakte Menge $K_n \in \mathcal{F}_{J_n}$ mit

$$K_n \subseteq C_n$$

und

$$Q_{J_n}[C_n \setminus K_n] < \frac{\alpha}{2^n}$$

(2) Für alle $k \in \mathbb{N}$ sei $B_k := \bigcap_{n=1}^k \pi_{J_n}^{-1}(K_n)$. Dann gilt $B_k \in \mathcal{Z}_{J_k}$ sowie

$$B_k = \bigcap_{n=1}^k \pi_{J_n}^{-1}(K_n) \subseteq \bigcap_{n=1}^k \pi_{J_n}^{-1}(C_n) = \bigcap_{n=1}^k A_n = A_k$$

und

$$A_k \setminus B_k = \left(\bigcap_{n=1}^k \pi_{J_n}^{-1}(C_n)\right) \setminus \left(\bigcap_{n=1}^k \pi_{J_n}^{-1}(K_n)\right) \subseteq \bigcup_{n=1}^k \pi_{J_n}^{-1}(C_n \setminus K_n)$$

und damit

$$\begin{aligned} Q[A_k \setminus B_k] &\leq Q\left[\bigcup_{n=1}^k \pi_{J_n}^{-1}(C_n \setminus K_n)\right] \\ &\leq \sum_{n=1}^k Q[\pi_{J_n}^{-1}(C_n \setminus K_n)] \\ &= \sum_{n=1}^k Q_{J_n}[C_n \setminus K_n] \\ &< \sum_{n=1}^k \frac{\alpha}{2^n} \\ &= \alpha \\ &\leq Q[A_k] \end{aligned}$$

Daher gilt $Q[B_k] = Q[A_k] - Q[A_k \setminus B_k] > 0$ und damit

$$B_k \neq \emptyset$$

(3) Es gibt daher eine Folge $\{x_k\}_{k\in\mathbb{N}}$ mit $x_k \in B_k = \bigcap_{n=1}^k \pi_{J_n}^{-1}(K_n)$ für alle $k \in \mathbb{N}$. Dann gilt für alle $n \in \mathbb{N}$ und $k \in \mathbb{N}(n)$

$$x_k \in \pi_{J_n}^{-1}(K_n)$$

und damit gilt für alle $n \in \mathbb{N}$, $k \in \mathbb{N}(n)$ und $j \in J_n$

$$\begin{aligned} \pi_j(x_k) &= \pi_{j,J_n}(\pi_{J_n}(x_k)) \\ &\in \pi_{j,J_n}(K_n) \end{aligned}$$

(4) Sei nun

$$J := \bigcup_{n \in \mathbb{N}} J_n$$

und sei $\{j_h \mid h \in \mathbb{N}\}$ eine Abzählung von J. Dann gibt es zu jedem $h \in \mathbb{N}$ ein $n \in \mathbb{N}$ mit $j_h \in J_n$ und

$$\pi_{j_h}(x_k) \in \pi_{j_h,J_n}(K_n)$$

für alle $k \in \mathbb{N}(n)$. Da K_n kompakt und die Projektion π_{j_h,J_n} stetig ist, ist auch $\pi_{j_h,J_n}(K_n)$ kompakt. Durch sukzessive Verdünnung erhält man daher für jedes $h \in \mathbb{N}$ eine Teilfolge $\{y_{h,k}\}_{k\in\mathbb{N}}$ von $\{x_k\}_{k\in\mathbb{N}}$ derart, dass die Folge $\{\pi_{j_h}(y_{h,k})\}_{k\in\mathbb{N}}$ konvergiert. Für alle $h \in \mathbb{N}$ sei $y_h := y_{h,h}$. Dann ist für alle $j \in J$ die Folge $\{\pi_j(y_h)\}_{h\in\mathbb{N}}$ konvergent und wir setzen

$$z_j := \lim_{h\to\infty} \pi_j(y_h)$$

(5) Für alle $n \in \mathbb{N}$ sei

$$z_{J_n} := \bigcap_{j \in J_n} \pi^{-1}_{j,J_n}(\{z_j\})$$

Für alle $k \in \mathbb{N}(n)$ gilt $x_k \in \pi^{-1}_{J_n}(K_n)$ und damit auch $y_k \in \pi^{-1}_{J_n}(K_n)$. Wir erhalten daher

$$\begin{aligned} z_{J_n} &= \bigcap_{j \in J_n} \pi^{-1}_{j,J_n}(\{z_j\}) \\ &= \bigcap_{j \in J_n} \pi^{-1}_{j,J_n}\left(\left\{\lim_{k\to\infty} \pi_j(y_k)\right\}\right) \\ &= \lim_{k\to\infty} \bigcap_{j \in J_n} \pi^{-1}_{j,J_n}(\{\pi_j(y_k)\}) \\ &= \lim_{k\to\infty} \pi_{J_n}(y_k) \\ &\in K_n \end{aligned}$$

Wegen

$$\pi^{-1}_{J_n}(\{z_{J_n}\}) = \pi^{-1}_{J_n}\left(\bigcap_{j \in J_n} \pi^{-1}_{j,J_n}(\{z_j\})\right) = \bigcap_{j \in J_n} \pi^{-1}_j(\{z_j\})$$

ist die Folge $\{\pi_{J_n}^{-1}(\{z_{J_n}\})\}_{n\in\mathbb{N}}$ monoton fallend und es gilt

$$\bigcap_{n\in\mathbb{N}} \pi_{J_n}^{-1}(\{z_{J_n}\}) \neq \emptyset$$

und damit

$$\emptyset \neq \bigcap_{n\in\mathbb{N}} \pi_{J_n}^{-1}(\{z_{J_n}\}) \subseteq \bigcap_{n\in\mathbb{N}} \pi_{J_n}^{-1}(K_n) \subseteq \bigcap_{n\in\mathbb{N}} \pi_{J_n}^{-1}(C_n) = \bigcap_{n\in\mathbb{N}} A_n$$

Daher ist Q $\emptyset$–stetig. □

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum. Eine Familie von Zufallsvariablen $\{X_i\}_{i\in I}$ heißt *stochastischer Prozess* auf $(\Omega, \mathcal{F}, P)$.

Sei $\{X_i\}_{i\in I}$ ein stochastischer Prozess auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Dann ist die Abbildung $X : \Omega \to \mathbb{R}^I$ mit den Koordinaten $\pi_i \circ X = X_i$ messbar und für alle $J \in \mathcal{H}(I)$ ist die Abbildung $X_J : \Omega \to \mathbb{R}^J$ mit den Koordinaten $\pi_{i,J} \circ X_J = X_i$ ebenfalls messbar; das Bildmaß $P_X : \bigotimes_{i\in I} \mathcal{B}(\mathbb{R}) \to [0,1]$ heißt *Verteilung* des stochastischen Prozesses und für $J \in \mathcal{H}(I)$ heißt das Bildmaß $P_{X_J} : \mathcal{B}(\mathbb{R}^J) \to [0,1]$ *endlichdimensionale Randverteilung* des stochastischen Prozesses bezüglich J. Für alle $J \in \mathcal{H}(I)$ gilt $X_J = \pi_J \circ X$ und damit $(P_{X_J}) = (P_X)_{\pi_J}$. Daher ist die Familie $\{P_{X_J}\}_{J\in\mathcal{H}(I)}$ projektiv.

Aus dem Satz von Kolmogorov ergibt sich die Existenz eines stochastischen Prozesses mit einer gegebenen projektiven Familie von endlichdimensionalen Randverteilungen:

20.2.2 Folgerung. *Sei $\{Q_J\}_{J\in\mathcal{H}(I)}$ eine projektive Familie von Wahrscheinlichkeitsmaßen mit $Q_J : \mathcal{B}(\mathbb{R}^J) \to [0,1]$ für alle $J \in \mathcal{H}(I)$. Dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und einen stochastischen Prozess $\{X_i\}_{i\in I}$ auf $(\Omega, \mathcal{F}, P)$ mit $P_{X_J} = Q_J$ für alle $J \in \mathcal{H}(I)$.*

Beweis. Aus der Projektivität der Familie $\{Q_J\}_{J\in\mathcal{H}(I)}$ und dem Satz von Kolmogorov folgt die Existenz eines eindeutig bestimmten Wahrscheinlichkeitsmaßes $Q : \bigotimes_{i\in I} \mathcal{B}(\mathbb{R}) \to [0,1]$ mit $Q_{\pi_J} = Q_J$ für alle $J \in \mathcal{H}(I)$. Sei

$$(\Omega, \mathcal{F}, P) := \left(\mathbb{R}^I, \bigotimes_{i\in I} \mathcal{B}(\mathbb{R}), Q\right)$$

und für alle $i \in I$ sei

$$X_i := \pi_i$$

Dann ist $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $\{X_i\}_{i\in I}$ ist ein stochastischer Prozess auf $(\Omega, \mathcal{F}, P)$ mit $P_{X_J} = Q_J$ für alle $J \in \mathcal{H}(I)$. □

Aufgaben

20.2.A Eine Funktion $\sigma : I \times I \to \mathbb{R}$ heißt

– *symmetrisch*, wenn für alle $i, j \in I$
$$\sigma(i,j) = \sigma(j,i)$$
gilt.

– *positiv definit*, wenn für alle $J \in \mathcal{H}(I)$ die Matrix $\{\sigma(i,j)\}_{i,j\in J}$ positiv definit ist.

Ist $\sigma : I \times I \to \mathbb{R}$ eine positiv definite symmetrische Funktion, so gilt für alle $i, j \in I$ mit $i \neq j$
$$\bigl(\sigma(i,j)\bigr)^2 < \sigma(i,i)\,\sigma(j,j)$$

20.2.B **Gauß–Prozess:** Ein stochastischer Prozess $\{X_i\}_{i\in I}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ heißt *Gauß–Prozess*, wenn für alle $J \in \mathcal{H}(I)$ der Zufallsvektor $X_J = \{X_i\}_{i\in J}$ eine Normal–Verteilung besitzt.

(1) Ist $\{X_i\}_{i\in I}$ ein Gauß–Prozess, so ist die Funktion $\sigma : I \times I \to \mathbb{R}$ mit
$$\sigma(i,j) := \mathrm{cov}\,[X_i, X_j]$$
symmetrisch und positiv definit.

(2) Sei $\mu : I \to \mathbb{R}$ eine Funktion und sei $\sigma : I \times I \to \mathbb{R}$ eine positiv definite symmetrische Funktion. Für alle $J \in \mathcal{H}(I)$ sei $\mu_J := (\mu(i))_{i\in J}$ und $\Sigma_J := (\sigma(i,j))_{i,j\in J}$. Dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und einen Gauß–Prozess $\{X_i\}_{i\in I}$ auf $(\Omega, \mathcal{F}, P)$ mit $P_{X_J} = \mathbf{N}(\mu_J, \Sigma_J)$ für alle $J \in \mathcal{H}(I)$.

20.2.C **Poisson–Prozess:** Sei $\alpha \in (0,\infty)$. Ein stochastischer Prozess $\{X_t\}_{t\in\mathbb{R}_+}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ heißt *Poisson–Prozess* zum Parameter α, wenn $X_0 = 0$ und für alle $n \in \mathbb{N}$ und für alle $t_0, t_1, \dots, t_n \in \mathbb{R}_+$ mit $0 = t_0 < t_1 < \dots < t_n$ und alle $k_1, \dots, k_n \in \mathbb{N}_0$
$$P\left[\bigcap_{j=1}^n \{X_{t_j} - X_{t_{j-1}} = k_j\}\right] = \prod_{j=1}^n e^{-(t_j - t_{j-1})\alpha}\,\frac{((t_j - t_{j-1})\alpha)^{k_j}}{k_j!}$$
gilt. Beweisen Sie die Existenz eines Poisson–Prozesses zum Parameter α.

Anhang

K.D. Schmidt, *Maß und Wahrscheinlichkeit*, 2. Aufl., Springer-Lehrbuch,
DOI 10.1007/978-3-642-21026-6,

A

Fakultät und Gamma–Funktion

Wir geben hier einen Überblick über die Eigenschaften von Fakultäten und der Gamma–Funktion und die daraus resultierenden Eigenschaften von Binomial–Koeffizienten und der Beta–Funktion.

A.1 Fakultät und Binomial–Koeffizient

Für $n \in \mathbb{N}_0$ sei

$$n! := \prod_{j=0}^{n-1} (n-j)$$

Die Zahl $n!$ heißt n *Fakultät*. Es gilt $0! = 1$ und

$$(n+1)! = (n+1) \cdot n!$$

Interpretation: $n!$ ist die Anzahl der Möglichkeiten, n unterscheidbare Objekte anzuordnen.

Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}_0$ sei

$$\binom{\alpha}{k} := \prod_{j=0}^{k-1} \frac{\alpha - j}{k - j}$$

Die Zahl $\binom{\alpha}{k}$ heißt *Binomial–Koeffizient* α *über* k. Es gilt $\binom{\alpha}{0} = 1$.

Für $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$ gilt

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!\,(n-k)!} & \text{falls } k \leq n \\ 0 & \text{sonst} \end{cases}$$

und für $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$ mit $k \leq n$ gilt

$$\begin{aligned}\binom{n}{0} &= 1 \\ \binom{n}{k} &= \binom{n}{n-k} \\ \binom{n+1}{k+1} &= \binom{n}{k} + \binom{n}{k+1}\end{aligned}$$

Durch vollständige Induktion ergibt sich $\binom{n}{k} \in \mathbb{N}_0$ und aus dem Binomischen Satz folgt

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Interpretation: $\binom{n}{k}$ ist die Anzahl der Möglichkeiten, k von n unterscheidbaren Objekten (ohne Berücksichtigung der Reihenfolge) auszuwählen. Insbesondere ist $\binom{n}{k}$ die Anzahl der k–elementigen Teilmengen der Menge $\{1, \ldots, n\}$ und 2^n ist die Anzahl aller Teilmengen der Menge $\{1, \ldots, n\}$.

A.2 Gamma–Funktion und Beta–Funktion

Die Funktion $\Gamma : (0, \infty) \to (0, \infty)$ mit

$$\Gamma(\gamma) := \int_0^\infty e^{-z} z^{\gamma-1}\, dz$$

heißt *Gamma–Funktion*. Es gilt

$$\begin{aligned}\Gamma(\tfrac{1}{2}) &= \sqrt{\pi} \\ \Gamma(1) &= 1 \\ \Gamma(\gamma{+}1) &= \gamma\, \Gamma(\gamma)\end{aligned}$$

Die letzte Gleichung heißt *5–Gamma–Formel*. Insbesondere gilt für $n \in \mathbb{N}_0$

$$\begin{aligned}\Gamma(n{+}1) &= n! \\ \binom{\gamma + n - 1}{n} &= \frac{\Gamma(\gamma{+}n)}{\Gamma(\gamma)\, n!}\end{aligned}$$

Die Funktion $\mathrm{B} : (0, \infty) \times (0, \infty) \to (0, \infty)$ mit

$$\mathrm{B}(\alpha, \beta) := \int_0^1 z^{\alpha-1} (1{-}z)^{\beta-1}\, dz$$

heißt *Beta–Funktion*. Es gilt

$$\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\, \Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

B

Vektorräume, Ordnung und Topologie

Wir geben hier einen Überblick über die linearen, ordnungstheoretischen und topologischen Strukturen in einem Vektorraum. Alle Vektorräume seien reell.

B.1 Vektorräume

Im gesamten Abschnitt sei E ein Vektorraum.

Eine Menge $C \subseteq E$ heißt
- *konvex*, wenn $ax + by \in C$ für alle $x, y \in C$ und für alle $a, b \in \mathbb{R}_+$ mit $a+b = 1$ gilt.
- *Kegel*, wenn $ax + by \in C$ für alle $x, y \in C$ und für alle $a, b \in \mathbb{R}_+$ gilt.
- *affin*, wenn $ax+by \in C$ für alle $x, y \in C$ und für alle $a, b \in \mathbb{R}$ mit $a+b = 1$ gilt.
- *linear*, wenn $ax + by \in C$ für alle $x, y \in C$ und für alle $a, b \in \mathbb{R}$ gilt.

Jede nichtleere lineare Teilmenge von E ist selbst ein Vektorraum und wird als *Unterraum* von E bezeichnet.

Sei $C \subseteq E$. Eine Funktion $h : C \to \mathbb{R}$ heißt
- *konvex*, wenn für alle $x, y \in C$ und $a, b \in \mathbb{R}_+$ mit $a+b = 1$ und $ax+by \in C$
$$h(ax+by) \leq ah(x) + bh(y)$$
gilt.
- *konkav*, wenn für alle $x, y \in C$ und $a, b \in \mathbb{R}_+$ mit $a+b = 1$ und $ax+by \in C$
$$h(ax+by) \geq ah(x) + bh(y)$$
gilt.

Offenbar ist h genau dann konvex, wenn die Funktion $-h : C \to \mathbb{R}$ mit $(-h)(x) := -h(x)$ konkav ist.

Sei $C \subseteq E$ und sei F ein Vektorraum. Eine Abbildung $T : C \to F$ heißt

– *linear*, wenn für alle $x, y \in C$ und $a, b \in \mathbb{R}$ mit $ax + by \in C$

$$T(ax+by) = aT(x) + bT(y)$$

gilt.

– *positiv linear*, wenn für alle $x, y \in C$ und $a, b \in \mathbb{R}_+$ mit $ax + by \in C$

$$T(ax+by) = aT(x) + bT(y)$$

gilt.

– *affin*, wenn es ein $c \in F$ und eine lineare Abbildung $S : C \to F$ gibt mit

$$T(x) = c + S(x)$$

für alle $x \in C$.

Eine lineare Abbildung $h : C \to \mathbb{R}$ wird auch als *Funktional* bezeichnet.

B.2 Ordnung

Sei E eine Menge. Eine Relation $\leq$ auf E heißt *Ordnungsrelation* (oder kurz *Ordnung*) auf E, wenn sie die folgenden Eigenschaften besitzt:

(i) Für alle $x \in E$ gilt $x \leq x$ (*Reflexivität*).

(ii) Für alle $x, y \in E$ mit $x \leq y$ und $y \leq x$ gilt $x = x$ (*Antisymmetrie*).

(iii) Für alle $x, y, z \in E$ mit $x \leq y$ und $y \leq z$ gilt $x \leq z$ (*Transitivität*).

In diesem Fall heißt $(E, \leq)$ *geordnete Menge*.

Sei $(E, \leq)$ eine geordnete Menge. Die Ordnungsrelation $\leq$ heißt *vollständig*, wenn für alle $x, y \in E$ mindestens eine der Eigenschaften $x \leq y$ und $y \leq x$ erfüllt ist. In diesem Fall heißt $(E, \leq)$ *vollständig geordnet*.

Sei $(E, \leq)$ eine geordnete Menge. Für $C \subseteq E$ heißt $z \in E$

– *Supremum* von C, wenn einerseits $x \leq z$ für alle $x \in C$ gilt und andererseits $z \leq u$ für jedes $u \in E$ mit $x \leq u$ für alle $x \in C$ gilt; in diesem Fall schreiben wir $z = \sup C$.

– *Infimum* von C, wenn einerseits $x \geq z$ für alle $x \in C$ gilt und andererseits $z \geq u$ für jedes $u \in E$ mit $x \geq u$ für alle $x \in C$ gilt; in diesem Fall schreiben wir $z = \inf C$.

Für $x, y \in E$ setzen wir im Fall der Existenz des Supremums bzw. des Infimums der Menge $\{x, y\}$

$$x \vee y := \sup\{x, y\}$$
$$x \wedge y := \inf\{x, y\}$$

Die geordnete Menge $(E, \leq)$ heißt *Verband*, wenn für alle $x, y \in E$ das Supremum und das Infimum der Menge $\{x, y\}$ existiert.

Sei E ein Vektorraum und sei $\leq$ eine Ordnungsrelation auf E. Die geordnete Menge $(E, \leq)$ heißt

– *geordneter Vektorraum*, wenn für alle $x, y, z \in E$ und $c \in \mathbb{R}_+$ mit $x \leq y$
$$
\begin{aligned}
x + z &\leq y + z \\
cx &\leq cy
\end{aligned}
$$
gilt.

– *Vektorverband*, wenn $(E, \leq)$ ein geordneter Vektorraum und ein Verband ist.

Ist $(E, \leq)$ ein geordneter Vektorraum, so ist die Menge $E_+ := \{x \in E \mid 0 \leq x\}$ ein Kegel, der als *positiver Kegel* von E (bezüglich $\leq$) bezeichnet wird. Ist $(E, \leq)$ ein Vektorverband, so existieren für alle $x \in E$ die Suprema
$$
\begin{aligned}
x^+ &:= x \vee 0 \\
x^- &:= (-x) \vee 0 \\
|x| &:= x \vee (-x)
\end{aligned}
$$
und man bezeichnet x^+ als den *Positivteil*, x^- als den *Negativteil* und $|x|$ als den *Betrag* von x; es gilt $x = x^+ - x^-$ und $|x| = x^+ + x^-$ sowie $x^+ \vee x^- = |x|$ und $x^+ \wedge x^- = 0$.

Ist $(E, \leq)$ ein Vektorverband, so heißt ein Unterraum $C \subseteq E$

– *Untervektorverband*, wenn für alle $x, y \in C$ auch (für das in E gebildete Supremum) $x \vee y \in C$ gilt.

– *Ideal*, wenn für alle $y \in C$ und $x \in E$ mit $|x| \leq |y|$ auch $x \in C$ gilt.

Jedes Ideal ist ein Untervektorverband.

B.3 Topologie

Sei E eine Menge. Eine Abbildung $d : E \times E \to \mathbb{R}_+$ heißt *Metrik* auf E, wenn sie die folgenden Eigenschaften besitzt:

(i) Für alle $x, y \in E$ gilt $d(x, y) = 0$ genau dann, wenn $x = y$ gilt (*Definitheit*).

(ii) Für alle $x, y \in E$ gilt $d(x, y) = d(y, x)$ (*Symmetrie*).

(iii) Für alle $x, y, z \in E$ gilt $d(x, z) \leq d(x, y) + d(y, z)$ (*Dreiecksungleichung*).

In diesem Fall heißt (E, d) *metrischer Raum*.

Sei (E, d) ein metrischer Raum. Eine Folge $\{x_n\}_{n \in \mathbb{N}} \subseteq E$ heißt *Cauchy–Folge*, wenn es zu jedem $\varepsilon \in (0, \infty)$ ein $p \in \mathbb{N}$ gibt mit $d(x_m, x_n) \leq \varepsilon$ für alle $m, n \in \mathbb{N}(p)$. Die Metrik d und damit auch der metrische Raum (E, d) heißt *vollständig*, wenn es zu jeder Cauchy–Folge $\{x_n\}_{n \in \mathbb{N}} \subseteq E$ ein $x \in E$ gibt mit $\lim_{n \to \infty} d(x_n, x) = 0$.

Sei E ein Vektorraum. Eine Abbildung $\|\,.\,\| : E \to \mathbb{R}_+$ heißt *Norm* auf E, wenn sie die folgenden Eigenschaften besitzt:

(i) Für alle $x \in E$ gilt $\|x\| = 0$ genau dann, wenn $x = 0$ gilt (*Definitheit*).
(ii) Für alle $x \in E$ und $c \in \mathbb{R}$ gilt $\|cx\| = |c|\,\|x\|$ (*absolute Homogenität*).
(iii) Für alle $x, y \in E$ gilt $\|x+y\| \leq \|x\| + \|y\|$ (*Dreiecksungleichung*).
In diesem Fall heißt $(E, \|\,.\,\|)$ *normierter Raum* und eine Menge $B \subseteq E$ heißt *beschränkt*, wenn es ein $c \in (0,\infty)$ gibt mit $B \subseteq \{x \in E \mid \|x\| \leq c\}$.

Sei $(E, \|\,.\,\|)$ ein normierter Raum. Dann ist die Abbildung $d : E \times E \to \mathbb{R}_+$ mit

$$d(x,y) := \|x-y\|$$

eine Metrik. Der normierte Raum $(E, \|\,.\,\|)$ heißt *vollständig*, wenn der metrische Raum (E, d) vollständig ist. Ein vollständiger normierter Raum heißt *Banach–Raum*.

Sei E ein Vektorraum. Eine Abbildung $\langle .\,,.\rangle : E \times E \to \mathbb{R}$ heißt *Skalarprodukt* auf E, wenn sie die folgenden Eigenschaften besitzt:
(i) Für alle $x \in E$ gilt $\langle x, x\rangle \geq 0$.
(ii) Für alle $x \in E$ gilt $\langle x, x\rangle = 0$ genau dann, wenn $x = 0$ gilt (*Definitheit*).
(iii) Für alle $x, y \in E$ gilt $\langle x, y\rangle = \langle y, x\rangle$ (*Symmetrie*).
(iv) Für alle $x, y \in E$ und $c \in \mathbb{R}$ gilt $\langle cx, y\rangle = c\,\langle x, y\rangle$ (*Homogenität*).
(v) Für alle $x, y, z \in E$ gilt $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ (*Additivität*).
In diesem Fall heißt $(E, \langle .\,,.\rangle)$ *Raum mit Skalarprodukt*.

Sei $(E, \langle .\,,.\rangle)$ ein Raum mit Skalarprodukt. Dann gilt für alle $x, y \in E$

$$\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$$

(*Ungleichung von Cauchy/Schwarz*). Aus dieser Ungleichung folgt, dass die Abbildung $\|\,.\,\| : E \to \mathbb{R}_+$ mit

$$\|x\| := \langle x, x\rangle^{1/2}$$

eine Norm ist. Der Raum mit Skalarprodukt $(E, \langle .\,,.\rangle)$ heißt *vollständig*, wenn der normierte Raum $(E, \|\,.\,\|)$ vollständig ist. Ein vollständiger Raum mit Skalarprodukt heißt *Hilbert–Raum*.

B.4 Ordnung und Topologie

Sei E ein Vektorraum und sei $\leq$ eine Ordnungsrelation und $\|\,.\,\|$ eine Norm auf E derart, dass
– $(E, \leq)$ ein Vektorverband ist,
– $(E, \|\,.\,\|)$ ein Banach–Raum ist, und
– $\|x\| \leq \|y\|$ für alle $x, y \in E$ mit $|x| \leq |y|$ gilt.
Dann heißt $(E, \leq, \|\,.\,\|)$ *Banach–Verband*.

Ein Banach–Verband, dessen Norm durch ein Skalarprodukt induziert wird, heißt *Hilbert–Verband*.

C

Der Euklidische Raum

Wir geben hier einen Überblick über die linearen, ordnungstheoretischen und topologischen Strukturen des Euklidischen Raumes.

C.1 Vektoren und Matrizen

Wir bezeichnen mit $\mathbb{R}^m$ den reellen Vektorraum aller (Spalten–)Vektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

(mit der koordinatenweise definierten Addition und Skalarmultiplikation). Wir bezeichnen das neutrale Element der Addition mit $\mathbf{0}$ und nennen $\mathbf{0}$ den *Nullvektor*. Mit Hilfe der *Einheitsvektoren* $\mathbf{e}_1, \ldots, \mathbf{e}_m \in \mathbb{R}^m$ lässt sich jeder Vektor $\mathbf{x} \in \mathbb{R}^m$ in der Form

$$\mathbf{x} = \sum_{i=1}^{m} x_i \mathbf{e}_i$$

darstellen. Wir setzen

$$\mathbf{1} := \sum_{i=1}^{m} \mathbf{e}_i$$

und nennen $\mathbf{1}$ den *Einsvektor*.

Wir bezeichnen mit $\mathbb{R}^{m \times n}$ den reellen Vektorraum aller Matrizen

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

(mit der koordinatenweise definierten Addition und Skalarmultiplikation). Wir bezeichnen das neutrale Element der Addition mit $\mathbf{O}$ und nennen $\mathbf{O}$ die *Nullmatrix*. Für eine Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ mit

$$\mathbf{A} = (a_{ij})_{i=1,\ldots,m,\, j=1,\ldots,n}$$

heißt die Matrix $\mathbf{A}' \in \mathbb{R}^{n \times m}$ mit

$$\mathbf{A}' := (a_{ji})_{j=1,\ldots,n,\, i=1,\ldots,m}$$

die zu $\mathbf{A}$ *transponierte Matrix*.

Eine Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ ist
– genau dann linear, wenn es eine Matrix $\mathbf{D} \in \mathbb{R}^{m \times n}$ gibt mit $T(\mathbf{x}) = \mathbf{D}\mathbf{x}$.
– genau dann affin, wenn es einen Vektor $\mathbf{c} \in \mathbb{R}^m$ und eine Matrix $\mathbf{D} \in \mathbb{R}^{m \times n}$ gibt mit $T(\mathbf{x}) = \mathbf{c} + \mathbf{D}\mathbf{x}$.

Ist $T : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung und $\mathbf{D} \in \mathbb{R}^{m \times n}$ eine Matrix mit $T(\mathbf{x}) = \mathbf{D}\mathbf{x}$, so sind die Spaltenvektoren der Matrix $\mathbf{D}$ gerade die Bilder der Einheitsvektoren des $\mathbb{R}^n$ unter T.

Wir betrachten nun den Fall $m = n$.

Wir bezeichnen das neutrale Element der Matrizenmultiplikation in $\mathbb{R}^{m \times m}$ mit $\mathbf{I}$ und nennen $\mathbf{I}$ die *Einheitsmatrix*. Eine Matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ heißt *invertierbar*, wenn es eine Matrix $\mathbf{B} \in \mathbb{R}^{m \times m}$ gibt mit $\mathbf{AB} = \mathbf{I}$ oder $\mathbf{BA} = \mathbf{I}$; in diesem Fall gelten sogar beide Gleichungen und die Matrix $\mathbf{B}$ ist eindeutig bestimmt und wird als *Inverse* von $\mathbf{A}$ und mit $\mathbf{A}^{-1}$ bezeichnet. Eine Matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ heißt *singulär*, wenn sie nicht invertierbar ist.

Eine Matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ heißt
– *Permutationsmatrix*, wenn sie in jeder Zeile und in jeder Spalte genau eine Eins und ansonsten nur Nullen enthält.
– *Orthogonalmatrix*, wenn $\mathbf{A}'\mathbf{A} = \mathbf{I}$ gilt.
– *Diagonalmatrix*, wenn für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$a_{ij} = 0$$

gilt.

Jede Permutationsmatrix ist eine Orthogonalmatrix, und jede Orthogonalmatrix ist invertierbar. Für einen Vektor $\mathbf{x} \in \mathbb{R}^m$ bezeichnen wir mit

$$\mathrm{diag}(\mathbf{x})$$

die Diagonalmatrix $\mathbf{A}$ mit

$$a_{ii} = x_i$$

für alle $i \in \{1, \ldots, m\}$. Es gilt $\mathbf{I} = \mathrm{diag}(\mathbf{1})$.

Für $k, l \in \{1, \ldots, m\}$ sei die Matrix $\mathbf{F}^{(kl)} \in \mathbb{R}^{m \times m}$ definiert durch

$$f_{ij}^{(kl)} := \begin{cases} 1 & \text{falls } (i,j) = (k,l) \\ 0 & \text{sonst} \end{cases}$$

Eine Matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ heißt *Elementarmatrix*, wenn es
- eine Matrix $\mathbf{F}^{(kl)}$ gibt mit $k \neq l$ und $\mathbf{A} = \mathbf{I} + \mathbf{F}^{(kl)}$ oder
- eine Matrix $\mathbf{F}^{(kk)}$ und ein $c \in \mathbb{R} \setminus \{0\}$ gibt mit $\mathbf{A} = \mathbf{I} + (c-1)\mathbf{F}^{(kk)}$.

Jede Elementarmatrix ist invertierbar.

C.1.1 Proposition. *Für eine Matrix* $\mathbf{A} \in \mathbb{R}^{m \times m}$ *sind äquivalent*:
(a) $\mathbf{A}$ *ist invertierbar.*
(b) $\mathbf{A}$ *ist ein Produkt von Elementarmatrizen.*
(c) $\mathbf{A}$ *ist ein Produkt von Permutationsmatrizen und Elementarmatrizen der Form* $\mathbf{I} + \mathbf{F}^{(12)}$ *oder* $\mathbf{I} + (c-1)\mathbf{F}^{(11)}$ *mit* $c \in \mathbb{R} \setminus \{0\}$.

Für einen Beweis der Äquivalenz von (a) und (b) in Proposition C.1.1 vgl. Koecher [1997; Kapitel 2, §6, Satz A]; die Äquivalenz von (b) und (c) ist dann klar.

Eine Matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ heißt
- *symmetrisch*, wenn $\mathbf{A}' = \mathbf{A}$ gilt.
- *positiv semidefinit*, wenn $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ für alle $\mathbf{x} \in \mathbb{R}^m$ gilt.
- *positiv definit*, wenn $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ für alle $\mathbf{x} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ gilt.

Eine symmetrische Matrix ist genau dann positiv definit, wenn sie positiv semidefinit und invertierbar ist.

C.1.2 Proposition. *Für eine Matrix* $\mathbf{A} \in \mathbb{R}^{m \times m}$ *sind äquivalent*:
(a) $\mathbf{A}$ *ist symmetrisch und positiv semidefinit.*
(b) *Es gibt eine Matrix* $\mathbf{B} \in \mathbb{R}^{m \times m}$ *mit* $\mathbf{A} = \mathbf{B}\mathbf{B}'$.

Für einen Beweis von Proposition C.1.2 vgl. Harville [1997; Corollary 14.3.10].

C.1.3 Proposition. *Für eine Matrix* $\mathbf{A} \in \mathbb{R}^{m \times m}$ *sind äquivalent*:
(a) $\mathbf{A}$ *ist symmetrisch und positiv definit.*
(b) *Es gibt eine invertierbare Matrix* $\mathbf{B} \in \mathbb{R}^{m \times m}$ *mit* $\mathbf{A} = \mathbf{B}\mathbf{B}'$.

Für einen Beweis von Proposition C.1.3 vgl. Harville [1997; Corollary 14.3.13].

C.2 Ordnung

Für $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ schreiben wir

$$\mathbf{x} \leq \mathbf{y}$$

wenn für alle $i \in \{1, \ldots, m\}$

$$x_i \leq y_i$$

gilt. Dann ist $\leq$ eine Ordnungsrelation und $(\mathbb{R}^m, \leq)$ ist ein Vektorverband. Die

Ordnungsrelation $\leq$ heißt *koordinatenweise Ordnung* auf $\mathbb{R}^m$. Für $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ schreiben wir

$$\mathbf{x} < \mathbf{y}$$

wenn für alle $i \in \{1, \ldots, m\}$

$$x_i < y_i$$

gilt.

C.3 Topologie

Die Abbildung $\langle . , . \rangle : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ mit

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}'\mathbf{y} = \sum_{i=1}^{m} x_i y_i$$

ist ein Skalarprodukt und heißt *Euklidisches Skalarprodukt*. Der Raum mit Skalarprodukt $(\mathbb{R}^m, \langle . , . \rangle)$ heißt m–dimensionaler *Euklidischer Raum*.

Die Abbildung $\| . \| : \mathbb{R}^m \to \mathbb{R}_+$ mit

$$\|\mathbf{x}\| := \langle \mathbf{x}, \mathbf{x} \rangle^{1/2} = \left(\sum_{i=1}^{m} x_i^2 \right)^{1/2}$$

ist eine Norm und heißt *Euklidische Norm*. Die Euklidische Norm ist zu jeder Norm auf $\mathbb{R}^m$ äquivalent.

Die Abbildung $d : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}_+$ mit

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$$

ist eine Metrik und heißt *Euklidische Metrik*.

Der metrische Raum $(\mathbb{R}^m, d)$ ist vollständig. Daher ist $(\mathbb{R}^m, \| . \|)$ ein Banach–Raum und $(\mathbb{R}^m, \langle . , . \rangle)$ ist ein Hilbert–Raum.

C.3.1 Proposition (Heine/Borel). *Für $K \subseteq \mathbb{R}^m$ sind folgende Aussagen äquivalent*:
(a) *K ist beschränkt und bezüglich der Normtopologie abgeschlossen.*
(b) *K ist bezüglich der Normtopologie kompakt.*

C.4 Ordnung und Topologie

Der Euklidische Raum $(\mathbb{R}^m, \leq, \| . \|)$ ist ein Banach–Verband (und sogar ein Hilbert–Verband).

Literaturverzeichnis

Aliprantis, C. D., and Burkinshaw, O. [1990]: *Principles of Real Analysis.* Boston: Academic Press.

Bauer, H. [1990]: *Maß– und Integrationstheorie.* Berlin – New York: DeGruyter.

Bauer, H. [1991]: *Wahrscheinlichkeitstheorie.* Berlin – New York: DeGruyter.

Behrends, E. [1987]: *Maß– und Integrationstheorie.* Berlin – Heidelberg – New York: Springer.

Billingsley, P. [1995]: *Probability and Measure.* Third Edition. New York – Chichester: Wiley.

Dudley, R. M. [1989]: *Real Analysis and Probability.* Pacific Grove (California): Wadsworth & Brooks/Cole.

Elstrodt, J. [1996]: *Maß– und Integrationstheorie.* Berlin – Heidelberg – New York: Springer.

Forster, O. [1983]: *Analysis 1.* Braunschweig: Vieweg.

Halmos, P. R. [1974]: *Measure Theory.* Berlin – Heidelberg – New York: Springer.

Harville, D. A. [1997]: *Matrix Algebra from a Statistician's Perspective.* Berlin – Heidelberg – New York: Springer.

Johnson, N. L., Kotz, S., and Balakrishnan, N. [1994]: *Continuous Univariate Distributions. Volume 1.* New York: Wiley.

Johnson, N. L., Kotz, S., and Balakrishnan, N. [1995]: *Continuous Univariate Distributions. Volume 2.* New York: Wiley.

Johnson, N. L., Kotz, S., and Balakrishnan, N. [1997]: *Discrete Multivariate Distributions.* New York: Wiley.

Johnson, N. L., Kotz, S., and Kemp, A.W. [1992]: *Univariate Discrete Distributions.* New York: Wiley.

Koecher, M. [1997]: *Lineare Algebra und Analytische Geometrie.* Vierte Auflage. Berlin – Heidelberg – New York: Springer.

Kolmogorov, A. N. [1933]: *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlin – Heidelberg – New York: Springer.

König, H. [1997]: *Measure and Integration.* Berlin – Heidelberg – New York: Springer.

Kotz, S., Balakrishnan, N., and Johnson, N. L. [2000]: *Continuous Multivariate Distributions. Volume 1.* New York: Wiley.

Krengel, U. [2002]: *Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Braunschweig – Wiesbaden: Vieweg.

Müller, P. H. (Hrsg.) [1991]: *Wahrscheinlichkeitstheorie und Mathematische Statistik – Lexikon der Stochastik.* Berlin: Akademie–Verlag.

Neveu, J. [1972]: *Martingales à Temps Discret.* Paris: Masson.

Schaefer, H. H. [1974]: *Banach Lattices and Positive Operators.* Berlin – Heidelberg – New York: Springer.

Schmidt, K. D. [1996]: *Lectures on Risk Theory.* Stuttgart: Teubner.

Schmidt, K. D. [2009]: *Versicherungsmathematik.* Berlin – Heidelberg – New York: Springer.

Schmitz, N. [1996]: *Vorlesungen über Wahrscheinlichkeitstheorie.* Stuttgart: Teubner.

Schubert, H. [1971]: *Topologie.* Stuttgart: Teubner.

Walter, W. [1990]: *Analysis I.* Berlin – Heidelberg – New York: Springer.

Symbolverzeichnis

Zahlenbereiche

$\mathbb{C}$	die Menge der komplexen Zahlen
$\mathbb{N}$	die Menge $\{1, 2, \dots\}$
$\mathbb{N}_0$	die Menge $\{0, 1, 2, \dots\}$
$\mathbb{N}(n)$	die Menge $\{n, n+1, n+2, \dots\}$
$\mathbb{Q}$	die Menge der rationalen Zahlen
$\mathbb{R}$	die Menge der reellen Zahlen
$\mathbb{R}_+$	die Menge $[0, \infty)$ der positiven reellen Zahlen
$\bar{\mathbb{R}}$	die Menge $[-\infty, \infty]$ der erweiterten reellen Zahlen
$\mathbb{Z}$	die Menge der ganzen Zahlen

Reelle Zahlen

$x \vee y$	$:= \max\{x, y\}$ (Maximum von $x, y \in \mathbb{R}$)
$x \wedge y$	$:= \min\{x, y\}$ (Minimum von $x, y \in \mathbb{R}$)
x^+	$:= x \vee 0$ (Positivteil von $x \in \mathbb{R}$)
x^-	$:= (-x) \vee 0$ (Negativteil von $x \in \mathbb{R}$)
$\lvert x \rvert$	$:= x \vee (-x)$ (Betrag von $x \in \mathbb{R}$)

Vektoren

$\mathbf{0}$	Nullvektor des Euklidischen Raumes
$\mathbf{1}$	Einsvektor des Euklidischen Raumes
$\mathbf{e}_i$	Einheitsvektor des Euklidischen Raumes
$\mathbf{x}$	Vektor des Euklidischen Raumes

Matrizen

$\mathbf{O}$	Nullmatrix
$\mathbf{I}$	Einheitsmatrix
$\mathbf{A}'$	Transponierte der Matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse einer invertierbaren Matrix $\mathbf{A}$
$\mathrm{diag}(\mathbf{x})$	Diagonalmatrix zum Vektor $\mathbf{x}$

Mengen

χ_A	Indikatorfunktion der Menge A
2^A	Potenzmenge der Menge A
$\overline{A}$	Komplement der Menge A (wenn die Grundmenge klar ist)
$\|A\|$	Anzahl der Elemente der Menge A
$\{a_i\}_{i\in I} \subseteq A$	Familie von Elementen aus der Menge A
$\prod_{i\in I} A_i$	Produkt der Familie $\{A_i\}_{i\in I}$
$\sum_{i\in I} A_i$	Vereinigung einer disjunkten Familie $\{A_i\}_{i\in I}$
A^m	$:= \prod_{i=1}^m A$
A^I	$:= \prod_{i\in I} A$
$A \setminus B$	$:= A \cap \overline{B}$ (Differenz)
$A \triangle B$	$:= (A \setminus B) \cup (B \setminus A)$ (symmetrische Differenz)

Mengensysteme

$\delta(\mathcal{E})$	das von $\mathcal{E}$ erzeugte Dynkin–System
$\varrho(\mathcal{E})$	der von $\mathcal{E}$ erzeugte Ring
$\sigma(\mathcal{E})$	die von $\mathcal{E}$ erzeugte σ–Algebra
$\tau(\mathcal{E})$	die von $\mathcal{E}$ erzeugte Topologie
$\mathcal{B}(\mathbb{R})$	Borelsche σ–Algebra auf $\mathbb{R}$
$\mathcal{B}(\bar{\mathbb{R}})$	Borelsche σ–Algebra auf $\bar{\mathbb{R}}$
$\mathcal{B}(\mathbb{R}^m)$	Borelsche σ–Algebra auf $\mathbb{R}^m$
$\mathcal{H}(I)$	Familie der endlichen nichtleeren Teilmengen von I
$\mathcal{J}(\mathbb{R})$	Halbring der halboffenen Intervalle auf $\mathbb{R}$
$\mathcal{J}(\bar{\mathbb{R}})$	Halbring der halboffenen Intervalle auf $\bar{\mathbb{R}}$
$\mathcal{J}(\mathbb{R}^m)$	Halbring der halboffenen Intervalle auf $\mathbb{R}^m$

Mengenfunktionen

$\boldsymbol{\delta}_\omega$	Dirac–Maß
$\boldsymbol{\zeta}$	Zählmaß
$\boldsymbol{\zeta}_C$	lokales Zählmaß
$\boldsymbol{\lambda}$	Lebesgue–Maß
μ^+	Positivteil von μ
μ^-	Negativteil von μ
μ_f	Bildmaß von μ unter f
$\mu \ll \nu$	μ ist ν–stetig
$\mu \approx \nu$	μ und ν sind äquivalent
$\mu \perp \nu$	μ und ν sind singulär

Messbare Funktionen

$\mathcal{L}^0(\mathcal{F})$	96
$\mathcal{L}^0(\mathcal{F},\mu)$	103
$\mathcal{L}^p_\mu(\mathcal{F})$	136
$\mathcal{L}^\infty(\mathcal{F})$	96
$=_\mu$	103

$[f]_\mu$	103
$L^0(\mathcal{F},\mu)$	104
$L^p(\mathcal{F},\mu)$	136
$L^\infty(\mathcal{F},\mu)$	106

Wahrscheinlichkeitstheorie

$P[A]$	Wahrscheinlichkeit von A
$E[X]$	Erwartungswert von X
$\mathrm{var}\,[X]$	Varianz von X
$\mathrm{cov}\,[X,Y]$	Kovarianz von X und Y
C_X	kumulantenerzeugende Funktion von X
F_X	Verteilungsfunktion von X
G_X	Überlebensfunktion von X
m_X	wahrscheinlichkeitserzeugende Funktion von X
M_X	momenterzeugende Funktion von X
P_X	Verteilung von X
ψ_X	charakteristische Funktion von X

Bedingte Wahrscheinlichkeit nach einem Ereignis

$P[A\|C]$	bedingte Wahrscheinlichkeit von A unter C
$E[X\|C]$	bedingter Erwartungswert von X unter C

Bedingte Wahrscheinlichkeit nach einer Unter–$\boldsymbol{\sigma}$–Algebra

$P^\mathcal{G}(A)$	bedingte Wahrscheinlichkeit von A unter $\mathcal{G}$
$E^\mathcal{G}(X)$	bedingte Erwartung von X unter $\mathcal{G}$
$\mathrm{var}^\mathcal{G}(X)$	bedingte Varianz von X unter $\mathcal{G}$
$\mathrm{cov}^\mathcal{G}(X,Y)$	bedingte Kovarianz von X und Y unter $\mathcal{G}$
$P_X^\mathcal{G}$	bedingte Verteilung von X unter $\mathcal{G}$
$L^0(\mathcal{F},\mathcal{G})$	410
$L^{1,\mathcal{G}}(\mathcal{F})$	417
$L^{2,\mathcal{G}}(\mathcal{F})$	452

Verteilungen

$\boldsymbol{\delta}_z$	Dirac–Verteilung
χ^2_n	χ^2–Verteilung
$\mathbf{t}_n$	t–Verteilung
$\mathbf{B}(\vartheta)$	Bernoulli–Verteilung
$\mathbf{B}(n,\vartheta)$	Binomial–Verteilung
$\mathbf{Be}(\alpha,\beta)$	Beta–Verteilung
$\mathbf{Ca}(\alpha,\beta)$	Cauchy–Verteilung
$\mathbf{Dir}(\eta,\boldsymbol{\eta})$	Dirichlet–Verteilung
$\mathbf{Exp}(\alpha)$	Exponential–Verteilung
$\mathbf{Ga}(\alpha,\gamma)$	Gamma–Verteilung

$\mathbf{Geo}(n,\vartheta)$	geometrische Verteilung
$\mathbf{H}(n,N,K)$	hypergeometrische Verteilung
$\mathbf{Log}(\vartheta)$	logarithmische Verteilung
$\mathbf{M}(n,\boldsymbol{\vartheta})$	Multinomial–Verteilung
$\mathbf{N}(\mu,\sigma^2)$	Normal–Verteilung
$\mathbf{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$	Normal–Verteilung
$\mathbf{NB}(\alpha,\vartheta)$	Negativbinomial–Verteilung
$\mathbf{NM}(\alpha,\boldsymbol{\vartheta})$	Negativmultinomial–Verteilung
$\mathbf{P}(\alpha)$	Poisson–Verteilung
$\mathbf{Pa}(\alpha,\beta)$	Pareto–Verteilung europäischer Art
$\mathbf{Pa}^*(\alpha,\beta)$	Pareto–Verteilung amerikanischer Art
$\mathbf{PE}(n,\eta,\boldsymbol{\eta})$	Pólya/Eggenberger–Verteilung
$\mathbf{PH}(n,N,\mathbf{K})$	polyhypergeometrische Verteilung
Pólya(n,α,β)	Pólya–Verteilung
$\mathbf{U}(a,b)$	uniforme Verteilung
$\mathbf{U}(C)$	uniforme Verteilung

Spezielle Funktionen

B	Beta–Funktion
Γ	Gamma–Funktion
Φ	Verteilungsfunktion der Standardnormal–Verteilung

Sachverzeichnis