

# Stochastische Probleme aus elementarer Sicht

Dr. Wolfgang Riemer  
Köln



Wissenschaftsverlag  
Mannheim/Wien/Zürich

CIP-Titelaufnahme der Deutschen Bibliothek

**Riemer, Wolfgang:**

Stochastische Probleme aus elementarer Sicht / von Wolfgang Riemer. – Mannheim; Wien; Zürich: BI-Wiss.-Verl., 1991

(Lehrbücher und Monographien zur Didaktik der Mathematik;

Bd. 18)

ISBN 3-411-14791-1

NE: GT

Gedruckt auf säurefreiem Papier  
mit neutralem pH-Wert (bibliotheksfest)

Alle Rechte, auch die der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieses Werkes darf ohne schriftliche Einwilligung des Verlages in irgendeiner Form (Fotokopie, Mikrofilm oder ein anderes Verfahren), auch nicht für Zwecke der Unterrichtsgestaltung, reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

© Bibliographisches Institut & F.A. Brockhaus AG, Mannheim 1991

Druck: Druckerei Krembel, Speyer

Bindearbeit: Progressdruck GmbH, Speyer

Printed in Germany

ISBN 3-411-14791-5

## VORWORT

Wahrscheinlichkeitsrechnung und Statistik sind seit langem in den Lehrplänen unserer Schulen verankert. Trotz vieler Erfahrungen mit Stochastikunterricht offenbaren sich jedoch bei genauerem Hinsehen viele bisher noch ungelöste didaktische Probleme.

Dieses Buch ist aus Arbeiten zu solchen Problemen hervorgegangen. Die Grundlage bilden langjährige Unterrichtserfahrungen, Unterrichtsexperimente und Veranstaltungen zur Lehrerfortbildung. Die Thematik ist weit gespannt: sie reicht von ersten Einführungsstunden, die auch von Studenten, Referendaren und Lehrern mit wenig stochastischer Unterrichtserfahrung erfolgreich getestet wurden, bis zu anspruchsvollen Themen, die gleichermaßen in Leistungskursen und Seminaren auf sehr positive Resonanz stießen.

Wir hoffen, dem Leser viele neue und faszinierende Anregungen geben zu können, die zum Nachdenken über "eingefahrene Gleise" und zum Beschreiten neuer Wege ermuntern. Daher haben wir uns bemüht, die Darstellung durch Beispiele so lebendig zu gestalten, daß sich die Ideen meist nur mit wenig Mühe im Unterricht, in Klausuraufgaben oder in Projekte umsetzen lassen.

Ich danke Herrn Professor Dr. Scheid für zahlreiche Hinweise und für die Anregung zu diesem Buch. Herrn Professor Dr. Michel, Herrn Professor Dr. Pfanzagl und Herrn Professor Dr. Schrage danke ich für wertvolle Anmerkungen. Schließlich danke ich meinen Schülerinnen und Schülern für vergnügliche Unterrichtsstunden und lebhaftige Diskussion. Dem Leser wünsche ich neben Freude an der Lektüre auch Spaß an eigenen Experimenten zu den vorgestellten Themen.

Köln, im Herbst 1990

Wolfgang Riemer



## INHALT

## Vorwort

0	Einführung	11
	- Zielsetzungen	11
	- Inhaltsübersicht	12
	- Computereinsatz	14
	- Statistisches Paradigma	15

Teil I Integration statistischer Aspekte in die  
Wahrscheinlichkeitsrechnung

1	Begriffsbildungen im Dienste der Statistik	16
1.1	Wahrscheinlichkeiten als Hypothesen	16
1.2	Hypothesen aufstellen	20
	- Hypothesen schätzen	20
	- Hypothesen berechnen	25
1.3	Stellung der Wahrscheinlichkeitsrechnung	28
	- Wartezeiten	29
	- Quader contra Würfel	31
	- Binomialverteilungen als Hypothesen	32
1.4	Hypothesen als Modelle der Wirklichkeit	33
1.5	Lernpsychologischer Rückblick	34
2	Abwägen zwischen Hypothesen (durch wiederholte Anwendung der Regel von Bayes)	39
2.1	Vorbemerkung	39
2.2	Merkwürdige Würfel	40
2.3	Bayessche Regel	44
2.4	Intuitive Entscheidungen quantitativ fundiert	48
2.5	Wahrscheinlichkeit, Information, Regel von Bayes	52
	- Entropie	52
	- Information	54
	- Experiment	54
2.6	Resümee	57
	- Subjektivisten - Objektivisten	57
	- Alternativtests aus Bayesscher Sicht	58

Teil II Chi-Quadrat-Testverfahren mit Anwendungen(Aspekte der linearen Geometrie in der Stochastik)

3	Chi-Quadrat-Anpassungstest	63
3.1	Grenzen Bayesscher Betrachtungsweise	63
3.2	Experimenteller Teil	67
	- Testgröße	67
	- Computersimulation	68
	- Versuchsumfang	72
3.3	Theoretischer Teil	73
	- Vektorielle Deutung der Testgröße	74
	- Zweidimensionale Irrfahrt	76
	- Zweidimensionale Normalverteilung	80
	- Chi-Quadrat-Verteilung	81
3.4	Verallgemeinerungen	82
	- Anpassungstests mit beliebigen Freiheitsgraden	82
	- Reduktion des Freiheitsgrades	85
3.5	Historische Anmerkungen	87
4.	Chi-Quadrat-Unabhängigkeitstest	89
4.1	Experimenteller Teil	89
	- Testgröße	91
	- Computersimulation	92
4.2	Theoretischer Teil	96
	- vektorielle Deutung der Testgröße	96
	- zweidimensionale Irrfahrt	99
	- zweidimensionale Normalverteilung	103
4.3	Verallgemeinerung	106
5	Wahrscheinlichkeiten gezinkter Quader	110
5.1	Wurftechnik	110
5.2	Boltzmann-Verteilung	111
5.3	Potenzverteilung	116

Teil III Eindimensionale Irrfahrten und Grenzwertsätze  
(Aspekte der Analysis in der Stochastik)

6	Zentraler Grenzwertsatz	119
6.1	Rekursionsformel für "Summenverteilungen"	120
6.2	Differentialgleichung der Normalverteilung	121
7	Eindimensionale Irrfahrten und Chi-Quadrat-Verteilung	128
7.1	Problemstellung	128
7.2	Differentialgleichung der Chi-Quadrat-Verteilung	131
	- Verallgemeinerung	134
	- Gammaverteilung	135
7.3	Anpassungstest	136
8	Eindimensionale Irrfahrten und Arcsin-Verteilung	138
8.1	Einführung	138
8.2	Irrfahrt	139
8.3	Rekursionsformel	142
8.4	Differentialgleichung der Arcsin-Verteilung	144
8.5	Länge des letzten Runs	148
	- Exkurs	149
	- Resümee	151
9	Mathematischer Hintergrund	153
9.1	Brownsche Bewegung und Zentraler Grenzwertsatz	153
9.2	Diffusionsprozesse und Chi-Quadrat-Verteilung	157
	- Kolmogoroffs Differentialgleichungen	157
	- Differentialgleichung der Chi-Quadrat-Verteilung	158
	- Lösung von Kolmogoroffs Gleichungen	159
	- Konvergenzbeweis	161
9.3	Arcsin-Gesetz und Renewal-Theorie	165

Teil IV Statistik und Datenbanken

10	Partnersuche	168
10.1	Testfragen	168
10.2	"Erotischer Abstand"	168
10.3	Modellbildung	173
	- Zufriedene Leser	173
	- Unabhängigkeitstest	174
	- Anpassungstest	174
	- Vermittlungschancen	175
	Literatur	177

## 0 Einführung

### 0.1 Zielsetzungen

Wer sich mit didaktischen Fragen der Stochastik auseinandersetzt, findet sich schnell in ein spannungsreiches Beziehungsgefüge eingebunden. Der Blick schweift zwischen Wahrscheinlichkeitstheorie und Statistik, er berührt die unterschiedlichen Auffassungen von Subjektivisten und Objektivisten, wendet sich von der Mathematik zu Lehrplänen und unterschiedlichen Realisierungen. Er schweift von psychologischen Untersuchungsergebnissen zu lerntheoretischen Paradigmen und macht bei Interaktionsanalysen nicht halt.

Viele dieser Aspekte werden uns in den folgenden Kapiteln beschäftigen. Den Schwerpunkt werden jedoch Analysen und neue curriculare Ansätze bilden, deren Ziel darin besteht, die ungenetische Trennung zwischen Wahrscheinlichkeitstheorie und Statistik zu überwinden. Daß wir es hier tatsächlich mit einem didaktischen Kernproblem der Stochastik zu tun haben, macht Nemetz auf der First International Conference on Teaching Statistics deutlich, indem er mit Blick auf Bruners Spiralprinzip schreibt:

"The starting point of my dreams is a perfect harmony among the main disciplines of the area called stochastics: the descriptive, the mathematical statistics and the probability theory. ... There is hardly any historical or didactical reason for separating these tree branches, their teaching goes parallel, following a kind of spiral structure (Nemetz 1983 S. 199).

Auch Schupp konstatiert, "daß in den meisten der an den bundesrepublikanischen Schulen verbreiteten Mathematikwerke statistische und wahrscheinlichkeitstheoretische Passagen nahezu beziehungslos (und zuweilen mit falschem Bezug) hintereinandergesetzt sind". "Statistik ohne Wahrscheinlichkeitsrechnung führt nicht zu wesentlich neuen, insbesondere

zu prognostischen Aussagen, Wahrscheinlichkeitsrechnung ohne Statistik ist leer ..." (Schupp 1982, S. 210, S. 214 mit Bezug auf die Untersuchung Steinbring/Strässer 1981).

Tatsächlich reicht die unglückliche Trennung von Wahrscheinlichkeitstheorie und Statistik weit in die Vergangenheit zurück. So haben Kombinatorik und der Laplacesche Wahrscheinlichkeitsbegriff in Schulen eine lange Tradition (Preußische Lehrpläne 1901, Meraner Pläne 1905, Lietzmanns Didaktik von 1923 - vgl. hierzu etwa Inhetveen 1976 S. 206 und Kütting 1981 S. 138). Aspekte der Beurteilenden Statistik wurden erst im Zuge der Oberstufenreform in den Sechziger und Siebziger Jahren in die Lehrpläne (hauptsächlich der Sekundarstufe II) übernommen. Sie standen aber zu Zeiten der "New-Math"-Bewegung (schon wegen der beschränkten Unterrichtszeit) in einer (ungewollten) Konkurrenz zur axiomatischen Grundlegung der Wahrscheinlichkeitsrechnung (Mengenalgebra, Kolmogoroff-Axiome).

Mit einem Rückzug der Strukturmathematik wächst das Gewicht der Statistik, wenngleich ihre Stellung immer noch einem "Anhängsel" an die Wahrscheinlichkeitsrechnung gleicht. Das vorliegende Buch möchte helfen, diese "Anhängsel-Funktion" zu überwinden. Es möchte innovativ wirken und im Sinne von Wittmanns "Teaching Units as the integrating Core of Mathematics Education" (Wittmann 1984) zeigen, wie man genuin statistische Gedankengänge schon von den ersten Unterrichtsstunden an zu einem Bestandteil von Stochastikunterricht machen kann. Es möchte zeigen, daß sich auch so "schwierige" Themengebiete wie Chi-Quadrat-Verfahren und stochastische Grenzwertsätze elementaren Darstellungen nicht entziehen.

## 0.2 Inhaltsübersicht

Die einzelnen Kapitel lassen sich unabhängig voneinander lesen, sie sind jedoch durch Beispiele, Querverweise und einen "roten Faden" miteinander verbunden, der sich als eine Folge von Antworten auf bisher noch ungelöste didaktische

Probleme lesen läßt:

In Kapitel 1 beschäftigen wir uns mit dem Begriff der Wahrscheinlichkeitshypothese, dessen Bereitstellung für einen konsistenten Aufbau von Wahrscheinlichkeitsrechnung und Statistik schon in sehr frühen Phasen unerlässlich ist. Wir zeigen, wie man diesen Begriff durch neue Sichtweisen und geeignetes Unterrichtsmaterial aus den in der Alltagserfahrung geprägten Primärintuitionen "herausschälen" kann. Statt einer Beschränkung auf das zu enge Laplacesche Konzept lernen Schüler schon sehr früh Aspekte des Modellbildens kennen, die mit jeder Art des Mathematisierens intrinsisch verbunden sind.

Diese Ansätze werden lernpsychologisch begründet und in Kapitel 2 konsistent weitergeführt, indem wir mit einer für den Schulbereich neuen Sicht der Bayesschen Regel ein quantitatives Hilfsmittel zum Abwägen zwischen Alternativhypothesen bereitstellen. So dient die Wahrscheinlichkeitsrechnung dem Ausbau des statistischen Gedankengebäudes auf eine Weise, die ihre Spuren inzwischen auch in der Lehrplanarbeit hinterläßt.

In Kapitel 3 werden wir eine weitergehende Vertiefung vornehmen, indem wir in einer auch für Schüler nachvollziehbaren Weise die Grenzen Bayesscher Betrachtungsweise aufzeigen. Als Antwort auf diese Grenzen werden wir uns mit Elementarisierungen des Chi-Quadrat-Anpassungstests beschäftigen, die wir in Kapitel 4 auf den Unabhängigkeitstest übertragen und in den Kapiteln 5 und 10 durch interessante Anwendungen bereichern.

In Kapitel 6 widmen wir uns einer heuristischen Begründung des Zentralen Grenzwertsatzes, dem Fundament vieler statistischer Testverfahren, dem traditionell aber auch unabhängig von der Statistik ein starkes Eigengewicht zukommt: Wir überführen Rekursionsgleichungen, die zu Summen von Zufallsgrößen gehören, in eine Differentialgleichung mit der Nor-

malverteilung als Lösung.

In den folgenden Kapiteln 7 und 8 gelingt uns die Übertragung dieser Technik auf eindimensionale Irrfahrtprobleme mit interessanten Anwendungen im Bereich des Anpassungstests und des Arcsin-Gesetzes. So werden elementare Wege zu recht anspruchsvollen Themen gebahnt, die über bloße Computersimulationen hinausgehen.

In Kapitel 9 skizzieren wir, wie sich die heuristischen Überlegungen der Kapitel 6, 7 und 8 mathematisch absichern lassen. Wir stellen Beziehungen her zu Diffusionsprozessen und der Renewal-Theorie.

Ein "heiteres" Kapitel 10 schließt das Buch ab. Anhand des Themas "Partnersuche" soll exemplarisch gezeigt werden, wie sich Aspekte der informationstechnologischen Grundbildung mit Aspekten von Wahrscheinlichkeitsrechnung und Statistik zu Unterrichtsprojekten verbinden lassen, die auch mathematisch weniger interessierte Schüler motivieren dürften.

### 0.3 Computereinsatz

Über die geschilderten Intentionen hinaus lassen sich die Kapitel über weite Strecken auch als inhaltliche Beiträge zu einer Didaktik des Computereinsatzes lesen. Sie zeigen exemplarisch, wie der Computer in der Stochastik über Simulationszwecke hinaus zur Verarbeitung eigener nicht simulierter Daten benutzt werden kann. Es gelingt, "emotionale Betroffenheit" zu erzeugen und damit auch affektive Komponenten in Lernprozessen anzusprechen (Kap. 2, 5, 10). Computersimulationen (Kap. 2, 3, 4, 6, 7, 8) dienen uns zur intuitiven Vorbereitung komplexer Zusammenhänge, allerdings bleiben wir bei Simulationen nicht stehen.

#### 0.4 Statistisches Paradigma

Egal, ob wir bei unseren Untersuchungen reale oder simulierte Daten verwenden: Das Formulieren von Fragen, das Erzeugen von Erwartungshaltungen, das Aufstellen von Hypothesen vor den Experimenten und die Bewertung der Prognosen durch Gegenüberstellung von Voraussagen und Ergebnissen erst im nachhinein muß in unseren Augen zu einem Paradigma des Stochastikunterrichts werden. Seine konsequente Berücksichtigung hält statistische und wahrscheinlichkeitstheoretische Aspekte wie eine Klammer zusammen. Darüber hinaus werden Schüler unmittelbar ins Unterrichtsgeschehen einbezogen und emotional angesprochen, da ihre eigenen Hypothesen zur Disposition stehen. Varga (1983, S. 73) unterstützt diese Sicht, indem er feststellt:

"Young children seem to be more responsive to essential, deep questions than occasional superficial ones ... they like personal involvement. In their experiences about random phenomena their involvement is greatly increased by first predicting the outcome then performing the experiment, finally comparing the result with their prediction. A further step is an attempt to explain the pattern of results by some theory. This carries them further towards probability. The order is hypothesis -> experience -> theory."

Neben anderen lernpsychologischen Leitlinien wird uns dieses Paradigma in verschiedenen Ausprägungen und variierenden Formen durch das Buch begleiten. Wir werden seine Tragfähigkeit ausloten und feststellen, daß es uns dem gesteckten Ziel einer Integration von wahrscheinlichkeitstheoretischen und statistischen Aspekten ein gutes Stück näherbringen wird.

Wie man dem Inhaltsverzeichnis entnimmt, haben wir das Buch in vier Teile gegliedert, denen weitere Informationen und didaktische Hinweise vorangestellt sind.

## Teil I Integration statistischer Aspekte in die Wahrscheinlichkeitsrechnung

In den beiden folgenden Kapiteln untersuchen wir, wie man mit geeignetem Unterrichtsmaterial schon in kleinen Kindern tief verwurzelte (und teilweise subjektivistisch geprägte) statistische Intuitionen mobilisiert. Wir werden diese wertvollen Vorstellungen handelnd (enaktiv) systematisieren, bis zum Begriff der Hypothese ausbauen und anschließend mit Hypothesen arbeiten. Unser Material besteht aus partiell symmetrischen Objekten (Quadern, Prismen), die im Gegensatz zu den klassischen Würfeln und Glücksrädern mehrere sinnvolle Hypothesen zulassen und erst dadurch den hypothetischen Charakter von Wahrscheinlichkeiten verdeutlichen. Andererseits machen sie - im Gegensatz zu völlig unsymmetrischen Reißnägeln oder gezinkten Würfeln - die Zusammenhänge zwischen Wahrscheinlichkeiten und relativen Häufigkeiten im wahrsten Sinne des Wortes greifbar. "Gedankliches Chaos" (Schönwald 1984), das sich häufig beim Verschwimmen der genannten Begriffe breitmacht, wird so vermieden (Schülerzitat: "relative Wahrscheinlichkeiten").

### 1 Begriffsbildungen im Dienste der Statistik

#### 1.1 Wahrscheinlichkeiten als Hypothesen

Die Integration statistischer Aspekte in die Wahrscheinlichkeitsrechnung muß "im Herzen" der Stochastik ansetzen, bei den Begriffen Wahrscheinlichkeit, Wahrscheinlichkeitshypothese und ihren Interpretationen. Anders als bei vielen Begriffen in den "klassischen" Lernbereichen Algebra und Geometrie ist die Bedeutung des Wahrscheinlichkeitsbegriffs nicht im Rahmen innermathematischer Systematik erfahrbar. Die Entwicklung einer adäquaten Vorstellung erfordert vielfältige experimentelle Erfahrungen, die auf unteren Curriculumstufen durchaus spielerischen Charakter haben dürfen. So stellen Fischbein/Gazit (1984, S. 3) in Übereinstimmung mit

unserem statistischen Paradigma (0.4) fest:

"New intuitive attitudes can be developed only through the personal involvement of the learner in a practical activity. Intuitions (cognitive beliefs) cannot be modified by verbal explanations only. Therefore a teaching program which intends to develop an improved and efficient intuitive background for probability concepts and strategies, along with the corresponding formal knowledge, must provide the learner with frequent opportunities to experience actively, even emotionally, with stochastic situations. In such situations the learner will confront his plausible expectations with empirically obtained outcomes."

Laplacescher Wahrscheinlichkeitsbegriff:

In der Schulpraxis dominieren traditionsgemäß nahezu ausschließlich Laplacesche Experimente, bei denen sich Wahrscheinlichkeiten aufgrund von Symmetrien als Quotienten von "günstigen" und "möglichen" Fällen berechnen lassen. Hier gibt es zu einer richtigen Hypothese keine sinnvollen Alternativen und damit verschwindet der hypothetische Charakter von Wahrscheinlichkeiten. Experimente benutzt man ausschließlich, um die Konvergenz relativer Häufigkeiten gegen diese (a priori bekannten) Wahrscheinlichkeiten zu demonstrieren. Genauso wenig wie der Begriff der Hypothese hat hier das in Alltagserfahrungen gewonnene, vielfach subjektivistisch geprägte Vorwissen einen Platz. Im Gegenteil, es wird durch das zu enge Laplacesche Konzept "erdrückt". Auch in der Fülle psychologischer Untersuchungen zur Genese des Wahrscheinlichkeitsbegriffs spiegelt sich diese Tatsache wider. Während man im Anschluß an Piagets bahnbrechende Untersuchungen lange Zeit ausschließlich die Entwicklung Laplacescher und kombinatorischer Konzepte studierte, hat in neuerer Zeit die Untersuchung alltäglicher, subjektivistisch geprägter Wahrscheinlichkeitsvorstellungen ein sehr viel reichhaltigeres Bild von den kindlichen Fähigkeiten entstehen lassen (Hawkins/Kapida 1984, Schrage 1980), die wir im folgenden ausnutzen werden. Fischbein faßt zusammen:

"An important theoretical conclusion may be drawn. Probabilistic thinking and proportional reasoning (damit ist das Laplacesche Konzept gemeint) are based on two distinct mental schemata. Certainly, probability computations may require ratio comparisons and calculations, but probability, as a specific mental attitude, does not necessarily imply a formal understanding of proportion concepts. This finding is in accordance with other results obtained before. Yost et al. (1962), Goldberg (1966) have shown that one can identify, even in preschool children, correct probability estimations while proportional reasoning is still deficient." (Fischbein 1982, S. 750.)

#### Frequentistischer Wahrscheinlichkeitsbegriff:

In der Unterrichtspraxis wird der Laplacesche Wahrscheinlichkeitsbegriff meist um den frequentistischen Aspekt bereichert, der Wahrscheinlichkeiten als Grenzwerte relativer Häufigkeiten in unendlich langen Versuchsserien beschreibt (Beispiel: Reißnagel). Wegen der Endlichkeit allen Experimentierens und des Wunsches diesen (in der Laplaceschen Welt eindeutig bestimmten) Grenzwert möglichst genau zu ermitteln, sind Zufallsschwankungen (so paradox das klingt) vom Lehrer emotional unerwünscht. Das folgende Zitat beschreibt leider die Realität recht genau:

"Die konkrete Durchführung der Zufallsexperimente zeigte deutlich ... der Lehrer muß merken, wann die Ergebnisse außerhalb jeglicher Erwartung, was durchaus vorkommen kann, liegen. Der Lehrer muß auf den zufälligen Ausgang der Zufallsexperimente angemessen vorbereitet werden<sup>1)</sup>." (Schröder 1981 S. 222.)

---

1) In diesem Zusammenhang spielt das Arcsin-Paradoxon (Kap. 8) eine wichtige Rolle: in einer langen Versuchsserie pendeln die relativen Häufigkeiten oft nur extrem selten um die Wahrscheinlichkeiten. Bei vielen kurzen Serien ist das anders.

Es drängt sich die Einsicht auf, daß es aus dem Blickwinkel der Statistik viel wichtiger wäre, ein Gefühl für die Schwankungen der relativen Häufigkeiten in endlichen Versuchsserien zu entwickeln ("schwaches Gesetz der großen Zahlen") als den Blick auf die Konvergenz in "unendlichen" Versuchsserien zu fixieren ("starkes Gesetz der großen Zahlen").

"Prognostischer" Wahrscheinlichkeitsbegriff:

Es zeigt sich, daß sowohl die Laplacesche Wahrscheinlichkeitsinterpretation (Wahrscheinlichkeit als Verhältnis der Anzahl günstiger zur Anzahl möglicher Fälle) als auch die frequentistische Deutung (Wahrscheinlichkeit als Grenzwert der relativen Häufigkeit) zu kurz greifen, wenn man nicht den "prognostischen Aspekt" herausarbeitet. Wahrscheinlichkeiten sind Prognosen relativer Häufigkeiten in zukünftigen langen Versuchsserien. Oder noch besser:

Wahrscheinlichkeiten sind Prognosen, um die die relativen Häufigkeiten in zukünftigen Versuchsserien schwanken werden.

Im Hinblick auf die Grundgedanken Beurteilender Statistik ist aber ebenso unverzichtbar, daß man die Möglichkeit verschiedener (manchmal gleichberechtigter, mitunter aber auch verschieden glaubwürdiger) Wahrscheinlichkeiten und damit deren hypothetischen Charakter in der Vorstellungswelt der Schüler verankert.

Wir halten fest: Ein für die Statistik tragfähiger Wahrscheinlichkeitsbegriff muß

- (1) prognostischen Charakter haben, d. h. es muß deutlich werden, daß Wahrscheinlichkeiten Prognosen darstellen, um die die relativen Häufigkeiten schwanken werden.
- (2) Er muß hypothetischen Charakter haben, d. h. Wahrscheinlichkeitsannahmen, die sich in Versuchen nicht bewähren, müssen verworfen und durch bessere ersetzt werden.
- (3) Sollten aus lernpsychologischen Gründen subjektive Aspekte nicht ausgeschlossen werden. Aussagen der folgenden Form müssen erlaubt sein: "Nach Lage der Versuchsergebnisse halte ich Hypothese A für glaubwürdiger als Hypothese B."

Solche Aussagen bereiten intuitiv Ideen vor, die dem Testen von Hypothesen zugrunde liegen.

## 1.2 Hypothesen aufstellen

### 1.2.1 Hypothesen schätzen

Die Attribute "prognostisch", "hypothetisch" und "subjektivistisch" mögen anspruchsvoll klingen, wenn man sie mit dem Wahrscheinlichkeitsbegriff auf der Sekundarstufe I verbinden möchte. Sie sind es nicht, wenn man sie an konkretem Material "festmacht" und ihre Bedeutung mit den Kindern aushandelt. Statt gleich zu Beginn formal zu erklären, was Wahrscheinlichkeiten "sind", lernt das Kind, wie man mit Wahrscheinlichkeiten und Wahrscheinlichkeitshypothesen umgeht. Wir müssen "wegkommen von direkter Unterweisung in der herkömmlichen Art und hinkommen zu interaktiven Prozessen." (Weinert/Kluve 1984, S. 102). "Der Lehrer erzählt dem Kind nicht, was es zu tun hat, sondern er beginnt eine Interaktion, in der beide, Kind und Lehrer für die Bewältigung der Aufgabe verantwortlich sind."

Diese Sicht der Einübung in neue Begriffe, die erfahrene Lehrer intuitiv stets befolgen, wurde durch Interaktionsanalysen (Bauersfeld/Voigt) auch empirisch belegt. Im Hinblick auf eine am Handeln orientierte Genese des von uns angestrebten Wahrscheinlichkeitsbegriffs haben sich Quader und andere Zufallsobjekte mit partiellen Symmetrien ausgezeichnet bewährt. Einerseits lassen sie im Gegensatz zu Würfeln oder Glücksrädern in gewissen Grenzen mehrere sinnvolle Wahrscheinlichkeitsannahmen zu. Der hypothetische Charakter von Wahrscheinlichkeiten wird sinnfällig. Andererseits wird (im Gegensatz zum Reißnagel) der Unterschied zwischen Wahrscheinlichkeiten und relativen Häufigkeiten greifbar - alle akzeptablen Wahrscheinlichkeitshypothesen sind partiell symmetrisch, die relativen Häufigkeiten nur in seltenen Ausnahmefällen. Der prognostische Aspekt, der das Verhalten auf lange Sicht beschreibt, prägt sich ein.

Primarstufe, Orientierungsstufe (bis ca. 12 Jahre):

So lassen sich schon in der Primarstufe und am Anfang der Sekundarstufe I bekannte Würfelspiele wie "Fang den Hut" oder "Mensch ärgere Dich nicht" modifizieren, indem man neben dem normalen Laplace-Würfel auch verschieden beschriftete Holzquader bereitstellt (Abb. 1-1 bis 1-3).

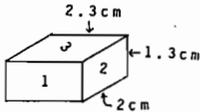


Abb. 1-1

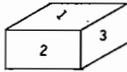


Abb. 1-2

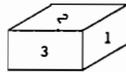


Abb. 1-3

Vor jedem Wurf darf der Spieler den "Würfel" wählen, der in der momentanen Spielsituation den größten Erfolg verspricht. Ohne Zwang zur Verbalisierung handeln Kinder in dieser stochastischen Situation sinnvoll. Sehr schnell erkennen sie, daß sich die Ergebnisse bei keinem der Quader vorhersagen lassen (Zufallsbegriff). Sie treffen aber ihre Wahl so, daß gewünschte Zahlen auf möglichst großen, unerwünschte auf möglichst kleinen Seiten liegen. Sie besitzen klare Vorstellungen von ihren Chancen. Diese Beobachtung wird gestützt durch neuere Untersuchungen, die belegen, daß auch Primarstufenkinder über (subjektivistisch geprägte) Wahrscheinlichkeitsvorstellungen verfügen - lange bevor sie das Stadium der formalen Operationen erreicht haben. Die im Anschluß an Piaget/Inhelder (1951) lange Zeit vorherrschende "Lehrmeinung", daß ein Wahrscheinlichkeitsbegriff erst mit diesem Stadium erworben werden kann, gründet darin, daß in den Untersuchungen fast ausschließlich mit "Laplaceschem Gedankengut" gearbeitet wurde, das den Verhältnissbegriff und die Verfügbarkeit kombinatorischer Aspekte erfordert.

Sekundarstufe I

Kinder, die (ab Klassenstufe 6/7) mit der Prozentrechnung vertraut sind, quantifizieren ihre subjektiven

Erwartungen ("Hypothesen") über das Verhalten der Quader gerne durch "prozentuale Angaben".

"Stelle Dir vor, Du würfelst mit dem Quader (Abb. 1-1) sehr sehr oft. Was schätzt Du: in wieviel Prozent aller Fälle wirst Du auf lange Sicht die einzelnen Augenzahlen von 1 bis 6 erhalten?" Hier einige Antworten (Klasse 7):

1	2	3	4	5	6		Glaubwürdigkeit
10%	5%	35%	35%	5%	10%	René	32%
15%	10%	25%	25%	10%	15%	Stefan	18%
10%	12%	35%	20%	15%	8%	Alexandra	0%
15%	15%	25%	25%	15%	15%	Joanna	0%
15%	5%	30%	30%	5%	15%	Jasmin	50%

Tab. 1-4 fünf verschieden glaubwürdige Hypothesen über den Quader aus Abb. 1-1

Alexandra handelt sich heftige Kritik ein. Offensichtlich dachte sie (wegen der Unsymmetrie) nur an eine ganz spezielle Würfelserie, nicht aber an eine Schätzung auf lange Sicht. Auch mit Joanna sind die Klassenkameraden nicht einverstanden, da die instabileren Lagen 1 und 6 seltener erwartet werden als die etwas stabileren 2 und 5. Übrig bleiben die Verteilungen von René, Stefan und Jasmin. Sie drücken Erwartungen über das Verhalten der Quader aus, die zunächst stark subjektiven Charakter haben. Wir nennen sie Wahrscheinlichkeiten oder Wahrscheinlichkeitshypothesen. Während diese Hypothesen als Schätzungen auf lange Sicht die Symmetrien des Quaders widerspiegeln, erwarten wir für die Versuchsergebnisse nur in Ausnahmefällen exakt symmetrische Ergebnisse. Wir erwarten, daß sie um symmetrische Zahlen schwanken.

Bei einer Abstimmung in der Klasse zeigt sich (noch vor jeglichem Experiment), daß die Hypothesen von Jasmin und René am glaubwürdigsten sind. Diejenigen von Alexandra und Joanna werden "verworfen" (vgl. letzte Spalte).

Nach diesen "Spekulationen" werden Versuche mit Spannung erwartet. Wer hat am besten geschätzt? Wessen Hypothese wird

einer Überprüfung standhalten? Hier die Versuchsergebnisse (je 100 Versuche):

1	2	3	4	5	6	
13	10	32	27	5	13	Jola
12	3	34	36	8	7	Sabine I
11	4	41	35	5	4	Sabine II
5	5	36	33	8	13	Martina
15	3	28	32	6	16	Silke
10	6	42	31	4	7	Stefanie I
11	5	33	34	8	9	Jasmin
19	6	30	34	5	6	Stefan I
16	2	46	29	4	3	Stefan II
9	7	28	33	5	18	Denise
9	5	32	37	7	10	Stefanie II
8	5	39	36	5	7	Robert
14	7	31	36	4	8	Marc
13	3	34	28	6	16	Jens
13	5	31	25	10	16	Thao
10	9	30	30	7	14	Roman
12	8	30	34	6	10	Joanna
10	7	28	38	7	10	Fehim

Tab. 1-5 Versuchsergebnisse zum Quader aus Abb. 1-1

Die Tabelle belegt: die Hypothesen von René und Jasmin waren beide recht brauchbar. Sie läßt darüber hinaus ein Gefühl für die Größe der Schwankungen (100 Versuche) entstehen. Von sich aus fassen Schüler die Häufigkeiten zunächst gruppenweise zusammen, um eine Entscheidung zwischen René und Jasmin herbeizuführen: Die relativen Häufigkeitsverteilungen aus je 500 (300) Versuchen sehen wie folgt aus (die Schwankungen werden geringer):

1	2	3	4	5	6	
11.2%	5.0%	34.2%	32.6%	6.4%	10.6%	(500)
13.0%	5.2%	35.8%	32.2%	5.2%	8.6%	(500)
11.4%	5.0%	33.4%	32.4%	6.4%	11.4%	(500)
10.7%	8.0%	29.3%	34.0%	6.7%	11.3%	(300)

Tab. 1-6

Nach einer Gegenüberstellung mit den Vorhersagen entschließt man sich, die Hypothesen der beiden Kontrahenten zu verwerfen und einigt sich auf eine Hypothese:

11%	6%	33%	33%	6%	11%	,	(*)
-----	----	-----	-----	----	-----	---	-----

die in der Tat "zwischen" René und Jasmin liegt. Aber auch andere (symmetrische) Hypothesen, die hiervon nicht allzu sehr abweichen, wären akzeptabel, beispielsweise:

11.5%	6%	32.5%	32.5%	6%	11.5%	.
-------	----	-------	-------	----	-------	---

Alle 1800 Versuche zusammen liefern die relative Häufigkeitsverteilung

11.7%	5.6%	33.6%	32.7%	6.1%	10.4%	,
-------	------	-------	-------	------	-------	---

die mit den eben formulierten Hypothesen intuitiv auch recht gut vereinbar ist. Dennoch ist uns die Information, die in den Teilversuchen mit dem Umfang 500 (300) steckt, ungleich wertvoller als die eine Information über die 1800 Versuche, da sie hilft, das für statistisches Denken unerläßliche Gefühl für Zufallsschwankungen zu entwickeln.

Wir halten fest: Wahrscheinlichkeiten lassen sich (auch durch noch so viele Versuche) nicht eindeutig bestimmen; es gibt meist mehrere brauchbare Hypothesen, um die die relativen Häufigkeiten von Versuch zu Versuch schwanken.

### 1.2.2 Hypothesen berechnen

Schüler "lassen an dieser Stelle nicht locker". "Kann man die Wahrscheinlichkeiten der Quader berechnen?" Solche Fragen verraten Neugier und sind uns höchst willkommen, bieten sie doch die Chance, auf dem Niveau der Sekundarstufe I ein wenig zu "forschen", den statistischen Denkansatz weiterzuentwickeln und den hypothetischen Charakter von Wahrscheinlichkeiten greifbar zu machen. "Berechnen" wir also die Wahrscheinlichkeiten der Quader:

#### Hypothese Ia

Nicole konstatiert völlig richtig: "je größer die Flächen des Quaders (Abb. 1-1) desto häufiger fallen die zugehörigen Augenzahlen". Sind die Wahrscheinlichkeiten  $P$  proportional zu den Flächen  $F$ ?

Seiten 1 und 6:  $2.99 \text{ cm}^2$

Seiten 2 und 5:  $2.60 \text{ cm}^2$

Seiten 3 und 4:  $4.60 \text{ cm}^2$

Die gesamte Oberfläche des Quaders ist  $F=20.38 \text{ cm}^2$  (=100%).  
Damit erhalten wir die Hypothese:

		(1800 Versuche)
$P(1) = P(6) = 2.99/20.38 = 14.7\%$		(11.7% 10.4%)
$P(2) = P(3) = 2.6 / 20.38 = 12.8\%$		( 5.6% 6.1%)
$P(3) = P(4) = 4.6 / 20.38 = 22.5\%$		(33.6% 32.7%)

Zwar wachsen die Wahrscheinlichkeiten mit der Größe der Grundfläche, doch sagen sie unsere experimentellen Ergebnisse (rechte Spalte) noch sehr viel schlechter voraus als unsere intuitiven Schätzungen. Trotz der schönen Rechnung müssen wir Nicoles Hypothese verwerfen.

## Hypothese Ib

Je mehr der Quader "hochkant" steht, desto instabiler ist seine Lage, desto seltener kommen die zugehörigen Augenzahlen. Sollten wir es mit einer umgekehrten Proportionalität zwischen Schwerpunkthöhe  $h$  und Wahrscheinlichkeit versuchen?

	$h$	$1/h$
Seiten 1 und 6:	1	1
Seiten 2 und 5:	1.15	0.87
Seiten 3 und 4:	0.65	1.538

Mit  $S = 2(1+0.87+1.538) = 6.816$  ergibt sich

$$P(1) = P(6) = 1.000/S = 14.7\%$$

$$P(2) = P(3) = 0.870/S = 12.8\%$$

$$P(3) = P(4) = 0.538/S = 22.5\%$$

"Schon wieder die gleiche Hypothese wie oben? Na klar, das Produkt aus Grundfläche und Schwerpunkthöhe ergibt immer das halbe Quadervolumen." Wenn die Wahrscheinlichkeiten zu den Grundflächen proportional sind, dann sind sie zu den Schwerpunkthöhen automatisch umgekehrt proportional. Allerdings läßt sich Hypothese Ib im Gegensatz zu Ia auf gezinkte Quader verallgemeinern.

## Hypothese II

Versuchen wir, Nicoles Hypothese I zu verbessern: Wie wir sehen, sind (verglichen mit den experimentellen Ergebnissen) die Unterschiede zwischen den berechneten Wahrscheinlichkeiten  $P$  zu klein, wenn wir sie zu den Flächen  $F$  proportional machen. Sie werden größer, wenn wir die Stabilität nicht durch die Grundfläche  $F$  sondern durch das Quadrat oder höhere Potenzen von  $F$  messen und die Wahrscheinlichkeiten proportional zu diesen Potenzen einrichten. Wir erhalten:

	F	F <sup>2</sup>	F <sup>3</sup>	F <sup>4</sup>	F <sup>5</sup>
Seiten 1 und 6	2.99	8.94	26.73	79.93	238.98
Seiten 2 und 5	2.60	6.76	17.58	45.70	118.81
Seiten 3 und 4	4.60	21.16	97.34	447.75	2059.63
2 · Summe	20.38	73.72	283.29	1146.76	4834.84

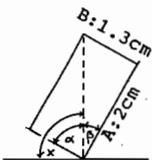
Daraus ergeben sich folgende Hypothesen:

	(1800)					
P(1) = P(6)	14.7%	12.1%	9.4%	7.0%	5.0%	(11.7% 10.4%)
P(2) = P(3)	12.8%	9.2%	6.2%	4.0%	2.5%	( 5.6% 6.1%)
P(3) = P(4)	22.5%	28.7%	34.4%	39.0%	42.5%	(33.6% 32.7%)

Tatsächlich werden die Wahrscheinlichkeiten für die stabilen Seiten immer größer, die für die kleinen Seiten immer kleiner, je mehr wir potenzieren. Die Hypothese, die wir aus einer Proportionalitätsannahme zwischen Wahrscheinlichkeit und dritter Potenz der Grundfläche erhalten, paßt noch am besten zu unseren experimentellen Ergebnissen (letzte Spalte). Aber ein Vergleich mit Tab. 1-6 zeigt, daß die relativen Häufigkeiten nicht um diese Wahrscheinlichkeiten schwanken. Wir müssen wohl auch die Hypothese II fallenlassen. (Wir greifen dieses Beispiel in den Abschnitten 2.1 und 5.3 auf.)

### Hypothese III

Wir geben noch nicht auf und machen unser Quaderproblem zweidimensional, indem wir die instabilen Seiten 2 und 5 völlig außer acht lassen. (Wir würfeln einfach weiter, wenn wir solche Zahlen erhalten.) Der Quader kann dann die stabile Lage A (3 oder 4) oder die



$$\alpha = 57.0^\circ$$

$$\beta = 33.0^\circ$$

Abb. 1-7

Lage B (1 oder 6) annehmen.

Abb. 1-7 zeigt ihn im Aufriß.

Wir machen die Modellannahme, daß der Quader stets mit der Kante aufschlägt und der Aufschlagwinkel  $x$  ( $\alpha \leq x \leq 90^\circ + \alpha$ ) über die Endlage A oder B entschei-

det. Gilt  $x > 90^\circ$ , so wird der Quader in Stellung A kippen, sonst in Stellung B. Wenn alle Winkel  $x$  gleiche Wahrscheinlichkeit besitzen (weitere plausible Modellannahme), dann müßten sich die Wahrscheinlichkeiten  $P(B)$  und  $P(A)$  zueinander verhalten wie die Winkel  $\beta = 33^\circ$  und  $\alpha = 57^\circ$ . Wir erhalten die Wahrscheinlichkeitshypothese:  $P(A) = 63.3\%$ ,  $P(B) = 36.7\%$ . In unseren 1800 Versuchen trat die Lage A (3;4) 1193mal, die Lage B (1;6) 397mal auf. Da wir die Lage (2;5) nicht zulassen, sind das 75.0% für A und 25.0% für B. Diese Ergebnisse sind mit unserer Hypothese intuitiv nicht vereinbar. Wir müssen sie zusammen mit den zugrunde liegenden Modellannahmen als unrealistisch verwerfen. (Dies gilt übrigens auch, wenn man statt eines Quaders einen langen Stab mit rechteckigem Querschnitt rollt, also tatsächlich ein zweidimensionales Problem untersucht.)

Es gibt auf dem Niveau der Sekundarstufe I nicht viele Probleme, die den hypothetischen Charakter von Wahrscheinlichkeiten und den engen Zusammenhang mit dem Prozeß des Modellbildens ähnlich prägnant hervortreten lassen wie die oben genannten Beispiele.

### 1.3 Stellung der Wahrscheinlichkeitsrechnung

Bisher haben wir Hypothesen aufgestellt und diese durch Vergleich mit experimentellen Ergebnissen verworfen, modifiziert oder beibehalten. Da wir unserem Ziel gemäß einen konsistenten Aufbau von Wahrscheinlichkeitsrechnung und Statistik anstreben, wird es bei der Wahrscheinlichkeitsrechnung nicht mehr um eine "voraussetzungslose" "exakte" Berechnung von Wahrscheinlichkeiten gehen. Vielmehr werden wir aus Wahrscheinlichkeitshypothesen (Primärhypothesen) andere Hypothesen (Sekundärhypothesen) ableiten, die uns interessierende Fragen beantworten. Die zugehörigen Prognosen werden sich wiederum in Experimenten zu bewähren haben. Auch die funktionalen Abhängigkeiten zwischen Primär- und

Sekundärhypothesen werden wir gemäß der Frage studieren: "Wie wirken sich Modifikationen der Primärhypothesen auf die Sekundärhypothesen aus?" Auf diese Weise wächst ein Gefühl für die "Stabilität" stochastischer Modelle.

### 1.3.1 Wartezeiten (Beispiel 1)

Wir werfen unseren Quader (Abb. 1-1) so lange bis die erste "3" erscheint. Wie lange müssen wir auf lange Sicht warten? Hier die spontan geäußerten Hypothesen dreier Schüler:

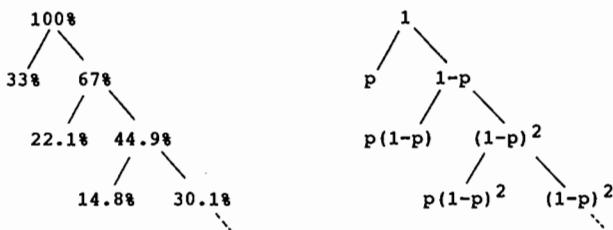
1x	2x	3x	4x	5x	6x	7x	8x	9x	≥10x	
5%	15%	25%	15%	10%	8%	7%	6%	5%	4%	Sabine
10%	10%	10%	10%	10%	10%	10%	10%	10%	10%	Jens
21%	17%	14%	12%	10%	8%	6%	5%	4%	3%	Joanna

Tab. 1-8 Hypothesen über die Wartezeitenverteilung

René schlug eine Verteilung vor, deren Wahrscheinlichkeiten mit wachsender Versuchszahl wachsen, zog diesen Vorschlag aber schnell wieder zurück - worauf Joanna eine monoton fallende Verteilung vorschlug und die einzelnen Wahrscheinlichkeiten so wählte, daß sich insgesamt 100% ergaben. Wieder haben wir eine Fülle von Hypothesen, eine "prickelnde" Erwartungshaltung, den "statistischen Spannungsbogen" zwischen Vorhersage und Experiment.

Denise: "Die Hypothesen sind alle falsch! Bei 1x muß 33% stehen! In 33% aller Fälle erhalte ich sofort die "3". In 67% muß ich weiterwürfeln." (Sie erinnert sich an die bewährte Hypothese (\*) mit  $P(3)=33\%$ .)

Von hier aus ist es nur noch ein kleiner Schritt zum Baumdiagramm (Taschenrechner, Zahlen gerundet; p bezeichnet die Trefferwahrscheinlichkeit):



Es ergibt sich  $P(\text{Wartezeit}=n)=p(1-p)^{n-1}$ . Mit Denise ( $p=0.33$ ) erhalten wir:

1x	2x	3x	4x	5x	6x	7x	8x	9x	>10x
33.0%	22.1%	14.8%	9.9%	6.6%	4.5%	3.0%	2.0%	1.3%	2.7%

Wir unterteilen unsere 1800 Würfelresultate in 180 "Zehnerpäckchen" und zählen aus, beim wievielten Wurf erstmalig die "3" erschien. Hier die mit Spannung erwarteten Ergebnisse:

34.3%	21.8%	16.8%	12.9%	4.6%	3.9%	2.1%	2.1%	0%	1.4%
-------	-------	-------	-------	------	------	------	------	----	------

Tatsächlich schwanken die relativen Häufigkeiten um die "nach Denise" vorausgesagten Werte. Die Hypothesen von Jens und Sabine müssen wir verwerfen, diejenige von Joanna hat sich wenigstens qualitativ bestätigt. Die Wahrscheinlichkeiten werden mit wachsender Wartezeit kleiner.

Ergebnis: Die Hypothese von Denise stimmt! Oder doch nicht? Was wäre, wenn wir für "3" nicht die Wahrscheinlichkeit  $p=33\%$  sondern  $32\%$ ,  $30\%$  oder  $35\%$  angenommen hätten? (Auch diese Hypothesen waren von den Schülern vorgeschlagen worden.) Führen wir drei Modellrechnungen durch:

1x	2x	3x	4x	5x	6x	7x	8x	9x	>10x
30.0%	21.0%	14.7%	10.3%	7.2%	5.0%	3.5%	2.5%	1.7%	4.0%
32.0%	21.8%	14.8%	10.1%	6.8%	4.7%	3.2%	2.3%	1.5%	3.1%
35.0%	22.8%	14.8%	9.6%	6.2%	4.1%	2.6%	1.7%	1.1%	2.1%

Tab. 1-9 Funktionale Abhängigkeit der Hypothesen bei Variation der Trefferwahrscheinlichkeit

Auch diese Wartezeitenverteilungen wären mit unseren experimentellen Ergebnissen gefühlsmäßig vereinbar. Wegen des geringen Versuchsumfangs 180 können wir keine verwerfen. An der Struktur (exponentielles Verhalten)  $P(\text{Wartezeit}=n) = p(1-p)^{n-1}$ , die sich zwangsläufig aus der Einsicht in das Baumdiagramm ergibt, besteht jedoch kein Zweifel. Untersuchungen anderer Quaderseiten, der Vergleich mit der Wartezeitenverteilung beim Laplace-Würfel, das Studium der mittleren Wartezeit und ihrer Schwankungen um den Erwartungswert (=Kehrwert der Trefferwahrscheinlichkeit) bieten sich als mögliche Vertiefungen an.

### 1.3.2 Quader contra Würfel (Beispiel 2)

Auch "klassische Einführungsaufgaben" zu Pfad- und Additionsregel gewinnen an Reiz, wenn man neben Glücksrädern und Laplace-Würfeln Quader benutzt und zu vergleichenden Prognosen auffordert. Wieder wird die Abhängigkeit der Antworten von den zugrunde liegenden Modellhypothesen ins Bewußtsein gerückt. Alle abgeleiteten Hypothesen lassen sich durch erneutes Befragen der Urlisten mit wenig Aufwand experimentell überprüfen.

(a) Du hast sechs Würfe. Beim ersten Wurf darf keine 1, beim zweiten keine 2, ..., beim sechsten keine 6 auftreten. Sind die Chancen für einen solchen "Durchmarsch" beim Quader oder beim Würfel größer? Antwort:

$$\text{Würfel: } P(\text{Durchmarsch}) = (5/6)^6 = 0.335$$

$$\text{Quader: } P(\text{Durchmarsch}) = 0.89^2 \cdot 0.94^2 \cdot 0.67^2 = 0.314$$

(wenn man Hypothese (\*) verwendet)

$$= 0.9^2 \cdot 0.95^2 \cdot 0.65^2 = 0.308$$

(wenn man Renés Hypothese verwendet)

Der Würfel verspricht die größeren Chancen.

(b) Möchte man einen Pasch (zwei gleiche Zahlen) würfeln, so erweist sich der Quader als günstiger. Hier gilt  $P(\text{Pasch}) = 2(0.33^2 + 0.11^2 + 0.06^2) = 24.92\%$ , falls wir die Hypothese (\*) verwenden. Verwenden wir René's Hypothese, so ergibt sich  $P(\text{Pasch}) = 2(0.35^2 + 0.10^2 + 0.05^2) = 27\%$ . Der normale Würfel liefert dagegen nur  $P(\text{Pasch}) = 1/6 = 16.66\%$ .

(c) Braucht man eine 10, so erweist sich dagegen der Laplace-Würfel mit  $P(10) = P(6;4) + P(5;5) + P(4;6) = 3/36 = 8.33\%$  als günstiger. Der Quader liefert die 10 nur mit der Wahrscheinlichkeit  $P(10) = 2 \cdot 0.11 \cdot 0.33 + 0.06^2 = 7.62\%$  (Hypothese (\*)) bzw.  $P(10) = 2 \cdot 0.10 \cdot 0.35 + 0.05^2 = 7.25\%$  (Hypothese von René).

(d) Möchte man beim Würfeln besonders schnell vorwärts kommen, sind Quader und Würfel gleich schlecht (Erwartungswert 3.5). Man sollte den Quader anders beschriften und auf die Regel "Summe der Gegenseiten ist 7" verzichten.

### 1.3.3 Binomialverteilungen als Hypothesen (Beispiel 3)

Mehrfachzüge aus Urnen gehören zu den Standardthemen der Wahrscheinlichkeitsrechnung. "Wir ziehen aus einer Urne mit  $1/4$  roten Kugeln. Wie groß ist die Wahrscheinlichkeit, daß ... ?" Wieviel spannender sind solche Fragestellungen unter statistischen Aspekten, wenn man etwa Urnen mit unbekanntem Inhalt erforschen möchte: "In dieser Kiste sind 5 Kugeln. Wieviele davon sind rot?" Wir ziehen 4 mal mit Zurücklegen: 2 rote - wieder 2 rote - eine rote - keine rote ... 20 Vierfach-Experimente lieferten folgende Anzahlen roter Kugeln:

0	1	2	3	4
7	8	3	2	0

Sicher enthält die Urne weniger rote als weiße Kugeln. Stefanie tippt auf eine rote, Nicole auf 2 rote (Hypothesen:  $p=0.2$  und  $p=0.4$ ) Diese Hypothesen führen zu folgenden Binomialverteilungen:

k	0	1	2	3	4
$B(4;0.2;k)$	40.96%	40.96%	15.36%	2.25%	0.26%
$B(4;0.4;k)$	12.96%	34.56%	34.56%	15.36%	2.56%

Wegen der höheren Häufigkeiten für 0 und 1 scheinen die experimentellen Ergebnisse eher für Stefanies Hypothese zu sprechen. Vielleicht hatten wir nur Pech und Nicoles Hypothese stimmt doch? Wir tragen die Ergebnisse zusammen und erhalten für 100 Vierfachzüge:

0	1	2	3	4	
44	39	14	2	1	(100)

Stefanie hat recht! Ihre Hypothese sagt die experimentellen Ergebnisse wesentlich besser voraus. Nicoles Hypothese müssen wir verwerfen. (Das kann man natürlich auch so begründen: Die 100 Vierfachzüge sind 400 Einfachzüge, in denen insgesamt 77 mal (19.25%) eine rote Kugel gezogen wurde. Nur eine von den fünf Kugeln kann rot gewesen sein.) Wer könnte der Versuchung widerstehen, jetzt einen Blick in die Urne zu riskieren?

#### 1.4 Hypothesen als Modelle von der Wirklichkeit

Der Leser erkennt an den genannten Beispielen unschwer, wie die Gegenüberstellung Laplacescher und nicht Laplacescher Objekte den Reiz der "klassischen Wahrscheinlichkeitsrechnung" erhöht. Wir benutzen kombinatorische Zählprinzipien, Pfadregel und Summenregel unter verändertem Blickwinkel. Sie dienen nicht mehr zur "genauen Berechnung von Wahrscheinlichkeiten", sondern (in statistischer Sicht) zur Herleitung von Hypothesen aus einfacheren Hypothesen. Diese Blickpunktverschiebung ist unter unserer Zielperspektive einer Integration statistischer Gedanken in die elementare Wahrscheinlichkeitsrechnung grundlegend. Hätten wir ausschließlich Laplace-Würfel statt der Quader verwendet, so wäre niemand auf die Idee gekommen, beispielsweise die Wartezeit-

tenverteilung für verschiedene Trefferwahrscheinlichkeiten zu berechnen, die Ergebnisse miteinander zu vergleichen und so den hypothetischen Charakter präsent zu halten. Mathematik - und insbesondere Stochastik - betreiben heißt eben nicht: absolute Wahrheiten über die Wirklichkeit ableiten, sondern in sich stimmige Modelle von der Wirklichkeit erarbeiten. Die Vernachlässigung dieses Aspektes hat - gerade die Schulmathematik - etwas in Verruf gebracht. Schupp (1982) formuliert:

"Daß ein Modell nicht aus der Wirklichkeit abstrahiert, sondern an sie herangetragen wird (Schreiber (1980) spricht von Ideation), daß somit jede mittels des Modells gewonnene Aussage über Realität in gewisser Weise hypothetisch ist, gilt grundsätzlich für jede Modellbildung. Doch scheint der stochastische Regelkreis wegen der jedermann ersichtlichen Schwierigkeiten beim Aufstellen der Primärwahrscheinlichkeiten besonders gut geeignet, erstmals auf diese prinzipielle Grenze der Wirklichkeitserfassung (durch Mathematik) hinzuweisen."

#### 1.5 Lernpsychologischer Rückblick

Vergleichen wir den Aufforderungsgehalt, den unsere Quader (und in noch stärkerem Maße die im folgenden Kapitel benutzten Objekte mit partiellen Symmetrien) im Hinblick auf eine adäquate Begriffsbildung bieten, mit den Möglichkeiten der "klassischen" Zufallsobjekte, so fällt auf:

(a) Laplacesche Objekte (mit vollständige Symmetrien) lassen nur eine Wahrscheinlichkeitsverteilung zu. Niemand bezweifelt ernsthaft, daß bei einer Münze Kopf und Zahl, bei einem Würfel die 6 Augenzahlen die gleiche Wahrscheinlichkeit besitzen, d. h. auf lange Sicht gleich häufig auftreten werden. Damit verschwindet die fundamentale Idee der Statistik, daß es für eine Situation verschiedene Modelle geben kann, daß Wahrscheinlichkeiten stets hypothetischen Charakter haben, nur Annahmen darstellen, die mitunter auch ver-

worfen werden müssen. Bei Verwendung teilweise symmetrischer Objekte drängt sich diese Idee von selbst auf. Andererseits verschwindet

(b) bei völlig unsymmetrischen Objekten wie dem Reißnagel der begriffliche Unterschied zwischen relativen Häufigkeiten und Wahrscheinlichkeiten. Wahrscheinlichkeiten sind Schätzungen der relativen Häufigkeiten auf lange Sicht. Wenn man bei 1000 Würfeln 692 mal die Lage "1" erhielt, warum sollte man durch eine von 0.692 verschiedene Zahl diese Prognose möglicherweise willkürlich verschlechtern? Muß damit die relative Häufigkeit in den Augen der Schüler nicht notwendig zur Wahrscheinlichkeit werden? (Man vergleiche das zweite der untenstehenden Transkripte). Die beiden Begriffe, die hier durch gleiche Zahlen beschrieben werden, liegen so dicht beieinander, daß eine Unterscheidung viele Schüler überfordert. Bei teilweise symmetrischen Objekten dagegen ist der Unterschied offensichtlich: Wahrscheinlichkeiten spiegeln die Symmetrien wider, relative Häufigkeiten nur in seltenen Ausnahmefällen, die in der Tat höchstes Erstaunen hervorrufen.

(c) Auch unter lernpsychologischen Gesichtspunkten spricht einiges für die partiell symmetrischen Objekte:

- Da man aufgrund der Primärintuition: "mit wachsender Stabilität wachsen die Wahrscheinlichkeiten" einen Anhaltspunkt für Schätzungen hat, ist man bereit, quantitative Prognosen auf lange Sicht abzugeben (1.2.1). Es entsteht der in der Einleitung paradigmatisch erwähnte "Spannungsbogen zwischen Erwartung und Experiment". Beim Würfel dagegen erwarten alle das gleiche, beim Reißnagel erwartet man - ehrlich gesagt - gar nichts (vgl. die untenstehenden Unterrichtsszenen).

- Wie oben beschrieben, nutzen wir mit Hilfe der Quader das aus, was Schüler auf natürliche Weise (spielerisch) von sich aus tun. Sie erfahren handelnd, wie man mit Wahrscheinlich-

keiten umgeht, wie sie mit relativen Häufigkeiten zusammenhängen, was man unter Hypothesen versteht, was es bedeutet, sie zu verwerfen... Durch Handlungen voneinander klar abgegrenzte Begriffe werden schließlich nur noch mit Namen versehen. Ein solches Vorgehen ist gemeint, wenn Bauersfeld vom "Aushandeln der Begriffe" spricht: Mathematische Begriffe werden eingeführt zur Strukturierung dessen, was Schüler tun - und nicht als Selbstzweck oder aus formallogischen Gründen.

Zur Konkretisierung unserer Ausführungen beschließen wir diesen Abschnitt durch Unterrichtsszenen aus Einführungsstunden in die Wahrscheinlichkeitsrechnung mit Hilfe der klassischen Objekte Würfel und Reißnagel. Bei diesen Unterrichtsszenen handelt es sich um Auszüge aus Transkripten von Videoaufzeichnungen, die zum Zwecke der Interaktionsanalyse angefertigt wurden.

#### 1. Szene:

(Einführung in die Wahrscheinlichkeitsrechnung, 7. Klasse, Realschule: Bauersfeld/Voigt 1986)

Der Lehrer notiert die Ergebnisse von 100 (Laplace-) Würfelversuchen an der Tafel

1	2	3	4	5	6
15	16	19	14	19	17

und fragt: "Was fällt auf?"

Den Schülern fällt nichts auf. Vermutlich wird auch dem Leser nichts auffallen, weil die Ergebnisse recht genau in das Bild hineinpassen, das man sich gemeinhin von einem normalen Würfel macht. Der Lehrer strebt die Formulierung an, daß die gewürfelten Zahlen "vom Zufall abhängen", aber das ist für die Schüler nichts auffälliges. So kommen Antworten wie:

Martina: "Alle Zahlen liegen über 10."

Michael: "Es ist selbstverständlich, daß die Ergebnisse unterschiedlich sind."

Lehrer: "Warum hast Du erwartet, daß nicht alle Ergebnisse gleich sind?"

Michael: "100 geht nicht durch 6."  
(Lehrer wirkt konsterniert.)

Nach einem längeren Frage- und Antwortspiel gelangt man zu der (vom Lehrer) angestrebten Formulierung, daß die Würfel-ergebnisse "vom Zufall" abhängen. Statt vor dem Experiment beispielsweise eine Vermutung über die Größe der von den Schülern erwarteten Schwankungen zu formulieren, die Vermutungen einer Serie von Experimenten gegenüberzustellen und so einen "statistischen Spannungsbogen" aufzubauen, wurden langweilige Selbstverständlichkeiten zum Unterrichtsthema. Die experimentellen Daten beantworten keine Frage, sie führen zu keinen Konsequenzen. Die Chance, Statistik zu erleben, wurde vertan.

## 2. Szene:

Klasse 8, Gymnasium, 2. Stunde einer Einführung in die Wahrscheinlichkeitsrechnung. Vorausgegangen war eine Stunde, in der man sich mit Laplace-Versuchen beschäftigte. (Voigt 1983 S. 24 ff.)

Diese Szene veranschaulicht die Probleme mit der Begriffsbildung, wenn man völlig unsymmetrische Objekte wie die Reißzwecke benutzt:

An der Tafel stehen (aufsummierte) relative Häufigkeiten für das Ergebnis "1" beim Werfen einer Reißzwecke.

Lehrer: "54%, 60%, 54%, 51%, 55%, 60% ... sind das Wahrscheinlichkeiten?"

Schüler: "weiß nicht".

Lehrer: "Was sind denn das für Dinge?"

Schüler: "Wahrscheinlichkeiten"

... Einige Minuten später...

Schüler: "Die Wahrscheinlichkeit hängt vom Zufall ab."

Lehrer: "Die Wahrscheinlichkeit hängt nicht so sehr vom Zufall ab. Die Versuchsausgänge hängen vom Zufall ab."

Lehrer: Was hat jetzt die relative Häufigkeit mit Wahrscheinlichkeit zu tun?

Schüler: Erklärt, wie man die relative Häufigkeit ausrechnet.

Lehrer: "Ja, aber ist jetzt Wahrscheinlichkeit dasselbe wie relative Häufigkeit?"

Stefan: Ich würde sagen ja. Von der Herkunft würde ich sagen ja. Nur relative Häufigkeit ist im Vergleich zur Wahrscheinlichkeit ungenauer."

Lehrer: (deutet auf eine relative Häufigkeit): "Bist Du denn sicher, daß das die richtige Wahrscheinlichkeit ist?"

Stefan: "Nein. Die kann man gar nicht feststellen"

Lehrer: "Die kann man gar nicht feststellen. So etwas gibt es, aber feststellen können wir sie nicht. Was können wir machen für die Wahrscheinlichkeit?"

Schüler: "Überhaupt nichts."

Die folgende Szene belegt, daß bei klassischen Objekten Schülern der hypothetische Charakter von Wahrscheinlichkeiten nicht präsent ist und höchstens vom gewissenhaften Lehrer in Nebensätzen eingeflochten wird.

### 3. Szene:

(5. Stunde des oben genannten Klasse 8. Bauersfeld u. a. 1983 S. 190)

Schüler: "Ja, bei den Reißzwecken da waren die beiden Seiten verschieden, aber - "

Lehrer: "möglicherweise verschieden, und beim Würfel?"

Schüler: "Ja, da sind sie alle gleich."

Lehrer: "Sofern der Würfel nicht gezinkt ist."

## 2 Abwägen zwischen Hypothesen (durch wiederholte Anwendung der Regel von Bayes)

### 2.1 Vorbemerkung

In der Beurteilenden Statistik versucht man, von experimentellen Daten ("Indizien") auf die Gültigkeit von Hypothesen zurückzuschließen. Der Präfiguration dieser "fundamentalen Idee" (Bruner 1980) haben wir uns - auf nicht formalem Niveau - in Kapitel 1 gewidmet. Wir wollen ihr nun eine erste quantitative Grundlage geben, indem wir die Bayessche Regel, ein klassisches Thema der Wahrscheinlichkeitsrechnung, in den Dienst der Statistik stellen und mit ihrer Hilfe aufgrund experimenteller Daten zwischen mehreren Alternativen entscheiden. Wir stimmen Dinges (1980) zu, wenn er die zentrale Bedeutung solcher Entscheidungssituationen betont.

"I should like to propose to begin stochastics instruction always with decisions under uncertainty naively conceived by the learners. Subsequently they are to formulate the procedure of decision-making. Finally there is a check how far mathematical methods can be developed to criticise the procedures of decision-making. Hence the starting point will always be a situation in which the facts to be explained and the concepts to be introduced are concretely given and a decision is required. After that there is an exercise at more precise formulation and finally mathematics with an apparatus of signs." (Zitiert nach Hawkins/Kapadia 1984).

Die mit bevorstehenden Entscheidungen verbundene Spannung, ja "emotionale Betroffenheit" erzeugt man allerdings nicht durch einmalige Verwendung der Bayesschen Regel. Erst wiederholte Anwendungen lassen uns erleben, wie wir durch hinzukommende Information lernen, wie wir an "Sicherheit" gewinnen, um schließlich bereit zu werden, uns zwischen verfügbaren Alternativen zu entscheiden. So verstanden liefert die Bayessche Regel den Schlüssel zum Verständnis induktiven Schließens (Dinges 1978). Sie modelliert die Informationsverarbeitung beim Lernen durch Erfahrung und bereitet die

Testverfahren der Beurteilenden Statistik vor - ohne sie ersetzen zu wollen. Insbesondere wollen wir mit der Bayes-schen Regel nicht in den Streit zwischen Subjektivisten und Objektivisten eingreifen. Wir setzen sie aus pädagogischen Gründen ein:

- weil sie intuitive Alltagsvorstellungen modelliert und sich ausgezeichnet in das Curriculum beider Sekundarstufen einfügt,
  - weil sie deutlich macht, daß Wahrscheinlichkeitsrechnung kein Selbstzweck ist, sondern im Dienste statistischer Entscheidungen stehen kann,
  - weil man mit Pfadregel und Additionsregel auskommt, man ist nicht auf Verteilungen von Testgrößen angewiesen und kann daher sehr viel früher als im klassischen Rahmen Entscheidungsprobleme quantitativ angehen,
  - weil sie hilft, späteren Fehlinterpretationen testtheoretischer Aussagen vorzubeugen.
- (Vgl. auch Riemer 1981 und 1985.)

"The mathematical calculus underlying the Bayesian approach is of simple structure. Here again the problem is not situated in mathematics but in the model's interpretation, in the way it is applied. ... From an educational perspective this is not simply a matter of deciding whether the Bayesian approach is right or wrong - i.e. of contributing to the controversy about the subjective and the objective concept of probability - it is intended to help us obtain new insights into the relationship between probability theory and statistics ... " (Steinbring/v.Harten 1983).

## 2.2 Merkwürdige Würfel

Ein Subjektivist mag versucht sein, mit Hilfe der Bayesschen Regel zwischen verschiedenen Modellen zu einem gegebenen stochastischen Problem abzuwägen (beispielsweise zwischen verschiedenen Hypothesen über einen bestimmten Quader.) Ein Objektivist steht solchen Vorhaben skeptisch gegenüber.

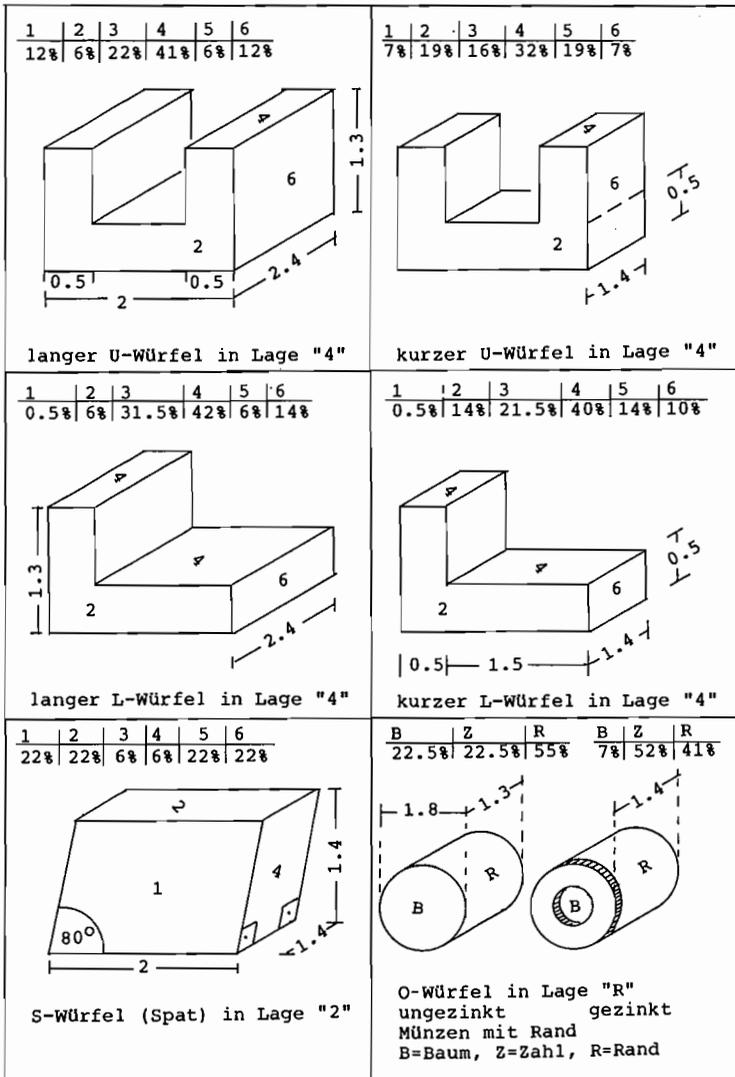


Abb. 2-1 merkwürdige "Würfel" (Maße in cm) mit Hypothesen  
(die Summe gegenüberliegender Seitenzahlen ist 7)

Hypothesen oder Modelle sind nämlich keine Zufallsgrößen, man kann ihnen keine Wahrscheinlichkeiten zuordnen, die sich frequentistisch (als Vorhersagen relativer Häufigkeiten) deuten lassen. Wir untersuchen im folgenden Situationen, in denen Hypothesen tatsächlich Zufallscharakter besitzen, die also auch für Objektivisten unbedenklich sind: Wir wählen zwischen verschiedenen Zufallsobjekten, die durch unterschiedliche Hypothesen gekennzeichnet werden. Mit einer Entscheidung zwischen den Hypothesen ist dann eine Entscheidung zwischen den Objekten verbunden. Die unmittelbare Nachprüfbarkeit solcher Entscheidungen ist für die Motivation von unschätzbarem Wert.

In der Unterrichtspraxis haben sich "Würfel" bewährt, die schon wegen ihrer ungewöhnlichen Form auf reges Interesse stoßen. Sie sind aus den in Kapitel 1 genannten Gründen wieder partiell symmetrisch und liegen paarweise in verschiedenen Ausführungen (kurz-lang, gezinkt-ungezinkt) vor (Riemer 1988). Wir bezeichnen sie als U-, L-, Spat (S-) und O-Würfel. Abb. 2-1 zeigt ihre Abmessungen und brauchbare Hypothesen.

Experiment:

Oliver wählt zwischen folgenden Würfeln:

	1	2	3	4	5	6
N: Laplace	16.6%	16.6%	16.6%	16.6%	16.6%	16.6%
L: kurzes L	0.5%	14.0%	21.5%	40.0%	14.0%	10.0%
S: Spat	22.0%	22.0%	6.0%	6.0%	22.0%	22.0%

Er teilt uns nach jedem Wurf die gefallene Augenzahl mit und wir wollen herausfinden, welchen Würfel Oliver benutzt. Anfangs sind wir sehr unsicher. Jeder Würfel kommt in Frage, wir besitzen für alle drei die gleiche (subjektive) Wahrscheinlichkeit:

N	L	S
33.3%	33.3%	33.3%

Im ersten Wurf würfelt Oliver die 4. Das spricht am meisten für L und sehr gegen S. Denn 4 tritt bei L sehr häufig (40%) auf, bei S dagegen sehr selten (6%). Wir drücken unser Gefühl durch folgende Verteilung aus:

N	L	S
30%	60%	10%

Dann wird 1 gewürfelt. Das spricht extrem stark gegen L, denn hier tritt 1 nur mit der Wahrscheinlichkeit 0.5% auf. Wir müssen obige Verteilung ganz erheblich revidieren:

N	L	S
75%	5%	20%

Der nächste Versuch liefert das Indiz 3, es spricht sehr gegen S. N gewinnt an Glaubwürdigkeit:

N	L	S
85%	6%	9%

Die nächste 1 macht uns noch sicherer:

N	L	S
89%	1%	10%

nach einer weiteren 3:

N	L	S
95%	1%	4%

riskieren wir die Entscheidung: "Es handelt sich um den normalen Laplace-Würfel" (Die Verteilungen stammen aus einem Unterrichtsexperiment.)

Die Wahrscheinlichkeit 5% für L und S zusammen ist unsere subjektive Irrtumswahrscheinlichkeit. Obwohl wir uns für N entschieden haben, glauben wir ja, daß L oder S mit 5% vorliegen, daß wir uns also geirrt haben. Nur weitere Versuche könnten unsere Sicherheit erhöhen.

### 2.3 Bayessche Regel

Verschiedene Personen nennen trotz gleicher Information intuitiv verschiedene (subjektive) Wahrscheinlichkeiten für das Vorliegen der drei möglichen Würfel (N L S). Die Bayessche Regel ist eine rational begründbare Methode zur Revision dieser Wahrscheinlichkeiten. Die folgenden Zeilen zeigen, wie man sie auf dem Niveau der Sekundarstufe I (auch ohne formalen Beweis) verstehen kann:

Hätten wir vor dem ersten Versuch gefragt: "mit welcher Wahrscheinlichkeit erwarten wir das Indiz 4?", so hätten wir (mit Pfad- und Additionsregel) folgende Antwort erhalten:

$$\begin{aligned} P(4) &= P(N) \cdot P(4|N) + P(L) \cdot P(4|L) + P(S) \cdot P(4|S) \\ &= 0.33 \cdot 0.166 + 0.33 \cdot 0.40 + 0.33 \cdot 0.06 = 0.21. \end{aligned}$$

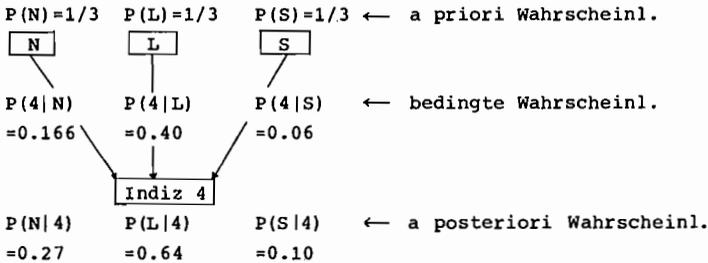
(Dabei bezeichnet z. B.  $P(N)$  die Wahrscheinlichkeit für den normalen Laplace-Würfel,  $P(4|N)$  ist die (bedingte) Wahrscheinlichkeit für 4 bei Verwendung eines normalen Würfels. Nun ist aber 4 eingetreten. Und der Anteil an  $P(4)$ , der von N herrührt, ist gleichzeitig die Wahrscheinlichkeit, daß N vorliegt, also die bedingte Wahrscheinlichkeit  $P(N|4)$ . Und das ist die Bayessche Regel:

$$P(N|4) = \frac{P(N) \cdot P(4|N)}{P(N) \cdot P(4|N) + P(L) \cdot P(4|L) + P(S) \cdot P(4|S)} = \frac{0.05}{0.21} = 27\%$$

Analog erhält man  $P(L|4)=64\%$  und  $P(S|4)=10\%$ .

(Alle Zahlenangaben in diesem Kapitel wurden intern mit acht Stellen berechnet und anschließend gerundet.) Durch Beobachtung des Indizes 4 hat sich also die (a priori) Wahr-

scheinlichkeitsverteilung (0.33, 0.33, 0.33) verändert zu (0.27, 0.64, 0.10). Man nennt sie a posteriori Wahrscheinlichkeitsverteilung. Wir fassen in einem "invertierten Baumdiagramm" zusammen:

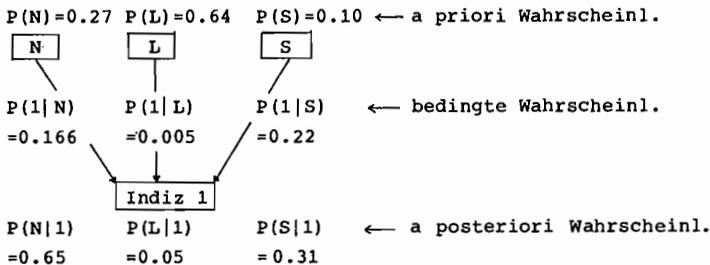


Für den zweiten Wurf übernimmt diese a posteriori Verteilung die Rolle der a priori Verteilung. Die Bayessche Regel liefert:

$$P(N|1) = \frac{P(N) \cdot P(1|N)}{P(N) \cdot P(1|N) + P(L) \cdot P(1|L) + P(S) \cdot P(1|S)}$$

$$= \frac{0.27 \cdot 0.166}{0.27 \cdot 0.166 + 0.64 \cdot 0.005 + 0.10 \cdot 0.22} = \frac{0.045}{0.070} = 0.65$$

und analog  $P(L|1)=0.05$ ,  $P(S|1)=0.13$ . Wir veranschaulichen wieder durch ein Diagramm:



Wir setzen unser Experiment fort und wenden in jedem Schritt die Bayessche Regel an:

$$P(A_i | I) = \frac{P(A_i) \cdot P(I | A_i)}{\sum_k P(A_k) \cdot P(I | A_k)}$$

Dabei ist I das Indiz, also die gewürfelte Augenzahl. Für die Alternativen schreiben wir:  $A_1=N$ ,  $A_2=L$ ,  $A_3=S$ . Wir erhalten folgende Tabelle (gerundet):

Bayessche Regel				vorher intuitiv geschätzt		
I	N	L	S	N	L	S
	0.33	0.33	0.33	0.33	0.33	0.33
4	0.27	0.64	0.10	0.30	0.60	0.10
1	0.65	0.05	0.31	0.75	0.05	0.20
3	0.79	0.07	0.14	0.85	0.06	0.09
1	0.81	0.00	0.18	0.89	0.01	0.10
3	0.92	0.00	0.07	0.95	0.01	0.04

Tab. 2-2

Tatsächlich lag ein normaler Würfel (N) vor. Es ist immer wieder reizvoll, die errechneten mit den vorher intuitiv geschätzten Wahrscheinlichkeiten zu vergleichen. Meist ist die Übereinstimmung nicht schlecht, die Bayessche Regel modelliert unsere Intuitionen also recht gut.

In obigem Beispiel brauchten wir zur Entscheidung nur wenige Schritte. Das lag daran, daß durch die Einsen der L-Würfel rasch ausschied, ebenso schied der S-Würfel durch das Auftreten von 3 und 4 schnell aus. Die Versuche können länger dauern. Wir geben zwei Beispiele:

Versuchsreihe 2				Versuchsreihe 3			
Wahrscheinlichkeit				Wahrscheinlichkeit			
I	N	L	S	I	N	L	S
2	0.33	0.33	0.33	3	0.33	0.33	0.33
2	0.32	0.27	0.42	3	0.38	0.49	0.14
3	0.39	0.42	0.19	5	0.39	0.42	0.19
6	0.44	0.29	0.28	4	0.27	0.69	0.05
1	0.54	0.01	0.45	4	0.14	0.85	0.01
1	0.48	0.00	0.52	4	0.06	0.94	0.00
3	0.72	0.00	0.28	5	0.07	0.92	0.00
5	0.66	0.00	0.34	4	0.03	0.97	0.00
4	0.84	0.00	0.16	5	0.04	0.96	0.00
2	0.80	0.00	0.20	3	0.03	0.97	0.00
3	0.92	0.00	0.08	5	0.04	0.96	0.00
2	0.89	0.00	0.10	3	0.03	0.97	0.00
1	0.87	0.00	0.13	3	0.02	0.98	0.00
5	0.83	0.00	0.17	3	0.02	0.98	0.00
1	0.79	0.00	0.21				
4	0.91	0.00	0.09				
2	0.89	0.00	0.11				
6	0.86	0.00	0.14				
1	0.82	0.00	0.18				
6	0.77	0.00	0.23				
1	0.72	0.00	0.28				
6	0.66	0.00	0.34				
4	0.84	0.00	0.16				
4	0.94	0.00	0.06				
6	0.92	0.00	0.08				
4	0.97	0.00	0.03				
4	0.99	0.00	0.01				
4	1.00	0.00	0.00				
1	0.99	0.00	0.01				
5	0.99	0.00	0.01				
2	0.99	0.00	0.01				

Entscheidung:  
der (kurze) L-Würfel wird benutzt

Entscheidung: der normale Würfel (N) wird benutzt

Tab. 2-3

Ergänzung: "komprimierte" Anwendung der Bayesschen Regel  
Wenn man sich für den "Prozeß des Erkennens" im Zuge wiederholter Anwendung der Bayesschen Regel nicht interessiert, kommt man auch in einem Schritt zum gleichen Endergebnis. Man betrachtet "zusammengesetzte" Indizien, bei unserem ersten Würfelbeispiel etwa  $I=(4,1,3,1,3)$ . Dieses Indiz hat die Wahrscheinlichkeit (Pfadregel):  $P(1)^2 \cdot P(3)^2 \cdot P(4)$ . Für den normalen Würfel (N), das kurze L (L) und den Spat-Würfel (S) ergeben sich unterschiedliche Werte:  
 $P(I|N)=0.0001286$   $P(I|L)=0.0000005$   $P(I|S)=0.0000105$ .

Die Bayessche Regel liefert mit

$$P(I) = (1/3)P(I|N) + (1/3)P(I|L) + (1/3)P(I|S) = 0.0000465$$

$$P(N|I) = (1/3)P(I|N) / P(I) = 0.9217542 \approx 0.92$$

$$P(L|I) = (1/3)P(I|L) / P(I) = 0.0033132 \approx 0.00$$

$$P(S|I) = (1/3)P(I|S) / P(I) = 0.0749325 \approx 0.07$$

das gleiche Endergebnis wie bei schrittweiser Anwendung in Tab. 2-2. Noch einfacher werden die Rechnungen, wenn man sich nur für die Verhältnisse der a posteriori Wahrscheinlichkeiten interessiert, die Nenner in der Bayesschen Regel kürzen sich. Bleiben wir beim obigen Beispiel:

$$\begin{aligned} \frac{P(N|I)}{P(S|I)} &= \frac{(1/3) \cdot P(I|N) / P(I)}{(1/3) \cdot P(I|S) / P(I)} = \frac{(1/3) \cdot P(I|N)}{(1/3) \cdot P(I|S)} \\ &= \frac{(1/3) \cdot (0.16)^2}{(1/3) \cdot (0.22)^2} \cdot \frac{(0.16)^2}{0.06} = 12.3 \end{aligned}$$

Beim Entscheiden zwischen nur zwei Alternativen hat man also (in jedem Schritt) den Quotienten der a priori Wahrscheinlichkeiten nur mit dem Quotienten der bedingten Wahrscheinlichkeiten zu multiplizieren, um den Quotienten der a posteriori Wahrscheinlichkeiten zu erhalten.

#### 2.4 Intuitive Entscheidungen quantitativ fundiert

Folgende Frage wurde von 39 Schülern (Klassen 8 und 9) vor ihrer ersten Stochastikstunde beantwortet. Der Antwort gingen keinerlei Experimente voraus: "Schau Dir das kurze U, das lange U (Abb. 2-1) und den normalen Würfel genau an. Wolfi hat mit diesen drei Würfelsorten je 50 mal gewürfelt. Die folgende Tabelle zeigt seine Ergebnisse. Kreuze für jede der sechs Zeilen (a) bis (f) an, mit welchem Würfel Wolfi Deiner Meinung nach gewürfelt hat."

Wolfis Würfelerggebnisse						Auswertung der Umfrage				
	1	2	3	4	5	6	kurzer U-Würfel	langer U-Würfel	normaler Würfel	weiß nicht
(a)	8	0	<u>12</u>	<u>21</u>	3	6	5%	<u>90%</u>	0%	5%
(b)	7	10	9	6	8	10	10%	3%	<u>87%</u>	0%
(c)	1	9	14	17	7	2	<u>74%</u>	15%	0%	10%
(d)	5	6	9	12	15	3	<u>54%</u>	13%	13%	20%
(e)	9	9	8	7	10	7	3%	3%	<u>92%</u>	3%
(f)	7	2	11	20	4	6	13%	<u>77%</u>	5%	5%

Tab. 2-4

(Richtige Antworten sind unterstrichen.) Es zeigt sich, daß Kinder intuitiv sinnvoll argumentieren: die Hypothese, die sich mit den experimentellen Daten am besten verträgt, wird ausgewählt. In dem Fall (d), in dem die Daten weniger eindeutig für das kurze U sprechen, erkennt man eine stark erhöhte Unsicherheit. Die Schüler wurden gebeten, ihre Antworten zu kommentieren:

1. Kommentare zu den Zeilen (a) und (f), die mit dem langen U gewürfelt wurden

- "Langes U, weil 2 und 5 nicht oft beim langen U kommt."
- "Langes U, weil 4 beim langen U auf der günstigsten Seite liegt, die glatte Fläche unten liegt und der Würfel die beste Balance hat."
- "Langes U. Bei diesem Würfel liegen 3 und 4 auf breiten Flächen."
- "Beim langen U kommt wahrscheinlich die 4 am meisten vor. Es ist schwer, eine 2 oder 5 zu würfeln."
- "Langes U, hier ist die 3 und die 4 an der günstigsten Seite. Bei dem normalen Würfel wären die Unterschiede nicht so groß."
- "Die Zahlen 2 und 5 sind fast gar nicht vorhanden. Beim langen U sind sie wegen ihrer Form sehr schwer zu würfeln."
- "Langes U. Weil 1 und 6 öfter geworfen wurden, sind die Seiten wahrscheinlich größer als 2 und 5."

2. Kommentare zu den Zeilen (b) und (e), die mit dem normalen Würfel erzeugt wurden:

- "Ausgeglichen."
- "Weil alle gleich verteilt sind und beim normalen Würfel alle Seiten gleich groß sind."
- "weil die Zahlen alle ziemlich gleich oft kommen."
- "Normaler Würfel, dort sind die Zahlen verteilt, der Würfel bleibt fast nie auf einer Zahl wie bei den anderen."
- "Es kann weder das kurze noch das lange U sein, denn die 2 und die 6 liegen an den schmalen Seiten und dort landet der Würfel bestimmt ganz selten."

3. Kommentare zu Zeile (c), die mit dem kurzen U gewürfelt wurde:

- "Kurzes U, weil die 1 und die 6 ziemlich selten beim kurzen U kommen."
- "kurzes U. Bei diesem Würfel sind 3 und 4 auf einer langen Seite und die 2 liegt auch auf einer Seite, die häufig gewürfelt werden kann."
- "Beim kurzen U wird viel 5 oder 2 gewürfelt, 1 und 6 nicht so oft."
- "Es würde gut zum kurzen U passen, da die Seiten mit den Zahlen 1 und 6 nicht oft gewürfelt wurden."
- "Kurzes U, 3 und 4 höchstwahrscheinlich, 2 und 5 zweitwahrscheinlich."

4. Kommentare zu Zeile (d), die mit dem kurzen U gewürfelt wurde; die Daten sind allerdings weniger eindeutig:

- "Es irritiert mich, daß 4 und 5 häufig gewürfelt wurden, die 2 sollte eigentlich genauso oft sein, weil es dann ein kurzes U wäre."
- "normaler Würfel, die Zahlen sind in etwa ausgeglichen"
- "Weiß nicht. Weil 4 und 5 am häufigsten geworfen wurden, müßte es normalerweise das kurze U sein, dann müßte aber 2 genauso oft geworfen worden sein, denn die ist genauso groß wie die 5. Es wäre am ehesten das kurze U".
- "Weiß nicht, weil es ungenau ist, mit welchem Würfel man gewürfelt haben könnte."

Gibt es eine bessere Möglichkeit, Kindern anhand eigener Daten statistische Grundgedanken vom Hypothesen-Verwerfen bis zu höheren Fehlentscheidungsquoten bei nicht eindeutigen Daten schon sehr frühzeitig bewußt zu machen? Mit Hilfe der Bayesschen Regel und brauchbarer Hypothesen über unsere Würfel können wir die Entscheidungen in einer Vertiefungsphase quantitativ untermauern. Wir benutzen folgende Verteilungen:

	1	2	3	4	5	6
kurzes U	0.07	0.19	0.16	0.32	0.19	0.07
langes U	0.12	0.06	0.22	0.41	0.06	0.12

und berechnen für die Zeilen (a) und (d) (Tab. 2.4) das Verhältnis der a posteriori Wahrscheinlichkeiten, für das lange bzw. kurze U. Bei der Verhältnisbildung kürzen sich die Nenner und die jeweils gleichen a priori Wahrscheinlichkeiten in den Bayes-Formeln. Als (zusammengesetztes) Indiz dient die Häufigkeitsverteilung (8,0,12,21,3,6) aus Zeile (a). Wir erhalten:

$$\frac{P(\text{langes U} | I)}{P(\text{kurzes U} | I)} = \frac{P(I | \text{langes U})}{P(I | \text{kurzes U})} .$$

$$\left(\frac{0.12}{0.07}\right)^8 \cdot \left(\frac{0.06}{0.19}\right)^0 \cdot \left(\frac{0.22}{0.16}\right)^{12} \cdot \left(\frac{0.41}{0.32}\right)^{21} \cdot \left(\frac{0.06}{0.19}\right)^3 \cdot \left(\frac{0.12}{0.07}\right)^6 = 496000$$

In Zeile (a) ist also das lange U 496000 mal so wahrscheinlich wie das kurze U. In der Zeile (d) ist die Lage weniger eindeutig. Wie man nachrechnet, ist das kurze U hier "nur" 26 mal so wahrscheinlich wie der normale Würfel. Das schlägt sich in der gegenüber (a) sehr viel stärkeren Unsicherheit bei den Schülerantworten nieder.

Wieder zeigen die Beispiele die Nähe der Bayesschen Regel zu unseren intuitiven Vorstellungen und belegen, wie gering der Rechenaufwand ist, wenn man sich auf Verhältnisse zwischen den a posteriori Wahrscheinlichkeiten beschränkt.

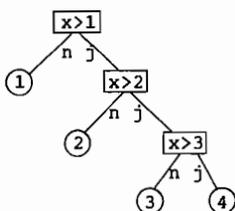
## 2.5 Wahrscheinlichkeit, Information und die Regel von Bayes

Wir untersuchen den Prozeß des "Lernens" im Zuge wiederholter Anwendungen der Bayesschen Regel mit Hilfe des Entropiebegriffs etwas detaillierter.

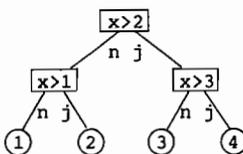
## 2.5.1 Entropie

Zunächst eine kurze Begriffserklärung: Wir erraten gedachte Zahlen  $x$  im Bereich 1, 2, 3, 4 durch eine Folge von  $j/n$ -Fragen. Hier zwei Fragestrategien:

Strategie A



Strategie B



Welche der Strategien ist effektiver? Bei Strategie A braucht man im Mittel

$$m_A = (1/4) 1 + (1/4) 2 + (1/4) 3 + (1/4) 3 = 2.25$$

Fragen, bei Strategie B dagegen nur

$$m_B = (1/4) 2 + (1/4) 2 + (1/4) 2 + (1/4) 2 = 2.00. \quad (1)$$

Strategie B ist günstiger, eine bessere gibt es nicht. Wie sieht die Situation aus, wenn nicht mehr alle Zahlen  $x$  mit gleicher Wahrscheinlichkeit erzeugt (ausgedacht) werden? Wenn wir beispielsweise  $p(1)=1/2$ ,  $p(2)=1/4$ ,  $p(3)=p(4)=1/8$  wählen, dann erweist sich Strategie A als optimal.

$$m_A = (1/2) \cdot 1 + (1/4) \cdot 2 + (1/8) \cdot 3 + (1/8) \cdot 3 = 1.75 \quad (2)$$

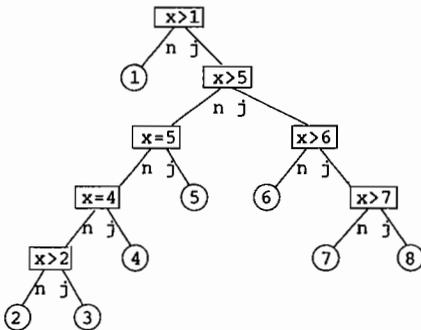
$$m_B = (1/2) \cdot 2 + (1/4) \cdot 2 + (1/8) \cdot 2 + (1/8) \cdot 2 = 2.00$$

(Die Ratezeit ist bei Strategie B wegen gleicher Pfadlängen verteilungsunabhängig.) Bei den optimalen Entscheidungsbäumen (1) und (2) gilt für die Pfadlänge  $l_i$ , die zum Erraten der Zahl  $i$  durchlaufen wird,  $l_i = \log_2(1/p(i))$ . (Bei Strategie A gilt beispielsweise  $l_3 = \log_2(1/0.125) = 3$ .) Die mittlere Ratezeit beträgt dann folglich

$$E := \sum_i p(i) \log_2(1/p(i)) .$$

Diese Größe  $E$  bezeichnet man (auch für den Fall, daß die Wahrscheinlichkeiten  $p(i)$  keine Potenzen von  $1/2$  sind) als Entropie der Wahrscheinlichkeitsverteilung. Sie ist ein Unbestimmtheitsmaß und gibt die untere Grenze für die Anzahl der ja/nein-Fragen an, die man im Mittel stellen muß um herauszufinden, welche Zahl ein Zufallsgenerator geliefert hat.

Ein weiteres Beispiel: Die Wahrscheinlichkeitsverteilung  $p(1)=1/2$ ,  $p(2)=p(3)=1/32$ ,  $p(4)=1/16$ ,  $p(5)=p(6)=1/8$ ,  $p(7)=p(8)=1/16$  hat die Entropie  $E=37/16=2.3125$ . Der zugehörige optimale Entscheidungsbaum hat die Struktur



Die mittlere Ratezeit ist tatsächlich  $m=2.3125$ . (Der ausgeglichene Entscheidungsbaum liefert dagegen nur  $m=3$ .) Möchte man zwischen drei gleichwahrscheinlichen Zahlen raten, dann erreicht die mittlere Ratezeit auch beim optimalen Baum (5/3) den Wert der Entropie (1.58) nicht.

### 2.5.2 Information

Verändert sich die Wahrscheinlichkeitsverteilung durch ein Indiz (eine Nachricht) so, daß die Entropie  $E$  abnimmt, dann ordnen wir der Nachricht die Differenz  $dE:=E_{\text{alt}}-E_{\text{neu}}$  als Informationsgehalt zu (Maßeinheit Bit).  $dE$  kann auch negativ sein, wenn uns die Nachricht verunsichert. Wir wollen diesen Prozeß im Zusammenhang mit der Bayesschen Regel studieren:

### 2.5.3 Experiment

Sven wählt (mit verbundenen Augen) eine von drei Urnen aus. Alle Urnen enthalten 50 Kugeln. In Urne A ist die Farbe Rot 15mal (30%), in Urne B 20mal (40%), in Urne C 35mal (70%) vertreten. Die restlichen Kugeln sind weiß. Wir ziehen ohne Zurücklegen und versuchen herauszufinden, welche Urne Sven erwischt hat. Anfangs besitzen alle Urnen die gleiche Wahrscheinlichkeit 1/3 (Entropie 1.883). Tabelle 2-5 zeigt ein Versuchsprotokoll: In den ersten 6 Zügen ergaben sich die Farben "r r w r w r" (zusammengesetztes Indiz I, Spalte 2). Dieses Indiz besitzt für die drei Urnen verschiedene Wahrscheinlichkeiten (Pfadregel)

$$p(I|A) = \frac{15}{50} \frac{14}{49} \frac{35}{48} \frac{13}{47} \frac{34}{46} \frac{12}{45} = 0.003407338$$

$$p(I|B) = \frac{20}{50} \frac{19}{49} \frac{30}{48} \frac{18}{47} \frac{29}{46} \frac{17}{45} = 0.008841964$$

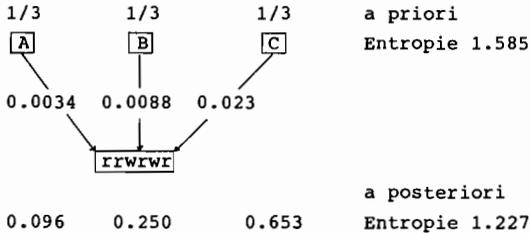
$$p(I|C) = \frac{35}{50} \frac{34}{49} \frac{15}{48} \frac{33}{47} \frac{14}{46} \frac{32}{45} = 0.023065063$$

	1	2	3	4	5	6	7	8	ohne Rücklegen				mit Rücklegen			
									15 r 35 w	20 r 30 w	35 r 15 w	Entro- pie	15 r 35 w	20 r 30 w	35 r 15 w	Entro- pie
1	1	1	0	0.333	0.333	0.333	1.585	0.333	0.333	0.333	1.585	0.333	0.333	0.333	1.585	
2	1	1	1	0.214	0.286	0.500	1.493	0.214	0.286	0.500	1.493	0.214	0.286	0.500	1.493	
3	1	1	2	0.118	0.213	0.669	1.228	0.122	0.216	0.662	1.241	0.122	0.216	0.662	1.241	
4	1	1	3	0.201	0.311	0.488	1.494	0.206	0.314	0.480	1.502	0.206	0.314	0.480	1.502	
5	1	1	4	0.107	0.231	0.662	1.228	0.118	0.240	0.642	1.268	0.118	0.240	0.642	1.268	
6	1	1	5	0.186	0.341	0.473	1.492	0.197	0.343	0.460	1.507	0.197	0.343	0.460	1.507	
7	1	1	6	0.096	0.250	0.653	1.227	0.114	0.265	0.621	1.292	0.114	0.265	0.621	1.292	
8	1	1	7	0.042	0.158	0.800	0.870	0.060	0.184	0.756	0.997	0.060	0.184	0.756	0.997	
9	0	5	3	0.085	0.273	0.641	1.225	0.110	0.292	0.598	1.312	0.110	0.292	0.598	1.312	
10	0	5	4	0.153	0.414	0.432	1.465	0.178	0.406	0.416	1.498	0.178	0.406	0.416	1.498	
11	0	5	5	0.234	0.531	0.234	1.466	0.253	0.494	0.253	1.506	0.253	0.494	0.253	1.506	
12	0	5	6	0.310	0.586	0.103	1.314	0.323	0.539	0.138	1.402	0.323	0.539	0.138	1.402	
13	0	6	6	0.207	0.586	0.207	1.392	0.237	0.527	0.237	1.471	0.237	0.527	0.237	1.471	
14	0	6	7	0.274	0.641	0.085	1.225	0.300	0.572	0.128	1.362	0.300	0.572	0.128	1.362	
15	0	6	8	0.332	0.639	0.029	1.091	0.355	0.580	0.065	1.243	0.355	0.580	0.065	1.243	
16	0	6	9	0.386	0.605	0.009	1.029	0.403	0.565	0.032	1.151	0.403	0.565	0.032	1.151	
17	0	6	10	0.440	0.558	0.002	1.011	0.447	0.538	0.015	1.092	0.447	0.538	0.015	1.092	
18	1	7	11	0.496	0.503	0.001	1.006	0.489	0.504	0.007	1.053	0.489	0.504	0.007	1.053	
19	1	7	12	0.387	0.611	0.001	0.977	0.415	0.571	0.014	1.075	0.415	0.571	0.014	1.075	
20	1	8	11	0.280	0.717	0.003	0.886	0.343	0.629	0.027	1.091	0.343	0.629	0.027	1.091	
21	0	8	12	0.330	0.669	0.001	0.922	0.384	0.603	0.013	1.051	0.384	0.603	0.013	1.051	
22	0	8	13	0.386	0.614	0.000	0.964	0.424	0.570	0.006	1.032	0.424	0.570	0.006	1.032	
23	0	8	14	0.449	0.551	0.000	0.993	0.463	0.534	0.003	1.022	0.463	0.534	0.003	1.022	
24	1	9	14	0.322	0.678	0.000	0.907	0.392	0.603	0.006	1.012	0.392	0.603	0.006	1.012	
25	1	10	14	0.206	0.794	0.000	0.735	0.324	0.665	0.011	0.989	0.324	0.665	0.011	0.989	
26	1	11	14	0.115	0.885	0.000	0.517	0.262	0.717	0.021	0.966	0.262	0.717	0.021	0.966	
27	1	12	14	0.054	0.945	0.001	0.313	0.207	0.755	0.038	0.956	0.207	0.755	0.038	0.956	
28	0	12	15	0.070	0.930	0.000	0.368	0.238	0.743	0.019	0.918	0.238	0.743	0.019	0.918	
29	0	12	16	0.092	0.908	0.000	0.442	0.269	0.722	0.009	0.911	0.269	0.722	0.009	0.911	
30	1	13	16	0.036	0.964	0.000	0.226	0.215	0.768	0.017	0.869	0.215	0.768	0.017	0.869	
31	0	13	17	0.049	0.951	0.000	0.281	0.244	0.748	0.008	0.867	0.244	0.748	0.008	0.867	
32	1	14	17	0.014	0.986	0.000	0.109	0.194	0.791	0.015	0.818	0.194	0.791	0.015	0.818	
33	0	14	18	0.020	0.980	0.000	0.141	0.221	0.772	0.007	0.822	0.221	0.772	0.007	0.822	
34	0	14	19	0.028	0.972	0.000	0.184	0.249	0.747	0.004	0.843	0.249	0.747	0.004	0.843	
35	0	14	20	0.040	0.960	0.000	0.243	0.280	0.719	0.002	0.872	0.280	0.719	0.002	0.872	
36	1	15	20	0.007	0.993	0.000	0.060	0.225	0.772	0.003	0.800	0.225	0.772	0.003	0.800	
37	0	15	21	0.010	0.990	0.000	0.083	0.253	0.745	0.002	0.833	0.253	0.745	0.002	0.833	
38	0	15	22	0.016	0.984	0.000	0.118	0.284	0.715	0.001	0.869	0.284	0.715	0.001	0.869	
39	1	16	22	0.000	1.000	0.000	0.000	0.229	0.769	0.001	0.791	0.229	0.769	0.001	0.791	
40	0	16	23	0.000	1.000	0.000	0.000	0.258	0.742	0.001	0.831	0.258	0.742	0.001	0.831	
41	0	16	24	0.000	1.000	0.000	0.000	0.288	0.711	0.000	0.871	0.288	0.711	0.000	0.871	
42	0	16	25	0.000	1.000	0.000	0.000	0.321	0.679	0.000	0.908	0.321	0.679	0.000	0.908	
43	1	17	25	0.000	1.000	0.000	0.000	0.262	0.738	0.000	0.833	0.262	0.738	0.000	0.833	
44	0	17	26	0.000	1.000	0.000	0.000	0.293	0.707	0.000	0.874	0.293	0.707	0.000	0.874	
45	0	17	27	0.000	1.000	0.000	0.000	0.325	0.674	0.000	0.911	0.325	0.674	0.000	0.911	
46	1	18	27	0.000	1.000	0.000	0.000	0.266	0.734	0.000	0.837	0.266	0.734	0.000	0.837	
47	0	18	28	0.000	1.000	0.000	0.000	0.297	0.703	0.000	0.878	0.297	0.703	0.000	0.878	
48	0	18	29	0.000	1.000	0.000	0.000	0.330	0.670	0.000	0.915	0.330	0.670	0.000	0.915	
49	0	18	30	0.000	1.000	0.000	0.000	0.365	0.635	0.000	0.947	0.365	0.635	0.000	0.947	
50	1	19	30	0.000	1.000	0.000	0.000	0.301	0.699	0.000	0.883	0.301	0.699	0.000	0.883	
51	1	20	30	0.000	1.000	0.000	0.000	0.244	0.756	0.000	0.803	0.244	0.756	0.000	0.803	

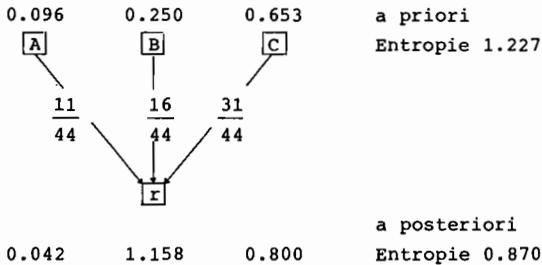
Tab. 2-5 Aus welcher Urne stammen die gezogenen Kugeln?

Spalte 2: gezogene Farbe (1=rot, 0=weiß). Spalten 3 und 4: kumulierte absolute Häufigkeiten für rot und weiß.  
 Spalten 5 bis 7: Wahrscheinlichkeiten für die Urnen A (30% rot) B(40% rot) C (70% rot), die sich aus der anfänglichen Gleichverteilung nach Bayes ergeben. Spalte 8: Entropie.

Es macht Urne C etwa 6.8 mal, Urne B etwa 2.6 mal so wahrscheinlich wie Urne A. Die Bayessche Regel liefert mit



das Ergebnis der Zeile 6. Das zusammengesetzte Indiz "rrwrwr" besitzt folglich die Information  $dE=1.585-1.227=0.358$  Bit. Wir sind uns "recht sicher" (65.3%), daß Sven Urne C (mit 70% roten Kugeln) erwischt hat. Im siebenten Zug erhalten wir schon wieder "rot":



Dieses Indiz spricht noch mehr für Urne C (80%). Es besitzt den Informationsgehalt  $dE=1.227-0.87=0.357$  Bit (Zeile 7). In den nächsten vier Zügen erhalten wir viermal eine weiße Kugel, was uns sehr "irritiert". (Zeile 11: Verteilung  $p(A)=31\%$ ,  $p(B)=58.6\%$ ,  $p(C)=10.3\%$  mit der Entropie 1.314.) Die Verunsicherung drückt sich durch eine negative Information  $dE=0.870-1.314=-0.444$  Bit aus. Die folgenden Züge, in denen sich Urne B langsam "durchsetzt", bringen Sicherheit. Im 21ten Zug erhält man die 13te weiße Kugel, wodurch die Wahrscheinlichkeit für C (15 weiße, 35 rote) erstmalig unter

1 Promille absinkt. Im 28ten Zug wird die 16te weiße Kugel gezogen, wodurch C gänzlich ausscheidet. Im 38ten Zug scheidet mit der 16ten roten Kugel auch A (15 rote, 35 weiße) aus. Mit Sicherheit (Entropie 0) hat Sven die Urne B erwischt.

Der rechte Teil von Tab. 2-5 zeigt zum Vergleich exemplarisch, wieviel Information (Bit) wir "verschenken", wenn wir die gezogenen Kugeln wieder zurücklegen. Die gleiche a priori Verteilung wurde hier Schritt für Schritt mit der gleichen Indizienfolge (die allerdings zum Ziehen ohne Zurücklegen gehört) revidiert. Nach dem 50ten Zug entspricht die relative Häufigkeit genau der Wahrscheinlichkeit, von Urne B (40% rot, 60% weiß). Stärker kann ein Indiz nicht für B sprechen. Dennoch besitzt Urne A, deren Inhalt demjenigen der Urne B nahekommt, immer noch die Wahrscheinlichkeit 24.4% (Entropie 0.803). Anfangs macht es wenig Unterschied, ob eine Indizienfolge vom Ziehen mit oder ohne Zurücklegen stammt. Erst mit leerer werdender Urne fallen die Unterschiede deutlich ins Gewicht.

## 2.6 Resümee

### 2.6.1 Subjektivisten - Objektivisten

"Der objektivistische Wahrscheinlichkeitsbegriff, der sich an Urne und Glücksrad orientiert, scheint gut geeignet, am Anfang des Stochastikunterrichts zu stehen. Er hat aber Grenzen. Die Bayessche Formel paßt nicht dazu. Sie braucht den subjektiven Wahrscheinlichkeitsbegriff. Die höhere Wahrscheinlichkeitstheorie braucht den Begriff der Hypothese" (Dinges 1978).

Im ersten Kapitel haben wir dargelegt, wie man objektivistische und hypothetische Aspekte des Wahrscheinlichkeitsbegriffs auch im praktischen Unterricht sehr gut miteinander vereinbaren kann.

Wahrscheinlichkeiten prognostizieren relative Häufigkeiten. Wie lassen sich aber die Wahrscheinlichkeiten interpretieren, die im Zuge wiederholter Anwendungen der Bayesschen Regel berechnet wurden? Für die meisten Menschen haben sie subjektivistische Züge. "Ich schätze: mit 95% Wahrscheinlichkeit würfeln Oliver mit dem L-Würfel. Mit 56%iger Wahrscheinlichkeit zieht Sven aus der Urne mit 40% roten Kugeln." Solche Aussagen wollen in Anlehnung an den Alltagssprachgebrauch einen "Grad der Sicherheit", ein "Vertrauen" quantifizieren, das wir in gewisse Hypothesen setzen. Ein Objektivist lehnt solche Interpretationen ab, er deutet auch die Ergebnisse der Bayesschen Regel frequentistisch: Man stellt sich eine große Zahl von Versuchen vor, in denen man in der ersten Stufe zunächst zufällig einen der Würfel auswählt (Vorversuch mit wohldefinierter a priori Verteilung) und versucht, das gegebene Indiz zu würfeln. Immer dann, wenn dies gelungen ist, schaut man nach, welcher Würfel zugrunde lag. In 95% aller Fälle wird dann auf lange Sicht der "L-Würfel" zum Vorschein kommen. In dieser Interpretation prognostizieren auch die Wahrscheinlichkeiten, die sich aus einer Anwendung der Bayesschen Regel ergeben, relative Häufigkeiten (für Hypothesen, die jetzt auch Zufallsgrößen sind). Wir überlassen es dem Leser, ob er neben der objektivistischen auch eine subjektivistische Interpretation zulassen möchte. Die Nähe zu intuitiven Alltagsvorstellungen lassen sie unter motivationspsychologischen Aspekten durchaus vertretbar erscheinen. (Man vergleiche hierzu auch Borovcnik 1984, Hawkins/Kapadia 1984, Riemer 1981, Schrage 1980 und 1981, Scozzafava 1987, Steinbring/v. Harten 1983.)

#### 2.6.2 Alternativtests aus Bayesscher Sicht.

Sowohl durch wiederholte Anwendung der Bayesschen Regel als auch mit dem Alternativtest versucht man, aus zwei (oder mehr) Hypothesen "die richtige" herauszufiltern. Während man bei der Bayesschen Regel jedoch ein "punktueller" Indiz berücksichtigt, das nach einer Versuchsserie zum Abwägen

zwischen Hypothesen benutzt wird, haben die bei Alternativtests verwendeten Indizien die Form kritischer Bereiche. Sie sind vor Testdurchführung festzulegen und besitzen Entscheidungsfunktion.

Wollen wir beispielsweise in 10 Versuchen zwischen zwei "Trefferwahrscheinlichkeiten"  $p=0.3$  und  $p=0.4$  entscheiden (Abb. 2-6), so stellt ein "Testtheoretiker", der auf dem 5%-Signifikanzniveau arbeiten möchte, seine Entscheidungsregeln mit Hilfe umgekehrter Baumdiagramme wie folgt dar:

Person A testet  
Hypothese  $H_0: p=0.3$  gegen  
Alternative  $H_1: p=0.4$

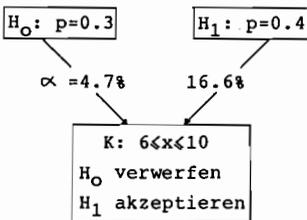


Abb. 2-6a

Person B testet  
Hypothese  $H_0: p=0.4$  gegen  
Alternative  $H_1: p=0.3$

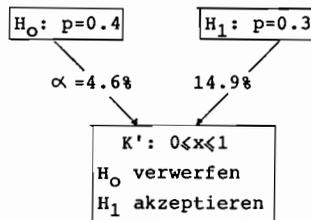


Abb. 2-6b

Falls in 10 Versuchen 5 Treffer auftraten, wird jede Person bei ihrer Nullhypothese bleiben, A bei  $p=0.3$  und B bei  $p=0.4$ . Ein unvoreingenommener Bayesianer wird dagegen der Hypothese  $p=0.3$  (mit 33.9%) nur halbsoviel Vertrauen schenken wie der Hypothese  $p=0.4$  (66.1%).

Nehmen wir aber an, das Indiz  $K$  (6 oder mehr Treffer in 10 Versuchen) wäre eingetreten, dann muß A seine Nullhypothese  $p=0.3$  gegen die Alternative  $p=0.4$  auf dem 5%-Signifikanzniveau verwerfen. Denn dieses Indiz tritt bei Gültigkeit von  $p=0.3$  in weniger als 5% aller Fälle ein (Abb. 2-6a). Oft wird diese Aussage fälschlich so interpretiert, als würde die Alternative  $p=0.4$  nun mit 95%iger Wahrscheinlichkeit gelten. Das konstatiert auch Schrage (1981 S. 197), wenn er

schreibt: "Besonders häufig taucht die Fehldeutung beim Testen von Hypothesen auf: Aufgrund des vorliegenden Stichprobenergebnisses ist mit einer Wahrscheinlichkeit von 95% zu erwarten, daß die Hypothese  $H_0$  falsch ist."

Der erfahrene Lehrer weiß, daß mit "ad hoc" eingeflochtenen Erklärungen - seien sie noch so eindringlich vorgebracht - solchen tiefsitzenden Fehlvorstellungen nicht beizukommen ist. ("Intuitions cannot be modified by verbal explanations only" Fischbein 1984, S. 2). Allein eine fundierte Kenntnis der Bayesschen Betrachtungsweise, wie wir sie in diesem Kapitel skizziert haben, schafft Abhilfe. Denn wer gelernt hat, nach Bayes Alternativen gegeneinander abzuwägen, wird aus einer Stichprobe keine Wahrscheinlichkeitsaussagen über die Gültigkeit von Hypothesen mehr ableiten. Er wird sie bestenfalls zur Revision schon vorgegebener a priori Wahrscheinlichkeiten heranziehen - wenn solche a priori Wahrscheinlichkeiten Sinn machen. Dann spricht ein Ergebnis im kritischen Bereich  $K$  beispielsweise 3.5 ( $=0.166/0.047$ ) mal stärker für die Alternative  $p=0.4$  als für die Hypothese  $p=0.3$  (Abb. 2-6a). Das Verhältnis der a priori Wahrscheinlichkeiten  $P(H_1)/P(H_0)$  wird also durch Beobachtung von  $K$  um den Faktor 3.5 größer. Im Sinne einer "Stützungslogik" wird die Alternative  $p=0.4$  durch ein Indiz im kritischen Bereich  $K$  gegenüber der Hypothese  $p=0.3$  gestützt. Um so weniger wird aber die Alternative "statistisch bewiesen" je näher sie bei  $H_0$  liegt, weil der "Stützungsfaktor"  $P(K|H_1)/P(K|H_0)$  sich dann dem Wert 1 nähert.

Schrage (1981 S. 197) schreibt: "Es ist abzulehnen, wenn - wie es in vielen Büchern geschieht - das Ablehnen der Nullhypothese als statistischer Beweis für die Gegenhypothese gewertet wird, nach einem Schema, bei dem die Analogie zum indirekten Beweis herausgestellt wird."

Wenn es gelingt, mit Hilfe der Bayesschen Regel ein Verständnis für solche Zusammenhänge zu erzeugen, dann stellt sie - neben der Motivation, die insbesondere iterierte An-

wendungen bei Schülern entfesseln können - ein solides Fundament zur adäquaten Interpretation statistischer Aussagen dar, ohne Verfahren der Beurteilenden Statistik ersetzen zu wollen. Näheres hierzu im folgenden Kapitel.

Teil II Chi-Quadrat-Testverfahren mit Anwendungen  
(Aspekte der linearen Geometrie in der  
Stochastik)

Im Zuge jeder stochastischen Modellbildung müssen wir Wahrscheinlichkeitshypothesen formulieren. Stehen mehrere Hypothesen mit möglicherweise unterschiedlichen "subjektiven Vorbewertungen" zur Auswahl oder lassen sich die Hypothesen selbst als Zufallsgrößen mit gegebener a priori Verteilung auffassen, dann kann uns die Bayessche Regel in Verbindung mit "Indizien" bei der Auswahl geeigneter Hypothesen helfen, wie wir im vorigen Kapitel gesehen haben. Vielfach sind diese Voraussetzungen aber nicht gegeben und man möchte nur wissen, ob eine bestimmte Hypothese mit gewissen experimentellen Daten vereinbar ist, oder ob sich verschiedene Phänomene durch eine gemeinsame Hypothese modellieren lassen. In solchen Fällen bieten sich Chi-Quadrat-Tests an, die auch im Schulbereich verbreitet sind. Allerdings gelten die mathematischen Hintergründe als derart "verwickelt", daß diese Verfahren im Unterricht meist nur "rezepthaft" zur Anwendung kommen. Bisher gibt es erst wenige Ansätze für Elementarisierungen, die sich entweder auf Computersimulationen oder auf Tests vom Freiheitsgrad 1 beschränken: Schneider/Stein (1980), Strick (1981), Borovcnik (1984), Lind/Scheid (1984), Scheid (1986).

Wir werden in Teil II (Kapitel 3 bis 5) versuchen, diese Lücke zu schließen und zeigen, wie man die Testverfahren auf elementarem Niveau verstehen kann. Dabei kommen Mittel der linearen Geometrie zum Einsatz, die den schulischen Standard nicht überschreiten. (Ebenen im Raum, Projektionen, Freiheitsgrad als Dimension von Irrfahrten.)

Wir schließen mit der Entdeckung, daß sich unsere (gezinkten) Quader durch Boltzmann-Verteilungen beschreiben lassen, egal mit welcher Wurftechnik wir arbeiten.

### 3 Chi-Quadrat-Anpassungstest

#### 3.1 Grenzen Bayesscher Betrachtungsweise

Der unter motivationspsychologischen und erkenntnisleitenden Gesichtspunkten hohe Wert der Bayesschen Betrachtungsweise wird nicht geschmälert, wenn man auch im Unterricht ihre Grenzen verdeutlicht. Mit folgenden Beispielen führen wir in die Thematik des Chi-Quadrat-Anpassungstests ein und vertiefen gleichzeitig die in 2.1 und 2.6 vorgebrachten Argumente.

##### Beispiel 1:

Sandra benutzt ihren Holzbleistift zum "Würfeln" der Zufallszahlen 1, 2 und 3: Die Aufschrift "Castell" bedeutet 1, "Germany" bedeutet 2 und "HB" bedeutet 3. Sie erhielt bei zwei Serien mit je 30 Experimenten:

1 Castell	2 Germany	3 HB	
8	9	13	t=1.4
5	13	12	t=3.8

Tab. 3-1

Sind die Bleistift - Zufallszahlen gleichverteilt? Ist das seltenere Auftreten von 1 nur Resultat einer ungünstigen Stichprobe oder kann man schon auf eine Unsymmetrie im Bleistift schließen?

##### Beispiel 2:

Frank findet "Würfeln mit Bleistiften" unzeitgemäß. Nach einer Idee der Firma Texas Instruments (Anleitung zu einem programmierbaren Taschenrechner) hat er seinen Computer programmiert: Die Zuweisung  $x := \text{frac}((x+\pi)^n)$  (n fest) liefert eine Folge von Zufallszahlen x zwischen 0 und 1 (frac bedeutet "Nachkommateil"). Durch die Vereinbarung

$$y := \begin{cases} 1, & \text{falls } 0 \leq x < 1/3 \\ 2, & \text{falls } 1/3 \leq x < 2/3 \\ 3, & \text{falls } 2/3 \leq x < 1 \end{cases}$$

erhalten wir eine "elektronische Variante" von Sandras Bleistiftwürfel. Frank potenziert bei seinem Zufallsgenerator mit  $n=2$  und erhielt in zwei Serien mit je 900 Versuchen und den Startzahlen 0.98 bzw. 0.16:

1	2	3	
274	353	273	$t=14.046$
296	335	269	$t= 7.340$

Tab. 3-2

Ist das häufige Auftreten von 2 eine zufällige Erscheinung unserer Stichproben - oder sollte man die Hypothese einer Gleichverteilung doch besser fallenlassen?

### Beispiel 3:

Nicole behauptet: Beim Quader ist die Wahrscheinlichkeit  $P$  einer Seite proportional zu (einer Potenz) der Fläche  $F$  ( $H_i: P \sim F^i, i=0,1,\dots,5$ ). Wir vergleichen die experimentellen Ergebnisse von Jola und Sabine (je 100 Versuche) und Herrn Remeir (1800 Versuche) mit diesen Wahrscheinlichkeitshypothesen. (Wegen der experimentellen Ergebnisse und der Rechnungen vergleiche man Tab. 1-5 und Abschnitt 1.2.2).

Seiten	1;6	2;5	3;4			
Fläche $F$	2.99	2.60	4.60	( $\text{cm}^2$ )		
<b>Experimente:</b>						
Jola	26	15	59	(100)		
Sabine	19	11	70	(100)		
Remeir	397	210	1193	(1800)		
	(22%)	12%	66%			
<b>Hypothesen:</b>						
	(%)			Jola	Sabine	Remeir
$H_0: P \sim F^0$	33.3	33.3	33.3	$t=31.46$	$t=61.46$	$t=908.3$
$H_1: P \sim F$	29.3	25.5	45.1	$t= 8.96$	$t=25.59$	$t=346.0$
$H_2: P \sim F^2$	24.3	18.3	57.4	$t= 0.78$	$t= 6.84$	$t= 72.0$
$H_3: P \sim F^3$	18.9	12.4	68.7	$t= 4.61$	$t= 0.19$	$t= 12.0$
$H_4: P \sim F^4$	13.9	8.0	78.1	$t=21.30$	$t= 3.82$	$t=148.1$
$H_5: P \sim F^5$	9.9	4.9	85.2	$t=55.02$	$t=18.65$	$t=512.3$

Tab. 3-3 Vergleich von sechs Hypothesen mit drei Versuchen

Die Vermutung "je größer die Grundfläche desto größer die Wahrscheinlichkeit" wird bestätigt, aber kann man beispielsweise die Hypothese  $H_1$  einer Proportionalität noch aufrecht erhalten?

Zur Beantwortung solcher Fragen benutzt man den Chi-Quadrat-Anpassungstest. Ein Arbeiten mit der Bayesschen Regel nach Kapitel 2 ist hier dem Problem nicht angemessen. Bei Sandras Bleistift und Heiners Zufallsgenerator gibt es (zumindest vor Versuchsdurchführung) zur Gleichverteilungshypothese keine plausiblen Alternativen. Insofern macht es wenig Sinn, diese Hypothese gegen irgendwelche Alternativen abzuwägen. Man möchte eben nur wissen, ob die experimentellen Daten mit einer Gleichverteilung vereinbar sind oder nicht. Ebenso steht es beim Quader mit Nicoles Proportionalitätsannahme. Hier haben wir zwar sechs Alternativhypothesen (allgemeiner ein Kontinuum  $H_b$ ,  $b \in \mathbb{R}$ ). Ein Subjektivist könnte versucht sein, diesen Alternativen gewisse a priori Wahrscheinlichkeiten zuzuordnen und diese unter Benutzung der Indizien (Stichproben) revidieren. So würden sich aus einer a priori Gleichverteilung für Jola, Sabine und Herrn Remeir folgende a posteriori Wahrscheinlichkeiten (auf sechs Nachkommastellen gerundet) ergeben:

	a priori	a posteriori		
		Jola(100)	Sabine(100)	Remeir (1800)
$H_0$	16.66%	0.000030%	0.000000%	0.000000%
$H_1$	16.66%	1.174954%	0.000168%	0.000000%
$H_2$	16.66%	84.440645%	2.415851%	0.000000%
$H_3$	16.66%	14.370933%	82.176574%	100.000000%
$H_4$	16.66%	0.013438%	15.358426%	0.000000%
$H_5$	16.66%	0.000000%	0.048981%	0.000000%

Tab. 3-4 Revision einer Gleichverteilung nach Bayes

Jola würde der Hypothese  $H_2$ , Sabine der Hypothese  $H_3$  am meisten vertrauen. Herr Remeir müßte von  $H_3$  "felsenfest" überzeugt sein.

Zwar wird auch ein Objektivist der Feststellung zustimmen, daß die experimentellen Daten sehr viel stärker für  $H_3$  sprechen als für die fünf anderen Hypothesen. Zur Festlegung auf irgend welche Wahrscheinlichkeiten für  $H_0, \dots, H_5$  könnte man ihn aber nicht bewegen, da man solche Wahrscheinlichkeiten nicht frequentistisch interpretieren kann. Ihnen entsprechen keine relativen Häufigkeiten. Die Hypothesen sind hier keine Zufallsgrößen. Man kann schließlich nicht wie in Kapitel 2 in einem Vorversuch zwischen sechs Quadern wählen, die durch die sechs Hypothesen  $H_0, \dots, H_5$  beschrieben werden, um dann aus einer Stichprobe zu erschließen, welchen Quader man erwischt hat.

Aber auch ein Subjektivist wird beispielsweise die Wahrscheinlichkeit 100% für  $H_3$  nach 1800 Versuchen (Tab. 3-4) sehr vorsichtig in Abhängigkeit von der a priori Verteilung interpretieren. Zwar ist von den sechs betrachteten Hypothesen  $H_3$  noch am besten mit den experimentellen Ergebnissen vereinbar, aber Hypothesen der Form  $H_b: P \sim F^b$  mit beliebigem reellem Parameter  $b$  (oder ganz andere Familien von Verteilungen) waren ja durch die a priori Verteilung von vornherein ausgeschlossen. So wird die Frage, ob das stochastische Verhalten des Quaders wirklich mit  $H_3$  oder irgend einer anderen Verteilung der Familie  $H_b$  vereinbar ist, durch Tab. 3-4 nicht beantwortet. Man ist auf einen Anpassungstest angewiesen. Wir werden ihn (wie in den obigen Beispielen) zunächst für drei Wahrscheinlichkeiten  $p_1, p_2, p_3$  untersuchen, da dieser Fall besonders anschaulich ist und doch alle wesentlichen Aspekte in sich birgt. Verallgemeinerungen folgen in 3.4.

### 3.2 Experimenteller Teil

#### 3.2.1 Testgröße

Um zu testen, ob die (drei) möglichen Ausgänge eines Zufallsexperimentes durch die Wahrscheinlichkeiten  $p_1$ ,  $p_2$ ,  $p_3$  beschrieben werden, macht man  $n$  Experimente und vergleicht die erwarteten absoluten Häufigkeiten  $np_i$  mit den beobachteten Häufigkeiten  $n_i$  ( $i=1,2,3$ ) durch Berechnung der Testgröße

$$t = \sum_{i=1}^3 \frac{(n_i - np_i)^2}{np_i}$$

So ergibt sich für Jolas Quaderexperiment und Nicoles Proportionalitätshypothese  $H_1$  ( $p_1=0.293$ ,  $p_2=0.255$ ,  $p_3=0.451$ ) bei  $n=100$  Versuchen

$$t = \frac{(26-29.3)^2}{29.3} + \frac{(15-25.5)^2}{25.5} + \frac{(59-45.1)^2}{45.1} = 8.96$$

(Die Testwerte  $t$  entnimmt man auch für die anderen Beispiele den Tabellen 7-1 bis 7-3.) Ist  $t$  sehr groß, so sind die Abweichungen zwischen den erwarteten und den tatsächlich eingetretenen Häufigkeiten beträchtlich. Und das spricht gegen die angenommenen Wahrscheinlichkeiten, die wir dann "verwerfen". Doch was bedeutet "sehr groß" ?

Wir suchen zunächst eine experimentelle Antwort durch Computersimulation und bestimmen danach die "asymptotische" Verteilung der Testgröße  $t$  mit Wahrscheinlichkeitsrechnung und linearer Geometrie.

### 3.2.2 Computersimulation

Wir programmieren einen Zufallsgenerator so, daß er die Zahlen 1, 2 und 3 mit vorgegebenen Wahrscheinlichkeiten  $p_1=p_2=p_3=1/3$  (bzw.  $p_1=0.29$ ,  $p_2=0.26$ ,  $p_3=0.45$ ) liefert. In den Spalten von Tab. 3-5 können wir studieren, wie dann die Häufigkeitsverteilungen und die zugehörigen Testgrößen (bei  $n=30$ ,  $n=900$ ,  $n=100$  Versuchen) schwanken:  $t$ -Werte bis 4.5 sind "nicht außergewöhnlich",  $t$ -Werte über 6 treten dagegen nur "sehr selten" auf, wenn die Verteilung, die man testet, "stimmt". Das belegt Tab. 3-6 noch sehr viel deutlicher. Sie zeigt, wie sich bei einer Computersimulation je 1000 Testwerte auf Intervalle der Breite 0.5 verteilen.

Tabelle 3-6 legt folgende Vermutung nahe: Wenn die Wahrscheinlichkeitsverteilung  $P=(p_1, p_2, p_3)$ , die man testet, "stimmt", dann ist die Verteilung der Testgröße

- (a) weitgehend unabhängig von der Versuchszahl  $n$
- (b) weitgehend unabhängig von den Wahrscheinlichkeiten  $p_1, p_2, p_3$ .

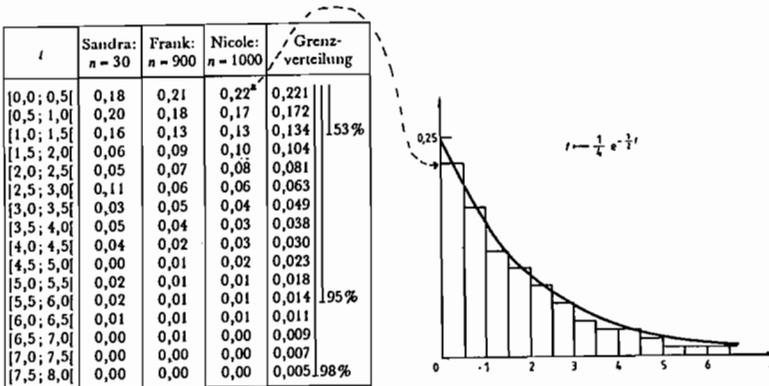
Weitere Simulationen erhärten die Vermutung:

- (c) Mit wachsendem Versuchsumfang ( $n \rightarrow \infty$ ) strebt die Verteilung der Testgröße  $t$  (unabhängig von  $p_1, p_2, p_3$ ) einer Grenzverteilung zu.

Das ist der (überzeugende) Grund dafür, daß sich die Testgröße  $t$  in der Mathematik "durchgesetzt" hat.

Test auf Gleichverteilung				Quader ( $H_1$ )							
Bleistift		Computer									
Sandra: $p_1 = p_2 = p_3 = 1/3$ Versuchsumfang $n = 30$		Frank: $p_1 = p_2 = p_3 = 1/3$ Versuchsumfang $n = 900$		Nicole: $p_1 = 0,29$ $p_2 = 0,26$ $p_3 = 0,45$ Versuchsumfang $n = 100$							
$n_1$	$n_2$	$n_3$	$t$	$n_1$	$n_2$	$n_3$	$t$	$n_1$	$n_2$	$n_3$	$t$
10	11	9	0,20	277	311	312	2,65	36	22	42	2,51
10	12	8	0,80	296	286	318	1,79	30	20	50	1,97
11	7	12	1,40	309	300	291	0,54	27	25	48	0,38
9	10	11	0,20	301	304	295	0,14	20	32	48	4,38
12	3	15	7,80	301	309	290	0,61	38	24	38	4,04
8	12	10	0,80	296	296	308	0,32	27	31	42	1,30
12	6	12	2,40	303	284	313	1,45	30	33	37	3,34
12	9	9	0,60	298	296	306	0,19	33	21	46	1,54
12	11	7	1,40	271	322	307	4,58	33	29	38	1,99
11	13	6	2,60	306	326	268	5,79	24	32	44	2,27
6	13	11	2,60	283	292	325	3,26	36	21	43	2,74
13	10	7	1,80	291	314	295	1,01	28	28	44	0,21
12	9	9	0,60	280	319	301	2,54	29	30	41	0,97
11	10	9	0,20	289	322	289	2,42	27	31	42	1,30
9	8	13	1,40	303	324	273	4,38	26	33	41	2,55
11	9	10	0,20	308	308	284	1,28	29	24	47	0,24
9	13	8	1,40	299	300	301	0,01	23	30	47	1,95
10	9	11	0,20	317	274	309	3,49	38	16	46	6,66
8	12	10	0,80	279	317	304	2,49	24	23	53	2,63
10	11	9	0,20	309	303	288	0,78	25	25	50	1,15

Tab. 3-5 Schwankungen der absoluten Häufigkeiten und der Testgröße. (Die Verteilungen, die man testet, werden vom Computer tatsächlich erzeugt.)



Tab. 3-6 Relative Häufigkeitsverteilungen (1000 Testwerte) und Chi-Quadrat-Grenzverteilung der Testgröße  $t$  bei Test auf die richtigen Wahrscheinlichkeiten

Exkurs (im Vorgriff auf 3.3)

Die Grenzverteilung ist die Chi-Quadrat-Verteilung mit zwei Freiheitsgraden. Sie besitzt, wie wir in 3.3 zeigen werden, die Dichte  $f(t) = 0.5 e^{-0.5t}$ . Folglich gilt

$$P(a < t < b) \approx 0.5 \int_a^b e^{-0.5t} dt.$$

Für Intervalle der Länge 0.5 erhalten wir die letzte Spalte von Tab. 3-6

$$P(a-0.25 \leq t < a+0.25) \approx 0.5 \int_{a-0.25}^{a+0.25} e^{-0.5t} dt \approx 0.25 e^{-0.5a}.$$

Wer die lineare Geometrie in Abschnitt 3.3 überschlagen, seinen Schülern die analytische Form der Grenzverteilung aber trotzdem nicht vorenthalten möchte, zeigt: Die relativen Häufigkeiten verringern sich "von Intervall zu Intervall" (bei Schrittweite 0.5) um durchschnittlich den gleichen Faktor 0.78 ( $e^{-0.25}$ ), bei Schrittweite 1 also um durchschnittlich 0.62 ( $e^{-0.5}$ ). Die Wahrscheinlichkeiten zeigen exponentielles Verhalten

$$P(a-0.5 \leq t < a+0.5) \approx c \cdot e^{-0.5a}$$

Die Konstante  $c=0.5$  bestimmt man aus  $c \int e^{-0.5t} dt = 1$ .

Dieses Vorgehen ist allerdings nicht verallgemeinerungsfähig, da die Testgröße  $t$  nur bei Tests auf drei Wahrscheinlichkeiten exponentiell verteilt ist (vgl. Abschnitt 3.4).

Wir kehren zu Tab. 3-6 zurück und erkennen (auch ohne Benutzung der Grenzverteilung): Wenn die Wahrscheinlichkeitsverteilung stimmt, dann liegt die Testgröße  $t$  in ca. 95% aller Fälle unter dem Wert 6. Anders ausgedrückt: Größere Werte als 6 sprechen sehr gegen die Richtigkeit der getesteten Wahrscheinlichkeiten. Man verwirft dann die Wahrscheinlichkeitshypothese auf dem 5%-Signifikanzniveau. Blicken wir mit diesen Erkenntnissen auf die Eingangsbeispiele zurück:

(1) Bei Sandras Bleistift sprechen die Testgrößen 1.4 und 3.8 nicht gegen eine Gleichverteilung der Zufallszahlen.

(2) Jola, Sabine und Herr Remeir haben dagegen bei den Testwerten 8.96, 25.59 und 346.0 die Hypothese  $H_1$ , daß sich bei einem Quader die Wahrscheinlichkeiten zu den Grundflächen proportional verhalten, auf dem 5%-(ja auf dem 1%) Signifikanzniveau zurückzuweisen.

Jola kann dagegen  $H_2$  und  $H_3$ , Sabine  $H_3$  und  $H_4$  nicht zurückweisen (vgl. Tab. 3-3). Herr Remeir muß alle Hypothesen verwerfen. (Bei Bayesscher Betrachtungsweise wäre er nach 3.1 von  $H_3$  "felsfest" überzeugt gewesen.) Streng genommen sind allerdings solche Signifikanz-Aussagen nur für die Hypothese  $H_1$  zulässig, die als einzige vor jeglichem Experiment aufgestellt worden war. Für  $H_0, H_2, \dots, H_5$  sind sie bedenklich, da diese Hypothesen im Unterricht erst nach Kenntnis der experimentellen Ergebnisse aufgestellt wurden. Man versuchte, die Hypothesen den experimentellen Ergebnissen anzupassen.

(3) Auch Frank hat mit den Testgrößen 14.046 und 7.340 die Hypothese auf dem 5%-Signifikanzniveau zu verwerfen, daß bei seinem Zufallsgenerator die Zahlen 1, 2 und 3 gleichverteilt sind. (Zur Ehrenrettung der Firma Texas-Instruments: sie empfiehlt, beim Erzeugen der Zufallszahlen mit  $n=8$  zu potenzieren, nicht mit  $n=2$ , wie Frank es aus Gründen der Rechengeschwindigkeit tat. Außerdem hat sich Frank vertippt: statt  $\pi:=3.1415926$  findet sich in seinem Programm die Zeile  $\pi:=1.1415926$  - und dieser Fehler hatte fatale Folgen.)

## 3.2.3 Versuchsumfang

Zur Abrundung des experimentellen Teils untersuchen wir den Einfluß des Versuchsumfanges  $n$  für einen Anpassungstest, bei dem die Wahrscheinlichkeiten nicht stimmen.

n = 30				n = 900			
$n_1$	$n_2$	$n_3$	$t$	$n_1$	$n_2$	$n_3$	$t$
12	4	14	5,60	233	230	437	93,86
8	9	13	1,40	224	259	417	70,49
8	6	16	5,60	250	239	411	61,81
8	9	13	1,40	225	246	429	83,94
12	9	9	0,60	241	242	417	68,45
8	7	15	3,80	234	255	411	62,34
5	10	15	5,00	251	234	415	66,61
10	6	14	3,20	263	233	404	55,58
13	8	9	1,40	254	224	422	75,92
10	11	9	0,20	260	232	408	59,63
9	9	12	0,60	251	246	403	53,09
15	3	12	7,80	250	226	424	77,84
7	10	13	1,80	261	210	429	87,54
6	10	14	3,20	258	236	406	56,99
10	5	15	5,00	260	241	399	49,61

Tab. 3-7 Einfluß des Versuchsumfanges bei Test auf falsche Wahrscheinlichkeiten

fang überschreitet  $t$  trotz falscher Hypothese "praktisch nie" die kritische Marke 6. Die falsche Hypothese läßt sich somit nur in Ausnahmefällen verwerfen. Ganz anders sieht die Lage bei einem höheren Versuchsumfang  $n=900$  aus (zweite Spalte von Tabelle 3-7). Die falsche Hypothese wird wegen der sehr hohen Testwerte stets verworfen.

Die gleiche Beobachtung können wir bei Sandras Bleistift machen. Fassen wir die beiden Versuchsserien aus Tab. 3-1 zusammen, so erhalten wir die Häufigkeitsverteilung

1	2	3		
13	22	25	$n=60$	$t=3.9$

Ein Anpassungstest auf Gleichverteilung liefert mit  $t=3.9$

In Tabelle 3-7 haben wir einen Zufallsgenerator mit  $p_1=0.29$ ,  $p_2=0.26$ ,  $p_3=0.45$  programmiert. Wir testen aber die (natürlich falsche) Hypothese, 1, 2 und 3 seien gleichverteilt. Die erste Spalte zeigt einige Häufigkeitsverteilungen für den Versuchsumfang  $n=30$  und die zugehörigen Testgrößen bei Test auf Gleichverteilung. Sie sind zwar (verglichen mit der ersten Spalte von Tab. 3-5) etwas erhöht, aber bei so kleinem Versuchsum-

einen höheren Testwert als jede der einzelnen Versuchsserien mit 30 Versuchen ( $t=1.4$  und  $t=3.8$ ). Das liegt daran, daß beide Serien die gleiche Tendenz ("1 zu selten") aufweisen. In der Tat können wir die Hypothese der Gleichverteilung auf dem 1%-Signifikanzniveau zurückweisen ( $t > 9.2$ ), wenn wir Sandras Bleistift noch öfter rollen lassen:

1	2	3		
24	43	53	n=120	t=17.5

Bei hinreichend hohem Versuchsumfang  $n$  läßt sich durch den Anpassungstest praktisch jede falsche Hypothese signifikant verwerfen - auch wenn sie nur sehr wenig (unbedeutend) von den wahren Wahrscheinlichkeiten abweicht. Nicht immer sind signifikante Abweichungen bedeutsam. In der Umgangssprache werden die beiden Begriffe leider oft synonym verwendet. Vgl. hierzu auch Diepgen (1985).

### 3.3 Theoretischer Teil

Wir wollen nun die experimentell gewonnenen Aussagen theoretisch untermauern, indem wir zeigen: Wenn die Wahrscheinlichkeiten  $p_1, p_2, p_3$  "stimmen", dann ist die Testgröße  $t$  für hinreichend großes  $n$  (unabhängig von diesen Wahrscheinlichkeiten) Chi-Quadrat-verteilt mit der Dichte

$$f(t) = 0.5e^{-0.5t}, \text{ es gilt also } P(a \leq t < b) \approx 0.5 \int_a^b e^{-0.5t} dt.$$

Hiermit wurde die letzte Spalte von Tab. 3-6 berechnet. Auch ist es nun leicht, die kritischen Grenzen der Testgröße  $t$  zu berechnen. Für die 95%-Grenze  $x$  erhalten wir

$$0.5 \int_0^x e^{-0.5t} dt = 0.95, \text{ also } x = -2 \ln(0.05) \approx 5.99.$$

Entsprechend ergibt sich für die 99%-Grenze ("Hypothese hochsignifikant verwerfen")  $x \approx 9.21$  und für die 90%-Grenze ("skeptisch werden")  $x \approx 4.61$ .

## 3.3.1 Vektorielle Deutung der Testgröße

Wir schreiben die Testgröße  $t$  wie folgt

$$\begin{aligned} t &= \frac{(n_1 - np_1)^2}{np_1} + \frac{(n_2 - np_2)^2}{np_2} + \frac{(n_3 - np_3)^2}{np_3} \\ &= \frac{1}{n} \left[ \left( \frac{n_1 - np_1}{\sqrt{p_1}} \right)^2 + \left( \frac{n_2 - np_2}{\sqrt{p_2}} \right)^2 + \left( \frac{n_3 - np_3}{\sqrt{p_3}} \right)^2 \right] \\ &= (1/n) \left\| \vec{v}_n \right\|^2 = \left\| (1/\sqrt{n}) \vec{v}_n \right\|^2 \end{aligned}$$

und deuten sie als das Normquadrat eines "Zustandsvektors"

$$\vec{v}_n := \left( \frac{n_1 - np_1}{\sqrt{p_1}} \mid \frac{n_2 - np_2}{\sqrt{p_2}} \mid \frac{n_3 - np_3}{\sqrt{p_3}} \right) \in \mathbb{R}^3$$

Wir wechseln nun den Standpunkt und stellen uns die Durchführung eines Anpassungstests als  $n$ -stufigen zeitabhängigen Prozeß vor ( $n$ =Versuchsumfang). Wir erhalten eine Folge von Zustandsvektoren  $\vec{v}_0 = \vec{0}, \vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ , also eine räumliche Irrfahrt in  $\mathbb{R}^3$  und untersuchen, wie sich diese Vektoren "von Schritt zu Schritt" ändern.

Ergibt sich im  $n$ -ten Schritt beispielsweise das Ergebnis "3", so erhöht sich die absolute Häufigkeit  $n_3$  (und natürlich auch der bisherige Versuchsumfang  $n$ ) um 1. Der "alte" Zustandsvektor

$$\vec{v}_n = \left( \frac{n_1 - np_1}{\sqrt{p_1}} \mid \frac{n_2 - np_2}{\sqrt{p_2}} \mid \frac{n_3 - np_3}{\sqrt{p_3}} \right)$$

verändert sich zum "neuen" Zustandsvektor

$$\vec{v}_{n+1} = \left( \frac{n_1 - (n+1)p_1}{\sqrt{p_1}} \mid \frac{n_2 - (n+1)p_2}{\sqrt{p_2}} \mid \frac{n_3 + 1 - (n+1)p_3}{\sqrt{p_3}} \right)$$

Die Differenz

$$\vec{e}_3 := \vec{v}_{n+1} - \vec{v}_n = \left( \frac{-p_1}{\sqrt{p_1}} \mid \frac{-p_2}{\sqrt{p_2}} \mid \frac{1-p_3}{\sqrt{p_3}} \right)$$

ist vom bisherigen Testverlauf unabhängig. Wir bezeichnen sie als "Schrittvektor". Ergibt sich in einem Schritt des Anpassungstests (statt 3) das Ergebnis 1 oder 2, so erhalten wir völlig analog den "neuen" Zustandsvektor aus dem "alten" durch Addition der Schrittvektoren

$$\vec{e}_1 := \left( \frac{1-p_1}{\sqrt{p_1}} \mid \frac{-p_2}{\sqrt{p_2}} \mid \frac{-p_3}{\sqrt{p_3}} \right) \quad \vec{e}_2 := \left( \frac{-p_1}{\sqrt{p_1}} \mid \frac{1-p_2}{\sqrt{p_2}} \mid \frac{-p_3}{\sqrt{p_3}} \right).$$

Wie man (wegen  $1-p_1-p_2-p_3=0$ ) unmittelbar erkennt, stehen die Schrittvektoren  $\vec{e}_i$  orthogonal zu  $\vec{s} := (\sqrt{p_1} \mid \sqrt{p_2} \mid \sqrt{p_3})$ , es gilt  $\langle \vec{e}_i \mid \vec{s} \rangle = 0$  ( $i=1,2,3$ ). Gleiches gilt für die Folge der Zustandsvektoren:  $\langle \vec{v}_n \mid \vec{s} \rangle = 0$  ( $n=0,1,2,\dots$ ). Die Irrfahrt ist also nicht dreidimensional, wie es auf den ersten Blick erscheinen mag, sondern nur zweidimensional. Daher spricht man bei diesem Anpassungstest vom Freiheitsgrad zwei. Die "Reduktion des Freiheitsgrades" resultiert letztendlich aus der einschränkenden Bedingung, daß die Besetzungszahlen sich zum Versuchsumfang  $n$  aufsummieren. Entsprechende Reduktionen beobachten wir auch beim Unabhängigkeitstest (4.4.1) und bei Anpassungstests mit Parametern (vgl. die Bemerkung am Anfang von 5).

Während eines Anpassungstests (für die Wahrscheinlichkeiten  $p_1, p_2, p_3$ ) läuft eine zufallsgesteuerte Irrfahrt in einer Ursprungsebene  $E$  des  $\mathbb{R}^3$  ab. Die Ebene steht orthogonal zu  $\vec{s} = (\sqrt{p_1} \mid \sqrt{p_2} \mid \sqrt{p_3})$ . Sie wird aufgespannt durch je zwei der drei Schrittvektoren  $\vec{e}_1, \vec{e}_2, \vec{e}_3$ . Mit Wahrscheinlichkeit  $p_1$  bewegt man sich um  $\vec{e}_1$ , mit Wahrscheinlichkeit  $p_2$  um  $\vec{e}_2$  und mit Wahrscheinlichkeit  $p_3$  bewegt man sich um  $\vec{e}_3$  fort. Normiert man den momentanen Zustandsvektor  $\vec{v}_n$  mit  $1/\sqrt{n}$ , dann ergibt sich die Anpassungstestgröße als dessen Normquadrat:  $t = \left\| \frac{1}{\sqrt{n}} \vec{v}_n \right\|^2$ .

## 3.3.2 Zweidimensionale Irrfahrt

Zum Studium der Irrfahrt in der Ebene  $E$  ( $L\vec{s}$ ) des  $\mathbb{R}^3$  untersuchen wir die gegenseitige Lage der Schrittvektoren. Mit

$$\langle \vec{e}_1 | \vec{e}_2 \rangle = \left\langle \left( \frac{1-p_1}{\sqrt{p_1}} \mid \frac{-p_2}{\sqrt{p_2}} \mid \frac{-p_3}{\sqrt{p_3}} \right) \mid \left( \frac{-p_1}{\sqrt{p_1}} \mid \frac{1-p_2}{\sqrt{p_2}} \mid \frac{-p_3}{\sqrt{p_3}} \right) \right\rangle$$

$$= -(1-p_1) - (1-p_2) + p_3 = -2 + p_1 + p_2 + p_3 = -1$$

und aus Gründen der Symmetrie

$$\langle \vec{e}_i | \vec{e}_j \rangle = -1 \quad \text{für } i \neq j \quad (1)$$

erhalten wir ein unerwartet einfaches Ergebnis. Die Skalarprodukte zwischen je zwei Schrittvektoren sind gleich, sie hängen von  $p_1$ ,  $p_2$  und  $p_3$  gar nicht ab. Für die Länge der Schrittvektoren ergibt eine kurze Rechnung

$$\|\vec{e}_i\| = \sqrt{\langle \vec{e}_i | \vec{e}_i \rangle} = \sqrt{(1/p_i) - 1} \quad (i=1, 2, 3) \quad (2)$$

Je größer die Wahrscheinlichkeit  $p_i$  desto kürzer ist also der zugehörige Schrittvektor  $\vec{e}_i$ . Aus (1) folgt:

Die Projektionen von je zwei Schrittvektoren auf den dritten gleichen einander. (1a)

So erhält man für die Projektion von  $\vec{e}_2$  auf  $\vec{e}_1$ :

$$\frac{\langle \vec{e}_1 | \vec{e}_2 \rangle}{\|\vec{e}_1\|^2} \vec{e}_1 = \frac{-\vec{e}_1}{\|\vec{e}_1\|^2} = \frac{-p_1}{1-p_1} \vec{e}_1$$

und für die Projektion von  $\vec{e}_3$  auf  $\vec{e}_1$  ergibt sich der gleiche Wert (Abb. 3-8). Auch die Winkel zwischen den Schrittvektoren lassen sich jetzt berechnen:

$$\chi(\vec{e}_i, \vec{e}_j) = \arccos \left( \frac{-1}{\|\vec{e}_i\| \|\vec{e}_j\|} \right) = \arccos \left( -\sqrt{\frac{p_i p_j}{(1-p_i)(1-p_j)}} \right) \quad (3)$$

So ergibt sich beispielsweise für die Wahrscheinlichkeiten  $p_1 = 0.29$ ,  $p_2 = 0.26$ ,  $p_3 = 0.45$  (vgl. Hypothese  $H_1$  aus Tab. 3-3)

$$\begin{aligned} \|\vec{e}_1\| &= 1.565 & \angle(\vec{e}_1, \vec{e}_2) &= 112.3^\circ \\ \|\vec{e}_2\| &= 1.687 & \angle(\vec{e}_1, \vec{e}_3) &= 125.3^\circ \\ \|\vec{e}_3\| &= 1.106 & \angle(\vec{e}_2, \vec{e}_3) &= 122.4^\circ \end{aligned}$$

Abb. 3-8 veranschaulicht die Lage dieser Schrittvektoren.

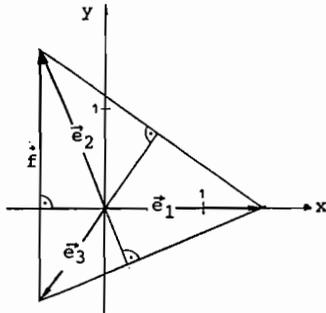


Abb. 3-8  
Die Projektionen je  
zweier Schrittvektoren  
auf den dritten  
gleichem einander

Möchte man die ebene Irrfahrt auf einem Computerbildschirm mit Polarkoordinaten ("Turtelgraphik") verfolgen, so reichen diese Informationen über Längen und Winkel. Wer lieber mit kartesischen Koordinaten arbeitet, lege die x-Achse der Ebene  $E$  in Richtung eines Schrittvektors (etwa  $\vec{e}_1$ ) und die y-Achse orthogonal dazu in Richtung  $\vec{f} := \vec{e}_2 - \vec{e}_3$  (denn  $\langle \vec{e}_2 - \vec{e}_3 | \vec{e}_1 \rangle = 0$ ). Mit den Beziehungen

$$\langle \vec{e}_2 | \vec{f} \rangle = \langle \vec{e}_2 | \vec{e}_2 - \vec{e}_3 \rangle = \|\vec{e}_2\|^2 - \langle \vec{e}_2 | \vec{e}_3 \rangle = 1/p_2 - 1 + 1 = 1/p_2 ,$$

$$\langle \vec{e}_3 | \vec{f} \rangle = -1/p_3 ,$$

$$\|\vec{f}\|^2 = \langle \vec{e}_2 - \vec{e}_3 | \vec{e}_2 - \vec{e}_3 \rangle = \langle \vec{e}_2 | \vec{f} \rangle - \langle \vec{e}_3 | \vec{f} \rangle = 1/p_2 + 1/p_3$$

berechnet man die Koordinatendarstellungen der Schrittvektoren bezüglich dieses Koordinatensystems zu

$$\vec{e}_1 = \begin{pmatrix} \Delta x_1 \\ \Delta y_1 \end{pmatrix} = \begin{pmatrix} |\vec{e}_1| \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{1-p_1}{p_1}} \\ 0 \end{pmatrix} \quad (4)$$

$$\vec{e}_2 = \begin{pmatrix} \Delta x_2 \\ \Delta y_2 \end{pmatrix} = \begin{pmatrix} \frac{\langle \vec{e}_2 | \vec{e}_1 \rangle}{\|\vec{e}_1\|} \\ \frac{\langle \vec{e}_2 | \vec{f} \rangle}{\|\vec{f}\|} \end{pmatrix} = \begin{pmatrix} -\sqrt{\frac{p_1}{1-p_1}} \\ \sqrt{\frac{p_3}{p_2(1-p_1)}} \end{pmatrix} \quad (5)$$

$$\vec{e}_3 = \begin{pmatrix} \Delta x_3 \\ \Delta y_3 \end{pmatrix} = \begin{pmatrix} \frac{\langle \vec{e}_3 | \vec{e}_1 \rangle}{\|\vec{e}_1\|} \\ \frac{\langle \vec{e}_3 | \vec{f} \rangle}{\|\vec{f}\|} \end{pmatrix} = \begin{pmatrix} -\sqrt{\frac{p_1}{1-p_1}} \\ -\sqrt{\frac{p_2}{p_3(1-p_1)}} \end{pmatrix} \quad (6)$$

In Abb. 3-9 hat ein Plotter den Verlauf einer Irrfahrt protokolliert, die zu einem Anpassungstest auf  $p_1=0.29$ ,  $p_2=0.26$   $p_3=0.45$  gehört ( $n=100$  Schritte). Man erkennt die Schrittvektoren aus Abb. 3-8 in verkleinertem Maßstab wieder. Die Endposition ist durch "P" markiert. Die Testgröße  $t$  erhält man, indem man den Zustandsvektor  $\vec{v}_{100} = \vec{OP}$  mit  $1/\sqrt{100}$  normiert und das zugehörige Normquadrat berechnet. In Abb. 3-9 gilt

$$\|\vec{v}_{100}\| = 23.3, \text{ also } t = \left\| \frac{1}{\sqrt{100}} \vec{v}_{100} \right\|^2 = 5.43.$$

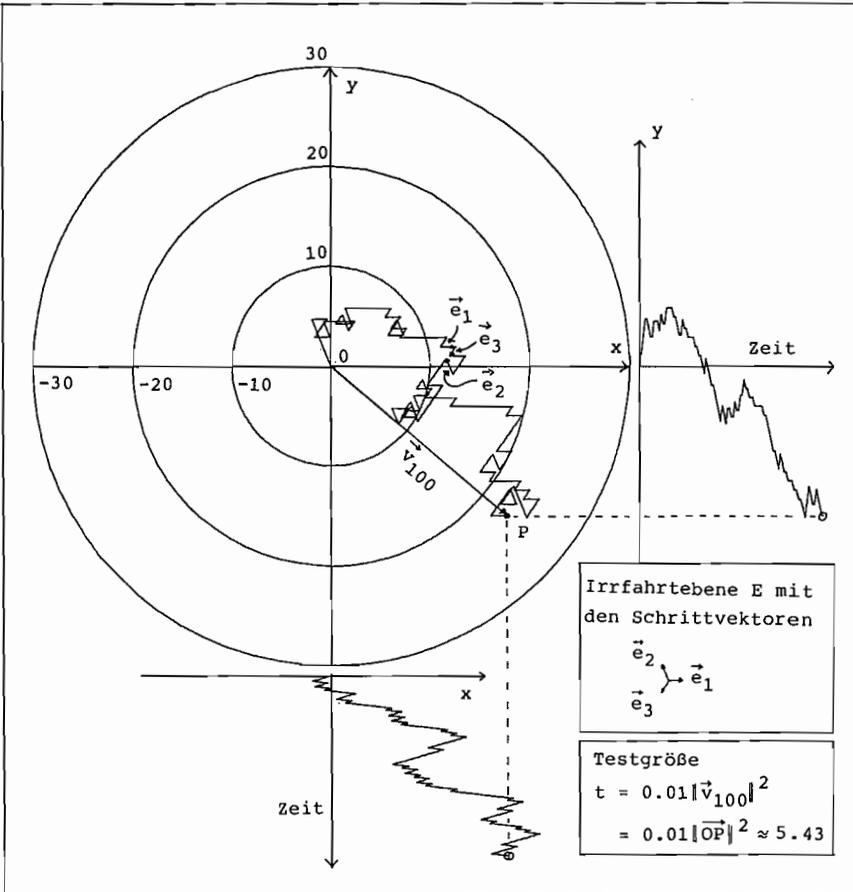


Abb. 3-9

Ebene Irrfahrt zum Anpassungstest für  $p_1=0.29$ ,  $p_2=0.26$ ,  $p_3=0.45$  (Freiheitsgrad 2) mit Projektionen auf die Koordinatenachsen. Man macht 100 Schritte mit einem der Schrittvektoren  $\vec{e}_1$ ,  $\vec{e}_2$ ,  $\vec{e}_3$ . Die Irrfahrt beginnt mit 2,2,2,1,3,3,1,1,2,1,... Die Projektion auf die x-Achse gehört zu einem Anpassungstest, bei dem man die Ergebnisse 2 und 3 nicht mehr unterscheidet, also  $p_2$  und  $p_3$  zusammenfaßt (Freiheitsgrad 1).

## 3.3.3 Zweidimensionale Normalverteilung

Um herauszufinden, wie sich die Endpositionen unserer Irrfahrt auf die Ebene E verteilen, untersuchen wir die zugehörigen x- und y- Koordinaten (Abb. 3-9). Die Schrittvektoren (4), (5) und (6) zeigen, um welche Beträge  $\Delta x$ ,  $\Delta y$  sich diese Koordinaten schrittweise in Abhängigkeit von den Ergebnissen 1, 2, 3 verändern. Der Erwartungswert von  $\Delta x$  ist

$$\mu(\Delta x) = p_1 \sqrt{\frac{1-p_1}{p_1}} - p_2 \sqrt{\frac{p_1}{1-p_1}} - p_3 \sqrt{\frac{p_1}{1-p_1}} = 0$$

Der Erwartungswert von  $\Delta y$  ist ebenfalls Null. Für die Varianz von  $\Delta x$  erhalten wir

$$\sigma^2(\Delta x) = p_1 \frac{1-p_1}{p_1} + p_2 \frac{p_1}{1-p_1} + p_3 \frac{p_1}{1-p_1} = 1$$

Der gleiche Wert ergibt sich für die Varianz von  $\Delta y$ .

Wenn wir die zweidimensionale Irrfahrt auf die Koordinatenachsen der Ebene E projizieren, so erhalten wir eindimensionale Irrfahrten, deren Schritte unabhängig von  $p_1$ ,  $p_2$ ,  $p_3$  den Erwartungswert 0 und die Varianz 1 haben.

Nach dem Zentralen Grenzwertsatz (Kapitel 6) sind dann die x- und y- Koordinaten des normierten Zustandsvektors  $(1/\sqrt{n}) \vec{v}_n$  asymptotisch standardnormalverteilt, denn sie sind Summen unabhängiger Zufallsgrößen:

$$x = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Delta x_i \quad y = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Delta y_i$$

Die x-Koordinate liegt also mit Wahrscheinlichkeit

$$g(x) dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

in einem Intervall um x mit Breite dx, analog für y.

Da die Projektionen der Schrittvektoren  $\vec{e}_2$  und  $\vec{e}_3$  - die allein eine Bewegung in  $y$ -Richtung verursachen - auf die  $x$ -Achse ( $\vec{e}_1$ ) gleich groß sind, kann man aus der  $x$ -Koordinate keine Rückschlüsse auf die  $y$ -Koordinate ziehen. Die Koordinaten am Ende der Irrfahrt sind also voneinander unabhängig. Nach der Produktregel erhalten wir: Die Wahrscheinlichkeit, daß die normierte Irrfahrt in einem Rechteck um den Punkt  $P(x|y)$  der Ebene  $E$  mit Seitenlängen  $dx$  und  $dy$  endet, ist

$$\varphi(x,y)dx dy = g(x)dx \cdot g(y)dy = \frac{1}{2\pi} e^{-0.5(x^2+y^2)} .$$

### 3.3.4 Chi-Quadrat-Verteilung

Die Summe von Quadraten unabhängiger standardnormalverteilter Zufallsgrößen ist Chi-Quadrat-verteilt. Da die  $x$ - und  $y$ -Koordinaten standardnormalverteilt sind, ist auch die Testgröße  $t = \|(1/\sqrt{n}) \vec{v}_n\|^2 = x^2 + y^2$  Chi-Quadrat-verteilt. Ein Induktionsbeweis hierfür findet sich in vielen Büchern der Wahrscheinlichkeitsrechnung. Für die Schule ist eine anschauliche Begründung vorzuziehen. Zur Berechnung der Dichte der Testgröße  $t$  fragen wir: wo muß die (normierte) zweidimensionale Irrfahrt enden, damit sich ein Testwert im Intervall  $[t; t+dt[$  ergibt?

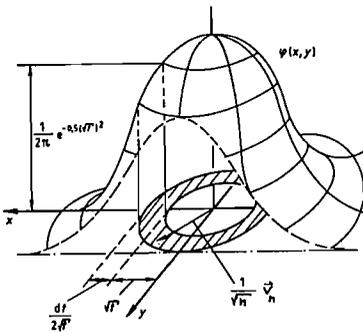


Abb. 3-10  
Bestimmung der Verteilung der Testgröße  $t$  durch Integration über Kreisringe

Antwort: in einem Kreisring um den Ursprung mit Radius  $r=\sqrt{t}$  und Breite  $dt/(2\sqrt{t})$  (denn  $(\sqrt{t}+dt/(2\sqrt{t}))^2 \approx t+dt$ ). Die Wahrscheinlichkeit dafür ist aber (vgl. Abb. 3-10)

$$\underbrace{\frac{1}{2\pi} e^{-0.5(\sqrt{t})^2}}_{\text{Höhe}} \cdot \underbrace{2\pi\sqrt{t}}_{\text{Umfang}} \cdot \underbrace{\frac{dt}{2\sqrt{t}}}_{\text{Breite}} = 0.5 e^{-0.5t} dt. \quad (7)$$

des Kreisringes

Damit ist  $f(t) = 0.5 e^{-0.5t}$  die gesuchte Dichte der Testgröße  $t$ . Es handelt sich, wie schon erwähnt, um die Chi-Quadrat-Verteilung mit Freiheitsgrad 2.

### 3.4 Verallgemeinerungen

Wir skizzieren, wie sich unsere Überlegungen auf Anpassungstests mit mehr oder weniger als drei Wahrscheinlichkeiten übertragen lassen und wie Anpassungstests mit verschiedenen Freiheitsgraden geometrisch zusammenhängen.

#### 3.4.1 Anpassungstests mit beliebigen Freiheitsgraden

Bisher haben wir den Anpassungstest nur für Hypothesen mit drei Wahrscheinlichkeiten  $p_1, p_2, p_3$  untersucht, weil dieser Fall besonders anschaulich ist und dennoch alles wesentliche in sich birgt. Zur Verallgemeinerung genügt eine kurze Skizze: Möchte man  $k$  Wahrscheinlichkeiten  $p_1, p_2, \dots, p_k$  ( $k > 2$ ) mit einem Anpassungstest überprüfen, benutzt man die Testgröße

$$t = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} = \left\| \frac{1}{\sqrt{n}} \vec{v}_n \right\|^2 \quad \text{mit}$$

$$\vec{v}_n = \left( \frac{n_1 - np_1}{\sqrt{np_1}} \mid \frac{n_2 - np_2}{\sqrt{np_2}} \mid \dots \mid \frac{n_k - np_k}{\sqrt{np_k}} \right) \in \mathbb{R}^k.$$

Während des Tests läuft jetzt eine Irrfahrt in dem  $k-1$  dimensionalen Unterraum  $E$  von  $\mathbb{R}^k$  (also einer Hyperebene) ab, die auf dem Vektor  $\vec{s} := (\sqrt{p_1} \mid \sqrt{p_2} \mid \dots \mid \sqrt{p_k})$  orthogonal steht. Deswegen spricht man nun vom Freiheitsgrad  $k-1$ .

Die  $k$  Schrittvektoren haben die Form

$$\vec{e}_1 = \left( \frac{1-p_1}{\sqrt{p_1}} \mid \frac{-p_2}{\sqrt{p_2}} \mid \frac{-p_3}{\sqrt{p_3}} \mid \dots \mid \frac{-p_k}{\sqrt{p_k}} \right)$$

$$\vec{e}_2 = \left( \frac{-p_1}{\sqrt{p_1}} \mid \frac{1-p_2}{\sqrt{p_2}} \mid \frac{-p_3}{\sqrt{p_3}} \mid \dots \mid \frac{-p_k}{\sqrt{p_k}} \right)$$

$\vdots$

$$\vec{e}_k = \left( \frac{-p_1}{\sqrt{p_1}} \mid \frac{-p_2}{\sqrt{p_2}} \mid \frac{-p_3}{\sqrt{p_3}} \mid \dots \mid \frac{1-p_k}{\sqrt{p_k}} \right)$$

Wieder gilt:  $\langle \vec{e}_i \mid \vec{e}_j \rangle = -1$  und  $\|\vec{e}_i\| = \sqrt{(1/p_i)-1}$ . Der Erwartungswert der Irrfahrt hat in jeder Richtung der Hyperebene  $E$  den Wert 0, die Varianz hat den Wert 1. (Das rechnet man leicht nach, indem man einen beliebigen Einheitsvektor  $\vec{u}$  der Ebene schreibt als  $\vec{u} = (a_1/\sqrt{p_1} \mid a_2/\sqrt{p_2} \mid \dots \mid a_k/\sqrt{p_k})$  mit  $\sum a_i = 0$  ( $\vec{u} \mid \vec{s}$ ) und  $\sum a_i^2/p_i = 1$  ( $\|\vec{u}\|^2 = 1$ ). Man zeigt  $\langle \vec{e}_i \mid \vec{u} \rangle = a_i/p_i$  und erhält  $\mu = \sum p_i \langle \vec{e}_i \mid \vec{u} \rangle = \sum a_i = 0$  sowie  $\sigma^2 = \sum p_i \langle \vec{e}_i \mid \vec{u} \rangle^2 = \sum a_i^2/p_i = 1$ .) Führt man in  $E$  wieder ein Orthogonalsystem ein, so ergibt ein Plausibilitätsargument (wie in 3.3.3) die Unabhängigkeit der einzelnen  $(k-1)$  Koordinaten. Die normierte Endposition der Irrfahrt  $((1/\sqrt{n}) \vec{v}_n)$  ist folglich wegen des Zentralen Grenzwertsatzes und der Produktregel  $k-1$  dimensional normalverteilt mit der Dichte

$$\begin{aligned} \varphi(x_1, x_2, \dots, x_{k-1}) &= \frac{1}{\sqrt{2\pi}^{k-1}} e^{-0.5(x_1^2 + x_2^2 + \dots + x_{k-1}^2)} \\ &= \frac{1}{\sqrt{2\pi}^{k-1}} e^{-0.5t} \end{aligned}$$

(Produkt von  $k-1$  eindimensionalen Standardnormalverteilungen). Die Verteilung der Testgröße  $t = \|(1/\sqrt{n}) \vec{v}_n\|^2 = \sum_{i=1}^{k-1} x_i^2$

ergibt sich durch Integration der Dichte  $\varphi$  jetzt nicht mehr über Kreisringe in der Ebene, sondern über Kugelschalen in  $\mathbb{R}^{k-1}$  mit der Dicke  $dt/(2\sqrt{t})$ : Die Wahrscheinlichkeit, daß sich ein Testwert zwischen  $t$  und  $t+dt$  ergibt, ist genauso groß wie die Wahrscheinlichkeit, daß die normierte Irrfahrt auf einer Kugelschale mit Radius  $\sqrt{t}$  und Dicke  $dt/(2\sqrt{t})$  endet. Sie ist analog zu (7) gegeben durch  $f(t) dt =$

$$O_{k-1}(\sqrt{t}) \quad \frac{dt}{2\sqrt{t}} \quad \frac{1}{\sqrt{2\pi}^{k-1}} e^{-0.5t}$$

Kugelfläche in $\mathbb{R}^{k-1}$ mit Radius $r = \sqrt{t}$	Dicke der Kugelschale	Wahrscheinl.- Dichte
--	--------------------------	-------------------------

Die "Kugelflächen" in  $\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3$  sind  $O_1(\sqrt{t})=2, O_2(\sqrt{t})=2\pi\sqrt{t}, O_3(\sqrt{t})=4\pi t$ . Für die höherdimensionalen Kugelflächen gilt <sup>1)</sup>

$$O_{k-1}(\sqrt{t}) = \begin{cases} \frac{(2\pi)^{(k-2)/2} \cdot 2}{(k-3)(k-5)\dots 1} \sqrt{t}^{k-2} & \text{falls } k \text{ gerade} \\ & k \geq 4 \\ \frac{(2\pi)^{(k-1)/2}}{(k-3)(k-5)\dots 2} \sqrt{t}^{k-2} & \text{falls } k \text{ ungerade} \\ & k \geq 5 \end{cases}$$

Hiermit ergibt sich für die Verteilung der Testgröße  $t$

$$f(t) = \begin{cases} \frac{\sqrt{t}^{k-3} (1/\sqrt{2\pi}) e^{-0.5t}}{(k-3)(k-5)\dots 1} & \text{falls } k \text{ gerade} \\ & k \geq 2 \\ \frac{\sqrt{t}^{k-3} 0.5 e^{-0.5t}}{(k-3)(k-5)\dots 2} & \text{falls } k \text{ ungerade} \\ & k \geq 3 \end{cases}$$

Dies ist die Dichte der Chi-Quadrat-Verteilung mit  $k-1$  Freiheitsgraden. Die letzten Formeln gelten auch für  $k=2, k=3$ , wenn man den Nenner 1 setzt. Die Chi-Quadrat-Verteilung mit 2 Freiheitsgraden ist eine Exponentialverteilung, die mit 3 Freiheitsgraden bezeichnet man auch als Maxwell'sche Geschwindigkeitsverteilung.

1) Man zeigt für das Volumen  $C_n(1)$  der  $n$ -dimensionalen Einheitskugel  $C_n(1)=C_{n-2}(1) 2\pi/n$ , benutzt  $C_n(r)=r^n C_n(1)$  und  $O_n(r)=C_n'(r)$ . (Vgl. z.B. Smirnow II S. 272.)

### 3.4.2 Reduktion des Freiheitsgrades durch Zusammenfassen von Wahrscheinlichkeiten

Betrachten wir nochmals die ebene Irrfahrt, die zum Anpassungstest für  $p_1, p_2, p_3$  gehört (Formeln (4) bis (6) sowie Abb. 3-8 und 3-9). Die x-Koordinate ändert sich beim Ergebnis 1 um

$$\Delta x = \sqrt{\frac{1-p_1}{p_1}} \quad (8a)$$

bei den Ergebnissen 2 oder 3 jeweils um

$$\Delta x = -\sqrt{\frac{p_1}{1-p_1}} \quad (8b)$$

Fassen wir nun die Wahrscheinlichkeiten  $p_2$  und  $p_3$  (zu  $p_2$ ) zusammen, indem wir auch das Ergebnis 3 in 2 umbenennen, und testen wir die Verteilung

1	2 oder 3	bzw. nach	1	2
$p_1$	$p_2 + p_3$	Umbenennung	$p_1$	$p_2 = 1 - p_1$

mit einem Anpassungstest vom Freiheitsgrad 1. Wir verwenden dann die Testgröße

$$t = \frac{(n_1 - np_1)^2}{np_1} + \frac{(n_2 - np_2)^2}{np_2} = \left\| \frac{1}{\sqrt{n}} \vec{v}_n \right\|^2$$

mit dem Zustandsvektor

$$\vec{v}_n = \left( \frac{n_1 - np_1}{\sqrt{p_1}} \quad \left| \quad \frac{n_2 - np_2}{\sqrt{p_2}} \right. \right) .$$

Die Irrfahrt, die zu diesem Anpassungstest gehört, verläuft in einem eindimensionalen Unterraum  $g \subset \mathbb{R}^2$ , der zu  $\vec{s} := (\sqrt{p_1} | \sqrt{p_2})$  orthogonal steht.

Mit Wahrscheinlichkeit  $p_1$  macht man den Schritt

$$\vec{e}_1 = \left( \frac{1-p_1}{\sqrt{p_1}} \mid \frac{-p_2}{\sqrt{p_2}} \right) \quad \|\vec{e}_1\| = \sqrt{\frac{1-p_1}{p_1}} \quad (9a)$$

mit Wahrscheinlichkeit  $p_2$  macht man den Schritt

$$\vec{e}_2 = \left( \frac{-p_1}{\sqrt{p_1}} \mid \frac{1-p_2}{\sqrt{p_2}} \right) \quad \|\vec{e}_2\| = -\sqrt{\frac{1-p_2}{p_2}} = -\sqrt{\frac{p_1}{1-p_1}} \quad (9b)$$

in entgegengesetzter Richtung zu  $\vec{e}_1$ . Aus einem Vergleich von (8a,b) mit (9a,b) entnimmt man (lax aber einprägsam formuliert): Der Anpassungstest zur Verteilung

1	2 oder 3	(Freiheitsgrad 1)
$p_1$	$p_2 + p_3$	

ist nichts anderes als die Projektion des Anpassungstests zur Verteilung

1	2	3	(Freiheitsgrad 2)
$p_1$	$p_2$	$p_3$	

auf die Gerade, die durch den Schrittvektor  $\vec{e}_1$  aufgespannt wird. Die zunächst willkürlich erscheinende Auszeichnung von  $\vec{e}_1$  als Richtungsvektor der x-Achse (in Abb. 3-8 und 3-9) hat also eine überraschende und inhaltlich sehr sinnvolle Begründung. Ebenso liefert die Projektion der zweidimensionalen Irrfahrt auf  $\vec{e}_2$  die eindimensionale Irrfahrt, die zum Anpassungstest auf die Verteilung

2	1 oder 3
$p_2$	$p_1 + p_3$

gehört. Analoge Aussagen gelten für Anpassungstests mit höheren Freiheitsgraden.

### 3.5 Historische Anmerkung

Der Anpassungstest wurde von Karl Pearson im Jahre 1900 entdeckt. Cochran (1952) analysiert Pearsons Arbeit und bezeichnet sie als eines der Fundamente moderner Statistik. Wir wollen skizzieren, wie Pearson in jener Arbeit die asymptotische Verteilung der Anpassungstestgröße - ohne das heute durchgängig benutzte Werkzeug der Charakteristischen Funktionen - erhält und Beziehungen zu unseren heuristischen Überlegungen herstellen. Pearson stellt fest:

Die absoluten Häufigkeiten  $n_1, n_2, \dots, n_k$ , die beim Test der Wahrscheinlichkeiten  $p_1, p_2, \dots, p_k$  auftreten, sind polynomial verteilt mit

$$\mu(n_i) = np_i, \quad \sigma^2(n_i) = np_i(1-p_i) \quad \text{und}$$

$$\text{kov}(n_i, n_j) = -np_i p_j \quad (1 \leq i \neq j \leq k)$$

Die zugehörige  $k \times k$ -Kovarianzmatrix  $S_k$  ist nicht regulär:

$$S_k = n \begin{pmatrix} p_1(1-p_1) & -p_1 p_2 & \dots & -p_1 p_k \\ -p_2 p_1 & p_2(1-p_2) & \dots & -p_2 p_k \\ \vdots & \vdots & \ddots & \vdots \\ -p_k p_1 & -p_k p_2 & \dots & p_k(1-p_k) \end{pmatrix} .$$

Sie wird regulär, wenn man nur die ersten  $k-1$  Zufallsgrößen betrachtet (Matrix  $S_{k-1}$ ). Die Inverse zu  $S_{k-1}$  ist

$$A = 1/n \begin{pmatrix} 1/p_1 + 1/p_k & 1/p_k & \dots & 1/p_k \\ 1/p_k & 1/p_2 + 1/p_k & \dots & 1/p_k \\ \vdots & \vdots & \ddots & \vdots \\ 1/p_k & 1/p_k & \dots & 1/p_{k-1} + 1/p_k \end{pmatrix} .$$

Nach dem Zentralen Grenzwertsatz in mehreren Veränderlichen ist dann der Zufallsvektor

$$\vec{x} := (n_1 - np_1 \mid n_2 - np_2 \mid \dots \mid n_{k-1} - np_{k-1}) \quad (10)$$

für großes  $n$  annähernd  $k-1$ -dimensional normalverteilt mit der Dichte

$$f(x_1, x_2, \dots, x_{k-1}) = C e^{-0.5 \vec{x}^t A \vec{x}} dx_1 dx_2 \dots dx_{k-1} \quad (11)$$

Pearson beweist dann, daß die quadratische Form  $Q = \vec{x}^t A \vec{x}$  Chi-Quadrat-verteilt ist mit  $k-1$  Freiheitsgraden. Er verformt das zu  $Q$  gehörige Ellipsoid in eine Kugel. (Für den Fall der Einheitsmatrix  $A=E$  stammt das Ergebnis von Helmert 1876.) Der Beweis endet mit der Beobachtung, daß die quadratische Form  $Q = \vec{x}^t A \vec{x}$  in die Anpassungstestgröße  $t$  überführt werden kann. Damit ist auch  $t$  Chi-Quadrat-verteilt mit  $k-1$  Freiheitsgraden.

Man kann wie folgt eine Beziehung zwischen Pearsons Gedanken und unserer heuristischen Begründung herstellen. Durch Weglassen der Komponente  $x_k = n_k - np_k$  in (10) projiziert Pearson die in 3.3 und 3.4 betrachtete Irrfahrt in die Ebene  $x_k=0$ , in der sich die (nicht rotationssymmetrische  $k-1$ -dimensionale) Normalverteilung (11) ergibt. Wir projizieren dagegen in die Ebene orthogonal zu  $\vec{s} = (\sqrt{p_1} \mid \sqrt{p_2} \mid \dots \mid \sqrt{p_k})$  und wählen die Koordinatenachsen so, daß die zugehörige Kovarianzmatrix zur Einheitsmatrix, die Grenzverteilung zur  $k-1$ -dimensionalen Standardnormalverteilung wird. Von dieser gelangt man unmittelbar zur Chi-Quadrat-Verteilung, ohne über allgemeine quadratische Formen nachdenken zu müssen.

Rückblickend erkennt man, daß für einen Schüler die aufgezeigten Zusammenhänge (ohne inhaltliche Verfälschung) erst durch eine sehr konkrete Darstellung und Beschränkung auf den dreidimensionalen Anschauungsraum erfahrbar werden.

#### 4 Chi-Quadrat-Unabhängigkeitstest

Die Methoden, mit denen es uns im vorigen Kapitel gelang, den Anpassungstest zu elementarisieren, lassen sich auch zur Untersuchung des nicht weniger wichtigen Unabhängigkeitstests nutzen. Obwohl uns die Parallelitäten zwischen beiden Verfahren interessieren, ist dieses Kapitel unabhängig von Kapitel 3 lesbar. Der experimentelle Teil 4.1 ist in sich abgeschlossen, der Abschnitt 4.2 dient wieder zur Vertiefung. Wir beginnen mit einigen Beispielen.

##### 4.1 Experimenteller Teil

###### Beispiel 1:

Ist die Einstellung zu Mathe geschlechtsabhängig? Im Rahmen einer Projektwoche wurde eine Datenbank zur Computer-Partnervermittlung programmiert (näheres in Kapitel 10). Etwa 40% aller Schülerinnen und Schüler eines Gymnasiums nahmen an dem Partnertest teil. Sie hatten zur Feststellung ähnlicher Interessenlagen auch die Frage zu beantworten "Magst Du das Schulfach Mathe gern (+) weniger (o) oder garnicht (-) ?" Ist die Wertschätzung geschlechtsabhängig? Eine Datenbankabfrage lieferte die Sechsfelder-Tafel:

	Mathe +	Mathe o	Mathe -		
Jungen	103 ( $n_{11}$ )	41 ( $n_{12}$ )	45 ( $n_{13}$ )	189 ( $a_1$ )	t=7.0
Mädchen	74 ( $n_{21}$ )	44 ( $n_{22}$ )	61 ( $n_{23}$ )	179 ( $a_2$ )	
	177 ( $b_1$ )	85 ( $b_2$ )	106 ( $b_3$ )	368 ( $n$ )	Tab. 4-1

Eine entsprechende Auswertung für Deutsch lieferte dagegen:

	Deutsch +	Deutsch o	Deutsch -		
Jungen	43	71	75	189	t=8.25
Mädchen	64	62	53	179	
	107	133	128	368	Tab. 4-2

Bei beiden Fächern scheinen die Tabellen auf Unterschiede

zwischen Jungen und Mädchen hinzuweisen. Sind sie möglicherweise nur ein Zufallsprodukt der Stichprobe oder steckt mehr dahinter?

Beispiel 2:

Ist die Einstellung zum Schummeln altersabhängig? Eine andere Testfrage lautet: "Was halten Sie vom Schummeln?"

- a) Dazu gehört auch etwas Geschicklichkeit
- b) Viel zu riskant
- c) Das ist charakterlos"

Die erste Zeile von Tab. 4-3 zeigt die Antworten von Schülerinnen und Schülern der Sekundarstufe I, die zweite bezieht sich auf die Sekundarstufe II:

	a)	b)	c)	
SI	151	20	22	193
SII	156	11	8	175
	307	31	30	368

$t=8.37$

Tab. 4-3

Ändert sich die Einstellung zum Schummeln im Laufe der Schulzeit oder sind die geringeren Häufigkeiten für b) und c) in der Sekundarstufe II nur zufällig zustande gekommen? (Nebenbei bemerkt: von 33 Lehrerinnen und Lehrern, die ebenfalls an dem Partnertest teilnahmen, hielten nur zwei Schummeln für zu riskant oder charakterlos.)

Beispiel 3:

Gleichheit von Wahrscheinlichkeitsverteilungen. Es ist bekannt, daß "ausgedachte Zufallszahlen" selten gleichverteilt sind. Ich bat Hiki, 50 mal die Zahlen 1, 2, und 3 rasch hintereinander "aus dem Gedächtnis" zu diktieren. Einige Tage später wiederholte ich das Experiment. Besitzen die Gedächtnis-Zufallsgeneratoren beider Tage die gleichen Wahrscheinlichkeitsverteilungen?

	1	2	3	
1. Versuch	19	17	14	50
2. Versuch	23	13	14	50
	42	30	28	100

t=0.91

Tab. 4-4

Solche und ähnliche Fragen beantwortet man mit dem Chi-Quadrat-Unabhängigkeitstest. Wir studieren ihn zunächst für Sechsfeldertafeln wie in Tab. 4-1 bis 4-4.

#### 4.1.1 Testgröße

Mit der Chi-Quadrat-Testgröße  $t$  vergleichen wir die tatsächlich eingetretenen Häufigkeiten mit den Häufigkeiten, die wir erwarten würden, wenn die beteiligten Merkmale voneinander unabhängig wären. Betrachten wir als Beispiel die Sechsfeldertafel aus Tab. 4-1. Wir denken uns dort die Randwerte fest vorgegeben. Es handelt sich um die Anzahl der am Test teilnehmenden Jungen ( $a_1=189$ ) und Mädchen ( $a_2=179$ ), die Anzahl der Teilnehmer, die dem Fach Mathematik die Wertschätzung "+" ( $b_1=177$ ), "o" ( $b_2=85$ ) und "-" ( $b_3=106$ ) entgegenbringen. Die Gesamtzahl der Teilnehmer ist  $n$  (368). Davon sind  $a_1/n$  (0.51) männlich und  $b_1/n$  (0.48) mögen Mathe (+). Wären die Merkmale Geschlecht und Wertschätzung unabhängig, so hätten wir nach der Multiplikationsregel im Feld  $n_{11}$  den Anteil  $(a_1/n)(b_1/n)$  (0.25) erwartet. Bei insgesamt  $n$  Beobachtungen ergibt sich daraus die erwartete absolute Häufigkeit  $n(a_1/n)(b_1/n) = (a_1 b_1)/n$  (90.9). Analog hätte man im Falle der Unabhängigkeit in der  $i$ -ten Zeile und  $j$ -ten Spalte die absolute Häufigkeit  $(a_i b_j)/n$  erwartet. Hier die numerischen Ergebnisse für unsere Erwartungen (Zahlen gerundet):

	Mathe +	Mathe o	Mathe -	
Jungen	90.9	43.7	54.4	189 ( $a_1$ )
Mädchen	86.1	41.3	51.6	179 ( $a_2$ )
	177 ( $b_1$ )	85 ( $b_2$ )	106 ( $b_3$ )	368 ( $n$ )

Tab. 4-5 Im Falle der Unabhängigkeit erwartete Tafel

Die Abweichungen zwischen erwarteter und beobachteter Tafel (Tab. 4-1 und Tab. 4-5) werden gemessen durch die Testgröße

$$t = \sum_{\substack{i=1..2 \\ j=1..3}} \frac{(n_{ij} - (a_i b_j)/n)^2}{(a_i b_j)/n}$$

In unserem Beispiel ergibt sich für das Fach Mathematik

$$t = \frac{(103-90.9)}{90.9} + \frac{(41-43.7)}{43.7} + \frac{(45-54.4)}{54.4} \\ + \frac{(74-86.1)}{86.1} + \frac{(44-41.3)}{41.3} + \frac{(61-51.6)}{51.6} = 7.0$$

Beim Fach Deutsch erhalten wir  $t=8.2$ . Die Testwerte zu den übrigen Beispielen entnimmt man den obigen Tabellen. Ist  $t$  sehr groß, so sind die Abweichungen zwischen den tatsächlichen und den erwarteten Häufigkeiten in den einzelnen Feldern beträchtlich. Und das spricht gegen die Hypothese einer Unabhängigkeit der beteiligten Merkmale - wir verwerfen die Unabhängigkeitshypothese. Wie schon in Kapitel 3 werden wir zunächst durch Computersimulation beantworten, was "sehr groß" bedeutet.

#### 4.1.2 Computersimulation

Wie sehen Sechsfeldertafeln aus, bei denen die beteiligten Merkmale mit Sicherheit voneinander unabhängig sind - und in welchen Grenzen schwanken dann die Testwerte "normalerweise"? Betrachten wir als Beispiel wieder Tab. 4-1 (Abhängigkeit zwischen Geschlecht und Neigung zur Mathematik). Wir denken uns eine Randbesetzung fest vorgegeben (etwa  $a_1=189$  Jungen,  $a_2=179$  Mädchen) und lassen beide Gruppen ihre Wertschätzung für Mathe mit dem gleichen Zufallsgenerator wählen. Dadurch erreichen wir, daß das Merkmal Wertschätzung tatsächlich nicht mehr vom Geschlecht abhängt.

Einstellung zu Mathe geschlechtsabhängig? $a_1=189$ Jungen $a_2=179$ Mädchen $p_1=0.48$ $p_2=0.23$ $p_3=0.29$				Einstellung zum Schummeln altersabhängig? $a_1=193$ SI Schüler $a_2=175$ SII Schüler $p_1=0.82$ $p_2=0.09$ $p_3=0.09$				Gleichheit zweier Wahrsch.-Verteilungen? $a_1=50$ 1. Versuch $a_2=50$ 2. Versuch $p_1=0.42$ $p_2=0.30$ $p_3=0.28$														
				t								t										
92	40	57	189	0.24	154	21	18	193	0.66	29	10	11	50	3.40	20	16	14	50	49	26	25	100
91	38	50	179		145	15	15	175		18	14	18	50		28	13	9	50				
183	78	107	368		299	36	33	368		46	27	27	100									
96	49	44	189	1.30	152	24	17	193	1.04	18	14	18	50	5.21	28	13	9	50	46	27	27	100
85	43	51	179		134	20	21	175		46	27	27	100									
181	92	95	368		286	44	38	368														
75	51	63	189	5.37	162	15	16	193	1.06	16	17	17	50	4.11	26	12	12	50	42	29	29	100
91	34	54	179		141	14	20	175		26	12	12	50									
166	85	117	368		303	29	36	368		42	29	29	100									
101	39	49	189	2.87	155	14	24	193	0.23	18	19	13	50	0.62	21	19	10	50	39	38	23	100
81	48	50	179		139	15	21	175		39	38	23	100									
182	87	99	368		294	29	45	368														
93	46	50	189	0.09	153	20	20	193	1.43	25	12	13	50	0.24	23	14	13	50	48	26	26	100
86	46	47	179		137	24	14	175		23	14	13	50									
179	92	97	368		290	44	34	368		48	26	26	100									
85	55	49	189	3.48	162	15	16	193	0.12	18	16	16	50	0.54	21	16	13	50	39	32	29	100
81	39	59	179		148	12	15	175		21	16	13	50									
166	94	108	368		310	27	31	368		39	32	29	100									
92	49	48	189	2.83	160	19	14	193	3.62	24	16	10	50	3.22	20	12	18	50	44	28	28	100
82	38	59	179		143	11	21	175		20	12	18	50									
174	87	107	368		303	30	35	368		44	28	28	100									
85	48	56	189	2.21	152	22	19	193	5.78	23	15	12	50	3.47	14	20	16	50	37	35	28	100
89	34	56	179		154	10	11	175		14	20	16	50									
174	82	112	368		306	32	30	368		37	35	28	100									
89	45	55	189	1.14	149	16	28	193	9.17	23	14	13	50	0.39	20	15	15	50	43	29	28	100
93	42	44	179		147	19	9	175		20	15	15	50									
182	87	99	368		296	35	37	368		43	29	28	100									
82	46	61	189	0.33	162	16	15	193	0.76	17	14	19	50	0.75	20	15	15	50	37	29	34	100
83	41	55	179		144	13	18	175		20	15	15	50									
165	87	116	368		306	29	33	368		37	29	34	100									

Tab. 4-6

Computersimulation: Schwankung von Besetzungszahlen und Testgröße  $t$  in simulierten Sechsfeldertafeln, bei denen die beteiligten Merkmale voneinander unabhängig sind. In der ersten und zweiten Zeile einer jeden Tafel wurde der gleiche Zufallsgenerator verwendet.

Bei unserem Zufallsgenerator verwenden wir die Wahrscheinlichkeiten  $p_1=0.48$  (für Mathe +),  $p_2=0.23$  (für Mathe o) und  $p_3=0.29$  (für Mathe -). Diese Wahrscheinlichkeiten entsprechen etwa den relativen Häufigkeiten von Tab. 4-1. Analog verfahren wir in den anderen Beispielen "Schummeln" und "Gedächtniszufallszahlen".

Den Spalten von Tab. 4-6 entnimmt man, wie die Sechsfelder- tafeln und die zugehörigen Testwerte schwanken. Tab. 6-7 zeigt noch genauer, wie sich bei einer Computersimulation je 1000 Testwerte auf die Intervalle der Breite 0.5 verteilen. Nur in ca. 5% aller Fälle ergaben sich Testwerte über 6.

t	Mathe $a_1=189$ $a_2=179$	Schummeln $a_1=193$ $a_2=177$	Z.-Zahlen $a_1=50$ $a_2=50$	Grenz- verteilung	
[0.0 ; 0.5[	0.22	0.24	0.22	0.221	53%
[0.5 ; 1.0[	0.16	0.16	0.14	0.172	
[1.0 ; 1.5[	0.14	0.10	0.13	0.134	
[1.5 ; 2.0[	0.10	0.11	0.11	0.104	
[2.0 ; 2.5[	0.08	0.07	0.08	0.081	
[2.5 ; 3.0[	0.06	0.07	0.06	0.063	
[3.0 ; 3.5[	0.05	0.04	0.05	0.049	
[3.5 ; 4.0[	0.03	0.05	0.03	0.038	
[4.0 ; 4.5[	0.03	0.02	0.03	0.030	
[4.5 ; 5.0[	0.03	0.04	0.02	0.023	
[5.0 ; 5.5[	0.02	0.03	0.02	0.018	95%
[5.5 ; 6.0[	0.01	0.01	0.02	0.014	
[6.0 ; 6.5[	0.01	0.01	0.01	0.011	
[6.5 ; 7.0[	0.01	0.01	0.02	0.009	
[7.0 ; 7.5[	0.01	0.01	0.01	0.007	98%
[7.5 ; 8.0[	0.01	0.01	0.00	0.005	

Tab. 4-7 Relative Häufigkeitsverteilungen (1000 Testwerte) und Grenzverteilung der Testgröße t bei Unabhängigkeit der Merkmale in einer Sechsfeldertafel

Tab. 4-7 läßt uns (völlig analog zu Tab. 3-6) vermuten:

Wenn die beteiligten Merkmale unabhängig voneinander sind, dann ist die Verteilung der Testgröße  $t$

- (a) weitgehend unabhängig vom Stichprobenumfang ( $a_1, a_2, n$ ),
- (b) weitgehend unabhängig von den zugrunde liegenden Wahrscheinlichkeiten, mit denen die Merkmale verteilt sind.
- (c) In ca. 95% aller Fälle liegt der Wert der Testgröße unter 6, in ca. 99% aller Fälle liegt er unter 9.
- (d) Mit wachsendem Versuchsumfang ( $a_1, a_2, n \rightarrow \infty$ ) strebt die Verteilung der Testgröße  $t$  einer Grenzverteilung zu.

Blicken wir mit diesen Erkenntnissen auf die Eingangsbeispiele zurück:

Beispiel 1: Bei den Tafeln, die die Zusammenhänge zwischen Geschlecht und Wertschätzung der Schulfächer Deutsch und Mathematik beschreiben, ergeben sich Testwerte über 6. Wir müssen die Hypothese einer Unabhängigkeit (auf dem 5%-Signifikanzniveau) zurückzuweisen. (Für das Fach Englisch ergab unsere Untersuchung übrigens einen sehr kleinen Testwert  $t=0.39$ . Hier brauchen wir die Hypothese der Unabhängigkeit zwischen Geschlecht und Wertschätzung nicht zu verwerfen.)

Beispiel 2: Wenn die Einstellung zum Schummeln von der Schulstufe (SI oder S II) unabhängig wäre, dann wäre ein Testwert über 8 (8.37) extrem unwahrscheinlich. Auch hier haben wir die Hypothese der Unabhängigkeit zu verwerfen.

Beispiel 3: Bei unserem Zufallszahlenexperiment erhielten wir mit  $t=0.91$  einen kleinen Wert. Die Hypothese, daß in beiden Versuchen die gleiche Wahrscheinlichkeitsverteilung verwendet wurde, läßt sich nicht zurückweisen. Eine spätere Kontrolle mit 200 Zufallszahlen je Experiment lieferte einen noch geringeren Wert  $t=0.58$ . Hiki hat also einen "recht stabilen" Gedächtnis-Zufallszahlengenerator, der allerdings keine gleichverteilten Zahlen liefert.

## 4.2 Theoretischer Teil

Wir werden nun die in 4.1 experimentell gewonnenen Aussagen theoretisch untermauern, indem wir zeigen: Wenn bei einem Sechsfeldertest die Merkmale voneinander unabhängig sind, dann ist die Testgröße  $t$  (genau wie beim Anpassungstest auf drei Wahrscheinlichkeiten) für hinreichend großen Versuchsumfang Chi-Quadrat-verteilt mit 2 Freiheitsgraden (Dichtefunktion  $f(t) = 0.5 e^{-0.5t}$ ).

### 4.2.1 Vektorielle Deutung der Testgröße

Wie schon bei den Simulationen in Tab. 4-6 betrachten wir die Zeilensummen (Teilgruppen vom Umfang  $a_1, a_2$ , etwa 189 Jungen, 179 Mädchen) als fest vorgegeben. Die relativen Gruppengrößen kürzen wir ab mit  $q_i := a_i/n$ ,  $i=1,2$ . In beiden Gruppen zählen wir, wie oft sich die fraglichen Merkmale (etwa  $1 \hat{=} +$ ,  $2 \hat{=} o$ ,  $3 \hat{=} -$ ) ergeben. Bei Zutreffen der Unabhängigkeitshypothese sind die (unbekannten) Wahrscheinlichkeiten für diese Merkmale gleich groß:

$$\begin{aligned} p_1 & \text{ für } + \text{ (Spaltensumme } b_1) & b_1 \approx np_1 \\ p_2 & \text{ für } o \text{ (Spaltensumme } b_2) & b_2 \approx np_2 \\ p_3 & \text{ für } - \text{ (Spaltensumme } b_3) & b_3 \approx np_3 \end{aligned}$$

Wir können dann die Chi-Quadrat-Testgröße

$$t = \sum_{\substack{i=1..2 \\ j=1..3}} \frac{(n_{ij} - (a_i b_j)/n)^2}{(a_i b_j)/n} = \sum_{\substack{i=1..2 \\ j=1..3}} \frac{(n_{ij} - q_i b_j)^2}{q_i b_j}$$

approximieren durch die Hilfs-Testgröße

$$u = \sum_{\substack{i=1..2 \\ j=1..3}} \frac{(n_{ij} - q_i b_j)^2}{q_i (np_j)} = \frac{1}{n} \sum_{\substack{i=1..2 \\ j=1..3}} \left( \frac{n_{ij} - q_i b_j}{\sqrt{q_i p_j}} \right)^2,$$

indem wir nicht die Spaltensummen  $b_j$  sondern deren Erwartungswerte  $np_j \approx b_j$  zur Normierung im Nenner heranziehen.

Die Grenzverteilung ( $n \rightarrow \infty$ ) bleibt davon unberührt. Wenn wir zeigen können, daß  $u$  asymptotisch Chi-Quadrat-verteilt ist mit 2 Freiheitsgraden, dann wird das gleiche auch für  $t$  gelten. (Wegen der in der Praxis unbekanntenen Werte für  $p_i$  ist  $u$  nur von theoretischem Interesse.) Wir deuten  $u$  (ähnlich wie in Kapitel 3) vektoriell

$$u = (1/n) \|\vec{v}_n\|^2 = \|(1/\sqrt{n}) \vec{v}_n\|^2$$

über das Normquadrat eines "Zustandsvektors"  $\vec{v}_n \in \mathbb{R}^6$ .

$$\vec{v}_n = \begin{pmatrix} \frac{n_{11}-q_1b_1}{\sqrt{q_1p_1}} & \frac{n_{12}-q_1b_2}{\sqrt{q_1p_2}} & \frac{n_{13}-q_1b_3}{\sqrt{q_1p_3}} \\ \frac{n_{21}-q_2b_1}{\sqrt{q_2p_1}} & \frac{n_{22}-q_2b_2}{\sqrt{q_2p_2}} & \frac{n_{23}-q_2b_3}{\sqrt{q_2p_3}} \end{pmatrix}$$

und stellen uns die Durchführung eines Unabhängigkeitstests nun als  $n (=a_1+a_2)$  - stufigen zeitabhängigen Prozeß vor. In jeder der  $a_1+a_2$  Stufen (= Befragungen) erhöht sich eine Besetzungszahl  $n_{11}, \dots, n_{23}$  der Sechsfeldertafel und eine der Spaltensummen  $b_1, \dots, b_3$ ;  $q_1$  und  $q_2$  bleiben konstant. Die Folge der Sechsfeldertafeln definiert auf diese Weise eine Folge von Zustandsvektoren  $\vec{v}_0 = \vec{0}, \vec{v}_1, \dots, \vec{v}_n$ , also eine Irrfahrt in  $\mathbb{R}^6$ . Wir untersuchen, wie sich diese Vektoren "von Schritt zu Schritt" ändern. Antwortet etwa einer der  $a_1$  Jungen mit "+" ( $\hat{=}$ 1), so erhöhen sich die absolute Häufigkeit  $n_{11}$  und die Spaltensumme  $b_1$  je um 1. Der "alte" Zustandsvektor

$$\vec{v} = \begin{pmatrix} \frac{n_{11}-q_1b_1}{\sqrt{q_1p_1}} & \frac{n_{12}-q_1b_2}{\sqrt{q_1p_2}} & \frac{n_{13}-q_1b_3}{\sqrt{q_1p_3}} \\ \frac{n_{21}-q_2b_1}{\sqrt{q_2p_1}} & \frac{n_{22}-q_2b_2}{\sqrt{q_2p_2}} & \frac{n_{23}-q_2b_3}{\sqrt{q_2p_3}} \end{pmatrix}$$

verändert sich folglich zum "neuen" Zustandsvektor

$$\vec{v}' = \begin{pmatrix} \frac{(n_{11}+1)-q_1(b_1+1)}{\sqrt{q_1 p_1}} & \frac{n_{12}-q_1 b_2}{\sqrt{q_1 p_2}} & \frac{n_{13}-q_1 b_3}{\sqrt{q_1 p_3}} \\ \frac{n_{21}-q_2(b_1+1)}{\sqrt{q_2 p_1}} & \frac{n_{22}-q_2 b_2}{\sqrt{q_2 p_2}} & \frac{n_{23}-q_2 b_3}{\sqrt{q_2 p_3}} \end{pmatrix}.$$

Die Differenz bezeichnen wir als Schrittvektor:

$$\vec{e}_{11} := \vec{v}' - \vec{v} = \begin{pmatrix} \frac{1-q_1}{\sqrt{q_1 p_1}} & 0 & 0 \\ -\frac{q_2}{\sqrt{q_2 p_1}} & 0 & 0 \end{pmatrix} = q_2 \begin{pmatrix} \frac{1}{\sqrt{q_1 p_1}} & 0 & 0 \\ -\frac{1}{\sqrt{q_2 p_1}} & 0 & 0 \end{pmatrix} =: q_2 \vec{e}_1.$$

Ergibt sich allgemeiner bei einer Befragung in der Teilstichprobe mit Umfang  $a_i$  (bzw. Anteil  $q_i$ ) das Ergebnis  $j$  ( $j=1 \hat{=} +$ ,  $j=2 \hat{=} 0$ ,  $j=3 \hat{=} -$ ), so ändert sich der Zustandsvektor nur in der Spalte  $j$ . Wir erhalten den "neuen" Zustandsvektor aus dem "alten", indem wir den Schrittvektor  $\vec{e}_{ij}$  gemäß Tab. 4-8 addieren:

$\vec{e}_{11} = q_2 \vec{e}_1$	$\vec{e}_{12} = q_2 \vec{e}_2$	$\vec{e}_{13} = q_2 \vec{e}_3$	mit
$\vec{e}_{21} = -q_1 \vec{e}_1$	$\vec{e}_{22} = -q_1 \vec{e}_2$	$\vec{e}_{23} = -q_1 \vec{e}_3$	
$\vec{e}_1 = \begin{pmatrix} 1/\sqrt{q_1 p_1} & 0 & 0 \\ -1/\sqrt{q_2 p_1} & 0 & 0 \end{pmatrix}$	$\vec{e}_2 = \begin{pmatrix} 0 & 1/\sqrt{q_1 p_2} & 0 \\ 0 & -1/\sqrt{q_2 p_2} & 0 \end{pmatrix}$	$\vec{e}_3 = \begin{pmatrix} 0 & 0 & 1/\sqrt{q_1 p_3} \\ 0 & 0 & -1/\sqrt{q_2 p_3} \end{pmatrix}$	

Tab. 4-8 Räumliche Schrittvektoren

Wie man sieht, sind die Schrittvektoren  $\vec{e}_{2j}$  (die zu den  $a_2$  Mädchen gehören) zu den Schrittvektoren  $\vec{e}_{1j}$  (der  $a_1$  Jungen) entgegengesetzt gerichtet. Während eines Unabhängigkeitstests läuft also eine  $n$ -stufige Irrfahrt in dem durch  $\vec{e}_1$ ,  $\vec{e}_2$  und  $\vec{e}_3$  aufgespannten dreidimensionalen Unterraum  $S$  ( $\subset \mathbb{R}^6$ ) ab. Tatsächlich kann die Irrfahrt aber nicht "irgendwo" in  $S$  enden: Da sich wegen  $\sum_j n_{1j} = a_1$  und  $\sum_j q_1 b_j = a_1$  die

Zähler von  $\vec{v}_n$  in jeder Zeile zu Null addieren, steht der Zustandsvektor  $\vec{v}_n$  am Ende der Irrfahrt orthogonal zu

$$\vec{o}_1 := \begin{pmatrix} \sqrt{P_1} & \sqrt{P_2} & \sqrt{P_3} \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad \vec{o}_2 := \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{P_1} & \sqrt{P_2} & \sqrt{P_3} \end{pmatrix},$$

also auch orthogonal zu deren Linearkombination

$$\begin{aligned} \vec{n} &:= \sqrt{q_2} \vec{o}_1 - \sqrt{q_1} \vec{o}_2 = \sqrt{q_1 q_2} \begin{pmatrix} \frac{P_1}{\sqrt{q_1 P_1}} & \frac{P_2}{\sqrt{q_1 P_2}} & \frac{P_3}{\sqrt{q_1 P_3}} \\ \frac{-P_1}{\sqrt{q_2 P_1}} & \frac{-P_2}{\sqrt{q_2 P_2}} & \frac{-P_3}{\sqrt{q_2 P_3}} \end{pmatrix} \\ &= \sqrt{q_1 q_2} (P_1 \vec{e}_1 + P_2 \vec{e}_2 + P_3 \vec{e}_3). \end{aligned}$$

(Die Kombination wurde so gewählt, daß  $\vec{n}$  selber im dreidimensionalen Raum  $S$  liegt und normiert ist.)

Die dreidimensionale Irrfahrt, die während eines Sechsfeldertests abläuft, endet also mit Sicherheit in der Ursprungsebene  $E$  von  $S$ , die zu  $\vec{n}$  orthogonal liegt ("Freiheitsgrad 2"). Deswegen reicht es, wenn wir im folgenden die zweidimensionale Irrfahrt verfolgen, die aus der dreidimensionalen durch Projektion auf  $E$  entsteht. Der Endpunkt der Irrfahrt (Ortsvektor  $\vec{v}_n$ ) und die Testgröße  $u = \|(1/\sqrt{n}) \vec{v}_n\|^2$  (für die wir uns ja interessieren) bleiben hiervon unberührt, da  $\vec{v}_n$  sowieso in  $E$  liegt.

#### 4.4.2 Zweidimensionale Irrfahrt

Die "räumlichen Schrittvektoren"  $\vec{e}_{ij}$  sind nach Tab. 4-8 skalare Vielfache der (orthogonalen) Basisvektoren  $\vec{e}_j$ . Wir berechnen deren Projektionen  $\vec{f}_j := \vec{e}_j - \langle \vec{e}_j | \vec{n} \rangle \vec{n}$  und erhalten die "ebenen Schrittvektoren"  $\vec{f}_{ij}$  durch skalare Vervielfachung von  $\vec{f}_j$  ( $1 \leq i \leq 2$ ,  $1 \leq j \leq 3$ ). Mit

$$\|\vec{e}_j\|^2 = \frac{1}{p_j} \frac{1}{q_1 q_2} \quad \text{und} \quad (1)$$

$$\langle \vec{e}_j | \vec{n} \rangle = \langle \vec{e}_j | \sqrt{q_1 q_2} (p_1 \vec{e}_1 + p_2 \vec{e}_2 + p_3 \vec{e}_3) \rangle = \sqrt{q_1 q_2} p_j \|\vec{e}_j\|^2 = \frac{1}{\sqrt{q_1 q_2}}$$

ergibt sich für die Projektionen

$$\vec{f}_j := \vec{e}_j - \langle \vec{e}_j | \vec{n} \rangle \vec{n} = \vec{e}_j - (p_1 \vec{e}_1 + p_2 \vec{e}_2 + p_3 \vec{e}_3). \quad (2)$$

Hier eine Koordinatendarstellung als Beispiel:

$$\vec{f}_1 = (1-p_1) \vec{e}_1 - p_2 \vec{e}_2 - p_3 \vec{e}_3 = \begin{pmatrix} \frac{1-p_1}{\sqrt{q_1 p_1}} & \frac{-p_2}{\sqrt{q_1 p_2}} & \frac{-p_3}{\sqrt{q_1 p_3}} \\ \frac{-(1-p_1)}{\sqrt{q_1 p_1}} & \frac{p_2}{\sqrt{q_1 p_2}} & \frac{p_3}{\sqrt{q_1 p_3}} \end{pmatrix}. \quad (3)$$

Mit (2) oder (3) erhalten wir für Längen und Skalarprodukte

$$\|\vec{f}_j\| = \frac{1}{\sqrt{q_1 q_2}} \sqrt{\frac{1}{p_j} - 1} \quad (1 \leq j \leq 3) \quad (4)$$

$$\langle \vec{f}_j | \vec{f}_k \rangle = \frac{-1}{q_1 q_2} \quad (1 \leq j, k \leq 3; j \neq k) \quad (5)$$

Für die Winkel ergibt sich

$$\angle(\vec{f}_j, \vec{f}_k) = \arccos \left( \frac{\langle \vec{f}_j | \vec{f}_k \rangle}{\|\vec{f}_j\| \|\vec{f}_k\|} \right) = \arccos \left( -\sqrt{\frac{p_j p_k}{(1-p_j)(1-p_k)}} \right). \quad (6)$$

Die ebenen Schrittvektoren  $\vec{f}_{1j} \in E$  sind skalare Vielfache:

$$\vec{f}_{1j} = q_2 \vec{f}_j \quad \text{und} \quad \vec{f}_{2j} = -q_1 \vec{f}_j.$$

Zur Veranschaulichung lassen wir in dem Beispiel "Geschlecht und Wertschätzung von Mathe" unseren Computer per Simulation 100 Personen befragen und zwar 25 Jungen und 75

Mädchen ( $q_1=0.25$ ,  $q_2=0.75$ ). Wir nehmen an, unabhängig vom Geschlecht werden die drei möglichen Antworten mit den Wahrscheinlichkeiten  $p_1=0.48$ ,  $p_2=0.23$ ,  $p_3=0.29$  gegeben. Dann erhält man für die ebenen Schrittvektoren (Zahlen gerundet):

$$\begin{aligned} \angle(\vec{f}_1, \vec{f}_2) &= 121.7^\circ & \|\vec{f}_1\| &= 2.4 & \|\vec{f}_{11}\| &= 1.8 & \|\vec{f}_{21}\| &= 0.6 \\ \angle(\vec{f}_1, \vec{f}_3) &= 110.4^\circ & \|\vec{f}_2\| &= 4.2 & \|\vec{f}_{12}\| &= 3.2 & \|\vec{f}_{22}\| &= 1.0 \\ \angle(\vec{f}_2, \vec{f}_3) &= 127.9^\circ & \|\vec{f}_3\| &= 3.6 & \|\vec{f}_{13}\| &= 2.7 & \|\vec{f}_{23}\| &= 0.9 \end{aligned} \quad (7)$$

die wir in Abb. 4-9 veranschaulichen

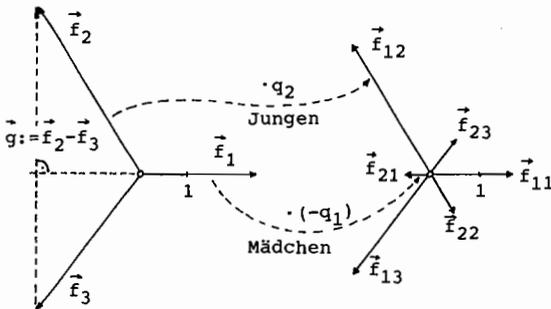


Abb. 4-9a

Projektionen der Basisvektoren  $\vec{e}_1, \vec{e}_2, \vec{e}_3$  auf die Ebene E

Abb. 4-9b

Schrittvektoren in E entstehen durch Vervielfachen

Die Winkel zwischen den Schrittvektoren werden allein durch die Antwortwahrscheinlichkeiten  $p_1, p_2, p_3$  bestimmt (die nach Hypothese geschlechtsunabhängig sind). Die relativen Stichprobenumfänge  $q_1, q_2$  beeinflussen dagegen die Längen der Schrittvektoren. So werden in unserem Beispiel dreimal so viele Mädchen befragt, "zum Ausgleich" sind die Schrittvektoren  $\vec{f}_{1j}$  der Jungen dreimal so lang wie die Schrittvektoren  $\vec{f}_{2j}$  der Mädchen. In Abb. 4-10 hat ein Plotter 100 Schritte protokolliert, die zu den Daten aus (7) und Abb. 4-9 gehören. Zuerst wurden die 25 Jungen "befragt" (lange Schrittvektoren), danach die 75 Mädchen (kurze

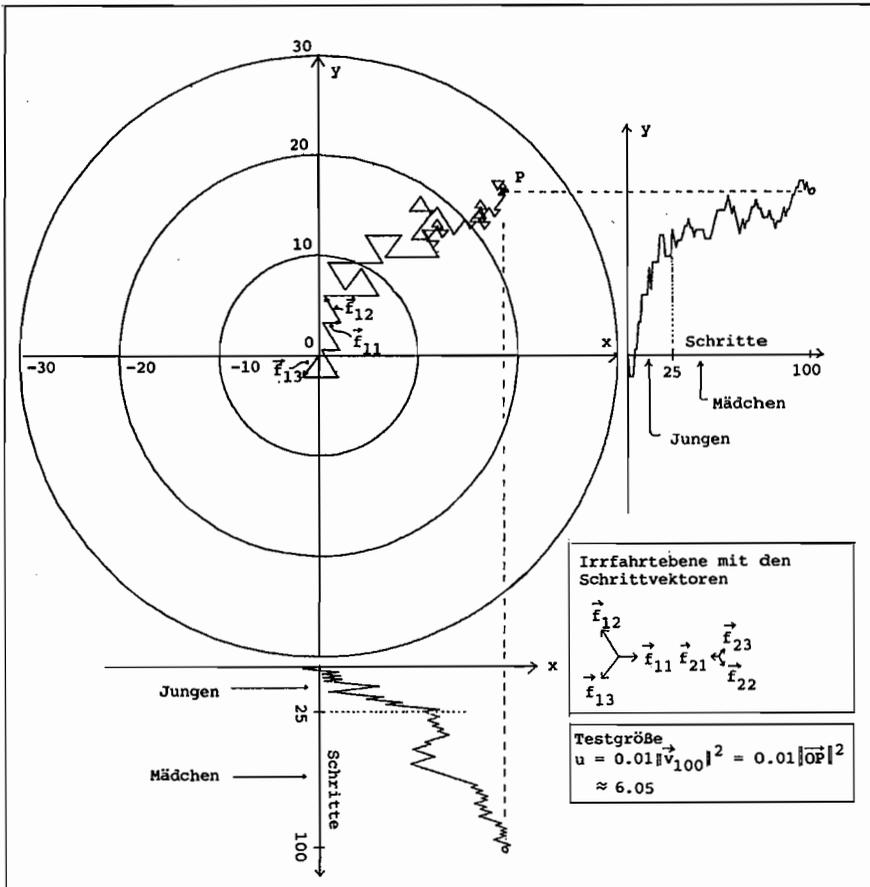


Abb. 4-10 Ebene Irrfahrt zum Unabhängigkeitstest mit Projektionen auf die Koordinatenachsen. Man macht 25 Schritte mit einem der Schrittvektoren  $f_{1j}$  (Jungen) und 75 Schritte mit einem der Schrittvektoren  $f_{2j}$  (Mädchen). Die Antworten der Jungen sind: 3112121211123211121131112.  
 Die Mädchen antworteten mit: 2231133132123111121113333  
 1111332232223323133132121 1321231133231331313112321.

Schrittvektoren). Natürlich kann man bei der Befragung die Geschlechter auch mischen. Durch Auszählen des Irrfahrtprotokolls zu Abb. 4-10 erhalten wir die Sechsfeldertafel

	1 $\hat{=}$ +	2 $\hat{=}$ o	3 $\hat{=}$ -	
Jungen	15	7	3	$q_1=0.25$
Mädchen	29	18	28	$q_2=0.75$
	$b_1=44$	$b_2=25$	$b_3=31$	(experimentell)
	$p_1=0.48$	$p_2=0.23$	$p_3=0.29$	(theoretisch)

$$\text{mit dem Testwert } u = \frac{1}{100} \sum_{\substack{i=1..2 \\ j=1..3}} \frac{(n_{ij} - q_i b_j)^2}{q_i p_j} = 6.05 .$$

Ein Vergleich mit Abb. 4-10 bestätigt dieses Ergebnis. Hier liefert der Endpunkt P der Irrfahrt mit  $|\overline{OP}| = 24.6$  ebenfalls den Testwert  $u=0.01|\overline{OP}|^2 = 6.05$ .

Anmerkung: Wenn man Abb. 4-10 mit Abb. 3-9 vergleicht, dann erkennt man, daß sich beim Unabhängigkeitstest zwei "entgegengesetzt gerichtete", ihren Anteilen an der Gesamtstichprobe entsprechend gewichtete ebene Irrfahrten überlagern. Die Irrfahrten haben die gleiche Struktur wie die Irrfahrt zum Anpassungstest. Das zeigt auch ein Vergleich der Formeln (4) bis (6) mit den Formeln (1) bis (3) aus Kapitel 3.

#### 4.2.3 Zweidimensionale Normalverteilung

Die Endpositionen  $P(x|y)$  unserer zweidimensionalen Irrfahrt sind normalverteilt. Um das (analog zu 3.3.3) einzusehen, wählen wir die x-Achse eines Koordinatensystems für E in Richtung  $\vec{f}_1$ , die y-Achse orthogonal dazu in Richtung  $\vec{g} = \vec{f}_2 - \vec{f}_3$  und erhalten mit (4) und (5) die Koordinatendarstellungen

Jungen	Mädchen
$\vec{f}_{11} = \sqrt{\frac{q_2}{q_1}} \begin{pmatrix} \sqrt{\frac{1-p_1}{p_1}} \\ 0 \end{pmatrix}$	$\vec{f}_{21} = \sqrt{\frac{q_1}{q_2}} \begin{pmatrix} \sqrt{\frac{1-p_1}{p_1}} \\ 0 \end{pmatrix}$
$\vec{f}_{12} = \sqrt{\frac{q_2}{q_1}} \begin{pmatrix} -\sqrt{\frac{p_1}{1-p_1}} \\ \sqrt{\frac{p_3}{p_2(1-p_1)}} \end{pmatrix}$	$\vec{f}_{22} = -\sqrt{\frac{q_1}{q_2}} \begin{pmatrix} -\sqrt{\frac{p_1}{1-p_1}} \\ \sqrt{\frac{p_3}{p_2(1-p_1)}} \end{pmatrix}$
$\vec{f}_{13} = \sqrt{\frac{q_2}{q_1}} \begin{pmatrix} -\sqrt{\frac{p_1}{1-p_1}} \\ -\sqrt{\frac{p_2}{p_3(1-p_1)}} \end{pmatrix}$	$\vec{f}_{23} = -\sqrt{\frac{q_1}{q_2}} \begin{pmatrix} -\sqrt{\frac{p_1}{1-p_1}} \\ -\sqrt{\frac{p_2}{p_3(1-p_1)}} \end{pmatrix}$

Tab. 4-11 Ebene Schrittvektoren in Koordinatendarstellung

Bei Projektion unserer ebenen Irrfahrt auf die x-Achse machen wir  $a_1=nq_1$  voneinander unabhängige Schritte gemäß der Wahrscheinlichkeitsverteilung (Jungen)

$$\frac{p_1}{\sqrt{\frac{q_2}{q_1}} \sqrt{\frac{1-p_1}{p_1}}} \quad \left| \quad \frac{p_2}{-\sqrt{\frac{q_2}{q_1}} \sqrt{\frac{p_1}{1-p_1}}} \quad \left| \quad \frac{p_3}{-\sqrt{\frac{q_2}{q_1}} \sqrt{\frac{p_1}{1-p_1}}} \quad . \quad (8)$$

Erwartungswert und Varianz berechnen sich zu

$$\mu_1 = \sqrt{\frac{q_2}{q_1}} \left( \sqrt{p_1(1-p_1)} - (p_2+p_3) \sqrt{\frac{p_1}{1-p_1}} \right) = 0$$

$$\sigma_1 = \frac{q_2}{q_1} \left( (1-p_1) + (p_2+p_3) \frac{p_1}{1-p_1} \right) = \frac{q_2}{q_1} .$$

Die restlichen  $a_2 = nq_2$  Schritte auf der x-Achse (die zu den Mädchen gehören) werden gesteuert durch die Verteilung

$$-\sqrt{\frac{p_1}{q_2}} \sqrt{\frac{1-p_1}{p_1}} \quad \left| \quad \sqrt{\frac{q_1}{q_2}} \sqrt{\frac{p_1}{1-p_1}} \quad \left| \quad \sqrt{\frac{q_1}{q_2}} \sqrt{\frac{p_1}{1-p_1}} \quad (9)$$

mit Erwartungswert  $\mu_2 = 0$  und Varianz  $V_2 = q_1/q_2$ . Die x-Koordinate am Ende der Irrfahrt ist eine unabhängige Summe von  $a_1$  Zufallsgrößen (8) und  $a_2$  Zufallsgrößen (9). Erwartungswert und Varianz ergeben sich daher zu

$$\mu(x) = a_1 \mu_1 + a_2 \mu_2 = 0,$$

$$V(x) = a_1 V_1 + a_2 V_2 = nq_1(q_2/q_1) + nq_2(q_1/q_2) = n.$$

Die gleichen Resultate erhält man für die y-Koordinate am Ende der Irrfahrt.

Ergebnis:

Wenn wir die zweidimensionale Irrfahrt, die während eines Unabhängigkeitstests abläuft, auf die Koordinatenachsen der Ebene E projizieren, entstehen eindimensionale Irrfahrten mit Erwartungswert 0 und Varianz n. Nach dem Zentralen Grenzwertsatz sind die Koordinaten des normierten Zustandsvektors am Ende der Irrfahrt

$$\frac{1}{\sqrt{n}} \vec{v}_n = \begin{pmatrix} x/\sqrt{n} \\ y/\sqrt{n} \end{pmatrix}$$

als Summen unabhängiger Zufallsgrößen annähernd standardnormalverteilt. Wie in 3.3 macht man hiermit plausibel, daß  $(1/\sqrt{n}) \vec{v}_n$  zweidimensional normalverteilt und folglich die Testgröße  $u = \|(1/\sqrt{n}) \vec{v}_n\|^2$  Chi-Quadrat-verteilt ist mit zwei Freiheitsgraden.

## 4.3 Verallgemeinerung

Aus Gründen der Anschaulichkeit haben wir den Unabhängigkeitstest bisher nur für  $2 \times 3$  - Feldertafeln untersucht. Wir skizzieren abschließend, wie sich die Überlegungen auf beliebige  $k \times l$  - Feldertafeln verallgemeinern lassen.

$n_{11}$	$n_{12}$	$\dots$	$n_{1l}$	$a_1$	$q_1$	$t := \sum_{\substack{i=1..k \\ j=1..l}} \frac{(n_{ij} - q_i b_j)^2}{q_i b_j} \quad (10)$
$n_{21}$	$n_{22}$	$\dots$	$n_{2l}$	$a_2$	$q_2$	
:						
:						
$n_{k1}$	$n_{k2}$	$\dots$	$n_{kl}$	$a_k$	$q_k$	$u := \sum_{\substack{i=1..k \\ j=1..l}} \frac{(n_{ij} - q_i b_j)^2}{n q_i p_j} \quad (11)$
$b_1$	$b_2$	$\dots$	$b_l$	$n$		
$p_1$	$p_2$	$\dots$	$p_l$		$1$	

Tab. 4-12  $k \times l$ -Feldertafel mit Randsummen  $a_i$ ,  $b_j$  und Testgrößen  $t$ ,  $u$  ( $q_i := a_i/n$ )

Die Testgröße  $t$  ist jetzt für großes  $n$  Chi-Quadrat-verteilt mit  $m=(k-1)(l-1)$  Freiheitsgraden, besitzt also die Dichte

$$f(t) = c \sqrt{t}^{m-2} e^{-0.5t}$$

mit einer geeigneten Konstante  $c$ . Um das einzusehen, denken wir uns wieder eine Randverteilung  $a_1, \dots, a_k$  mit den Anteilen  $q_1, \dots, q_k$  als fest vorgegeben. Die Besetzungszahlen  $b_1, \dots, b_l$  sind zufallsabhängig. Bei Gültigkeit der Unabhängigkeitshypothese werden die Besetzungszahlen in jeder Zeile (also auch die Spaltensummen  $b_1, \dots, b_l$ ) durch die gleichen (unbekannten) Wahrscheinlichkeiten  $p_1, \dots, p_l$  gesteuert. Wir approximieren die Chi-Quadrat-Testgröße  $t$  wieder durch die Hilfstestgröße  $u$ , indem wir im Nenner von (10)  $b_j$  durch  $n p_j$  ersetzen. Die Grenzverteilung bleibt davon unberührt, d. h. wenn wir zeigen, daß  $u$  Chi-Quadrat-verteilt ist mit  $m$  Freiheitsgraden, dann gilt das gleiche auch für  $t$ .

Wir deuten  $u$  wieder als das Normquadrat eines Zustandsvektors:  $u = (1/n) \|\vec{v}_n\|^2 = \|(1/\sqrt{n})\vec{v}_n\|^2$  mit

$$\vec{v}_n = \begin{pmatrix} \frac{n_{11}-q_1b_1}{\sqrt{q_1p_1}} & \frac{n_{12}-q_1b_2}{\sqrt{q_1p_2}} & \dots & \frac{n_{11}-q_1b_1}{\sqrt{q_1p_1}} \\ \frac{n_{21}-q_2b_1}{\sqrt{q_2p_1}} & \frac{n_{22}-q_2b_2}{\sqrt{q_2p_2}} & \dots & \frac{n_{21}-q_2b_1}{\sqrt{q_2p_1}} \\ \vdots & \vdots & & \vdots \\ \frac{n_{k1}-q_kb_1}{\sqrt{q_kp_1}} & \frac{n_{k2}-q_kb_2}{\sqrt{q_kp_2}} & \dots & \frac{n_{k1}-q_kb_1}{\sqrt{q_kp_1}} \end{pmatrix}$$

Wenn sich im Verlauf eines Unabhängigkeitstests in der  $i$ -ten Teilstichprobe (vom Umfang  $a_i$ ) das Ergebnis  $j$  einstellt, so verändert sich der Zustandsvektor in der Spalte  $j$  und zwar um den Schrittvektor (vgl. auch Tab. 4-8 als Spezialfall)

$$\vec{e}_{ij} = \begin{pmatrix} 0 \dots 0 & \frac{-q_1}{\sqrt{q_1p_j}} & 0 \dots 0 \\ \vdots & \vdots & \\ 0 \dots 0 & \frac{1-q_i}{\sqrt{q_i p_j}} & 0 \dots 0 \\ \vdots & \vdots & \\ 0 \dots 0 & \frac{-q_k}{\sqrt{q_k p_j}} & 0 \dots 0 \end{pmatrix} \leftarrow i$$

$\uparrow$   
 $j$

Während eines Unabhängigkeitstests läuft also wieder eine Irrfahrt (im Raum  $\mathbb{R}^{k1}$ ) ab. Der Zustandsvektor  $\vec{v}_n$  ergibt sich am Ende als zufallsabhängige Summe von  $n$  Schrittvektoren. Da sich die Zähler von  $\vec{v}_n$  (am Ende der Irrfahrt) in allen  $k$  Zeilen und 1 Spalten zu Null aufaddieren, steht der Zustandsvektor  $\vec{v}_n$  am Ende der Irrfahrt orthogonal zu den Vektoren

$$\vec{o}_i = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \\ \sqrt{p_1} & \sqrt{p_2} & \dots & \sqrt{p_1} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \leftarrow i \quad \vec{w}_j = \begin{pmatrix} 0 \dots 0 \sqrt{q_1} & 0 \dots 0 \\ 0 \dots 0 \sqrt{q_2} & 0 \dots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 \dots 0 \sqrt{q_k} & 0 \dots 0 \end{pmatrix} \uparrow j$$

( $i=1..k$ ,  $j=1..1$ ). Von diesen  $k+1$  Vektoren sind  $k+1-1$  linear unabhängig. Folglich liegt der Zustandsvektor  $\vec{v}_n$  am Ende der Irrfahrt in dem zugehörigen orthogonalen Komplement, einem  $k1-(k+1-1) = (k-1)(1-1)$ -dimensionalen Unterraum  $E$  von  $\mathbb{R}^{k1}$  - egal wie die Irrfahrt im einzelnen abläuft. Daher erklärt sich die Sprechweise vom Freiheitsgrad  $(k-1)(1-1)$ .

Aus diesem Grund brauchen wir im folgenden nicht die Irrfahrt in  $\mathbb{R}^{k1}$  zu untersuchen, es reicht, wenn wir ihre Projektion auf den Unterraum  $E$  bzw. auf beliebige Geraden  $g$  ("Koordinatenachse") von  $E$  verfolgen.  $g$  werde durch den Einheitsvektor  $\vec{g}$  aufgespannt, den wir ohne Einschränkung schreiben wollen als

$$\vec{g} = \begin{pmatrix} \frac{c_{11}}{\sqrt{q_{1P1}}} & \frac{c_{12}}{\sqrt{q_{1P2}}} & \dots & \frac{c_{11}}{\sqrt{q_{1P1}}} \\ \frac{c_{21}}{\sqrt{q_{2P1}}} & \frac{c_{22}}{\sqrt{q_{2P2}}} & \dots & \frac{c_{21}}{\sqrt{q_{2P1}}} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \frac{c_{k1}}{\sqrt{q_{kP1}}} & \frac{c_{k2}}{\sqrt{q_{kP2}}} & \dots & \frac{c_{k1}}{\sqrt{q_{kP1}}} \end{pmatrix}$$

$$\text{Mit } \|\vec{g}\| = 1 \text{ gilt } \sum_{i,j} \frac{c_{ij}^2}{q_{iPj}} = 1 \quad (12)$$

$$\text{Wegen } \vec{g} \perp \vec{o}_i \text{ gilt } \sum_j c_{ij} = 0 \quad (i=1, \dots, k) \quad (13)$$

$$\text{Aus } \vec{g}_1 \vec{w}_j \text{ folgt } \sum_i c_{ij} = 0 \quad (j=1, \dots, l) \quad (14)$$

Eine kurze Rechnung liefert mit (14) die Beziehung

$$\langle \vec{e}_{ij} | \vec{g} \rangle = \frac{c_{ij}}{q_i p_j}$$

Der Teil der Irrfahrt, der zu der  $i$ -ten Teilstichprobe vom Umfang  $a_i$  gehört, besitzt nach Projektion auf  $g$  den Erwartungswert

$$\mu_i = a_i \sum_j p_j \langle \vec{e}_{ij} | \vec{g} \rangle = a_i \sum_j \frac{c_{ij}}{q_i} = 0$$

und die Varianz

$$\begin{aligned} v_i &= a_i \sum_j p_j \langle \vec{e}_{ij} | \vec{g} \rangle^2 = a_i \sum_j p_j \frac{c_{ij}^2}{q_i^2 p_j^2} = \\ &= \frac{a_i}{q_i} \sum_j \frac{c_{ij}^2}{q_i p_j} = n \sum_j \frac{c_{ij}^2}{q_i p_j} \end{aligned}$$

( $i=1, \dots, k$ ). Die gesamte Irrfahrt hat nach Projektion auf die Gerade  $g$  folglich ebenfalls den Erwartungswert 0

$$\text{und die Varianz } v = \sum_i v_i = n \sum_{i,j} \frac{c_{ij}^2}{q_i p_j} = n .$$

Normiert man, indem man den Zustandsvektor  $\vec{v}_n$  am Ende der Irrfahrt durch  $\sqrt{n}$  teilt, dann ist die Projektion auf  $g$   $\langle (1/\sqrt{n})\vec{v}_n | \vec{g} \rangle$  nach dem Zentralen Grenzwertsatz standardnormalverteilt. Führt man in  $E$  wieder ein geeignetes Orthogonalsystem ein, so ergibt ein Plausibilitätsargument (wie in 3.3.3) die Unabhängigkeit der  $m=(k-1)(l-1)$  Koordinaten, die normierte Endposition der Irrfahrt  $(1/\sqrt{n})\vec{v}_n$  ist dann wegen der Produktregel  $m$ -dimensional normalverteilt. Hieraus folgt wie in 3.4, daß die Testgröße  $u = \|(1/\sqrt{n})\vec{v}_n\|^2$  Chi-Quadratverteilt ist mit  $m$  Freiheitsgraden.

## 5 Wahrscheinlichkeiten gezinkter Quader

Die (gezinkten und ungezinkten) Quader haben mit ihren partiellen Symmetrien bei einer konsistenten statistischen Begriffsentwicklung in den Kapiteln 1 und 2 ausgezeichnete Dienste geleistet. Wir wollen sie nun mit Hilfe der Chi-Quadrat-Verfahren etwas eingehender untersuchen. Dabei kommen (wie auch bei der Partnersuche in Kapitel 10) Anpassungstests mit Parametern zum Einsatz. Wir werden die Tatsache, daß sich die Anzahl der Freiheitsgrade um die Anzahl der geschätzten Parameter reduziert, nicht weiter thematisieren. Eine gängige Plausibilitätsbetrachtung sagt uns aber: Jeder Parameter, welcher aus der Stichprobe geschätzt wurde, stellt eine einschränkende Bedingung für die möglichen Besetzungszahlen dar und zieht eine Reduktion des Freiheitsgrades nach sich - genau wie die durch den Versuchsumfang vorgegebene Summe der Besetzungszahlen (eine Bedingung) beim Anpassungstest zum Freiheitsgrad  $k-1$  führte und die gegebenen Spalten- und Zeilensummen ( $k+1-1$  Bedingungen) beim Unabhängigkeitstest den Freiheitsgrad  $k_1-(k+1-1)$  lieferten (vgl. 3.4.1 und 4.3).

## 5.1 Wurftechnik

Hat die Unterlage einen Einfluß auf das stochastische Verhalten unserer Quader (Abb. 1-1)? Wir würfeln mit einem 0.31 Becher einmal auf einer glatten Holzunterlage, ein anderes mal auf einem Teppich. Beide Male wird der Becher auf die Unterlage gestülpt, so daß die Quader nicht rollen. Wegen der Symmetrie fassen wir die Gegenseiten des Quaders zusammen. Es ergab sich die Sechsfeldertafel:

Seiten	1;6	2;5	3;4	
Tisch	217	109	674	1000
Teppich	114	72	314	500
	331	181	988	1500

t=4.6

Tab. 5-1 Ein Quader wird mit einem Becher auf verschiedenen Unterlagen gewürfelt

Der Testwert  $t=4.6$  liegt deutlich unterhalb der 95%-Grenze (6.0) der Chi-Quadrat-Verteilung mit 2 Freiheitsgraden. Wir brauchen die Hypothese der Unabhängigkeit des Würfelergbnisses von der Unterlage auf dem 5%-Signifikanzniveau nicht zu verwerfen.

Anders sieht es aus, wenn wir die Wurftechnik ändern. Wir demonstrieren das für einen gezinkten Quader, das lange U (Abb. 2-1), das wir einerseits wieder mit einem Becher auf einer Holzunterlage würfeln, andererseits aus 50 cm Höhe auf einen Teppich frei fallenlassen. Die instabilen Seiten, die sonst durch den Becher gestützt werden, treten jetzt sehr viel seltener auf. Tatsächlich zwingt uns ein Zwölfel-dertest

Seiten	1	2	3	4	5	6	
Becher	42	33	111	197	38	60	481
freier Fall	7	3	42	86	2	10	150
	49	36	153	283	40	70	631

$t=26.4$

Tab. 5-2 Langes U: Becher auf Holz / freier Fall auf Teppich

die Hypothese der Unabhängigkeit von Wurfsergebnissen und Wurftechnik fallenzulassen. Die Testgröße  $t=26.4$  liegt weit oberhalb der 99%-Grenze (15.1) der Chi-Quadrat-Verteilung mit 5 Freiheitsgraden. Auch wenn wir zur Erhöhung der Belegungszahlen die Gegenseiten 1-6 und 2-5 zusammenfassen (3 und 4 sind nicht symmetrisch), dann ergibt sich ein Testwert (26.3), der weit oberhalb der 99%-Grenze (11.3) der Chi-Quadrat-Verteilung (jetzt mit 3 Freiheitsgraden) liegt.

## 5.2 Boltzmann-Verteilung

In den vorangehenden Kapiteln hatten wir "aus dem Stegreif" einige plausible Hypothesen über das stochastische Verhalten unserer Quader aufgestellt, wie etwa eine Proportionalität zwischen Wahrscheinlichkeit und Grundfläche bzw. Kippwinkel.

Der Anpassungstest zwingt uns, alle diese Hypothesen zu verwerfen. In der Tat kann man wegen der eben festgestellten Abhängigkeit der Wahrscheinlichkeiten von der Wurftechnik nicht erwarten, Quader durch eine universell gültige Wahrscheinlichkeitshypothese beschreiben zu können. Wir erinnern jedoch an den Boltzmannschen Verteilungssatz, der das Verhalten statistischer Systeme (etwa idealer Gase in Kraftfeldern) beschreibt. Der Satz besagt, daß die Wahrscheinlichkeit eines Zustandes  $z$  in Abhängigkeit von der Energie  $E(z)$  des Zustandes gegeben ist durch

$$P(z) = c e^{-E(z)/(kT)}$$

Dabei ist  $k$  die Boltzmannkonstante,  $T$  die absolute Temperatur und  $c$  dient zur Normierung. Je höher also die Energie  $E$ , desto unwahrscheinlicher ist der Zustand  $z$ . Beschreibt der Boltzmannsche Verteilungssatz unsere (gezinkten) Quader? Da die Energie, die zur Lage eines Quaders gehört, direkt proportional zur Höhe  $h$  des Schwerpunktes ist, impliziert der Boltzmannsche Verteilungssatz: Je höher der Schwerpunkt, desto unwahrscheinlicher ist die Lage. Während wir aber früher bei unseren (ungezinkten) Quadern mit der Proportionalität zwischen Wahrscheinlichkeit  $P(i)$  und Fläche  $F(i)$  gleichzeitig einen umgekehrt proportionalen Zusammenhang mit der Schwerpunkthöhe  $h(i)$  unterstellten (denn die Schwerpunkthöhe ist beim Quader wegen der Formel " $h(i) \cdot F(i) = \text{halbes Quadervolumen}$ " zur Grundfläche umgekehrt proportional):

$$P(i) = c \cdot 1/h(i) ,$$

liefert der Boltzmannsche Verteilungssatz einen exponentiellen Zusammenhang:

$P(i) = c \cdot e^{-bh(i)}$
-----------------------------

Anpassungstest mit Parameter:

Unser Problem besteht nun darin, daß wir keinen Wert für die Konstante  $b$  kennen. (Bei Gasen hängt  $b=1/kT$  von der Temperatur  $T$  ab.) Betrachten wir als Beispiel den langen U-Würfel mit den Schwerpunkthöhen:

$i$	1	2	3	4	5	6
$h(i)$	1.0	1.2	0.761	0.539	1.2	1.0

und schauen wir, welche Hypothesen sich für die Werte  $b=2$ ,  $b=3$  und  $b=4$  ergeben. (Der Normierungsfaktor  $c$  ergibt sich aus der Bedingung  $P(1)+P(2)+\dots+P(6)=1$ .)

$b$	$c$	1	2	3	4	5	6	$t$	$m$
2	1	0.134	0.090	0.216	0.337	0.090	0.134	19.2	-760.81
3	2.2	0.109	0.060	0.224	0.437	0.060	0.109	7.7	-754.47
4	4.6	0.085	0.038	0.220	0.535	0.038	0.085	56.7	-775.46

Tab. 5-3 Verschiedene Boltzmann-Verteilungen

Wir vergleichen diese Hypothesen mit den experimentellen Ergebnissen aus der ersten Zeile von Tab. 5-2, die wir mit dem Würfelbecher auf einem Tisch erhielten:

$i$	1	2	3	4	5	6	
$a(i)$	42	33	111	197	38	60	(481)
$r(i)$	0.087	0.685	23.02	0.408	0.078	0.124	(100%) ,

indem wir die zugehörigen Anpassungstestgrößen berechnen. Die Testwerte  $t$  sind in Tabelle 5-3 aufgeführt. Für  $b=3$  ist die Übereinstimmung besser als für  $b=2$  oder  $b=4$ . Der Anpassungstest (mit Parameter) besteht nun darin, den Wert für  $b$  zu finden, bei dem die Übereinstimmung zwischen Hypothese und Experiment optimal ist. Man schätzt also  $b$  aus der Stichprobe. Dann untersucht man, ob der Testwert  $t$  für diesen Parameter oberhalb der kritischen Grenze der Chi-Quadrat-Verteilung (jetzt mit 4 Freiheitsgraden, weil wir einen Parameter aus der Stichprobe bestimmt haben) liegt.

Es gibt zwei Methoden zur Schätzung des Parameters:

(a) die Chi-Quadrat-Minimum-Methode, bei der man  $b$  so wählt, daß der Testwert  $t$  minimal wird;  $b$  wird zwischen 2 und 4 liegen.

(b) die maximum likelihood-Methode: Für jeden Wert von  $b$  bestimmt man die Wahrscheinlichkeiten  $P(1), \dots, P(6)$  und berechnet (mit der Pfadregel) die Wahrscheinlichkeit  $Q = P(1)^{a(1)} P(2)^{a(2)} \dots P(6)^{a(6)}$ , mit der man das gerade beobachtete Versuchsergebnis  $a(1), \dots, a(6)$  erwartet hätte. Nun wählt man für  $b$  den Wert, bei dem  $Q$  am größten wird (daher der Name maximum likelihood).  $Q$  ist bei hohem Stichprobenumfang sehr klein. Daher maximieren wir in Tab. 5-3 den Logarithmus:  $m = \log Q = a(1) \log P(1) + \dots + a(6) \log P(6)$ . Wie man sieht, liefert auch diese Methode einen Parameter  $b$  zwischen 2 und 4.

Die Tabelle 5-4 stellt einen "verfeinerten Ausschnitt" von Tabelle 5-3 dar. Wir entnehmen ihr mit ausreichender Genauigkeit: bei  $b=2.7$  wird  $t$  minimal (und  $m$  maximal). Der Testwert beträgt  $t=4.8$ , er liegt deutlich unter der 95%-Grenze (9.5) der Chi-Quadrat-Verteilung mit 4 Freiheitsgraden.

$i$	1	2	3	4	5	6		
$h(i)$	1.000	1.200	0.761	0.539	1.200	1.000		
$a(i)$	42	33	111	197	38	60	(481)	
$r(i)$	0.087	0.069	0.231	0.410	0.079	0.125	$t$	$m$
$b=2.30$	0.127	0.080	0.220	0.366	0.080	0.127	9.5	-755.86
$b=2.40$	0.124	0.077	0.221	0.376	0.077	0.124	7.4	-754.80
$b=2.50$	0.122	0.074	0.222	0.386	0.074	0.122	6.0	-754.03
$b=2.60$	0.120	0.071	0.223	0.396	0.071	0.120	5.1	-753.55
$b=2.70$	0.117	0.068	0.223	0.406	0.068	0.117	4.8	-753.36
$b=2.80$	0.115	0.065	0.224	0.416	0.065	0.115	5.1	-753.45
$b=2.90$	0.112	0.063	0.224	0.426	0.063	0.112	6.0	-753.82
$b=3.00$	0.109	0.060	0.224	0.437	0.060	0.109	7.5	-754.47

Tab. 5-4 Boltzmann-Verteilungen für verschiedene Parameterwerte  $b$  und zugehörige Testgrößen (Langes U auf Holz mit Würfelbecher)

Das bedeutet: Die Wahrscheinlichkeiten des langen U-Würfels sind (bei Verwendung eines Würfelbechers und einer Holzunterlage) mit der Hypothese einer Boltzmann-Verteilung vereinbar. Die zugehörige Verteilung lautet:

i	1	2	3	4	5	6
P(i)	0.117	0.068	0.223	0.406	0.068	0.117

oder analytisch:  $P(i) = 1.74 e^{-2.7 h(i)}$ .

Auch bei freiem Fall auf einen Wollteppich sind die experimentellen Ergebnisse mit einer Boltzmann-Hypothese vereinbar:

i	1	2	3	4	5	6		
h(i)	1.000	1.200	0.761	0.539	1.200	1.000		
a(i)	7	3	42	86	2	10	(150)	
r(i)	0.047	0.020	0.280	0.573	0.013	0.067	t	m
b=4.60	0.071	0.028	0.212	0.590	0.028	0.071	6.1	-173.36
b=4.70	0.069	0.027	0.211	0.598	0.027	0.069	5.9	-173.19
b=4.80	0.066	0.025	0.209	0.607	0.025	0.066	5.8	-173.09
b=4.90	0.064	0.024	0.207	0.616	0.024	0.064	5.8	-173.05
b=5.00	0.062	0.023	0.206	0.624	0.023	0.062	5.9	-173.06
b=5.10	0.060	0.022	0.204	0.632	0.022	0.060	6.2	-173.14
b=5.20	0.058	0.021	0.202	0.640	0.021	0.058	6.5	-173.27

Tab. 5-5 Boltzmann-Verteilungen für verschiedene Parameterwerte b und zugehörige Testgrößen (Langes U bei freiem Fall auf einen Teppich)

Die Parameterschätzung liefert jetzt einen wesentlich höheren Wert  $b=4.9$ , aber auch hier liegt die Testgröße  $t=5.8$  deutlich unter der 95%-Grenze (9.5). Die zugehörige Boltzmann-Verteilung ist

i	1	2	3	4	5	6
P(i)	0.064	0.024	0.207	0.616	0.024	0.064

oder analytisch  $P(i) = 8.63 e^{-4.9 h(i)}$ .

Wenn man die Konstante  $b=1/(kT)$  "thermodynamisch" deutet, kann man sagen: durch freies Fallen auf den Teppich ist die

Temperatur  $T$  des Systems gesunken (je größer  $b$  desto kleiner  $T$ ). In der Tat sinkt bei kleiner Temperatur die Aufenthaltswahrscheinlichkeit für hochenergetische Zustände, die stabilen Lagen mit niedrigen Lageenergien werden bevorzugt. Diese Ergebnisse sind kein Produkt spezieller Stichproben. Sie gelten auch für unsere Quader, den kurzen U-Würfel und den S-Würfel, nicht jedoch für unsere L-Würfel und Rundhölzer. Näheres hierzu in Riemer (1988).

#### Zusammenfassung:

- (a) Die Wahrscheinlichkeiten gezinkter und ungezinkter Quader hängen in beträchtlichem Maße von den Versuchsbedingungen ab.
- (b) Die Wahrscheinlichkeiten werden (unabhängig von den Versuchsbedingungen) durch Boltzmann-Verteilungen beschrieben:  $P(i) = c e^{-b h(i)}$ .
- (c) Je "freier" die Quader geworfen werden und je unebener die Unterlage ist, desto unwahrscheinlicher werden instabile Lagen (große Schwerpunkthöhen) angenommen. Desto größer ist der Parameter  $b$ .

#### 5.3 Potenzverteilung

Wir greifen abschließend unser Quaderbeispiel aus Tab. 3-3 auf. Wir untersuchten dort, ob die Wahrscheinlichkeiten der Seiten zu einer ganzzahligen Potenz ( $b=0,1,2,\dots,5$ ) der Grundfläche  $F$  direkt (also zu einer Potenz der Schwerpunkthöhe  $h$  umgekehrt) proportional sind:

$$P(i) \sim F(i)^b \sim 1/(h(i)^b) \quad (\text{"Potenzverteilung"})$$

Wir variieren nun  $b$  wie bei der Boltzmann-Verteilung kontinuierlich und erkennen aus der folgenden Tabelle 5-6, daß der Testwert (1800 Versuche) nicht unter das Minimum 10.6 absinkt (Parameter  $b=2.8$ ). Er liegt über der 95%-Grenze (9.5) der Chi-Quadrat-Verteilung mit 4 Freiheitsgraden.

i	1	2	3	4	5	6	(1800)	
h(i)	1.000	1.150	0.650	0.650	1.150	1.000		
a(i)	210	100	605	588	110	187		
r(i)	0.117	0.056	0.336	0.327	0.061	0.104	t	m
b= 2.50	0.108	0.076	0.316	0.316	0.076	0.108	19.5	-2799.04
b= 2.60	0.105	0.073	0.322	0.322	0.073	0.105	14.6	-2796.32
b= 2.70	0.102	0.070	0.327	0.327	0.070	0.102	11.6	-2794.62
b= 2.80	0.100	0.067	0.333	0.333	0.067	0.100	10.6	-2793.91
b= 2.90	0.097	0.065	0.338	0.338	0.065	0.097	11.4	-2794.19
b= 3.00	0.094	0.062	0.344	0.344	0.062	0.094	14.2	-2795.42

Tab. 5-6 "Potenzverteilungen" für verschiedene  
Parameterwerte b und zugehörige Testgrößen  
(Quader auf Holz mit Würfelbecher)

Wir müssen die Hypothese einer "Potenzverteilung" für unsere  
Quader auf dem 5%-Signifikanzniveau zurückweisen.

Teil III Eindimensionale Irrfahrten und Grenzwertsätze  
(Aspekte der Analysis in der Stochastik)

Den folgenden Kapiteln liegt eine gemeinsame Struktur zugrunde: Wir lassen ein Teilchen nach gewissen Regeln zufallsgesteuert auf der Zahlengeraden wandern. Wenn wir die Position am Ende der Irrfahrt geeignet normieren, so beobachten wir mit wachsender Schrittzahl Konvergenz gegen eine Grenzverteilung. Wir bestimmen deren analytische Form, indem wir Rekursionsgleichungen durch Taylorentwicklung in Differentialgleichungen überführen und deren Lösungen ermitteln. Durch diesen didaktisch neuen Ansatz werden Themengebiete wie der Zentrale Grenzwertsatz, Chi-Quadrat-Verteilungen (mit Anwendung auf den Anpassungstest) und das Arcsin-Gesetz einer elementaren Darstellung zugänglich.

Vom didaktischen Hintergrund her sind unsere Ausführungen vergleichbar mit Engels Arbeiten (1976), mit denen endliche Markoffketten im Schulunterricht heimisch wurden: Man berechnet - ohne deren Existenz zu hinterfragen - diskrete Grenzverteilungen oder mittlere Wartezeiten bis zur Absorption in bestimmten Zuständen über lineare Gleichungssysteme. Wir haben es im folgenden mit "stetigen" Phänomenen zu tun und müssen statt linearer Gleichungssysteme lineare Differentialgleichungen betrachten.

Trotz der verbindenden Struktur kann man die folgenden Kapitel unabhängig voneinander lesen. Wer die Brücke, die die Differentialgleichungen zur Analysis schlagen, nicht beschreiten möchte, beschränke sich auf das computerunterstützte Studium der Phänomene jeweils in den ersten Abschnitten der folgenden Kapitel.

Für den mathematisch interessierten Leser sind Beziehungen zu Kolmogoroffs Differentialgleichungen, zur Theorie der Diffusionsprozesse und zur Renewal-Theorie in Kapitel 9 skizziert.

## 6 Zentraler Grenzwertsatz

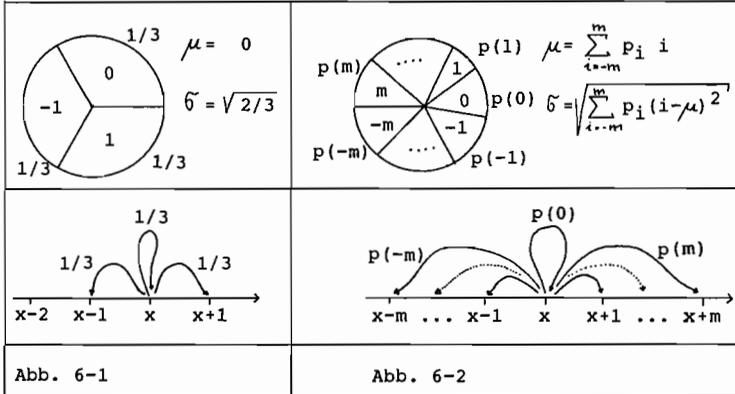
Warum sind so verschiedenartige Dinge unserer Alltagswelt wie Meßfehler, Produktionstoleranzen, Fruchtgewichte, Testergebnisse... "glockenförmig" verteilt? Hinter diesem wohlbekannten Phänomen steckt der Zentrale Grenzwertsatz, der besagt, daß Summen unabhängiger Zufallsgrößen näherungsweise normalverteilt sind. In den obigen Beispielen kann man sich nämlich die Schwankungen der genannten Größen aus einer Reihe unabhängig voneinander wirkender Komponenten zusammengesetzt denken. In der Tat kommt diesem Grenzwertsatz wegen seines Beitrages zur "Umwelterschließung" auch im schulischen Bereich ein beachtliches Eigengewicht zu. In der Statistik gewinnt er eine zentrale Rolle, weil er die Grundlage vieler Test- und Schätzverfahren bereitet (vgl. z. B. 3.3, 4.2, 10.3).

Wie schon erwähnt, gehen wir in diesem Kapitel von der (empirisch gesicherten) Existenz einer Grenzverteilung für eine Folge von Summen geeignet normierter Zufallsgrößen aus und zeigen durch Aufstellen einer linearen Differentialgleichung, daß von den vielen glockenförmigen Funktionen einzig die Dichte  $f(x) = ce^{-0.5x^2}$  der Normalverteilung als Grenzverteilung in Frage kommt.

Das Schöne an dieser heuristischen Überlegung ist, daß sie nicht von den speziellen Verteilungen der Summanden abhängt und dadurch die Universalität der Normalverteilung unterstreicht.

## 6.1 Rekursionsformel für "Summenverteilungen"

Die zeitliche Entwicklung einer Summe von Zufallsgrößen (wie das Entstehen der Punktzahl in einem Test, der Trefferzahl in einer Bernoulli-Kette, einer Augensumme beim wiederholten Würfeln ...) läßt sich als Irrfahrt auf der Zahlengeraden deuten.



Haben wir beispielsweise beim wiederholten Drehen des Glücksrades aus Abb. 6-1 die Punktzahl  $x$  erreicht, so erhalten wir im nächsten Schritt die Summe  $x-1$ ,  $x$  oder  $x+1$  je mit Wahrscheinlichkeit  $1/3$ . Wir springen also auf dem Zahlenstrahl eine Einheit nach links, nach rechts oder wir bleiben stehen wo wir sind. Die in Abb. 6-2 skizzierte Situation, auf die wir uns im folgenden beziehen werden, ist etwas allgemeiner. Wenn wir mit  $P_n(x)$  die Wahrscheinlichkeit bezeichnen, daß man nach  $n$  Schritten die Position (=Punktzahl)  $x$  erreicht hat, dann gilt folgende Rekursionsformel

$$P_{n+1}(x) = \sum_{i=-m}^m P_n(x-i) \cdot p(i) \quad (1)$$

Genau dann steht man nämlich nach  $n+1$  Schritten auf der Stelle  $x$ , wenn man nach  $n$  Schritten auf der Stelle  $x-i$  stand und (mit Wahrscheinlichkeit  $p(i)$ ) einen Sprung der Weite  $i$  ausführte. Man bezeichnet  $P_{n+1}$  auch als Faltung von  $P_n$  mit

$p$  (und schreibt  $P_{n+1}=P_n * p$ ). In Tab. 6-3 wurde diese Rekursionsformel (für  $p(-1)=p(0)=p(1)=1/3$ , Abb. 6-1) zur Berechnung der Wahrscheinlichkeitsverteilungen  $P_1, P_2, \dots, P_{10}$  verwendet. Die Wahrscheinlichkeitsverteilung  $P_n$  des Standortes nach  $n$  Schritten hat bekanntlich den Erwartungswert  $n\mu$  und die Standardabweichung  $\sigma\sqrt{n}$ . Die zugehörigen Histogramme werden mit wachsendem  $n$  immer "breiter" und "flacher". Man normiert daher den Standort  $x$  gemäß der Vorschrift  $y:=(x-n\mu)/(\sigma\sqrt{n})$ . Der normierte Standort  $y$  besitzt dann den Erwartungswert 0 und die Standardabweichung 1. Wir beschränken uns im folgenden auf den Fall  $\mu=0$  und erhalten für die Stufenhöhe  $f_n$  des normierten Histogramms:

$$f_n(y) = f_n\left(\frac{x}{\sigma\sqrt{n}}\right) = \sigma\sqrt{n} P_n(x) \quad \text{oder}$$

$$P_n(x) = \frac{1}{\sigma\sqrt{n}} f_n\left(\frac{x}{\sigma\sqrt{n}}\right) \quad (2)$$

Das Histogramm, das zu dieser normierten Größe  $y$  gehört, entsteht aus dem Histogramm der nicht normierten Größe  $x$ , indem man die Breite der Histogrammstufen mit dem Faktor  $1/(\sigma\sqrt{n})$  staucht und die Stufenhöhe zum Ausgleich mit dem Faktor  $\sigma\sqrt{n}$  streckt. Die Histogrammfläche 1 bleibt davon unberührt. Abb. 6-4 veranschaulicht diesen Vorgang für die Verteilung  $P_{10}$  aus Tab. 6-3).

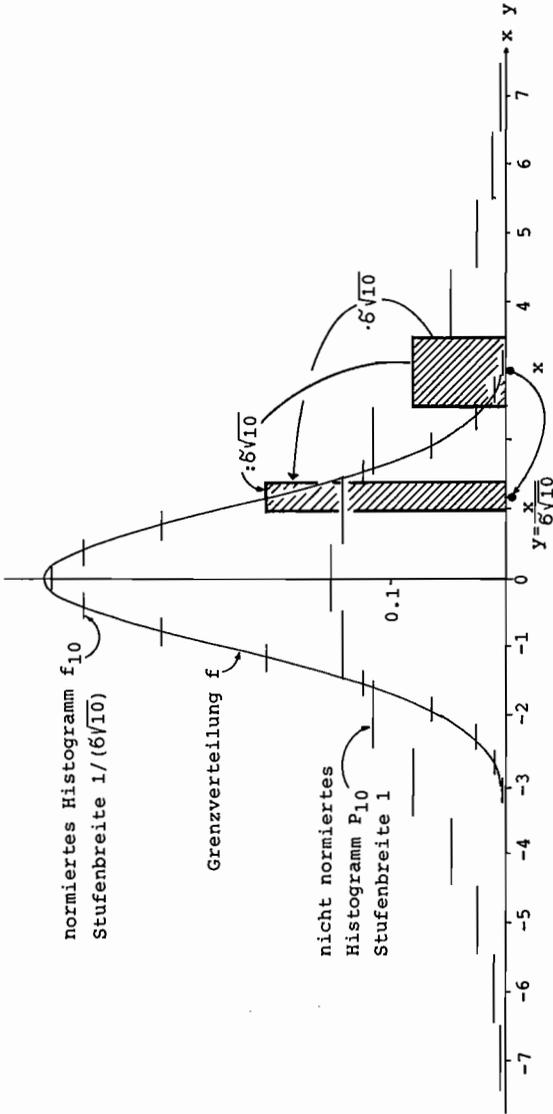
## 6.2 Differentialgleichung der Normalverteilung

Zeichnet man die normierten Histogramme für verschiedene Werte von  $n$ , so erkennt man, daß sie "übereinander" liegen und sich einer glockenförmigen Funktion  $f$  "anschmiegen". Wir wollen den analytischen Ausdruck dieser Funktion bestimmen. Dazu setzen wir (2) in die Rekursionsformel (1) ein und erhalten eine Rekursionsformel für die normierten Histogramme:

$$\frac{1}{\sigma\sqrt{n+1}} f_{n+1}\left(\frac{x}{\sigma\sqrt{n+1}}\right) = \sum_{i=-m}^m p(i) \frac{1}{\sigma\sqrt{n}} f_n\left(\frac{x-i}{\sigma\sqrt{n}}\right), \quad \text{also}$$

	Menner	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10					
P <sub>1</sub>	3																										
*p		↘																									
P <sub>2</sub>	9									1	2	3	2	1													
*p		↘																									
P <sub>3</sub>	27								1	3	6	7	6	3	1												
*p		↘																									
P <sub>4</sub>	81							1	4	10	16	19	16	10	4	1											
*p		↘																									
P <sub>5</sub>	243							1	5	15	30	45	51	45	30	15	5	1									
*p		↘																									
P <sub>6</sub>	729							1	6	21	50	90	126	141	126	90	50	21	6	1							
*p		↘																									
P <sub>7</sub>	2187							1	7	28	77	161	266	357	393	357	266	161	77	28	7	1					
*p		↘																									
P <sub>8</sub>	6561							1	8	36	112	266	504	784	1016	1107	1016	784	504	266	112	36	8	1			
*p		↘																									
P <sub>9</sub>	19683							1	9	45	156	414	882	1554	2304	2907	3139	2907	2304	1554	882	414	156	45	9	1	
*p		↘																									
P <sub>10</sub>	59049	1	10	55	210	615	1452	2850	4740	6765	8350	8953	8350	6765	4740	2850	1452	615	210	55	10	1					
P <sub>10</sub>		.000	.000	.001	.004	.010	.025	.048	.080	.115	.141	.152	.141	.115	.080	.048	.025	.010	.004	.001	.000	.000	.000	.000	.000	.000	.000
Normalverteilung		.000	.000	.001	.004	.010	.024	.047	.079	.114	.143	.155	.143	.114	.079	.047	.024	.010	.004	.001	.000	.000	.000	.000	.000	.000	.000

Tab. 6-3 Rekursive Berechnung der Wahrscheinlichkeitsverteilung P<sub>10</sub> nach (1).  
Die Zähler in einer Zeile sind durch den angegebenen Nenner zu teilen.

Abb. 6-4 Das Histogramm  $P_{10}$  wird zu  $f_{10}$  normiert

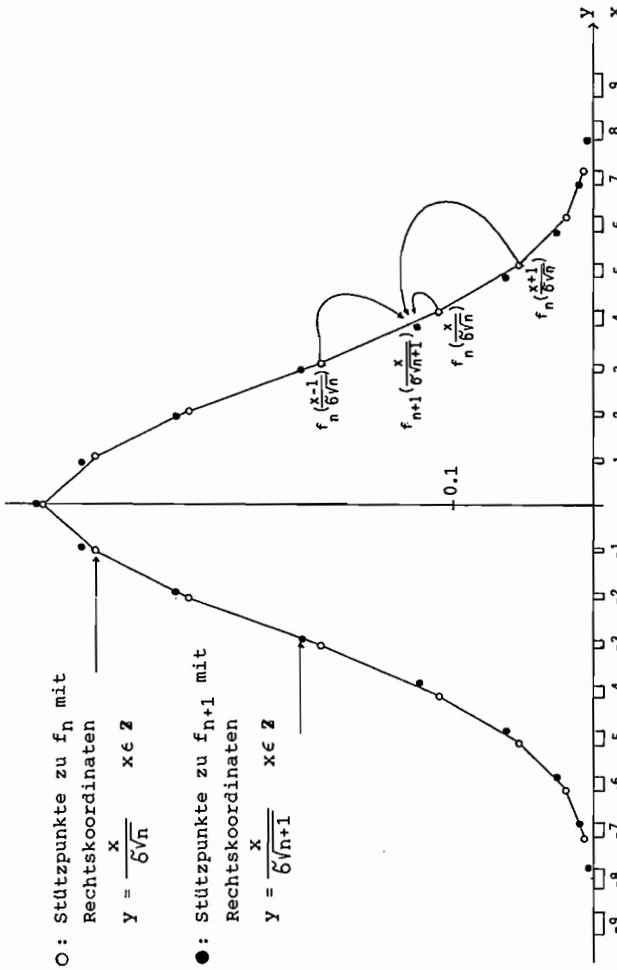


Abb. 6-5 Normierte Histogramme  $f_n$  ( $\circ$ ) und  $f_{n+1}$  ( $\bullet$ )  
 mit Deutung der Rekursionsformel (3) für  $n=9$

$$\underbrace{f_{n+1}\left(\frac{x}{\sqrt[n+1]{6n+1}}\right)}_C = \underbrace{\sqrt{\frac{n+1}{n}}}_{B} \underbrace{\sum_{i=-m}^m p(i) f_n\left(\frac{x-i}{\sqrt[n]{6n}}\right)}_A. \quad (3)$$

Die Rekursionsformel ist in Abb. 6-5 (wieder für  $p(-1)=p(0)=p(1)=1/3$ ) veranschaulicht. Dort haben wir die Stufenmittelpunkte der normierten Histogramme  $f_n$  für  $n=9$  und  $n=10$  eingezeichnet (und für  $n=9$  durch ein "Wahrscheinlichkeitspolygon" verbunden. Die Rekursionsformel (3) besagt anschaulich: wenn man das (durch  $p$  gewichtete) Mittel nebeneinander liegender "Stützwerte" von  $f_n$  (A) mit dem Faktor  $\sqrt{(n+1)/n}$  verlängert (B) und von  $x/(\sqrt[n]{6n})$  auf  $x/(\sqrt[n+1]{6n+1})$  verschiebt, dann erhält man einen Stützwert (C) von  $f_{n+1}$ . Dieser liegt, wie man sieht, wieder sehr nahe bei  $f_n$ . Wenn nun die Histogramme  $f_n$  mit wachsendem  $n$  gegen eine Grenzfunktion  $f$  konvergieren ( $f_n \approx f_{n+1} \approx f$ ), dann muß auch  $f$  eine Form haben, für die das "verlängerte" und "verschobene" arithmetische Mittel aus Funktionswerten wieder einen Funktionswert ergibt. Nun liegen bei hohem  $n$  alle Stellen, die zur Bildung des Mittelwertes benötigt werden, sehr nahe beieinander, und wir können ihn auch durch eine Taylorentwicklung von  $f$  um die Stelle  $y = \frac{x}{\sqrt[n]{6n}}$  berechnen. Wir erhalten auf der rechten Seite von (3)

$$A = \sum_{i=-m}^m p(i) f\left(y - \frac{i}{\sqrt[n]{6n}}\right) \approx$$

$$f(y) \underbrace{\sum_{i=-m}^m p(i)}_1 - \frac{1}{\sqrt[n]{6n}} f'(y) \underbrace{\sum_{i=-m}^m p(i) i}_0 + \frac{1}{2\sqrt[n]{6n}^2} f''(y) \underbrace{\sum_{i=-m}^m p(i) i^2}_{\sqrt[n]{6n}^2} =$$

$$f(y) + \frac{1}{2n} f''(y).$$

Wegen  $B = \sqrt{1+(1/n)} \approx 1+1/(2n)$  ergibt sich hiermit für die rechte Seite von (3)

$$BA \approx \left(1 + \frac{1}{2n}\right) \left(f(y) + \frac{1}{2n} f''(y)\right) \approx f(y) + \frac{1}{2n} f(y) + \frac{1}{2n} f''(y). \quad (4)$$

Auf der linken Seite von (3) ergibt sich wegen

$$\frac{x}{\delta\sqrt{n+1}} = \frac{x}{\delta\sqrt{n}} \sqrt{\frac{n}{n+1}} \approx y \left(1 - \frac{1}{2n}\right) = y - \frac{1}{2n} y$$

$$C = f\left(\frac{x}{\delta\sqrt{n+1}}\right) \approx f(y) - \frac{1}{2n} y f'(y) \quad (5)$$

und als Bedingung an die Grenzfunktion  $f$  erhält man durch Vergleich von (4) und (5) die Differentialgleichung

$$f(y) + yf'(y) + f''(y) = 0. \quad (6)$$

Die linke Seite ist die Ableitung von  $yf(y) + f'(y)$  und damit ergibt sich durch Integration eine Differentialgleichung erster Ordnung

$$yf(y) + f'(y) = \bar{c} \quad (\bar{c} \in \mathbb{R}). \quad (7)$$

Aus der Symmetrie von  $f$  folgern wir  $f'(0)=0$ . Setzen wir das in (7) ein, so erhalten wir für die Konstante  $c=0$  und (7) vereinfacht sich zu  $f'(y)/f(y) = -y$ . Durch Integration erhalten wir  $\ln(f(y)) = -y^2/2 + k$  ( $k \in \mathbb{R}$ ) und  $f(y) = c e^{-0.5y^2}$ . Wegen  $\int f(y) dy = 1$  muß gelten  $c = 1/\sqrt{2\pi}$ . Das kann man nachweisen, indem man das Volumen des Körpers berechnet, der durch Rotation des Graphen von  $x \rightarrow e^{-0.5x^2}$  um die  $y$ -Achse entsteht. Berechnet man das Volumen mit Polarkoordinaten, so ergibt sich  $2\pi$ , führt man die Berechnung in kartesischen Koordinaten durch, ergibt sich  $(\int e^{-0.5x^2} dx)^2$ . (Vgl. auch Scheid S. 114 f).

Ergebnis: Ist  $x$  die Summe aus  $n$  Zufallsgrößen, die alle die gleiche Verteilung (mit Erwartungswert  $\mu$  und Standardabweichung  $\delta$ ) besitzen, dann kommt für die Grenzverteilung ( $n \rightarrow \infty$ ) der normierten Summe  $y = (x - n\mu) / (\delta\sqrt{n})$  nur die Normalverteilung  $f(y) = ce^{-0.5y^2}$  in Frage.

Meist benutzt man den Zentralen Grenzwertsatz für nicht normierte Summen  $x$  wie folgt zur Berechnung von Näherungswerten

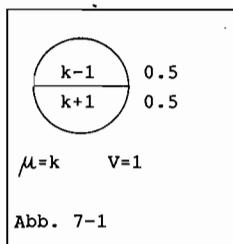
$$P(a < x < b) = P\left(\frac{a-n\mu}{\sigma\sqrt{n}} < Y < \frac{b-n\mu}{\sigma\sqrt{n}}\right) \approx \int_{\frac{a-n\mu}{\sigma\sqrt{n}}}^{\frac{b-n\mu}{\sigma\sqrt{n}}} f(y) dy = \Phi\left(\frac{b-n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{a-n\mu}{\sigma\sqrt{n}}\right),$$

wobei  $\Phi$  eine (tabellierte) Stammfunktion der Normalverteilung  $f$  ist. Anwendungsbeispiele finden sich in 10.3. Der Leser, der die heuristischen Überlegungen auf den allgemeinen Fall nicht identisch verteilter Summanden (und den Fall  $\mu \neq 0$ ) übertragen möchte, sei auf Riemer (1985) verwiesen.

## 7 Eindimensionale Irrfahrten und Chi-Quadrat-Verteilung

Wir modifizieren Irrfahrten, die nach dem Zentralen Grenzwertsatz zur Normalverteilung führen und erhalten überraschend Chi-Quadrat-Verteilungen mit beliebigen Freiheitsgraden. Diese Entdeckung nutzen wir für eine erneute Begründung des Anpassungstests, indem wir die Entwicklung der Testgröße als eindimensionale Irrfahrt deuten. Wer sich mit Computersimulationen begnügen möchte, beschränke sich auf die Abschnitte 7.1 und 7.3. Der Abschnitt 7.2 dient (nach dem Vorbild von Kapitel 6) zur analytischen Vertiefung mit Hilfe von Rekursionsformeln und linearen Differentialgleichungen.

## 7.1 Problemstellung



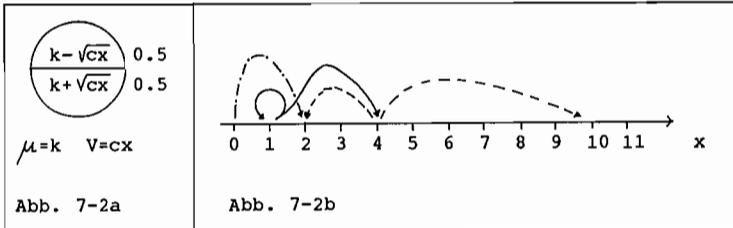
Wenn eine eindimensionale Irrfahrt in jedem Schritt durch eine beliebige Verteilung mit Varianz  $V$  und Erwartungswert  $\mu$  (Abb. 7-1) gesteuert wird, dann gehorcht der Standort  $x_n$  nach  $n$  Schritten annähernd einer Normalverteilung mit Varianz  $nV$  und Erwartungswert  $n\mu$ . Die normierte Zufallsgröße  $y := (x_n - n\mu) / \sqrt{nV}$  genügt dann annähernd der Standardnormalverteilung mit Dichte

$$\varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-0.5 y^2}.$$

Dies ist die Aussage des Zentralen Grenzwertsatzes.

Wir modifizieren nun den Ablauf der Irrfahrt wie folgt: Bei konstantem Erwartungswert  $\mu=k$  sei die Varianz  $V$  der Verteilung, die den jeweils folgenden Schritt der Irrfahrt steuert, proportional zur momentanen Position  $x$  der Irrfahrt (und nicht mehr konstant). Die Verteilung, die den nächsten Schritt steuert, wenn man sich auf der Stelle  $x$  befindet,

könnte aussehen wie in Abb. 7-2a. Für  $k=2$  und Proportionalitätsfaktor  $c=4$  zeigt Abb. 7-2b mögliche Schritte in Abhängigkeit von der momentanen Position.



Man rechnet leicht nach, daß für  $k \geq c/4$  alle Stationen der Irrfahrt in  $\mathbb{R}_0^+$  liegen, die Irrfahrt also definiert ist. Im Mittel befindet sich unsere Irrfahrt nach  $m-1$  Schritten auf der Stelle  $(m-1)k$ . Die Varianz nimmt (wegen der Proportionalität zur momentanen Position) beim  $m$ -ten Schritt um  $c(m-1)k$  zu, ist also nach  $n$  Schritten  $ck(1+2+\dots+n-1) = 0.5 ckn(n-1)$ . (Eine exakte Begründung findet sich in Riemer (1987, Teil I S. 122.)) Während in Abb. 7-1 die Varianz der Irrfahrt proportional zur Schrittzahl  $n$  war, nimmt sie jetzt quadratisch mit  $n$  zu. Deswegen normieren wir nicht mehr mittels Division durch  $\sqrt{n}$  wie beim Zentralen Grenzwertsatz, sondern mittels Division durch  $n$ . Die normierte Position  $y := x_n/n$  unserer Irrfahrt hat dann den Erwartungswert  $k$  und die Varianz  $(c/2)k(1-1/n) \approx (c/2)k$  für großes  $n$ . Wählen wir  $c=4$ , so besitzt die Wahrscheinlichkeitsverteilung von  $y = x_n/n$  für hohes  $n$  den gleichen Erwartungswert ( $k$ ), die gleiche Varianz ( $2k$ ) und den gleichen Definitionsbereich ( $\mathbb{R}_0^+$ ) wie die Chi-Quadrat-Verteilung  $f_k$  mit  $k$  Freiheitsgraden, deren Dichte definiert ist durch

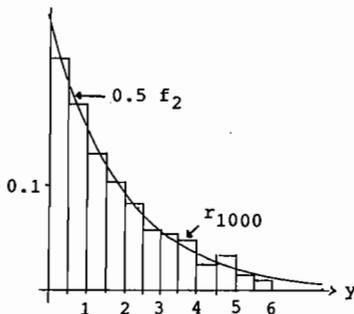
$$f_k(y) := c_k \sqrt{y}^{k-2} e^{-0.5y} \quad \text{mit}$$

$$c_k = \begin{cases} \frac{1}{(k-2)(k-4) \dots 2} \cdot \frac{1}{2} & \text{falls } k \text{ gerade} \\ \frac{1}{(k-2)(k-4) \dots 1} \cdot \frac{1}{\sqrt{2\pi}} & \text{falls } k \text{ ungerade.} \end{cases}$$

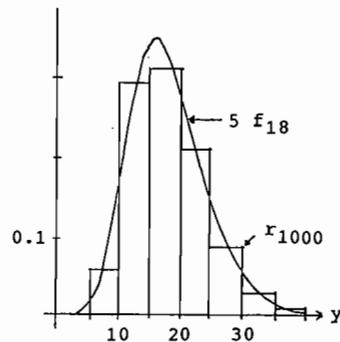
Abb. 7-3 Wir vergleichen die relative Häufigkeit  $r_{1000}$  ( $r_{10000}$ ), mit der die "normierte Endstelle"  $y$  der Irrfahrt nach 50 Schritten im Intervall  $I = [a;b[$  liegt, mit der zugehörigen Wahrscheinlichkeit

$$p = \int_a^b f_k(y) dy \quad (\text{Chi-Quadrat-Verteilung})$$

k=2				k=18			
I	$r_{1000}$	$r_{10000}$	P	I	$r_{1000}$	$r_{10000}$	P
[0.0;0.5[	0.219	0.223	0.221	[ 0; 5[	0.000	0.000	0.001
[0.5;1.0[	0.176	0.175	0.172	[ 5;10[	0.062	0.077	0.067
[1.0;1.5[	0.126	0.130	0.134	[10;15[	0.292	0.264	0.270
[1.5;2.0[	0.101	0.106	0.104	[15;20[	0.315	0.327	0.329
[2.0;2.5[	0.082	0.084	0.081	[20;25[	0.213	0.208	0.208
[2.5;3.0[	0.056	0.060	0.063	[25;30[	0.082	0.094	0.087
[3.0;3.5[	0.055	0.051	0.049	[30;35[	0.024	0.022	0.028
[3.5;4.0[	0.052	0.038	0.038	[35;40[	0.011	0.007	0.007
[4.0;4.5[	0.023	0.032	0.030	[40;45[	0.001	0.001	0.002
[4.5;5.0[	0.033	0.023	0.023	[40;50[	0.000	0.000	0.000
[5.0;5.5[	0.014	0.017	0.018	⋮	⋮	⋮	⋮
[5.5;6.0[	0.009	0.013	0.014				
[6.0;6.5[	0.015	0.012	0.011				
[6.5;7.0[	0.006	0.008	0.009				
⋮	⋮	⋮	⋮				



Chi-Quadrat-Verteilung  
mit  $k=2$  Freiheitsgraden



Chi-Quadrat-Verteilung  
mit  $k=18$  Freiheitsgraden

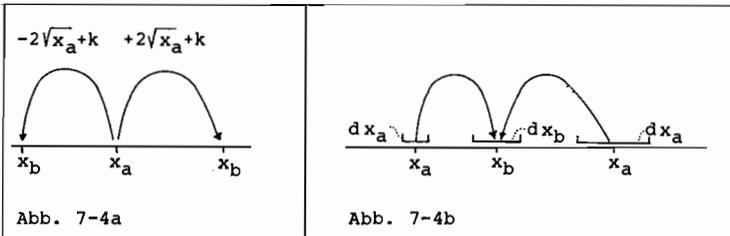
Die Frage, ob  $y$  für hohes  $n$  Chi-Quadrat-verteilt ist mit Freiheitsgrad  $k$ , drängt sich auf. Simulationen (vgl. Abb. 7-3 für  $k=2$  und  $k=18$  mit 1000 und 10000 Irrfahrten der Länge  $n=50$ ) bestätigen unsere Vermutung.

## 7.2 Differentialgleichung der Chi-Quadrat-Verteilung

Wir wollen das experimentelle Ergebnis analytisch untermauern, indem wir (ähnlich wie zuvor beim Zentralen Grenzwertsatz) zeigen, daß die Grenzverteilung der Zufallsgröße  $y = x_n/n$  durch eine Differentialgleichung beschrieben wird, die tatsächlich durch die Chi-Quadrat-Dichte gelöst wird. Wir gehen wieder aus von der rekursiven Definition unserer Irrfahrt: wenn wir die Stelle  $x_a$  erreicht haben, dann springen wir im nächsten Schritt auf

$$x_b := x_a + 2u_i \sqrt{x_a} + k, \quad i=1,2 \text{ mit } u_1=-1, u_2=1$$

je mit Wahrscheinlichkeit  $1/2$  (Abb. 7-4a).



Wir fragen nun umgekehrt: von wo ( $x_a$ ) starten Sprünge, die in  $x_b$  landen (Abb. 7-4b)? Wir haben folgende Gleichung nach  $x_a$  aufzulösen:  $x_a + 2u_i \sqrt{x_a} + k = x_b$ . Es ergibt sich

$$x_a = x_b - k + 2 - 2u_i \sqrt{x_b - k + 1} \quad (1)$$

So können beispielsweise für  $k=2$  die Sprünge, die bei  $x_b=2$  landen, nur von  $x_a=0$  oder  $x_a=4$  gestartet sein (Abb. 7-2b).

Wenn  $P_n(x)$  die Wahrscheinlichkeit bezeichnet, daß man nach  $n$  Schritten die Stelle  $x$  erreicht hat, ( $n \in \mathbb{N}$ ,  $x \in \mathbb{R}^+$ ), dann erhält man (aus Pfad und Additionsregel mit (1)) folgende Rekursionsformel:

$$P_{n+1}(x_b) = \sum_{i=1}^2 \frac{1}{2} P_n(x_b - k + 2 - 2u_i \sqrt{x_b - k + 1}). \quad (2)$$

$\underbrace{\hspace{10em}}_{x_a}$

Da es aber nicht sinnvoll ist, bei "stetigen Prozessen" Punkten Wahrscheinlichkeiten zuzuordnen, arbeiten wir mit Intervallen und Wahrscheinlichkeitsdichten: Bezeichnet  $g_n(x)$  die Wahrscheinlichkeitsdichte für den Standort der Irrfahrt nach  $n$  Schritten, so ist  $g_n(x)dx$  die Wahrscheinlichkeit, daß man sich nach  $n$  Schritten in einem Intervall um  $x$  der Breite  $dx$  befindet. Wir legen nun um die "Zielstelle"  $x_b$  unserer Irrfahrt ein Intervall der Breite  $dx_b$  (Abb. 7-4b). Wie breit ( $dx_a$ ) kann das Intervall um die "Startstelle"  $x_a$  sein, damit man noch in  $dx_b$  landet? Aus (1) folgt durch Differenzieren

$$dx_a = \left(1 - \frac{u_i}{\sqrt{x_b - k + 1}}\right) dx_b \quad (3)$$

und anstelle von (2) erhalten wir für die Dichten

$$g_{n+1}(x_b) dx_b = \sum_{i=1}^2 \frac{1}{2} \left(1 - \frac{u_i}{\sqrt{x_b - k + 1}}\right) g_n(x_b - k + 2 - 2u_i \sqrt{x_b - k + 1}) dx_b \quad (4)$$

$\underbrace{\hspace{10em}}_{dx_a} \quad \underbrace{\hspace{10em}}_{x_a}$

Für große Werte von  $x$  können wir nähern:

$$\sqrt{x - k + 1} \approx \sqrt{x} + \frac{1}{2\sqrt{x}}(-k + 1) \approx \sqrt{x} \quad ;$$

wir setzen  $x_b = s^2$  und (4) vereinfacht sich zu

$$g_{n+1}(s^2) = \sum_{i=1}^2 \frac{1}{2} \left(1 - \frac{u_i}{s}\right) g_n(s^2 - k + 2 - 2u_i s) \quad (5)$$

Die Dichte  $f_n$  der Zufallsgröße  $Y=x_n/n$ , an der wir ja interessiert sind, entsteht aus der Dichte  $g_n$  der Zufallsgröße  $x_n$  durch Strecken mit Faktor  $1/n$  in horizontaler und mit Faktor  $n$  in vertikaler Richtung:

$$f_n\left(\frac{x}{n}\right) = n g_n(x); \quad g_n(x) = \frac{1}{n} f_n\left(\frac{x}{n}\right).$$

Einsetzen in (5) liefert für die normierten Dichten die Rekursionsformel

$$\frac{1}{n+1} f_{n+1}\left(\frac{s^2}{n+1}\right) = \frac{1}{n} \sum_{i=1}^2 \frac{1}{2} \left(1 - \frac{u_i}{s}\right) f_n\left(\frac{s^2 - 2u_i s - k + 2}{n}\right). \quad (6)$$

Wir setzen  $y := x_p/n = s^2/n$ . Wenn die Dichten  $f_n$  gegen die Dichte einer Grenzverteilung  $f$  konvergieren ( $n \rightarrow \infty$ ), so können wir im Grenzfall  $f_n = f_{n+1} = f$  setzen. Es ergibt sich

$$\underbrace{f\left(y\left(\frac{n}{n+1}\right)\right)}_A = \underbrace{\frac{n+1}{n}}_C \sum_{i=1}^2 \underbrace{\frac{1}{2} \left(1 - \frac{u_i}{\sqrt{y}} \cdot \frac{1}{\sqrt{n}}\right)}_B \underbrace{f\left(y - 2u_i \sqrt{y} \frac{1}{\sqrt{n}} + \frac{2-k}{n}\right)}_D. \quad (7)$$

Wir entwickeln die Terme A, B, C und D auf beiden Seiten bis zur ersten Ordnung in  $1/n$  (also bis zur zweiten Ordnung in  $1/\sqrt{n}$ ). Wegen

$$y\left(\frac{n}{n+1}\right) \approx y - \frac{y}{n} \quad \text{gilt} \quad A \approx f(y) - \frac{y}{n} f'(y).$$

Weiter hat man:

$$B \approx f(y) + (-2u_i \sqrt{y} \frac{1}{\sqrt{n}} + \frac{2-k}{n}) f'(y) + (2u_i^2 \cdot \frac{y}{n}) f''(y).$$

Hieraus ergibt sich mit  $\sum u_i = 0$ ,  $\sum \frac{1}{2} u_i^2 = 1$  und unter Vernachlässigung von Termen höherer Ordnung (etwa  $\frac{1}{n\sqrt{n}}$ )

$$C \approx f(y) + \left( \frac{2}{n} + \frac{2-k}{n} \right) f'(y) + \frac{2y}{n} f''(y).$$

Multiplikation mit  $(1 + \frac{1}{n})$  liefert wieder unter Vernachlässigung von Termen höherer Ordnung

$$D \approx f(y) + \frac{1}{n} f(y) + \frac{4-k}{n} f'(y) + \frac{2y}{n} f''(y).$$

Der Vergleich von A mit D liefert für f die lineare Differentialgleichung:

$$-y f'(y) = f(y) + (4-k)f'(y) + 2yf''(y)$$

oder

$$f(y) + (4-k+y)f'(y) + 2yf''(y) = 0 \quad (8)$$

Wie man nachrechnet, wird diese Differentialgleichung gelöst durch

$$f_k(y) = c(\sqrt{y})^{k-2} e^{-0.5y},$$

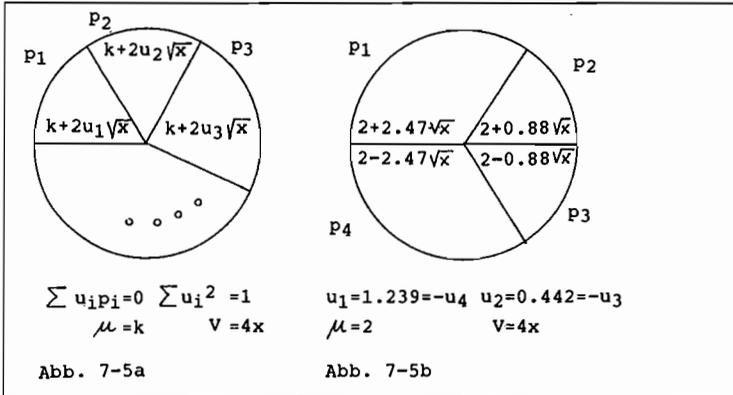
also durch die Dichte der Chi-Quadrat-Verteilung mit k Freiheitsgraden.

Ergebnis: Für die Grenzverteilung der normierten Irrfahrt kommt nur die Chi-Quadrat-Verteilung mit k Freiheitsgraden in Frage. Damit sind die Ergebnisse der Computersimulation kein Zufall.

### 7.2.1 Verallgemeinerung auf beliebige Sprungverteilungen

Auch wenn man die Irrfahrt nicht durch die spezielle Verteilung aus Abb. 7-2, steuert, sondern allgemeiner durch eine Verteilung gemäß Abb. 7-5a (mit  $\mu=k$  und  $V=4x$ ), erhält man die Chi-Quadrat-Verteilung als Grenzverteilung. (In der Tat

stammt die Spalte  $r_{10000}$  (Abb. 7-3) von einer Irrfahrt, die durch Abb. 7-5b gesteuert wurde.) Man überzeugt sich leicht, daß auch die Herleitung der Differentialgleichung (8) auf den allgemeinen Fall (Abb. 7-5a) übertragbar ist.



### 7.2.2 Gammaverteilung

Wählen wir in unserer Irrfahrt die Proportionalitätskonstante  $c$  (zwischen Standort  $x$  und Varianz  $V$ ) beliebig (also nicht notwendig  $c=4$  wie in Abb. 7-2 oder 7-5), so erhalten wir statt (8) die Differentialgleichung

$$f(y) + (c-k+y)f'(y) + \frac{c}{2} y f''(y) = 0 \quad (9)$$

Sie wird gelöst durch die Dichte der Gammaverteilung

$$f(y) = \frac{b^p}{\Gamma(p)} y^{p-1} e^{-by} \quad , \quad y > 0 \quad \text{mit } p = \frac{2k}{c} \quad \text{und} \quad b = \frac{2}{c} .$$

Auf diese Weise können wir mit unserer Irrfahrt beliebige Gammaverteilungen erzeugen.

## 7.3 Anpassungstest

Mit der Anpassungs-Testgröße

$$y := \frac{1}{n} \sum_{i=1}^{k_0} \frac{(n_i - np)^2}{p} \quad (10)$$

überprüft man, ob die Hypothese einer Gleichverteilung ( $p=1/k$ ) mit den in einem Versuch ermittelten absoluten Häufigkeiten  $(n_1, n_2, \dots, n_k)$  verträglich ist ( $n_1 + n_2 + \dots + n_k = n$ ). Hat  $y$  einen zu hohen Wert, so weist man die Hypothese zurück (vgl. Kapitel 3). Was haben unsere eindimensionalen Irrfahrten mit dem Chi-Quadrat-Anpassungstest zu tun?

Um diese Frage zu beantworten, setzen wir

$$x := \sum_{i=1}^{k_0} \frac{(n_i - np)^2}{p} = \sum_{i=1}^{k_0} \frac{z_i^2}{p} \quad \text{mit } z_i = n_i - np. \quad (11)$$

In jedem Schritt eines Anpassungstests nimmt der Versuchsumfang  $n$  und eine der absoluten Häufigkeiten  $n_i$  um 1 zu. Dadurch verändert sich  $x$  um einen gewissen Betrag  $d$ . Es läuft also während eines Anpassungstests eine eindimensionale Irrfahrt ab. Wir zeigen: Wenn sie die Stelle  $x$  erreicht hat, dann ist der Zuwachs  $d$  eine Zufallsgröße mit konstantem Erwartungswert  $k-1$  und der Varianz  $4x$ , die zur momentanen Position proportional ist. Die Irrfahrt, die durch (11) definiert wird, hat also eine Struktur, wie wir sie in den vorangegangenen Abschnitten 7.1 und 7.2 studiert haben. Folglich wird die normierte Irrfahrtposition  $y:=x/n$ , also die Anpassungstestgröße (10), für hohes  $n$  eine Chi-Quadrat-Verteilung mit  $k-1$  Freiheitsgraden besitzen. Damit haben wir den Anpassungstest erneut (und unabhängig von Kapitel 3) begründet.

Wir tragen nun den Beweis der Aussagen nach, die wir über Erwartungswert und Varianz des Zuwachses  $d$  soeben benutzt haben. Die Irrfahrt (11) habe (nach  $n$  Schritten) die Stelle

$$x_a = \sum_{i=1}^k \frac{z_i^2}{p}$$

erreicht. Falls im  $n+1$ -ten Schritt "1" fällt, erhöht sich  $z_1$  ( $=n_1 - np$ ) um  $1-p$  (dennsowohl  $n$  als auch  $n_1$  werden um 1 erhöht).  $z_2, \dots, z_k$  erniedrigen sich je um  $p$ . Folglich gilt für die neue Position  $x_b$  der Irrfahrt (wegen  $kp=1$ )

$$\begin{aligned} x_b &= (1/p) (z_1 + (1-p))^2 + (z_2 - p)^2 + \dots + (z_k - p)^2 \\ &= \underbrace{(z_1^2 + z_2^2 + \dots + z_k^2)}_{x_a} / p - 2 \underbrace{(z_1 + z_2 + \dots + z_k)}_0 + (2/p) (z_1 - p) + 1/p + kp \\ &= x_a + 2z_1/p - 2 + k + 1 = x_a + 2z_1/p + k - 1 \\ &= x_a + d_1 \quad \text{mit } d_1 = 2z_1/p + k - 1. \end{aligned}$$

Falls im  $n+1$ -ten Schritt "i" fällt, ändert sich (analog) die Irrfahrtposition  $x_a$  um  $d_i = 2z_i/p + k - 1$ . Der Erwartungswert der Änderung  $d$  (Zufallsgröße) ist konstant:

$$\mu(d) = \sum_{i=1}^k p d_i = 2 \underbrace{\sum_{i=1}^k z_i}_0 + k - 1 = k - 1.$$

Die Varianz der Änderung ist mit

$$v(d) = \sum_{i=1}^k p (d_i - (k-1))^2 = \sum_{i=1}^k (2z_i)^2 = 4 \sum_{i=1}^k z_i^2 = 4x_a$$

tatsächlich proportional zur momentanen Position  $x_a$ .

## 8 Eindimensionale Irrfahrten und Arcsin-Verteilung

"Wer durch Zufall nach vorne kam, hält sich oft lange an der Spitze."

Wir beschäftigen uns mit dem Wechsel der Führung bei Wettkämpfen gleichguter Kontrahenten. Es zeigt sich, daß die seit dem letzten Führungswechsel verstrichene Zeit Arcsin-verteilt ist. Dieses Ergebnis ist noch relativ neu, es wurde 1967 von Feller bewiesen und hat überraschende Konsequenzen, die unserer Intuition zu widersprechen scheinen. Uns gelingt eine elementare Darstellung dadurch, daß wir das Problem der wechselnden Führung als Spezialfall einer Irrfahrt deuten, auf die sich die in den vorangehenden Kapiteln entwickelte "Technik der linearen Differentialgleichungen" anwenden läßt.

Wer nur die Phänomene studieren möchte, beschränke sich auf die Abschnitte 8.1 bis 8.3 sowie den Anfang und das Ende von 8.5. Vertiefungsmöglichkeiten bietet 8.4.

### 8.1 Einführung

"Steffi wird besser benotet, sie lag in den letzten Monaten beständig vor Nicole." Das Bild, das wir uns beispielsweise von Schülern machen, ergibt sich aus einer Vielzahl von Beobachtungen, die sich zu einem Gesamtbild "aufsummieren". Die Schüler, die im Zeitraum vor der Notegebung vorne lagen, werden besser beurteilt. Wir werden im folgenden ein "Paradoxon" studieren, das unserer im Alltag geprägten Intuition widerspricht. Einem merkwürdigen "Gesetz des Durchschnitts" zufolge haben wir nämlich das Gefühl, daß zwei gleichgute Schüler (zwei gleichstarke Spieler - oder noch unverfänglicher die Seiten einer ungezinkten Münze) während einer langen Beobachtungsphase etwa gleichhäufig in Führungsposition liegen werden. Weiterhin sind wir überzeugt,

daß die Führung des öfteren wechseln wird. Tatsächlich ist das Gegenteil wahr. Ein einfaches Beispiel möge erläutern, was gemeint ist: Wir werfen eine normale Münze 38mal und notieren nach jedem Wurf, um welchen Betrag Zahl vor Kopf führt. Z. B. bedeutet "5", daß Zahl momentan 5 Positionen vor Kopf liegt, "0" bedeutet "Ausgleich" und bei "-1" liegt Kopf eine Position hinter Zahl. Da wir mit "0" starten, enthält das Protokoll 39 Positionen:

0 1 2 1 2 1 0 -1 0 -1 0 1 2 3 4 5 6 7 8 7  
8 7 8 7 8 7 8 7 6 5 4 3 4 5 4 3 2 3 4 ... Tab. 8-1

Zahl liegt am Ende in Führung - und hatte diese Führung im Laufe der letzten 28 Schritte inne. Würden wir die Münze 10000 mal werfen, so hätte die Seite, die am Ende führt, in über 11% aller Fälle die Führung in den letzten 9700 Würfungen nicht abgegeben. Es passiert im Mittel immerhin einem von 7 Schülern, daß er - möglicherweise bis auf die ersten 5 Stunden - während der restlichen 95 Stunden eines Schuljahres ständig hinter einem Kameraden zurücklag, obwohl er gleichgütig war und die unterschiedliche Bewertung durch puren Zufall zustande kam. Würden wir beide Schüler gleich benoten?

## 8.2 Irrfahrt: "Wer weit weg ist, der wird selten gefangen"

Wir untersuchen (zunächst unabhängig von 8.1) die Irrfahrt eines Teilchens im Bereich des positiven Zahlenstrahls. Das Teilchen versucht, mit Sprüngen der Weite  $e$  nach rechts "auszureißen". Es wird aber mit einer gewissen Wahrscheinlichkeit "eingefangen", muß dann zurück zum Ursprung und neu starten (Abb. 8-2). Je weiter das Teilchen schon vom Ursprung entfernt ist (Stelle  $x=ie$ ) desto größer sei die Wahrscheinlichkeit  $r(x)$ , beim nächsten Schritt weiter nach rechts zu springen, desto kleiner sei die Wahrscheinlichkeit  $l(x)$ , beim nächsten Schritt im Ursprung zu landen. (Wer sofort wissen möchte, wie sich die in 8.1 angesprochene Thematik als Spezialfall dieser Irrfahrt deuten läßt, werfe einen kurzen Blick auf den Anfang von Abschnitt 8.5.)

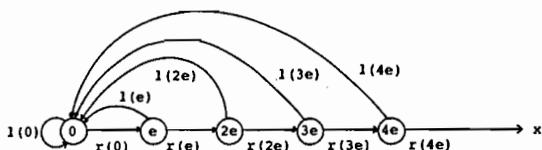


Abb. 8-2 Flucht eines Teilchens mit Sprungweite e

Wir wollen verlangen: Wenn sich das Teilchen bei  $x$  befindet, dann sei der Erwartungswert für die nächste Position:  $x+d$ . Das Teilchen komme also im Mittel bei jedem Schritt um den konstanten Betrag  $d$  vorwärts. Aus dieser Forderung ergibt sich:

$$l(x) \cdot 0 + r(x) \cdot (x+e) = x+d, \text{ also}$$

$$r(x) = \frac{x+d}{x+e} \quad \rightarrow 1 \quad \text{für } x \rightarrow \infty$$

$$l(x) = 1-r(x) = \frac{e-d}{x+e} \quad \rightarrow 0 \quad \text{für } x \rightarrow \infty$$

Der Erwartungswert für die Position  $x$  des Teilchens nach  $n$  Schritten ist gemäß Konstruktion  $n \cdot d$ . Als Beispiel wählen wir  $e=2$  und  $d=1$ , also  $r(x)=(x+1)/(x+2)$  und  $l(x)=1/(x+2)$ . Hier das Protokoll einer Irrfahrt mit  $n=19$  Schritten (also 20 Positionen):

0 2 4 0 0 0 2 4 6 8 10 12 14 16 18 20 22 24 26 28

Je weiter das Teilchen schon gekommen ist, desto geringer wird die Wahrscheinlichkeit, daß es im nächsten Schritt auf 0 zurückgeworfen wird. Für  $x=28$  beträgt diese Wahrscheinlichkeit nur noch  $l(28)=3.3\%$ . Wir wollen vom Einzelfall abstrahieren und fragen: Wo wird das Teilchen am Ende einer langen Irrfahrt mit  $n$  Schritten stehen? Eher in der Nähe des Ursprunges, eher "weit draußen" (in der Nähe von  $n \cdot e$ ) oder irgendwo in der Mitte bei  $(n/2) \cdot e$  ?

Um eine erste Antwort zu erhalten, simulieren wir je 1000 Irrfahrten mit  $n=19$ ,  $n=99$  und  $n=199$  Schritten (20, 100 bzw. 200 mögliche Endpositionen). Um die Positionen  $x$  der Teilchen am Ende der Irrfahrt besser miteinander vergleichen zu können, normieren wir sie, indem wir durch  $ne$  teilen:  $y:=x/(ne)$ . Während  $x$  die Werte  $0, e, 2e, \dots, ne$  annehmen kann, befindet sich die normierte Endposition  $y$  im Intervall  $[0;1]$ . Auch der Erwartungswert von  $y$  ist mit  $d/e$  (in unserem Beispiel 0.5) von der Irrfahrtdauer  $n$  unabhängig. Die Spalten 2 bis 4 von Tab. 8-3 zeigen, wie sich die normierten Endpositionen der (je 1000) Irrfahrten auf die in Spalte 1 angegebenen Teilintervalle von  $[0;1]$  verteilen. Wir erkennen:

- Die häufigsten Werte liegen an den Rändern. Am Ende der Irrfahrt steht man also am wahrscheinlichsten in der Nähe von 0 oder  $ne$ , das Teilchen ist erst vor sehr kurzer Zeit oder schon sehr lange nicht mehr eingefangen worden.
- Die Dauer  $n$  der Irrfahrt hat keinen merklichen Einfluß auf die Verteilung der normierten Endposition. Es gibt eine Grenzverteilung.
- Die Verteilungen sind für  $e=2$  und  $d=1$  symmetrisch.

Intervall	relative Häufigkeit der normierten Endposition $y$ (bei je 1000 Irrfahrten)			zugehörige Wahrscheinlichkeit			Grenzverteilung
	$n = 19$	$n = 99$	$n = 199$	$n = 19$	$n = 99$	$n = 199$	
$[0,0;0,1]$	0,188	0,194	0,192	0,195	0,204	0,203	0,205
$[0,1;0,2]$	0,095	0,101	0,087	0,095	0,091	0,091	0,090
$[0,2;0,3]$	0,070	0,075	0,078	0,076	0,074	0,074	0,074
$[0,3;0,4]$	0,076	0,054	0,067	0,069	0,067	0,067	0,067
$[0,4;0,5]$	0,069	0,083	0,067	0,066	0,064	0,064	0,064
$[0,5;0,6]$	0,057	0,051	0,073	0,066	0,064	0,064	0,064
$[0,5;0,7]$	0,081	0,073	0,070	0,069	0,067	0,067	0,067
$[0,7;0,8]$	0,071	0,082	0,076	0,076	0,074	0,074	0,074
$[0,8;0,9]$	0,085	0,109	0,088	0,095	0,091	0,091	0,090
$[0,9;1,0]$	0,208	0,178	0,202	0,195	0,204	0,203	0,205

$n$  = Dauer der Irrfahrt  
 $x$  = Endposition der Irrfahrt  
 $y = x/(ne)$  normierte Endposition der Irrfahrt

Tab. 8-3 Ein Teilchen springt um  $e=2$  nach rechts oder zum Ursprung zurück. Im Mittel kommt es um  $d=1$  voran.

## 8.3 Rekursionsformel

Wir begnügen uns nicht mit Simulationen und berechnen rekursiv (wie in den Kapiteln 6 und 7) die Wahrscheinlichkeitsverteilung für die Endposition unserer Irrfahrt nach  $n$  Schritten (Spalten 5 bis 7 in Tab. 8-3).

$P_n(x_i)$  bezeichne die Wahrscheinlichkeit, daß man nach  $n$  Schritten auf der Stelle  $x_i := ie$  ( $i \in \mathbb{N}$ ) steht. Da man nur von der Stelle  $x_{i-1}$  auf die Stelle  $x_i$  ( $i > 0$ ) gelangt sein kann, und zwar durch einen Sprung der Weite  $e$  mit Wahrscheinlichkeit  $r(x_{i-1}) = (x_{i-1} + d)/x_i$ , gilt

$$P_{n+1}(x_i) = P_n(x_{i-1}) \frac{x_{i-1} + d}{x_i} \quad 0 < i \leq n+1. \quad (1)$$

Analog hat man mit  $l(x_i) = (e-d)/x_{i+1}$

$$P_{n+1}(0) = \sum_{i=0}^n P_n(x_i) \cdot l(x_i) = \sum_{i=0}^n P_n(x_i) \frac{e-d}{x_{i+1}}. \quad (2)$$

Mit den Gleichungen (1) und (2) berechnet man nacheinander  $P_1, P_2, P_3, \dots$ . Tab. 8-4 zeigt die ersten 19 Berechnungsschritte (wieder für  $e=2, d=1$ ). Fassen wir in der 19ten Zeile jeweils 2 Wahrscheinlichkeiten zu einer Klasse zusammen, so erhalten wir Spalte 5 von Tab. 8-3. So liegt beispielsweise die normierte Endposition  $y = x_i / (19e) = i/19$  genau dann im Intervall  $[0.0; 0.1[$ , wenn  $i=0$  oder  $i=1$  gilt. Das passiert mit Wahrscheinlichkeit  $0.129 + 0.066 = 0.195$ . Die Fortsetzung der Tabelle und Zusammenfassen von jeweils 10 Wahrscheinlichkeiten für  $n=99$  (20 Wahrscheinlichkeiten für  $n=199$ ) zu einer Klasse liefert die Spalten 6 und 7 von Tab. 8-3. Tab. 8-4 und die genannten Spalten von Tab. 8-3 belegen wieder die Symmetrie der Verteilungen (für  $e=2, d=1$ ) und die Tatsache, daß die größten Wahrscheinlichkeiten an den Rändern auftreten.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	$\leftarrow i$
0	1																				
1	0,500	0,500																			
2	0,375	0,250	0,375																		
3	0,313	0,188	0,188	0,313																	
4	0,273	0,156	0,141	0,156	0,273																
5	0,246	0,137	0,117	0,117	0,137	0,246															
6	0,226	0,123	0,103	0,098	0,103	0,123	0,226														
7	0,209	0,113	0,092	0,085	0,085	0,092	0,113	0,209													
8	0,196	0,105	0,085	0,077	0,075	0,077	0,085	0,105	0,196												
9	0,185	0,098	0,079	0,070	0,067	0,067	0,070	0,079	0,098	0,185											
10	0,176	0,093	0,074	0,065	0,062	0,061	0,062	0,065	0,074	0,093	0,176										
11	0,168	0,088	0,070	0,061	0,057	0,056	0,056	0,057	0,061	0,070	0,088	0,168									
12	0,161	0,084	0,066	0,058	0,054	0,052	0,051	0,052	0,054	0,058	0,066	0,084	0,161								
13	0,155	0,081	0,063	0,055	0,051	0,048	0,047	0,047	0,048	0,051	0,055	0,063	0,081	0,155							
14	0,149	0,077	0,060	0,053	0,048	0,046	0,044	0,044	0,044	0,046	0,048	0,053	0,060	0,077	0,149						
15	0,144	0,075	0,058	0,050	0,046	0,043	0,042	0,041	0,041	0,042	0,043	0,046	0,050	0,058	0,075	0,144					
16	0,140	0,072	0,056	0,048	0,044	0,041	0,040	0,039	0,039	0,040	0,041	0,044	0,048	0,056	0,072	0,140					
17	0,136	0,070	0,054	0,047	0,042	0,040	0,038	0,037	0,036	0,036	0,037	0,038	0,040	0,042	0,047	0,054	0,070	0,136			
18	0,132	0,068	0,052	0,045	0,041	0,038	0,036	0,035	0,034	0,034	0,035	0,036	0,038	0,041	0,045	0,052	0,068	0,132			
19	0,129	0,066	0,051	0,044	0,040	0,037	0,035	0,034	0,033	0,033	0,033	0,033	0,034	0,035	0,037	0,040	0,044	0,051	0,066	0,129	

$$P_{19}(y \in 0; 0,1) = P_{19}(y \in 0_{19}^0; \frac{1}{2}) = 0,129$$

$$P_{19}(y \in 0_{19}^0; \frac{1}{2}) = P_{19}(x_8) + P_{19}(x_9) = 0,066$$

$$P_{19}(y \in 10,4; 0,5) = P_{19}(y \in 10_{19}^0; \frac{1}{2}) = 0,066$$

Tab. 8-4 Rekursive Berechnung der Wahrscheinlichkeit  $P_n(x_i)$ , daß man nach n Schritten auf der Stelle  $x_i=2i$  steht

Die Wahrscheinlichkeiten der Tabelle 8-4 genügen der Formel

$$P_n(x_i) = \binom{2i}{i} \binom{2(n-i)}{n-i} 2^{-2n} \quad (3)$$

Dabei gibt der Term  $\binom{2i}{i} 2^{-2i}$  die Wahrscheinlichkeit an, daß in den letzten  $i$  Schritten keine Rückkehr zum Ursprung erfolgte, der Term  $\binom{2(n-i)}{n-i} 2^{-2(n-i)}$  ist die Wahrscheinlichkeit, daß man sich zur Zeit  $n-i$  im Ursprung aufhielt. Ein Beweis dieser Formel ergibt sich im Zusammenhang mit der Münzwurfinterpretation unserer Irrfahrt am Ende von 8.5.1. Für  $d/e \neq 0.5$  ist eine explizite Formel für  $P_n$  nicht bekannt.

#### 8.4 Differentialgleichung der Arcsin-Verteilung

Wir gehen von der in Tab. 8-3 gemachten Beobachtung aus, daß die Verteilung der normierten Standorte  $y$  für wachsendes  $n$  einer Grenzverteilung zustrebt. Unser Ziel ist es, deren Dichte  $f$  zu bestimmen (und damit die letzte Spalte von Tab. 8-3 zu berechnen). Wir benutzen die Rekursionsformel

$$P_{n+1}(x) = P_n(x-e) \cdot \frac{x-e+d}{x} \quad 0 < x \leq (n+1)e. \quad (1a)$$

Die Dichte  $f_n$  der zugehörigen normierten Zufallsgröße  $y := x/(ne)$  ist definiert durch  $f_n(y) := ne \cdot P_n(x)$  oder

$$P_n(x) = \frac{1}{ne} f_n(y), \quad \text{wobei gilt} \quad x = ney. \quad (4)$$

(Anschaulich wird das Histogramm zur Verteilung  $P_n$  in horizontaler Richtung um den Faktor  $ne$  gestaucht und in vertikaler Richtung um den gleichen Faktor gestreckt. Die zugehörige Fläche "1" wird durch diese Transformation nicht verändert.) Wir setzen (4) in die Rekursionsformel (1a) ein und erhalten für die Wahrscheinlichkeitsdichten  $f_n$  der normierten Zufallsgröße :

$$\frac{1}{(n+1)e} \cdot f_{n+1}\left(\frac{x}{(n+1)e}\right) = \frac{1}{ne} \cdot \frac{x-e+d}{x} f_n\left(\frac{x-e}{ne}\right) \quad \text{oder}$$

$$f_{n+1}\left(\frac{x}{en}\right) \cdot \frac{en}{e(n+1)} = \left(1 + \frac{1}{n}\right) \left(1 - \frac{e-d}{x}\right) f_n\left(\frac{x-e}{ne}\right).$$

mit  $y=x/(ne)$  ( $0 < y < 1$ ) ergibt sich

$$f_{n+1} \left( y \frac{n}{n+1} \right) = \left( 1 + \frac{1}{n} \right) \cdot \left( 1 - \frac{e-d}{ney} \right) \cdot f_n \left( y - \frac{1}{n} \right).$$

Für große  $n$  können wir nähern

$$y \frac{n}{n+1} = y \frac{1}{1+(1/n)} \approx y - \frac{y}{n}, \quad f_n \approx f_{n+1} \approx f$$

und auf beiden Seiten eine Taylorentwicklung um die Stelle  $y$  vornehmen. Es ergibt sich

$$f(y) - \frac{y}{n} \cdot f'(y) = \left( 1 + \frac{1}{n} \right) \cdot \left( 1 - \frac{e-d}{ney} \right) \cdot \left( f(y) - \frac{1}{n} f'(y) \right).$$

Ausmultiplizieren liefert unter Vernachlässigung von Termen höherer Ordnung in  $1/n$

$$f(y) - \frac{y}{n} \cdot f'(y) = f(y) + \frac{1}{n} f(y) - \frac{e-d}{ney} f(y) - \frac{1}{n} f'(y).$$

Wir vereinfachen und erhalten für die gesuchte Grenzverteilung  $f$  die lineare Differentialgleichung erster Ordnung:

$$(1-y) f'(y) = \left( 1 + \frac{(d/e)-1}{y} \right) f(y).$$

Mit der Partialbruchzerlegung

$$\frac{f'(y)}{f(y)} = \frac{y+(d/e)-1}{y(1-y)} = -\frac{a}{y} + \frac{b}{1-y}$$

und  $a=1-(d/e)$ ,  $b=d/e$  erhalten wir durch Integrieren

$$\int \frac{f'(y)}{f(y)} dy = -\int \frac{a}{y} dy + \int \frac{b}{1-y} dy$$

also

$$\ln(f(y)) = -a \cdot \ln(y) - b \cdot \ln(1-y) + c, \quad c \in \mathbb{R}.$$

Es folgt

$$f(y) = \frac{k}{y^a (1-y)^b} \quad (0 < y < 1).$$

Für  $k = \sin(a\pi)/\pi$  ist dies die Dichte der Beta-Verteilung. Mitunter bezeichnet man sie (wegen  $a+b=1$ ) auch als verallgemeinerte Arcsin-Verteilung. Diese Bezeichnung rührt daher, daß für  $a=b=1/2$  gilt

$$f(y) = \frac{1}{\pi \sqrt{y(1-y)}}$$

und diese Dichte läßt sich durch

$$F(y) = \frac{2}{\pi} \arcsin(\sqrt{y})$$

geschlossen integrieren (Feller II, S. 471, S. 50). Damit können wir auch die letzte Spalte von Tab. 8-3 berechnen. Beispielsweise gilt

$$P(0.0 \leq y < 0.1) = \frac{2}{\pi} \left( \arcsin(\sqrt{0.1}) - \arcsin(\sqrt{0}) \right) = 0.205.$$

Ergebnis: Die (normierte) Position  $y$  eines Teilchens, das gemäß Abb. 8-2 dem Ursprung zu entfliehen sucht (es springt in jedem Schritt um  $e$  Einheiten nach rechts oder zum Ursprung zurück und bewegt sich dabei im Mittel um  $d$  Einheiten vorwärts), wird nach hinreichend vielen Schritten beschrieben durch die Arcsin-Dichte:

$$f(y) = \frac{\sin(a\pi)}{\pi y^a (1-y)^b} \quad \text{mit } b=d/e, a=1-b, 0 < y < 1 \quad (5)$$

Der Verlauf dieser an den Rändern des Intervalls  $]0;1[$  unbeschränkten Dichte ist für  $a=1/2$  und  $a=1/3$  in Abb. 8-5 dargestellt.

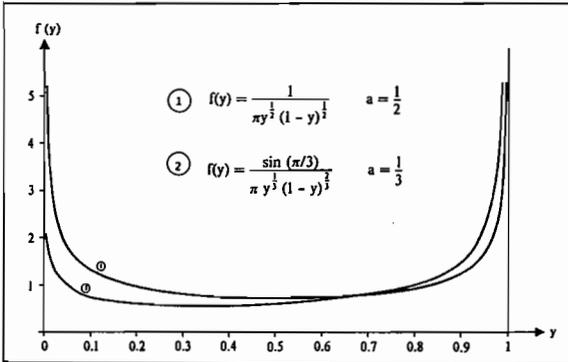


Abb. 8-5 Dichte der Arcsin-Verteilung

Zum Abschluß studieren wir in Tab. 8-6 das Verhalten eines Teilchens, das sich in jedem Schritt um  $e=6$  Einheiten nach rechts (oder zum Ursprung zurück) bewegt und dabei im Mittel um  $d=4$  Einheiten vorwärts kommt. Hier hat man also  $a=1/3$ .

Intervall	relative Häufigkeit der normierten Endposition $y$ (bei je 1000 Irrfahrten)			zugehörige Wahrscheinlichkeit			Grenzverteilung
	$n = 19$	$n = 99$	$n = 199$	$n = 19$	$n = 99$	$n = 199$	
$\{0,0;0,1[$	0,081	0,105	0,087	0,088	0,091	0,091	0,090
$[0,1;0,2[$	0,056	0,061	0,058	0,060	0,059	0,058	0,059
$[0,2;0,3[$	0,064	0,048	0,054	0,055	0,053	0,053	0,053
$[0,3;0,4[$	0,062	0,055	0,049	0,054	0,052	0,052	0,052
$[0,4;0,5[$	0,051	0,050	0,054	0,055	0,054	0,054	0,054
$[0,5;0,6[$	0,058	0,064	0,071	0,059	0,058	0,058	0,058
$[0,6;0,7[$	0,061	0,068	0,060	0,067	0,065	0,065	0,065
$[0,7;0,8[$	0,080	0,071	0,073	0,081	0,078	0,077	0,078
$[0,8;0,9[$	0,118	0,120	0,109	0,114	0,107	0,106	0,107
$[0,9;1,0[$	0,369	0,358	0,385	0,368	0,383	0,385	0,385

$n$  = Dauer der Irrfahrt  
 $x$  = Endposition der Irrfahrt mit  
 $y = x/(ne)$  normierte Endposition der Irrfahrt

Abb. 8-6 Ein Teilchen springt um  $e=6$  nach rechts oder zum Ursprung zurück. Im Mittel kommt es um  $d=4$  voran.

Der Vergleich mit Abb. 8-3 (wo  $e=2$ ,  $d=1$ , also  $a=1/2$  galt) ist lehrreich. Das Teilchen ist jetzt nicht mehr so stark an den Ursprung gebunden. Die Symmetrie geht verloren, aber die gute Konvergenz der Wahrscheinlichkeitsverteilungen mit wachsender Schrittzahl  $n$  gegen die Grenzverteilung ist wieder beeindruckend. Die letzte Spalte wurde durch numerische Integration der Arcsin-Dichte berechnet.

### 8.5 Länge des letzten Runs

Wir stellen nun die Beziehung her zwischen dem wiederholten Werfen einer Münze (Abschnitt 8.1) und den Irrfahrten, die wir in den Abschnitten 8.2 bis 8.4 untersucht haben: Die Zeit, die beim Münzwurf seit dem letzten Führungswechsel verstrichen ist ("Länge des letzten Runs", Abb. 8-1) läßt sich deuten als Position  $x$  der Irrfahrt eines Teilchens, das versucht, gemäß Abb. 8-2 dem Ursprung zu "entkommen".

Anfangs sind Kopf und Zahl gleichhäufig (0mal) gefallen, die seit dem letzten Ausgleich verstrichene Zeit ist folglich  $x=0$  (Startposition). Nun wirft man die Münze zweimal, man führt einen Doppelwurf aus. Sind dabei sowohl Kopf als auch Zahl gefallen (Wahrscheinlichkeit  $1/2$ ), steht das Spiel wieder ausgeglichen. Die seit dem letzten Ausgleich verstrichene Zeit ist dann schon wieder  $x=0$ . Man beginnt von vorne zu zählen. Ist dagegen (a) zweimal Kopf oder (b) zweimal Zahl gefallen (zusammen Wahrscheinlichkeit  $1/2$ ), hat man die Spieldauer  $x=2$  erreicht. Um von hier aus wieder Ausgleich herzustellen, braucht man im Fall (a) zweimal Zahl, im Fall (b) zweimal Kopf (jeweils Wahrscheinlichkeit  $1/4$ ). Mit Wahrscheinlichkeit  $3/4$  erreicht man also die Spieldauer 4.

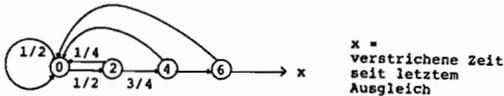


Abb. 8-7 Irrfahrt und verstrichene Zeit

Wir erkennen in Abb. 8-7 den Anfang der Irrfahrt aus Abb. 8-2 (mit  $e=2$ ,  $d=1$ ,  $l(2i)=1/(2i+2)$ ,  $r(2i)=(2i+1)/(2i+2)$ ).

### 8.5.1 Exkurs

(kann beim ersten Lesen überschlagen werden, man setze dann die Lektüre beim Resümee fort. Für das folgende vergleiche man auch Engel Bd. 2 S. 52 f und Feller Bd. 1 S. 67 f.)  
Zur Vervollständigung von Abb. 8-7 berechnen wir zunächst die Wahrscheinlichkeit, daß in  $n$  Doppelwürfen (abgesehen vom Start) Kopf beständig vor Zahl lag. Das passiert genau dann, wenn der Pfad, der den Münzwurf protokolliert, oberhalb der  $n$ -Achse verläuft.

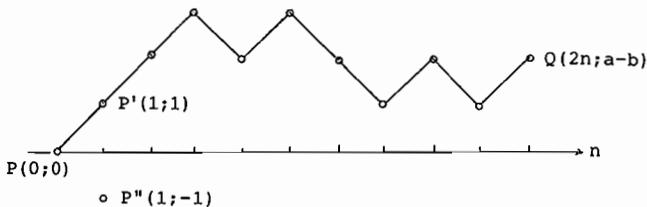


Abb. 8-8 Protokoll einer Münzwurfserie, bei der bis zur Zeit  $2n$  Kopf beständig vor Zahl liegt.  
(Hier  $2n=10$ , Serie: KKKZKZZKZK,  $a-b=2$ )

Er kann dann in den Punkten  $(2n;2)$ ,  $(2n;4)$ , ...  $(2n;2n)$  (allgemein  $Q(2n;a-b)$ ) enden, wobei  $a$  und  $b$  zählen, wie oft Kopf bzw. Zahl fielen ( $a>b$ ,  $a+b=2n$ ). Alle diese Pfade, gehen durch  $P'(1;1)$ . Von  $P'(1;1)$  nach  $Q(2n;a-b)$  gibt es insgesamt  $v = \binom{2n-1}{a-1}$  Pfade. Die Anzahl der ungünstigen Pfade von  $P'$  nach  $Q$ , die die  $n$ -Achse berühren oder schneiden, gleicht nach dem Spiegelungsprinzip der Anzahl aller Pfade von  $P''(1;-1)$  nach  $Q$ , ist also  $u = \binom{2n-1}{a}$ . Daher gibt es von  $P(0;0)$  nach  $Q(2n,a-b)$  genau  $v-u = \binom{2n-1}{a-1} - \binom{2n-1}{a}$  Pfade, die abgesehen vom Ursprung über der  $n$ -Achse verlaufen.

Läßt man nun  $a$  die Zahlen  $n+1, \dots, 2n$  durchlaufen, so heben sich in der Summe die negativen Teile mit den positiven des folgenden Summanden auf und die gesuchte Anzahl aller Pfade über der  $n$ -Achse ergibt sich zu

$$\binom{2n-1}{n} = \frac{1}{2} \binom{2n}{n}.$$

Die Anzahl aller Pfade, die die  $n$ -Achse meiden, ist doppelt so groß, und man erhält für die Wahrscheinlichkeit, daß in den ersten  $n$  Doppelwürfen kein Ausgleich stattfindet

$$\begin{aligned} t_{2n} &= \binom{2n}{n} 2^{-2n}. \\ &= 1/2 \cdot 3/4 \cdot 5/6 \dots (2n-1)/(2n) \\ &= r(0) \cdot r(2) \cdot r(4) \dots r(2(n-1)). \end{aligned}$$

Dieses Ergebnis zeigt: Die Wahrscheinlichkeit, daß der letzte Ausgleich zwischen Kopf und Zahl  $2n$  Zeiteinheiten zurückliegt, ist genauso groß wie die Wahrscheinlichkeit, daß sich die Irrfahrt aus Abb. 8-2 (mit  $d=1$ ,  $e=2$ ) nach  $n$  Schritten auf der Stelle  $2n$  befindet. Abb. 8-2 beschreibt also tatsächlich die beim Münzwurf seit dem letzten Ausgleich verstrichene Zeit.

Wir tragen nun den Beweis von Formel (3) aus 8.3 nach: Wenn bei insgesamt  $n$  Doppelwürfen der letzte Ausgleich vor  $i$  Doppelwürfen stattfand, dann stand das Spiel nach  $n-i$  Doppelwürfen ausgeglichen. Die Wahrscheinlichkeit dafür ist  $\binom{2(n-i)}{n-i} 2^{-2(n-i)}$ . Bei den folgenden  $i$  Doppelwürfen fand kein Ausgleich statt. Das passiert mit der Wahrscheinlichkeit  $t_{2i} = \binom{2i}{i} 2^{-2i}$ . Die Wahrscheinlichkeit, daß sich die Irrfahrt aus Abb. 8-2 nach  $2n$  Schritten auf der Position  $x_i=2i$  befindet, ergibt sich als Produkt:

$$P_n(2i) = \binom{2i}{i} \binom{2(n-i)}{n-i} 2^{-2n}.$$

## 8.5.2 Resümee

Wenn wir mit einer Münze  $n$  Doppelwürfe machen und mit der Zufallsgröße  $x$  die seit dem letzten Ausgleich verstrichene Zeit zählen ( $x$  kann die Werte  $0, 2, 4, \dots, 2n$ ) annehmen, dann läßt sich das Verhalten von  $x$  beschreiben als Irrfahrt:

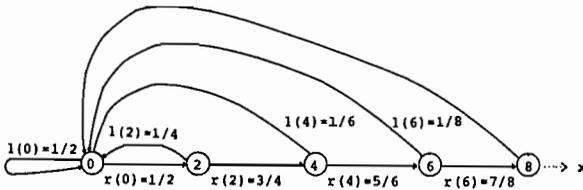


Abb. 8-9 Verstrichene Zeit als Irrfahrt eines Teilchens

Die normierte Zufallsgröße  $y := x/2n$  besitzt für großes  $n$  die Arcsin-Dichte

$$f(y) = \frac{1}{\pi \sqrt{y(1-y)}} \quad \text{und es gilt}$$

$$P\left(a < \frac{x}{2n} < b\right) \approx \frac{2}{\pi} (\arcsin(\sqrt{b}) - \arcsin(\sqrt{a})).$$

Am wahrscheinlichsten ist, daß der letzte Ausgleich vor sehr langer ( $x \approx 2n$ ) oder sehr kurzer Zeit ( $x \approx 0$ ) stattgefunden hat.

Historische Notiz:

Die Thematik dieses Kapitels ist noch relativ "neu". Die Tatsache, daß beim Münzwurf die seit dem letzten Ausgleich zwischen Kopf und Zahl verstrichene Zeit  $x$  eine symmetrische Verteilung besitzt, wurde empirisch durch Computereinsatz gefunden und von Blackwell/Deuel/Freedman (1964) ohne Kenntnis der exakten Verteilung bewiesen. Die Formel (3) wurde von Feller entdeckt und erstmals 1967 veröffentlicht (Feller Bd. I S. 78. und S. 82 ).

Abschließend belegen wir die in Abschnitt 8.1 erwähnten Ergebnisse, die sich mit unserer Intuition nur schwer vereinbaren lassen.

Erstes Beispiel:

Wir werfen eine Münze 10000mal ( $n=5000$  Doppelwürfe). Wenn am Ende der letzte Führungswechsel vor 9700 oder mehr Würfeln stattgefunden hat, steht man mit der Irrfahrt bei  $x \geq 9700$ . Die Wahrscheinlichkeit dafür ist mit

$$P(0.97 \leq y = x/(2n) \leq 1) \approx \frac{2}{\pi} \arcsin(\sqrt{1}) - \frac{2}{\pi} \arcsin(\sqrt{0.97}) \approx 0.1108$$

erstaunlich groß. Ebenso groß ist (wegen der nicht weniger überraschenden Symmetrie der Verteilung) die Wahrscheinlichkeit, daß der letzte Führungswechsel nicht mehr als 300 Zeiteinheiten zurückliegt.

Zweites Beispiel:

Die Wahrscheinlichkeit, daß von zwei gleichguten Schülern einer während der letzten 95 von 100 Stunden beständig vor dem anderen lag, beträgt (mit  $n=50$ ,  $x=95$ ,  $y=x/(2n) \geq 0.95$ )

$$P(0.95 \leq y \leq 1) \approx 1 - \frac{2}{\pi} \arcsin(\sqrt{0.95}) \approx 0.1435$$

mehr als 1 von 7. Die Begründung: "Steffi lag bis auf die ersten 5 Stunden des Halbjahres beständig vor Nicole, also erhält sie die bessere Note" ist immerhin in einem von sieben Fällen ungerechtfertigt. Wer möchte nach Kenntnis des Arcsin-Gesetzes noch guten Gewissens behaupten, unsere Noten seien "gerecht"?

## 9 Mathematischer Hintergrund

Zum Abschluß von Teil III wollen wir Beziehungen der vorangegangenen Kapitel zur Theorie der Diffusionsprozesse herstellen und skizzieren, wie sich die Ergebnisse der heuristischen Überlegungen mathematisch absichern lassen.

### 9.1 Brownsche Bewegung und Zentraler Grenzwertsatz

In den Zwanziger Jahren des letzten Jahrhunderts beobachtete der Englische Botaniker Robert Brown unter dem Mikroskop die nach ihm benannte unregelmäßige Bewegung kleinster in Flüssigkeiten suspendierter Teilchen. Nach Experimenten mit unterschiedlichen Materialien glaubte er aktive Moleküle gefunden zu haben. Erst Ende des letzten Jahrhunderts entdeckte man den wahren Grund der Bewegung: eine hohe Zahl von Zusammenstößen der suspendierten Teilchen mit den thermisch bewegten Flüssigkeitsmolekülen. In den Jahren 1905/1906 veröffentlichte Albert Einstein erstmals stochastisch fundierte Untersuchungen zur Brownschen Bewegung. Er leitete die Diffusionsgleichung  $\partial f / \partial t = D \cdot \partial^2 f / \partial x^2$  für die zeitabhängigen Übergangswahrscheinlichkeiten ab, die die Bewegung der Teilchen beschreiben. Diese Differentialgleichung wird durch die Normalverteilung (mit zeitlich veränderlicher Varianz) gelöst. Chung (1978, S. 265) spricht von einer "Ableitung des Zentralen Grenzwertsatzes mit Hilfe von Differentialgleichungen". (Aus Physikalischer Sicht vielleicht noch wichtiger als diese Herleitung ist Einsteins Formel für die Diffusionskonstante:  $D = (2RT) / (Nf)$ , wobei R die universelle Gaskonstante, T die absolute Temperatur, N die Avogadro-Konstante und f der Reibungskoeffizient ist. Diese Beziehung erlaubte später die Bestimmung der Avogadro-Konstante aus Beobachtungen zur Brownschen Bewegung, wofür Perrin 1926 den Nobelpreis erhielt. (Kac 1949 S. 370)). Hier der für uns relevante Abschnitt aus Einsteins Arbeit von 1905:

§ 4. Über die ungeordnete Bewegung von in einer Flüssigkeit suspendierten Teilchen und deren Beziehung zur Diffusion.

Wir gehen nun dazu über, die ungeordneten Bewegungen genauer zu untersuchen, welche, von der Molekularbewegung der Wärme hervorgerufen, Anlaß zu der im letzten Paragraphen untersuchten Diffusion geben.

Es muß offenbar angenommen werden, daß jedes einzelne Teilchen eine Bewegung ausführe, welche unabhängig ist von der Bewegung aller anderen Teilchen; es werden auch die Bewegungen eines und desselben Teilchens in verschiedenen Zeitintervallen als voneinander unabhängige Vorgänge aufzufassen sein, solange wir diese Zeitintervalle nicht zu klein gewählt denken.

Wir führen ein Zeitintervall  $\tau$  in die Betrachtung ein, welches sehr klein sei gegen die beobachtbaren Zeitintervalle, aber doch so groß, daß die in zwei aufeinanderfolgenden Zeitintervallen  $\tau$  von einem Teilchen ausgeführten Bewegungen als voneinander unabhängige Ereignisse aufzufassen sind.

Seien nun in einer Flüssigkeit im ganzen  $n$  suspendierte Teilchen vorhanden. In einem Zeitintervall  $\tau$  werden sich die  $X$ -Koordinaten der einzelnen Teilchen um  $\Delta$  vergrößern, wobei  $\Delta$  für jedes Teilchen einen anderen (positiven oder negativen) Wert hat. Es wird für  $\Delta$  ein gewisses Häufigkeitsgesetz gelten; die Anzahl  $dn$  der Teilchen, welche in dem Zeitintervall  $\tau$  eine Verschiebung erfahren, welche zwischen  $\Delta$  und  $\Delta + d\Delta$  liegt, wird durch eine Gleichung von der Form

$$dn = n \varphi(\Delta) d\Delta$$

ausdrückbar sein, wobei

$$\int_{-\infty}^{+\infty} \varphi(\Delta) d\Delta = 1$$

und  $\varphi$  nur für sehr kleine Werte von  $\Delta$  von Null verschieden ist und die Bedingung

$$\varphi(\Delta) = \varphi(-\Delta)$$

erfüllt.

Wir untersuchen nun, wie der Diffusionskoeffizient von  $\varphi$  abhängt, wobei wir uns wieder auf den Fall beschränken, daß die Anzahl  $\nu$  der Teilchen pro Volumeneinheit nur von  $x$  und  $t$  abhängt.

Es sei  $\nu = f(x, t)$  die Anzahl der Teilchen pro Volumeneinheit, wir berechnen die Verteilung der Teilchen zur Zeit  $t + \tau$  aus deren Verteilung zur Zeit  $t$ . Aus der Definition der Funktion  $\varphi(\Delta)$  ergibt sich leicht die Anzahl der Teilchen,

welche sich zur Zeit  $t + \tau$  zwischen zwei zur  $X$ -Achse senkrechten Ebenen mit den Abszissen  $x$  und  $x + dx$  befinden. Man erhält:

$$f(x, t + \tau) dx = dx \cdot \int_{\Delta = -\infty}^{\Delta = +\infty} f(x + \Delta) \varphi(\Delta) d\Delta.$$

Nun können wir aber, da  $\tau$  sehr klein ist, setzen:

$$f(x, t + \tau) = f(x, t) + \tau \frac{\partial f}{\partial t}$$

Ferner entwickeln wir  $f(x + \Delta, t)$  nach Potenzen von  $\Delta$ :

$$f(x + \Delta, t) = f(x, t) + \Delta \frac{\partial f(x, t)}{\partial x} + \frac{\Delta^2}{2!} \frac{\partial^2 f(x, t)}{\partial x^2} \dots \text{in inf.}$$

Diese Entwicklung können wir unter dem Integral vornehmen, da zu letzterem nur sehr kleine Werte von  $\Delta$  etwas beitragen. Wir erhalten:

$$\begin{aligned} f + \frac{\partial f}{\partial t} \cdot \tau &= f \cdot \int_{-\infty}^{+\infty} \varphi(\Delta) d\Delta + \frac{\partial f}{\partial x} \int_{-\infty}^{+\infty} \Delta \varphi(\Delta) d\Delta \\ &\quad + \frac{\partial^2 f}{\partial x^2} \int_{-\infty}^{+\infty} \frac{\Delta^2}{2} \varphi(\Delta) d\Delta \dots \end{aligned}$$

Auf der rechten Seite verschwindet wegen  $\varphi(x) = \varphi(-x)$  das zweite, vierte etc. Glied, während von dem ersten, dritten, fünften etc. Gliede jedes folgende gegen das vorhergehende sehr klein ist. Wir erhalten aus dieser Gleichung, indem wir berücksichtigen, daß

$$\int_{-\infty}^{+\infty} \varphi(\Delta) d\Delta = 1,$$

und indem wir

$$\frac{1}{\tau} \int_{-\infty}^{+\infty} \frac{\Delta^2}{2} \varphi(\Delta) d\Delta = D$$

setzen und nur das erste und dritte Glied der rechten Seite berücksichtigen:

$$(1) \quad \frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2}.$$

Dies ist die bekannte Differentialgleichung der Diffusion, und man erkennt, daß  $D$  der Diffusionskoeffizient ist.

Folgende Beziehungen bestehen zwischen Einsteins Ausführungen und Kapitel 6: Während Einstein mit kontinuierlicher Zeit  $t$  arbeitet und damit zeitliche Änderungen über die Ableitung  $\partial f/\partial t$  beschreiben kann, nimmt bei uns der Zeitparameter nur ganzzahlige Werte an. Dafür normieren wir die Verteilungen mit  $\frac{1}{\sigma\sqrt{n}}$ . Unterlassen wir die Normierung und nehmen wir in Kapitel 6.2 Gleichung (3) nur die Entwicklung des Terms  $A$  vor, so erhalten wir

$$f_{n+1}(x) \approx f_n(x) + 0.5 \sigma^2 f_n''(x) \text{ bzw.}$$

$$\frac{f_{n+1}(x) - f_n(x)}{n+1 - n} \approx 0.5 \sigma^2 f_n''(x).$$

Deuten wir die linke Seite als Näherung eines zeitlichen Differentials, so erhalten wir Einsteins Ergebnis:

$$\partial f/\partial t = D \partial^2 f/\partial x^2 \text{ mit } D=0.5 \sigma^2. \text{ Umgekehrt läßt sich}$$

Einsteins Differentialgleichung in unsere überführen: Wir nehmen - wie in Kapitel 6 implizit geschehen - an, daß die Lösungen für verschiedene Zeiten durch Normierung auseinander hervorgehen:  $f(t,x) = (1/\sqrt{t})g(x/\sqrt{t})$ . Dann erhalten wir mit

$$\frac{\partial f}{\partial t} = \frac{-1}{2\sqrt{t}^3} g\left(\frac{x}{\sqrt{t}}\right) - \frac{1}{2\sqrt{t}^3} \frac{x}{\sqrt{t}} g'\left(\frac{x}{\sqrt{t}}\right) \quad \text{und}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{\sqrt{t}^3} g''\left(\frac{x}{\sqrt{t}}\right)$$

aus  $\partial f/\partial t = D \partial^2 f/\partial x^2$  (für  $\sigma=1$ , also  $D=1/2$  und  $z=x/\sqrt{t}$ ) die gewöhnliche Differentialgleichung der Normalverteilung

$$g(z) + zg'(z) + g''(z) = 0,$$

die wir in Kapitel 6.2 (Formel 6) hergeleitet haben.

Diese heuristischen Überlegungen werden durch eine Arbeit Khintchines (1933) auf ein festes Fundament gestellt. Khintchine beweist dort den Zentralen Grenzwertsatz, indem er die Folge von Verteilungsfunktionen summierter Zufalls-

größen der zeitlichen Entwicklung einer Lösung von Einsteins Diffusionsgleichung gegenüberstellt. Diesen Beweis kann man in leicht "modernisierten" Fassungen auch bei Ito (1965, S. 10) und Schuss (1980, S. 48) nachlesen. Das Faszinierende an dem Beweis ist seine Übertragbarkeit auf beliebige Diffusionsprozesse. Wir werden Khintchines Methode nutzen, um die in Kapitel 7 beobachtete Konvergenz diskreter Irrfahrten gegen Chi-Quadrat-Verteilungen zu beweisen.

## 9.2 Diffusionsprozesse und Chi-Quadrat-Verteilung

### 9.2.1 Kolmogoroffs Differentialgleichungen

Wir betrachten  $n$  Schritte der Irrfahrt aus Kapitel 7.1, bei der der Erwartungswert des Zuwachses in jedem Schritt konstant ( $k$ ) und die Varianz ( $4u$ ) proportional zum momentanen Standort  $u$  ist (Abb. 9-1).

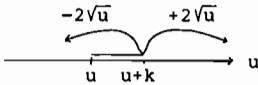


Abb. 9-1

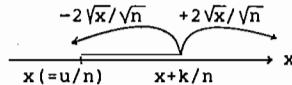


Abb. 9-2

Die Zuwächse der zugehörigen (mit  $1/n$ ) normierten Irrfahrt (Abb. 9-2) haben dann den Erwartungswert  $k/n$  und die Varianz  $4x/n$ . Wir stellen uns vor, daß das Teilchen zu den Zeitpunkten  $t_i = i/n$  ( $0 \leq i < n$ ) springt.  $F(x,y)$  bezeichne die Verteilungsfunktion, die zu einem Sprung von der Stelle  $x$  aus gehört (Atome vom Gewicht 0.5 bei  $x+k/n \pm 2\sqrt{x}/\sqrt{n}$ ). Es gilt

$$\int (y-x) F(x,dy) = k/n \quad (= a(x)/n + o(1/n)) \quad (1)$$

$$\int (y-x)^2 F(x,dy) = 4x/n + (k/n)^2 \quad (= b(x)/n + o(1/n)) \quad (2)$$

und die Lindebergbedingung

$$\int_{|y-x| > \mathcal{T}} (y-x)^2 F(x,dy) = o(1/n) \quad \text{für jedes } \mathcal{T} > 0 \quad (3)$$

(Für hinreichend große  $n$  verschwindet das letzte Integral gleichmäßig über kompakten Intervallen  $x \in [0; A]$ ). Die Bedingungen (1) bis (3) besagen, daß wir es mit dem diskreten Modell eines Diffusionsprozesses zu tun haben, bei dem die "Drift"  $a(x)=k$  konstant und die "Diffusion"  $b(x)=4x$  proportional zum Standort ist. Bei einem "echten" Diffusionsprozeß variiert die Zeit  $t$  kontinuierlich und nicht in Sprüngen der Größe  $1/n$ . Im Jahre 1931, also kurz vor Khintchines Arbeit leitete Kolmogoroff für die Übergangswahrscheinlichkeiten von Diffusionsprozessen zwei Differentialgleichungen ab. (Sie umfassen Einsteins Ergebnis als Spezialfall). Wenn  $f(t,x,y)$  die Wahrscheinlichkeitsdichte für einen Übergang von  $x$  nach  $y$  während der Zeit  $t$  bezeichnet, dann gilt unter gewissen Bedingungen die Rückwärtsdifferentialgleichung (man differenziert bei festem Ziel  $y$  nach der Startposition  $x$ ):

$$\frac{\partial}{\partial t} f(t,x,y) = \frac{1}{2} b(x) \frac{\partial^2}{\partial x^2} f(t,x,y) + a(x) \frac{\partial}{\partial x} f(t,x,y) \quad (4)$$

und die Vorwärtsdifferentialgleichung (man differenziert bei festem Start  $x$  nach dem Ziel  $y$ ):

$$\frac{\partial}{\partial t} f(t,x,y) = \frac{1}{2} \frac{\partial^2}{\partial y^2} (b(y)f(t,x,y)) - \frac{\partial}{\partial y} (a(y)f(t,x,y)) \quad (5)$$

Für  $a(x)=k$  und  $b(x)=4x$  erhalten wir

$$\frac{\partial}{\partial t} f(t,x) = 2x \frac{\partial^2}{\partial x^2} f(t,x) + k \frac{\partial}{\partial x} f(t,x) \quad (4a)$$

$$\frac{\partial}{\partial t} f(t,y) = 2y \frac{\partial^2}{\partial y^2} f(t,y) + (4-k) \frac{\partial}{\partial y} f(t,y) \quad (5a)$$

### 9.2.2 Differentialgleichung der Chi-Quadrat-Verteilung:

Wir haben uns in Kapitel 7 mit der unter stochastischen Gesichtspunkten naheliegenderen Vorwärts-Fragestellung beschäftigt. Wie dort implizit geschehen, machen wir die Annahme, daß die Übergangsdichten für verschiedene Zeiten durch Skalierung auseinander hervorgehen:

$$f(t, y) = \frac{1}{t} g\left(\frac{y}{t}\right). \quad (6)$$

Wegen  $\frac{\partial}{\partial y} f(t, y) = \frac{1}{t^2} g'\left(\frac{y}{t}\right),$

$$\frac{\partial^2}{\partial y^2} f(t, y) = \frac{1}{t^3} g''\left(\frac{y}{t}\right) \quad \text{und}$$

$$\frac{\partial}{\partial t} f(t, y) = \frac{-1}{t^2} g\left(\frac{y}{t}\right) - \frac{y}{t^3} g'\left(\frac{y}{t}\right)$$

ergibt sich mit  $z:=y/t$  aus Kolmogoroffs Vorwärtsgleichung (5a) die Differentialgleichung der Chi-Quadrat-Verteilung mit  $k$  Freiheitsgraden

$$g(z) + (4-k+z)g'(z) + 2zg''(z) = 0, \quad (7)$$

die wir in 7.2 hergeleitet haben.

### 9.2.3 Lösung von Kolmogoroffs Differentialgleichungen

Für den angekündigten Konvergenzbeweis benötigen wir die Lösungen von Kolmogoroffs Gleichungen (4a,5a). Wir erhalten sie durch eine Transformation aus der Lösung des Bessel-Prozesses. Dieser Diffusionsprozeß hat eine Drift, die mit  $a(x)=(k-1)/2x$  zum Standort  $x$  umgekehrt proportional ist und die konstante Diffusion  $b(x)=1$ . Die zugehörige Rückwärtsdifferentialgleichung

$$\frac{\partial}{\partial t} p(t, x, y) = \frac{1}{2} \frac{\partial^2}{\partial x^2} p(t, x, y) + \frac{k-1}{2x} \frac{\partial}{\partial x} p(t, x, y) \quad (8)$$

wird gelöst durch

$$p(t, x, y) = \exp\left(-\frac{x^2 + y^2}{2t}\right) \frac{y^{k-1}}{t^{(k-1)/2}} I_{k/2-1}\left(\frac{xy}{t}\right), \quad (9)$$

wobei  $I_d(x) = \sum_{i=0}^{\infty} \frac{1}{i! \Gamma(d+i+1)} \left(\frac{x}{2}\right)^{2i+d}$

die modifizierte Besselfunktion der Ordnung  $d$  ist.

keit der partiellen Ableitungen bis zur ersten Ordnung in  $t$  und bis zur zweiten Ordnung in  $x$  über beschränkten Rechtecken  $0 \leq t \leq 1$  und  $0 \leq x \leq A$ ,  $A$  beliebig <sup>1)</sup>. Eine entsprechende Konstruktion führen wir für unsere diskrete Irrfahrt durch:  $F(x, y)$  bezeichne (wie in 9.2.1 erläutert) die Verteilungsfunktion, die zur Stelle  $x$  gehört. Sie hat die Gewichte  $0.5$  an den Stellen  $x+k/n \pm 2\sqrt{x}/\sqrt{n}$ .  $v(t_i, x)$  bezeichne die Erwartung von  $h$  zur Zeit  $t_i = i/n$  unter der Bedingung, daß man sich zur Zeit  $0$  bei  $x$  befindet. Es gilt

$$v(t_{i+1}, x) = \int v(t_i, y) F(x, dy), \quad 0 \leq i < n \quad \text{und} \quad v(0, x) = h(x).$$

Sei nun  $\varepsilon > 0$  fest gewählt. Wir betrachten neben  $u$  die Funktion  $u_\varepsilon(t, x) := u(t, x) + \varepsilon t$ ,

1) Bezeichnet  $f_k(t, x, y)$  die Übergangsdichte (9b), die zur Drift  $k$  gehört, dann liefert eine kurze Rechnung

$$\frac{\partial}{\partial x} u(t, x) = \frac{1}{2t} \int (f_{k+2}(t, x, y) - f_k(t, x, y)) h(y) dy \quad \text{und}$$

$$\frac{\partial^2}{\partial x^2} u(t, x) = \frac{1}{4t^2} \int (f_{k+4}(t, x, y) - 2f_{k+2}(t, x, y) + f_k(t, x, y)) h(y) dy.$$

Hieraus folgt die Stetigkeit der Ableitungen nach  $x$  für  $t > 0$ . Zum Beweis der Stetigkeit in  $t=0$  benutzt man die Momente

$$\int (y-x) f_k(t, x, y) dy = kt$$

$$\int (y-x)^2 f_k(t, x, y) dy = 4xt + (2k+k^2)t^2,$$

$$\int (y-x)^3 f_k(t, x, y) dy = (24x+12kx)t^2 + (8k+6k^2+k^3)t^3$$

(Johnson/Kotz S. 134). Man entwickelt die Testfunktion  $h$  um  $x$  und rechnet nach, daß  $\partial u(t, x) / \partial x \rightarrow h'(x)$ ,  $\partial^2 u(t, x) / \partial x^2 \rightarrow h''(x)$  für  $t \rightarrow 0$ . Die Stetigkeit der partiellen Ableitung von  $u$  nach  $t$  ergibt sich dann aus (11). Beschränktheit und gleichmäßige Stetigkeit der Ableitungen folgen aus der Kompaktheit des Rechtecks  $0 \leq t \leq 1$ ,  $0 \leq x \leq A$ .

die die Differentialgleichung

$$\frac{\partial}{\partial t} u_\varepsilon = 2x \frac{\partial^2}{\partial x^2} u_\varepsilon + k \frac{\partial}{\partial x} u_\varepsilon + \varepsilon \quad (12)$$

löst und zeigen, daß für alle  $x$  aus einem beliebigen beschränkten Intervall  $[0;A]$  bei genügend hohem  $n$  gilt

$$u_\varepsilon(t + \frac{1}{n}, x) > \int u_\varepsilon(t, y) F(x, dy). \quad (13)$$

**Beweis:**

Einerseits gilt wegen der gleichmäßigen Stetigkeit von  $\partial u / \partial t$

$$u_\varepsilon(t + \frac{1}{n}, x) = u_\varepsilon(t, x) + \frac{1}{n} \frac{\partial}{\partial t} u_\varepsilon + o(\frac{1}{n}). \quad (14)$$

Andererseits hat man

$$u_\varepsilon(t, y) = u_\varepsilon(t, x) + (y-x) \frac{\partial}{\partial x} u_\varepsilon + \frac{1}{2} (y-x)^2 \frac{\partial^2}{\partial x^2} u_\varepsilon + \frac{1}{2} (y-x)^2 r(t, x, y)$$

$$\text{mit } r(t, x, y) = \frac{\partial^2}{\partial x^2} u_\varepsilon \Big|_{t, x+\vartheta(y-x)} - \frac{\partial^2}{\partial x^2} u_\varepsilon, \quad 0 < \vartheta < 1,$$

Dabei sind die nicht explizit hingeschriebenen Argumente von  $u_\varepsilon$  stets  $t$  und  $x$ . Durch Integration folgt mit (1) und (2)

$$\int u_\varepsilon(t, y) F(x, dy) = u_\varepsilon(t, x) + \frac{2x}{n} \frac{\partial^2}{\partial x^2} u_\varepsilon + \frac{k}{n} \frac{\partial}{\partial x} u_\varepsilon + J + o(\frac{1}{n}) \quad (15a)$$

$$\text{mit } J = \frac{1}{2} \int (y-x)^2 r(t, x, y) F(x, dy) = o(1/n). \quad (15b)$$

Die für  $0 \leq t \leq 1$ ,  $0 \leq x \leq A$  gleichmäßige Konvergenz  $J = o(1/n)$  in (15b) ergibt sich wie folgt: Man wählt  $\mathcal{T}$  beliebig klein und unterteilt das Integrationsintervall: Für  $|y-x| < \mathcal{T}$  nutzt man die Kleinheit von  $|y-x|$  und die gleichmäßige Stetigkeit von  $\partial^2 u_\varepsilon / \partial x^2$ , für  $|y-x| > \mathcal{T}$  folgt die Konvergenz aus der Lindbergbedingung (3) und der Beschränktheit von  $\partial^2 u_\varepsilon / \partial x^2$ . Da nämlich  $F(x, -)$  Atome nur in  $x+k/n \leq 2\sqrt{x}/\sqrt{n}$  besitzt, liegen mit  $0 \leq x \leq A$  auch die bei der Integration in (15b) auftretenden Argumente  $x+\vartheta(y-x)$  in einem kompakten Intervall.

Zusammen mit (12) und (14) folgt aus (15a,b)

$$u_{\xi}(t+\frac{1}{n}, x) = \int u_{\xi}(t, y) F(x, dy) + \xi \frac{1}{n} + o(\frac{1}{n}),$$

woraus sich für hinreichend hohes  $n$  die behauptete Ungleichung (13) ergibt. Setzt man nun

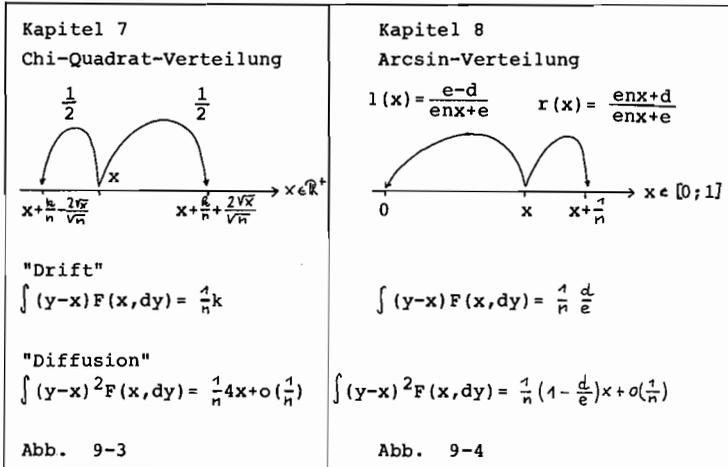
$$w(t_i, x) := u_{\xi}(t_i, x) - v(t_i, x),$$

so gilt  $w(0, x) = 0$  und wegen (13) für hinreichend großes  $n$  und alle  $x \in [0; A]$ :  $w(t_{i+1}, x) > \int w(t_i, y) F(x, dy)$ . Induktiv folgt hieraus  $w(1, x) > 0$ , also  $u(1, x) + \xi > v(1, x)$  für alle  $x \in [0; A]$ . Aus einer entsprechenden Abschätzung nach unten ergibt sich für jedes feste  $x$ :  $\lim_{n \rightarrow \infty} v(1, x) = u(1, x)$ . Aus der Konvergenz der Erwartungen für beliebige Testfunktionen folgt wie erwähnt die schwache Konvergenz der zugrundeliegenden Wahrscheinlichkeitsmaße. Insbesondere konvergieren für  $x=0$  die normierten Verteilungen unserer diskreten Irrfahrt gegen die Chi-Quadrat-Verteilung mit  $k$  Freiheitsgraden.

In 7.3 haben wir erläutert, daß auch während des Anpassungstests die Folge der Testgrößen als eindimensionale Irrfahrt mit konstanter Drift  $k-1$  und Diffusion  $4x$  gedeutet werden kann. Die Konvergenz der Testgröße gegen die Chi-Quadrat-Verteilung mit  $k-1$  Freiheitsgraden wird dadurch sehr plausibel. Es ist jedoch nicht möglich, den oben geführten Beweis zu einem exakten Konvergenzbeweis für den Anpassungstest auszubauen. Die zu einem Testverlauf gehörige Irrfahrt ist nämlich nicht markoffsch. Die Verteilung, die den "Sprung" von der Stelle  $x$  steuert, hängt nicht nur von dem momentanen Wert  $x$  der Testgröße ab, sondern von den Besetzungszahlen, also der "Geschichte" des bisherigen Testverlaufs. (Das Symbol  $F(x, -)$  macht keinen Sinn mehr.) Um die Markoff-Eigenschaft zurückzuerhalten, müßte man einen mehrdimensionalen Diffusionsprozess untersuchen. Der Konvergenzbeweis liefe auf einen Beweis des Zentralen Grenzwertsatzes in mehreren Dimensionen (Khintchine S. 11) hinaus - und damit auf die in 3.5 erläuterte Idee Pearsons.

## 9.3 Arcsin-Gesetz und Renewal-Theorie

Wir vergleichen die Irrfahrten der Kapitel 7 und 8 (Abb. 7-4a und Abb. 8-2). Nach Normierung mit  $1/n$  bzw.  $1/(en)$  sehen sie wie folgt aus:



Beide Irrfahrten haben eine "Drift", die mit  $a(x)=k$  bzw.  $a(x)=d/e$  konstant und eine "Diffusion", die mit  $b(x)=4x$  bzw.  $b(x)=(1-d/e)x$  zum Standort  $x$  proportional ist. Die Ausführungen der vorigen Abschnitte besagen, daß die Grenzverteilungen nur von "Drift" und "Diffusion" abhängen. (Nur diese Parameter gehen in Kolmogoroffs Differentialgleichungen ein.) Daher könnte man für beide Irrfahrten gleichartige Grenzverteilungen erwarten. Tatsächlich handelt es sich in Abb. 9-4 aber nicht mehr um das diskrete Modell eines Diffusionsprozesses. Wegen der möglicherweise sehr großen Sprungweiten (von  $x$  direkt in den Ursprung) ist die Lindebergbedingung (Kap. 9.2.1 Formel (3)) nicht mehr erfüllt. Khintchines Beweis greift nicht mehr und das Auftreten völlig andersartiger Grenzverteilungen (Arcsin-Verteilung statt Chi-Quadrat-Verteilung) wird trotz gleichartiger Parameter verständlich.

Um die in Kapitel 8 heuristisch begründete Konvergenz gegen die verallgemeinerte Arcsin-Verteilung abzusichern, ziehen wir einen (recht tief liegenden) Satz der Renewal-Theorie heran (Feller II S. 472): Die Fluchtirrfahrt (in Abb. 8-2 ist sie im Gegensatz zu Abb. 9-4 nicht normiert) beginnt nach Rückkehr in den Ursprung immer wieder von vorne, sie "erneuert" sich. Mit  $u_n$  bezeichnen wir die Wahrscheinlichkeit, daß bis zur Zeit  $n$  noch keine Erneuerung stattfand.  $Y_n$  sei die seit der letzten Erneuerung verstrichene Zeit. Der erwähnte Satz besagt: Unter der Voraussetzung, daß  $u_n$  eine Entwicklung

$$u_n = n^{-a} L(n) \text{ mit } \forall k > 1 \ L(kn)/L(n) \rightarrow 1 \text{ für } n \rightarrow \infty \quad (1)$$

besitzt, konvergiert die Verteilung der normierten Wartezeit  $Y_n/n$  bei wachsendem  $n$  gegen die Arcsin-Verteilung mit Dichte

$$f(x) = \frac{\sin(\alpha \pi)}{\pi x^\alpha (1-x)^{1-\alpha}}$$

Dabei besagt die in (1) genannte Eigenschaft, daß die Funktion  $L$  regulär variiert ("regularly varying tail"). Wir beweisen nun, daß unsere Fluchtirrfahrt (Abb. 8-2) die Voraussetzung (1) erfüllt. Dabei beschränken wir uns der Anschaulichkeit halber auf den in Tabelle 8-6 dargestellten Spezialfall mit Sprungweite  $e=6$  und "Drift"  $d=4$ . (Diese Irrfahrt ist nach Normierung identisch mit der Irrfahrt für  $e=1$  und  $d=2/3$ , bei der die Position mit der seit der letzten Erneuerung verstrichenen Zeit zusammenfällt.) Aus der Pfadregel folgt

$$u_n = \frac{2 \ 5 \ 8 \ \dots \ (3n-1)}{3 \ 6 \ 9 \ \dots \ 3n}$$

Daneben betrachten wir die Folgen

$$s_n = \frac{3 \ 6 \ 9 \ \dots \ 3n}{4 \ 7 \ 10 \ \dots \ (3n+1)} \quad t_n = \frac{4 \ 7 \ 10 \ \dots \ (3n+1)}{5 \ 8 \ 11 \ \dots \ (3n+2)}$$

Es gilt

$$u_n \cdot s_n \cdot t_n = \frac{2}{3n+2} = \frac{2}{3} \frac{1}{n} \frac{1}{\left(1 + \frac{2}{3n}\right)} \quad (2)$$

Die Quotienten  $q_n := s_n/u_n$  und  $q_n' := t_n/u_n$  sind monoton steigend und wegen  $u_n < s_n < t_n < (3/2)u_n$  beschränkt, also konvergent. Mit  $q_n'' := 1 + 2/(3n)$  folgt aus (2)  $u_n^3 (q_n \cdot q_n' \cdot q_n'') = \frac{2}{3n}$ , also

$$u_n = n^{-1/3} \sqrt[3]{\frac{2/3}{q_n q_n' q_n''}} = n^{-1/3} L(n). \quad (3)$$

Damit haben wir die Voraussetzung (1) (mit  $a=1/3$ ) bestätigt, wenn wir zeigen, daß

$$L(n) := \sqrt[3]{\frac{2/3}{q_n q_n' q_n''}}$$

regulär variiert. Diese Eigenschaft von  $L(n)$  ergibt sich aber aus der regulären Variation von  $q_n$ ,  $q_n'$  und  $q_n''$ . Z. B. läßt sie sich für  $q_n$  wie folgt nachweisen: Der Quotient

$$\frac{q_{kn}}{q_n} = \frac{9(n+1)^2}{9(n+1)^2-1} \cdot \frac{9(n+2)^2}{9(n+2)^2-1} \cdots \frac{9(kn)^2}{9(kn)^2-1}$$

ist für jedes feste  $k > 1$  ein "Reststück" des monoton steigenden und beschränkten Produktes  $q_n$ , muß also mit  $n \rightarrow \infty$  gegen 1 konvergieren. Analog argumentiert man für  $q_n'$ . Die reguläre Variation von  $q_n''$  ist unmittelbar einsichtig.

Damit konvergieren die Verteilungen der seit der letzten Erneuerung verstrichenen (normierten) Zeiten gegen eine verallgemeinerte Arcsin-Verteilung. Die gleiche Aussage gilt folglich auch für die normierten Positionen der in Kap. 8.2 untersuchten Irrfahrt.

## Teil IV Statistik und Datenbanken

## 10 Partnersuche

Mit einer gewissen Regelmäßigkeit finden sich in den Sommermonaten Partnertests in den Zeitschriften an unseren Kiosks. Aus naheliegenden Gründen stoßen solche Tests auch bei Schülern auf beträchtliches Interesse, so daß daraus ein Projektwochentema erwuchs. Ein Schwergewicht dieses Projekts lag im Bereich der Grundbildung Informatik (Programmierung einer Datenbank zur Erfassung der Testdaten, lineare Algorithmen mit Verzweigungen zur Erstellung von Personencharakterisierungen und Algorithmen mit Schleifen zum Durchsuchen von Dateien nach Minimalstellen ("optimalen Partnern"). Der andere Schwerpunkt lag in der Analyse der Datenbestände mit Wahrscheinlichkeitsrechnung und Statistik (Normalverteilung, Chi-Quadrat-Tests) im regulären Stochastikunterricht.

## 10.1 Testfragen

Der Test umfaßt 20 Fragen (Tab. 10.1), von denen sicher nicht alle ganz ernst zu nehmen sind. Sie stammen aus der Zeitschrift "Hör Zu" vom 4.7.86. Zusätzlich wurde nach der Wertschätzung für die Schulfächer Mathematik, Deutsch, Englisch, nach dem Sternzeichen, nach Name und Klasse gefragt.

## 10.2 "Erotischer Abstand"

Bei dem Test sollen nicht nur zwei Partner durch Vergleich ihrer Punktsommen (nach der Vorschrift von Tab. 10-2) den Grad ihrer Harmonie ermitteln können. Wir wollen den Partner in der Datei finden, der am besten "paßt". Hierzu ist die Punktsomme ein zu grobes Suchkriterium. Wir definieren den "erotischen Abstand" zwischen zwei Partnern als die Summe der (betraglichen) Punktdifferenzen bei den Testfragen 4,9, 10,11,14,16,20, den Wertschätzungen (von 1 bis 3) der Schul-



**1** Ein Sinnbild für Partnerschaft: zwei Uhren mit verschiedenen Zeiten, blühende Rosen. Was sagt das Foto über die Beziehung aus?

- a) Sie ist oberflächlich und materialistisch
- b) Sie ist romantisch und zärtlich
- c) Sie ist tolerant

**2** Was sagt dieses Bild mit den beiden Uhren über die Partner aus?

- a) Sie sind sich ähnlich
- b) Sie sind verschieden
- c) Sie passen nicht zusammen



**3** Welcher der folgenden Gedanken paßt Ihrer Meinung nach zu diesem Foto am besten?

- a) Die Sprache der Liebe kennt keine Grenzen
- b) Liebe ist (oft) nur ein Wort
- c) Wissen ist Macht

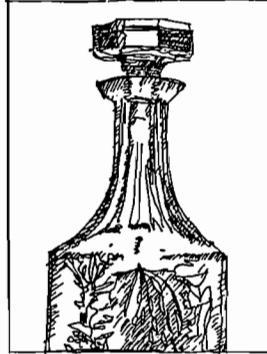
**4** Hätten Sie in der Schule ein Buch zum Mogen? Was halten Sie heute vom Schummeln?

- a) Dazu gehört auch etwas Geschicklichkeit
- b) Viel zu riskant!
- c) Das ist charakterlos!



**5** Welchen Begriff bringen Sie mit diesem Foto am ehesten in Verbindung?

- a) Unbefangenheit oder Verspiehtheit
- b) Zwiespältigkeit oder Unberechenbarkeit
- c) Trauer oder Verzweiflung



**6** Lassen Sie das Bild einen Moment auf sich wirken. Wie finden Sie das Kristallgefäß?

- a) Kitschig bis nichtsagend
- b) Strahlt irgendwie Wärme aus
- c) Wirkt recht harmonisch



**7** Legen Sie dem Kreuzbuben einen der folgenden Sätze in den Mund:

- a) „Ich mag keine Blumen“
- b) „Auch eine Spielkarte braucht mal einen freien Tag ...“
- c) „Ich sehne mich nach der Herzdame.“



**8** Versetzen Sie sich einmal in den Astronomen. Was mag er gerade empfinden?

- a) Angst, daß die Verbindungsschnur abreißen könnte
- b) Ärger über die Bevormundung durch die Bodenstation
- c) Stolz und Freude über seine Teilnahme an dieser Mission

**9** Ihre(r) Partner(in) flirtet auf einer Party mit einem attraktiven Gast. Wie fühlen Sie sich?

- a) Ich bin ziemlich ärgerlich und beobachte ihn (sie) heimlich
- b) Ich fühle mich verletzt und strafe ihn (sie) mit Nichtbeachtung
- c) Ich gönne ihm (Ihr) das kleine Vergnügen und flirte meinerseits

**10** Mal ehrlich: Können Sie von Ihrem Partner Kriuk verzeihen?

- a) Ich denke schon, solange die Kriuk sachlich ist
- b) Damit habe ich keine Probleme
- c) Ehrlich gesagt: nicht so gut

**11** Können Sie sich vorstellen, ein Kind zu adoptieren?

- a) Absolut undenkbar
- b) Unter bestimmten Bedingungen schon
- c) Auf jeden Fall

**12** Ihre(r) Partner(in) bringt Ihnen ein kleines Stofftier mit. Welche der folgenden Bemerkungen könnte von Ihnen stammen?

- a) „Ist das der Ersatz für Zärtlichkeit?“
- b) „Hat irgendwie Ähnlichkeit mit dir ...“
- c) „Willst du mich auf den Arm nehmen?“

**13** Ein gemütlicher Abend im Restaurant. Ihre(r) Partner(in) wirkt gedankenverloren. Was sagen Sie?

- a) „Du kannst mir ruhig erzählen, was dich beschäftigt.“
- b) „Ziemlich unbehöflich, dein Verhalten.“
- c) „Sag schon, wo du mit deinen Gedanken bist.“

**14** Welche der folgenden Eigenschaften finden Sie in der Partnerschaft am wichtigsten?

- a) Intelligenz b) Toleranz c) Zärtlichkeit

**15** Welcher Baum könnte Ihre Partnerschaft am treffendsten symbolisieren?

- a) Apfelbaum b) Birke c) Eiche

**16** Wenn Ehepartner getrennt in Urlaub fahren, dann ...

- a) ... ist ihre Beziehung kaputt
- b) ... sind sie verständnisvoll und aufgeschlossen
- c) ... brauchen sie etwas Abstand voneinander

**17** Der Seitensprung eines Ehepartners ist ...

- a) ... ein „Ausrutscher“, über den man reden kann
- b) ... unverzeihlich
- c) ... für den anderen ein Stöck ins Herz

**18** Womit könnte Ihre(r) Partner(in) Sie am meisten verletzen?

- a) Indem er (sie) mich ungerecht behandelt
- b) Indem er (sie) mich anläßt
- c) Eigentlich gar nicht, denn ich bin ein robuster Charakter

**19** Wie wirkt das Wort „Geborgenheit“ auf Sie?

- a) Klingt irgendwie alibacken
- b) Löst in mir positive Empfindungen aus
- c) Weckt bei mir eine gewisse Sehnsucht

**20** Man sagt: Schadenfreude ist die größte Freude. Auch bei Ihnen?

- a) Manchmal – wenn jemanden trifft, den ich nicht mag
- b) Na klar!
- c) Ich habe eher Mitleid mit dem Pechvogel

Tab. 10-1 Partnertest

Test-Auflösung			
Frage	Antwort		
	a	b	c
1	6	0	3
2	0	4	6
3	0	3	6
4	3	0	6
5	3	6	0
6	6	1	3
7	6	3	0
8	0	6	3
9	6	0	3
10	3	6	0
11	6	3	0
12	0	3	6
13	1	6	4
14	6	3	0
15	3	1	6
16	6	3	1
17	3	6	0
18	4	0	6
19	6	3	1
20	3	6	1

Gesamtpunktzahl SIE:

Gesamtpunktzahl ER:

Ziehen Sie nun das niedrigere Ergebnis vom höheren ab, und tragen Sie die Differenz hier ein.

Dieser Punktwert sagt Ihnen, in welcher der drei folgenden Auswertungsgruppen Sie die Beurteilung Ihrer Partnerschaft finden.

**Erste Auswertungsgruppe.**  
Weniger als 18 Punkte Unterschied: In Ihrer Partnerschaft hängt der Himmel voller Geigen. Sie lachen über dieselben Witze, begeistern sich für dieselben Bücher – das ist eine Wellenlänge. Trotzdem ist das noch kein Garantieschein für ungetrübtes Glück zu zweit. Falls nur einer von Ihnen mehr als 70 Punkte

hat, genügt schon eine versalzene Suppe oder eine Zigarettenkippe im Badezimmer, um es zum Krach kommen zu lassen. Ein Tip: Zeigen Sie etwas mehr Toleranz. Sollte einer von Ihnen auf unter 30 Punkte kommen, dann sind Sie eher konfliktstreu. Deshalb: mehr Mut zur Offenheit!

**Und so harmonieren Sie in der Liebe:** In Ihrer Partnerschaft spielt Erotik eine sehr große Rolle. Wer weniger als 30 Punkte hat, bei dem muß das erotische Gefühl manchmal erst geweckt werden. Kommt einer von Ihnen auf mindestens 70 Punkte, dann wirkt Ihre gegenseitige Anziehungskraft geradezu magisch.

**Zweite Auswertungsgruppe.**  
18 bis 36 Punkte Unterschied: Ihre Partnerschaft ist ein farbiges Wechselspiel unterschiedlicher Regungen: Sie ärgern sich über die Gleichgültigkeit des anderen, und Sie freuen sich, wenn er Ihnen mit einem unerwarteten Kompliment schmeichelt. Manchmal wünschen Sie sich vom anderen etwas mehr Einfühlungsvermögen – Sie machen ihm das auch unmißverständlich klar. Doch gerade das wechselnde Auf und Ab in Ihrer Beziehung reizt Sie besonders. Bei einer Differenz von über 30 Punkten ist allerdings Vorsicht geboten: Der Partner mit der höheren Gesamtpunktzahl neigt dazu, dem anderen seinen Willen aufzuzwingen. Allerdings nicht immer mit Erfolg. Denn wenn der Partner mit der niedrigeren Gesamtpunktzahl über 35 liegt, hat er trotzdem genug Selbstbewußtsein, um seine eigenen Vorstellungen gegenüber dem anderen durchzusetzen.

**Und so harmonieren Sie in der Liebe:** Erotische Abwechslung prägt Ihr Miteinander. Gewohnheit in der Liebe ist Ihnen beiden zuwider. Der Partner mit der höheren Gesamtpunktzahl ergreift gern die Initiative und animiert den anderen.

**Dritte Auswertungsgruppe.**  
Mehr als 36 Punkte Unterschied: Ein Leben ohne den anderen können Sie sich nur schwer vorstellen. Davon sind Sie beide überzeugt. Allerdings herrscht eine ganz klare Rollenverteilung: Der Partner mit der geringeren Gesamtpunktzahl ordnet sich unter. Er ist stets um „Burgfrieden“ bemüht. Der andere hingegen hält gern die Fäden in der Hand. Dieses harmonische Miteinander funktioniert bestens – solange sich beide mit den ungleichen Rollen zufriedengeben. Wer von Ihnen aber in seiner Gesamtpunktzahl unter 30 liegt, der hat als Partner nichts zu lachen. Denn die dominante „bessere Hälfte“ bestimmt letztlich sogar, ob der Kaffee mit oder ohne Zucker getrunken wird. Ein guter Rat an den unterdrückten Partner: Verleihen Sie Ihren Wünschen mehr Nachdruck, auch wenn Sie damit ein partnerschaftliches Donnerwetter provozieren.

**Und so harmonieren Sie in der Liebe:** Alles zu seiner Zeit – so ist Ihr „erotisches Uhrwerk“ eingestellt. Der Partner mit der höheren Punktzahl ergreift fast immer die Initiative. Das erzeugt oft Widerstand beim anderen, ruft sogar Aggressionen hervor und bewirkt genau das Gegenteil: Der „Herausgeforderte“ verzichtet auf Aktivitäten, hält sich meist vornehm zurück und schließt die Augen vor zuviel Sinnlichkeit. ■

Tab. 10-2 Auswertung

fächer Deutsch, Mathe, Englisch (mit Faktor 5 gewichtet) und die Differenz der Klassenstufe (mit Faktor 20 gewichtet – denn kein Junge aus Klasse 7 würde eine Abiturientin als Traumpartner akzeptieren). Natürlich hat man bei der Wahl des Abstandsmaßes freie Hand, man kann auch anders wichten oder quadrieren, statt Beträge zu bilden.

Die Datei wird nun nach folgendem Algorithmus durchsucht:

1. Suche in der Datei die erste Person des anderen Geschlechts.
2. Ermittle den erotischen Abstand dieser Person zu Dir und speichere ihn auf "Distanz". Merke Dir den Namen auf "Traumpartner".
3. Solange das Dateiende nicht erreicht:
  - nimm die nächste Person des anderen Geschlechts, ermittle den Abstand zu Dir;
  - ist dieser Abstand kleiner als "Distanz", so speichere ihn als neuen Wert von "Distanz" und den zugehörigen Namen als neuen Wert von "Traumpartner" ab.
4. Am Ende enthält "Traumpartner" den Namen der Person mit minimaler "Distanz" zu Dir.

Dieser Algorithmus ist leicht zu erarbeiten und als Datenbank-Abfrage zu formulieren, wenn man benutzerfreundliche Programmierumgebungen (etwa DBASE) einsetzt. Natürlich wünscht man neben dem Namen des Traumpartners auch eine Personencharakteristik. Hierzu fragt man in einem linearen Algorithmus interessant erscheinende Testantworten ab und leitet daraus Charaktermerkmale ab. Beispielsweise erhält jemand, der 3c ankreuzt, die Charakterisierung "ein wenig materialistisch". Jemanden, der 4b ankreuzt, nennen wir "manchmal etwas ängstlich". Die Sternzeichen liefern Charakterisierungen, die in einschlägigen Horoskopen in beliebiger Vielfalt zu finden sind. (Widder sind "selbständig und stolz" ... ) Man wird auch Kombinationen von Merkmalen abfragen und Formulierungen durch Zufallsgeneratoren abwechslungsreich gestalten. Hier das Beispiel eines Suchprotokolls. Es wird der Traumpartner gesucht von Bernd:

Petra:	Distanz	191
Iliana:	Distanz	190
Bärbel:	Distanz	15
Marita:	Distanz	13
Monika:	Distanz	2

Ihr Traumpartner Monika ist

- ehrgeizige Führernatur, aber nicht egoistisch
- eher naturwissenschaftlich interessiert
- ein verlässlicher Freund
- von Grund auf ehrlich
- sehnt sich nach Liebe
- tolerant.

Die Partnersuche innerhalb einer geschlossenen Gruppe führt zu einem "Soziogramm" (Abb.10-3). Die Kanten des gerichteten Graphen sind mit den "erotischen Abständen" beschriftet.

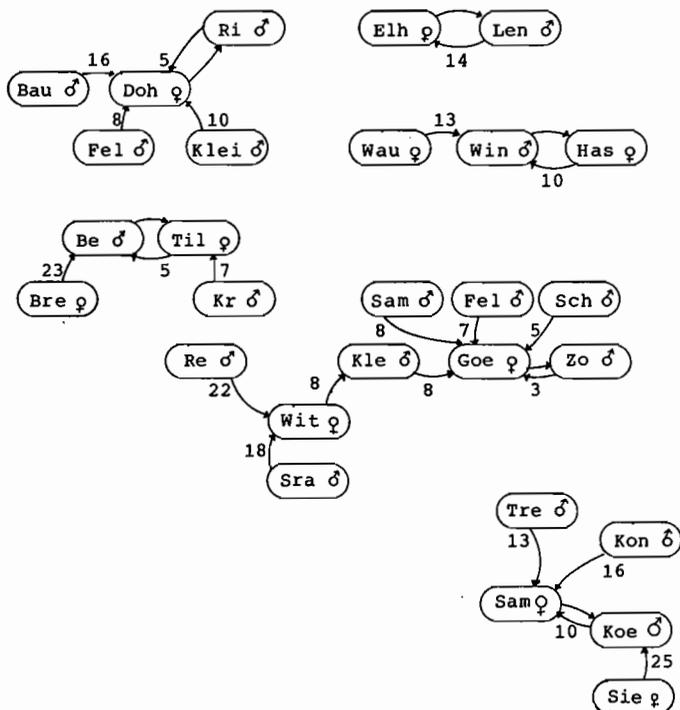


Abb. 10-3 "Soziogramm der Herzen"

Besonderes Interesse finden "stabile Pärchen" und "Attraktoren".

### 10.3 Modellbildung

#### 10.3.1 Zufriedene Leser

Kehren wir nach diesem Exkurs zurück zu Aspekten stochastischer Modellbildung: Mit welcher Wahrscheinlichkeit "erzeugt" sich "Hör Zu" zufriedene Kunden (das sind solche Paare, denen man mit weniger als 18 Punkten große Harmonie

verspricht)? Machen wir die Modellannahme, daß beide Partner die Fragen unabhängig voneinander aufs Geratewohl beantworten, also jede Alternative mit Wahrscheinlichkeit  $1/3$  ankreuzen. Wie aus Tabelle 10-4 ersichtlich, hätte dann das Testergebnis (=Punktsomme aus den 20 Items) den Erwartungswert 63 und die Standardabweichung 10.48, wird nach dem Zentralen Grenzwertsatz also annähernd  $N(63;10.48)$ -normalverteilt sein. Die Punktedifferenz zweier Partner besitzt dann den Erwartungswert 0 und die Standardabweichung  $\sqrt{2} \cdot 10.48=14.82$ . Die Wahrscheinlichkeit "zufriedener Kunden" ergibt sich dann zu

Nr.	Punkte	$\mu$	$\sigma^2$
	a b c		
1	6 0 3	3	6
2	0 4 6	10/3	168/27
3	0 3 6	3	6
4	3 0 6	3	6
5	3 6 0	3	6
6	6 1 3	10/3	114/27
7	6 3 0	3	6
8	0 6 3	3	6
9	6 0 3	3	6
10	3 6 0	3	6
11	6 3 0	3	6
12	0 3 6	3	6
13	1 6 4	11/3	114/27
14	6 3 0	3	6
15	3 1 6	10/3	114/27
16	6 3 1	10/3	114/27
17	3 6 0	3	6
18	4 0 6	10/3	168/27
19	6 3 1	10/3	114/27
20	3 6 1	10/3	114/27
$\Sigma$		63	109,77

Tab. 10-4 Kenngrößen der Items bei Gleichverteilung

$$\frac{1}{\sqrt{2\pi} \cdot 14.82} \int_{-17.5}^{17.5} e^{-0.5 \left( \frac{t-0}{14.82} \right)^2} dt = 76.2 \%$$

## 10.3.2 Unabhängigkeitstest

Überprüfen wir einen Aspekt unseres Modells, nämlich die Annahme, daß sich Jungen und Mädchen in ihrem Testergebnissen nicht unterscheiden, durch einen Chi-Quadrat-Unabhängigkeitstest. Wir verteilen die erreichten Punktskummen auf 10 Klassen. Eine Datenbankabfrage liefert:

<35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	71-75	>76	
5	14	19	36	47	31	12	13	8	4	189 ♂
10	22	18	45	30	25	14	12	3	2	181 ♀

Die Testgröße  $t=12.0$  liegt unterhalb der 95%-Grenze (16.9) der Chi-Quadrat-Verteilung mit 9 Freiheitsgraden. Die Daten widersprechen nicht der Annahme, daß die Punktskummen vom Geschlecht unabhängig sind.

## 10.3.3 Anpassungstest, Normalverteilung

Die weitergehende Modellannahme, die Punktskummen seien durch zufälliges Ankreuzen entstanden und gemäß 10.3.1 normalverteilt mit  $\mu=63$  und  $\sigma=10.48$  (Zeilen 1 und 4 in Tab. 10-5), wird durch einen Anpassungstest hochsignifikant widerlegt. Die Testgrößen (Jungen:  $t=230.8$  Mädchen:  $t=468.7$ ) liegen weit oberhalb der 99%-Grenze (21.7; 9 Freiheitsgrade).

<35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	71-75	>76	
Jungen										
0.4%	1.2%	3.2%	6.9%	12.1%	16.9%	18.9%	16.9%	12.1%	11.6%	N(63.0;10.48)   t=230.8
2.6%	7.4%	10.1%	19.0%	24.9%	16.4%	6.3%	6.9%	4.2%	2.1%	189 ♂
3.8%	6.3%	11.8%	17.2%	19.7%	17.7%	12.4%	6.8%	2.9%	1.3%	N(53.3;10.01)   t= 12.1
Mädchen										
0.4%	1.2%	3.2%	6.9%	12.1%	16.9%	18.9%	16.9%	12.1%	11.6%	N(63.0;10.48)   t=468.7
5.5%	12.2%	9.9%	24.9%	16.6%	13.8%	7.7%	6.6%	1.7%	1.1%	181 ♀
7.3%	8.9%	14.0%	17.8%	18.2%	15.0%	10.0%	5.4%	2.3%	1.1%	N(51.0;10.67)   t= 12.3

Tab. 10-5 Punktskummen und Normalverteilung

Schätzen wir dagegen die Parameter aus der Stichprobe (Jungen:  $\mu=53.3$ ,  $\sigma=10.01$ , Mädchen:  $\mu=51.0$ ,  $\sigma=10.67$ , Zeilen 3 und 6 in Tab.10-5), so liegen die Testwerte mit 12.1 und 12.3 unterhalb der 95%-Grenze (14.1) der Chi-Quadrat-Verteilung mit 7 Freiheitsgraden. (Wir haben zwei Parameter geschätzt, deswegen reduziert sich die Zahl der Freiheitsgrade von 9 auf 7.)

Wir gehen also weiterhin von normalverteilten Punktschümen bei Jungen und Mädchen aus und berechnen erneut (jetzt mit den Parameterwerten aus der Stichprobe) die Wahrscheinlichkeit "zufriedener Kunden". Die Punktdifferenz zwischen männlichem und weiblichem Partner eines Paares wird normalverteilt sein mit  $\mu = 53.3 - 51.0 = 2.3$  und  $\sigma = \sqrt{12.01^2 + 10.67^2} = 14.64$ . Hiermit erhalten wir für die Wahrscheinlichkeit zufriedener Kunden (weniger als 18 Punkte Unterschied)

$$\frac{1}{\sqrt{2\pi} \cdot 14.64} \int_{-17.5}^{17.5} e^{-0.5 \left( \frac{t - 2.3}{14.64} \right)^2} dt \approx 76.3\% ,$$

also einen Wert der von dem Ergebnis aus 10.3.1 kaum abweicht.

#### 10.3.4 Vermittlungschancen

Stefan hat 73 Punkte erreicht. Wie viele Partnerinnen (mit weniger als 18 Punkten Unterschied) wird die Datei (mit 181 Mädchen) für ihn bereithalten? Wir benutzen zu einer Prognose wieder die Normalverteilung für die Mädchen:

$$\frac{1}{\sqrt{2\pi} \cdot 10.67} \int_{73-17.5}^{73+17.5} e^{-0.5 \left( \frac{t - 51.0}{10.67} \right)^2} dt \approx 33.7\% .$$

Stefans Vermittlungschancen sind trotz seiner recht hohen Punktsumme beträchtlich. Die Datei wird ca. 60 Partnerangebote bereithalten. Natürlich kann man hier noch weiter in die Tiefe gehen und wieder mit Hilfe des "erotischen Abstandes" vermitteln: Man fragt die Verteilungen der relevanten Items ab, bestimmt Erwartungswerte und Varianzen für die betraglichen Punktdifferenzen zu Stefan und erhält mit einer Unabhängigkeitsannahme die Normalverteilung, der die erotischen Abstände zu Stefan genügen werden. Natürlich hat diese Normalverteilung für jede Person andere Parameter, kann aber nach Festlegung von Abweichungsschranken wieder zur Abschätzung von Vermittlungschancen benutzt werden. Das Faszinierende an diesen Beispielen ist erstens, daß wir mit Daten arbeiten, die die Schüler selbst betreffen (!) und zweitens die Tatsache, daß sich alle Prognosen durch kurze Datenbankabfragen unmittelbar überprüfen lassen.

## LITERATUR

- Bauersfeld, H. u. a.  
Lehren und Lernen von Mathematik. IDM Untersuchungen zum Mathematikunterricht Band 6. Köln 1983.
- Bauersfeld, H.  
Ergebnisse und Probleme von Mikroanalysen mathematischen Unterrichts. In: Empirische Untersuchungen zum Lehren und Lernen von Mathematik. Beiträge zum 4. Internationalen Symposium für Didaktik der Mathematik. Schriftenreihe Didaktik der Mathematik, Universität Klagenfurt, Bd. 10. Wien 1985.
- Bauersfeld, H. / Voigt, J.  
Den Schüler abholen, wo er steht! In: Lernen - Ereignis und Routine. Friedrich Jahresheft IV 1986: 18-20.
- Benz, H. J. / Borovcnik, M.  
Probleme bei empirischen Untersuchungen zum Wahrscheinlichkeitsbegriff. Journal für Mathematikdidaktik 4/1985: 241-242.
- Biehler, R.  
Die Neyman-Pearson-Theorie statistischer Hypothesentests in historischer und didaktischer Perspektive. Beiträge zum Mathematikunterricht 1985: 71-74.
- Biehler, R.  
Explorative Datenanalyse. IDM Materialien und Studien Bd. 24. Bielefeld 1982.
- Blackwell, D. / Deuel, P. / Freedman, D.  
Ann. Math. Statist., Vol 35 1964: 1344.
- Borovcnik, M.  
Eine Analyse des Chi-Quadrat-Unabhängigkeitstests - Möglichkeiten und Grenzen. In: Stochastik im Schulunterricht. Beiträge zum 3. Internationalen Symposium für Didaktik der Mathematik. Schriftenreihe Didaktik der Mathematik, Universität Klagenfurt, Bd. 3. Wien 1981: 35-48.
- Borovcnik, M.  
Was bedeuten statistische Aussagen. Stuttgart/Wien 1984.
- Borovcnik, M.  
Begünstigen - Eine stochastische Intuition. Praxis der Mathematik 27 1985: 327-333.
- Bruner, J. S.  
Der Prozeß der Erziehung. Berlin 5. Auflage 1980.
- Bruner, J. S.  
Entwurf einer Unterrichtstheorie. Berlin 1974.
- Chung, K.L.  
Elementare Wahrscheinlichkeitstheorie und stochastische Prozesse. Berlin, Heidelberg, New York 1978.

Cochran, W. G.  
The Chi-Square-Test of Goodness of Fit. The Annals of Mathematical Statistics 23 1952: 315-345.

Cohen, J. / Hansel, M.  
Risk and gambling: The study of subjective probability.  
New York 1956.

Davies, C. M.  
Development of the probability concept in children.  
Child Development 36 1965: 779-788.

Diepgen, R.  
Signifikanz - na und? Mathematiklehren 13 1985: 54-58.

Dinges, H.  
Schwierigkeiten mit der Bayesschen Regel. Mathematische und Physikalische Semesterberichte 1978: 113-156.

Dinges, H. / Rost, H.  
Prinzipien der Stochastik. Stuttgart 1982.

Doherty, J.  
Levels of four concepts of probability possessed by children of fourth, fifth and sixth grade before formal instruction.  
University of Missouri 1965.

Einstein, A.  
Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Ann. d. Phys. 17 1905: 549-560.

Einstein, A.  
Zur Theorie der Brownschen Bewegung. Ann. d. Phys. 19 1906: 371-381.

Engel, A.  
Wahrscheinlichkeitsrechnung und Statistik Bände 1, 2.  
Stuttgart 1976.

Falk, R. / Falk, R. / Levin, I.  
A potential for learning probability in young children.  
Educational Studies in Mathematics 11 1980: 181-202.

Feller, W.  
An Introduction to Probability Theory I, 3rd Edition.  
New York 1970.

Feller, W.  
An Introduction to Probability Theory II, 2nd Edition.  
New York 1970.

Feuerpfeil, J. / Heigl, F. / Wiedling, H.  
Praktische Stochastik. München 1983.

Fischbein, E.  
The intuitive sources of probabilistic thinking in children.  
Dortrecht 1975.

Fischbein, E.  
Image and concept in learning mathematics. Educational  
studies in Mathematics 8 1977: 153-165.

Fischbein, E.  
Intuition, structure and heuristic methods in the teaching  
of mathematics. Developments in Mathematical Education:  
Proceedings of the Second International Congress on Mathema-  
tical Education. Hrsg. A. G. Howson. Cambridge 1973: 222-  
232.

Fischbein, E. / Barbat, I. / Minzat, I.  
Intuitions primaires et intuitions secondaires dans  
l'initiation aux probabilités. Educational Studies in Mathe-  
matics 4 1971: 264-280.

Fischbein, E. / Gazit, A.  
Does the teaching of Probability improve Probabilistic  
Intuitions? Educational Studies in Mathematics 15/1 1984:  
1-24.

Fischbein, E. / Pampu, I. / Minzat, I.  
L'intuition probabiliste chez l'enfant. Enfance 2 1967:  
193-208.

Fischbein, E. / Pampu, I. / Minzat, I.  
Effects of age and instruction on combinatory ability in  
children. The British Journal of educational Psychology 40  
1970: 261-270.

Fisz, M.  
Wahrscheinlichkeitsrechnung und mathematische Statistik.  
Berlin 1971.

Freudenthal, H.  
The crux of course design in probability. Educational  
Studies in Mathematics 5 1974: 261-277.

Goldberg, S.  
Probability judgements by preschool children; task condi-  
tions and performance. Child Development 37 1966: 157-167.

Gratch, G.  
The development of expectation of the nonindependence of  
random events in children. Child development 30 1959: 217-  
227.

E.A. Guggenheim  
Boltzmann's Distribution Law, 4. Auflage. Amsterdam 1967.

- Hawkins, A. / Kapida, R.  
Childrens Conceptions of Probability - A Psychological and Pedagogical Review. Educational Studies in Mathematics 15 1984: 349-377.
- Heigel, F.  
Elementare Ableitung der Laplaceschen Formel der Wahrscheinlichkeitsrechnung. Elemente der Mathematik 1976: 53-57.
- Heitele, D.  
Über den Wert und die Problematik der vorliegenden Befunde zur Entwicklung stochastischer Ideen beim Kinde. In: Beiträge zum Mathematikunterricht. Hannover 1975: 49-53.
- Heitele, D.  
Didaktische Ansätze zum Stochastikunterricht in Grundschule und Förderstufe. Diss. Dortmund 1976.
- Helmert, F. R.  
Über die Wahrscheinlichkeit der Potenzsummen und über einige damit in Zusammenhang stehende Fragen. Z. Math. Phys. 21 1876: 192.
- Hoemann, H. W. / Ross, B. M.  
Children's understanding of probability concepts. Child Development 42 1971: 221-236.
- Ikeda, N. / Watanabe S.  
Stochastic Differential Equations and Diffusion Processes. New York 1981.
- Ineichen, R.  
Über die Behandlung der Normalverteilung. Der Mathematikunterricht 4 1966: 62-75.
- Inhetveen, H.  
Die Reform des gymnasialen Mathematikunterrichts zwischen 1890 und 1914. Eine sozioökonomische Analyse. Bad Heilbrunn 1976.
- Ito, K.  
Diffusion Processes and their Sample Paths. Berlin 1965.
- Johnson, N. / Kotz, S.  
Continuous Univariate Distributions - 2. New York 1970.
- Kac, M.  
Random Walk and the Theory of Brownian Motion  
Amer. Math. Monthly 54 1949: 369-391.
- Khintchine, A.  
Asymptotische Gesetze der Wahrscheinlichkeitsrechnung.  
Berlin 1933.
- Kolmogoroff, A.  
Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Mathematische Annalen 104 1931: S. 415-458.

- Kütting, H.  
Didaktik der Wahrscheinlichkeitsrechnung. Freiburg 1981.
- Leake, L.  
The status of three concepts of probability in children of the 7th, 8th and 9th grades. University of Wisconsin 1962.
- Leffin, W. W.  
A study of three concepts of probability possessed by children in the 4th, 5th and 6th grades. Techn. Report 170. University of Wisconsin 1971.
- Lietzmann, W.  
Methodik des Mathematikunterrichts Bd. 2, 2. Aufl. Heidelberg 1953.
- Lind, D. / Scheid, H.  
Abiturwissen Stochastik. Stuttgart 1984.
- Morgenstern, D.  
Didaktische Schwierigkeiten mit der Stochastik. Mathematische und Physikalische Semesterberichte 1/1978: 31-51.
- Nemetz, T.  
Teaching Statistics to 16-18 Year Olds: Dreams and Reality. In: Grey, D. R. e. a. (eds.) Proceedings of First International Conference on Teaching Statistics Vol. I. Sheffield 1983: 199-207.
- Pearson, K.  
On the Criterion that a given System of Deviations from the Probable in the Case of a correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling. Phil. Magazine Vol. 50 1900: 157-175.
- Pfanzagl, J.  
Allgemeine Methodenlehre der Statistik Band II. Berlin 1974.
- Pfanzagl, J.  
Elementare Wahrscheinlichkeitsrechnung. Berlin 1988.
- Piaget, J. / Inhelder, B.  
La genese de l'idée du hasard chez l'enfant Paris 1951.  
Übersetzung: The Origin of the Idea of Chance in Children. New York 1975.
- Riemer, W.  
Neue Ideen zur Stochastik. Mannheim 1985.
- Riemer, W.  
Der Varianz-Abakus. Mathematiklehren 12 1985: 39-43.
- Riemer, W.  
Eine neue Sicht der Bayesschen Regel. Stochastik in der Schule 3/1986: 4-13.

Riemer, W.  
Ein Partnertest. Beschreibende Statistik kann spannend sein.  
Mathematiklehren 22 1987: 42-45.

Riemer, W.  
Neue Ideen zum Chi-Quadrat-Verfahren, Teil I. Didaktik der  
Mathematik 2/1987: 115-126.

Riemer, W.  
Neue Ideen zum Chi-Quadrat-Verfahren, Teil II. Didaktik der  
Mathematik 3/1987: 211-218.

Riemer, W.  
Riemer-Würfel, Begleitheft, Disketten. Stuttgart 1988  
(Klett).

Riemer, W.  
Neue Aspekte in der beurteilenden Statistik mit dem Computer  
und der Regel von Bayes. Der Mathematikunterricht 4/1989:  
48-62.

Riemer, W.  
Das Arcsin-Gesetz der Wahrscheinlichkeitsrechnung. Der  
Mathematikunterricht 4/1989: 64-75.

Riemer, W.  
Der Chi-Quadrat-Anpassungstest und Irrfahrten in der Ebene.  
Der Mathematische und Naturwissenschaftliche Unterricht 42/6  
1989: 344-352.

Riemer, W.  
Der Chi-Quadrat-Unabhängigkeitstest und Irrfahrten in der  
Ebene. Erscheint in: Der Mathematikunterricht 1990.

Scheid, H.  
Stochastik in der Kollegstufe. Mannheim 1986.

Schneider, H. / Stein, G.  
Der Chi-Quadrat-Test im Stochastik-Grundkurs der Sekundar-  
stufe II. Didaktik der Mathematik 3/1980: 200-212.

Schönwald, H. G.  
Wahrscheinlichkeit und Wirklichkeiten. Eine Bemerkung zur  
Wahrscheinlichkeitsrechnung in Klasse 6. Stochastik in der  
Schule 1/1984: 29-34.

Schrage, G.  
Schwierigkeiten mit stochastischer Modellbildung. Journal  
für Mathematikdidaktik 1/1980: 86-101.

Schrage, G.  
Entscheiden und Begründen - Leitlinien für den Statistik-  
unterricht. In: Stochastik im Schulunterricht. Beiträge zum  
3. Internationalen Symposium für Didaktik der Mathematik.  
Schriftenreihe Didaktik der Mathematik, Universität  
Klagenfurt, Bd. 3. Wien 1981: 179-208.

- Schrage, G.  
Stochastische Trugschlüsse. *Mathematica Didactica* 3/1984:  
3-19.
- Schrage, G.  
Irrwege zur Stochastik. *Der Mathematikunterricht* 5/1984:  
50-53.
- Schreiber, A.  
Idealisierungsprozesse - Ihr logisches Verständnis und ihre  
didaktische Funktion. *Journal für Mathematikdidaktik* 1/1980:  
42.
- Schröder, J.  
Wahrscheinlichkeitsrechnung in der Sekundarstufe I: Indeter-  
ministische Problemsituationen im Unterricht. In: *Stochastik  
im Schulunterricht. Beiträge zum 3. Internationalen Sympo-  
sium für Didaktik der Mathematik, Universität Klagenfurt.*  
Wien 1981: 217-222.
- Schupp, H.  
Zum Verhältnis statistischer und wahrscheinlichkeitstheore-  
tischer Komponenten im Stochastik-Unterricht der Sekundar-  
stufe I. *Journal für Mathematikdidaktik* 3/4 1982: 207-226.
- Schuss, Z.  
*Theory and Applications of Stochastic Differential  
Equations.* New York 1980.
- Scozzafava, R.  
Subjective probability and Bayesian statistics in engineering  
mathematics education. *International Journal of Mathematical  
Education in Science and Tecnology.* 18/5 1987: 685-688.
- Shaughnessy, J. M.  
Misconceptions of probability. *Educational Studies in  
Mathematics* 8 1978: 295-315.
- Smirnow, W. I.  
*Lehrgang der Höheren Mathematik Band II.* 12. Aufl.  
Berlin 1968.
- Steinbring, H.  
Zur Entwicklung des Wahrscheinlichkeitsbegriffs - Das Anwen-  
dungsproblem in der Wahrscheinlichkeitstheorie aus didakti-  
scher Sicht. *IDM Materialien und Studien Bd. 18.* Bielefeld  
1980.
- Steinbring, H. / Strässer, R. (Hrsg.)  
Rezensionen von Stochastik-Lehrbüchern beider Sekundarstu-  
fen. *Zentralblatt für Didaktik der Mathematik* 13/1981.
- Steinbring, H. / v. Harten, G.  
Learning from Experience - Bayes Theorem: A Model for  
Stochastic Learning Situations? In: Grey, D. R. e. a. (eds.)  
*Proceedings of First International Conference on Teaching  
Statistics Vol. II.* Sheffield 1983: 701-713.

- v.d. Steinen, J.  
Impulse aus der Stochastik auf den Analysisunterricht.  
Der Mathematikunterricht 25/2 1979: 70-81.
- Strick, H. K.  
Der Chiquadrat-Anpassungstest im Mathematik- und Biologieunterricht der Sekundarstufe II. Der Mathematische und Naturwissenschaftliche Unterricht 25/2 1981: 139-147.
- Varga, T.  
Statistics in the Curriculum for everybody - how young children and how their teachers react. In: Grey, D. R. e. a. (eds.) Proceedings of First International Conference on Teaching Statistics Vol. I. Sheffield 1983: 71-78.
- Voigt, J (Hrsg.)  
Mathematikunterricht im 5. bis 11. Schuljahr. Transskripte zum Projekt "Routinen im Mathematikunterricht". IDM Materialien und Studien Bd. 33. Bielefeld 1983.
- Weinert, F.E. / Kluwe, R.H. (Hrsg.)  
Metakognition, Motivation und Lernen. Stuttgart 1984.
- Wickmann, D.  
Ein Plädoyer für die subjektivistische Stochastik. Habilitationsschrift TH Aachen 1984.
- Winter, H.  
Erfahrungen zur Stochastik in der Grundschule.  
Didaktik der Mathematik 1/1976: 22-37.
- Witting, H. / Nölle, G.  
Angewandte mathematische Statistik. Stuttgart 1970.
- Wittmann, E.  
Grundfragen des Mathematikunterrichts. 6. Aufl. Braunschweig 1983.
- Wittmann E.  
Teaching Units as the Integrating Core of Mathematics Education. Educational Studies in Mathematics 15/1 1984: 25-36.
- Yost, P. A. / Siegel, A. / Andrews, J.  
Nonverbal probability judgements by young children. Child Development 33 1962: 769-780.

## Index

- Alternativtest 58
- Arcsin-Gesetz 138 ff, 146, 165
- Arcsin-Verteilung, Differentialgleichung 141, 145, 166
- Baumdiagramm 29
- Bayes, Regel von 39 ff, 44 ff
- Besseldiffusion 159
- Boltzmann-Verteilung 111 ff
- Brownsche Bewegung 153 ff
- Chi-Quadrat
  - Anpassungstest 62 ff, 136 f
  - Anpassungstest mit Parameter 110 ff, 174
  - Differentialgleichung 134, 159, 161
  - Unabhängigkeitstest 89 ff, 174
  - Verteilung 69, 84
- Computereinsatz 47, 55, 68, 79, 92, 102, 130, 139 ff, 171 ff
- Datenbank 89, 168 ff
- Diffusionsprozesse 157
- Entropie 52
- Erotischer Abstand 168
- Freiheitsgrad 62, 79, 82, 108, 110
- Gammaverteilung, Differentialgleichung 135
- Glaubwürdigkeit (von Hypothesen) 22, 65
- Hypothese 16, 20, 24, 25 ff, 28, 32
- Indiz s. Bayessche Regel
- induktives Schließen 39, s. a. Bayessche Regel
- Information 54
- Interaktionsanalyse 36
- Intuition 21, 49
- Irrfahrt, Ebene 74, 76, 99
- Irrfahrt, Gerade 120, 128, 140
- Kolmogoroffs Differentialgleichungen 157

Likelihoodquotienten 51  
Modellbildung 28, 33, 173  
Normalverteilung, Differentialgleichung 121 ff, 156, 174  
Normalverteilung, zweidimensionale 81, 103  
Paradigma, statistisches 15  
Partnertest 168  
Pearson 87  
Potenzverteilung 116  
Quader 21 ff, 110 ff  
Reißnagel 35, 38  
Rekursionsformeln 120  
Renewal Theorie 166  
Riemer-Würfel 41 ff, 110 ff  
Schrittvektor 75, 83, 98, 104  
Soziogramm, "erotisches" 172  
Strukturmathematik 12  
Wahrscheinlichkeitsbegriff  
- frequentistischer 18  
- Laplacescher 17, 34  
- objektivistischer 57, 66  
- prognostischer 19, 20  
- subjektivistischer 17, 21 ff, 57, 66  
Wartezeit 29  
Wurftechnik 110  
Zentraler Grenzwertsatz 80, 118 ff, 153 ff  
Zufallszahlen 63, 90  
Zustandsvektor 74, 107

## Experimentier-Set

---

Der Autor bietet ein Experimentier-Set an, bestehend aus

- einem Klassensatz mit 30 Quadern für die Experimente aus den Kapiteln 1, 3, 5,
- einer Begleitdiskette mit zwei Programmen zur Bayesschen Regel nach Kapitel 2 (MS-DOS, 3 1/2" oder 5 1/4" Text- und Code),
- einer Kurzanleitung.

Sie erhalten dieses Set direkt vom Autor für 35,-- DM (Verrechnungsscheck oder Überweisung auf das Konto: W. Riemer, 170635-509, Postgiro Köln, BLZ 370 100 50) und gegen Einsendung eines mit 1,80,-- DM frankierten Rückumschlages (Warensendung) an:

Dr. W. Riemer  
August-Bebel-Str. 80  
a5024 Pulheim 2

Änderungen vorbehalten

---

## Lehrbücher und Monographien zur Didaktik der Mathematik:

Entsprechend dem zunehmenden Interesse an didaktischen Fragestellungen ist die Zahl der Zeitschriften zur Didaktik der Mathematik inzwischen stark angewachsen. Diese Entwicklung ist grundsätzlich zu begrüßen. Für den Studenten wie für den in der Praxis stehenden Lehrer und Dozenten ist es dadurch aber auch schwieriger geworden, einen Überblick über alle relevanten Aspekte der Originalliteratur zu erwerben bzw. zu behalten.

Hier soll die neue Schriftenreihe „Lehrbücher und Monographien zur Didaktik der Mathematik“

Information und Orientierung ermöglichen. Die einzelnen Bände wenden sich nicht nur an den Spezialisten. Sie sollen dem interessierten Leser, auch wenn er nicht über Kenntnisse der Originalliteratur verfügt, Informationen über alle aktuelle und relevante Fragestellungen und über deren Diskussionsstand vermitteln und ihn in die Lage versetzen, diese Diskussion wertend zu analysieren.

Band 1:

Fischer, R./G. Malle

### **Mensch und Mathematik. Eine Einführung in didaktisches Denken und Handeln**

Unter Mitarbeit von Doz.

Dr. Heinrich Bürger

Hinweise für einen Unterricht, der neben der Vermittlung wichtiger Kenntnisse und Fertigkeiten auch einen Beitrag zur Entwicklung eines reflektierten Verhältnisses zur Mathematik leisten will.

367 Seiten. 1985. Wv.

Band 2:

Blankenagel, J.

### **Numerische Mathematik im Rahmen der Schulmathematik.**

#### **Ansätze zu einer Didaktik**

Vieles, was heute den Computer im Unterricht wichtig erscheinen läßt, kann auch von der Numerischen Mathematik aus erschlossen werden.

192 Seiten. 1985. Wv.



**Wissenschaftsverlag**

Mannheim/Wien/Zürich

## Lehrbücher und Monographien zur Didaktik der Mathematik:

Band 3:

Riemer, W.

### **Neue Ideen zur Stochastik**

Mit Hilfe neuer heuristischer Modelle gelangt man unter Umgehung sinnleerer Formalismen zum Kern zentraler stochastischer Themenbereiche.

156 Seiten. 1985. Wv.

Band 4:

Knoche, N./H. Wippermann

### **Vorlesungen zur Methodik und Didaktik der Analysis**

Unterschiedliche Zugänge zu den grundlegenden Begriffen der Analysis werden vorgestellt und im Vergleich wertend diskutiert.

350 Seiten. 1986. Wv.

Band 5:

Laugwitz, D.

### **Zahlen und Kontinuum.**

#### **Eine Einführung in die Infinitesimalmathematik**

Eine Einführung in die Nichtstandard-Analysis mit Darstellung der historischen und philosophischen Grundlagen der Infinitesimalrechnung.

269 Seiten. 1986. Wv.

Band 6:

Scheid, H.:

### **Stochastik in der Kollegstufe**

Fachwissenschaftliche und didaktische Hintergründe der Stochastik und ihre Bezüge zur Analysis; auf die Einsatzmöglichkeiten programmierbarer Rechner wird eingegangen.

250 Seiten. 1986. Wv.

Band 7:

Padberg, F.

### **Didaktik der Arithmetik**

Eine praxisnahe Darstellung der Arithmetik der ersten vier Schuljahre. Im Vordergrund stehen die von Schülern benutzten Lösungsstrategien sowie typische Schülerfehler.

253 Seiten. 1986. Wv.



**Wissenschaftsverlag**

Mannheim/Wien/Zürich

