Elementare Wahrscheinlichkeitstheorie

Stochastik I

Prof. Dr. Uwe Küchler
Institut für Mathematik
Humboldt-Universität zu Berlin

Sommersemester 2006

19. April 2006

e-mail: kuechler@mathematik.hu-berlin.de
www.mathematik.hu-berlin.de/ kuechler

Inhaltsverzeichnis

1 Einleitung **3**

2 Zufällige Versuche und zufällige Ereignisse **9**
2.1 Zufällige Versuche 9
2.2 Zufällige Ereignisse 11
2.3 Verknüpfung von Ereignissen 13
2.4 Ereignisse und Mengen 14
2.5 Münzenwurf 15

3 Wahrscheinlichkeiten und Zufallsgrößen **19**
3.1 Axiomatik und erste Folgerungen 20
3.2 Laplace-Experimente 29
3.3 Münzenwurf, zum Zweiten 32
3.4 Was sagen Wahrscheinlichkeiten? 34
3.5 Elemente der Kombinatorik* 36
3.6 Zufallsgrößen 44
3.7 Verteilungsfunktionen 55

Kapitel 1

Einleitung

Diese Vorlesung handelt vom Zufall, genauer, von der Mathematik des Zufalls. Mit dem Zufall haben wir täglich zu tun. Wir sagen, ein Ereignis hänge vom Zufall ab und sprechen von einem *zufälligen Ereignis*, wenn es nicht gewiss ist, ob dieses Ereignis eintritt oder nicht. Eine Größe, deren Wert nicht genau vorhergesagt werden kann, bezeichnen wir als eine *zufällige Größe* oder *Zufallsgröße*.
So wird zum Beispiel die zeitliche Dauer des Weges von zu Hause zur Universität vom Zufall beeinflusst. Der finanzielle Schaden, den ein PKW-Unfall verursacht, hängt vom Zufall ab, Messungen physikalischer Größen werden vom Zufall beeinträchtigt.
Bedeutende Größen zufälliger Natur sind weiterhin z.B. die Lebenszeit von Menschen, das Geschlecht Neugeborener, Niederschlagsmengen eines Monats, aber auch Kurse am Aktienmarkt und das Ergebnis von Fußballspielen.

Zufall in reiner Form findet man bei Glücksspielen. Die Augenzahl beim Werfen eines Würfels, die Zahl, bei der die Roulettekugel nach dem Ausrollen liegen bleibt, das Skatblatt, das man nach dem Austeilen erhalten hat, sind rein zufällig. Hier ist der Zufall erwünscht.
Der Einfluss des Zufalls wird allerdings häufig als unangenehm und störend empfunden. Er verursacht Risiken, also die Gefahr, dass Schäden entstehen. Wir versuchen ihn deshalb zurückzudrängen, möglichst auszuschließen. Das Ergebnis einer Klausur oder Prüfung wollen wir möglichst nicht vom Zufall in negativer Hinsicht beeinflusst wissen, also bereiten wir uns möglichst gut vor (ein Restrisiko bleibt natürlich: Aufgaben, die wir nicht lösen können, nervliche Anspannung usw.) Unfällen, die häufig durch "Verkettung unglücklicher

Umstände" entstehen, beugt man durch technische Überwachung, Schulungen usw. vor.
Völlig unterdrücken lässt sich der schädliche Zufall meist nicht oder nur mit extrem großem Aufwand. Wenn man ihn aber schon nicht eliminieren kann, so möchte man die durch ihn verursachten Risiken aber einschätzen, um vorbereitet zu sein. Ein Mittel dafür ist die Stochastik, die Mathematik des Zufalls.

Der Zufall macht zukünftige Ereignisse ungewiss. Er schafft Risiken, aber auch Chancen. Die Stochastik stellt mathematische Verfahren zur Verfügung, mit deren Hilfe man zufällige Erscheinungen, Chancen und Risiken, rechnerisch bewerten kann. Sie gliedert sich in *Wahrscheinlichkeitstheorie* und *Mathematische Statistik*.

Unter der *Wahrscheinlichkeit eines Ereignisses* verstehen wir, kurz gesagt, den *Grad der Gewissheit seines Eintretens*. Sie kann Werte zwischen Null und Eins annehmen. Dabei ordnet man Ereignissen, die praktisch nicht eintreten können, den Wert Null zu, Ereignissen, die mit Sicherheit eintreten, den Wert Eins. Am meisten unbestimmt ist das Eintreten von Ereignissen mit der Wahrscheinlichkeit 1/2.

Mit dem Zufall haben die Menschen seit jeher zu tun. Wetter, Krankheit, Nahrungssuche, Tod waren fundamentale Größen, von denen die Menschen abhingen und die durch Zufall geprägt waren. Es gab vielfältige Versuche, hier dem Zufall auf die Spur zu kommen, seine Herrschaft einzudämmen bzw. sein Wirken aufzuklären. Die ersten mathematischen Ansätze haben ihre Ursprünge in Problemen, die sich bei der Organisation von Versicherungsgesellschaften (z.B. Berechnung von Versicherungsprämien) ergaben und solchen, die bei der Beurteilung von Fragen von Glücksspielen entstanden.
Im Laufe der Zeit hat die Bedeutung des richtigen Umganges mit dem Zufall, insbesondere seine angemessene quantitative Beschreibung, noch zugenommen. In modernen technischen Produkten wie Flugzeugen, Schiffen, Eisenbahnen sind außerordentlich viele Einzelteile vereint und müssen für einen reibungslosen Ablauf zuverlässig funktionieren. Kleine zufällige Störungen können Unfälle mit größten Sach- und Personenschäden verursachen. Ohne eine genaue mathematische Analyse aller auftretenden Risiken wäre der Betrieb solcher technische Produkte nicht mehr denkbar, Unglücke würden wesentlich öfter auftreten. Aber auch für das Verständnis vieler Naturvorgänge kommt man heute nicht mehr ohne Einbeziehung des Zufalls aus.

Denken wir an die Evolution im Pflanzen- und Tierreich, die man auf zufällige Mutationen mit anschließender natürlicher Auslese zurückführt, oder an die Begriffswelt der Quantenphysik, wo man den Ort eines Teilchens nicht mehr exakt bestimmen kann und als Wahrscheinlichkeitsverteilung modelliert.
Zufall, Chance und Risiko sind Begriffe, die im Gegensatz zu vielen physikalischen Größen wie Temperatur, Länge, Gewicht nicht sinnlich wahrnehmbar sind. Für viele Menschen haben sie heute noch etwas Unheimliches, Mystisches und auch Reizvolles an sich. Die Abhängigkeit vom Zufall weckt Hoffnung und Angst. Die Menschen spielen Lotto bzw. tragen ein Maskottchen. Das muss nicht schlimm sein. Problematisch ist es erst, wenn man Risiken und Chancen, auch für das persönliche Leben, nicht richtig abschätzt und damit eventuell Gefahren für Hab und Gut oder gar die Gesundheit und das Leben eingeht.

Was ist nun eigentlich Zufall? Diese alte und schwierige Frage gehört zur Philosophie und soll hier nicht behandelt werden.

Die Vorlesung soll die Hörer mit einigen grundlegenden Begriffen der Stochastik bekannt machen, typische Denk- und Schlussweisen vorstellen und einige wichtige Gesetzmäßigkeiten des Zufalls nahebringen.

Das Wissen um den Zufall, die damit verbundenen Begriffe und Methoden des Umganges mit dem Zufall ist heute Allgemeingut aller Wissenschaftszweige und gehört auch um notwendigen Alltagswissen. Vielen Menschen ist das genauere Wissen um den Zufall jedoch noch fremd.

Das vorliegende Skript entstand auf der Grundlage von Vorlesungen über Elemente der Stochastik, die ich in Abständen von Jahren mehrfach für Studierende der Mathematik gehalten habe. Dennoch ist die schriftliche Ausarbeitung eines Skriptes immer auch eine aufwändige Arbeit. Ich danke unserer Institutssekretärin Frau S. Bergmann für die umfangreiche Arbeit am Skript und die unendliche Geduld gegenüber meinen zahlreichen Änderungswünschen.

Dieses Skript enthält sicher noch eine Reihe von Fehlern, insbesondere Druckfehler, aber auch andere Unzulänglichkeiten. Dafür bin ich allein verantwortlich. Ich bitte die Hörer der Vorlesung und andere Leser um Nachsicht und bin sehr dankbar für kritische und helfende Hinweise.
Vieles wird in der Vorlesung vorkommen, zum Beispiel Bilder, Graphiken, interessante Beispiele und Anwendungen, das im Skript noch nicht Platz ge-

funden hat. Andererseits werden in der Vorlesung manche Ausführungen des Skriptes nur gestreift werden, zum Beispiel gewisse Elemente der Maß- und Integrationstheorie. Sie wurden teilweise nur aufgenommen, um eine gewisse Geschlossenheit der Darstellung und eine Festlegung der Terminologie zu erreichen.

An der Lehrveranstaltung Stochastik I werden mehrere erfahrene wissenschaftliche Mitarbeiter und Studierende höherer Semester als Übungsleiter und Korrektoren der schriftlich anzufertigenden Übungen beteiligt sein. Wir sind gespannt, freuen uns auf die Arbeit mit den Studierenden im Sommersemester 2006 und wünschen uns allen viel Erfolg!

Uwe Küchler

Berlin, 13. April 2006

Literaturverzeichnis

[1] Bauer, H.: Maß- und Integrationstheorie, de Gruyter, 1990

[2] Bauer, H.: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie. 4. Auflage, de Gruyter, Berlin, 1991

[3] Dacunha-Castelle, D. und Duflo, M.: Probability and Statistics, Vol. I + II, Springer, 1986

[4] Dinges, H.; Rost, H.: Prinzipien der Stochastik, B.G. Teubner, Stuttgart, 1982

[5] Elstrodt, J.: Maß- und Integrationstheorie, Springer, 1999

[6] Feller, W.: An Introduction to Probability Theory and its Applications. Vol. 1, Wiley, N.Y., 3. Auflage, 1968

[7] Henze, N.: Stochastik für Einsteiger, 4. Auflage, 2003

[8] Hesse, Ch.: Angewandte Wahrscheinlichkeitstheorie, Vieweg, 2003

[9] Krengel, U.: Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg, 6. Auflage, 2002

[10] Jacod, J. and Protter, Ph.: Probability Essentials, Springer 2000

[11] Müller, P.H. u.a.: Lexikon der Stochastik, Akademie-Verlag, Berlin, 5. Auflage, 1991

[12] Müller, P.H.; Neumann, R. und Storm, R.: Tafeln zur Mathematischen Statistik, Fachbuchverlag, Leipzig 1973

[13] Pfanzagl, J.: Elementare Wahrscheinlichkeitsrechnung, de Gruyter Berlin, 2. Auflage, 1991

[14] Pitmann, J.: Probability, Springer 1993

[15] Renyi, A.: Briefe über die Wahrscheinlichkeit, Dt. Verlag der Wiss., Berlin, 1969

[16] Ripley, B.D.: Stochastic Simulation, Wiley, 1987

[17] Siraev, A.N.: Wahrscheinlichkeit, Dt. Verlag der Wiss., Berlin, 1988

[18] Winkler, M.: Vorlesungen zur Mathematischen Statistik, Teubner, Leipzig, 1983

Kapitel 2

Zufällige Versuche und zufällige Ereignisse

In diesem Kapitel führen wir zunächst anschaulich die grundlegenden Begriffe des zufälligen Versuchs und des zufälligen Ereignisses ein und stellen danach eine Verbindung zur Mengenlehre her. Damit wird die Grundlage einer mathematischen Theorie des Zufalls, der Wahrscheinlichkeitstheorie gelegt. Wir erfahren, dass die mit einem zufälligen Versuch verbundenen zufälligen Ereignisse eine σ-Algebra bilden.

2.1 Zufällige Versuche

Unter einem *zufälligen Versuch* versteht man einen Versuch (im weitesten Sinne des Wortes), dessen Ausgang unter bestimmten wesentlichen und fixierten Bedingungen im Rahmen bestimmter Möglichkeiten ungewiss ist.

Bezeichnung:

Mögliche Versuchsausgänge (-ergebnisse): ω
Menge aller möglichen Versuchsausgänge: Ω

Beispiele:

a) Werfen einer Münze: $\Omega = \{Z, W\}$ oder $\Omega = \{-1, +1\}$,

b) Werfen zweier unterschiedlicher Münzen: $\omega = (\omega_1, \omega_2)$
$\omega_i =$ Wurfergebnis der i-ten Münze,
$\omega_i \in \{Z, W\}, \quad i = 1, 2$
$\Omega = \{(Z, Z,), (Z, W), (W, Z), (W, W)\}$

c) n-maliges Werfen einer Münze, $\omega = (\omega_1, \cdots, \omega_n), \omega_i \in \{Z, W\}, i = 1, 2, \cdots, n$

d) Werfen eines Würfels:
$\Omega = \{1, 2, \cdots, 6\}$

e) Werfen zweier unterscheidbarer Würfel: $\omega = (i, j), \quad i$ Augenzahl des ersten, j Augenzahl des zweiten Würfels
$\Omega = \{(i, j) : i, j \in \{1, 2, \cdots, 6\}\}$

f) Sonntagsziehung im Lotto "6 aus 49" (ohne Zusatzzahl):
$\Omega = \{\{i_1, \cdots, i_6\} : i_1, \cdots, i_6 \in \{1, \cdots, 49\}, \quad i_j \neq i_k (j \neq k)\}$

g) Geburt eines Kindes, es werden registriert Gewicht η in g, Größe ξ in cm, Geschlecht τ:
$\omega = (\eta, \xi, \tau), \quad \Omega = (0, \infty) \times (0, \infty) \times \{m, w\}$

h) Niederschlagsmenge pro Quadratmeter in mm am 30. 10. 2006 auf dem Alexanderplatz: $\Omega = [0, \infty)$

i) Schadenhöhe ω bei einem PKW-Unfall, die der Versicherer in Euro zu zahlen hat:
$\Omega = [0, \infty)$

j) Anzahl aller polizeilich gemeldeten Kfz-Unfälle an einem bestimmten Tag auf der Rudower Chaussee in Adlershof: $\Omega = \{0, 1, \cdots, n, \cdots\}$

Diskussion:

1. Ω ist nicht eindeutig festgelegt. Einzige Bedingung: nach Ausführung des Versuches muss genau ein ω aus Ω als Versuchsergebnis feststehen. Insbesondere ist es nicht notwendig, dass alle $\omega \in \Omega$ auch tatsächlich auftreten können, Ω kann also größer gewählt werden als unbedingt notwendig (vgl. Beispiele g) - j)).
Erweiterungen von Ω sind aus mathematischen Gründen oder wegen Übersichtlichkeit häufig vorteilhaft.

2. Vor dem Versuch ist der tatsächlich auftretende Ausgang $\omega \in \Omega$ ungewiss. Sicher ist nur, dass genau eines der ω aus Ω auftreten wird. Nach dem Versuch liegt der aufgetretene Ausgang ω fest. Die Ungewißheit ist verschwunden, das Ergebnis steht fest. Der Versuch wurde *realisiert, verwirklicht*. Das nach dem Versuch erschienene ω, eine Zahl in a), d), h), i), j) oder allgemeinere Ergebnisse in den anderen Beispielen, heißt *Realisierung dieses Versuches*. Bei erneuter Ausführung des Versuches tritt i.A. ein anderer Ausgang in Erscheinung, es erscheint eine andere Realisierung.
   Wird der Versuch mehrmals durchgeführt, ergibt sich eine Folge von Realisierungen, eine sogenannte (konkrete) $(\omega, \eta, \cdots, \kappa)$ *Stichprobe*, ein *Datensatz*.

2.2 Zufällige Ereignisse

Ein *zufälliges Ereignis* (oder kurz Ereignis) ist ein Ereignis, das (im Rahmen eines bestimmten zufälligen Versuches) eintreten kann, aber nicht eintreten muss.
Zufällige Ereignisse beschreibt man häufig verbal durch eine logische Aussage.
Bezeichnung: $A, B, C, \cdots$

Betrachten wir einige Ereignisse im Zusammenhang mit den Beispielen aus Punkt 2.1.

a) $A :=$ "Es erscheint das Wappen"

e) $A :=$ "Die Summe der Augenzahlen ist gerade"

f) $A :=$ "Es erscheint mindestens ein Zahlenzwilling"
   $B :=$ "Der abgegebene Tippschein enthält 3 Richtige"

h) $A :=$ "Es regnet mehr als 10 mm"

i) $A :=$ "Der Schaden ist größer als 100 000 EUR"

Man sagt: *Das Ereignis A* tritt (mit Versuchsdurchführung und dem Versuchsausgang ω) *ein*, wenn die zugehörige logische Aussage bei diesem ω wahr wird,

es tritt nicht ein, wenn sie bei diesem ω falsch ist.

Man sagt, *das Ereignis A zieht das Ereignis B nach sich* oder sei ein Teil von B, falls aus dem Eintreten von A folgt, dass auch B eintritt. Symbolisch: $A \subseteq B$.

Im Beispiel e) gilt mit

A : = "Die Summe der Augenzahlen ist gerade" und

B : = "Es erscheint eine der Zahlen 4, 6, 8, 10, 11, 12" die Beziehung $A \subseteq B$.

Zwei Ereignisse A und B heißen einander *gleich* ($A = B$), wenn das Eintreten des einen Ereignisses das Eintreten des anderen nach sich zieht: ($A \subseteq B$ und $B \subseteq A$)

Im Beispiel d) gilt mit

A : = "Es erscheint eine ungerade Zahl"

B : = "Die gewürfelte Augenzahl ist nicht 2, 4 oder 6" die Beziehung $A = B$.

Ein zufälliges Ereignis A heißt *mit einem gegebenen zufälligen Versuch verbunden*, falls man für jeden Versuchsausgang ω entscheiden kann, ob er zum Eintreten von A führt oder nicht.
Anders formuliert:

Das Ereignis B : = "Der abgegebene Tippschein enthält drei Richtige" ist mit dem zufälligen Versuch einer Sonntagsziehung im Zahlenlotto verbunden.

Das Ereignis A heißt mit dem zufälligen Versuch Ω verbunden, wenn man nach Ausführung des Versuches entscheiden kann, ob A eingetreten ist oder nicht. Das Ereignis "Morgen scheint die Sonne mindestens zwei Stunden" ist nicht mit dem zufälligen Versuch des Werfens eines Würfels verbunden.

Um die Notationen abzurunden, führen wir als zufällige Ereignisse auch folgende zwei Ereignisse ein. Ein Ereignis S heißt *sicheres Ereignis*, falls es bei jedem Versuchsausgang eintritt, ein Ereignis U nennt man *unmögliches Ereig-*

nis, wenn es bei keinem Versuchsausgang eintritt.

Wir erinnern daran, dass wir mit jedem zufälligen Versuch eine Menge Ω festlegen, die alle möglichen Ausgänge des Versuches enthält.

Die Menge aller Ereignisse A, die mit einem gegebenen zufälligen Versuch verbunden sind, wird mit $\mathfrak{A}$ bezeichnet.

Das Paar $(\Omega, \mathfrak{A})$ ist für uns das vorläufige Modell eines zufälligen Versuches.

2.3 Verknüpfung von Ereignissen

Es sei $(\Omega, \mathfrak{A})$ ein zufälliger Versuch, d. h. Ω enthalte die Menge aller möglichen Versuchsausgänge ω und $\mathfrak{A}$ sei das System der mit dem Versuch verbundenen Ereignisse.

Aus gegebenen Ereignissen $A, B \in \mathfrak{A}$ lassen sich neue Ereignisse bilden.

Das Ereignis $A \cup B$ tritt ein, falls A eintritt oder B eintritt (oder beide).
$A \cup B$ nenn man die *Vereinigung von A und B*.

Das Ereignis $A \cap B$ tritt ein, falls A und B beide eintreten.
$A \cap B$ nennt man den *Durchschnitt von A und B*.

$\overline{A}$ tritt genau dann ein, falls A nicht eintritt.
$\overline{A}$ heißt das zu A *komplementäre Ereignis*. Es gilt $\overline{U} = S$ und $\overline{S} = U$.

$A \backslash B$ tritt genau dann ein, wenn A eintritt und B nicht eintritt.
$A \backslash B$ heißt die *Differenz von A und B*. Es gilt $A \setminus B = A \cap \overline{B}$.

Das Ereignis $A \triangle B := (A \setminus B) \cup (B \setminus A)$ heißt *symmetrische Differenz* von A und B.

Wenn $A \cap B = U$ gilt, so heißen A und B *disjunkt* oder *unvereinbar*.
Es gilt stets: $A \cap \bar{A} = U, A \cup \bar{A} = S$.

Sind $A_k, k = 1, \cdots, m$, Ereignisse aus $\mathfrak{A}$, so bezeichne $\bigcup\limits_{k=1}^{m} A_k$ das Ereignis, das genau dann eintritt, wenn mindestens eines der Ereignisse A_k eintritt, und $\bigcap\limits_{k=1}^{m} A_k$ das Ereignis, das genau dann eintritt, wenn alle A_k eintreten.

Analog definiert man zu jeder Folge $A_k, k \geq 1$, aus $\mathfrak{A}$ die Ereignisse $\bigcup\limits_{k=1}^{\infty} A_k$ und $\bigcap\limits_{k=1}^{\infty} A_k$.

Folgerung: Die Menge $\mathfrak{A}$ aller mit einem zufälligen Versuch verbundenen Ereignisse hat also die Eigenschaften:

1) $U, S \in \mathfrak{A}$,

2) Für jedes $A \in \mathfrak{A}$ ist auch $\overline{A} \in \mathfrak{A}$,

3) Für jedes $n \geq 2$ und alle $A_1, A_2, \cdots, A_n \in \mathfrak{A}$ gilt $\bigcup\limits_{k=1}^{n} A_k \in \mathfrak{A}$.

   Auf Grund dieser Eigenschaften und der Definition von U und S heißt $\mathfrak{A}$ eine *Algebra* (bez. der Operationen $\bigcup, ^-$) mit Nullelement U und Einselement S.

   Da außerdem

4) Für alle $A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A}$ gilt $\bigcup\limits_{k=1}^{\infty} A_k \in \mathfrak{A}$

   erfüllt ist, nennt man $\mathfrak{A}$ auch eine *σ-Algebra*.

2.4 Ereignisse und Mengen

Es sei $(\Omega, \mathfrak{A})$ ein zufälliger Versuch.
Jedes mit diesem Versuch verbundene Ereignis A, d.h. jedes A aus $\mathfrak{A}$, wird durch eine Teilmenge A' von Ω charakterisiert:

$A \longleftrightarrow A' = \{\omega \in \Omega :$ das Eintreten von ω zieht das Eintreten von A nach sich $\}$

Wenn A eintritt, so ist ein Versuchsausgang ω eingetreten, der zu A' gehört. Wenn ein $\omega \in A'$ als Versuchsausgang auftritt, so tritt auch A ein. A und "Es erscheint ein $\omega \in A'$ als Versuchsergebnis" sind somit einander gleiche Ereignisse, d. h. entweder treten sie beide ein oder beide nicht. Insofern charakterisiert die Teilmenge A' von Ω das Ereignis A. Identifiziert man A mit seiner zugehörigen Menge A', so können wir feststellen:

Merke:

Zufällige Ereignisse, die mit einem zufälligen Versuch verbunden sind, sind Teilmengen von Ω, m. a. W., die σ-Algebra $\mathfrak{A}$ ist ein System von Teilmengen von Ω. Bei dieser Entsprechung wird das sichere Ereignis mit Ω, das unmögliche Ereignis mit der leeren Menge $\emptyset$ identifiert. Die Korrespondenz $A \longleftrightarrow A'$ ist bezüglich der Operationen Vereinigungs-, Durchschnitts-, Differenz- und Komplementbildung für Ereignisse bzw. für Mengen ein Isomorphismus.

In den Beispielen a) bis f) und j) ist $\mathfrak{A} = \mathfrak{P}(\Omega)$ (Potenzmenge von Ω) in den Beispielen g) bis i) wählt man i. a. $\mathfrak{A} \subsetneq \mathfrak{P}(\Omega)$, die Begründung werden wir später kennen lernen.

2.5 Münzenwurf

Wir formulieren zum Ende dieses Abschnittes noch ein mathematisches Modell, auf das wir später mehrfach zurückkommen werden. Eine Münze werde n-mal geworfen. Die möglichen Ausgänge dieser Wurfserie sind die n-tupel $\omega = (x_1, x_2, \cdots, x_n)$ mit $x_k \in \{-1, 1\}, k = 1, 2, \cdots, n$, wobei $x_k = +1 (= -1)$ gesetzt wird, falls beim k-ten Wurf die Zahl (bzw. das Wappen) oben liegt.
Die Menge Ω aller möglichen Ausgänge ω der Wurfserie besteht aus 2^n Elementen.
Für jedes $\omega = (x_1, x_2, \cdots, x_n) \in \Omega$ definieren wir

$$s_0 = 0, s_k = \sum_{l=1}^{k} x_l, k = 1, 2, \cdots, n$$

und nennen die Folge

$$s = ((k, s_k), k = 0, 1, \cdots, n)$$

den zu ω gehörenden Pfad. Wir veranschaulichen jeden Pfad s durch die Punkte (k, s_k) in der Ebene und verbinden die benachbarten Punkte linear. Diese Pfade s haben die Eigenschaft

$$s_0 = 0 \text{ und } |s_k - s_{k-1}| = 1, k = 1, 2, \cdots, n$$

und entsprechen den Versuchsausgängen ω eineindeutig.

Vereinbart man ein Spiel, in dem der Spieler A von einer Bank den Betrag +1 erhält, falls im Ergebnis eines Münzenwurfes die Zahl erscheint, und er den Betrag 1 zu zahlen hat, wenn Wappen oben liegt, so ist sein Gewinn nach k Würfen gleich s_k.

Das Ereignis

A: = "Spieler A hat zum Schluß einen positiven Betrag gewonnen"

tritt genau dann ein, wenn der Versuchsausgang $\omega = (x_1, x_2, \cdots, x_n)$ zur Menge

$$A = \{\omega \in \Omega | s_n = \sum_{k=1}^{n} x_k > 0\}$$

gehört. Zu A gehören also alle Pfade, die nach n Schritten im Positiven enden.

Das Ereignis

B:= "Das Guthaben des Spielers A sinkt im Verlauf des Spieles niemals unter Null"

tritt genau dann ein, wenn ein Versuchsausgang $\omega \in \Omega$ mit

$$\min_{k=1,2,\cdots,n} s_k \geq 0$$

auftritt, d. h., wenn der zugehörige Pfad niemals die -1 berührt.

Zur Vereinfachung der Bezeichnung führen wir folgende Notation ein:

$X_k(\omega) := x_k, k = 1, 2, \cdots n,$

$$S_m(\omega) := \sum_{k=1}^{m} X_k(\omega) = \sum_{k=1}^{m} x_k, m = 1, 2, \cdots, n.$$

Die "Zufallsgröße" X_k gibt das Ergebnis des k-ten Wurfes an, die "Zufallsgröße" $S_m(\omega)$ ist der Gewinn des Spielers A nach m Würfen, $m = 1, 2, \cdots, n$.

Für die oben eingeführten Ereignisse A und B, gilt dann

$A = \{\omega \in \Omega | S_n(\omega) > 0\}$, kurz geschrieben $A = \{S_n > 0\}$, und

$B = \{\omega \in \Omega | \min_{k=1,\cdots,n} S_k(\omega) \geq 0$, oder kurz $B = \{\min_{k=1,\cdots,n} S_k \geq 0\}$.

Kapitel 3

Wahrscheinlichkeiten und Zufallsgrößen

Zufällige Ereignisse unterscheiden sich im Grad der Gewissheit ihres Eintretens. Als Gradmesser für diese Gewissheit wird die Wahrscheinlichkeit des betrachteten Ereignisses angesehen.

Es ist eine Erfahrungssache, dass sich die relative Häufigkeit, mit der ein Ereignis in einer langen Reihe von Versuchen, die immer wieder neu unter gleichartigen Bedingungen ausgeführt werden, um einen festen Wert stabilisiert. Diesen Wert könnte man als Grad der Gewissheit des Eintretens des Ereignisses ansehen. Ausgehend von dieser Vorstellung formulieren wir einige plausible Eigenschaften von Wahrscheinlichkeiten, die sich dann auch wieder finden im Axiomensystem der Wahrscheinlichkeitstheorie.

Zu den mathematisch übersichtlichen zufälligen Versuchen gehören die Laplace-Versuche. Sie besitzen nur endlich viele und dabei gleichwahrscheinliche Ausgänge. Die Bestimmung von Wahrscheinlichkeiten der mit ihnen zusammenhängenden Ereignisse läuft auf das Abzählen gewisser Fälle, häufig unter Verwendung kombinatorischer Formeln, hinaus.

Der Begriff der Zufallsgröße gehört ebenfalls zum Grundbestand der Wahrscheinlichkeitstheorie. Zufallsgrößen vermitteln stets Teilaspekte eines zufälligen Versuchs und fungieren als beobachtbare (bzw. interessierende) Größen, wenn der Ausgang des Versuches selbst nicht beobachtbar ist (bzw. nicht von Interesse ist). Von Wichtigkeit sind die von ihnen induzierten Wahrscheinlich-

keitsverteilungen.

3.1 Axiomatik und erste Folgerungen

Wir betrachten einen zufälligen Versuch und nehmen an, Ω sei die Menge seiner möglichen Versuchsausgänge ω und $\mathfrak{A}$ die σ-Algebra der mit diesem Versuch verbundenen Ereignisse, d. h. eine σ-Algebra von Teilmengen von Ω.

Wir wiederholen noch einmal ihre charakteristischen Eigenschaften:

1. $\Omega, \in \mathfrak{A}$,
2. Für alle $A \in \mathfrak{A}$ gilt $\overline{A} \in \mathfrak{A}$,
3. Für alle $n \geq 2$ und $A_1, A_2, \cdots, A_n \in \mathfrak{A}$ ist auch $\bigcup\limits_{k=1}^{n} A_k \in \mathfrak{A}$,
4. Für alle $A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A}$ gilt auch $\bigcup\limits_{k=1}^{\infty} A_k \in \mathfrak{A}$.

Wir wissen, dass jedes Ereignis A aus $\mathfrak{A}$ bei der Versuchsausführung eintreten kann, aber nicht eintreten muss. Sein Eintreten ist ungewiss. Unterschiedliche Ereignisse können sich allerdings im *Grad der Gewissheit* ihres Eintretens unterscheiden.

Beispiel: Werfen einer Zündholzschachtel.

Versuchsausgang sei eine der drei Seiten, auf denen die Schachtel zu liegen kommt: Große Seite (Ober- bzw. Unterseite) / Mittlere Seite (Seiten mit Reibflächen) / Kleine Seite (Stirn- bzw. Hinterseite): Als Ω wählen wir die Menge $\Omega = \{G, M, K\}$.
Wir bemessen den Grad der Gewissheit des Eintreffens jeder der möglichen Fälle aus der Erfahrung. Dieser Grad wird umso höher eingeschätzt, je häufiger bei längerer Versuchsreihe die Schachtel auf der entsprechenden Seite zu liegen kommt.

Formal: A sei ein Ereignis, das mit einem zufälligen Versuch verbunden ist. Der Versuch werde n mal durchgeführt. Jedes Mal unter im Wesentlichen gleichartigen Bedingungen und unabhängig voneinander.

In $n(A)$ Fällen trete A ein. Dann zeigt die *relative Häufigkeit* des Eintretens von A in n Versuchen, nämlich $\frac{n(A)}{n}$, mit wachsendem n eine bemerkenswerte Stabilität: $\frac{n(A)}{n}$ verändert sich immer weniger, sie scheint gegen einen Grenzwert zu konvergieren, wenn n unbegrenzt wächst.
Wir nennen diese Erscheinung das *empirische Gesetz der großen Zahlen.*

Beispiel: Werfen eines Kronenverschlusses:

Relative Häufigkeit dafür, dass die offene Seite nach oben zeigt:

Zahl der Versuche

100	200	300	400	500	600	700
0,7300	0,7750	0,7767	0,7750	0,7800	0,7900	0,7943
800	900	1000				
0,8012	0,7967	0,7910				

(Nach Nawrotzki, K., Lehrmaterial zur Ausbildung von Diplomlehrern Mathematik, Jena 1984)

Beispiel:
Skatspiel: Die relativen Häufigkeiten bestimmter Konstellationen prägen sich beim Spieler ein, Erfahrungstatsachen: Zwei Buben im Skat sind relativ selten. Daraus wird geschlossen, dass auch im nächsten Spiel nur mit geringer Chance zwei Buben im Skat zu finden sein werden. Es ergibt sich eine Verhaltensgrundlage: "Auf den Skat reizt man nicht".

Aus dem genannten Gesetz der großen Zahlen leitet man die Auffassung ab, dass es zu jedem zufälligen Ereignis A eine Zahl $P(A)$ gibt, die *Wahrscheinlichkeit* von A, die den Grad der Gewissheit des Eintretens von A (in einem einzelnen Versuch) ausdrückt.

Welche Eigenschaften hat die Zuordnung $A \to P(A)$, $A \in \mathfrak{A}$, und wie bestimmt man $P(A)$?

Man könnte $\frac{n(A)}{n}$ als Grad für die Gewissheit des Eintretens von A in einem einzelnen (im nächsten) Versuch ansehen.

Diese Verwendung ist problematisch. Bei einer zweiten Versuchsreihe kann $\frac{n(A)}{n}$ einen anderen Wert annehmen, dasselbe passiert, wenn man n unterschiedlich groß wählt. Andererseits zeigt sich bei großem und immer weiter wachsendem n eine bemerkenswerte Stabilisierung der relativen Häufigkeit $\frac{n(A)}{n}$. Man könnte also $\lim\limits_{n\to\infty} \frac{n(A)}{n}$ nehmen. Aber wie groß ist dieser Wert? Und existiert er überhaupt?

Man hat versucht, auf dieser Beobachtung eine Wahrscheinlichkeitstheorie aufzubauen. Ein prominenter Vertreter dieser Richtung war Richard von Mises, (1883 - 1953) (siehe Wahrscheinlichkeit, Statistik, Wahrheit, 1936, Springer Verlag Wien).
Sie wurde längere Zeit von Ingenieuren und Naturwissenschaftlern mit Erfolg verwendet, scheiterte aber an der Untersuchung komplexer zufälliger Erscheinungen, wie zufälliger Prozesse.
Allerdings vermittelt das empirische Gesetz der großen Zahlen erste Vorstellungen, welche Eigenschaften Wahrscheinlichkeiten aufweisen sollten.

Wir gehen dabei von folgender *Grundüberzeugung* aus:

Jedes zufällige Ereignis A besitzt im Rahmen des zugrunde liegenden zufälligen Versuches als Grad der Gewissheit seines Eintretens (in einem einzelnen Versuch) eine Zahl $P(A)$, die Wahrscheinlichkeit des Eintretens von A. Für lange Versuchsreihen sollte das die Zahl sein, um die sich $\frac{n(A)}{n}$ stabilisiert.

$\frac{n(A)}{n} \approx P(A)$.

Daraus ergeben sich plausible Eigenschaften für $P(A)$:

$0 \leq P(A) \leq 1$,

$P(\Omega) = 1, \quad P(\emptyset) = 0$

$P(A \cup B) = P(A) + P(B), \text{falls } A \cap B = \emptyset$

Aus diesen Vorstellungen hat sich ein Axiomensystem entwickelt, das 1933 A.N. Kolmogorov in einer berühmten Arbeit eingeführt hat. (Kolmogorov, Grundlagen der Wahrscheinlichkeitstheorie, Springer, Berlin 1933).

Dieses Axiomensystem ist heute die Grundlage der Wahrscheinlichkeitstheorie und Mathematischen Statistik und lautet wie folgt.

Axiomensystem der Wahrscheinlichkeitstheorie

Wir stellen uns wieder $(\Omega, \mathfrak{A})$ als einen zufälligen Versuch vor, die Elemente ω von Ω bilden die Versuchsausgänge, $\mathfrak{A}$ sei die σ-Algebra der mit dem Versuch verbundenen Ereignisse, also eine σ-Algebra von Teilmengen von Ω.

Als *Wahrscheinlichkeitsverteilung* $P(\cdot)$ auf der σ-Algebra $\mathfrak{A}$ von Teilmengen einer nichtleeren Menge Ω bezeichnet man jede Abbildung P von $\mathfrak{A}$ in $[0,1]$ mit

1. $P(\Omega) = 1$ und $P(\emptyset) = 0$

2. Für jedes $n \geq 2$ und jede Folge

   $(A_k, k = 1, \cdots, n)$ aus $\mathfrak{A}$ mit

   $A_k \cap A_l = \emptyset, k \neq l$ (paarweise Unvereinbarkeit) gilt

   $$P\Big(\bigcup_{k=1}^{k} A_k\Big) = \sum_{k=1}^{k} P(A_k)$$

   (Endliche Additivität der Wahrscheinlichkeitsverteilung P)

2.' Für jede abzählbar unendliche Folge $(A_k, k \geq 1)$ aus $\mathfrak{A}$ mit

   $A_k \cap A_l = \emptyset \quad , \quad k \neq l$ (paarweise Unvereinbarkeit) gilt

$$P\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} P(A_k)$$

(σ-Additivität, abzählbare Additivität der Wahrscheinlichkeitsverteilung)

$(\Omega, \mathfrak{A}, P)$ ist mit dieser Definition ein (normierter) Maßraum, P heißt auch ein *Wahrscheinlichkeitsmaß* auf $\mathfrak{A}$. Statt *Wahrscheinlichkeitsverteilung* $P(\cdot)$ *auf* $\mathfrak{A}$, sprechen wir einfach auch von einer *Verteilung* $P(\cdot)$ *auf* $\mathfrak{A}$.

Definition: *Sind Ω eine nichtleere Menge, $\mathfrak{A}$ eine σ-Algebra von Teilmengen von Ω und P eine Abbildung von $\mathfrak{A}$ in $[0,1]$ mit den Eigenschaften 1., 2. und 2'., so heißt das Tripel $(\Omega, \mathfrak{A}, P))$ ein Wahrscheinlichkeitsraum.*

Jeder Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ ist das mathematische Modell eines zufälligen Versuches. Ω enthält dabei die Menge der möglichen Versuchsergebnisse, $\mathfrak{A}$ entspricht der Menge der mit dem Versuch verbundenen Ereignisse, P ist die Wahrscheinlichkeitsverteilung des Versuches. P legt fest, mit welcher Wahrscheinlichkeit jedes mit dem Versuch verbundene Ereignis bei der Versuchsdurchführung eintritt.

Die Axiome bilden kein minimales System.
Offenbar folgt 2. aus 2.' Weiterhin ergibt sich für jedes $A \in \mathfrak{A}$ wegen $A \cup \bar{A} = \Omega$ und $A \cap \bar{A} = \emptyset$ die Gleichung

$$P(\bar{A}) = 1 - P(A),$$

und damit ist auch jede der beiden Forderungen in 1. eine Folgerung der anderen.

Weitere unmittelbare Folgerungen aus den Axiomen:

Wenn $A, B, A_1, A_2, \cdots \in \mathfrak{A}$, dann gilt:

1. $P(A) = P(A \cap B) + P(A \backslash B)$
   wegen $A = (A \cap B) \cup (A \cap \bar{B})$ und Axiom 2.

2. $B \subseteq A \Longrightarrow P(B) \leq P(B) + P(A \backslash B) = P(A)$ (Monotonie der Verteilung P) wegen Folgerung 1.

3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
   wegen

$$\begin{aligned} A \cup B &= \Big((A \cap \bar{B}) \cup (A \cap B)\Big) \cup \Big((B \cap \bar{A}) \cup (B \cap A)\Big) \\ &= (A \backslash B) \cup (A \cap B) \cup B \backslash A; \end{aligned}$$

   Die drei Ereignisse der rechten Seite sind paarweise disjunkt, unter Verwendung von Folgerung 1. folgt die Behauptung.

4. $P(A \cup B) \leq P(A) + P(B)$ (Subadditivität)
   $P\Big(\bigcup_1^n A_k) \leq \sum_1^n P(A_k)$ (endliche Subadditivität)

   Das ergibt sich aus Folgerung 3 bzw. mittels vollständiger Induktion.

   Das Axiom 2.' ermöglicht es, die Wahrscheinlichkeiten von Ereignissen zu bestimmen, die im Zusammenhang mit unendlichen Folgen von Ereignissen stehen. Um dies zu illustrieren, beginnen wir mit zwei zu 2.' äquivalenten Eigenschaften.

Lemma: *(σ-Stetigkeit von Wahrscheinlichkeitsverteilungen)*

Wenn für die Abbildung P von $\mathfrak{A}$ in $[0,1]$ nur die Axiome 1. und 2. gelten, so ist 2.' äquivalent mit jeder der folgenden Aussagen:

a) Für jede monoton fallende Folge $(A_n, n \geq 1)$ aus $\mathfrak{A}$ mit $\bigcap_{n=1}^{\infty} A_n = \emptyset$ gilt $\lim_{n\to\infty} P(A_n) = 0$

b) Für jede monoton wachsende Folge $(A_n, n \geq 1)$ aus $\mathfrak{A}$ mit $\bigcup_{n=1}^{\infty} A_n = \Omega$ gilt $\lim_{n\to\infty} P(A_n) = 1$

Beweis:

2.' $\Longrightarrow$ a): Mit $B_n = A_n \backslash A_{n+1},\ n \geq 1$ ist $(B_n, n \geq 1)$ nach Voraussetzung eine Folge paarweise disjunkter Ereignisse mit $\bigcup_{n=1}^{\infty} B_n = A_1$. Folglich gilt mit Axiom 2' die Gleichung $P(A_1) = \sum_{n=1}^{\infty} P(B_n)$ und $P(A_m) = \sum_{n=m}^{\infty} P(B_n)$.
Somit haben wir $\lim_{m\to\infty} P(A_m) = 0$.

a) $\Longrightarrow$ b): $(\overline{A}_n, n \geq 1)$ ist monoton fallend mit $\bigcap_{n=1}^{\infty} \overline{A}_n = \emptyset$, somit gilt $\lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} (1 - P(\overline{A}_n)) = 1 - 0 = 1$.

b) $\Longrightarrow$ 2.': Ist $(C_n, n \geq 1)$ eine Folge paarweise disjunkter Ereignisse, so definieren wir $C'_n = \bigcup_{m=1}^{n} C_m, C' = \bigcup_{m=1}^{\infty} C_m, A_n = C'_n \cup \overline{C'}, n \geq 1$. Damit folgt $A_n \subseteq A_{n+1}, n \geq 1$ und $\bigcup_{n=1}^{\infty} A_n = \Omega$, und somit nach Voraussetzung $\lim_{n\to\infty} P(A_n) = 1$. Wegen Axiom 2. ergibt sich $1 = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} P(C'_n) + P(\overline{C'})$, also
$P(C') = \lim_{n\to\infty} P(C'_n) = \lim_{n\to\infty} \sum_{m=1}^{n} P(C_m) = \sum_{m=1}^{\infty} P(C_m)$.

Folgerungen aus dem Lemma:

a) Ist (A_n) eine monoton fallende (monoton wachsende) Folge aus $\mathfrak{A}$, so gilt

$\lim_{n\to\infty} P(A_n) = P\Big(\bigcap_{n=1}^{\infty} A_n\Big)$ bzw.

$\lim_{n\to\infty} P(A_n) = P\Big(\bigcup_{n=1}^{\infty} A_n)\Big)$.

Zum Beweis: Man wende das vorausgehende Lemma auf $\Big(A_n \setminus \Big(\bigcap_{k=1}^{\infty} A_k\Big)\Big)$ bzw. auf $\Big(A_n \cup \Big(\overline{\bigcup_{k=1}^{\infty} A_k}\Big)\Big)$ an.

b) Für jede Folge (A_k) aus $\mathfrak{A}$ gilt:

$$P\Big(\bigcup_{1}^{\infty} A_k\Big) \overset{\leq}{=} \sum_{k=1}^{\infty} P(A_k) \text{ (abzählbare Subadditivität)}$$

Beweis: $B_n = \bigcup_{1}^{n} A_k \uparrow B = \bigcup_{1}^{\infty} A_k$, beachte Folgerung 4.

c) <u>1. Lemma von Borel-Cantelli:</u>

Falls $A_n \in \mathfrak{A}, n \geq 1, und \quad \sum_{n=1}^{\infty} P(A_n) < \infty$, so gilt

$$P\Big(\lim_{n\to\infty} \sup A_n\Big) = 0$$

Beweis: $P\Big(\bigcap_{n=1}^{\infty} \bigcup_{m\geq n} A_m\Big) = \lim_{n\to\infty} P\Big(\bigcup_{m\geq n} A_m\Big)$

$$\leq \lim_{n\to\infty} \sum_{m=n}^{\infty} P(A_m) = 0.$$

In Worten: Gilt $\sum_{n=1}^{\infty} P(A_n) < \infty$, so ist die Wahrscheinlichkeit dafür, dass unendlich viele der Ereignisse A_n eintreten, gleich Null. Anders ausgedrückt, mit Wahrscheinlichkeit Eins treten höchstens endlich viele der Ereignisse A ein.

d) Wir geben noch eine nützliche Formel zur Berechnung von $P(\bigcup_{k=1}^{\infty} A_k)$ an, bei der die $(A_k, k = 1, \cdots, n)$ nicht paarweise disjunkt zu sein brauchen. Sie ist eine Verallgemeinerung der Folgerung 3.

(Ein- und Ausschlussformel)

$$P\Big(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k) - \sum_{1\le i<k\le n} P(A_i \cap A_k)$$
$$+ \sum_{1\le i<j<k\le n} P(A_i \cap A_j \cap A_k) - \ldots (-1)^{n+1} P(A_1 \cap \cdots \cap A_n)$$
$$= \sum_{r=1}^{n} (-1)^{r+1} \sum_{\substack{J_r \subseteq \{1,\ldots,n\} \\ card J_r = r}} P(A_{k_1} \cap \ldots \cap A_{k_r})$$

wobei $\mathfrak{J}_r$ alle r-elementigen Teilmengen von $\{1, \cdots, n\}$ durchläuft.

Beweis mittels vollst. Induktion, siehe z. B. Henze (2003), Kap. 11.

Die Ein- und Ausschlussformel vereinfacht sich wesentlich, falls die Wahrscheinlichkeiten $P(A_{k_1} \cap \cdots \cap A_{k_r})$ nur von r und nicht von der Wahl des Tupels $(k_1, \cdots, k_r)$ abhängen. Wir definieren:

Definition: *Es seien $(\Omega, \mathfrak{A}, P)$ ein W-Raum und $A_1, \cdots, A_n$ Ereignisse aus $\mathfrak{A}$. Diese Ereignisse heißen (untereinander) austauschbar, falls $P(A_{k_1} \cap \ldots \cap A_{k_r}) = P(A_1 \cap \ldots \cap A_r)$ gilt für alle r mit $1 \le r \le n$ und alle r-elementigen Teilmengen $\{k_1, \cdots, k_r\}$ von $\{1, \cdots, n\}$.*

Aussage: *Sind $A_1, \cdots, A_n$ austauschbar, so gilt*

$$P(\bigcup_{k=1}^{n} A_k) = \sum_{r=1}^{n} (-1)^{r+1} \binom{n}{r} P(A_1 \cap \cdots \cap A_r).$$

Beweis: Es gibt $\binom{n}{r}$ Teilmengen $\mathscr{J}_r$ von $\{1, \cdots, n\}$ mit r Elementen.

e) Bonferroni-Ungleichungen:

Falls $A_i \in \mathfrak{A}, i = 1, \cdots, n$, dann gilt

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) \le \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$
$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$

Beweis: mittels vollständiger Induktion.

f) Ebenfalls mittels vollständiger Induktion zeigt man, dass für $A_1, A_2, \cdots, A_n \in \mathfrak{A}$ und

$A_{n,m} :=$ "Es treten genau m der n Ereignisse $A_1, A_2, \cdots, A_n$ ein" gilt:

$$P(A_{n,m}) = \sum_{k=m}^{n} (-1)^{r-m} \binom{r}{m} \sum_{\substack{\mathfrak{A}_r \subseteq \{1,\cdots,n\} \\ card J_r = r}} P(A_{K_1} \cap \cdots \cap A_{k_r})$$

3.2 Laplace-Experimente

Eine mathematisch übersichtliche Klasse zufälliger Versuche bilden die Laplace-Experimente.

Definition: *Als Laplace-Experiment (L-Experiment) bezeichnet man einen zufälligen Versuch mit:*

1. *der Versuch hat nur endlich viele* $(= N)$ *mögliche Ausgänge.*
2. *Alle Ausgänge haben die gleiche Wahrscheinlichkeit.*

Als mathematisches Modell eines L-Experimentes wählt man einen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ *mit:*

1. Ω *ist endlich:* $\Omega = \{1, 2, \ldots, N\}$

2. $\mathfrak{A} = \mathfrak{P}(\Omega)$ *und alle Versuchsausgänge sind gleichwahrscheinlich:*

$$P(\{\omega\}) \equiv: p \ , \ \omega \in \Omega.$$

N heißt der Parameter des L-Experimentes.

Folgerungen: a) $P(\{\omega\}) = p = \dfrac{1}{N}$, denn

$$1 = P(\Omega) = P(\bigcup_{k=1}^{N} \{\omega_k\}) = \sum_{k=1}^{N} P(\{\omega_k\}) = N \cdot p$$

b) Für jede Teilmenge A von Ω gilt:

$$\begin{aligned} P(A) = \quad & P(\bigcup_{\omega \in A} \{\omega\}) = \sum_{\omega \in A} P(\{\omega\}) = \sharp A \cdot p \\ & = \frac{\text{Anzahl der Elemente von} A}{\text{Anzahl der Elemente von} \Omega} \\ & = \frac{\text{Anzahl der (für A) günstigen Fälle}}{\text{Anzahl der möglichen Fälle}} \\ & = \frac{N(A)}{N} \end{aligned}$$

Die Gesamtwahrscheinlichkeit 1 (d.h., die Wahrscheinlichkeit des sicheren Ereignisses "Irgendein Versuchsausgang tritt ein") ist bei einem Laplace-Experiment gleichmäßig auf die Versuchsgänge ω verteilt: Man nennt P auch die *gleichmäßige Wahrscheinlichkeitsverteilung* auf $\{1, 2, \ldots, N\}$ oder einfach die *Gleichverteilung* auf $\{1, 2, \ldots, N\}$. Die Gleichverteilung auf $\{1, 2, \cdots, N\}$ ist ein Spezialfall sogenannter diskreter Wahrscheinlichkeitsverteilungen, die wir in Kapitel 4 behandeln.

Bei Laplace-Experimenten spricht man auch davon, dass das Versuchsergebnis "auf gut Glück" oder "rein zufällig" ausgewählt wird, um die Gleichwahrscheinlichkeit aller möglichen Versuchsausgänge hervorzuheben.

Fazit: Bei L-Experimenten lassen sich die Wahrscheinlichkeiten aller mit dem Versuch verbundener Ereignisse durch Abzählen der für diese Ereignisse günstigen Ausgänge bestimmen.

Beispiel: Der zufällige Versuch bestehe im Werfen zweier regulärer Würfel und im Registrieren, welche Augenzahl der erste und welche der zweite Würfel zeigt. $\Omega = \{(i,j) : i,j \in \{1,2,\ldots,6\}\}$.
i = Augenzahl des ersten Würfels,
j = Augenzahl des zweiten Würfels.

Dabei wirft man die Würfel nacheinander oder verwendet verschiedenfarbige Würfel, um zu wissen, welches Ergebnis zum ersten, welches zum zweiten gehört. Alle Ergebnisse sind aus Symmetriegründen gleichwahrscheinlich, also gilt:

$P((i,j) \text{ tritt auf }) = \frac{1}{36} \quad \text{für alle} \quad i,j \in \{1,\cdots,6\}.$

$A :=$ "Augensumme gleich 6"

$P(A) = \frac{N(A)}{N} = \frac{5}{36}$,

$B :=$ "Die Augenzahlen sind verschieden"

$P(B) = \frac{N(B)}{N} = \frac{30}{36} = \frac{5}{6}$

Problem:
In Anwendungsbeispielen mit endlichem Ω muß man genau prüfen, ob es sich tatsächlich um ein Laplace-Experiment handelt:

Beobachtet man im obigen Beispiel nur die Augensumme, so ist dies ein neuer zufälliger Versuch. Man wählt $\Omega = \{2,\ldots,12\}$. Jetzt sind aber nicht alle Ausgänge gleichberechtigt, d.h. gleichwahrscheinlich:
"Augensumme = 2" hat eine kleinere Wahrscheinlichkeit ($= \frac{1}{36}$) als "Augensumme = 4" ($= \frac{1}{12}$), denn das erste Ereignis tritt nur beim Versuchsausgang (1,1), das zweite dagegen bei jedem der Ausgänge (1,3), (2,2), (3,1) ein.

Wir kehren zurück zum Modell des n-maligen Münzenwurfes aus Abschnitt 2.4.

3.3 Münzenwurf, zum Zweiten

Die Münze, die wir für das Spiel verwenden, sei symmetrisch. Das bedeute, beide Seiten erscheinen bei einem Wurf mit gleicher Chance. Dann ist das n-malige Werfen ein Laplaceexperiment mit 2^n gleichwahrscheinlichen Ausgängen. Der entsprechende Wahrscheinlichkeitsraum ist $(\Omega, \mathfrak{P}(\Omega), P)$ mit $P(\{\omega\}) = 2^{-n}, \omega \in \Omega$ und $P(A) = \sum_{\omega \in A} 2^{-n} = \frac{\#A}{2^n}, A \subseteq \Omega$.

Wir berechnen die Wahrscheinlichkeit dafür, dass nach diesen n Würfen der Pfad $(k, S_k(\omega)), k \geq 0$, mit $S_0(\omega) := 0$ bei r endet: $P(\{\omega \in \Omega | S_n(\omega) = r\})$.

Aussage: *Für $P(S_n = r)$ gelten die Formeln*

a) Ist n gerade, also $n = 2m$ für ein $m \geq 1$, so ist

$P(S_{2m} = 2l) = \binom{2m}{m+l} 2^{-2m}$ *falls* $|l| \leq m$,

$P(S_{2m} = r) = 0$ *für alle anderen ganzzahligen* r.

b) Ist n ungerade, also $n = 2m + 1$ für ein $m \geq 1$, so ist

$P(S_{2m+1} = 2l + 1) = \binom{2m+1}{m+l+1} 2^{-2m-1}$ *falls* $-m - 1 \leq l \leq m$,
$P(S_{2m+1} = r) = 0$ *für alle anderen ganzzahligen* r.

Bemerkung: In beiden Fällen handelt es sich um eine um den Nullpunkt symmetrische Verteilung mit Null bzw. ± 1 als Punkte maximaler Wahrscheinlichkeit, vgl. Abschnitt 4.

Beweis:

a) $\{S_{2m} = 2l\}$ tritt genau dann ein, wenn in $\omega = (x_1, \cdots, x_{2m})$ genau $l + m$ mal die Eins enthalten ist.

b) $\{S_{2m+1} = 2l + 1\}$ tritt genau dann ein, wenn $\omega = (x_1, \cdots, x_{2m+1})$ genau $m + l + 1$ Einsen enthält.

Für den Spieler A (siehe Abschnitt 2.4) ist es von Interesse, wann er zum ersten Mal ein negatives Guthaben hat.

Aussage: *Die Wahrscheinlichkeit dafür, dass* $S_1 \geq 0, S_2 \geq 0, \cdots, S_{n-1} \geq 0$ *und* $S_n = -1$ *gilt, ist für ungerades* $n = 2m+1, m \geq 0$ *gleich dem Wert*

$$\frac{(2m-1)!!}{(m+1)!2^{m+1}}$$

mit $(2m-1)!! = (2m-1)(2m-3)\cdots 3 \cdot 1$
und für gerades n *gleich Null.*

Beweis: Für jeden der für das betrachtete Ereignis günstigen Pfade gilt $S_{n-1} = 0$. Wir bestimmen deshalb die Zahl aller Pfade von $(0,0)$ nach $(2m,0)$ die -1 niemals berühren.

Es gibt insgesamt $\binom{2m}{m}$ Pfade von $(0,0)$ nach $(2m,0)$. Wir setzen $n = 2m$. Zur Berechnung der gesuchten Zahl der Pfade bedienen wir uns des sogenannten *Spiegelungsprinzips*. Jedem durch ein ω erzeugten Pfad $(s_0, s_1, \cdots, s_n)$, der die Zahl -1 jemals vor n erreicht, wird der Pfad ω' zugeordnet, der bei -2 startet und bis zur Zeit $T_{-1}(\omega) = \min\{k \geq 1 | s_k = -1\}$ spiegelbildlich bezüglich der Horizontalen der Höhe -1 zu $(s_1, s_2, \cdots s_n)$ verläuft, sowie danach mit $(s_1, s_2, \cdots, s_n)$ übereinstimmt.
Die Zuordnung ist eineindeutig. Folglich ist die Zahl der Pfade, die -1 vor der Zeit n berühren und zur Zeit $n-1$ bei Null sind gleich der Zahl der Pfade, die bei -2 starten und zur Zeit $n-1$ in Null enden.

Davon gibt es $\binom{2m}{m+1}$ Exemplare. Somit ist die gesuchte Anzahl gleich

$$\binom{2m}{m} - \binom{2m}{m+1} = \binom{2m}{m}\frac{1}{m+1}.$$

Die gesuchte Wahrscheinlichkeit ist also $2^{-2m-1}\binom{2m}{m}\frac{1}{m+1}$. Eine einfache Umformung liefert das zu beweisende Resultat.

Zur Vereinfachung der Bezeichnung führen wir folgende Notation ein, wobei $\omega = (x_1, x_2, \cdots, x_n) \in \Omega$ gelte:

$$X_k(\omega) := x_k \quad , \quad k = 1, 2, \cdots, n,$$

$$S_m(\omega) := \sum_{k=1}^{m} X_k(\omega) = \sum_{k=1}^{m} x_k, m = 1, 2, \cdots, n.$$

Die "Zufallsgröße" X_k gibt das Ergebnis des k-ten Wurfes an, die "Zufallsgröße" $S_m(\omega)$ ist der Gewinn des Spielers A nach m Würfen, $m = 1, 2, \cdots, n$.

Für die Ereignisse A und B gilt dann

$A = \{\omega \in \Omega | S_n(\omega) > 0\}$, oder abgekürzt

$A = \{S_n > 0\}$ und

$B = \{\omega \in \Omega | \min_{k=1,\cdots,n} S_k(\omega) \geq 0\}$, oder kurz

$B = \{\min_{k=1,\cdots,n} S_k \geq 0\}$.

3.4 Was sagen Wahrscheinlichkeiten?

Welche anschauliche Bedeutung hat die Wahrscheinlichkeit $P(A)$ eines zufälligen Ereignisses A? Es gibt zwei grundlegende Erfahrungen im Umgang mit dem Zufall.

a) Empirisches Gesetz der großen Zahlen: In einer langen Reihe gleichartiger, voneinander unabhängiger Versuche ist die relative Häufigkeit $\frac{n(A)}{n}$ des Eintretens von A etwa gleich $P(A)$. Wenn $P(A) > \frac{1}{2}$ gilt, so kann man auf das Eintreten von A Wetten abschließen und wird bei fortlaufenden Wetten dieser Art schließlich im Vorteil sein.

b) Es ist neben dem empirischen Gesetz der großen Zahlen eine zweite Erfahrungstatsache, dass zufällige Ereignisse mit sehr kleinen Wahrscheinlichkeiten bei einmaliger Versuchsdurchführung praktisch nicht eintreten.

Genauer gesagt: Man muss bei einem zufälligen Versuch mit dem Eintreten von A nicht rechnen, falls $P(A)$ sehr klein ist. Diese Erfahrung hat jeder Mensch verinnerlicht.

(Auf diese Erfahrungen weist z. B. Kolmogorov (1933) hin.)

Beispiel für a):

1) Werfen Sie einen regulären Spielfwürfel mehrere Mal und beobachten Sie das Verhalten der relativen Häufigkeit des Auftretens einer "Sechs" im Verlaufe der Würfe. Sie tendiert zu $\frac{1}{6}$.

Beispiele für b):

1) Man erhält keinen Kredit von der Bank, wenn man als Sicherheit anbietet, dass man auf seinen wöchentlichen Tippschein im Lotto "6 aus 49" innerhalb eines Jahres einen "Sechser" erzielt. Die Wahrscheinlichkeit dieses Ereignisses ist so gering, dass man mit seinem Eintreten nicht wirklich rechnet.

2) Man rechnet nicht damit, dass jemand durch maximal dreimaliges zufälliges Raten der PIN bei einer EC-Karte die richtige PIN errät.

   $P(A_1 \cup A_2 \cup A_3) \approx \frac{3}{10^4} = 0,0003$

3) Wenn ein vorgegebenes Ereignis, das eine sehr kleine Wahrscheinlichkeit hat, tatsächlich eintritt, so zweifelt man mitunter daran, dass der "reine Zufall" zu diesem Ereignis geführt hat. Man stellt eher die Richtigkeit der zugrunde liegenden Annahmen in Frage und prüft sie sorgfältig: Den Ausspruch "Das kann kein Zufall sein!" hat jeder schon mal gehört.

   Beispiel seltener Ziehungen beim Lotto wie (1, 2, 3, 4, 5, 6) führen regelmäßig zur Aufmerksamkeit der Medien und der Frage, ob hier nicht der Zufall außer Kraft gesetzt sei. Diese Zahlenkombination hat aber die gleiche Wahrscheinlichkeit wie jede andere.

4) Aus Dorothy L. Sayers "Keines natürlichen Todes", rororo, 1991:

   a) S. 58_8: Ein merkwürdiger Zufall, sagte er (der Chef von Scotland Yard) geduldig, und ich kann verstehen, dass Sie sich darüber aufregen.

   b) S. 62_{14}: Schon wieder ein Reinfall sagte Winsey: "Aber ein sonderbarer Zufall ist das schon."

5) Wenn jemand beim Skatspiel dreimal hintereinander alle vier Buben erhält, glaubt man nicht mehr an reinen Zufall, obwohl dieses Ereignis eine positive Wahrscheinlichkeit hat.

3.5 Elemente der Kombinatorik*

Bei der Abzählung der "günstigen" Fälle bei L-Experimenten erweisen sich Formeln der Kombinatorik häufig als günstig.
Wir geben hier vier Grundaufgaben der Kombinatorik an, sie werden häufig auch in Form sogenannter Urnenprobleme formuliert.

Wir beginnen mit einer elementaren aber wichtigen Feststellung.

Aussage: *Es seien $M_1, \ldots, M_m$ m Mengen mit $m_1, \ldots, m_m$ Elementen. Dann hat die Menge M aller m-Tupel $(i_1, \ldots, i_m)$ mit $i_k \in M_k$ für $k = 1, \ldots, m$ genau $m_1, m_2, \ldots, m_m$ Elemente.*

Beweis: Mittels vollständiger Induktion

Als Nächstes kommen wir zu den vier angekündigten Aufgaben der Kombinatorik.
Aus einer Menge $M = \{a_1, a_2, \ldots, a_m\}$ von m Elementen ($m \geq 1$) werden r Elemente ausgewählt, $r \geq 1$. Man spricht von einer Stichprobe vom Umfang r aus der Menge M. Die Entnahme von Stichproben kann auf unterschiedliche Weise erfolgen:
Mit Wiederholung oder ohne Wiederholung
Mit Berücksichtigung der Reihenfolge oder ohne Berücksichtigung der Reihenfolge.
(d. h. geordnet oder ungeordnet)

Dementsprechend unterscheiden wir vier Fälle.

Anzahl von möglichen Stichproben des Umfanges r aus einer Menge M vom Umfang m

(In den Fällen ohne Wiederholung ist $r \leq m$ vorauszusetzen.)

	mit Wiederholung	ohne Wiederholung
geordnet	r-Tupel $(a_1, \ldots, a_r)$ mit $a_i \in M, i = 1, \ldots, r$ r-Permutation mit W, m^r **A1**	r-Tupel $(a_1, \ldots, a_r)$ mit $a_i \in M, i \in 1, \ldots, r)$, paarw. verschieden r-Permutation ohne W. $m(m-1)\ldots(m-r+1) =: (m)_r$ **A2**
ungeordnet	$[a_1, a_2, \ldots, a_r]$, Anordnung von r Elementen $\{a_1, \ldots, a_r\} \subset M$ $a_i \in M, i = 1, \ldots, r$ r-Kombination mit W. $\binom{m+r-1}{r}$ **A3**	Teilmenge vom Umfang r r-Kombination ohne W. $\binom{m}{r} = \frac{(m)_r}{r!}$ **A4**

Die Fälle A1, A2 und A4 sind leicht zu beweisen.

Der Fall A3:
Jede ungeordnete Stichprobe vom Umfang r mit Wiederholung aus der Menge M ist eindeutig charakterisiert durch eine Folge $(i_1, i_2, \ldots, i_m)$ natürlicher Zahlen $i_k \geq 0$ mit $\sum_{k=1}^{m} i_k = r$, wobei i_k angibt, wie oft das Element a_k aus M in der Stichprobe vorkommt.

Diese Vektoren $(i_1, \ldots, i_m)$ lassen sich eineindeutig auf die Menge aller Anordnungen der Form $\bullet\bullet\bullet|\bullet\bullet||\bullet\bullet\bullet$ von r Punkten und $(m-1)$ Strichen abbilden, wobei vor dem ersten Strich i_1 Punkte stehen, zwischen dem k-ten und $(k+1)$-ten Strich i_{k+1} Punkte stehen, und nach dem $(m-1)$-ten Strich i_m Punkte platziert sind. Insgesamt gibt es $\binom{m+r-1}{r}$ solcher Anordnungen.

Zu jedem der vier Fälle der Entnahme von Stichproben vom Umfang r aus einer Menge vom Umfang m gibt es ein sogenanntes "duales Problem" der Verteilung von r Kugeln auf m Urnen.

A_1: Duales Problem: r unterscheidbare Kugeln werden auf m Urnen verteilt, wobei in jeder Urne beliebig viele Kugeln liegen dürfen. Auf wieviel verschiedene Weisen ist dies möglich? Jede Verteilung der Kugeln ist charakterisiert durch die r Nummern $(a_1, a_2, \ldots, a_r)$ der Urnen, in die die erste zweite, ..., r-te Kugel zu liegen kommt. Für jede dieser Nummern gibt es m Möglichkeiten. Im Ergebnis entsteht wieder eine Stichprobe $(a_1, a_2, \ldots, a_r)$ vom Umfang r mit Wiederholung aus einer Menge vom Umfang m.

A_2: Duales Problem: r unterscheidbare Kugeln werden auf m Urnen verteilt, wobei in jeder Urne höchstens eine Kugel liegen darf. Auf wieviel verschiedene Weisen ist dies möglich? Jede Verteilung der Kugel ist charakterisiert durch die r Nummern $(a_1, a_2, \ldots, a_r)$ der Urnen, in die die erste, zweite k-te, r-te Kugel zu liegen kommt. Für die erste Kugel gibt es m, für die zweite $m-1$, für die r-te Kugel $m-r+1$ Möglichkeiten. Im Ergebnis entsteht wieder eine Stichprobe $(a_1, a_2, \ldots, a_r)$ vom Umfang r ohne Wiederholung aus einer Menge vom Umfang m.

A_3: Duales Problem: r ununterscheidbare Kugeln werden auf m Urnen aufgeteilt, wobei jede Urne auch mehrfach besetzt werden kann.
Jede Aufteilung ist charakterisiert durch die Anzahl i_k der Kugeln, die in die k-te Urne fallen, $k = 1, \ldots, m, \sum_{k=1}^{m} i_k = r$.

A_4: Duales Problem: r ununterscheidbare Kugeln sind auf m Urnen so aufzuteilen, daß in jeder Urne höchstens eine Kugel zu liegen kommt. Die Aufteilung ist charakterisiert durch die Menge $\{a_1, a_2, \ldots, a_r\}$ der Urnen, die durch eine Kugel besetzt werden, also durch eine r-elementige Teilmenge von M.

Anzahl der Aufteilungen von r Kugeln auf m Urnen

	mit Mehrfach- -besetzung	ohne Mehrfach- -besetzung
unterscheidbare Kugeln	m^r **A1**	$m(m-1)\dots(m-r+1)$ **A2**
ununterscheidbare Kugeln	$\binom{m+r-1}{r}$ **A3**	$\binom{m}{r}$ **A4**

Beispiele:

1. Wie groß ist die Wahrscheinlichkeit, beim Werfen von vier Würfeln mindestens eine "Sechs" zu erzielen?
   Der zufällige Versuch "Werfen von vier Würfeln" ist ein Laplace-Experiment mit 6^4 möglichen Ausgängen. Es bezeichne A das Ereignis "Es erscheint mindestens eine Sechs". Dann gibt es 5^4 günstige Ausgänge für das komplimentäre Ereignis $\overline{A}$ = "Es erscheint keine Sechs". Also gilt

   $$P(A) = 1 - \left(\tfrac{5}{6}\right)^4 = 0,52.$$

2. Es werden k Kugeln auf n Urnen aufgeteilt, $k \leq n$. Jede Kugel habe die gleiche Wahrscheinlichkeit in jede Urne zu gelangen. Wie groß ist die Wahrscheinlichkeit für das Ereignis A, daß es nach der Aufteilung Urnen gibt, in der mehr als eine Kugel liegt?

   Lösung: Es gibt n^k Möglichkeiten der geschilderten Aufteilung und $(n)_k$ Möglichkeiten, die günstig sind für das komplementäre Ereignis $\overline{A}$ = "In keiner Urne liegt mehr als eine Kugel". Daraus folgt

   $$P(A) = 1 - \tfrac{(n)_k}{n^k}.$$

3. In einem Raum mögen sich k Personen befinden. Wie groß ist die Wahrscheinlichkeit für das Ereignis A, dass mindestens zwei dieser Personen am gleichen Tag Geburtstag haben? (Jeder Tag des Jahres komme bei jeder Person mit gleicher Wahrscheinlichkeit als Geburtstag in Frage, Schaltjahre bleiben unberücksichtigt.)

Lösung: $\overline{A}$ = "Alle k Personen haben an verschiedenen Tagen Geburtstag". Es gibt $N = (365)^k$ möglich Fälle für Geburtstage und $N(\overline{A}) = (365)_k$ für $\overline{A}$ günstige Fälle, somit ist $P(A) = 1 - \frac{(365)_k}{365^k}$.

Diese Wahrscheinlichkeit wächst mit k und ist gleich $0,507$ für $k = 23$.

4. Koinzidenzproblem:
n Briefe werden auf rein zufällige Weise in n adressierte Umschläge gesteckt. Wie groß ist die Wahrscheinlichkeit, dass mindestens ein Brief in den richtigen Umschlag kommt?

Lösung: Mögliche Versuchsausgänge $\omega = (a_1, \ldots, a_n)$ sind die Permutation von $(1, \ldots, n)$ mit a_k gleich der Nummer des Umschlages, in den der k-te Brief kommt. Wir setzen

$$A_k := \{\omega | a_k = k\}, k = 1, \ldots, n.$$

Das interessierende Ereignis A ist gleich $\bigcup_{k=1}^{n} A_k$. Zur Anwendung der Ein- und Ausschlussformel berechnen wir

$$P(A_{k_1} \cap \ldots \cap A_{k_r}) = P(\{\omega = (a_1, \ldots, a_n) | a_{k_1} = k_1, \ldots, a_{k_r} = k_r\}) =$$

$$\frac{card\{\omega : a_{k_1} = k_1, \ldots, a_{k_r} = k_r\}}{n!} = \frac{(n-r)!}{n!}.$$

Die Ein- und Ausschlussformel liefert

$$P(A) = \sum_{r=1}^{n} (-1)^{r+1} \binom{n}{r} \frac{(n-r)!}{n!} = \sum_{r=1}^{n} (-1)^{r+1} \frac{1}{r!} \underset{n\to\infty}{\longrightarrow} 1 - \frac{1}{\mathrm{e}} = 0,632$$

Für $n \geq 7$ ist die Näherung auf drei Stellen genau. (Weitere Bemerkungen zum Koinzidenzproblem und anderen kombinatorischen Aufgaben findet man in Henze, Kap. 9-11.)

5. Am Eröffnungstag eines großen Kongresses mit 500 Teilnehmern soll jeder teilnehmenden Person, die an diesem Tag Geburtstag hat, ein Blumenstrauß überreicht werden. Wieviele Sträuße braucht man mindestens, wenn man mit Sicherheit auschließen will, dass man zu wenige Sträuße hat?
   Wie groß muss die Zahl der Sträuße mindestens sein, wenn man mit der Wahrscheinlichkeit von 0,95 diesen blamablen Fall vermeiden will?
   Wir nehmen näherungsweise an, dass für jede Person die Wahrscheinlichkeit, an einem bestimmten Tag Geburtstag zu haben, gleich ist für alle Tage des Jahres. Schaltjahre werden nicht berücksichtigt. Dann ist die Feststellung der Geburtstage aller Teilnehmer ein Laplace-Experiment mit den möglichen Ausgängen $\omega = (i_1, \ldots, i_{500})$, wobei i_k die Nummer des Tages angibt, an denen der k-te Teilnehmer Geburtstag hat. Die Menge aller möglichen Versuchsausgänge hat den Umfang $N = 365^{500}$. Es gibt nämlich $N = 365^{500}$ Möglichkeiten der Verteilung der Geburtstage der 500 Personen auf das Jahr.

   Für das Ereignis $A_k :=$ "Genau k Personen haben am Eröffnungstag Geburtstag" gibt es $N(A_k) = \binom{500}{k} \cdot 364^{500-k}$ "günstige" Versuchsausgänge.

   Es gilt $P(A_k) = \frac{N(A_k)}{N}$, und folglich ergibt sich

k	0	1	2	3	4	5
$P(A_k)$	0,2532	0,3478	0,2384	0,1087	0,0371	0,0101

   Deshalb ist

$$P\Big(\bigcup_{k=1}^{5} A_k\Big) = \sum_{k=1}^{5} P(A_k) = 0,9953.$$

   Das heißt, mit an Eins grenzender Wahrscheinlichkeit haben höchstens fünf der Personen am Eröffnungstag Geburtstag.
   (Zur Berechnung wurde die Näherung

$$P(A_k) \approx \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} \text{ mit } \lambda = \frac{500}{365}$$

benutzt.)

Rein zufällige Wahl eines Punktes aus $[0, 1)$

Das Laplace-Experiment lässt sich nicht unmittelbar auf das in der Überschrift genannte Problem anwenden, da $[0, 1)$ unendlich viele Punkte enthält. Wir müssen hier den Begriff der "rein zufälligen Wahl" etwas modifizieren.

Rein zufällige Wahl soll bedeuten, dass für jedes Intervall $[a, b) \subseteq [0, 1)$ die Wahrscheinlichkeit, dass der gewählte Punkt aus $[a, b)$ stammt, unabhängig von der Lage des Intervalls sein soll. Das heißt

$$P([a, b)) = P([a + x, b + x)) \tag{1}$$

für alle x mit $a + x \geq 0, b + x \leq 1$.

Daraus folgt $P([a, b)) = b - a$. (2)

(Beweisen Sie (2)).

Es existiert allerdings kein Wahrscheinlichkeitsmaß P auf der Potenzmenge von $[0, 1)$ mit den Eigenschaften (1) und (2), siehe Elstrodt, III. 2. Man kann aber zeigen, dass ein Wahrscheinlichkeitsmaß P mit (1) und (2) existiert auf der kleinsten σ-Algebra $\mathscr{B}_{[0,1)}$ von Teilmengen von $[0, 1)$, die alle Intervalle der Form $[a, b)$ mit $0 \leq a < b \leq 1$ enthält (σ-Algebra der Borelmengen aus $[0, 1)$). Dieses Maß ist eindeutig bestimmt und heißt Lebesgue-Borel-Maß auf $([0, 1), \mathscr{B}_{[0,1)})$ oder einfach Lebesguemaß. Wir werden es mit $\lambda_{[0,1)}$ bezeichnen.

Die Tatsache, dass man $\lambda_{[0,1)}$ unter Beibehaltung von (1) nicht auf $\mathfrak{P}([0, 1))$ erweitern kann, führt zu der auf den ersten Blick eigenartigen Situation, dass man nicht jede Teilmenge C von $[0, 1)$ als zufälliges Ereignis bei der rein zufälligen Wahl eines Punktes aus $[0, 1)$ ansehen kann.

Der Wahrscheinlichkeitsraum $([0, 1), \mathscr{B}_{[0,1)}, \lambda_{[0,1)})$ ist das mathematische Modell des zufälligen Versuches, einen Punkt aus dem Intervall "rein zufällig"

oder "auf gut Glück" auszuwählen.

Die Wahrscheinlichkeitsverteilung $\lambda_{[0,1)}$ "verteilt" die Gesamtwahrscheinlichkeit Eins "gleichmäßig" auf das Intervall $[0,1)$. Sie heißt *gleichmäßige Verteilung* auf $[0,1)$. Insbesondere hat dann auch jeder Punkt $x \in [0,1)$ als Ereignis $\{x\}$ die gleiche Wahrscheinlichkeit, die folglich gleich Null sein muss. Das folgt auch aus $\{x\} = \bigcap_{k=1}^{\infty} \left[x, x + \frac{1}{k}\right)$ und der σ-Stetigkeit von $\lambda_{[0,1)}$. Diese Verteilung ist Beispiel einer stetigen Verteilung.

In dem eben eingeführten Wahrscheinlichkeitsraum gibt es Ereignisse, die nicht unmöglich (bzw. nicht sicher) sind, aber dennoch die Wahrscheinlichkeit Null (bzw. Eins) haben. Das führt uns auf folgende Definition.

Definition: *Es sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum. Jedes Ereignis $A \in \mathfrak{A}$ mit $P(A) = 1 (P(A) = 0)$ heißt fast sicheres Ereignis (bzw. fast unmögliches Ereignis).*

Bei der rein zufälligen Wahl eines Punktes aus $[0,1)$ ist das Ereignis $A :=$ "Es wird ein irrationaler Punkt gewählt" ein fast sicheres Ereignis und $\overline{A} =$ "Es wir ein rationaler Punkt gewählt" ein fast unmögliches Ereignis.

3.6 Zufallsgrößen

Unter einer Zufallsgröße versteht man umgangssprachlich eine Größe, die im Rahmen gewisser zufälliger Erscheinungen einen Wert annimmt, der nicht von vornherein feststeht, sondern vom Zufall abhängt. Beispiele findet man überall. In der Natur (Wetter), der Wirtschaft (Aktienkurse), der Technik (Ausfallzeitpunkte von Konsumgütern). Ihre mathematische Erfassung und Untersuchung ist ein zentraler Punkt der Wahrscheinlichkeitstheorie.

Im Allgemeinen sind zufällige Erscheinungen von sehr komplexer Natur. Man denke nur an das Wetter oder das Geschehen an einer Aktienbörse. Durch die Konzentration auf Zufallsgrößen, wie Tageshöchsttemperatur, monatliche Niederschlagsmenge bzw. Aktientagesschlusskurse oder wöchentliche Rendite bestimmter Unternehmen werden Teilaspekte der zugrunde liegenden zufälli-

gen Prozesse herausgestellt, für die man sich besonders interessiert.

Ein sehr einfaches Beispiel ist das Werfen zweier Würfel, wobei man sich nicht für den Ausgang des gesamten Experimentes interessiert, sondern nur für die Augensumme. Diese ist eine Zufallsgröße im Rahmen des Werfens dieser beiden Würfel. Offenbar ist ihr (vom Zufall abhängender) Wert eine Funktion des Ausganges des zugrunde liegenden Würfelexperimentes. Ausgehend von diesen Bemerkungen führen wir den mathematischen Begriff der Zufallsgröße gleich in voller Allgemeinheit ein. Das hat den Vorteil, dass seine wesentlichen Aspekte klar zum Ausdruck kommen.

Es seien $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum und X eine Abbildung von Ω in eine Menge E, z.B. $E = R_1$, $E = R_n$ oder $E = N_0 = \{0, 1, 2, \cdots, n, ...\}$. Indem man nicht den Versuchsausgang $\omega \in \Omega$ zur Kenntnis nimmt, sondern nur den Wert $X(\omega)$ beobachtet, den die Funktion X in Abhängigkeit von ω annimmt, ist ein neues zufälliges Experiment definiert mit möglichen Ausgängen x, die aus E stammen. Die mit diesem neuen Experiment verbundenen Ereignisse sind nunmehr Teilmengen von E. Sie bilden eine σ-Algebra $\mathfrak{E}$ von Teilmengen von E. Das Ereignis B aus $\mathfrak{E}$ tritt für den neuen Versuch offenbar genau dann ein, wenn der ursprüngliche Versuch zu einem ω führt, für das $X(\omega) \in B$ gilt.

Beispiel: Wir betrachten das Laplace-Experiment des Werfens zweier Würfel und wählen

$$\Omega = \{\omega = (i, j) : i, j \in \{1, 2, \ldots, 6\}\}, \quad \mathfrak{A} = \mathfrak{P}(\Omega),$$

$$P(\{\omega\}) = 36^{-1}, \quad \omega \in \Omega,$$

$$X(\omega) = i + j \quad , \quad \omega = (i, j), \in \Omega,$$

Hier wählt man den Bildraum E als die Menge $\{2, 3, \cdots, 12\}$ und für $\mathfrak{E}$ die Potenzmenge $\mathfrak{P}(E)$.

X ist also gleich der Augensumme der zwei Würfel.

Das Ereignis "Augensumme ist gleich 4" entspricht der Menge $\{4\}$ aus E und tritt genau dann ein, wenn ein $\omega = (i, j)$ mit $i + j = 4$ Ergebnis des Würfels ist, also wenn $(1, 3), (2, 2)$ oder $(3, 1)$ gewürfelt wurde.

Wir kehren zurück zum allgemeinen Fall und wollen auf $(E, \mathfrak{E})$ eine Wahrscheinlichkeitsverteilung P^X einführen, die den Ereignissen des neuen Versuches ihre Wahrscheinlichkeiten zuordnet. Das geschieht durch

$$P^X(B) := P(\{\omega \in \Omega | X(\omega) \in B\}) = P(X^{-1}(B)),\ E \in \mathfrak{E}.$$

Üblicherweise ist $\mathfrak{E}$ zusammen mit E von vornherein festgelegt. Damit die Definition von P^X dann sinnvoll ist, müssen wir eine Forderung an X stellen, die wir in der nächsten Definition formulieren.

Definition: *Die Abbildung X von $(\Omega, \mathfrak{A})$ in $(E, \mathfrak{E})$ heißt eine Zufallsgröße über $(\Omega, \mathfrak{A}, P)$ mit Werten in $(E, \mathfrak{E})$, falls gilt*

$$X^{-1}(B) := \{\omega \in \Omega | X(\omega) \in B\} \in \mathfrak{A}, \textit{ für alle } B \textit{ aus } \mathfrak{E} \tag{1}$$

m.a.W., falls die Abbildung X (in der Sprache der Maßtheorie)eine $\mathfrak{A} - \mathfrak{E}$-messbare Abbildung ist.

Für $\{\omega \in \Omega | X(\omega) \in B\}$ schreiben wir häufig kürzer $\{X \in B\}$ und statt $P^X(B)$ bzw. $P(X^{-1}(B))$ verwenden wir die Schreibweise $P(X \in B)$.
Die Eigenschaft (1) drücken wir mitunter kürzer auch in der Form

$$X^{-1}(\mathfrak{E}) \subseteq \mathfrak{A} \tag{1'}$$

aus.

Wir merken uns also: Zufallsgrößen über einen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten in $(E, \mathfrak{E})$ sind die $\mathfrak{A} - \mathfrak{E}$-messbaren Abbildungen von $(\Omega, \mathfrak{A})$ in $(E, \mathfrak{E})$.

Aussage: *Durch*

$$P^X(B) := P(X \in B), \qquad B \in \mathfrak{E}$$

ist auf $\mathfrak{E}$ eine Wahrscheinlichkeitsverteilung P^X gegeben. Die Verteilung P^X nennt man die Wahrscheinlichkeitsverteilung der Zufallsgröße X oder die durch X induzierte Wahrscheinlichkeitsverteilung.

Zum Beweis prüft man die Axiome 1.-2.' nach.

Fortsetzung des Beispiels:
Die möglichen Werte der Zufallsgröße X sind die Zahlen $2, 3, \cdots, 12$. Für die Wahrscheinlichkeitsverteilung P^X der Zufallsgröße X gilt mit

$$E = \{2, \ldots, 12\}, \quad \mathfrak{E} = \mathfrak{P}(E),$$

die Gleichung

$$P^X(\{x\}) = P(\{\omega \in \Omega | X(\omega) = x\}) =$$

$$P(\{(i,j) \in \Omega | i + j = x\}) = \tfrac{6-|7-x|}{36}, x \in E$$

und

$$P^X(B) = \sum_{x \in B} P^X(\{x\}), \quad B \in \mathfrak{E}. \tag{4}$$

Die Forderung $X^{-1}(\mathfrak{E}) \subseteq \mathfrak{A}$ ist in diesem Fall natürlich erfüllt.

Wir kehren zurück zum Beispiel c) aus Abschnitt 2.1 und ersetzen die Zahl 100 durch ein allgemeines $n \geq 1$.
Hier führen wir die Zufallsgrößen $X_k(\omega) = +1(-1)$ ein, falls $\omega_k = Z(= W)$ wobei $\omega = (\omega_1, \omega_2, \cdots, \omega_n)$ sei. Nach Definition sind die möglichen Werte von X_k die Zahlen $+1$ und -1, und die Verteilung P^{X_k} ist gegeben durch $P^{X_k}(\{i\}) = \frac{1}{2}, i = +1, -1$.

Vermittels der X_k bilden wir die Summe

$$S_n(\omega) = \sum_{k=1}^{n} X_k(\omega), \omega \in \Omega,$$

die wiederum eine Zufallsgröße ist. Sie hat die möglichen Werte $-n, -n + 2m, \cdots, +n$ und für ihre Wahrscheinlichkeitsverteilung gilt

$P^X(\{-n + 2m\}) = \binom{n}{m} 2^{-n}, m = 0, 1, \ldots, n.$
(Man überzeuge sich davon!)

Was wir an diesem Beispiel gesehen haben, gilt auch allgemein:

Ist $(E, \mathfrak{E}) = (R_n, \mathfrak{B}_n)$, so sind mit Zufallsgrößen $X_1, X_2, \cdots$ über $(\Omega, \mathfrak{A}, P)$ auch die Summen $S_n := \sum_{k=1}^{n} X_k$, die Funktionen $M_n : \max(X_1, X_2, \cdots, X_n)$, der Grenzwert $\lim_{n\to\infty} X_n$ (sofern er existiert) wieder Zufallsgrößen. (Die Operationen verstehen sich dabei punktweise, also ω-weise.) Diese Tatsache ergibt sich sofort aus den entsprechenden Eigenschaften meßbarer Funktionen, die in der Maßtheorie bewiesen werden.

Aussage: *Es seien $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum, $(E, \mathfrak{E})$ und $(F, \mathfrak{F})$ meßbare Räume.*
Ist X eine Zufallsgröße über $(\Omega, \mathfrak{A}, P)$ mit Werten in $(E, \mathfrak{E})$ und Y eine Zufallsgröße über $(E, \mathfrak{E}, P^X)$ mit Werten in $(F, \mathfrak{F})$, so ist $Z = Y \circ X$ eine Zufallsgröße über $(\Omega, \mathfrak{A}, P)$ mit Werten in $(F, \mathfrak{F})$ und ihre Wahrscheinlichkeitsverteilung P^Z auf $\mathfrak{F}$ ist gegeben durch

$$P^Z(C) = P^X(Y^{-1}(C)) = P(X^{-1}(Y^{-1}(C))), C \in \mathfrak{F}. \tag{5}$$

Beweis: Nach Definition gilt

$Y^{-1}(\mathfrak{F}) \subseteq \mathfrak{E}$ und

$X^{-1}(\mathfrak{E}) \subseteq \mathfrak{A}$, folglich gilt $Z^{-1}(\mathfrak{F}) \subseteq \mathfrak{A}$. Also ist Z eine Zufallsgröße über $(\Omega, \mathfrak{A}, P)$ mit Werten in $(F, \mathfrak{F})$. Die Formel (5) ergibt sich unmittelbar aus der Definition der Verteilung von Z.

Definition: *Jede Zufallsgröße über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten in $(R_1, \mathscr{B}_1)((R_n, \mathscr{B}_n))$ heißt eine reellwertige Zufallsgröße (bzw. ein n-dimensionaler zufälliger Vektor).*
$\mathfrak{B}_1$ bzw. $\mathfrak{B}_n$ bezeichnen dabei die σ-Algebren der Borelmengen aus R_1 bzw. R_n (siehe Maßtheorie).

Die Wahrscheinlichkeitsverteilung P^X sagt uns, mit welcher Wahrscheinlichkeit die Zufallsgröße X ihren Wert in Mengen $B \in \mathfrak{E}$ annimmt:

$P^X(B) = P(\{\omega \in \Omega | X(\omega) \in B\}),\ B \in \mathfrak{E}$

Die Verteilung P^X ist als Wahrscheinlichkeitsmaß ein kompliziertes i. A. und nur schwer zu handhabbares Objekt. (Für jedes $B \in \mathfrak{E}$ muss man $P^X(B)$ angeben.) Bei reellwertigen Zufallsgrößen und zufälligen Vektoren gelingt es, die Eigenschaften von P^X durch die Untersuchung eines einfacheren mathematischen Objektes, die sogenannte Verteilungsfunktion von P^X bzw. X, zu erhalten.

Am Ende des Abschnittes 3.2 haben wir die Zufallsgröße T_{-1} im Zusammenhang mit einer symmetrischen Irrfahrt eingeführt und de facto die Wahrscheinlichkeitsverteilung der Zufallsgröße

$T_{-1}(\omega) = \min\{k \geq 1 | S_k(\omega) = -1\}$

bei unbegrenzter Fortführung des Münzexperimentes berechnet:

$q_{2m+1} := P(\tau_{-1}(\omega) = 2m+1) = \frac{(2m-1)!!}{(m+1)!2^{m+1}}, \quad m \geq 1,$

$P(T_{-1}(\omega) = 1) = \frac{1}{2}.$

Wir werden später sehen, dass $\sum_{m=0}^{\infty} q_{2m+1} = 1$ gilt, d. h., bei unbegrenzter Fortführung des Münzwurfes wird der Spieler A mit Wahrscheinlichkeit Eins irgendwann "ruiniert", d. h. sein Guthaben wird irgendwann negativ.

Völlig analog kann man aber auch schlussfolgern, dass er mit Wahrscheinlichkeit Eins irgendwann mindestens einen Betrag der Größe Eins auf seinem Konto hat. (Wir setzen dabei voraus, dass er in der Zwischenzeit, wenn sein Guthaben im Negativen ist, immer genügend Finanzmittel besitzt, das Spiel fortzusetzen.) Wenn er die Strategie verfolgt, in dem Moment aufzuhören zu spielen, wenn er das erste Mal einen Gesamtgewinn der Höhe Eins hat, so gewinnt er bei dem vereinbarten durchaus fairen Spiel des Münzenwurfes ohne Zeitlimit mit Wahrscheinlichkeit Eins eine Geldeinheit. Ein Paradoxon.

Verteilungsfunktionen reellwertiger Zufallsgrößen

Definition: *Ist X eine reellwertige Zufallsgröße über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$, so bezeichnet man die durch*

$$F_X(x) := P(\{\omega \in \Omega | X(\omega) \leq x\}) = P(X \leq x), x \in R_1$$

auf R_1 gegebene Funktion F_X als Verteilungsfunktion der Zufallsgröße X.

Offenbar gilt nach Definition der Wahrscheinlichkeitsverteilung P^X die Gleichung $F_X(x) = P^X((-\infty, x]), x \in R_1$.

Wenn keine Verwechslungen möglich sind, schreiben wir F anstelle F_X. Ist $a < b$ so haben wir $(-\infty, a] \subseteq (-\infty, b]$ und wegen $(a, b] = (-\infty, b] \backslash (-\infty, a]$ gilt

$$P(a < X \leq b) = P^X((a, b]) = F(b) - F(a). \tag{6}$$

Aussage: *Die Verteilungsfunktion $F = F_X$ hat folgende Eigenschaften:*

1. *F ist monoton nichtfallend: $x \leq y \Longrightarrow F(x) \leq F(y)$*
2. $\lim\limits_{x \to -\infty} F(x) = 0, \lim\limits_{x \to \infty} F(x) = 1$
3. *F ist an jeder Stelle $x \in R_1$ von rechts stetig:*

   $$F(x + 0) := \lim_{y \downarrow x} F(y) = F(x)$$

4. *Für jedes $x \in R_1$ gilt mit $F(x - 0) := \lim\limits_{y \uparrow x} F(y)$*

   $$F(x) - F(x - 0) = P(X = x) = P^X(\{x\})$$

Bemerkung: Jede Funktion F auf R_1 mit den Eigenschaften 1. - 3. heißt eine Verteilungsfunktion auf R_1. Wir sehen weiter unten, dass jede Verteilungsfunktion von einer Wahrscheinlichkeitsverteilung erzeugt wird.

Beweis der Aussage:

1. $x \leq y \Longrightarrow (-\infty, x] \subseteq (-\infty, y] \Longrightarrow F(x) \leq F(y)$

2. $x_n \downarrow -\infty \Longrightarrow (-\infty, x_n] \downarrow \emptyset \Longrightarrow \lim\limits_{n\to\infty} F(x_n) = 0,$

   $x_n \uparrow \infty \Longrightarrow (-\infty, x_n] \uparrow R_1 \Longrightarrow \lim\limits_{n\to\infty} F(x_n) = 1$

3. Wegen $(-\infty, x] = \bigcap\limits_n (-\infty, x_n]$ für jede Folge (x_n) mit $x_n \downarrow x \Longrightarrow F(x_n) \downarrow F(x)$

4. $(x_n, x] \downarrow \{x\}$ für jede Folge (x_n) mit $x_n \uparrow x$.
   $$\begin{aligned} \Longrightarrow F(x) - F(x-0) &:= \lim_{n\to\infty} (F(x) - F(x_n)) \\ &= \lim_{n\to\infty} P^X((x_n, x]) = P^X(\{x\}). \end{aligned}$$

Die Bedeutung der Verteilungsfunktion F_X für die Wahrscheinlichkeitsverteilung P^X der Zufallsgröße X besteht in Folgendem.

Kennt man $F = F_X$, so ist P^X zumindestens auf Mengen A der Form $A = (a, b], a < b$ festgelegt:

$$P^X((a, b]) = F(b) - F(a), \quad a < b. \tag{7}$$

Das Mengensystem $\gamma = \{(a, b] | -\infty < a < b < \infty\}$ ist aber ein Semiring, und die Mengenfunktion

$$Q((a, b]) := F(b) - F(a), \quad (a, b] \in \gamma \tag{8}$$

ist auf γ σ-stetig (d. h., ist $[a_n, b_n)$ eine monoton fallende Folge halboffener Intervalle mit $\bigcap\limits_{n\geq 1} [a_n, b_n) = \emptyset$, so haben wir $\lim\limits_{n\to\infty} Q((a_n, b_n]) = 0$).
(Der Beweis ist sehr technisch, siehe z. B. Siraev, II 3).

In der Maßtheorie wird gezeigt, dass man Q auf eine und nur eine Weise zu einem Maß auf die σ-Algebra $\mathfrak{B}_1$ der Borelmengen aus R_1 erweitern kann. Da P^X mit Q auf γ übereinstimmt, ist P^X diese Erweiterung.
Damit haben wir gezeigt:

Aussage: *Zu jeder Wahrscheinlichkeitsverteilung Q auf R_1 gehört eine Verteilungsfunktion F, definiert durch*

$$F(x) = Q((-\infty, x]), \quad x \in R_1.$$

Jede Verteilungsfunktion F auf R_1, d. h. jede Funktion F mit den Eigenschaften 1.-3., ist die Verteilungsfunktion einer Wahrscheinlichkeitsverteilung Q auf $(R_1, \mathfrak{B}_1)$. Wahrscheinlichkeitsverteilung und Verteilungsfunktion entsprechen sich eineindeutig.

Beispiele:

a) Ist $\Omega, \mathfrak{A}, P)$ ein Laplace-Experiment mit $\Omega = \{w_1, w_2, \cdots w_N\}, \mathfrak{A} = \mathfrak{P}(\Omega)$, $P(\{\omega\}) = \frac{1}{N}, \omega \in \Omega$, und ist X eine reellwertige Zufallsgröße über $(\Omega, \mathfrak{A}, P)$ mit $X(\omega_k) = x_k, k = 1, \cdots, N, \quad x_k \neq x_j$ für $k \neq j$, so lautet die Verteilungsfunktion $F = F_X$ wie folgt:

$$F(x) = \frac{1}{N} \sum_{k: x_k \leq x} \mathbb{1} = \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{(-\infty, x_k]}(x), x \in R_1$$

F ist in diesem Fall eine stückweise konstante rechtsseitig stetige, nichtfallende Funktion auf R_1 mit den Sprungstellen x_k, den Sprunghöhen $\frac{1}{N}$ und $F(x) = 0$ für $x < \min\limits_{k=1,\cdots,N} x_k$ und $F(x) = 1$ für $x \geq \max\limits_{k=1,\ldots,N} x_k$.

b) Ist $(\Omega, \mathfrak{A}, P) = ([0,1), \mathfrak{B}_{[0,1)}, \lambda_{[0,1)})$ der Wahrscheinlichkeitsraum, der die rein zufällige Wahl eines Punktes ω aus $[0,1)$ modelliert, und ist $X(\omega) = \omega, \omega \in \Omega$, so gilt für $F = F_X$

$$F(x) = \begin{cases} 0 & x \leq 0 \\ x & x \in [0,1] \\ 1 & x \geq 1 \end{cases}$$

Eine Möglichkeit, sich von der Lage einer Wahrscheinlichkeitsverteilung auf R_1 und ihre Ausbreitung eine Vorstellung zu verschaffen, besteht in der Berech-

nung ihrer Quantile.

Definition: *Es seien Q eine Wahrscheinlichkeitsverteilung auf $(R_1, \mathfrak{B}_1)$ und F ihre Verteilungsfunktion:*

$F(x) = Q((-\infty, x]), \quad , x \in R_1.$

Weiterhin sei p irgendeine Zahl mit $0 < p < 1$.
Als p-Quantil der Teilung bezeichnet man jede Zahl $q_p \in R_1$ mit $F(q_p - 0) \leq p \leq F(q_p)$.

Die Menge aller p-Quantile von Q ist nichtleer und bildet ein beschränktes abgeschlossenes Intervall. Sie ist einelementig genau dann, wenn es keine zwei Zahlen $x < y$ gibt mit $F(x) = F(y) = p$.

Definition: *Jedes Quantil $q_{\frac{1}{2}}$ heißt Median. Jedes $q_{\frac{1}{2}}$ hat also die Eigenschaft $Q((-\infty, q_{\frac{1}{2}})) = F(q_{\frac{1}{2}} - 0) \leq \frac{1}{2} \leq q_{\frac{3}{4}} F(q_{\frac{1}{2}}) = Q((\infty, q_{\frac{1}{2}}])$. Das Quantil $q_{\frac{1}{4}}$ heißt unteres und oberes Quantil. Die Differenz $q_{\frac{3}{4}} - q_{\frac{1}{4}}$ ist ein Maß für die "Ausbreitung" der Verteilung Q. Es gilt $Q([q_{\frac{1}{4}}, q_{\frac{3}{4}}]) = Q((-\infty, q_{\frac{3}{4}}]) - Q((-\infty, q_{\frac{1}{4}})) \geq \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$.*

Verteilungsfunktion zufälliger Vektoren

Es sei X ein n-dimensionaler zufälliger Vektor über $(\Omega, \mathfrak{A}, P)$, d. h. $X(\omega) = (X_1(\omega), X_2(\omega), \cdots, X_n(\omega)), \quad \omega \in \Omega$, und die Abbildung X ist $\mathfrak{A} - \mathfrak{B}_n$-messbar. Äquivalent dazu ist, dass alle Komponentenabbildungen X_k reellwertige Zufallsgrößen, also $\mathfrak{A} - \mathfrak{B}_1$-messbare Funktionen sind.

Definition: *Die Funktion*

$F_X(x_1, \cdots, x_n) :=$

$P(\{\omega \in \Omega | X_1(\omega) \leq x_1, \cdots, X_n(\omega) \leq x_n\}) =$

$P(X_1 \leq x_1, \cdots, X_n \leq x_n) =$

$P(X \in (-\infty, x_1] \times \cdots \times (-\infty, x_n]) =$

$= P^X((-\infty, x_1] \times \cdots \times (-\infty, x_n]), x = (x_1, \cdots, x_n) \in R_n$

heißt Verteilungsfunktion des zufälligen Vektors X bzw. seiner Wahrscheinlichkeitsverteilung P^X.

Aussage: *Die Verteilungsfunktion F_X hat folgende Eigenschaften:*

1. $0 \leq F_X(x) \leq 1, x \in R_n$
2. $\lim\limits_{x_k \downarrow -\infty} F_X(x_1, \cdots, x_n) = 0$ *für jedes* $k = 1, \cdots, n$
3. $\lim\limits_{x_1, \cdots, x_n \uparrow \infty} F_X(x_1, \cdots, x_n) = 1$
4. *F_X ist von rechts stetig:*
   $\lim\limits_{\substack{h_i \downarrow 0 \\ i=1,\cdots,n}} F_X(x_1 + h_1, \cdots, x_n + h_n) = F_X(x_1, \cdots, x_n)$
5. *Mit der Definition*
   $\triangle_{h_i} F_X(x) = F_X(x_1, \cdots, x_{i-1}, x_i + h_i, x_{i+1}, \cdots, x_n) - F_X(x_1, \cdots, x_n)$ *gilt*
   $0 \leq \triangle_{h_1} \cdot \ldots \cdot \triangle_{h_n} F_X(x)$ *(Verallgemeinerung der Monotonie im Fall $n = 1$)*

Beispiel zu 5: Für $n = 2$ gilt

$$F_X(x_1 + h_1, x_2 + h_2) - F_X(x_1, x_2 + h_2) - F_X(x_1 + h_1, x_2) + F_X(x_1, x_2) \geq 0$$

Analog zum eindimensionalen Fall kann man zeigen, dass die Wahrscheinlichkeitsverteilung P^X von X durch die Verteilungsfunktion F_X eindeutig bestimmt ist. Das entsprechende Mengensystem γ besteht jetzt aus den "Rechtecken" $\prod\limits_{k=1}^{n} \Big(a_k, b_k\Big]$.

Wie im Fall $n = 1$ werden wir jede Funktion F auf R_n mit den Eigenschaften 1. - 5. als Verteilungsfunktion auf R_n bezeichnen. Vermittels

$$Q\Big(\prod_{k=1}^{n}(a_k, a_k + h_k]\Big) = \triangle_{k_1} \cdots \triangle_{k_n} F(x)$$

ist eine eindeutig bestimmte Wahrscheinlichkeitsverteilung Q auf $(R_n, \mathfrak{B}_n)$ definiert, für die F die zugehörige Wahrscheinlichkeitsverteilung ist. Zu jeder Verteilungsfunktion F auf R_n d. h. jeder Funktion F mit 1. - 5. gilt es eine Wahrscheinlichkeitsverteilung P auf $(R_n, \mathfrak{B}_n$, die F als Verteilungsfunktion hat.

3.7 Verteilungsfunktionen

Verteilungsdichten

Wir haben gesehen, dass sich Wahrscheinlichkeitsverteilungen auf $(R_1, \mathfrak{B}_1)$ bzw. $(R_n, \mathfrak{B}_n)$ durch ihre Verteilungsfunktionen charakterisieren lassen. Diese wiederum besitzen in Spezialfällen eine besonders einfache Struktur. Zum einen handelt es sich dabei um sogenannte *diskrete Verteilungen* bzw. *diskret verteilte Zufallsgrößen*. Diesen Verteilungen ist das Kapitel 4 gewidmet. Zum anderen heben sich Verteilungen mit einer *Verteilungsdichte*, man spricht auch einfach von *Dichten*, heraus. Diese Verteilungen werden wir in voller Allgemeinheit unter Verwendung des Begriffs des Lebesgueintegrals erst in Kapitel 7 behandeln. Vorab wollen wir jedoch Fälle, in denen man mit dem vertrauten Riemannintegral auskommt, vorstellen.

Definition: *Es sei $n \geq 1$ und F eine Verteilungsfunktion auf R_n. Weiterhin sei f eine nichtnegative, Riemann-integrierbare Funktion auf R_n mit*

$$\int\limits_{R_n} f(x)dx = \underbrace{\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty}}\, f(x_1, x_2, \cdots, x_n)dx_1, dx_2, \cdots, dx_n = 1 \tag{1}$$

Man sagt, F besitze eine Verteilungsdichte und diese sei gleich f, falls

$$F(x) = \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{x_2} \cdots \int\limits_{-\infty}^{x_n} f(s_1, s_2, \cdots, s_n)ds_1, ds_2, \cdots, ds_n \tag{2}$$

für alle $x = (x_1, x_2, \cdots, x_n) \in R_n$ *gilt.*

Die folgende Aussage enthält zwei unmittelbare Folgerugnen dieser Definition.

Aussage:

a) Jede nichtnegative Riemannintegrierbare Funktion f *auf* R_n *mit der Eigenschaft (1) definiert durch (2) eine Verteilungsfunktion* F *auf* R_n.

b) Ist f *eine Dichte der Verteilungsfunktion* F *auf* R_n*, und ist* f *in Umgehung von* $x = (x_1, x_2, \cdots, x_n)$ *stetig, so gilt*

$$f(x_1, x_2, \cdots, x_n) = \frac{\partial^n}{\partial_{x_1}\partial_{x_2}\cdots\partial_{x_n}} F(x_1, x_2, \cdots, x_n) \tag{3}$$

für dieses $x = (x_1, x_2, \cdots, x_n)$.

Beweis:

a) Man weist leicht alle die Verteilungsfunktion definierenden Eigenschaften für F nach.

b) Man wende den Mittelwertsatz der Integralrechnung an.

Beispiele:

1) Die gleichmäßige Verteilung $U([0,1))$ auf $([0,1), \mathfrak{B}_{[0,1)})$ (Beispiel b) aus 3.6) besitzt die Verteilungsfunktion

$$F(x) = \begin{cases} 0 & : \quad x \leq 0 \\ x & : \quad x \in [0,1] \\ 1 & : \quad x \geq 1 \end{cases}$$

und die Dichte

$$f(x) = \mathbb{1}_{[0,1]}(x), \quad x \in R_1,$$

2) Es sei $\lambda > 0$ und f_λ die durch

$$f_\lambda(x) = \left\{ \begin{array}{c} 0 \quad , x < 0 \\ \lambda \mathrm{e}^{-\lambda x} \, , x \geq 0 \end{array} \right\} = \mathbb{1}_{[0,\infty)}(x) \lambda \mathrm{e}^{-\lambda x}, \; x \in R_1$$

definierte Funktion. Dann ist f_λ die Dichte einer Wahrscheinlichkeitsverteilung auf $(R_1, \mathfrak{B}_1)$, die man als *Exponentialverteilung* $EXP(\lambda)$ *mit dem Parameter* λ bezeichnet. Ihre Verteilungsfunktion F_λ lautet

$$F_\lambda(x) = \left\{ \begin{array}{c} 0 \quad , x \leq 0 \\ 1 - \mathrm{e}^{-\lambda x} \, , x > 0 \end{array} \right.$$

3) Die *eindimensionale Normalverteilung* $N(\mu, \sigma^2)$ (auch Gaußsche Verteilung genannt) besitzt zwei Parameter $\mu \in R_1$ sowie $\sigma^2 > 0$ und ist definiert durch ihre Verteilungsdichte

$$\varphi_{\mu,\sigma^2}(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\Big[-\tfrac{1}{2\sigma^2}(x-\mu)^2 \Big], \; x \in R_1$$

Die Verteilungsdichte hat als Graphik die Form der sogenannten Gaußschen Glockenkurve mit ihrem Maximum im Punkt μ und den Wendepunkten bei $\mu + \sigma$ und $\mu - \sigma (\sigma = \sqrt{\sigma^2} > 0$. Sie ist symmetrisch bez. $\mu : \varphi_{\mu,\sigma^2}(\mu + z) = \varphi_{\mu,\sigma^2}(\mu - z) \; , \; z > 0$, und fällt für $x \to +\infty$ und $x \to -\infty$ sehr schnell gegen Null ab.

Die Verteilungsfunktion φ_{μ,σ^2} ist gegeben durch $\Phi_{\mu,\sigma^2}(x) = \int\limits_{-\infty}^{x} \varphi_{\mu,\sigma^2}(s) ds$ und nicht explizit durch elementare Funktionen ausdrückbar. $\varphi = \varphi_{0,1}$ heißt Dichte und $\Phi = \Phi_{0,1}$ Verteilungsfunktion der *Standard-Normalverteilung* $N(0,1)$.

Es gilt

$$\varphi_{\mu,\sigma^2}(x) = \tfrac{1}{\sigma} \varphi_{0,1}(\tfrac{x-\mu}{\sigma}).$$

Eigenschaften von φ:

1. φ ist bezüglich Null symmetrisch: $\varphi(-x) = \varphi(x), x \in R_1$.
2. φ ist unimodal und hat ihr Maximum bei Null.
3. φ hat zwei Wendepunkte, und zwar bei $+1$ und -1.
4. Die Verteilungsfunktion $\Phi = \Phi_{0,1}$ der Standardnormalverteilung

   ist vertafelt (siehe Vorlesungsanlagen). Es gilt

   $1 - \Phi(x) = \Phi(-x), \quad x \in R_1, \quad \Phi(0) = 0,5$ und

   $1 - \Phi(x) \leq \frac{1}{x\sqrt{2\pi}} \mathrm{e}^{-\frac{\mathrm{x}^2}{2}} \quad (\mathrm{x} > 0),$

   $\Phi_{\mu,\sigma^2}(x) = \Phi\left(\frac{x-\mu}{\sigma}\right),\ x \in R_1.$

4) Es sei $\mu \in R_n$ und $\sum$ eine positiv definierte symmetrische $n \times n$-Matrix. Dann ist die Funktion $f_{\mu,\sum}$, definiert durch

   $f_{\mu,\sum}(x) = \frac{1}{(2\pi)^{n\backslash 2}\sqrt{\det \sum}} \exp\left[-\frac{1}{2}(x-\mu)^T \sum^{-1}(x-\mu)\right],\ x \in R_n$

   die Dichte der sogenannten n-dimensionalen Normalverteilung $N(\mu, \sum)$.

   Im Fall $n = 2$ hat sie mit der Notation

   $$\sum = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\rho_2 & \sigma_2^2 \end{pmatrix}, \sigma_1, \sigma_2 > 0, |\rho| < 1, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

   die Form

   $f_{\mu,\sum}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}.$

   $\exp\left[-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\right]$

Als Höhenlinie H_c der Dichte f_X bezeichnet man die Menge aller Punkte (x_1, x_2), in denen f_X den Wert c besitzt $\left(0 < c < \frac{1}{2\pi\sigma_1,\sigma_2\sqrt{1-\varrho^2}}\right)$.

Offenbar ist

$$H_0 = \{x = (x_1, x_2) : f_X(x) = c\} =$$

$$\{x : (x-\mu)^T \textstyle\sum_X^{-1}(x-\mu) = c^*\}$$

für ein gewisses $c^* > 0$. (Genauer: $c^* = -2(1-\varrho^2)ln(2\pi\sigma_1, \sigma_2\sqrt{1-\varrho^2}c)$. Die Menge H_c ist eine Ellipse mit dem Mittelpunkt μ. Wir setzen $\mu = 0$. Die Form und Lage der Ellipsen H_c hängt von σ_1, σ_2 und ϱ ab.

Gilt $\varrho = 0$, so sind die Hauptachsen der Ellipsen H_c die Koordinatenachsen und die Länge der Hauptachsenabschnitte ist $\sqrt{c^*} \cdot \sigma_i, i = 1, 2$.

Gilt $\sigma_1^2 = \sigma_2^2 = \sigma^2$ und $\varrho \neq 0$, so sind die Hauptachsen um $\frac{\pi}{4}$ gegenüber den Koordinatenachsen gedreht und die Längen der Hauptachsenabschnitte $\sqrt{c^*} \cdot \sigma \cdot \sqrt{1 \pm \varrho}^{-1}$. (Das Vorzeichen $-$ steht für die Hauptachse im 1. und 3. Quadranten.)

Im allgemeinen Fall $\sigma_1^2 \neq \sigma_2^2$ und $\varrho \neq 0$ sind die Hauptachsen der Ellipsen um den Winkel

$$\gamma = 0,5 \cdot \arctan\left(\frac{2\varrho\sigma_1\sigma_2}{\sigma_2^2-\sigma_1^2}\right)$$

gegenüber den Koordinatenachsen gedreht. Die Hauptachsenabschnitte sind

$$a = \sigma_1\sigma_2\left(\frac{c^*(1-\varrho^2)}{\sigma_1^2\sin^2\gamma+\sigma_2^2\cos^2\gamma+2\varrho\sigma_1\sigma_2\sin\gamma\cos\gamma}\right)^{\frac{1}{2}}$$

bzw.

$b = \sigma_1\sigma_2\left(\frac{c^*(1-\varrho^2)}{\sigma_1^2\cos^2\gamma+\sigma_2^2\sin^2\gamma-2\varrho\sigma_1\sigma_2\sin\gamma\cos\gamma}\right)^{\frac{1}{2}}$

(a steht für den Hauptachsenabschnitt im ersten Quadranten).

(Siehe Beyer et. al., Wahrscheinlichkeitsrechnung und mathematische Statistik, 7. Auflage, Teubner 1995, Abschnitt 2.3)

It X eine reellwertige Zufallsgröße mit der Dichte f_X und ist ψ eine streng monotone und stetig differenzierbare Funktion von R_1 in sich, so gilt die folgende

Aussage: $Y = \psi(X)$ *besitzt ebenfalls eine Dichte* f_Y*, und es gilt*

$f_Y(y) = f_X(\psi^{-1}(y))|\frac{d\psi^{-1}}{dy}(y)|\ ,\ y \in Wb(X).$

Beweis: Zunächst sei ψ streng wachsend. Dann ist

$F_Y(y) = P(Y \leq y) = P(\psi(X) \leq y) = P(X \leq \psi^{-1}(y))$

und somit

$f_>(y) = f_X(\psi^{-1}(y)) \cdot \frac{d\psi^{-1}}{dy}(y).$

Ist dagegen ψ streng fallend, so haben wir

$F_Y(y) = P(X \geq \psi^{-1}(y)) = 1 - P(X < \psi^{-1}(y))$ und

$f_Y(y) = f_X(\psi^{-1}(y))|\frac{d\psi^{-1}}{dy}(y)|.$

Beispiel: X sei gleichmäßig auf $[0,1)$ verteilt und es sei $Y = \lambda^{-1}\, lnX$ für ein $\lambda > 0$. Dann ist Y eine mit dem Parameter λ exponentialverteilte Zufallsgröße.