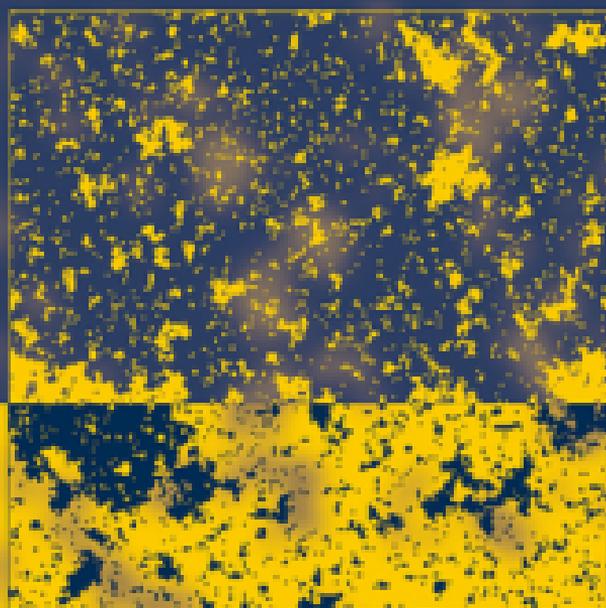


Thomas Müller-Gronbach
Erich Novak · Klaus Ritter

Monte Carlo-Algorithmen



 Springer

Springer-Lehrbuch

Thomas Müller-Gronbach · Erich Novak ·
Klaus Ritter

Monte Carlo-Algorithmen

 Springer

Prof. Dr. Thomas Müller-Gronbach
Passau, Germany

Prof. Dr. Klaus Ritter
Kaiserslautern, Germany

Prof. Dr. Erich Novak
Jena, Germany

ISSN 0937-7433

ISBN 978-3-540-89140-6

e-ISBN 978-3-540-89141-3

DOI 10.1007/978-3-540-89141-3

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.

© Springer-Verlag Berlin Heidelberg 2012

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Einbandentwurf: WMXDesign GmbH, Heidelberg

Gedruckt auf säurefreiem Papier

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

Vorwort

Monte Carlo-Methoden sind Algorithmen, die Zufallszahlen benutzen. Sie werden für unterschiedlichste Fragestellungen der diskreten und der kontinuierlichen Mathematik eingesetzt und finden dementsprechend vielfältige Anwendungen. Oft liegt ein komplexes stochastisches Modell vor, und man möchte dieses simulieren, um typische Zustände oder zeitliche Verläufe zu erhalten, oder man ist an der Berechnung deterministischer Kenngrößen wie Erwartungswerten interessiert, die sich nicht oder nur sehr schwer analytisch bestimmen lassen. Man kann aber auch grundsätzlich untersuchen, ob und wie sich für eine algorithmische Fragestellung Zufallszahlengeneratoren als Ressource nutzen lassen.

Dieses Lehrbuch gibt eine Einführung in die Mathematik und die Einsatzmöglichkeiten der Monte Carlo-Methoden und verwendet dazu durchgängig die Sprache der Stochastik. Es soll den Leser in den Stand versetzen, dieses wichtige algorithmische Werkzeug kompetent anwenden und die Ergebnisse interpretieren zu können. Die vorgenommene Themenauswahl zielt dabei nicht auf ein breites Überblickswissen ab, sondern stellt die Berechnung von Erwartungswerten und Integralen in den Vordergrund. Über die Einführung hinaus wird dem Leser anhand ausgewählter Fragestellungen ein erster Eindruck von aktuellen Forschungsthemen in diesem Bereich vermittelt, um ihn so für eine vertiefte Beschäftigung mit diesem Teilgebiet der Angewandten Mathematik zu gewinnen.

Dem Text liegen Vorlesungen zugrunde, die wir in den letzten Jahren mehrfach an den Universitäten Erlangen, Augsburg, Jena, Darmstadt, Magdeburg, Passau und Kaiserslautern für Studierende der Mathematik im Haupt- und Nebenfach und im Lehramt gehalten haben. Der Leser sollte Grundkenntnisse der Analysis, Linearen Algebra und Stochastik sowie, in etwas geringerem Umfang, der Numerischen Mathematik besitzen.

In den Kapiteln 1 bis 5 behandeln wir die direkte Simulation, d. h. den klassischen Monte Carlo-Algorithmus, sowie die grundlegenden Methoden zur Simulation von Verteilungen und zur Varianzreduktion. Begleitend präsentieren wir Anwendungsbeispiele aus der Teilchenphysik und der Finanz- und Versicherungsmathematik, die sich auf elementare Weise modellieren lassen, und wir zeigen, wie Monte Carlo-Methoden zur Lösung elliptischer Randwertprobleme eingesetzt werden.

Markov Chain Monte Carlo-Algorithmen sind das Thema von Kapitel 6, das mit einem Abschnitt über die Grenzen der direkten Simulation beginnt. Nach einer Einführung in die Theorie endlicher Markov-Ketten behandeln wir Grundprinzipien der Markov Chain Monte Carlo-Methode, klassische Verfahren wie den Metropolis-Algorithmus und den Gibbs-Sampler und illustrieren den Einsatz dieser Verfahren anhand des hard core model und des Ising-Modells aus der Statistischen Physik.

Gegenstand von Kapitel 7 ist die numerische Integration. Hier studieren wir insbesondere den Einfluß von Glattheit und Dimension, und wir zeigen, wie sich die Frage nach der Optimalität von deterministischen Verfahren oder von Monte Carlo-Algorithmen formulieren und beantworten läßt. Mit der Integration bezüglich des Wiener-Maßes lernen wir ein unendlich-dimensionales Problem kennen, bei dessen Lösung stochastische Multilevel-Algorithmen zum Einsatz kommen.

Die Kapitel 6 und 7 geben außerdem Ausblicke auf aktuelle Forschungsfragen und -ergebnisse, deren detaillierte Behandlung mehr als die genannten Vorkenntnisse erfordern würde. Zu den Themen gehören schnelle Mischung sowie Fehler-schranken und Aufwärmzeiten für Markov Chain Monte Carlo-Verfahren, der Fluch der Dimension bei hochdimensionalen Integrationsproblemen und die Berechnung von Integralen im Kontext stochastischer Differentialgleichungen.

Die praktische Erzeugung von Zufallszahlen und die spannende Frage nach einer mathematischen Definition der Zufälligkeit einer Zahlenfolge sind nicht Gegenstand des Buches. Die Behandlung dieser beiden Themen erfordert zum Teil andere als die für diesen Text benötigten Vorkenntnisse, und sie führt letztlich in andere Bereiche der Mathematik. Generatoren für gleichverteilte Zufallszahlen werden in diesem Lehrbuch folglich als black box betrachtet.

Zahlreiche Freunde, Kollegen, Mitarbeiter und Studenten haben das Manuskript durchgesehen, uns wertvolle Hinweise gegeben und Verbesserungsvorschläge gemacht. Für Hinweise und Anregungen danken wir Ingo Althöfer, Ehrhard Behrends, Jakob Creutzig, Norbert Gaffke, Michael Gnewuch, Siegfried Graf, Peter Hellekalek, Albrecht Irlé, Achim Klenke, Peter Kloeden, Peter Mathé und Harald Niederreiter. Wir danken Florian Blöthner, Nico Döhring, Thomas Ehlenz, Andreas Fromkorth, Ivo Hedtke, Mario Hefter, Felix Heidenreich, Daniel Henkel, Florian Hübsch, Eva Kapfer, Dominique Küpper, Harald Pfeiffer, Christian Richter, Daniel Rudolf, Mehdi Slassi, Regina Spichtinger, Mario Ullrich, Markus Weimar und Larisa Yaroslavtseva als kritischen Lesern von Teilen des Manuskriptes.

Die Rechnungen und graphischen Darstellungen bei der Simulation zur Neutronenbewegung in Kapitel 4 stammen von Martin Amft. Alle weiteren Simulationen in den Kapiteln 2, 3, 4 und 5 wurden von Robert Offinger durchgeführt und illustriert. Die in Kapitel 6 und auf dem Buchumschlag gezeigten Rechnungen und Visualisierungen stammen von Stephan Brummer und Mario Ullrich. Von Thorsten Sickenger, Andreas Weyhausen und Heidi Weyhausen stammen frühere Beiträge. Wir danken allen für ihre Unterstützung.

Passau, Jena, Kaiserslautern,
November 2011

*Thomas Müller-Gronbach
Erich Novak
Klaus Ritter*

Inhaltsverzeichnis

1	Vorbereitungen	1
1.1	Zufallszahlengeneratoren	2
1.2	Wahrscheinlichkeitstheoretische Grundbegriffe	4
1.3	Zufallszahlen und gleichverteilte Zufallsvariablen	8
	Aufgaben	10
2	Algorithmen, Fehler und Kosten	13
2.1	Was ist ein Algorithmus?	14
2.2	Fehler und Kosten einer Monte Carlo-Methode	16
2.3	Ein Zählproblem	20
	Aufgaben	24
3	Das Verfahren der direkten Simulation	27
3.1	Direkte Simulation	28
3.2	Erste Beispiele	30
3.2.1	Die Berechnung von π	31
3.2.2	Numerische Integration	32
3.2.3	Eigenschaften von zufälligen Polyedern	35
3.2.4	Chancen bei Patience-Spielen	37
3.2.5	Die Bewertung von Finanzderivaten	38
3.2.6	Ein Ruinproblem	42
3.3	Grenzwertsätze der Stochastik	42
3.3.1	Das starke Gesetz der großen Zahlen	43
3.3.2	Der zentrale Grenzwertsatz	48
3.4	Die Hoeffding-Ungleichung	53
3.5	Varianzschätzung und Konfidenzintervalle	56
3.6	Stoppzeiten und Unabhängigkeit	63
	Aufgaben	70

4	Simulation von Verteilungen	75
4.1	Die Inversionsmethode	76
4.2	Durchgang von Neutronen durch Materie	79
4.3	Die Verwerfungsmethode	87
4.4	Simulation von Normalverteilungen	92
4.5	Simulation stochastischer Prozesse	96
4.5.1	Zeit-diskrete Markov-Prozesse	97
4.5.2	Die Brownsche Bewegung und der Poisson-Prozeß	101
4.5.3	Bewertung von Optionen im Black-Scholes-Modell	110
4.5.4	Ruinwahrscheinlichkeiten im Cramér-Lundberg-Modell	114
4.5.5	Der sphärische Prozeß und das Dirichlet-Problem	116
	Aufgaben	127
5	Varianzreduktion	133
5.1	Antithetic Sampling	134
5.2	Control Variates	139
5.3	Stratified Sampling	146
5.4	Importance Sampling	155
5.5	Varianzreduktion zur Verbesserung der Konvergenzordnung	166
	Aufgaben	173
6	Die Markov Chain Monte Carlo-Methode	179
6.1	Die Grenzen der direkten Simulation	180
6.2	Endliche Markov-Ketten	188
6.2.1	Grundbegriffe	188
6.2.2	Grenzwertsätze	202
6.3	Markov Chain Monte Carlo-Algorithmen	207
6.4	Ausblick	216
6.4.1	Schnelle Mischung	216
6.4.2	Fehlerschranken und Aufwärmzeit	219
6.4.3	Schnelle Mischung beim Ising-Modell	222
6.4.4	Perfekte Simulation	226
6.4.5	Markov-Ketten auf konvexen Körpern	236
	Aufgaben	239
7	Numerische Integration	245
7.1	Deterministische Algorithmen	247
7.1.1	Grundbegriffe	247
7.1.2	Untere Schranken für minimale Fehler	254
7.1.3	Multivariate polynomiale Interpolation	260
7.1.4	Asymptotisch optimale Algorithmen für F_d^r	267
7.1.5	Optimalität von Quadraturformeln	270
7.2	Randomisierte Algorithmen	273
7.2.1	Grundbegriffe	274
7.2.2	Untere Schranken für minimale Fehler	278

7.2.3	Asymptotisch optimale Algorithmen für F_d^r	285
7.3	Unendlich-dimensionale Integration	287
7.4	Ausblick	297
7.4.1	Funktionen mit beschränkten gemischten Ableitungen	298
7.4.2	Tractability und der Fluch der Dimension	299
7.4.3	Zur Überlegenheit randomisierter Algorithmen	302
7.4.4	Average Case-Analyse	303
7.4.5	Stochastische Multilevel-Verfahren	304
	Aufgaben	307
	Literaturverzeichnis	311
	Symbolverzeichnis	319
	Sachverzeichnis	321

Kapitel 1

Vorbereitungen

Monte Carlo-Methoden sind Algorithmen, die Zufallszahlen benutzen. Neben den üblichen Befehlen (die Grundrechenarten, Vergleich von Zahlen, ein Orakel zur Berechnung von Funktionswerten) ist also zusätzlich der Aufruf eines *Zufallszahlengenerators* zugelassen, d. h. Befehle der Art

- „Wähle $x \in [0, 1]$ zufällig“ oder
- „Wähle $x \in \{0, \dots, N - 1\}$ zufällig“

sind erlaubt. In der Literatur werden *Monte Carlo-Methoden* auch als *stochastische* oder *randomisierte Algorithmen* bezeichnet, und das Gebiet der Monte Carlo-Methoden heißt auch *Stochastische Simulation* oder *Experimentelle Stochastik*. Wir verwenden die Begriffe Algorithmus, Methode und Verfahren synonym.

Warum benutzt man Monte Carlo-Methoden? Oder genauer: Für welche Probleme ist es sinnvoll, Monte Carlo-Methoden zu benutzen? Wir werden sehen, daß für einige Probleme Monte Carlo-Methoden zur Verfügung stehen, die besser als alle deterministischen Verfahren sind.

Praktische Berechnungen erfordern oft eine große Anzahl von Zufallszahlen, und ein Zufallszahlengenerator sollte idealerweise die Realisierung einer unabhängigen Folge von auf $[0, 1]$ oder $\{0, \dots, N - 1\}$ gleichverteilten Zufallsvariablen liefern. Uns geht es vor allem darum zu zeigen, was man mit einem *idealen Zufallszahlengenerator* erreichen kann – letztlich werden mathematische Sätze über Zufallsvariablen behandelt. Bedenken der Art, daß es keine idealen Generatoren gibt, stellen wir zurück und verweisen auf die kurze Diskussion im folgenden Abschnitt.

Beim Einsatz von Monte Carlo-Methoden sollte sich der verwendete Generator so verhalten, wie man es bei echten Realisierungen von Zufallsvariablen erwarten würde. Da auch Generatoren verbreitet sind, deren Qualität nicht ausreichend gesichert ist, darf die Wahl eines Zufallszahlengenerators nicht unbedacht geschehen.

Wir werden in Abschnitt 1.1 drei empfehlenswerte Generatoren vorstellen, mit denen der Leser selbst Monte Carlo-Methoden programmieren und anwenden kann. In Abschnitt 1.2 stellen wir wahrscheinlichkeitstheoretische Grundbegriffe bereit und präzisieren dann in Abschnitt 1.3, was wir unter von einem idealen Zufallszahlengenerator erzeugten Zufallszahlen verstehen wollen.

1.1 Zufallszahlengeneratoren

Die Entwicklung des Computers und militärische Anwendungen setzten zur Mitte des letzten Jahrhunderts eine rapide Entwicklung randomisierter Algorithmen in Gang. Zu den Pionieren dieser Entwicklung zählt John von Neumann, der im Jahr 1949 auf einer Konferenz mit dem Namen „Monte Carlo Method“ einen Vortrag zur Problematik der Erzeugung von Zufallszahlen hielt. Wir zitieren aus seinem Beitrag für den zugehörigen Tagungsband. John von Neumann [143] war der folgenden Ansicht:

We see then that we could build a physical instrument to feed random digits directly into a high-speed computing machine and could have the control call for these numbers as needed.

Aber von Neumann fährt fort mit einem Einwand gegen diesen „echten“ Zufall, der uns auch heute noch einleuchtet:

The real objection to this procedure is the practical need for checking computations. If we suspect that a calculation is wrong, almost any reasonable check involves repeating something done before.

Wiederholbarkeit ist eine Anforderung an jede wissenschaftliche Arbeit, die sich schlecht mit Rechenergebnissen verträgt, welche tatsächlich zufällig und damit nicht reproduzierbar sind. Deshalb benutzen randomisierte Algorithmen meist sogenannte *Pseudo-Zufallszahlen*, die von deterministischen Algorithmen erzeugt werden.

Generatoren für Pseudo-Zufallszahlen in $Z = [0, 1]$ oder $Z = \{0, \dots, N - 1\}$ sind durch endliche Mengen A und B sowie durch Abbildungen $g: A \rightarrow B$, $f: B \rightarrow B$ und $h: B \rightarrow Z$ gegeben. Ein typischer Wert für die Mächtigkeit von B ist $|B| = 2^{128}$, aber es werden auch viel größere Werte wie etwa $|B| = 2^{19937}$ beim Mersenne Twister verwendet. Der Benutzer wählt zur Initialisierung des Generators einen Startwert $s \in A$, indem er ein Unterprogramm etwa in der Form `init(s)` aufruft. Damit wird durch $b_1 = g(s)$ und $b_i = f(b_{i-1})$ eine Folge in B definiert. Durch sukzessive Aufrufe des Generators, etwa von der Form `x := rand()`, werden Pseudo-Zufallszahlen $x_i = h(b_i)$ erzeugt. Will man die Reproduzierbarkeit einer Monte Carlo-Simulation gewährleisten, sollte man den verwendeten Generator und seine Initialisierung angeben.

Die Güte von Generatoren wird einerseits empirisch durch statistische Tests überprüft, siehe Gentle [63], Knuth [107] und Niederreiter [144]. Andererseits wird mit Methoden der Komplexitätstheorie analysiert, welcher Aufwand nötig ist, um die Folge $(x_i)_{i \in \mathbb{N}}$ von einer Realisierung einer unabhängigen Folge gleichverteilter Zufallsvariablen zu unterscheiden, siehe Goldreich [71] und Stinson [186].

Zum Einsatz solcher auf deterministischen Methoden basierender Generatoren sagt von Neumann in derselben Arbeit:

Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin.

Wir sind keine Spezialisten für Sünden. Dennoch stellt sich die Frage, ob man zum Beispiel die in Kapitel 3 behandelte Methode der direkten Simulation und Aussagen wie die Fehlerabschätzungen aus den Sätzen 3.2 oder 3.5 überhaupt anwenden kann, da diese Aussagen für Zufallsvariablen und ihre Realisierungen gelten. Die gängigen Computer arbeiten aber deterministisch, und die erzeugten Pseudo-Zufallszahlen sind somit nicht zufällig. In der Praxis sollte man deshalb

- nur solche Zufallszahlengeneratoren benutzen, die sich für ähnliche Probleme bereits gut bewährt haben und
- wichtige Rechnungen nach Möglichkeit mit vom Typ her verschiedenen Generatoren durchführen.

Mit diesen naheliegenden Vorsichtsmaßnahmen kann man die *prinzipielle Problematik* der Monte Carlo-Methoden, daß der Zufall mit Hilfe von deterministischen Algorithmen simuliert werden soll, nicht aus der Welt schaffen. Aber man kann praktisch mit ihr umgehen.

An dieser Stelle wollen wir Zufallszahlengeneratoren empfehlen, die bisher sowohl theoretisch eingehend analysiert wurden als auch bei zahlreichen Tests besonders gut abgeschnitten haben und außerdem schnell und portabel sind.

Der Mersenne Twister und der Generator TT800 Der Mersenne Twister, der von Matsumoto, Nishimura [131] vorgestellt wurde, gilt als besonders guter Zufallszahlengenerator. Für „kleine“ Anwendungen ist auch der Vorläufer TT800 des Mersenne Twisters, der von Matsumoto, Kurita [130] entwickelt wurde, zu empfehlen. In den Simulationsbeispielen des vorliegenden Textes wurde stets einer dieser beiden Generatoren eingesetzt. Der Mersenne Twister ist der Standard-Generator in MATLAB, R und MAPLE, und er ist auch in MATHEMATICA verwendbar. Weitere Implementationen sind in vielen höheren Programmiersprachen verfügbar. Panneton, L'Ecuyer, Matsumoto [155] führen als Weiterentwicklung die WELL-Generatoren ein.

Ein Generator für paralleles Rechnen Beim Arbeiten auf Parallelrechnern benötigt man eventuell gleichzeitig viele Zufallszahlengeneratoren, die unabhängig voneinander Zufallszahlen liefern. Bei solchen Anwendungen hat sich besonders das Software-Paket von L'Ecuyer, Simard, Chen, Kelton [112] bewährt.

AES (advanced encryption standard) Hierbei handelt es sich, im Gegensatz zu den beiden bisher genannten Generatoren, um einen nicht-linearen Zufallszahlengenerator. Er basiert auf einem Verfahren zur Kryptographie, ist ausreichend schnell und hat bei statistischen Tests sehr gut abgeschnitten, siehe Hellekalek, Wegenkittl [85].

L'Ecuyer [111] stellt in einer Übersichtsarbeit Grundprinzipien und Methoden der Erzeugung von Zufallszahlen vor. Weitere Hinweise und Links zu diesem Thema sowie zu verfügbarer Software finden sich bei Press, Teukolsky, Vetterling, Flannery [158], auf der Internetseite <http://random.mat.sbg.ac.at> von Hellekalek und der Internetseite <http://www.iro.umontreal.ca/~lecuyer> von L'Ecuyer.

1.2 Wahrscheinlichkeitstheoretische Grundbegriffe

Im folgenden geben wir eine kurze Zusammenstellung elementarer Sachverhalte aus der Wahrscheinlichkeitstheorie. Eine kompakte Darstellung dieser und weiterer Grundlagen bietet Kapitel 1 von Brémaud [27]. Für eine umfassende Einführung in die Wahrscheinlichkeitstheorie und Statistik verweisen wir auf Georgii [66], Irle [91] und Krengel [108]. Weiterführende Texte sind Gänsler, Stute [61] und Klenke [105].

Zufallsvariablen und Erwartungswerte Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum und Ω' eine weitere Menge mit einer σ -Algebra \mathfrak{A}' . Eine meßbare Abbildung $X: \Omega \rightarrow \Omega'$ heißt *Zufallsvariable mit Werten in (Ω', \mathfrak{A}')* . Die Funktionswerte $X(\omega)$ mit $\omega \in \Omega$ werden als *Realisierungen* von X bezeichnet. Durch

$$P_X(A) = P(\{\omega \in \Omega \mid X(\omega) \in A\}) = P(\{X \in A\}), \quad A \in \mathfrak{A}',$$

wird ein Wahrscheinlichkeitsmaß P_X auf (Ω', \mathfrak{A}') definiert, die sogenannte *Verteilung* von X . Von besonderem Interesse sind Zufallsvariablen mit Werten in $(\mathbb{R}^d, \mathfrak{A}')$, wobei \mathfrak{A}' die Borelsche σ -Algebra ist. Man spricht dann im Falle $d = 1$ von *reellwertigen Zufallsvariablen* und für $d \geq 1$ auch von *Zufallsvektoren*. Im folgenden wird für \mathbb{R}^d stets die Borelsche σ -Algebra verwendet, und Borel-meßbare Mengen $A \subset \mathbb{R}^d$ werden kurz als meßbare Mengen bezeichnet. Wir identifizieren Zufallsvektoren X und Y auf $(\Omega, \mathfrak{A}, P)$, die $P(\{X = Y\}) = 1$ erfüllen. Zur Definition eines Zufallsvektors X genügt es also, die Werte $X(\omega)$ für ω aus einer Menge $A \in \mathfrak{A}$ mit $P(A) = 1$ festzulegen.

Die für unsere Zwecke benötigten Verteilungen auf \mathbb{R}^d sind entweder *diskret*, d. h. durch höchstens abzählbar viele Punktmassen auf \mathbb{R}^d spezifiziert, oder über eine Wahrscheinlichkeitsdichte bezüglich des Lebesgue-Maßes auf \mathbb{R}^d , kurz eine *Lebesgue-Dichte*, definiert. Ist P_X diskret, so gibt es Punkte $x_i \in \mathbb{R}^d$ und Gewichte $\alpha_i \geq 0$ mit $\sum_{i=1}^{\infty} \alpha_i = 1$, so daß

$$P_X(A) = \sum_{i=1}^{\infty} \alpha_i 1_A(x_i)$$

für alle meßbaren Mengen $A \subset \mathbb{R}^d$ gilt. Eine Lebesgue-Dichte ist eine Lebesgue-integrierbare, nicht-negative Funktion h auf \mathbb{R}^d , die $\int_{\mathbb{R}^d} h(x) dx = 1$ erfüllt. Ist h die Lebesgue-Dichte der Verteilung von X , kurz die Lebesgue-Dichte von X , so gilt

$$P_X(A) = \int_A h(x) dx$$

für alle meßbaren Mengen $A \subset \mathbb{R}^d$.

Im folgenden betrachten wir reellwertige Zufallsvariablen. Die reelle Zahl

$$E(X) = \int_{\Omega} X(\omega) dP(\omega) = \int_{\Omega} X dP$$

heißt *Erwartungswert* von X , falls das Integral von X bezüglich P existiert. In diesem Fall wird X als integrierbare Zufallsvariable bezeichnet, und wir schreiben kurz $X \in L^1$. Insbesondere ergibt sich

$$E(1_A) = P(A)$$

für jede Menge $A \in \mathfrak{A}$. Der Erwartungswert ist linear und monoton, für $X, Y \in L^1$ und $a, b \in \mathbb{R}$ gilt also $aX + bY \in L^1$ mit

$$E(aX + bY) = a E(X) + b E(Y)$$

und

$$X \leq Y \quad \Rightarrow \quad E(X) \leq E(Y).$$

Zufallsvariablen $X \geq 0$ erfüllen ferner

$$E(X) = 0 \quad \Leftrightarrow \quad P(\{X = 0\}) = 1.$$

Ist P_X diskret und durch Punkte x_i und Gewichte α_i definiert, so ist $X \in L^1$ äquivalent zur absoluten Konvergenz der Reihe $\sum_{i=1}^{\infty} \alpha_i x_i$, und in diesem Fall gilt

$$E(X) = \sum_{i=1}^{\infty} \alpha_i x_i.$$

Besitzt P_X die Lebesgue-Dichte h , so ist $X \in L^1$ äquivalent zu $\int_{\mathbb{R}} |x| h(x) dx < \infty$, und dann gilt

$$E(X) = \int_{\mathbb{R}} x h(x) dx.$$

Ist X quadratisch integrierbar, d. h. X^2 ist integrierbar, so schreiben wir $X \in L^2$. In diesem Fall heißt

$$\sigma^2(X) = E(X - E(X))^2$$

die *Varianz* von X und die Zahl $\sigma(X) = (E(X - E(X))^2)^{1/2}$ wird als *Standardabweichung* von X bezeichnet. Für $X \in L^2$ und $a, b \in \mathbb{R}$ folgt aus der Linearität und Monotonie des Erwartungswertes

$$\sigma^2(X) \geq 0$$

sowie

$$\sigma^2(X) = E(X^2) - (E(X))^2$$

und $aX + b \in L^2$ mit

$$\sigma^2(aX + b) = a^2 \sigma^2(X).$$

Nun sei $X: \Omega \rightarrow \Omega'$ eine Zufallsvariable und $f: \Omega' \rightarrow \mathbb{R}$ eine meßbare Abbildung. Dann ist $f(X) = f \circ X$ eine reellwertige Zufallsvariable, und der *Transformationsatz* sichert

$$E(f(X)) = \int_{\Omega'} f(x) dP_X(x) = \int_{\mathbb{R}} y dP_{f(X)}(y),$$

sofern bereits eines der auftretenden Integrale existiert. Insbesondere zeigt die Wahl von $\Omega' = \mathbb{R}$ und $f = \text{Id}$, daß der Erwartungswert einer reellwertigen Zufallsvariablen nur von ihrer Verteilung abhängt. Gleiches gilt für die Varianz.

Die Tschebyschev-Ungleichung Mit Hilfe der Varianz erhält man eine sehr einfache Abschätzung für die Konzentration einer quadratisch integrierbaren Zufallsvariablen X um ihren Erwartungswert. Für alle $\varepsilon > 0$ gilt

$$P(\{|X - E(X)| \geq \varepsilon\}) \leq \frac{\sigma^2(X)}{\varepsilon^2},$$

und diese Abschätzung heißt *Tschebyschev-Ungleichung*.

Zum Beweis setzen wir

$$A = \{|X - E(X)| \geq \varepsilon\}$$

und erhalten

$$\sigma^2(X) = \int_{\Omega} |X - E(X)|^2 dP \geq \int_A |X - E(X)|^2 dP \geq \int_A \varepsilon^2 dP = \varepsilon^2 \cdot P(A).$$

Daraus folgt die Behauptung.

Zur Anwendung der Tschebyschev-Ungleichung benötigt man nur die quadratische Integrierbarkeit der Zufallsvariablen X . Unter zusätzlichen Annahmen an X liefert die Hoeffding-Ungleichung, die wir in Abschnitt 3.4 kennenlernen werden, viel schärfere Abschätzungen für $P(\{|X - E(X)| \geq \varepsilon\})$.

Unabhängigkeit und bedingte Wahrscheinlichkeiten Eine Familie $(X_i)_{i \in I}$ von Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$ mit Werten in (Ω', \mathfrak{A}') heißt *unabhängig*, falls für beliebige paarweise verschiedene Indizes $i_1, \dots, i_n \in I$ und Mengen $A_1, \dots, A_n \in \mathfrak{A}'$ die Produktformel

$$P\left(\bigcap_{j=1}^n \{X_{i_j} \in A_j\}\right) = \prod_{j=1}^n P(\{X_{i_j} \in A_j\})$$

gilt. In entsprechender Weise definiert man Unabhängigkeit, falls die Zufallsvariablen X_i Werte in unterschiedlichen Räumen $(\Omega_i, \mathfrak{A}_i)$ annehmen.

Für zwei Zufallsvariablen X_1 und X_2 auf $(\Omega, \mathfrak{A}, P)$ mit Werten in $(\Omega_1, \mathfrak{A}_1)$ bzw. $(\Omega_2, \mathfrak{A}_2)$, und Mengen $A_i \in \mathfrak{A}_i$ mit $P(\{X_1 \in A_1\}) > 0$ heißt

$$P(\{X_2 \in A_2\} | \{X_1 \in A_1\}) = \frac{P(\{X_1 \in A_1\} \cap \{X_2 \in A_2\})}{P(\{X_1 \in A_1\})}$$

die *bedingte Wahrscheinlichkeit von $\{X_2 \in A_2\}$ gegeben $\{X_1 \in A_1\}$* . Die beiden Zufallsvariablen sind also genau dann unabhängig, wenn