René L. Schilling
**Martingale und Prozesse**
De Gruyter Studium

## Weitere empfehlenswerte Titel

Das Lehrbuch-Set *Moderne Stochastik* umfasst die drei Bände *Maß und Integral*, *Wahrscheinlichkeit* sowie *Martingale und Prozesse* und behandelt damit detailliert aber kompakt den gesamten Lernstoff der Wahrscheinlichkeitstheorie nicht nur für Studierende der Mathematik, sondern auch der Natur-, Wirtschafts- und Ingenieurwissenschaften.

*Moderne Stochastik. Eine Einführung für Bachelor-Studenten*
*Lehrbuch-Set*
René L. Schilling, 2018
ISBN 978-3-11-053749-9

*Maß und Integral*
René L. Schilling, 2015
ISBN 978-3-11-034814-9, e-ISBN (PDF) 978-3-11-035064-7,
e-ISBN (EPUB) 978-3-11-038332-4

*Wahrscheinlichkeit*
René L. Schilling, 2017
ISBN 978-3-11-035065-4, e-ISBN (PDF) 978-3-11-035066-1,
e-ISBN (EPUB) 978-3-11-038750-6

*Martingale und Prozesse*
René L. Schilling, 2018
ISBN 978-3-11-035067-8, e-ISBN (PDF) 978-3-11-035068-5,
e-ISBN (EPUB) 978-3-11-038751-3

René L. Schilling

# Martingale und Prozesse

Eine Einführung für Bachelor-Studenten

DE GRUYTER

**Mathematics Subject Classification 2010**
Primary: 60-01. Secondary: 60G42; 60G50; 60G40; 60J45; 60Fxx.

**Autor**

Prof. Dr. René L. Schilling
Technische Universität Dresden
Institut für Mathematische Stochastik
D-01062 Dresden
Germany

rene.schilling@tu-dresden.de
www.math.tu-dresden.de/sto/schilling

**Weiterführendes Material**

www.motapa.de/maps

ISBN 978-3-11-035067-8
e-ISBN (PDF) 978-3-11-035068-5
e-ISBN (EPUB) 978-3-11-038751-3

**Library of Congress Control Number: 2018939058**

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

© 2018 Walter de Gruyter GmbH, Berlin/Boston
Einbandabbildung: »*Advantages of the Martingale*« – Holzstich eines unbekannten englischen Künstlers des 19. Jhdts. aus dem Privatbesitz des Autors
Druck und Bindung: CPI books GmbH, Leck

www.degruyter.com

# Vorwort

Dieses Lehrbuch ist eine Einführung in die Theorie der Martingale und Irrfahrten (*random walk*) in diskreter Zeit. Es richtet sich an Studierende der Mathematik und Physik ab dem zweiten Studienjahr. Mein Ziel ist es, in kompakter und eingängiger Weise zentrale Techniken und Resultate der Stochastik darzustellen und so eine Grundlage für weiterführende Vorlesungen zu geben.

Der Text folgt meinen Vorlesungen an der TU Dresden, er kann als Grundlage oder Begleittext für eine Vorlesung aber auch zum Selbststudium verwendet werden. Voraussetzung für das Verständnis des vorliegenden Bandes sind Grundlagen der Maß- & Integrationstheorie und der Wahrscheinlichkeitsrechnung, wie ich sie in den ersten beiden Bänden dieser Reihe entwickelt habe. Relevant sind vor allem Kapitel 1–16 aus *Maß und Integral* und Kapitel 1–7 & 9 aus *Wahrscheinlichkeit*. Diese Bände werden im Text als MI und WT zitiert, zahlreiche Querverweise auf die entsprechenden Sätze vereinfachen die Lektüre.

In Kapitel 1–7 und 9 werden die Grundlagen der Theorie der diskreten Martingale entwickelt, die dann in Kapitel 8 auf klassische Sätze der Wahrscheinlichkeitstheorie (Null-Eins–Gesetze, Summen von unabhängigen Zufallsvariablen, zentraler Grenzwertsatz) mit Martingalmethoden bewiesen und für Folgen von Martingaldifferenzen erweitert. Zusammen mit WT Kapitel 1–7, 9 kann man so eine Einführung in die Wahrscheinlichkeitstheorie auf der Grundlage von Martingalen geben. Diesem Teil schließt sich das Studium zufälliger Irrfahrten (*random walks*) an, das exemplarisch in die Gedankenwelt von zeitdiskreten stochastischen Prozessen einführt. Der Schwerpunkt liegt auf der Untersuchung von Rekurrenz und Transienz sowie dem Zusammenhang mit (diskreten) Randwertproblemen.

Die Auswahl der Themen und Techniken ist natürlich subjektiv, dennoch will ich dem Leser ein breites Spektrum von klassischen und modernen Methoden vorstellen, das auf eine weitere Spezialisierung optimal vorbereitet. Dabei war mein Leitmotiv die Frage „Was wird später im Studium und in Anwendungen wirklich benötigt“, wobei meine eigenen Forschungsinteressen – die Theorie der stochastischen Prozesse – im Vordergrund stehen.

Für das tiefere Verständnis ist es wichtig, dass der Leser sich mit der Materie selbständig auseinandersetzt. Zum einen sind dafür die Übungsaufgaben gedacht (vollständige Lösungen gibt es unter www.motapa.de/maps), andererseits weise ich im laufenden Text mit dem Symbol [✍] auf (bisweilen nicht ganz so offensichtliche) Lücken hin, die der Leser selbst ausfüllen sollte. Auf

---

- wichtige Schreibweisen,
- Gegenbeispiele,
- typische Fallen und versteckte Schwierigkeiten

---

wird durch derart markierte Absätze aufmerksam gemacht.

Vom Umfang entsprechen die Kapitel 1–7 & 9, abgerundet um einige Wahlthemen, einer dreistündigen Vorlesung; etwa 4–5 Textseiten können in einer Vorlesungs-Doppelstunde durchgenommen werden. Die mit dem Symbol ⧫ gekennzeichneten Abschnitte sind als Ergänzung gedacht und können je nach Zeit und Zielsetzung ausgewählt werden. Sie sind auch als Themen für ein Seminar geeignet. Eine Übersicht über die Abhängigkeit der einzelnen Kapitel findet sich auf Seite VII.

Dieser Text ist aus Vorlesungen entstanden, und ich danke meinen Studenten, Schülern und Kollegen für ihr Interesse und ihre Mitarbeit. Ich danke ganz besonders Dr. Wojciech Cygan, Dr. Victoria Knopova und Dr. Franziska Kühn für die genaue Durchsicht des Texts und ihre kritischen und hilfreichen Kommentare. Herr Dr. Böttcher hat in bewährter Weise die Grafiken erstellt und das Kapitel zur Kopplung gelesen. Herrn Prof. Jacob danke ich für zahlreiche Diskussionen und die Möglichkeit, seine Privatbibliothek zu nutzen.

Die Zusammenarbeit mit dem Verlag de Gruyter, allen voran Frau Schedensack und Herr Lindenhain, war wieder sehr angenehm und hat wesentlich zum Gelingen dieses Buchs beigetragen. Meine Frau machte es möglich, dass ich „ungezügelt“ an den Martingalen arbeiten konnte – danke!

Dresden, Februar 2018 René L. Schilling

## Mathematische Grundlagen, weiterführende Literatur

Für die Lektüre dieses Texts werden Grundlagen der Maß- und Integrationstheorie und der Wahrscheinlichkeitstheorie benötigt, etwa im Umfang der Kapitel 1–16 meines Lehrbuchs *Maß und Integral* und Kapitel 1–7 und 9 meines Lehrbuchs *Wahrscheinlichkeit*. Beide Bände sind in gleicher Ausstattung wie dieses Buch beim Verlag de Gruyter erschienen.

### Grundlagen

Schilling, R.L.: *Maß und Integral*. De Gruyter, Berlin 2015 (zitiert als MI).

Schilling, R.L.: *Measures, Integrals and Martingales*. Cambridge University Press, Cambridge $^{2}$2017 (zitiert als MIMS).

Schilling, R.L.: *Wahrscheinlichkeit*. De Gruyter, Berlin 2017 (zitiert als WT).

Georgii, H.-O.: *Stochastik. Einführung in die Wahrscheinlichkeitstheorie und Statistik*. De Gruyter, Berlin $^{4}$2009.

Krengel, U.: *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Vieweg, Wiesbaden $^{8}$2007.

**Weiterführende Literatur**

Khoshnevisan, D., Schilling, R.L.: *From Lévy-Type Processes to Parabolic SPDEs*. Birkhäuser, Cham 2016.

Norris, J.: *Markov Chains*. Cambridge University Press, Cambridge 1997.

Schilling, R.L., Partzsch, L.: *Brownian Motion. An Introduction to Stochastic Processes*. De Gruyter, Berlin $^2$2014.

## Abhängigkeit der einzelnen Kapitel

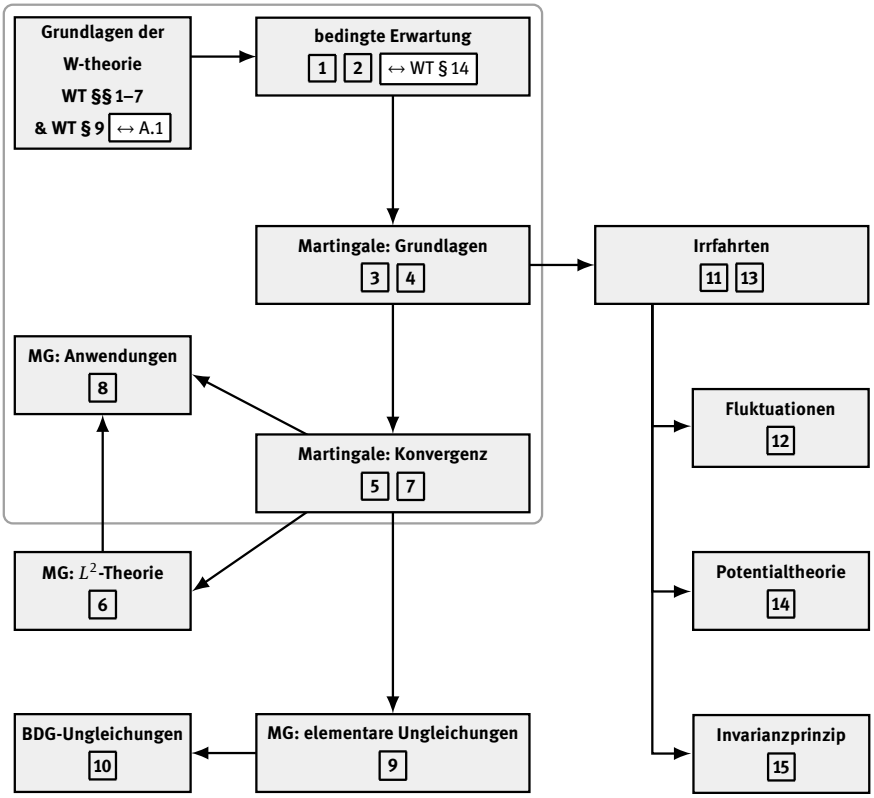


**Abb. 1:** Die Grafik zeigt die Abhängigkeit der Themen dieses Buchs, gestrichelte Pfeile stehen für kleinere oder indirekte Abhängigkeiten; „↔ WT *n*“ bedeutet, dass das entsprechende Kapitel durch Kapitel *n* meines Lehrbuchs „Wahrscheinlichkeit“ (WT) ersetzt werden kann. Vorkenntnisse aus der Wahrscheinlichkeitstheorie werden etwa im Umfang der Kapitel WT 1–7 vorausgesetzt, eine knappe Übersicht über die Konvergenzarten der Wahrscheinlichkeitstheorie (vgl. WT Kapitel 9) wird im Anhang A.1 & A.2 gegeben. Die mit der gestrichelten Linie zusammengefassten Kapitel sind eine „Einführung in die Stochastik“ mit Martingalmethoden, die den klassischen Zugang (WT Kapitel 8–13) ersetzen können.

# Bezeichnungen

### Allgemeines & Konventionen

| | |
|---|---|
| [MI, Satz $n.m$] | Verweis *Maß und Integral* |
| [WT, Satz $n.m$] | Verweis *Wahrscheinlichkeit* |
| positiv | stets im Sinne $\geqslant 0$ |
| negativ | stets im Sinne $\leqslant 0$ |
| $\mathbb{N}$, $\mathbb{N}_0$ | $1, 2, 3, \ldots$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ |
| $\inf \emptyset$ | $\inf \emptyset = +\infty$ |
| $a \vee b$, $a \wedge b$ | $\max\{a, b\}$, $\min\{a, b\}$ |
| $\lfloor x \rfloor$ | $\max\{n \in \mathbb{Z} : n \leqslant x\}$ |
| $\lvert x \rvert$ | Euklidische Norm in $\mathbb{R}^d$ |

### Mengen

| | |
|---|---|
| $\# A$, $\lvert A \rvert$ | Kardinalität der Menge $A$ |
| $\subset$ | Teilmenge (inkl. „=“) |
| $\dot{\bigcup}$ | Vereinigung paarweise disjunkter Mengen |
| $A^c$ | Komplement der Menge $A$ |
| $\overline{A}$ | Abschluss der Menge $A$ |
| $A \overset{\text{f.s.}}{=} B$ | $\mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A) = 0$, 48 |
| $B_r(x)$ | offene Kugel um $x$, Radius $r$ |
| $A_n \uparrow A$ | $A_n \subset A_{n+1} \subset \ldots$ & $A = \bigcup_n A_n$ |
| $B_n \downarrow B$ | $B_n \supset B_{n+1} \supset \ldots$ & $B = \bigcap_n B_n$ |
| $\liminf_{n\to\infty} A_n$ | $\bigcup_{k\in\mathbb{N}} \bigcap_{n\geqslant k} A_n$ |
| $\limsup_{n\to\infty} B_n$ | $\bigcap_{k\in\mathbb{N}} \bigcup_{n\geqslant k} B_n$ |
| $A \perp\!\!\!\perp F$ | $A$ und $F$ sind unabhängig (analog für Mengensysteme) |
| $\mathscr{A}$, $\mathscr{F}$ | generische $\sigma$-Algebren |
| $\mathscr{A} \times \mathscr{F}$ | $\{A \times F : A \in \mathscr{A},\ F \in \mathscr{F}\}$ „Rechtecke“ |
| $\mathscr{A} \otimes \mathscr{F}$ | Produkt-$\sigma$-Algebra |
| $\mathscr{B}(E)$ | Borelmengen in $E$ |
| $\sigma(\mathscr{G})$, $\sigma(X)$ | erzeugte $\sigma$-Algebra |
| $(\mathscr{F}_n)_{n\in\mathbb{N}_0}$ | Filtration, 17 |
| $\mathscr{F}_\infty$ | $\sigma(\mathscr{F}_n, n \in \mathbb{N}_0) = \sigma\big(\bigcup_n \mathscr{F}_n\big)$ |

### Maße & Verteilungen

| | |
|---|---|
| $\mathbb{P}$, $\mathbb{E}$ | W-Maß, Erwartungswert |
| $\mathbb{P}^x$, $\mathbb{E}^x$ | $-$, $-$ wenn $X_0 = x$, 147 |
| $\mathbb{P}_X$, $\mathbb{P}(X \in \bullet)$ | Verteilung der ZV $X$ |
| $\mathbb{P}(\bullet \mid \mathscr{F})$ | bedingte W-keit, 14 |
| $\mathbb{E}(\bullet \mid \mathscr{F})$ | bedingte Erwartung, 7 |
| $\delta_x$ | Dirac-Maß in $x$ |
| $\mu \otimes \nu$ | Produkt von Maßen |
| $\mathsf{N}(\mu, \sigma^2)$ | Normalverteilung, Mittel $\mu$, Varianz $\sigma^2$ |
| $\mathsf{N}(m, \Gamma)$ | Normalverteilung in $\mathbb{R}^d$, Mittel $m$, Kovarianzmatrix $\Gamma$ |

### Zufallsvariable & Funktionen

| | |
|---|---|
| $X, Y, Z, \xi$ | Zufallsvariable (ZV) |
| $X^+, X^-$ | Positivteil, Negativteil |
| $\{X \in B\}$ | $\{\omega : X(\omega) \in B\}$ |
| $\{X \geqslant a\}$ | $\{\omega : X(\omega) \geqslant a\}$ usw. |
| $X \perp\!\!\!\perp Y$ | $X$, $Y$ sind unabhängig |
| $X \sim Y$ | $X$ ist wie $Y$ verteilt |
| $X \sim \mu$ | $X$ hat Verteilung $\mu$ |
| $X = (X_n)_{n\in\mathbb{N}_0}$ | stochastischer Prozess, 109 |
| $(X_n, \mathscr{F}_n)_{n\in\mathbb{N}_0}$ | adaptierter stochastischer Prozess, 17 |
| $X^T$ | gestoppter Proz. $X_n^T := X_{n\wedge T}$, 31 |
| $\langle X \rangle$ | Kompensator, 25 |
| $[X]$ | quadratische Variation, 93 |
| $X^*$ | $\sup_n \lvert X_n \rvert$, 89 |
| $C \bullet X$ | Martingaltransformation, 26 |
| $T$ | Stoppzeit $\{T \leqslant n\} \in \mathscr{F}_n$, 31 |
| $T_x$ | $\inf\{n : X_n = x\}$, 113*f.*, 128 |
| $u_n$ | Rückkehrwahrscheinlichkeit $\mathbb{P}(X_n = 0)$, 116 |
| $f_n$ | $\mathbb{P}(T_0 = n)$, 116 |
| $\mathbb{1}_A$ | $\mathbb{1}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ |
| $C(E)$ | stetige Funktionen auf $E$ |
| $C_b(E)$ | beschränkte stetige Funktionen |
| $L^p$, $L^\infty$ | ZV mit $\mathbb{E}\left[\lvert X \rvert^p\right] < \infty$ |
| $L^p(\mathscr{F})$, $L^\infty(\mathscr{F})$ | betont die $\mathscr{F}$-Messbarkeit |

### Abkürzungen

| | |
|---|---|
| BL | Beppo Levi |
| CLT | zentraler Grenzwertsatz |
| f.s. | fast sicher |
| ggi | gleichgradig integrierbar |
| iid | unabhängig und identisch verteilt |
| MG | Martingal(e), 17 |
| SLLN | starkes Gesetz d. großen Zahlen |
| (S)RW | (einfache) Irrfahrt, 110 |
| W– | Wahrscheinlichkeit(s)– |
| ZV | Zufallsvariable(n) |
| $\cap$/$\cup$-stabil | Familie enthält endliche Schnitte/Vereinigungen |
| [✍] | selbst rechnen! |

# Inhalt

**Vorwort —— V**

**Mathematische Grundlagen, weiterführende Literatur —— VI**

**Abhängigkeit der einzelnen Kapitel —— VII**

**Bezeichnungen —— VIII**

**1 Fair Play —— 1**
⧫Ein unvorteilhaftes „faires“ Spiel —— **5**

**2 Bedingte Erwartung —— 7**
Bedingte Wahrscheinlichkeiten —— **14**

**3 Martingale —— 17**
Die Doob–Zerlegung —— **23**
Der Kompensator —— **25**
Die Martingaltransformation —— **26**

**4 Stoppen und Lokalisieren —— 31**
⧫Lokale Martingale —— **36**
⧫Verallgemeinerte Martingale —— **40**

**5 Konvergenz von Martingalen —— 43**
Rückwärtsmartingale —— **46**
⧫Konvergenzmengen von Martingalen und Submartingalen —— **48**

**6 ⧫$L^2$-Martingale —— 51**

**7 Gleichgradig integrierbare Martingale —— 57**
Gleichgradige Integrierbarkeit und *optional stopping* —— **63**
Rückwärtsmartingale und gleichgradige Integrierbarkeit —— **65**

**8 ⧫Einige klassische Resultate der W-Theorie —— 67**
8.1 Lévys Konvergenzsatz und Kolmogorovs 0-1–Gesetz —— **67**
8.2 Rückwärtsmartingale und Kolmogorovs $L^1$-SLLN —— **68**
8.3 Das 0-1–Gesetz von Hewitt–Savage —— **69**
8.4 Variationen zu einem Thema von Borel–Cantelli —— **71**
8.5 Lévys Konvergenzsatz für Reihen von unabhängigen ZV —— **73**

8.6 Kolmogorovs Drei-Reihen-Satz — **75**
8.7 Ein einfacher zentraler Grenzwertsatz — **78**
8.8 Martingale mit allgemeiner Indexmenge — **81**
8.9 Der Satz von Radon-Nikodým — **82**

**9 Elementare Ungleichungen für Martingale — 87**
♦Die Ungleichung von Azuma–Hoeffding — **91**

**10 ♦Die Burkholder–Davis–Gundy Ungleichungen — 93**
10.1 Die adaptierte quadratische Variation — **97**
10.2 Die Krickeberg-Zerlegung — **99**
10.3 Die Ungleichungen von Burkholder — **101**
10.4 Die Zerlegung und die Ungleichungen von Davis — **105**

**11 Zufällige Irrfahrten auf $\mathbb{Z}^d$ – erste Schritte — 109**
*Drunkard's walk* und *gambler's ruin* — **109**
Rekurrenz und Transienz — **116**

**12 ♦Fluktuationen einer einfachen Irrfahrt auf $\mathbb{Z}$ — 121**

**13 Rekurrenz und Transienz allgemeiner Irrfahrten — 131**
Der Satz von Chung–Fuchs für integrierbare Irrfahrten — **131**
Das Chung–Fuchs Kriterium — **139**

**14 ♦Irrfahrten und Analysis — 147**
Irrfahrten mit beliebigen Startwerten — **147**
Elemente der Potentialtheorie für einfache Irrfahrten — **150**
Randwertprobleme und einfache Irrfahrten — **154**
Kopplung von Irrfahrten — **160**

**15 ♦Donskers Invarianzprinzip und die Brownsche Bewegung — 169**

**A Anhang — 179**
A.1 Konvergenz in Wahrscheinlichkeit — **179**
A.2 Konvergenz in Verteilung — **181**
A.3 Das Faktorisierungslemma — **182**
A.4 Der Projektionssatz im Hilbertraum — **183**
A.5 Zwei nützliche Integralformeln — **187**

**Literatur — 189**

**Stichwortverzeichnis — 193**

# 1 Fair Play

*You have not played as yet? Do not do so; above all avoid a martingale, if you do. Play ought not to be an affair of calculation, but of inspiration. I have calculated infallibly, and what has been the effect?*

William Makepeace Thackeray
The Newcomes, Chapter XXVIII

Die Theorie der Martingale gehört zu den bedeutendsten Entwicklungen der Wahrscheinlichkeitstheorie im 20. Jahrhundert. Ursprünglich wurden Martingale 1937 von Paul Lévy [30] für das Studium von nicht-unabhängigen Zufallsvariablen eingeführt, der Begriff *Martingal* geht auf Jean Ville [45] zurück, der auch den Zusammenhang von Martingalen mit der (mathematischen) Theorie von Spielen herstellte. Die zentrale Bedeutung von Martingalen für die Wahrscheinlichkeitstheorie hat Joseph Doob erkannt; viele der wichtigsten Sätze und Definitionen der Martingaltheorie gehen auf ihn zurück und finden sich bereits in seinem Buch [19] aus dem Jahr 1953.

Oft werden Martingale als Modelle für „faire" Spiele verwendet. Dabei stellen wir uns einen Spieler vor, der auf die Ausgänge von nacheinander ausgeführten Zufallsexperimenten („Spiele") wettet; die Auszahlungen können wir durch reelle Zufallsvariable (ZV) $\xi_1, \xi_2, \xi_3, \dots$ darstellen. Wir interessieren uns dafür, wie hoch der Einsatz $e_1, e_2, e_3, \dots$ jeweils sein muss, damit das Spiel „fair" ist.

Wir nehmen zunächst an, dass die Auszahlungen (und damit die Spiele) $\xi_i$ unabhängig und identisch verteilt (iid) sind, wobei die Einsätze in jedem Spiel gleich bleiben $e = e_i$. Wenn die ZV $\xi_i$ einen endlichen Erwartungswert $\mu = \mathbb{E}\xi_i$ haben, können wir das starke Gesetz der großen Zahlen [WT, Satz 12.4] anwenden, und erhalten für den Reingewinn des Spielers nach $n$ Spielen

$$X_n - ne = \underbrace{(\xi_1 + \cdots + \xi_n - n\mu)}_{\approx 0,\ \text{für } n \to \infty} + n(\mu - e).$$

Das zeigt, dass für $\mu > e$ das Spiel für den Spieler vorteilhaft, und für $\mu < e$ nachteilig ist. Traditionell nennt man das Spiel *fair*, wenn $\mu = e$ gilt.

---

Derartige „faire" Spiele können einen Spieler so benachteiligen, dass er nahezu mit Wahrscheinlichkeit 1 bis zum $n$ten Spiel einen Nettoverlust von $en/\log n$ erleidet, s.u. Typischerweise ist das der Fall, wenn wir die Fluktuation von $X_n - ne$ um den Nullpunkt nicht kontrollieren können, z.B. wenn der Zentrale Grenzwertsatz [WT, Kapitel 13] nicht greift, weil die Varianz $\sigma^2 = \mathbb{V}\xi_i$ unendlich ist.

---

Trotzdem ist dieser Begriff von *fairness* recht intuitiv, wie folgendes Beispiel zeigt: Beim Münzwurf mit einer fairen Münze

$$p = \mathbb{P}(\eta = 1) = \frac{1}{2} \quad \text{und} \quad q = 1 - p = \mathbb{P}(\eta = 0) = \frac{1}{2}$$

wettet Spieler $A$ auf das Erscheinen von „1" und $B$ auf „0"; die Einsätze sind $e$ bzw. $e'$, und der Gewinner erhält als Auszahlung $e + e'$. Weil die Fälle $\eta = 1$ und $\eta = 0$

https://doi.org/10.1515/9783110350685-001

gleich wahrscheinlich sind, sollte bei einem fairen Spiel $e = e'$ sein. Wir können auch die Wette auf eine unfaire Münze mit $\mathbb{P}(\eta = 1) = p \neq q = \mathbb{P}(\eta = 0)$ durch gewichtete Wetteinsätze zu einem fairen Spiel machen:

$$\text{Einsatz auf „1“} : \text{Einsatz auf „0“} = p : q = \mathbb{P}(\eta = 1) : \mathbb{P}(\eta = 0).$$

Das sieht man so: wenn z.B. $p = 1/4$ und $q = 3/4$, dann hat die Wette auf „0“ dreimal höhere Gewinnchancen. Wenn wir die unfaire Münze durch einen Tetraeder mit drei „0“ und einer „1“ ersetzen und uns vorstellen, dass Spieler B für drei Spieler mit Gewinnwahrscheinlichkeit von jeweils $p = 1/4$ steht, dann muss er auch drei Einsätze leisten. Daher verhalten sich die Einsätze wie

$$\text{Einsatz auf „1“} : \text{Einsatz auf „0“} = 1 : 3 = \frac{1}{4} : \frac{3}{4}.$$

Allgemein macht man sich schnell klar, dass das Verhältnis der Einsätze stets wie $p : q$ gewählt werden muss. Insbesondere ist die Auszahlung an Spieler A

$$\xi = \left\{ \begin{array}{ll} e + e', & \text{wenn } \eta = 1 \\ 0, & \text{wenn } \eta = 0 \end{array} \right\} = (e + e')\eta \overset{e:e'=p:q}{=} \frac{e}{p}\eta,$$

und wir erhalten wiederum $e = \mathbb{E}\xi$. Für Spieler B gilt eine entsprechende Überlegung. Wiederholen wir die Wette mehrfach, dann ist der Reingewinn von Spieler A nach $n$ Runden

$$R_n = \frac{e}{p} \sum_{i=1}^{n} (\eta_i - p) = \sum_{i=1}^{n} (\xi_i - e).$$

Für ein faires Spiel sollte außerdem gelten, dass kein Spieler durch…

---

**!**

- die Kenntnis des bisherigen Spielverlaufs (z.B. durch Kartenzählen beim „Black Jack“),
- eine geschickte Strategie (z.B. durch „Spielabbruch im richtigen Moment“),

---

Vorteile haben kann. Problematisch sind sehr unterschiedliche Vermögensverhältnisse der Spieler, die zum Ausscheiden durch Bankrott führen können, oder Restriktionen durch die Spielbank, etwa Tischlimits beim Roulette, die die Strategie der Spieler einschränken. Diese Betrachtungen werden wir hier zurückstellen.

Zumindest für Spiele mit endlich vielen verschiedenen Auszahlungen können wir die *Kenntnis der Vergangenheit* mit Hilfe der bedingten Wahrscheinlichkeiten modellieren:

$$\mathbb{P}(\xi_n = x \mid \xi_1 = x_1, \ldots, \xi_{n-1} = x_{n-1})$$

gibt die Wahrscheinlichkeit für die Auszahlung $x$ an, wenn wir die Auszahlungen (und damit die Ausgänge) $x_1, \ldots, x_{n-1}$ der ersten $n - 1$ Spiele kennen. Wenn die Spiele $\xi_1, \ldots, \xi_n$ unabhängig sind, dann ist das $\mathbb{P}(\xi_n = x)$.

Eine *Strategie* lässt sich durch variable Einsätze modellieren. Wir nehmen an, dass wir beim $n$ten Spiel den Einsatz $\mu$ mit

$$e_n = e_n(R_0, \xi_1, \dots, \xi_{n-1})$$

gewichten können, wobei auch die Auszahlung proportional zu $e_n$ ist. Der Einsatz $\mu e_n$ im $n$ten Spiel hängt vom Anfangskapital $R_0$ und den bisherigen Auszahlungen $e_1\xi_1, \dots, e_{n-1}\xi_{n-1}$ ab – insbesondere vom Vermögen

$$R_{n-1} = R_0 + \sum_{i=1}^{n-1} e_i \xi_i$$

vor dem $n$ten Spiel. Wir nehmen an, dass $e_n = e_n(r_0, x_1, \dots, x_{n-1})$ eine messbare Funktion ist.

Nun seien die ZV $\xi_1, \xi_2, \dots$ identisch verteilt mit Mittelwert $\mu = \mathbb{E}\xi_i$, aber nicht notwendigerweise unabhängig. Der Reingewinn im $n$ten Spiel ist

$$\Delta R_n = e_n(V_0, \xi_1, \dots, \xi_{n-1})\,(\xi_n - \mu)$$

und wir erhalten für ein festes Anfangskapital $R_0 = r_0$

$$\begin{aligned}
&\mathbb{E}\left[\Delta R_n \mid \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = \mathbb{E}\left[e_n(R_0, \xi_1, \dots, \xi_{n-1})\,(\xi_n - \mu) \mid \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = \mathbb{E}\left[e_n(r_0, x_1, \dots, x_{n-1})\,(\xi_n - \mu) \mid \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = e_n(r_0, x_1, \dots, x_{n-1})\mathbb{E}\left[\xi_n - \mu \mid \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = e_n(r_0, x_1, \dots, x_{n-1})\left(\mathbb{E}\left[\xi_n \mid \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] - \mu\right).
\end{aligned}$$

Diese Überlegung zeigt, dass wir für ein *faires* Spiel nicht nur $\mathbb{E}\xi_i = \mu = e$, sondern

$$\mathbb{E}\left[\xi_n \mid \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] = \mu \tag{1.1}$$

fordern sollten. Außerdem sehen wir, dass die Strategie die (ggf. fehlende) Fairness (1.1) nicht beeinflussen kann. Wir können die Bedingung (1.1) auch folgendermaßen ausdrücken

$$\begin{aligned}
&\mathbb{E}\left[R_n \mid \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = \mathbb{E}\left[R_{n-1} \,\middle|\, \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = \mathbb{E}\left[r_0 + \sum_{i=1}^{n-1} e_i(r_0, \xi_1, \dots, \xi_i)(\xi_i - \mu) \,\middle|\, \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = \mathbb{E}\left[r_0 + \sum_{i=1}^{n-1} e_i(r_0, x_1, \dots, x_i)(x_i - \mu) \,\middle|\, \xi_1 = x_1, \dots, \xi_{n-1} = x_{n-1}\right] \\
&\quad = r_0 + \sum_{i=1}^{n-1} e_i(r_0, x_1, \dots, x_i)(x_i - \mu).
\end{aligned}$$

Ein Spiel ist also genau dann *fair*, wenn das erwartete Vermögen $R_n$ eines Spielers zum zukünftigen Zeitpunkt $n$ gerade seinem derzeitigen Vermögen $r_{n-1}$ entspricht. Weil für geeignete $x_i$ bzw. $r_i$ die Beziehung $\bigcap_{i=1}^{n-1}\{\xi_i = x_i\} = \bigcap_{i=1}^{n-1}\{R_i = r_i\}$ gilt, können wir (1.1) auch folgendermaßen schreiben:

$$\begin{aligned} &\mathbb{E}\left[R_n \mid R_1 = r_1, \ldots, R_{n-1} = r_{n-1}\right] \\ &\quad = r_{n-1} = \mathbb{E}\left[R_{n-1} \mid R_1 = r_1, \ldots, R_{n-1} = r_{n-1}\right]. \end{aligned} \tag{1.2}$$

Eine Familie von ZV $(R_n)_{n\in\mathbb{N}_0}$, die der Beziehung (1.2) genügt, ist ein *Martingal*. Martingale sind zentrale Objekte der modernen Stochastik, die es uns in vielen Fällen ermöglicht, Unabhängigkeitsannahmen abzuschwächen. In den folgenden Kapiteln werden wir erst den Begriff der bedingten Erwartung genauer untersuchen und darauf die Theorie der Martingale aufbauen.

## Das klassische „Thackeraysche“ Martingal

Das klassische Martingal ist eine Spielstrategie beim Roulette, bei der ein Spieler so lange auf „Rot“ oder „Schwarz“ setzt, bis er zum ersten Mal gewinnt. Bei jedem Verlust verdoppelt er seinen Einsatz und spielt weiter. Dieses Spielsystem war schon lange vor dem Erscheinen von William Makepeace Thackerays Roman *Pendennis* bekannt, aber Thackerays Roman brachte das Martingal und seine verheerenden Folgen einem breiten Leserkreis nahe.

Aufgrund der folgenden Überlegung gilt das Martingal als sichere Gewinnstrategie: Wenn ein Spieler mit einem Einsatz von 1 Euro beginnt und $n$ Mal hintereinander verliert, ist der bis dahin aufgelaufene Verlust $1 + 2 + 2^2 + \cdots + 2^{n-1} = 2^n - 1$ Euro. Im nächsten Spiel beträgt der Einsatz $2^n$ Euro und, wenn der Spieler nun gewinnt, ist die Auszahlung $2 \cdot 2^n$ Euro. Damit ergibt sich ein Reingewinn von $2 \cdot 2^n - (2^n - 1) - 2^n = 1$ Euro. Weil das Eintreten des ersten Gewinns durch eine geometrische Verteilung beschrieben wird, vgl. [WT, Beispiel 2.4.f), S. 10], ist die Wahrscheinlichkeit des Eintretens eines Gewinns gleich eins:

$$\mathbb{P}(n\text{-mal hintereinander verlieren}) = (1-p)^n \xrightarrow[n\to\infty]{} 0,$$

wenn die Gewinnwahrscheinlichkeit $p > 0$ ist.

Das Problem ist, dass die Wartezeit $n$ bis zum ersten Gewinn sehr groß werden kann, und dass der Verlust $2^n - 1$ bzw. der neue Einsatz $2^n$ das Vermögen des Spielers und das Tischlimit des Casinos übersteigen kann, vgl. hierzu den Abschnitt zu *gambler's ruin*: Beispiele 11.7–11.10. Daher kann die Martingalstrategie i.Allg. nicht durchgeführt werden – und die „sichere“ Gewinnstrategie wird zur Illusion.

## ♦Ein unvorteilhaftes „faires“ Spiel[1]

Wir betrachten ein Spiel, dessen Auszahlungsprofil durch die folgende ZV $\xi$ beschrieben wird:

$$\mathbb{P}(\xi = 2^k) = p_k = \frac{1}{2^k k(k+1)}, \quad k \in \mathbb{N}, \quad \text{und} \quad \mathbb{P}(\xi = 0) = p_0 = 1 - \sum_{k=1}^{\infty} p_k.$$

Offensichtlich gilt

$$\mathbb{E}\xi = \sum_{k=1}^{\infty} \frac{2^k}{2^k k(k+1)} = \sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \left(1 - \tfrac{1}{2}\right) + \left(\tfrac{1}{2} - \tfrac{1}{3}\right) + \left(\tfrac{1}{3} - \tfrac{1}{4}\right) + \cdots = 1,$$

und das Spiel wird bei einem Einsatz von 1 Euro fair.

Wir bezeichnen mit $\xi_1, \xi_2, \xi_3, \ldots$ iid Wiederholungen des Spiels. Nach $n$ Runden ist der Gewinn bzw. Verlust $\xi_1 + \cdots + \xi_n - n$, aber es gilt für alle $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(X_n - n < -\frac{(1-\epsilon)n}{\log_2 n}\right) = 1 \tag{1.3}$$

($\log_2 x$ bezeichnet den Logarithmus zur Basis 2). Mit anderen Worten: Der Spieler erleidet durch wiederholtes Spielen fast sicher einen Verlust.

Um (1.3) zu zeigen, verwenden wir folgende Stutzungstechnik:

$$U_k := \xi_k \mathbb{1}_{\{\xi_k \leqslant n/\log_2 n\}} \quad \text{und} \quad V_k := \xi_k - U_k = \xi_k \mathbb{1}_{\{\xi_k > n/\log_2 n\}}.$$

Zunächst haben wir $\lim_{n\to\infty} \mathbb{P}(\xi_k = U_k \ \ \forall k \leqslant n) = 1$, weil für das Gegenereignis gilt

$$\mathbb{P}(V_1 + \cdots + V_n > 0) \leqslant n\mathbb{P}(V_1 > 0) = n\mathbb{P}(\xi_1 > n/\log_2 n) = n \sum_{k: 2^k > \frac{n}{\log_2 n}} \frac{1}{2^k k(k+1)}.$$

Die auftretende Summe lässt sich folgendermaßen abschätzen

$$\sum_{2^k > \frac{n}{\log_2 n}} \frac{1}{2^k k(k+1)} \leqslant \frac{1}{\left[\log_2\left(\frac{n}{\log_2 n}\right)\right]^2} \sum_{k \geqslant \log_2 \frac{n}{\log_2 n}} \frac{1}{2^k} \leqslant \frac{1}{\left[\log_2\left(\frac{n}{\log_2 n}\right)\right]^2} \frac{(\log_2 n)/n}{1 - \frac{1}{2}},$$

woraus

$$\lim_{n\to\infty} \mathbb{P}(V_1 + \cdots + V_n > 0) = 0 \tag{1.4}$$

folgt.

Andererseits gilt wegen der Chebyshev-Markov Ungleichung für jedes $\epsilon > 0$

$$\mathbb{P}\left(|U_1 + \cdots + U_n - n\mathbb{E}U_1| > \frac{\epsilon n}{\log_2 n}\right) \leqslant \frac{(\log_2 n)^2 \mathbb{V}U_1}{\epsilon^2 n^2} \leqslant \frac{(\log_2 n)^2 \mathbb{E}(U_1^2)}{\epsilon^2 n^2}.$$

**1** Dieses Beispiel geht auf William Feller [23] und [24, Problem 15, S. 262*f.*] zurück.

Weil aber

$$\mathbb{E}(U_1^2) = \sum_{2^k \leqslant \frac{n}{\log_2 n}} \frac{2^{2k}}{2^k k(k+1)} \leqslant \sum_{2^k \leqslant \frac{n}{\log_2 n}} \frac{2^k}{k(k+1)} \leqslant 1 + \sum_{2 \leqslant k \leqslant \log_2 \frac{n}{\log_2 n}} \frac{2^k}{k(k+1)}$$

ist, können wir die Monotonie des Summanden verwenden, und erhalten

$$\mathbb{E}(U_1^2) \leqslant 1 + \frac{\frac{n}{\log_2 n}}{\left[\log_2\left(\frac{n}{\log_2 n}\right)\right]^2}.$$

Daher gilt

$$\lim_{n\to\infty} \mathbb{P}\left(|U_1 + \cdots + U_n - n\mathbb{E}U_1| > \frac{\epsilon n}{\log_2 n}\right) = 0. \tag{1.5}$$

Schließlich rechnet man schnell nach, dass für hinreichend große Werte $n \in \mathbb{N}$

$$\mathbb{E}U_1 = \sum_{k \leqslant \log_2\left(\frac{n}{\log_2 n}\right)} \frac{2^k}{2^k k(k+1)} = \sum_{k=1}^{\log_2 \frac{n}{\log_2 n}} \left(\frac{1}{k} - \frac{1}{k+1}\right) \approx 1 - \frac{1}{\log_2 n - \log_2 \log_2 n + 1}$$

und somit auch für $\epsilon > 0$

$$1 - \frac{1+\epsilon}{\log_2 n} \leqslant \mathbb{E}U_1 \leqslant 1 - \frac{1}{\log_2 n} \tag{1.6}$$

gilt. Wenn wir die Beziehungen (1.4)–(1.6) kombinieren, folgt (1.3).

## Aufgaben

1. Zeigen Sie die Gleichheit $\bigcap_{i=1}^{n-1}\{\xi_i = x_i\} = \bigcap_{i=1}^{n-1}\{R_i = r_i\}$, die in (1.2) verwendet wird.

# 2 Bedingte Erwartung

In der Einleitung (Kapitel 1) haben wir gesehen, dass bedingte Erwartungswerte zentral für die Definition von Martingalen sind. Vor allem benötigen wir bedingte Erwartungen von ZV, die nicht notwendig diskret sind. In diesem Kapitel betrachten wir ausschließlich reelle oder komplexwertige ZV auf einem W-Raum $(\Omega, \mathscr{A}, \mathbb{P})$. Für eine ZV $X \in L^1(\mathscr{A})$ gilt

$$\mathbb{E}(X \mid A) = \frac{\mathbb{E}[X \mathbb{1}_A]}{\mathbb{P}(A)}, \quad A \in \mathscr{A},\ \mathbb{P}(A) > 0, \tag{2.1}$$

d.h. die bedingte Erwartung ist eine Mittelung der ZV $X$ relativ zu einer Menge $A \in \mathscr{A}$. Die Division durch $\mathbb{P}(A)$ dient der Normierung des Ausdrucks (2.1): wenn $X \equiv 1$, dann gilt $\mathbb{E}(1 \mid A) = \mathbb{E}(1 \cdot \mathbb{1}_A)/\mathbb{P}(A) = 1$.

Wenn $A_1, \dots, A_n \in \mathscr{A}$ eine Partition von $\Omega$ ist, $\Omega = A_1 \uplus A_2 \uplus \cdots \uplus A_n$, $\mathbb{P}(A_i) > 0$, dann können wir der ZV $X$ eine einfache ZV mit endlich vielen Werten zuordnen

$$Y := \sum_{i=1}^{n} \mathbb{E}(X \mid A_i) \mathbb{1}_{A_i} = \sum_{i=1}^{n} \frac{\mathbb{E}[X \mathbb{1}_{A_i}]}{\mathbb{P}(A_i)} \mathbb{1}_{A_i}, \tag{2.2}$$

und diese einfache Funktion hat offensichtlich folgende Eigenschaften

a) $Y$ ist $\mathscr{F}$-messbar, wobei $\mathscr{F} := \sigma(A_1, \dots, A_n)$;
b) $\int_F Y \, d\mathbb{P} = \int_F X \, d\mathbb{P}$ für alle $F \in \mathscr{F}$ (vgl. auch Beispiel 2.3);
c) $Y(\omega) = \mathbb{E}(X \mid A_i)$ für alle $\omega \in A_i$ und $i = 1, \dots, n$.

Für allgemeine $\sigma$-Algebren können wir a), b) als Definition der bedingten Erwartung verwenden.

**2.1 Definition.** Es sei $X \in L^1(\mathscr{A})$ und $\mathscr{F} \subset \mathscr{A}$ eine $\sigma$-Algebra. Eine ZV $Y \in L^1(\mathscr{F})$, d.h. $Y$ ist $\mathscr{F}$-messbar und $\mathbb{E}|Y| < \infty$, so dass gilt

$$\int_F X \, d\mathbb{P} = \int_F Y \, d\mathbb{P} \quad \forall F \in \mathscr{F}, \tag{2.3}$$

heißt (*Version* der) *bedingte*(*n*) *Erwartung* von $X$ unter $\mathscr{F}$. Wir schreiben $Y = \mathbb{E}(X \mid \mathscr{F})$.

---

Im Gegensatz zur klassischen bedingten Erwartung $\mathbb{E}(X \mid A)$, die eine *reelle Zahl* ist, ist $\mathbb{E}(X \mid \mathscr{F})$ eine *Zufallsvariable*. **!**

---

**2.2 Bemerkung.** Wenn $\mathscr{F} = \sigma(\mathscr{G})$ für einen $\cap$-stabilen Erzeuger $\mathscr{G}$ gilt, der eine Folge $G_i \uparrow \Omega$ enthält (z.B. wenn $G_i = \Omega \in \mathscr{G}$ gilt), dann können wir (2.3) ersetzen durch

$$\int_G X \, d\mathbb{P} = \int_G Y \, d\mathbb{P} \quad \forall G \in \mathscr{G}. \tag{2.4}$$

https://doi.org/10.1515/9783110350685-002

Um das zu zeigen, zerlegen wir die ZV in Positiv- und Negativteil, $Y = Y^+ - Y^-$ und $X = X^+ - X^-$, und schreiben

$$\mu(F) := \int_F (X^+ + Y^-)\, d\mathbb{P}, \quad \nu(F) := \int_F (Y^+ + X^-)\, d\mathbb{P}, \quad F \in \mathscr{F}.$$

Offensichtlich sind $\mu, \nu$ Maße auf $\mathscr{F}$, und (2.4) zeigt $\mu|\mathscr{G} = \nu|\mathscr{G}$. Weil $\mathscr{G}$ den Voraussetzungen des Eindeutigkeitssatzes für Maße [MI, Satz 4.5] genügt, folgt $\mu|\sigma(\mathscr{G}) = \nu|\sigma(\mathscr{G})$ und somit (2.3).

**2.3 Beispiel.** Es seien $B_1, \dots, B_N \in \mathscr{A}$, $N \in \mathbb{N} \cup \{\infty\}$, disjunkte Mengen, die $\Omega$ partitionieren, d.h. $\Omega = \dot{\bigcup}_{n=1}^N B_n$. Wir definieren $\mathscr{F} = \sigma(B_n, 1 \leqslant n \leqslant N)$. Wenn $\mathbb{P}(B_n) > 0$ für alle $1 \leqslant n \leqslant N$ gilt, dann ist für jede ZV $X \in L^1(\mathscr{A})$

$$\mathbb{E}(X \mid \mathscr{F}) = \sum_{n=1}^N \mathbb{E}(X \mid B_n) \mathbb{1}_{B_n}, \quad \mathbb{E}(X \mid B_n) := \frac{\mathbb{E}[X \mathbb{1}_{B_n}]}{\mathbb{P}(B_n)}.$$

Das sieht man so: Zunächst ist klar, dass die rechte Seite eine $\mathscr{F}$-messbare ZV ist. Weil $\mathscr{G} = \{B_n : 1 \leqslant n \leqslant N\} \cup \{\emptyset, \Omega\}$ ein $\cap$-stabiler Erzeuger von $\mathscr{F}$ ist, folgt die Behauptung aus (2.4) wegen

$$\begin{aligned}
\int_{B_i} \mathbb{E}(X \mid \mathscr{F})\, d\mathbb{P} &= \int_{B_i} \sum_{n=1}^N \mathbb{E}(X \mid B_n) \mathbb{1}_{B_n}\, d\mathbb{P} \\
&= \sum_{n=1}^N \mathbb{E}(X \mid B_n) \int \mathbb{1}_{B_n \cap B_i}\, d\mathbb{P} \\
&= \mathbb{E}(X \mid B_i)\mathbb{P}(B_i) = \mathbb{E}(X \mathbb{1}_{B_i}), \quad 1 \leqslant i \leqslant N.
\end{aligned}$$

Entsprechend zeigt man $\int_\Omega \mathbb{E}(X \mid \mathscr{F})\, d\mathbb{P} = \mathbb{E}X$.

Wenn $\mathbb{E}(X \mid B) := 0$ für $B \in \mathscr{A}$ mit $\mathbb{P}(B) = 0$ gesetzt wird, gilt die eben angestellte Rechnung sogar für *beliebige* abzählbare Partitionierungen von $\Omega$.

Wir wollen nun die Existenz und Eindeutigkeit der bedingten Erwartung zeigen.

**2.4 Lemma.** *Die bedingte Erwartung $\mathbb{E}(X \mid \mathscr{F})$ einer ZV $X$ ist bis auf Nullmengen eindeutig.*

*Beweis.* Es seien $Y, Y'$ zwei Versionen von $\mathbb{E}(X \mid \mathscr{F})$. Dann ist $\{Y > Y'\} \in \mathscr{F}$ und

$$\int_{\{Y > Y'\}} \underbrace{(Y' - Y)}_{<0}\, d\mathbb{P} \overset{(2.3)}{=} 0 \implies \mathbb{P}(Y > Y') = 0 \implies Y \leqslant Y' \text{ f.s.}$$

Analog zeigt man $Y \geqslant Y'$ f.s., und es folgt $Y = Y'$ f.s. □

**2.5 Satz.** *Es sei $\mathscr{F} \subset \mathscr{A}$ eine $\sigma$-Algebra und $X \in L^1(\mathscr{A})$. Dann existiert die bedingte Erwartung $X \mapsto \mathbb{E}(X \mid \mathscr{F})$ und ist ein stetiger linearer Operator von $L^1(\mathscr{A})$ nach $L^1(\mathscr{F})$:*

$$\mathbb{E}\,|\mathbb{E}(X \mid \mathscr{F}) - \mathbb{E}(Z \mid \mathscr{F})| \leqslant \mathbb{E}|X - Z|, \quad X, Z \in L^1(\mathscr{A}). \tag{2.5}$$

*Für $X \in L^2(\mathscr{A})$ ist $X \mapsto \mathbb{E}(X \mid \mathscr{F})$ die orthogonale Projektion $L^2(\mathscr{A}) \to L^2(\mathscr{F})$.*

*Beweis.* 1° Wir nehmen zunächst an, dass $X \in L^2(\mathscr{A})$. Nach Satz A.14 existiert die orthogonale Projektion $Y \in L^2(\mathscr{F})$ von $X$ auf den Raum $L^2(\mathscr{F})$. Nach Definition der Projektion gilt

$$\begin{aligned} X - Y \perp L^2(\mathscr{F}) &\overset{\text{Def.}}{\iff} \forall \Phi \in L^2(\mathscr{F}) \;:\; \mathbb{E}[(X-Y)\Phi] = 0 \\ &\iff \forall F \in \mathscr{F} \;:\; \mathbb{E}[(X-Y)\mathbb{1}_F] = 0. \end{aligned} \tag{2.6}$$

(für die Richtung „$\Leftarrow$“ approximieren wir $\Phi$ durch einfache ZV $\Phi_n$, die $|\Phi_n| \leqslant |\Phi|$ erfüllen, vgl. [MI, (Sombrero-)Lemma 7.12]). Die letzte Äquivalenz von (2.6) entspricht (2.3), d.h. $Y$ ist eine Version der bedingten Erwartung $\mathbb{E}(X \mid \mathscr{F})$.

2° Die Linearität von $X \mapsto \mathbb{E}(X \mid \mathscr{F})$, $X \in L^2(\mathscr{A})$, folgt entweder aus der Linearität der orthogonalen Projektion, oder durch direkte Rechnung: Für $X, Z \in L^2(\mathscr{A})$, $a, b \in \mathbb{R}$ und beliebige $F \in \mathscr{F}$ haben wir

$$\begin{aligned} \int_F [aX + bZ]\, d\mathbb{P} = a \int_F X\, d\mathbb{P} + b \int_F Z\, d\mathbb{P} &\overset{(2.3)}{=} a \int_F \mathbb{E}(X \mid \mathscr{F})\, d\mathbb{P} + b \int_F \mathbb{E}(Z \mid \mathscr{F})\, d\mathbb{P} \\ &= \int_F \big[a\mathbb{E}(X \mid \mathscr{F}) + b\mathbb{E}(Z \mid \mathscr{F})\big]\, d\mathbb{P}. \end{aligned}$$

Daher ist die ZV $a\mathbb{E}(X \mid \mathscr{F}) + b\mathbb{E}(Z \mid \mathscr{F})$ eine Version der bedingten Erwartung von $aX + bZ$. Wegen der Eindeutigkeit gilt $\mathbb{E}(aX + bZ \mid \mathscr{F}) = a\mathbb{E}(X \mid \mathscr{F}) + b\mathbb{E}(Z \mid \mathscr{F})$.

3° Für $X \in L^2(\mathscr{A})$ definieren wir $F^\pm := \{\pm\mathbb{E}(X \mid \mathscr{F}) > 0\} \in \mathscr{F}$, und beachten

$$\begin{aligned} \mathbb{E}\,|\mathbb{E}(X \mid \mathscr{F})| &= \int_{F^+} \mathbb{E}(X \mid \mathscr{F})\, d\mathbb{P} - \int_{F^-} \mathbb{E}(X \mid \mathscr{F})\, d\mathbb{P} \\ &\overset{(2.3)}{=} \int_{F^+} X\, d\mathbb{P} - \int_{F^-} X\, d\mathbb{P} \leqslant \int_{F^+} |X|\, d\mathbb{P} + \int_{F^-} |X|\, d\mathbb{P} \leqslant \mathbb{E}|X|. \end{aligned}$$

Indem wir $X$ durch $X - Z$ ersetzen und $\mathbb{E}(X - Z \mid \mathscr{F}) = \mathbb{E}(X \mid \mathscr{F}) - \mathbb{E}(Z \mid \mathscr{F})$ beachten, folgt (2.5) für $X, Z \in L^2(\mathscr{A})$.

4° Wir approximieren schließlich $X \in L^1(\mathscr{A})$ mit $X_n := (-n) \vee X \wedge n \in L^2(\mathscr{A})$ fast sicher und in $L^1$. Mit Hilfe von Schritt 3° erhalten wir

$$\mathbb{E}\,|\mathbb{E}(X_m \mid \mathscr{F}) - \mathbb{E}(X_n \mid \mathscr{F})| \leqslant \mathbb{E}|X_m - X_n| \xrightarrow[m,n\to\infty]{\text{dom. Konvergenz}} 0,$$

d.h. $(\mathbb{E}(X_n \mid \mathscr{F}))_{n\in\mathbb{N}}$ ist eine $L^1(\mathscr{F})$-Cauchyfolge.

Daher existiert der $L^1$-Limes $Y = \lim_{n\to\infty} \mathbb{E}(X_n \mid \mathscr{F})$. Weil $\mathbb{1}_F X = \lim_{n\to\infty} \mathbb{1}_F X_n$ und $\mathbb{1}_F Y = \lim_{n\to\infty} \mathbb{1}_F \mathbb{E}(X_n \mid \mathscr{F})$ in $L^1$ für alle $F \in \mathscr{F}$ gilt, folgt

$$\forall F \in \mathscr{F} \;:\; \int_F Y\, d\mathbb{P} = \lim_{n\to\infty} \int_F \mathbb{E}(X_n \mid \mathscr{F})\, d\mathbb{P} \overset{(2.3)}{=} \lim_{n\to\infty} \int_F X_n\, d\mathbb{P} = \int_F X\, d\mathbb{P},$$

also ist $Y$ eine Version von $\mathbb{E}(X \mid \mathscr{F})$ für $X \in L^1(\mathscr{A})$. Die Linearität und Stetigkeit (2.5) zeigt man nun genauso wie in 2° und 3°. □

**2.6 Satz** (Eigenschaften der bedingten Erwartung). *Es sei $(\Omega, \mathscr{A}, \mathbb{P})$ ein W-Raum und $\mathscr{F} \subset \mathscr{A}$ eine σ-Algebra. Dann gilt für $X, Z \in L^1(\mathscr{A})$ und $a, b \in \mathbb{R}$:*

a) $\mathbb{E}(X \mid \mathscr{F}) \in L^1(\mathscr{F})$ *und* $\|\mathbb{E}(X \mid \mathscr{F})\|_{L^1} \leqslant \|X\|_{L^1}$. (Kontraktion)

b) $\mathbb{E}(1 \mid \mathscr{F}) = 1$. (Konservativität)

c) $\mathbb{E}(aX + bZ \mid \mathscr{F}) = a\,\mathbb{E}(X \mid \mathscr{F}) + b\,\mathbb{E}(Z \mid \mathscr{F})$. (Linearität)

d) *Wenn $X \in L^2(\mathscr{A})$, dann ist $\mathbb{E}(X \mid \mathscr{F})$ das eindeutig bestimmte Element aus $L^2(\mathscr{F})$ mit* (Minimierer)

$$\inf_{\Phi \in L^2(\mathscr{F})} \|X - \Phi\|_{L^2} = \|X - \mathbb{E}(X \mid \mathscr{F})\|_{L^2}.$$

e) *Ist $\mathscr{H} \subset \mathscr{F}$ eine weitere σ-Algebra, dann gilt* (tower property)

$$\mathbb{E}\left[\mathbb{E}(X \mid \mathscr{F}) \mid \mathscr{H}\right] = \mathbb{E}(X \mid \mathscr{H}).$$

f) $\mathbb{E}(\Phi \cdot X \mid \mathscr{F}) = \Phi \cdot \mathbb{E}(X \mid \mathscr{F}) \quad \forall \Phi$ *$\mathscr{F}$-messbar und* $\mathbb{E}|\Phi \cdot X| < \infty$. (pull out)

g) $\mathbb{E}(\Phi \mid \mathscr{F}) = \Phi \quad \forall \Phi \in L^1(\mathscr{F})$. (pull out)

h) $0 \leqslant X \leqslant 1 \implies 0 \leqslant \mathbb{E}(X \mid \mathscr{F}) \leqslant 1$. (Markov-Eigenschaft)

j) $X \leqslant Z \implies \mathbb{E}(X \mid \mathscr{F}) \leqslant \mathbb{E}(Z \mid \mathscr{F})$. (Monotonie)

k) $|\mathbb{E}(X \mid \mathscr{F})| \leqslant \mathbb{E}(|X| \mid \mathscr{F})$. (Dreiecksungleichung)

l) $\mathbb{E}(X \mid \{\emptyset, \Omega\}) = \mathbb{E}X$.

m) $\mathbb{E}\left[\mathbb{E}(X \mid \mathscr{F})\right] = \mathbb{E}X$. (tower property)

---

**i** Die Eigenschaft 2.6.d hat eine sehr anschauliche Interpretation: $\mathbb{E}(X \mid \mathscr{F})$ ist der *least square predictor* für die $\mathscr{A}$-messbare ZV $X$, wenn nur die „Information" aus $\mathscr{F} \subset \mathscr{A}$ bekannt ist.

---

*Beweis von Satz 2.6.* a) und c) folgen aus Satz 2.5.

b) Weil die konstante ZV $X \equiv 1$ messbar bezüglich $\mathscr{F}$ ist, folgt die Behauptung direkt aus Definition 2.1 und der Eindeutigkeit (Lemma 2.4) der bedingten Erwartung.

d) Für $X \in L^2(\mathscr{A})$, $Y := \mathbb{E}(X \mid \mathscr{F}) \in L^2(\mathscr{F})$ und beliebiges $\Psi \in L^2(\mathscr{F})$ gilt

$$\begin{aligned}
\mathbb{E}\left(|X - \Psi|^2\right) &= \mathbb{E}\left(|(X - Y) + (Y - \Psi)|^2\right) \\
&= \mathbb{E}\left(|X - Y|^2\right) + \mathbb{E}\left(|\Psi - Y|^2\right) + \underbrace{2\mathbb{E}\big((X - Y)\underbrace{(Y - \Psi)}_{=\Phi}\big)}_{=0 \text{ wegen } (2.6)} \\
&\geqslant \mathbb{E}\left(|X - Y|^2\right).
\end{aligned}$$

Das zeigt, dass $Y$ der Minimierer des Funktionals $L^2(\mathscr{F}) \ni \Phi \mapsto \|X - \Phi\|_{L^2}$ ist.

e) Für $H \in \mathscr{H} \subset \mathscr{F}$ gilt

$$\int_H \underbrace{\mathbb{E}(X \mid \mathscr{F})}_{\text{originale ZV}}\, d\mathbb{P} \overset{(2.3)}{=} \int_H X\, d\mathbb{P} \overset{(2.3)}{=} \int_H \mathbb{E}(X \mid \mathscr{H})\, d\mathbb{P}.$$

Mithin ist die $\mathscr{H}$-messbare ZV $\mathbb{E}(X \mid \mathscr{H})$ eine Version der bedingten Erwartung (bezüglich der $\sigma$-Algebra $\mathscr{H}$) der ZV $\mathbb{E}(X \mid \mathscr{F})$.

f) & g) Zunächst nehmen wir an, dass $\Phi = \mathbb{1}_G$ für ein $G \in \mathscr{F}$ ist. Dann ist für alle $F \in \mathscr{F}$

$$\int_F \mathbb{1}_G \cdot X \, d\mathbb{P} = \int_{F \cap G} X \, d\mathbb{P} = \int_{F \cap G} \mathbb{E}(X \mid \mathscr{F}) \, d\mathbb{P} = \int_F \mathbb{1}_G \cdot \mathbb{E}(X \mid \mathscr{F}) \, d\mathbb{P}.$$

Wegen der Linearität des (bedingten) Erwartungswerts folgt daraus für $\mathscr{F}$-messbare einfache Funktionen $\Phi_n$

$$\int_F \Phi_n \cdot X \, d\mathbb{P} = \int_F \Phi_n \cdot \mathbb{E}(X \mid \mathscr{F}) \, d\mathbb{P} \quad \forall F \in \mathscr{F} \implies \mathbb{E}(\Phi_n \cdot X \mid \mathscr{F}) = \Phi_n \cdot \mathbb{E}(X \mid \mathscr{F}).$$

Mit Hilfe des Sombrero-Lemmas [MI, Korollar 7.12] können wir jede $\mathscr{F}$-messbare ZV $\Phi$ durch eine Folge von einfachen, $\mathscr{F}$-messbaren Funktionen $\Phi_n$ approximieren, die zudem die Ungleichung $|\Phi_n| \leqslant |\Phi|$ erfüllen. Wegen $\mathbb{E}|\Phi \cdot X| < \infty$ folgt

$$\Phi_n \cdot X \xrightarrow[n \to \infty]{} \Phi \cdot X \quad \text{in } L^1 \text{ und fast sicher.}$$

Weil die bedingte Erwartung stetig in $L^1$ ist, vgl. Satz 2.5, erhalten wir dann

$$\mathbb{E}(\Phi \cdot X \mid \mathscr{F}) = \lim_{k \to \infty} \mathbb{E}(\Phi_{n(k)} \cdot X \mid \mathscr{F}) = \lim_{k \to \infty} \Phi_{n(k)} \cdot \mathbb{E}(X \mid \mathscr{F}) = \Phi \cdot \mathbb{E}(X \mid \mathscr{F}).$$

h) & j) Es sei $X \geqslant 0$. Wir betrachten die Menge $\{\mathbb{E}(X \mid \mathscr{F}) < 0\} \in \mathscr{F}$. Nach Definition der bedingten Erwartung gilt

$$0 \leqslant \int_{\{\mathbb{E}(X|\mathscr{F})<0\}} X \, d\mathbb{P} \overset{(2.3)}{=} \int_{\{\mathbb{E}(X|\mathscr{F})<0\}} \mathbb{E}(X \mid \mathscr{F}) \, d\mathbb{P} \leqslant 0,$$

und wir erhalten $\mathbb{P}\left(\mathbb{E}(X \mid \mathscr{F}) < 0\right) = 0$, also $\mathbb{E}(X \mid \mathscr{F}) \geqslant 0$ f.s.
Insbesondere gilt für $X \leqslant Z$, dass $Z - X \geqslant 0$, und wir sehen

$$\mathbb{E}(Z - X \mid \mathscr{F}) \geqslant 0 \implies \mathbb{E}(Z \mid \mathscr{F}) \geqslant \mathbb{E}(X \mid \mathscr{F}).$$

Schließlich folgt mit $Z \equiv 1$ wegen $\mathbb{E}(1 \mid \mathscr{F}) = 1$ die zweite Ungleichung von h).

k) Wegen $\pm X \leqslant |X|$ folgt $\pm\mathbb{E}(X \mid \mathscr{F}) \leqslant \mathbb{E}(|X| \mid \mathscr{F})$, mithin $|\mathbb{E}(X \mid \mathscr{F})| \leqslant \mathbb{E}(|X| \mid \mathscr{F})$.

l) Weil die konstante ZV $\mathbb{E}X$ messbar bezüglich der $\sigma$-Algebra $\{\emptyset, \Omega\}$ ist, folgt die Behauptung aus

$$\int_F X \, d\mathbb{P} = \left\{ \begin{matrix} 0, & F = \emptyset \\ \mathbb{E}X, & F = \Omega \end{matrix} \right\} = \int_F \mathbb{E}X \, d\mathbb{P}.$$

m) Es gilt $\mathbb{E}\big(\mathbb{E}(X \mid \mathscr{F})\big) \overset{l)}{=} \mathbb{E}[\mathbb{E}(X \mid \mathscr{F}) \mid \{\emptyset, \Omega\}] \overset{e)}{=} \mathbb{E}(X \mid \{\emptyset, \Omega\}) \overset{l)}{=} \mathbb{E}X.$ □