Zufall! Zufall?

Gerhard Winkler[1] **12. Februar 2011**

Keywords and Phrases: Zufall, Wahrscheinlichkeit, Grundlagen der Wahrscheinlichkeitstheorie
Mathematical Subject Classification: 00A30, 60A05, 97E20

Zusammenfassung

In diesem Essay beschäftigen wir uns mit dem Zufall. Die Mathematik lassen wir dabei beiseite. Es geht vielmehr um eine Diskussion der vielfältigen Facetten dieses Begriffes, von der Intuition bis hin zur mathematischen Axiomatik.

In der Tat zählt der Zufall zu den oft mißverstandenen fundamentalen Begriffen. Dies gilt vom menschlichen Alltag angefangen bis in die verschiedensten Zweige der Wissenschaft und auch den schulischen Bereich hinein.

Inhaltsverzeichnis

[1] HelmholtzZentrum München, Institut für Biomathematik und Biometrie
Fakultät für Mathematik, Informatik und Statistik der Ludwig-Maximilians Universität München

1 Einleitung

Viele mathematische Gegenstände haben eine Abfolge von Entwicklungsstufen durchlaufen, beginnend mit subjektiver Einschätzung oder einem Gefühl auf der Basis realer Erscheinungen, forschreitend über einen erkenntnistheoretischen auf der Basis philosophischer Überlegungen und dann einen wissenschaftlichen Begriff, der sich u.a. auf Messungen und Beobachtungen stützt, bis schließlich hin zur abstrakten mathematischen Axiomatisierung und dem anschließenden mathematischen Glasperlenspiel gemäß den Regeln solcher Axiomatik.

Beispiele sind etwa die klassischen Disziplinen der Geometrie oder Algebra. Auch in der Thermodynamik finden wir eine solche Entwicklung,von der gefühlten Temperatur, über die objektive Messung mittels Thermometern[1] hin zu den experimentell gewonnenen Grundgesetzen und schließlich den Hauptsätzen der Thermodynamik und der Gibbsschen Fundamentalform.

Schließlich stellt sich die Frage, wie die daraus gewonnen Sätze wieder in Form von Entscheidungen auf die Realität zurückwirken sollen.

Dieser Aufsatz beschäftigt sich mit einem Begriff, der ungleich schillernder ist, nämlich dem Zufall. Die mathematische Lehre vom Zufall nennt man *Stochastik*. Das Wort kommt aus dem Griechischen und bedeutet „zum Erraten gehörende Kunst“. Wir zitieren frei nach PLATON (siehe PLATON (1957-1959)):

> SOKRATES: Wenn man zum Beispiel von allen Künsten die Rechenkunst und die Meßkunst und die Statik absonderte, so wäre doch das, was von jeder noch übrig bleibt, sozusagen bedeutungslos.
> PROTARCHOS: Jawohl, ganz bedeutungslos.
> SOKRATES: Wenigstens blieben uns danach nur noch Schätzungen übrig und die Übung unserer Wahrnehmungen mit Hilfe der Erfahrung und einer gewissen Routine, wobei wir dazu noch die *Kräfte des Vermutens* anwendeten, die viele als Künste bezeichnen, die aber *nur nach mühevoller Übung ihre Wirkung ausüben können.*

Heute ist Stochastik der Oberbegriff für *Wahrscheinlichkeitstheorie* (oder -rechnung) und *Statistik*; sie umfaßt alle theoretischen und quantifizierbaren Aspekte ungewisser Erscheinungen.

Erstere beschäftigt sich mit Axiomen und Folgerungen daraus in streng mathematischer Weise, allerdings stets mit dem Blick auf Anwendungen.

[1] deren Entwicklung fast parellel zuWahrscheinlichkeitrechnung verlief

Die (schließende) Statistik nutzt diese Erkenntnisse, um aufgrund erhobener Daten die Gültigkeit von Modellannahmen zu überprüfen, Vorhersagen zu treffen und letztlich Entscheidungen herbeizuführen. Hierbei stehen Daten und deren Analyse im Vordergrund. *Erhebungen* im Rahmen der *deskriptiven*

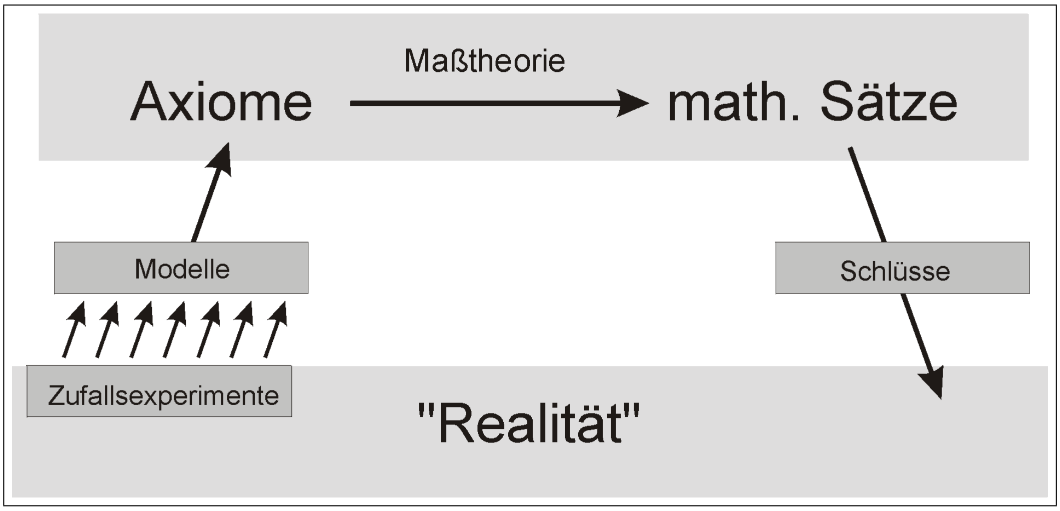


Abbildung 1: Schema zur Wahrscheinlichkeitstheorie

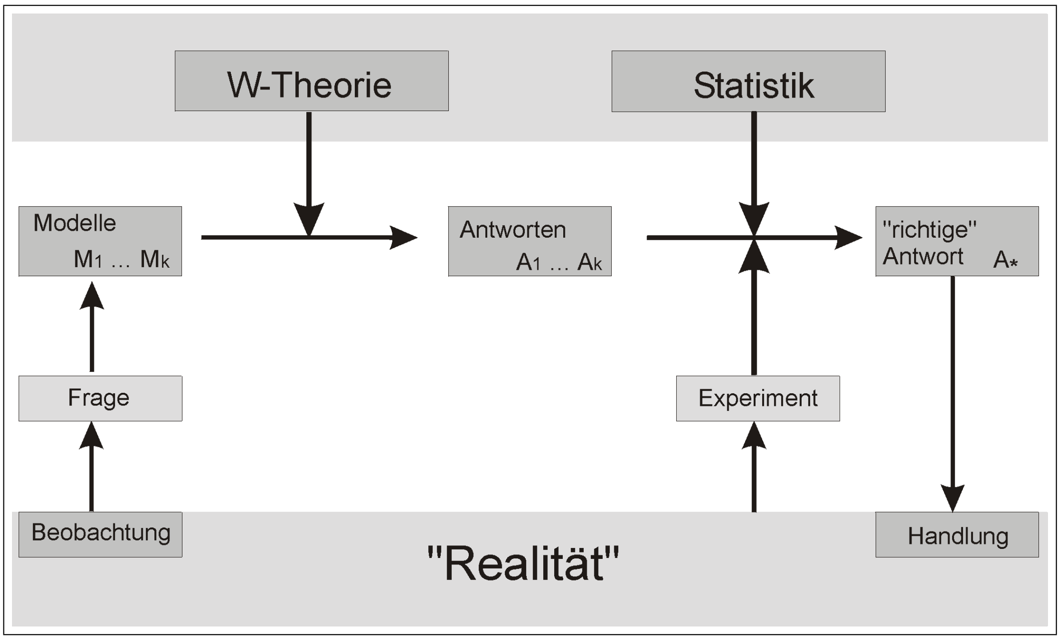


Abbildung 2: Schema zur Statistik

Statistik wurden schon in alter Zeit durchgeführt, z.B. in Form von Buch- und

Lagerhaltung oder von Volkszählungen. Sie beinhaltet neben der Tabellierung von Daten ihre summarische Beschreibung durch Kenngrößen wie Mittelwert, Median, Streuung und ähnliche Charakteristika. Dies ist gegenwärtig nicht unser Thema.

Die Anfänge einer wissenschaftlichen Betrachtung liegen in der Mitte des 17. Jahrhunderts. Die *schließende Statistik* in der strengen heutigen Form gibt es erst seit dem 20. Jahrhundert. Sie baut auf der *Wahrscheinlichkeitstheorie* auf, deren jetzige Form aus den dreißiger Jahren des 20. Jahrhunderts stammt. Die *Mathematisierung* dieser Gebiete ist also, verglichen mit Geometrie oder Algebra, die man bis ins Altertum zurückverfolgen kann, jung.

In den Abbildungen 1 und 2 versuchen wir symbolisch darzustellen, wie Stochastik funktioniert. Entscheidend ist, daß die Disziplinen Wahrscheinlichkeitstheorie und Statistik beide tief in der Realität verwurzelt sind. Ohne Blick auf die Wirklichkeit würde die Wahrscheinlichkeitstheorie auf Kombinatorik, Maßtheorie oder auch Funktionalanalysis zusammenschnurren. Obwohl die *mathematische Statistik* ein unverzichtbarer Bestandteil der Statistik ist, gewinnt die Statistik ihre Bedeutung haupsächlich aus der Analyse von realen Daten.

2 Psychologie des Zufalls

Zufall und Wahrscheinlichkeit sind schillernde Begriffe, die - vom wissenschaftlichen Standpunkt gesehen - zumeist unvollkommen oder sogar falsch verstanden benutzt werden. Unter Zufall versteht man gemeinhin ein unwägbares, unvorhersehbares Geschehen, etwas was einem *zu - fällt* oder sogar *zu - stößt*. Oft verbindet man mit Zufall ein eher seltenes Geschehen. Jedenfalls hätte es auch anders kommen können, es gehört nicht zu unserer eigentlichen Identität. Mit Altersschwäche rechnen wir, aber eine Krebserkrankung oder einen Unfall sehen wir nicht zwingend vor. Anders ausgedrückt: Wir haben einen Begriff von Normalität, von dem, was wesentlich zu uns gehört. Alles andere wird als Störung empfunden - eben als Zufall. Einige der folgenden Überlegungen sind durch dieses deutsche Wort „Zufall “inspiriert.

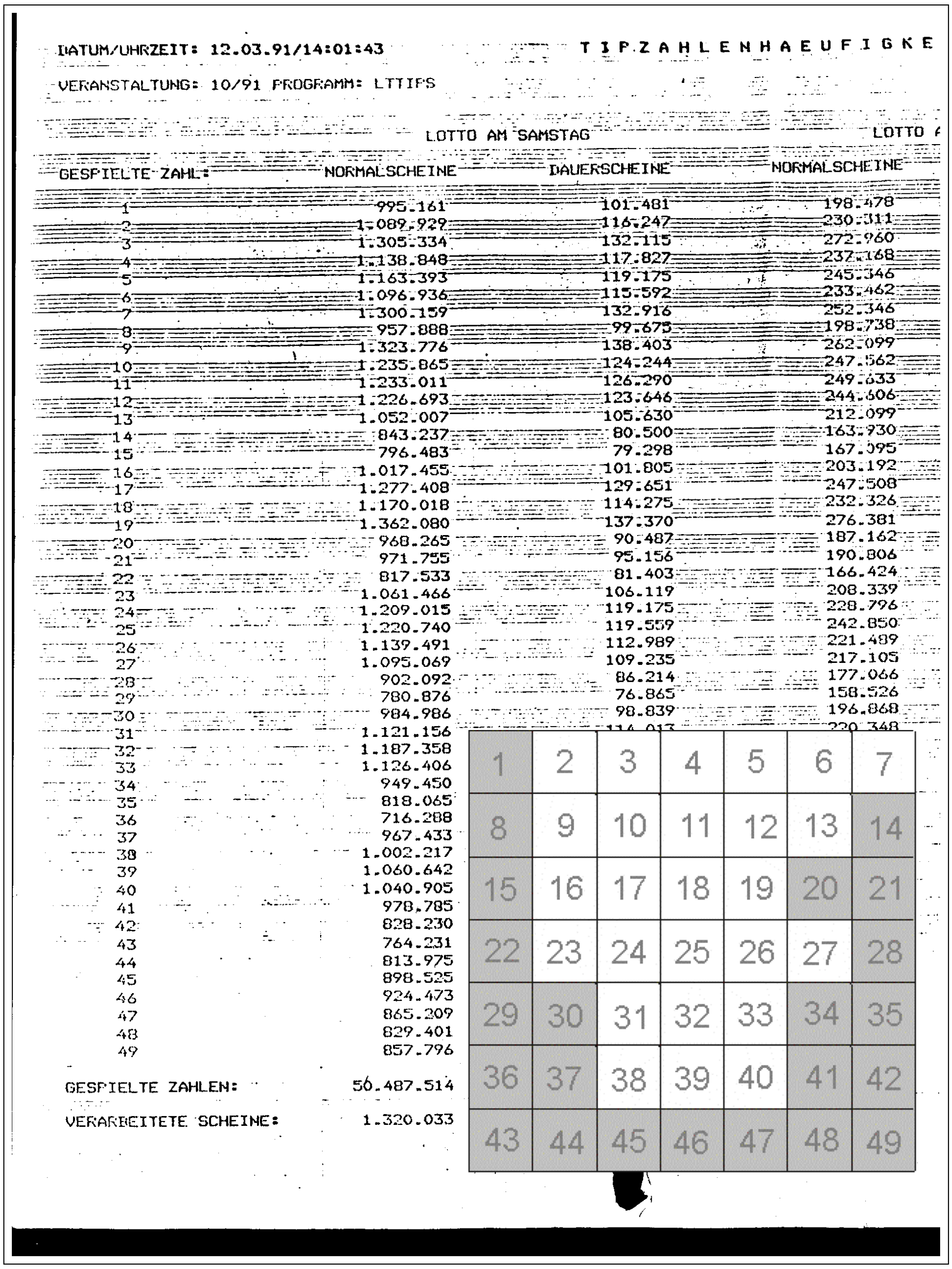
Im französischen finden wir die Worte *hasard*, *chance*, *accident*, *fortune*, *accident* oder auch *coincidence*. Hierzu sei bemerkt, da§ ursprünglich *chance* zunächst nichts mit dem Zufall oder Wahrscheinlichkeit zu tun hatte; es war einfach nur die Wahl, auf die man seinen Einsatz setzte, etwa *rouge* beim Roulette. Im Englischen finden wir im wesentlichen die französischen Wörter wieder: *hazard*, *chance*, *accident*, *fortune*, *accident*, *haphazard* usw.. Das wichtige Wort *hasard* scheint sus dem Arabischen zu kommen,vermutlich von *az zahr* für Würfel.

Die meisten täglichen Entscheidungen fällen wir aufgrund eines „(Wahrscheinlichkeits-) Gefühles“. Zum Beispiel laufen wir den Gehsteig entlang, obwohl uns ja theoretisch ein Dachziegel auf den Kopf fallen könnte. Wir wissen aber, daß dies in der Vergangenheit selten geschah, und vertrauen darauf, das es so bleibt. Ein solcher Schluß ist nicht unbedingt zielführend. Die meisten Hühner z.B. vertrauen darauf, daß sie den nächsten Tag gut gefüttert überleben. Die meisten verlieren aber schlußendlich Kopf und Kragen.

2.1 Zum Wahrscheinlichkeitsgefühl

Der wissenschaftlichen Präzisierung vieler Begriffe geht oft ein intuitiver Begriff voraus. Ein unverfängliches Beispiel ist das Gefühl für Temperaturen, das ja dem am Thermometer ablesbaren Wert nicht immer entspricht; deshalb spricht man ja zunehmend von der „gefühlten Temperatur“. Ähnlich, aber komplizierter, verhält es sich mit dem „Zufall“.

Der allgemein irrationale Umgang damit zeigt sich insbesondere im Verhalten bei Glücksspielen (aber auch in der Welt der Finanzen, was sehr ähnlich ist). Kommt beim Roulette etwa 10 mal rot, so erwarten viele Spieler beim 11. mal schwarz. Die Chance beträgt aber - unter dem üblichen mathematischen Modell - nach wie vor 1/2. Ein gewiefter Trader würde natürlich - dem entgegengesetzt - auf Rot tippen, da er einen Betrug vermuten würde. Beim Lotto ist die Gewinnerwartung lediglich 50% des Einsatzes, da nur die Hälfte des Einsatzes ausgeschüttet wird. Es handelt sich also um ein extrem unfaies Spiel zu ungunsten des Spielers. Trotzdem ist der Umsatz beim Lotto am Samstag alleine in Deutschland in der Größenordnung von 170 -200 Millionen Euro. Die Finanzämter freuen sich über Einnahmen von ca. 85 -100 Millionen Euro pro Samstag sehr (die Umsätze haben sich seit Gleichstellung des Mittwochs- und Samstagslottos etwas verschoben). Der Jackpot lag jedenfalls am Samstag, 1. Dezember 2007, bei ca. 38 Millionen Euro! Fragen Sie Bekannte, ob sie die Zahlenfolge 1,2,3,4,5,6 für besonders unwahrscheinlich oder wahrscheinlich halten! Man vergleiche auch, wie häufig etwa in Bayern die einzelnen Lottozahlen getippt werden. Die Tabelle und das Bild in Abb. 3 zeigen, daß die Häufigkeiten sowohl von der Zahl selbst, als auch von der Geometrie des Lottoscheines abhängen. Außer den mittleren Feldern hat die 19 die höchste Häufigkeit, weil sie in jedem Geburtsdatum vorkommt, usw. Es werden nicht nur einzelne Zahlen bevorzugt, sondern auch gewisse Tippmuster. Bei Lottofachleuten sind sie als „Strickmuster“bekannt. Diese führen regelmäßig zu niedrigen Quoten, denn man muß ja die Ausschüttung mit allen teilen, die denselben Tip abgegeben haben. Man erniedrigt also seine Gewinnhöhe systematisch (und drastisch) wenn man beliebte Zahlen



DATUM/UHRZEIT: 12.03.91/14:01:43 T I P Z A H L E N H A E U F I G K E

VERANSTALTUNG: 10/91 PROGRAMM: LTTIPS

GESPIELTE ZAHL:	NORMALSCHEINE (LOTTO AM SAMSTAG)	DAUERSCHEINE (LOTTO AM SAMSTAG)	NORMALSCHEINE (LOTTO A…)
1	995.161	101.481	198.478
2	1.089.929	116.247	230.311
3	1.305.334	132.115	272.960
4	1.138.848	117.827	237.168
5	1.163.393	119.175	245.346
6	1.096.936	115.592	233.462
7	1.300.159	132.916	252.346
8	957.888	99.675	198.738
9	1.323.776	138.403	262.099
10	1.235.865	124.244	247.562
11	1.233.011	126.290	249.633
12	1.226.693	123.646	244.606
13	1.052.007	105.630	212.099
14	843.237	80.500	163.930
15	796.483	79.298	167.095
16	1.017.455	101.805	203.192
17	1.277.408	129.651	247.508
18	1.170.018	114.275	232.326
19	1.362.080	137.370	276.381
20	968.265	90.487	187.162
21	971.755	95.156	190.806
22	817.533	81.403	166.424
23	1.061.466	106.119	208.339
24	1.209.015	119.175	228.796
25	1.220.740	119.559	242.850
26	1.139.491	112.989	221.489
27	1.095.069	109.235	217.105
28	902.092	86.214	177.066
29	780.876	76.865	158.526
30	984.986	98.839	196.868
31	1.121.156	114.013	220.348
32	1.187.358		
33	1.126.406		
34	949.450		
35	818.065		
36	716.288		
37	967.433		
38	1.002.217		
39	1.060.642		
40	1.040.905		
41	978.785		
42	828.230		
43	764.231		
44	813.975		
45	898.525		
46	924.473		
47	865.209		
48	829.401		
49	857.796		

GESPIELTE ZAHLEN: 56.487.514

VERARBEITETE SCHEINE: 1.320.033

Abbildung 3: Häufigkeit der in Bayern bei der 10. Veranstaltung 1991 getippten Lottozahlen

oder Muster tippt. In Abb. 4 sind tatsächlich häufig gewählte Tips in der oberen Reihe zufällig erzeugten in der unteren Reihe gegenübergestellt. Man

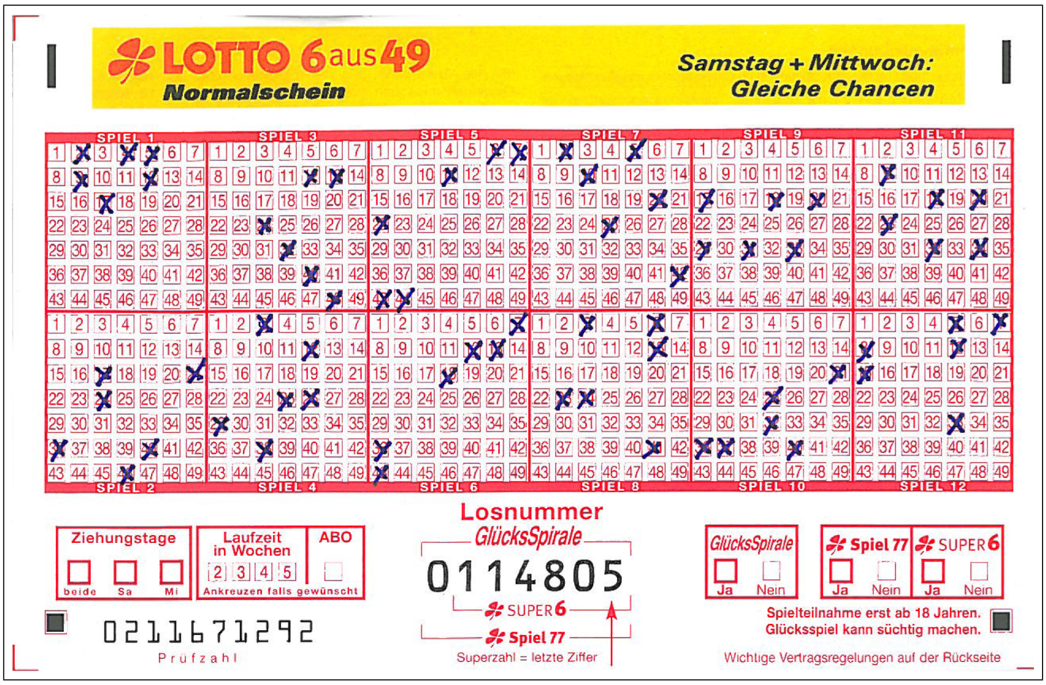


Abbildung 4: Tips beim Lotto 6 aus 49. Obere Reihe: von rechts nach links die Tips der Jahre 2008 - 2010 mit den 6 höchsten Gewinnerzahlen im Rang 2 (6 Richtige ohne Superzahl), nämlich 26, 16,15,14,13, und 12. Bei 26 Gewinnen betrug die Ausschüttung ca. 37.000 Euro . Die Gewinnsummen pro Tip betrugen 37, 89, 213, 169, 160, 84 tausend Euro entgegen einer üblichen Ausschüttung im einstelligen Millionenbereich. In der unteren Reihe sind 6 rein zufällige Tips (Ziehen ohne Zurücklegen) gezeigt.

erkennt in den subjektiven Tips diverse Muster. Ferner weisen die zufälligen Tips wesentlich mehr Nachbarpaare von Kreuzen auf, sowie Zahlen in den Randbereichen, usw. .

Verwunderung oder Entsetzen über Glücks- und Pechsträhnen beim Glücksspiel. Man kann aber mathematisch beweisen, daß unerwartet lange Serien sehr wahrscheinlich sind.

Zu diesem Thema gibt es eine Reihe systematischer Studien. B. Inhelder (1978) und J. Piaget and B. Inhelder (1978) z.B. untersuchen die Entwicklung des Wahrscheinlichkeitsgefühles bei Kindern. Sie folgern aus ihren Versuchsreihen, daß Kinder unter 11 Jahren mit dem Zufall noch gar nichts anfangen können (mir scheint diese Schranke persönlich recht hoch gegriffen, aber im Prinzip ist die Aussage sicher richtig). Die Kinder suchen andere - deterministische - Erklärungen, die z.B. auf ihrem Gerechtigkeitssinn oder ihren Vorlieben beruhen.

2.2 Zur Beherrschung des Zufalls

Genauer gehen Versicherungen mit dem Zufall um, und sie verdienen nicht schlecht dabei. Obwohl ein einzelner Autofahrer am 1. Januar weder weiß, ob er im folgenden Jahr einen Unfall verschulden wird und wie hoch dann die Schadensumme sein wird, können Versicherungen die Prämien sehr genau kalkulieren. Es muß also *Gesetzmäßigkeiten* geben, denen der Zufall gehorcht. Offensichtlich hat das bei den Versicherungsprämien damit zu tun, daß *sehr viele* Autofahrer versichert sind. In der Wärmelehre geht es noch extremer zu. Obwohl die Bewegung einzelner Moleküle sehr chaotisch verläuft und kaum vorherzusagen ist, können aufgrund der statistischen Eigenschaften makroskopische Größen wie Druck oder Temperatur hergeleitet werden. Diese benehmen sich - wie man z.B. in der Experimentalphysik lernt - vollkommen deterministisch, ähnlich wie Längen oder Massen, die *per se* deterministisch sind. Dies liegt wiederum am Zusammenwirken einer immens großen Zahl einzelner Moleküle; die Größenordnung ist etwa 10^{26}.

Somit kann sich die *Stochastik* nicht auf mathematische Aspekte beschränken. Sie will es auch nicht. Die eigentliche Stochastik rechtfertigt sich gerade aus dem Zusammenspiel von Beobachtung und Theorie, sowie der Rückwirkung der Theorie auf die Realität. In den Abbildungen 1 und 2 sind die Zusammenhänge anschaulich dargestellt.

3 Zur Geschichte

Wir befassen uns erst mit dem magischen und anschließend mit dem rationalen Aspekt des Zufalls. Es wird im wesentlichen um Würfelspiele gehen.

3.1 Würfeln: Magie und Spiel im Altertum

Die ersten nachgewiesenen Würfel-„Spiele“wurden mit „Astragali“(Abb. 5) durchgeführt. Wir fassen schematisch einige Angaben und Daten zusammen:

Zoologisch: Der Astragalus ist der oberste Fußgelenkknöchel am Hinterbein von Huftieren. Verwendet wurden vor allem Knochen von Schafen oder Hirschen.

Prähistorisch: Es gibt Gräberfunde aus der Zeit 20.000 - 30.000 v.Chr. Astragali wurden wohl zunächst meist als Amulett getragen, was man aufgrund der Bohrungen vermutet.

Historisch: Astragali wurden nachgewiesen u.a. in Babylon, Ägypten, Costa Rica und Sambia.

Spielregeln: Spielregen wurden wohl zuerst in Ägypten aufgestellt. Sie sind nicht direkt überliefert, sondern auf dem Umweg über Groß-Griechenland und von dort über die Römer, Araber und Spanier nach Frankreich und Italien.

Abbildung 5: Zwei Astragali. Abbildungen 5 und 6 aus HC. KÖRPER AND N.A. WHITNEY-DESAULTS (1999).

Das Spiel hatte magischen Charakter:

(a) Der Astragalus wurde in der Antike als Abbild des menschlichen Torsos empfunden.

(b) Ein Wurf-Ergebnis galt als Ausdruck des Willens der Götter.

(c) Astragali wurden vor allem von Priestern verwendet und gedeutet.

Die magische Bedeutung verhinderte eine Analyse, z.B. durch die hochentwickelte arabische Mathematik, vergleiche Abschnitt 3.2.
Eine empirische Statistik ist in folgender Tabelle zusammengefaßt; die konkreten Zahlen hängen von der Tierart und dem Individuum ab.

4 Seiten des Torsos, die oben liegen können	rechte Seite	linke Seite	Rücken	Bauch
antike Bezifferung	6	1	3	4
Wahrscheinlichkeiten	~ 7 %	~ 10 %	~ 35 %	~ 48%

Man beachte, daß schon hier die Summe der Augenzahlen gegenüberliegender Seiten 7 war. Man muß also nur noch die fehlenden Seiten mit den fehlenden Punkten 2 und 5 belegen, um im Prinzip zum heutigen Würfel zu kommen. Leichtes Anschleifen war in späterer Zeit erlaubt, was die Wahrscheinlichkeiten änderte. Besonders beliebt war das gleichzeitige Werfen von 4 Astragali mit 35 Kombinationsmöglichkeiten. In der Antike galt:

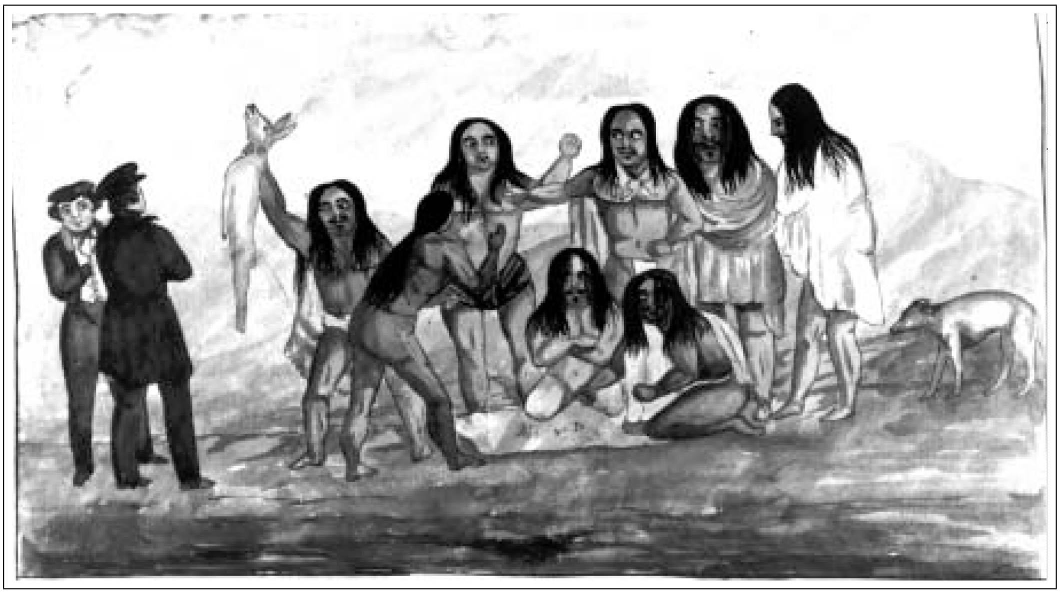


Abbildung 6: Indianer beim Würfeln mit Astragali, Aquarell.

Einschätzung	Bezeichnung	Kombination	p
bester Wurf	Venus = Eros	6, 1, 3, 4	p ∼ 2,8 %
schlechtester Wurf	Hund	1, 1, 1, 1	p ∼ 0,01 %
dagegen: seltenster Wurf	ohne Bezeichnung	6, 6, 6, 6	p ∼ 0,002 %

Kriterium für den Wert ist die *Harmonie*. Im Gegensatz hierzu sind bei heutigen Spielen die seltensten Kombinationen am wertvollsten.

Würfelspiele in verschiedenen Kulturkreisen manifestieren sich in vielen bildlichen Darstellungen. Abb. 6 zeigt Indianer beim Würfelspiel. In der Abb. 7 zeigt Breughel das Spiel als eine der zahlreichen (verwerflichen?) Zerstreuungen.

3.2 Religiöse Aspekte in nachchristlicher Zeit

Nachdem die Geometrie etwa in der griechischen und die Algebra in der arabischen Kultur einen hohen Reifegrad erlangten, stellt sich die Frage, warum trotz hochentwickelter Mathematik die Wahrscheinlichkeitsrechnung vor dem 17. Jahrhundert nicht vorkommt. Wir versuchen dies durch wenige Beispiele zu erklären.

In späterer, jüdisch-christlicher Zeit symbolisieren die Würfel die Erfüllung der Prophetie von Psalm 22.19 aus dem Alten Testament. In der LUTHERI-

Abbildung 7: Kinderspiele von Peter Breughel (Bruegel,Brueghel) (1528-1569), Pieter der Ältere (genannt Bauernbruegel), Vater von Jan Bruegel und Pieter Bruegel, gemalt 1560, Kunsthistorisches Museum Wien. Auf dem Bild sind zwei Frauen zu sehen, die mit Astragali spielen (wo ist der vierte?).

SCHEN Übersetzung lautet er:

> Sie teilen meine Kleider unter sich und werfen das Los um mein Gewand.

Die Erfüllung steht im Johannes Evangelium 19.24 im Zusammenhang mit der Kreuzigung Christi, unter dessen Kreuz die Kriegsknechte um sein Gewand gewürfelt haben sollen:

> Da sprachen sie untereinander:
> Lasset uns den nicht zerteilen[2], sondern darum losen, wes er sein soll.
> Auf daß erfuellet werde die Schrift, die da sagt[3]...
> Solches taten die Kriegsknechte.

[2]den ungenähten Rock nämlich, ebenda 19.23
[3]hier zitiert Luther Psalm 22.19

Auf zahlreichen Darstellungen der Kreuzigung Christi findet man deshalb neben Kreuz, Dornenkrone, Lanze, Schwammstab oder Nägeln auch Würfel dargestellt[4]. Ein gutes Beispiel ist ein Spätwerk des Meisters von Flémalle um 1440[5] Die Darstellung von Würfeln in Kreuzigungsszenen hat sich bis in die junge Vergangenheit erhalten. Zwei Wegekreuze sind in Abb. 8 dargestellt.

Das Spiel wurde zumindest in christlicher Zeit häufig durch Kirche und Staat verboten, z.B. durch Louis IX, genannt Ludwig der Heilige oder Saint Louis, 1254 in der „Grande ordonnance pur la réforme du royaume“: Er verbot sogar die Herstellung von Würfeln, zusammen mit Prostitution, Gotteslästerung und Wucher. Ludwig IX, 25. April 1214 - 25. August 1270 (gestorben bei Karthago) wurde 1297 heiliggesprochen, vgl. Abb. 9. Ebenfalls aus dem 12. Jahrhundert stammt ein Erlaß des englischen Königs Richard Löwenherz, daß niemand, der von geringerem Stand als ein Ritter war, um

[4] Ich danke Herrn Dr. Josef Fellermayr, der mich auf diese biblische Spur gebracht hat

[5] Leider ist es schwierig, solche Darstellungen ohne Urheberrechtsverlezunungen einzubinden

Abbildung 8: Auf beiden Wegkreuzen finden sich Würfel als Symbol der Prophezeihung aus Psalm 22.19 des Alten Testamentes und ihre Erfüllung, Johannes Evangelium 19.24. Die Reproduktion der Abbildungen erfolgt mit freundlicher Genehmigung von Herrn Dr. Ekkehard Kaier, `www.freiburg-schwarzwald.de`

Geld würfeln durfte. Es gibt auch ein Verbot im Koran, z.B. in den Suren 2:219, 5:90 („Werk des Satans“) oder 5:91, vgl Abb. 9. Diskussionen werden gegenwärtig in mehreren Internetforen geführt; es geht selbst darum, ob das Spiel „Mensch-Ärgere-Dich-Nicht“erlaubt sei. Es sei bemerkt, daß im Koran eher Empfehlungen gegeben werden. Die drastischen Verbote sind in Hadithe und Sunna ausgesprochen, also späteren Sammlungen.

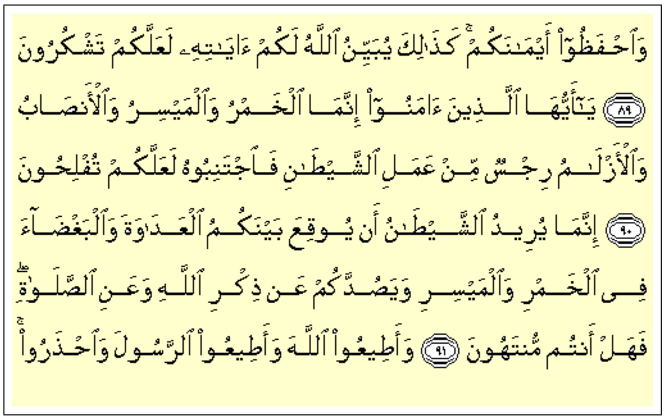


Abbildung 9: Der heilige Saint Louis IX und die Koran -Suren 5:89-91

Insgesamt wurden dem Glücksspiel hier schwere moralische und gesetzliche Hypotheken auferlegt. Dies erschwerte eine rationale Analyse des Gegenstandes erheblich.

3.3 Beginn der rationalen Analyse

Eine rationale Annäherung an zufällige Ereignisse mußte bis zur Zeit der Aufklärung, d.h. bis zur Mitte des 17. Jahrhunderts warten. Einen idealen Nährboden gab das exzessive Glücksspiel in den Salons der höheren Pariser Gesellschaft ab. Allerdings gibt es einige Vorläufer wie Cardano und Galilei, verschiedene Fehlinterpretationen setzten sich allerdings noch lange fort.

Noch D’ALEMBERT (1717-1783), siehe Seite 15, postulierte, daß bei zweimaligem Münzwurf die Kombinationen

„Kopf, Zahl“; „Kopf, Kopf“; „Zahl, Zahl“

jeweils die Wahrscheinlichkeit 1/3 hätten. Demselben Irrtum begegnen wir ein halbes Jahrhundert früher bei CHEVALIER DE MÉRÈ (1654), vgl. S. 21.

Jean-Baptiste le Rond, genannt d'Alembert (* 16. November 1717 in Paris; † 29. Oktober 1783 in Paris), Mathematiker und Physiker des 18. Jahrhunderts, Philosoph der Aufklärung. Mit Diderot Herausgeber des mathematischen Teiles der Encyclopédie.
Nach ihm benannt das D'Alembertsche Prinzip der Mechanik. Arbeitete in Funktionentheorie, löste 1747 die Gleichung der schwingenden Saite. Sein Quotientenkriterium wird auch D'Alembert-Kriterium genannt. Arbeiten in Wahrscheinlichkeitsrechnung.

Heute würden wir argumentieren: das Ereignis „Kopf, Zahl"(= „Zahl, Kopf") besteht aus den zwei zeitlich oder sonst irgendwie geordneten Paaren (Kopf,-Zahl) und (Zahl, Kopf) und hat somit die Wahrscheinlichkeit 1/4+1/4=1/2. Das stellt sich auch im Experiment als richtig heraus, wenn man gewöhnliche Münzen benutzt oder ein äquivalentes binäres Experiment durchführt. Also hat d'Alembert seine Behauptung wohl nie empirisch überprüft. Im Gegensatz dazu hat der Chevalier de Méré dies anhand penibler Aufzeichnungen von Spielverläufen beim Würfeln sehr wohl getan und damit die Diskussionen zwischen Fermat und Pascal angestoßen, siehe unten. Übrigens scheint bereits Galileo Galilei richtig gerechnet zu haben, er ging bei drei Würfeln schon von 216 Möglichkeiten aus.

Damit könnte man d'Alembert's Argument zur Seite legen und vergessen. Andererseits wäre sein Argument ja richtig, wenn man die Münzen der zwei Würfe wirklich nicht unterscheiden könnte. Dann würden ja die Fälle (Kopf, Zahl) und (Zahl, Kopf) zusammenfallen. In der Tat trifft man in der Quantenmechanik genau auf diese Situation. Photonen und Atome mit geradzahliger Anzahl von Teilchen - der Oberbegriff ist „Bosonen"- verhalten sich tatsächlich so, wie d'Alembert es angibt. Wir nehmen diesen Aspekt in Abschitt 3.4 noch einmal auf.

Im folgenden skizzieren wir in Form von Stichpunkten die Entwicklung der „modernen"Wahrscheinlichkeitstheorie. Ein Vorläufer ist Geronimo Cardano.

Geronimo Cardano, ∗24. September 1501 in Pavia, †21. September 1576 in Rom. Er stellte Tabellen über Astragali-Würfelkombination auf und zog teils richtige und teils falsche Folgerungen. Ratschläge für Erhöhung der ‘Chance’ des Spielers. ‘Chance’ ist keine Wahrscheinlichkeit, sondern ein gewisser Ausgang des Spieles, der in weiterem Spiel zu wiederholen ist. Man stellte fest, daß gewisse Chancen leichter zu erreichen waren, als andere. So wurde ‘Chance’ zu einem Gütebegriff. Siehe Liber de Ludo Alea, [3].

Als weiterer Vorläufer kann Galileo Galileo betrachtet werden. Jedenfalls ging er Probleme bei Würfelspielen mit rationalen Methoden an und löst sie anscheinend richtig.

Galileo Galilei, ∗ 15. Februar 1564 in Pisa, † 8. Januar 1642 in Arcetri bei Florenz. Er schrieb ‘De Motu’, mit der Idee, daß Theorien durch Experimente prüfbar sind. Baute Teleskope zwanzigfacher Vergrößerung und entdeckte die Bewegungs- und Fallgesetze. Er argumentierte, daß die Physik mit exakten Messungen anstatt mit metaphysischen Prinzipien begründete werden, und daß die Bibel nicht wörtlich ausgelegt werden solle, wo sie wissenschaftlichen Fakten widerspricht. Er propagierte die kopernikanische Theorie als physikalische Wirklichkeit.
Er lieferte die richtige Lösung einer umstrittenen Würfelaufgabe! Dies war eine gut bezahlte Auftragsarbeit, ansonsten hatte er kein weiteres Interesse daran.

Jetzt wurde es ernst mit der Wahrscheinlichkeitsrechnung. Ein Dreiergespann, bestehend aus zwei Mathematikern und Philosophen, sowie einem Berufsspieler, nahm sich der Sache an.

Blaise Pascal ∗ 19 Juni 1623 in Clermont, Auvergne, Frankreich, † 19. August 1662 in Paris. Pascal war nicht der erste, der das ‘Pascalsche Zahlendreieck studierte. Trotzdem war seine Arbeit dazu einer der wichtigsten Beiträge dazu. In seiner Korrespondenz mit Fermat legte er die Grundlagen der heutigen Wahrscheinlichkeitstheorie. Es handelt sich um fünf Briefe im Sommer 1654. Es ging um das Würfelproblem, das schon von Cardano studiert wurde.

Der Berufsspieler Chevalier de Méré fragt seinen Freund Fermat, weil “L’arithmétique se démentait”; er versteht die Ergebnisse mancher Würfelspiele nicht (siehe Ende des Abschnittes). Dies führt im Jahre 1654 zu einem intensiven Briefwechsel zwischen Pierre de Fermat und Blaise Pascal. Die beiden entdecken de Mérés Fehler und schaffen gleichzeitig die Grundlagen für den noch heute gültigen Umgang mit der elementaren Wahrscheinlichkeitsrechnung. Die eigentliche Mathematik dabei besteht aus simpler Kombinatorik; entscheidend ist, das richtige mathematische Modell aufzustellen.

Pierre de Fermat (1601-1665). Fermat ist einer der zur Zeit in den Medien am meisten genannte Mathematiker, wegen des von ihm formulierten 'Fermatschen Problemes'. Es wurde erst 1993 von Andrew Wiles und Richard Taylor bewiesen (1995 veröffentlicht).

In der Folge befaßten sich eine Reihe genialer Mathematiker mit der Wahrscheinlichkeitsrechnung, u.a Christiaan Hyg(h)ens.

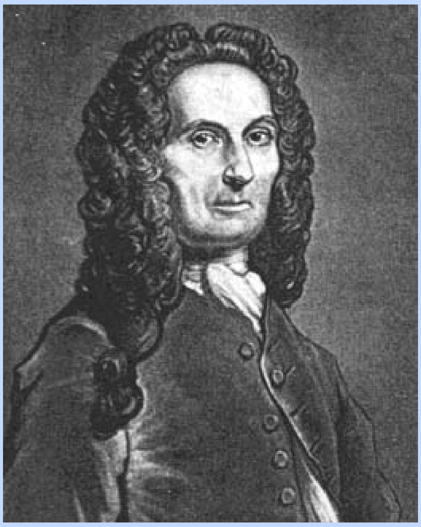
Christiaan Hyg(h)ens (1629 - 1695). Er versucht vergeblich, Schüler von Pascal zu werden. Er stellt selbständig drei Prinzipien der Chancenberechnung auf, die weitgehend die spätere Definition von Laplace vorwegnehmen. Er löst Würfelaufgaben richtig, aber sehr umständlich, da die mathematischen Hilfsmittel noch zu wenig entwickelt sind.

Es zeigt sich, daß in dem Briefwechsel zwischen Pascal und Fermat noch kein eigenständiger Wahrscheinlichkeitsbegriff konstituiert wurde. Dieser wurde erst von Jakob Bernoulli (1656-1705) und Gottfried Wilhelm Leibniz (1646-1716) explizit eingeführt. Ein wichtiges Beispiel der ersten

Lehrbücher ist die 'Ars conjectandi', JAKOB BERNOULLI (2002).

Jakob Bernoulli (1656-1705). Schweizer Mathematiker, Physiker, Autor von JAKOB BERNOULLI (2002), 'Ars conjectandi' (lateinisch: Die Kunst des Mutmaßens, Vermutens, Schließens), setzt den Wahrscheinlichkeitsbegriff als Quotient der Chancen. Mit stärkeren Hilfsmitteln kommt man zum sogenannten Theorem von Bernoulli.

Darauf konnte MARQUIS DE LAPLACE aufbauen, indem er den damaligen Entwicklungsstand der Differential- und Integralrechnung nutzte. Er nahm bewußt den subjektivistischen Standpunkt ein. Er schrieb ein modernes Werk über die Wahrscheinlichkeitrechnung, M. LE COMTE LAPLACE (1812).



Abraham de Moivre (1667 - 1754). Er entwickelte bereits eine geschlossene Theorie (Doctrine of Chances 1711?, 1718, 1738) der Wahrscheinlichkeit.

Auf das vorhergehende äußerst wichtige und einflußreiche Werk von ABRAHAM DE MOIVRE werden wir in Abschnitt 6 ausführlicher zurückkommen.

Pierre Simon Marquis de Laplace (1749-1827). Er schrieb den 'Traité analytique des probabilités', 1812, M. Le Comte Laplace (1812); das ist ein deduktives, rein mathematisches Werk im modernen Sinne.

Die spätere Entwicklung, etwa im zwanzigsten Jahrhundert, kann hier nicht gewürdigt werden. Wir beschränken uns auf wenige richtungsweisende Impulse, vertreten durch Namen. Zunächst ist Louis Bachelier zu erwähnen, der über Finanzmärkte arbeitete. Seine Doktorarbeit *Théorie de la spéculation* (1900), [1] wird heute hoch gefeiert. Er eröffnete den Weg zur modernen Martingaltheorie und gilt als ein Urheber des Wiener-Prozesses, der in der Mathematik auch als Brownsche Bewegung bekannt ist (Physiker verstehen unter letzterem meist etwas anderes). Der Durchbruch wird Albert Einstein zugeschrieben durch seine epochale Arbeit *Über die von der molekularkinetischen Theorie der Wärme gefordete Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, [9]. Dies geschah in Unkenntnis der Arbeit Bacheliers.

Wichtige weitere Schritte in der Theorie stochastischer Prozesse und der Stochastischen Analysis sind mit Namen wie Paul Lévy (1886 - 1971) und Kiyosi Itô (1915 - 2008), auf gleichem Rang vor den späteren, verbunden. Aufgrund der rasanten Entwicklung der Wahrscheinlichkeitstheorie wird es ab nun unübersichtlich. Zu erwähnen sind Entwicklungen wie die Vielteilchentheorie oder die Entdeckung der virtuellen Handlungsmöglichkeiten (etwa Optionen). Aber das ist eine andere Geschichte.

Louis Bachelier (1870-1946) arbeite über die mathemathische Theorie der Finanzmärkte. Sein bekanntestes Werk ist seine Doktorarbeit *Théorie de la spéculation* (1900), [1].

3.4 Diskussion des Beispieles von Chevalier de Méré

Einer unter mehreren Auslöser des Briefwechsel zwischen Pascal und Fermat war die folgende Frage: Chevalier de Méré wunderte sich: Beim gleichzeitigen Wurf mit 3 Würfeln ergibt sich die Augensumme 12 durch die 6 Kombinationen

$$6\,5\,1 \quad 6\,4\,2 \quad 6\,3\,3 \quad 5\,5\,2 \quad 5\,4\,3 \quad 4\,4\,4$$

und die Augensumme 11 ebenfalls durch 6 Kombinationen, nämlich

$$6\,4\,1 \quad 6\,3\,2 \quad 5\,5\,1 \quad 5\,4\,2 \quad 5\,3\,3 \quad 4\,4\,3.$$

In seinen Augen waren die Kombinationen gleich wahrscheinlich, also auch die 11 und die 12 als Augensumme. In der Praxis stellte er aufgrund seiner peniblen schriftlichen Aufzeichnungen von Spielergebnissen fest, daß die 11 öfter auftritt, als die 12. Er ist also dem selben Irrtum wie d'Alambert aufgesessen, hat aber im Gegensatz zu diesem Versuche durchgeführt, welche ihn zweifeln ließen. Pascal und Fermat lösten dann das Problem. Dies ist beispielhaft für wissenschaftliches Vorgehen. Wir erkennen folgende Schritte:

- Aufstellen eines Modells (durch Chevalier de Méré),

- empirische Datenerhebung (durch Chevalier de Méré),

- Anzweifeln des Modelles aufgrund der Daten (durch Chevalier de Mére), Ersetzen durch ein neues Modell (durch Pascal und Fermat),

- Prüfung des neuen Modelles anhand von Daten,

- daraus folgend Auswirkung auf die praktischen Handlungen, z.B. Änderung der Wettstrategie, oder zumindest deren schlüssige Begründung..

Im klassischen Laplace Modell lösen sich die Kombinationen auf in

Augensumme 12	6+6+3+3+6+1	=25 Fälle
Augensumme 11	6+6+3+6+3+3	=27 Fälle ’

da die Permutationen der Ergebnisse gezählt werden. Es gibt nämlich $6^3 = 216$ ‘mögliche’ und 25 bzw. 27 ‘günstige’ Fälle. Das ergibt für die Augensumme 11 die Wahrscheinlichkeit 1/8= 0,125 und für die 12 die Wahrscheinlichkeit 0,1157. Die Differenz beträgt lediglich 0,00993, also ca. 1 %.

Übrigens scheint bereits Galileo Galilei richtig gerechnet zu haben, er ging ein halbes Jahrhundert vorher bei drei Würfeln schon von 216 Möglichkeiten aus. Allerdings ist seine Arbeit nicht publiziert, da es sich um eine lukrative Auftragsarbeit handelte, die geheim bleiben mußte.

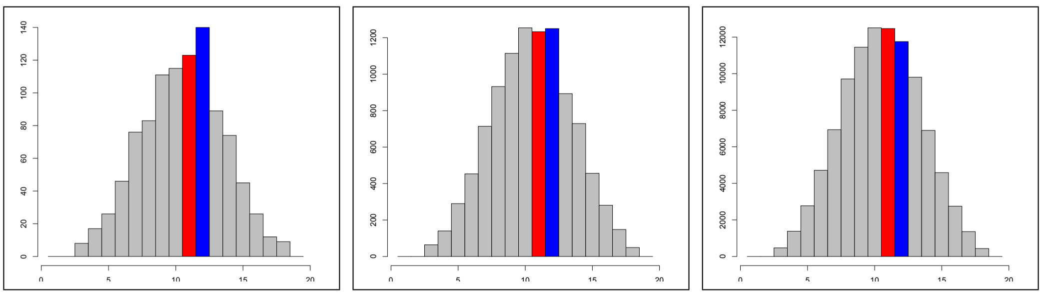


Abbildung 10: Histogramme für die Augensummen bei drei Würfen für 1000, 10.000 und 100.000 Wiederholungen. Diese sind zufallsabhängig. Es wurden bewußt negative Ergebnisse ausgewählt. Die 11 ist rot, die 12 blau eingefärbt

Bemerkung Es ist aufschlußreich zu erkunden, wie oft der Chevalier de Méré wohl gespielt und aufgezeichnet haben muß, um den Unterschied zwischen den Wahrscheinlichkeiten der Augensumme 11 und 12 zu entdecken. Ausgewählte Simulationen ergeben die Histogramme in Abb. 10

Um ein Gefühl für den notwendigen Stichprobenumfang zu bekommen, wurden für einen Signifikanztest, genauer einen t-Test, die (uns bekannten)

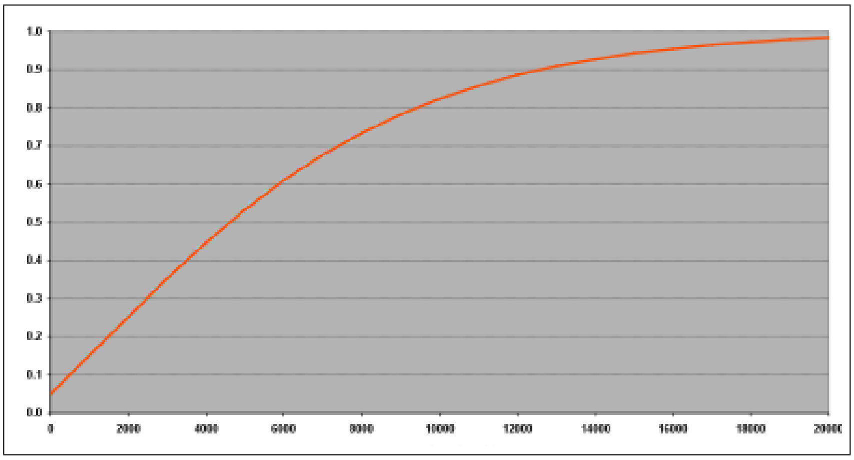


Abbildung 11: Power des t-Tests aus der Bemerkung in Abhängigkeit vom Stichprobenumfang

Wahrscheinlichkeiten für die Augensummen 11 und 12 vertauscht: Es wurde getestet die Nullhypothese $\mathcal{H}_0$ gegen die Alternative $\mathcal{H}_1$, wobei

$$\mathcal{H}_0 : \mathbb{P}(\text{Augensumme} = 12) = 1/8, \quad \mathcal{H}_1 : \mathbb{P}(\text{Augensumme} = 12) = 0,1157$$

Wir wissen, daß die Nullhypothese falsch und die Alternative richtig ist. Somit ist es erhellend zu wissen, mit welcher Wahrscheinlichkeit die Nullhypothese unter der (richtigen) Alternative abgelehnt wird. Die Wahrscheinlichkeiten sind in Abb. 11 in Abhängigkeit von der Zahl N der Wiederholungen des Spieles aufgetragen. Inspektion des Plots zeigt, daß für eine zuverläßige Ablehnung, z.B. mit Wahrscheinlichkeit 0,95, ca. 15.500 Spiele nötig sind. Der *chevalier* muß also sehr fleißig gespielt und bewundernswerte Aufzeichnungen angelegt haben [6].

Es ist aufschlußreich, die Frage der Modellbildung anhand dieses Beispieles zu überdenken. Wären die Würfel tatsächlich ununterscheidbar, so hätten DE MÉRÉ und D'ALAMBERT tatsächlich recht. Sie sind es natürlich nicht. Das Experiment besteht ja nicht nur aus den Würfeln, sondern auch ihrer Lage im Becher und dem Werfen, also der Funktionsweise des Armes.

Andererseits gibt es Gegebenheiten, bei denen Ununterscheidbarkeit, wie bei D'ALEMBERT oder DE MÉRÉ die richtige Annahme ist. In der Tat trifft man in der Quantenmechanik genau auf diese Situation.

[6] Ich danke Herrn Hagen Scherb für die Idee zu den Simulationen in den Abbildungen 10 und 11, und Herrn Karsten Rodenacker für seine Hilfe bei der Graphik.

Wir betrachten r physikalische Teilchen. Man unterteilt den Phasenraum in eine (große) Zahl n von Zellen; jedes Teilchen ist einer Zelle zugeordnet. Wie viele der so entstehenden Konfigurationen gibt es, welche Wahrscheinlichkeit haben sie unter Gleichverteilung? Man muß die Anzahl der möglichen Belegungen der Zellen berechnen.

Sind die Teilchen *unterscheidbar*, so gibt es n^r Möglichkeiten, d.h. die Einzelwahrscheinlichkeiten sind n^{-r}. In der statistischen Physik spricht man von einer *Maxwell-Boltzmann Statistik*. Diese trifft allerdings auf keine der bekannten Partikel zu.

Sind die Teilchen *ununterscheidbar*, so interessieren nur die Anzahlen von Teilchen in den jeweiligen Zellen etwa:

$$| * * * | * |||| * * * * |, \quad r = 8,\, n = 6. \tag{1}$$

Gemäß der symbolischen Darstellung in (1) müssen aus $n - 1 + r$ Stellen diejenigen r auswählen, an denen ein $*$ steht. Dies geht auf

$$A_{r,n} = \binom{n + r - 1}{r}$$

Arten. Die Wahrscheinlichkeiten sind $A_{r,n}^{-1}$ und man spricht von einer *Bose-Einstein Statistik* und *Bosonen*. Dies sind Photonen und Atome mit geradzahliger Anzahl von Elementareilchen.

Schließlich gibt es noch egoistische Teilchen, von denen sich höchstens eines in einer Zelle befinden kann. Hier muß man aus den n Zellen die r besetzten auswählen, wobei natürlich $r \leq n$ gelten muß. Es gibt

$$\binom{n}{r}$$

Möglichkeiten. Man spricht von der *Fermi-Dirac Statistik* und *Fermionen*. Dies sind Elektronen, Neutronen und Protonen. Nach letzterer Verteilung sind z.B. auch Druckfehler verteilt.

Zu diesem Themenkreis lese man bei W. Feller (1968) nach. Offensichtlich kommt man zum verschiedenen statistischen Verhalten etwa von Photonen und Protonen nicht ohne weiteres über *a priori* Überlegungen.

4 Zum Wesen des Zufalls

Innerhalb der verschiedenen Sichten und Theorien über die Welt kann man über das *eigentliche Wesen* des Zufalles nachdenken.

4.1 Es gibt keinen echten Zufall

Begreift man die Welt als Maschine, die den klassischen Gesetzen der Physik - also der Newtonschen Mechanik - gehorcht, so kann der Ablauf des Geschehens im Prinzip als Lösung eines (nicht gerade kleinen) Systems von *Hamiltonschen Differentialgleichungen* der genial einfachen Gestalt

$$\frac{\partial H}{\partial p} = \dot{q}, \quad \frac{\partial H}{\partial q} = -\dot{p},$$

mit dem Ort q, dem Impuls p und der Hamiltonfunktion $H(p, q)$ - also der Energie - zusammen mit den Anfangsbedingungen, erklärt werden. Danach läuft das Weltgeschehen - gegeben einen Anfangszustand - vollkommen deterministisch ab[7]. Dies gilt sowohl in der Zeit vorwärts, als auch rückwärts.

Diese Theorie gilt auf der „makroskopischen Skala“und ist somit eigentlich für unseren Alltag zuständig, wenn man nicht gerade mit moderner Elektronik oder Lasertechnik konfrontiert ist.

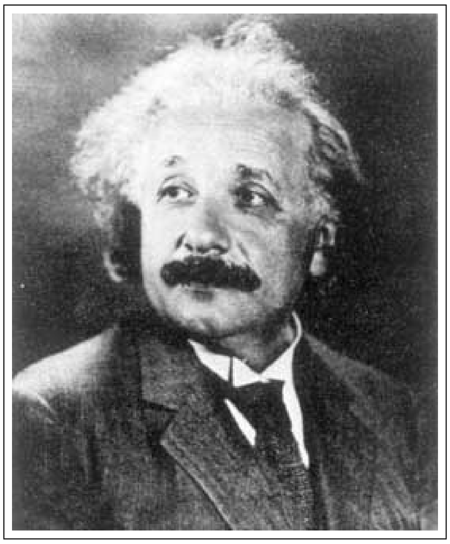
In einem solchen System hat der Zufall zunächst keinen Platz. Das deckt sich mit der Auffassung von René Descartes (1596-1650) (lateinisch: Renatus Cartesius), der die Welt als Maschine - analog einem Uhrwerk - versteht. Selbst Albert Einstein (1879-1955) wollte nicht an Zufall glauben; berühmt ist sein Ausspruch „Gott würfelt nicht“. Er drückte damit seine Skepsis gegenüber der Wahrscheinlichkeitsinterpretationen von Modellen der Quantenmechanik aus.

Eine solche Weltauffassung wird durch das folgende Zitat von Laplace treffend charakterisiert:

> Wir müssen also den gegenwärtigen Zustand des Universums als Folge eines früheren Zustandes ansehen und als Ursache des Zustandes, der danach kommt. Eine Intelligenz, die in einem gegebenen Augenblick alle Kräfte kennt, mit denen die Welt begabt ist, und die gegenwärtige Lage der Gebilde, die sie zusammensetzen, und die überdies umfassend genug wäre, diese Kenntnisse der Analyse zu unterwerfen, würde in der gleichen Formel die Bewegungen der größten Himmelskörper und die des leichtesten Atoms einbegreifen. Nichts wäre für sie ungewiß, Zukunft und Vergangenheit lägen klar vor ihren Augen.

Die angesprochene „Intelligenz“ ist heute unter dem Begriff „Laplacescher Dämon“ bekannt.

[7]Da stellt sich natürlich sofort die Frage nach der Willensfreiheit etc. Aber das ist eine andere Geschichte

	
René Descartes (1596-1650)	Albert Einstein (1879-1955)

4.2 Es gibt (vielleicht) doch echten Zufall

Den einzigen möglicherweise „echten Zufall“gibt es im sehr kleinen Bereich, also in der Quantenmechanik. Er ist Bestandteil der heute akzeptierten physikalischen Theorie. Das ist festgelegt in der *Kopenhagener Deutung der Quantentheorie*, siehe W. Heisenberg (2000). Dieser Zufall beruht auf der Unbestimmtheitsrelation, heute unter dem Namen *Heisenbergsches Unschärfeprinzip* bekannt, vgl. (2). Ob ein radioaktives Isotop in der nächsten Minute zerfällt oder nicht, ist per Postulat der Quantenmechanik zufällig. Ähnlich ist es mit der Streuung im Doppelspaltexperiment oder im Stern-Gerlach Experiment. Aus solchen Experimenten werden auch Daten gewonnen, die als „echte Zufallszahlen“dienen. Sie werden für spezielle Simulationen verwendet.

In der Quantenmechanik dienen *Wellenfunktionen* zur mathematischen Beschreibung des Zustandes eines physikalischen Systems. Dies sind komplexwertige quadratintegrierbare Funktionen ψ auf dem Ortsraum der Vektoren q. Normalerweise wird ψ auf $\int \psi\psi^* \, dq = 1$ normiert, wobei ψ^* die komplex konjugierte Funktion von ψ ist. In der schrödingerschen Quantenmechanik ergeben sich Wellenfunktionen als Lösungen der Schrödingergleichung. Im Unterschied zur Klassischen Physik ist eine exakte Aussage z.B. über den Aufenthaltsort eines Teilchens im allgemeinen nicht möglich. Das Betragsquadrat $|\psi|^2 = \psi\psi^*$ wird als Dichte der Wahrscheinlichkeit für den Ort des Teilchens interpretiert. Der Raumpunkt der klassischen Mechanik wird also durch eine Wahrscheinlichkeitsdichte ersetzt. Die Fouriertransformierte

$\hat{\psi}$ dient als Beschreibung des Impulses. *Eine* Form der Unschärferelation schreibt sich

$$\sigma_q \sigma_p \geq \frac{h}{4\pi}, \tag{2}$$

wobei h das Plancksche Wirkungsquantum bezeichnet. σ_q und σ_p sind die zu ψ bzw. $\hat{\psi}$ gehörigen Streuungen.

Allerdings haben viele Physiker Probleme mit einer wahrscheinlichkeitstheoretischen Interpretation. Es gibt Bemühungen, den Zufall aus der Quantenmechanik zu vertreiben. Die entsprechenden Stichworte sind *verborgene Parameter* oder *Bohmsche Mechanik*. Eine Diskussion dieser Begriffe geht weit über den Rahmen des vorliegenden Textes hinaus. Allerdings gibt es hervorragende Webseiten zu diesem Thema, insbesondere von einer Gruppe am mathematischen Institut der Ludwig-Maximilians Universität in München.

5 Deterministischer Zufall

Denken wir an einen Münzwurf im Rahmen der Newtonschen Mechanik, so kommt der „Zufall"ins Spiel, weil das System, bestehend aus dem menschlichen Körper, dem Würfel und dem Tisch, aufgrund seiner Komplexität praktisch nicht mit Differentialgleichungen behandelbar ist. Nicht einmal die Anfangsbedingungen können präzise bestimmt werden. Es handelt sich hier nicht um echten Zufall, sondern um unvollständiges Wissen, weshalb Prognosen unmöglich sind.

5.1 Dynamische Systeme

Iteriert man nichtlineare Funktionen, so entstehen in vielen Fällen vollständig determinierte Zahlenfolgen, welche mit zufälligen aber viele Eigenschaften gemeinsam haben.

Ein erstes Beispiel ist die Bahn eines Punktes unter der folgenden besonders einfachen (nichtlinearen!) Funktion. Es sei

$$f : [0,1] \longrightarrow [0,1], \;\; x \longmapsto f(x) = \begin{cases} 2x & \text{falls } 0 \leq x < 1/2 \\ 2x - 1 & \text{falls } 1/2 \leq x \leq 1 \end{cases} . \tag{3}$$

Der Graph ist in Abb. 12 links skizziert. Für fast alle x (außer denen an diadischen Punkten) ist die Folge

$$x,\; f(x),\; f^{(1)}(x),\; f^{(2)}(x),\; \ldots\;,\; f^{(n)}(x),\; \ldots$$

der Iterierten $f^{(n)}(x) = f \circ \cdots \circ f(x)$ über dem Einheitsintervall gleichmäßig verteilt, und man kann aus dem Wert $f^{(n)}(x)$ für große n kaum auf x zurück

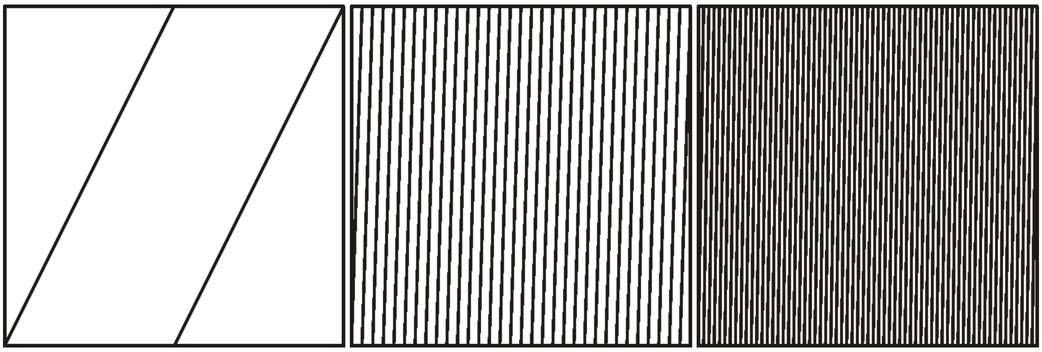


Abbildung 12: Graphen der Abbildungen f, $f^{(4)}$, $f^{(5)}$

schließen. Anschaulich kann man sich das mit Hilfe von Abb. 12, rechts, anschaulich machen. Ein heuristisches Argument findet sich im nächsten Abschnitt.

Ein interessanter Aspekt dieser Abbildung ist, daß ihre Anwendung das Komma in der Binärdarstellung des Argumentes eine Stelle nach rechts rückt (und eine eventuelle 1 links vom Komma in eine 0 verwandelt). Dies erleichtert die Berechnung der Bahn erheblich. Deshalb heißt die Abbildung 3 auch *bit-shift-map*.

Ein weiteres populäres Beispiel ist die *logistische Abbildung*

$$g_\alpha : [0,1] \longrightarrow [0,1], \quad x \longmapsto \alpha x(1-x), \quad \alpha \in (0,4]. \tag{4}$$

Plottet man für jedes α ein Stück der Bahn, welches sehr spät anfängt, so erhält man den wohlbekannten „Feigenbaum“, vgl. Abb. 13.
Für kleine α hat f einen stabilen Fixpunkt, auf den hin sich die Bahn zunehmend konzentriert. Bei wachsendem α ergeben sich stabile Perioden der Längen 2^n in aufsteigender Folge. Auf der rechten Seite herrscht Chaos. Allerdings erkennt man auch dort α mit stabiler Periode 3 oder 6. Bei $\alpha = 4$ folgen die Bahnpunkte für fast alle Anfangswerte einer kontinuierlichen Verteilung auf ganz $(0,1)$ mit Dichte

$$\varphi(x) = \frac{1}{\pi\sqrt{x(1-x)}}, \quad x \in (0,1). \tag{5}$$

Von dort nach links verdoppelt sich dies wiederum. Insgesamt ist das Verhalten höchst kompliziert. Man kann (5) als Transformierte der Gleichverteilung auffassen. In der Tat: bezeichnen wir die Bahnpunkte von (3) mit x_n und die von g_4 aus (4) mit y_n, so gilt

$$y_n = \sin^2(2\pi x_n)$$

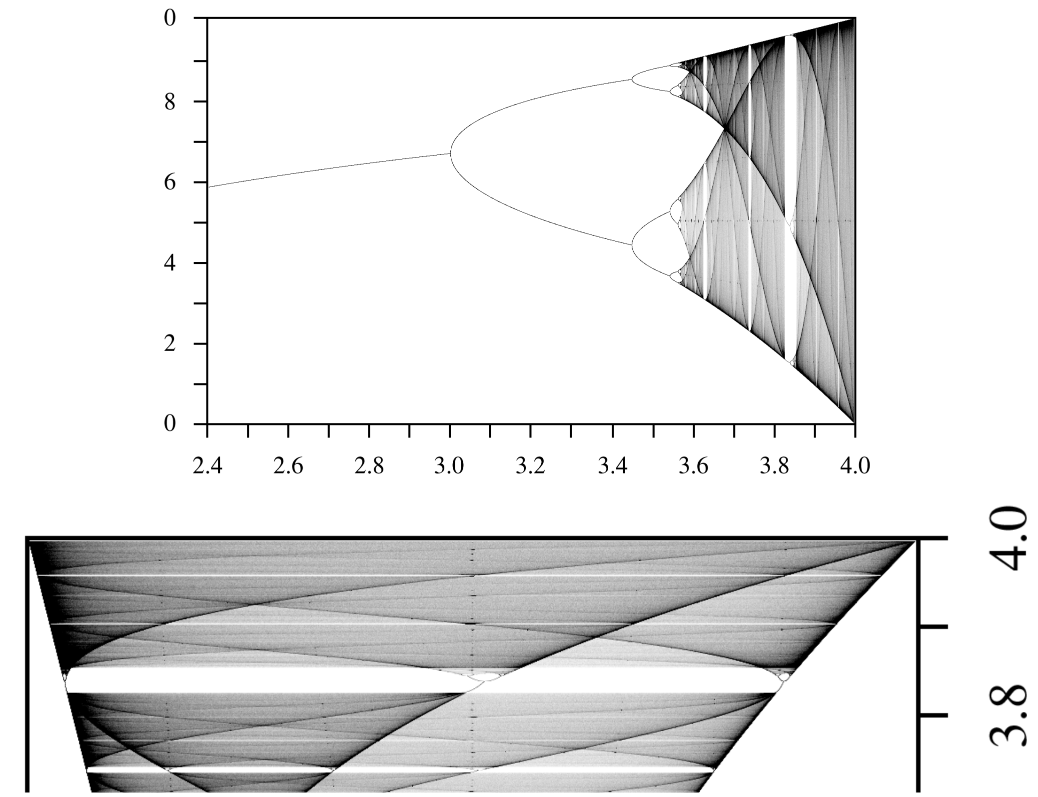


Abbildung 13: Ein Feigenbaumdiagramm, Detail vom rechten Rand

und wir werden in Abschnitt 5.2 argumentieren, daß fast alle Bahnen unter der Abbildung in (3) einer Gleichverteilung auf $(0,1)$ folgen.

Bemerkung 5.1 Die Definition (4) leitet sich von einem simplen Model für die Entwicklung einer Population ab, nämlich

$$y_{n+1} = \mu y_n - b y_n^2, \tag{6}$$

wobei μ die Wachstumsrate von einer Generation zur nächsten ist und b den Wettbewerb zwischen den Individuen widerspiegelt. Unter der Transformation $x_n = y_n b/\mu$ ergibt sich (4).

Ein einfaches aber attraktives (im wahrsten Sinn des Wortes) Beispiel mit kontinuierlichem Zeitparameter ist der *Lorenz Attraktor*. Er beruht auf dem formal einfachen (meteorologischen) System von Differenzialgleichungen:

$$\dot{x} = \sigma(y-x), \quad \dot{y} = x(r-z) - y, \quad \dot{z} = xy - bz.$$

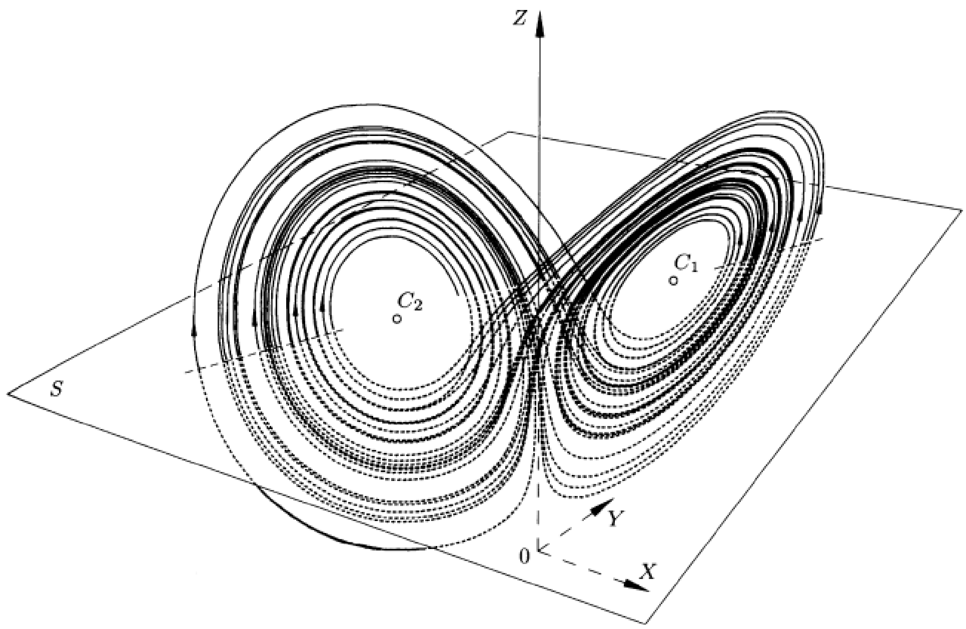


Abbildung 14: Der Lorenz Attraktor

Entscheidend ist wieder die Nichtlinearität des Systems und die Tatsache, daß in kontinuierlicher Zeit Chaos erst ab Raumdimension 3 auftritt. Die Lösungen nähern sich schließlich einem Attraktor, der in Abb. 14 skizziert ist. Man ahnt, daß bei minimaler Verschiedenheit der Anfangsbedingungen die Lösungen sich völlig verschieden verhalten. Dies wird gerne als ein Hinweis auf den wohlbekannten *Schmetterlingseffekt* benutzt.

Jedenfalls ist dies ein weiteres Indiz dafür, daß zukünftige Zustände auch bei deterministischen Systemen nicht vorhergesagt werden können und man in diesem Sinne einer Form des Zufalls gegenübersteht.

5.2 Pseudozufallszahlen

Pseudozufallszahlen sind für die Simulation stochastischer Vorgänge unabdingbar. Es sind Folgen von Zahlen, die durch deterministische Algorithmen erzeugt sind, sich jedoch in der Praxis wie zufällige verhalten, d. h. entsprechende Tests passieren. Sie beruhen auf den eben erwähnten Phänomenen.

Solche Generatoren sollten wenigstens die folgenden Bedingungen erfüllen:

(1) eine gute Approximation an auf dem Einheitsintervall $[0, 1)$ gleichverteilte Zufallsvariablen darstellen,

(2) in etwa unabhängig sein,

(3) leicht, schnell und exakt generiert werden können.

Die Entwicklung entsprechender Algorithmen ist trickreich. Wir zitieren aus B.D. Ripley (1988) §5:

> The whole history of pseudo-random numbers is riddled with myths and extrapolations from inadequate examples. A healthy scepticism is needed in reading the literature.

und aus §1 derselben Referenz:

> S.K. Park and K.W. Miller (1988), comment that examples of good generators are hard to find .... Their search was, however, in the computer science literature, and mainly in texts at that; random number generation seems to be one of the most misunderstood subjects in computer science!

Üblicherweise erzeugt man rekursiv Zahlenfolgen

$$u_0 = \texttt{Saat}, \;\; u_{k+1} = f(u_k)$$

mit einer $\texttt{Saat} \in [0, 1)$.

Die *lineare Kongruenzmethode* ist die populärste Methode zur Erzeugung von Pseudozufallszahlen. Dazu betrachtet man zunächst die Funktion

$$f : [0, 1) \to [0, 1), \quad f(u) = (au + b) \bmod 1 \tag{7}$$

mit natürlichen Zahlen a and b; $v \bmod 1$ ist die Differenz von v und seinem ganzzahligen Anteil. Die Funktion (3) kann man in dieser Form als $f(u) = 2u \bmod 2$ schreiben. Der Graph solcher Funktionen f besteht aus a Geradensegmenten mit Steigung a. Der Punkt (3) ist für solche Generatoren sicherlich erfüllt.

Wir geben einige informelle Argumente, daß auch (1) und (2) approximativ gelten.

Behauptung: *Seien I und J in $[0, 1)$ Intervalle der Länge $\lambda(I)$ und $\lambda(J)$, die wesentlich größer sind als a^{-1}.* Wir nehmen an *daß u_k auf $[0, 1)$ gleichverteilt ist. Dann gilt:*

$$\mathbb{P}(u_{k+1} \in J \,|\, u_k \in I) \sim \mathbb{P}(u_{k+1} \in J) \sim \lambda(J).$$

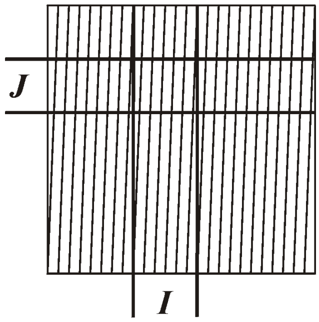


Abbildung 15: Graph einer Abbildung $f(u) = au + b \bmod 1$

Die zweite Beziehung würde bedeuten, daß u_{k+1} auf $[0,1)$ ungefähr gleichverteilt ist und die erste daß die Verteilung nicht von der Position von u_k abhängt.

Dies sieht man so ein: Nehmen wir der Einfachkeit halber $b = 0$ an. f ist linear auf den a elementaren Intervallen $I_k = [k/a, (k+1)/a)$, $0 \leq k < a$. Das Urbild $f^{-1}(J)$ besteht aus a Intervallen $I_k \cap f^{-1}(J)$, jeweils der Länge $\lambda(J)/a$. Sei I_e die Vereinigung aller $I_k \subset I$ und n die Anzahl dieser I_k. Dann gilt

$$\lambda(f^{-1}(J) \cap I) \sim \lambda(f^{-1}(J) \cap I_e) = \frac{n}{a}\,\lambda(J) \sim \lambda(I)\lambda(J)$$

und somit

$$\mathbb{P}(u_{k+1} \in J \,|\, u_k \in I) = \frac{\mathbb{P}(u_{k+1} \in J, u_k \in I)}{\mathbb{P}(u_k \in I)} \sim \frac{\lambda(I)\lambda(J)}{\lambda(I)} = \lambda(J).$$

Somit ist die Aussage für große n (oder a) richtig.

In der Praxis arbeitet man nur mit natürlichen Zahlen; man setzt

$$v_0 = \texttt{seed}, \;\; v_{k+1} = (av_k + b) \bmod c \tag{8}$$

mit einem Multiplikator a, einem Shift b und einem Modulus c, wobei alle Zahlen natürlich sind. Man startet mit einer Saat *seed* $\in \{0, 1, \ldots, c-1\}$ ($n \bmod c$ ist die Differenz von n und dem größten Vielfachen von c less or equal to n). Die so erhaltene Folge normiert man auf $\{0, 1, \ldots, c-1\}$ und erhält

$$u_k = \frac{v_k}{c}.$$

Die Folgen (v_k) bzw. (u_k) sind natürlich periodisch mit Periode höchstens c.

Beispiel 5.1 Die Wahl der Parameter ist eine eigene Wissenschaft. Bekannt ist der sehr schlechte Generator RANDU von IBM: $v_{k+1} = (2^{16} + 3)v_k$ mod

2^{31}; sukzessive Tripel (v_k, v_{k+1}, v_{k+2}) liegen auf nur 15 Hyperebenen, man vergleiche [42], Seite. 23, [26] oder [17]. Abb. 16 illustriert dieses Phänomen. In der Literatur wird u.a. der Generator $(69069\,v + 1) \bmod 2^{32}$ aus G. Marsaglia (1972) empfohlen.

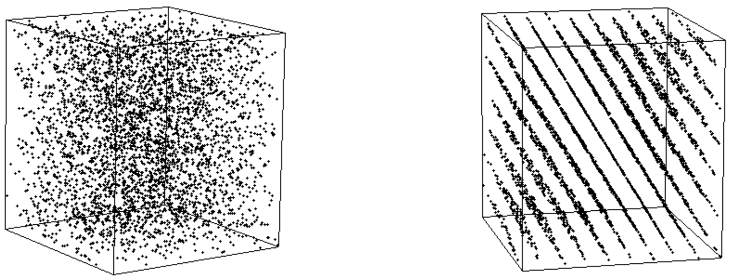


Abbildung 16: SukzessiveTripel von RANDU aus zwei Perspektiven

Abb. 16 illustriert die Aussage des Satzes von Marsaglia (1968): Bildet man die Zufallszahlen u_k/c mittels eines linearen Kongruenzgenerators der Form (8) und sodann sukzessive r-Tupel $(v_i, v_{i+1}, \ldots, v_{i+r-1})$, so liegen diese Punkte im $\mathbb{R}^r$ auf höchstens $(cr!)^{1/r}$ Hyperebenen.

5.3 Eine Abweichung: der Satz von Šarkovskii

Dieser Abschnitt beschäftigt sich nicht eigentlich mit dem Zufall. Er wirft aber ein Schlaglicht auf die Phänomene, die in Abschnitt 5.1 angerissen wurden.

Der Satz von Sarkovskii ist ein Satz der Mathematik, der eine wichtige Aussage Ÿber die mšglichen Perioden bei der Iteration einer stetigen Funktion macht. Ein Spezialfall des Satzes ist die Aussage, dass ein stetiges dynamisches System auf der reellen Geraden mit einem Punkt der Ordnung 3 bereits Punkte zu jeder Ordnung besitzt. Dies wird hŠufig kurz so formuliert, dass Periode 3 Chaos impliziert. es handelt sich um einen der zunächst erstaunlichsten Sätze der elementaren reellen Analysis.

Es geht um die Länge der Perioden reeller Funktionen $f : \mathbb{R} \to \mathbb{R}$. Ein Punkt $x \in \mathbb{R}$ hat die *Periode* k wenn $f^{(k)}(x) = x$ und $f^{(j)}(x) \neq x$ für jedes j mit $0 < j < k$.

Seien die natürlichen Zahlen folgendermaßen angeordnet:

$$\begin{array}{ll} 3 & \prec 5 \prec 7 \prec 9 \prec 11 \prec 13 \prec \cdots \\ 2 \cdot 3 & \prec 2 \cdot 5 \prec 2 \cdot 7 \prec 2 \cdot 9 \prec \cdots \end{array}$$

$$\begin{array}{ll}
4 \cdot 3 & \prec 4 \cdot 5 \prec 4 \cdot 7 \prec 4 \cdot 9 \prec \cdots \\
\vdots & \\
2^n \cdot 3 & \prec 2^n \cdot 5 \prec 2^n \cdot 7 \prec 2^n \cdot 9 \prec \cdots \\
\vdots & \\
\cdots & \prec 2^4 \prec 2^3 \prec 2^2 \prec 2 \prec 1
\end{array}$$

Der Satz besagt nun:

Satz von Šarkovskii *Habe die stetige Funktion $f : \mathbb{R} \to \mathbb{R}$ einen periodischen Punkt der Ordnung k, so hat es für jedes $m \succ k$ einen periodischen Punkt der Ordnung m.*

Das verblüffende an dieser weitreichenden Aussage ist, daß sie praktisch ohne Voraussetzungen auskommt. Ein sehr einfacher Beweis findet sich in G. Winkler (1986). Einige der erstaunlichen Konsequenzen sind:

- Hat f irgendeinen periodischen Punkt, so hat es auch einen *Fixpunkt* x, d.h. $f(x) = x$.
- Hat f nur endlich viele periodische Punkte, so sind alle Periodenlängen Potenzen von 2.
- hat f einen Punkt der Periode 3, so auch periodische Punkte jeder Ordnung („period 3 implies chaos").

Allerdings besagt der Satz nichts über die Stabilität; man kann also in der Simulation nicht alle Perioden sehen. Trotzdem findet man einige etwa in der Abb. 13 wieder.

6 Die Gaußverteilung und ihre Folgen

Die Gauß- oder Normalverteilungverteilung ist neben der Laplace- und der Binomialverteilung die bekannteste aller Verteilungen. Ihr Segen aber auch ihr Fluch liegt in ihrem phänomenalen Erfolg.

Dies ist eine Verteilung mit Dichte $1/\sqrt{2\pi\sigma^2}\exp(-(x-\mu)^2/(2\sigma^2))$. Dabei sind μ der Erwartungswert und σ^2 die Varianz. Sind $\mu = 0$ und $\sigma^2 = 1$, so spricht man von einer *Standardnormalverteilung* und bezeichnet ihre Verteilungsfunktion mit Φ.

6.1 Der Zentrale Grenzwertsatz

Die Bedeutung der Normalverteilung beruht auf auf dem Zentralen Grenzwertsatz. Zur Formulierung seien $\xi_1, \xi_2, \ldots$ Zufallsvariablen, jeweils mit Erwartungswert μ und Varianz $\sigma^2 < 0$; die *standardisierten Summen* seien

$$S_n^* = \frac{\sum_{k=1}^{n} \xi_k - n\mu}{\sqrt{n\sigma^2}}.$$

Eine einfache Variante liest sich

Satz 1 (Zentraler Grenzwertsatz) *Seien $\xi_1, \xi_2, \ldots$ unabhängige Zufallsvariablen, welche identisch verteilt sind und deren Varianz endlich ist. Dann konvergieren die Verteilungsfunktionen F_n der standardisierten Summen S_n^* punktweise gegen Φ.*

Die Wirkungsweise des Satzes ist in den Abbildungen 17 und 18 illustriert.

Der Satz wurde im Prinzip von Abraham de Moivre (1867-1754) für binäre Zufallsvariablen ξ_k bewiesen. Die Beweise sind nicht rigoros. Trotzdem ist er der Vater der Gaußverteilung. Eine von mehreren Originalarbeiten ist Abraham De Moivre (1733). Für bibliographische Notizen dazu vergleiche man R.H. Daw and E.S.. Pearson (1972). De Moivre ging es vor allem darum, die für große n immer aufwendiger werdende Berechnung von Binomialkoeffizienten und deren Summen zu umgehen. In diesem Zusammenhang bewies er eine Version der Sterlingformel.

Definitiv geht die Normalverteilung *nicht* auf C.F. Gauss zurück. Er entwickelte die Methode der kleinsten Quadrate. Somit ist auf der 10 DM Note die Dichte zu Unrecht dem Porträt von Gauß zur Seite gestellt, vgl Abb. 19.

Laplace kannte die Doctrine of Chances. Er hat die Normalverteilung verschiedentlich in seinen Arbeiten zur Fehlerrechnung benutzt und in seinem Hauptwerk zur Wahrscheinlichkeitstheorie der *Théorie analytique des probabilités* im Rahmen seiner Theorie der erzeugenden Funktionen abgeleitet. Die Liste der Wissenschaftler die weiter am Satz forschten liest sich wie ein *who is who* der Mathematik. Sie beinhaltet u.a. Laplace, Poisson, Bessel, Dirichlet, Cauchy, Ellis, Chebyshev, Markov und Liapunov.

Bei Gauß, der in einer seiner verschiedenen Arbeiten zur Begründung der Methode der kleinsten Quadrate die Normalverteilung aus einer Funktionalgleichung ableitete, ist die Normalverteilung bereits so weit etabliert, daß man dafür auf die 1798 erschienenen Tafeln der Normalverteilung von Christian Kramp zurückgreifen konnte.[8] Details zur Geschichte findet man in den Arbeiten und Mongraphien von Ivo Schneider, [46; 47] und anderen.

[8]Für diese Mitteilungen danke ich Herrn Ivo Schneider

Der Zentrale Grenzwertsatz ist eine rein asymptotische Aussage und somit für praktische Zwecke im Prinzip wertlos. Niemand käme auf die Idee, aus der Tatsache, daß die Folge der $1/n$ gegen 0 konvergiert, wenn n gegen Unendlich strebt, zu schließen, daß $1/6 = 0$ ist. Genau dies wird aber in Naturwissenschaften täglich praktiziert. Nicht selten wird ja der t-Test zum Vergleich von Mittelwerten bei Fallzahlen von 3 - 6 verwendet. Letzterer beruht aber auf der Annnahme von Gaußverteilungen, also letztlich auf dem Zentralen Grenzwertsatz. Dies ist Unfug!

Quantitative Abschätzungen von Wahrscheinlichkeiten liefert der Satz von Berry-Esséen, vgl. P. Gänssler and W. Stute (1977). Für binäre Variablen ξ_k mit Erfolgswahrscheinlichkeit p gilt z.B. die Abschätzung

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \le \frac{p^2 + (1-p)^2}{\sqrt{2\pi np(1-p)}}. \tag{9}$$

Im Nenner der Schranke stehen also im wesentlichen die Standardabweichung des Einzelexperimentes und die Wurzel der Anzahl. Man vergleiche dazu H. Ziezold (2001) und P. Eichelsbacher (2001).

Beispiel 6.1 Die (heute historische) „Zwölferregel“zur approximativen Erzeugung einer normalverteilten Zufallsvariablen lautet: erzeuge unabhängige, auf (0,1) gleichverteilte Zufallsvariablen $U_1, \ldots, U_{12}$. Dann ist

$$\xi = \sum_{k=1}^{12} U_k - 6$$

ungefähr standardnormalverteilt. In der Tat haben die U_k Erwartungswert 1/2 und Varianz 1/12, so daß $S_{12}^* = (\sum_{k=1}^{12} U_k - 6)/(12\cdot 1/12)^{-1/2} = \xi$ ist. Eine allgemeine Version des Satzes von Berry-Esséen liefert die Abschätzung

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \le \frac{0,8\gamma}{\sigma^3\sqrt{n}},$$

mit dem zentrierten dritten Moment γ von U_1. Standardrechnung ergibt die Schranke 0,3 für ξ. Zwar ist dies eine grobe Abschätzung, dennoch ist sie nicht sehr vertrauenseinflößend.

6.2 Der zentrale Grenzwertsatz gilt manchmal

In der Tat gibt es Bereiche der Wissenschaft, in denen der Zentrale Grenzwertsatz die Realität sehr genau widerspiegelt. Insbesondere ist dies in der Statistischen Physik der Fall.

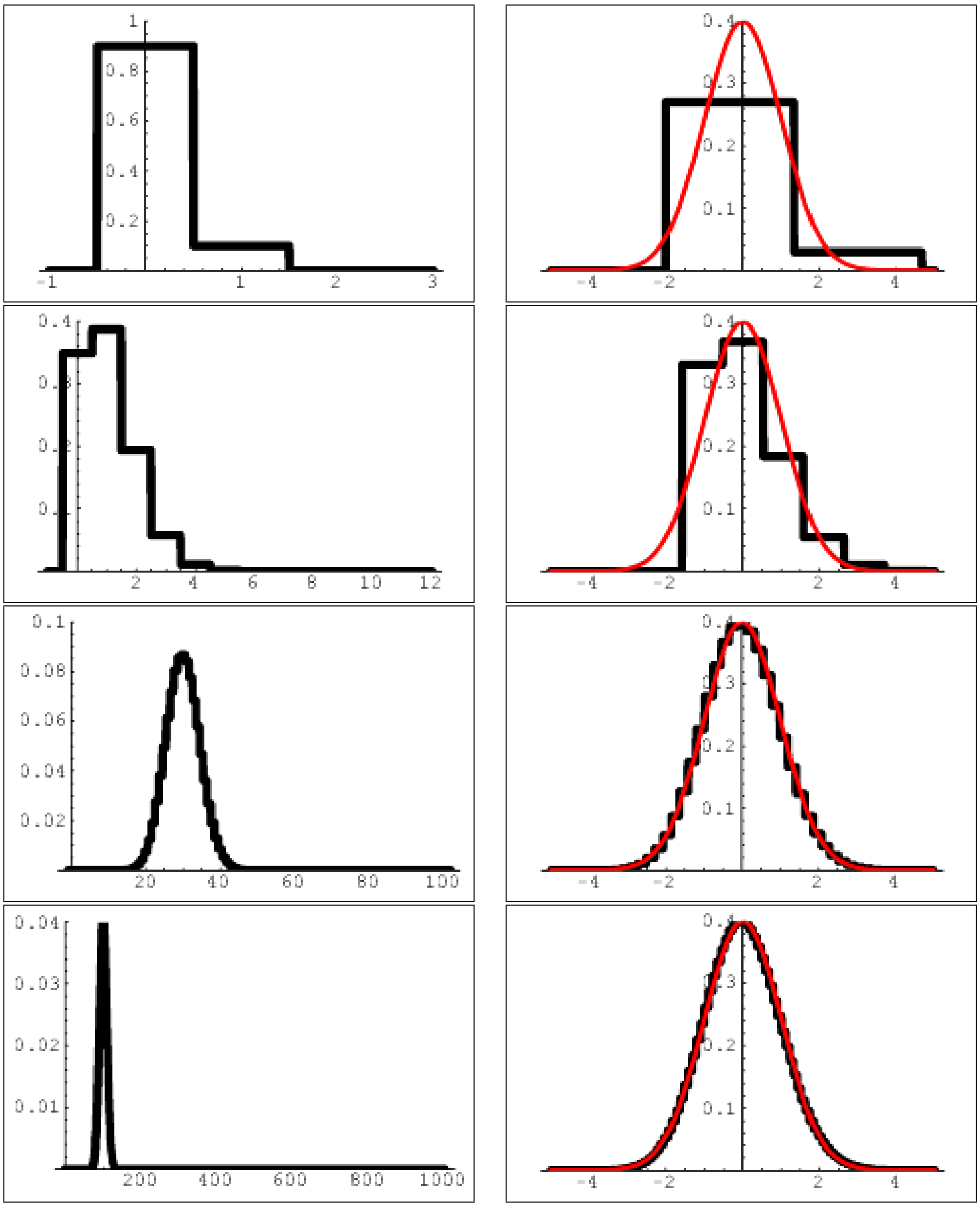


Abbildung 17: Histogramme der Binomialverteilung zu $p = 0,1$ und $n = 1, 10, 100, 1000$ (jeweils links), die standardisierten Histogramme im Vergleich zur Standardnormalverteilung (rechts).

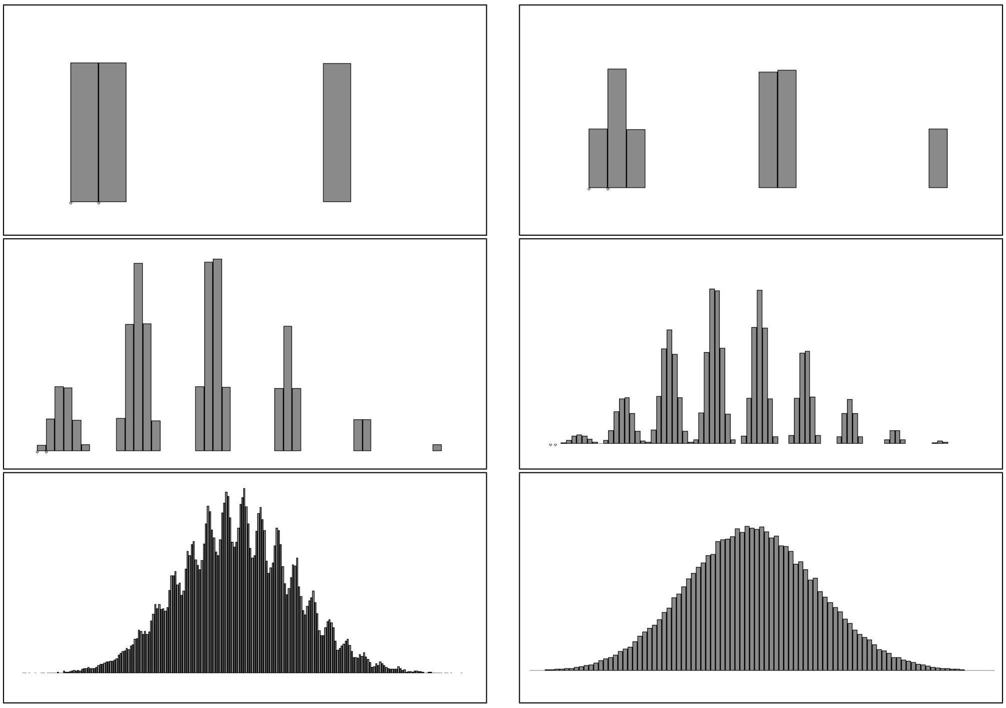


Abbildung 18: Seien $\xi_1, \ldots, \xi_n$ unabhängige Wiederholungen einer Zufallsvariablen ξ, welche auf $\{0, 1, 9\}$ gleichverteilt ist. μ_n seien die Verteilungen der Summen $\xi_1 + \cdots + \xi_n$. Für $n = 1, 2, 5, 10, 50$ und 1000 wurden jeweils 100 000 Realisierungen berechnet. Im Bild sind die jeweiligen empirischen Histogramme dargestellt.

Abbildung 19: Die Vorderseite des 10 DM Scheines der BRD mit dem Portrait von Carl Friedrich Gauß und der Standardnormalverteilung.

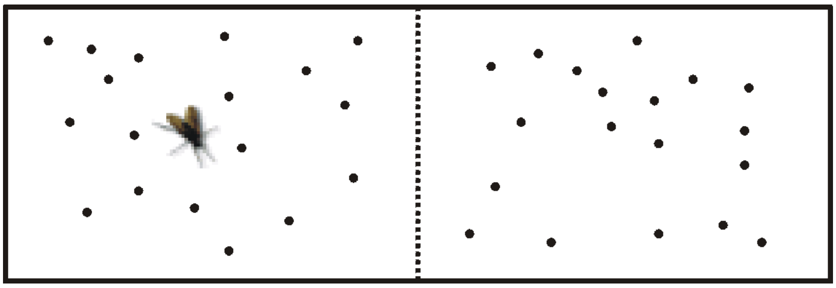


Abbildung 20: Zwei Hälften eines Gefäßes mit Luftmolekülen (und Fliege)

Beispiel 6.2 Das folgende Beispiel orientiert sich am *Urnenmodell* von P. und T. Ehrenfest, also deren *Stoßzahlansatz*. In einem Gefäß des Volumens 1 Kubikmeter seien etwa $0,2687 \cdot 10^{26}$ Gasmoleküle, vgl Abb. 20. Diese Zahl ist ungefähr die Loschmidt-Konstante für ein ideales Gas unter Normalbedingungen. Jedes Molekül befindet sich mit Wahrscheinlichkeit $1/2$ unabhängig von den anderen in der linken bzw. rechten Hälfte. In der linken Hälfte sitzt eine leicht paranoide Fliege, die schon mehrere Selbsthilfegruppen durchlaufen hat. Sie hat nun panische Angst, keine Luft mehr zu bekommen, da sich die Moleküle in der anderen Hälfte konzentrieren könnten.

Wir betrachten die Zufallsvariablen

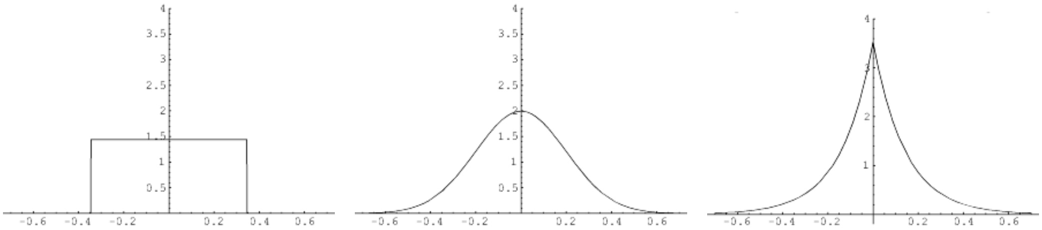


Abbildung 21: Uniforme Verteilung auf $[-\sqrt{3}/5, \sqrt{3}/5]$, Normal- und Laplaceverteilung, jeweils Streuung 0,2

$$\xi_i = \left\{ \begin{matrix} 0 \\ 1 \end{matrix} \right\} \text{ falls Molekül } \left\{ \begin{matrix} \text{links} \\ \text{rechts} \end{matrix} \right. , \quad 1 \leq i \leq n.$$

Wir nehmen ferner den Gleichgewichtszustand an, d.h. daß $(\xi_1, \ldots, \xi_n)$ auf $\{0,1\}^n$ gleichverteilt ist und daß die ξ_i unabhängig sind. Nach derUngleichung 6.1 ist die gleichmäßige Abweichung der Verteilungsfunktionen der standardisierten Summen von Φ von der Größenordnung 10^{-13} und somit die Wahrscheinlichkeit, daß sich die Moleküle gleichmäßig verteilen, praktisch 1.

6.3 Die ludische Verzerrung

Diesen Begriff hat N.N. Taleb durch seinen Bestseller [50] popularisiert. Ludische Verzerrung bedeutet, daß die klassische Stochastik um Urnenmodelle, Binomial- und Normalverteilung auf klassische Spiele wie Würfel oder Roulette zugeschnitten, aber (*ludus* lateinisch für Spiel) in der Realität, und insbesondere in den „Social Sciences“für die Modellierung ungeeignet ist und sogar katastrophal versagt . Dies gilt zum Beispiel im Bereich der Finanzmathematik In der Tat ist die Gaußverteilung nicht in der Lage, große Variabi-

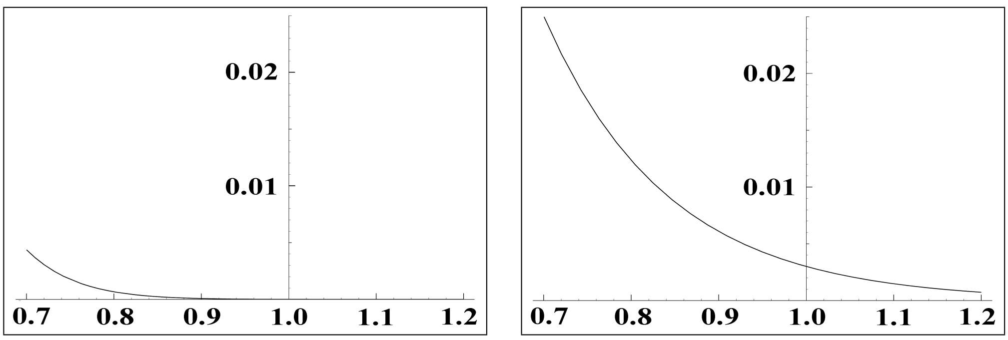


Abbildung 22: Details der Außenbereiche der Tails von Standardnormal- und Standardlaplaceverteilung

litäten zu erfassen.

Folgende Zahen entnehmen wir B.B. Mandelbrot and R.L. Hudson (2005) (der Autor kann die Berechnungen nicht nachvollziehen). Auf den Seiten 26 und 38 berichtet er:

Von 1916 bis 2003 sollte sich nach der Standardtheorie innerhalb dieses Zeitraumes der Dow-Jones Index an 58 Tagen um mehr als 3,4 % ändern, tatsächlich war dies an 1001 Tagen der Fall; ähnlich verhält es sich bei 4,5 % mit prognostizierten sechs Tagen gegenüber 366 beobachteten. Eine Indexänderung von 7,7 % sollte sich unter dem Standardmodell nur mit Wahrscheinlichkeit 1 zu 50 Milliarden ereignen. Tatsächlich gab es im Juli 2002 drei solcher steilen Abstürze innerhalb von sieben Handelstagen. Im 20. Jahrhundert gab es alleine 48 solcher Tage beim Dow-Jones Index. Der Einbruch am 19.10.1987 um 29,2 Prozent habe unter dem Standardmodell eine Wahrscheinlichkeit geringer als 10^{-50}. Mandelbrot bringt weitere solcher Beispiele.

Die Kritik ist in B.B. Mandelbrot and R.L. Hudson (2005) vorgetragen, einem populärwissenschaftlichen Werk zur Finanzmathematik. Dort werden auch alternative Zugänge vorgeschlagen. N.N. Taleb (2008) popularisiert diese Überlegungen bis zur Unkenntlichkeit.

Der hauptsächliche Grund ist, daß die äußeren Abschnitte der Normalverteilung, die *tails*, fast keine Wahrscheinlichkeitsmasse tragen, die Verteilung ist scharf um den Mittelpunkt konzentriert, sie ist *light tailed.* In Abb. 21 sind drei Dichten skizziert. Die Laplaceverteilung rechts ist *heavy tailed* und läßt

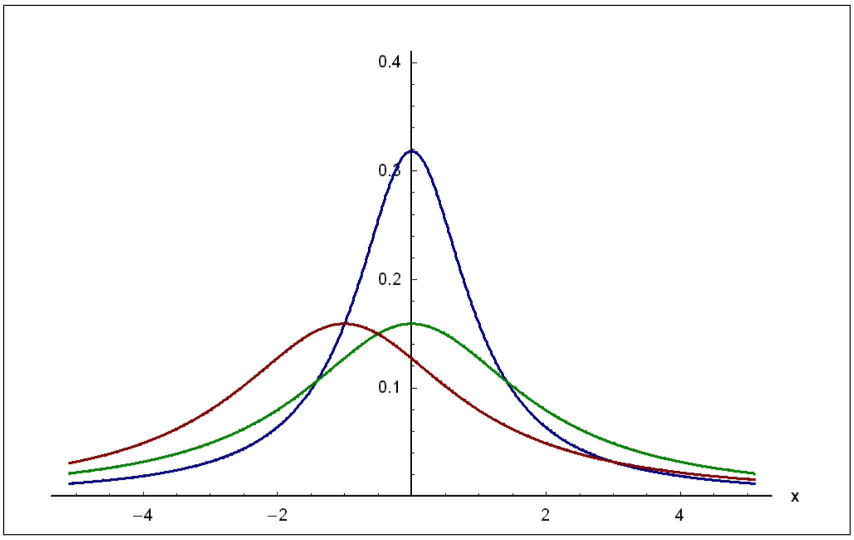


Abbildung 23: Dichten von Cauchy Verteilungen; die blaue Kurve stellt die Standardcauchydichte dar.

somit *robustere* Modelle als die Gaußverteilung zu. Dieser Effekt ist in Abb. 22 detaillierter erfaßt. Ein Kompromiß ist die *Huber-Verteilung*, bei welcher der Exponent der Dichte in der Mitte Gaußisch und außen linear verläuft, also ähnlich wie die Laplaceverteilung. Robuster noch als Laplaceverteilungen sind die Cauchyverteilungen mit Standarddichte f bzw. Verteilungsfunktion F der Form

$$f(x) = \frac{1}{\pi(1-x^2)}, \quad F(x) = \pi/2 + \arctan x.$$

Sie treten z.B. als Verteilungen der Quotienten normalverteilter Zufallsvariablen auf und somit als Spezialfall der Studentschen t-Verteilung. Allerdings besitzen sie keinen Erwartungswert und somit auch keine endliche Varianz. Insbesondere erfüllen Cauchyverteilungen *nicht die Gesetze der großen Zahlen*. Einige Dichten sind in Abb. 23 geplottet. Während das 99 % Quantil der Standardnormalverteilung ca. 2,33 beträgt liegt dieses für die Standard-Cauchy-Verteilung bei 32.

Die Verteilung wird oft als die des „blinden Bogenschützen“ bezeichnet; dieser schießt seinen Pfeil unter einem auf $[-\pi/2, \pi/2]$ gleichverteilten Winkel ab. Dann sind die Auftreffpunkte auf eine eine Einheit entfernte Mauer cauchyverteilt; dies liest man sofort aus der Verteilungsfunktion ab.

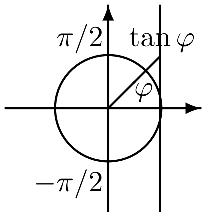


Tabelle 1 enthält einige Standardbereiche der Normalverteilung. Man erkennt deutlich, daß weit außen praktisch keine Masse mehr liegt. Tatsächlich ist die Laplaceverteilung proportional zu $\exp(-C|x|)$ während die Normalverteilung proportional zu $\exp(-Dx^2)$ ist und somit nach außen dramatisch schneller abfällt. In Tabelle 2 sind die 95 % und 99% Quantile aufgelistet.

von – bis	Wahrscheinlichkeit %
$\boldsymbol{m} \pm 1 \cdot \sigma$	68,27
$\boldsymbol{m} \pm 2 \cdot \sigma$	95,45
$\boldsymbol{m} \pm 3 \cdot \sigma$	99,73
$\boldsymbol{m} \pm 5 \cdot \sigma$	99,996

von – bis	Wahrscheinlichkeit %
$\boldsymbol{m} \pm 1,96 \cdot \sigma$	95
$\boldsymbol{m} \pm 2,58 \cdot \sigma$	99
$\boldsymbol{m} \pm 3,29 \cdot \sigma$	99,9

Tabelle 1: Wahrscheinlichkeiten symmetrischer Intervalle um den Mittelwert $\boldsymbol{m}$ der Normalverteilung

Dies sind die Werte q, für die das Intervall $[q, \infty)$ die Wahrscheinlichkeiten 0,05 bzw. 0,01 hat.

Auch abseits der Finanzmathematik gibt es erhebliche Kritik seitens moderner Statistiker. Dies führte unter anderem zur Entwicklung der robusten

Statistik, siehe P.J. Huber (1981) oder F.R. Hampel et al. (1986). J. Tuckey und P.L. Davies empfehlen absolute Abweichungen anstelle der quadratischen Abweichungen, was zur Laplaceverteilung führt. Dies ist eng mit obiger Diskussion verbunden.

Verteilung	95% Quantil	99% Quantil
Standardnormal-Verteilung	1,65	2,33
Laplace-Verteilung	2,30	3,91
Cauchy-Verteilung	6,31	31,82

Tabelle 2: Quantile verschieden robuster Verteilungen, jeweils in der Standardform

7 Grundbegriffe der Wahrscheinlichkeitsrechnung

Die letzte Stufe der Entwicklung vom Wahrscheinlichkeitsgefühl über Erkenntnistheorie und den wissenschaftlichen Wahrscheinlichkeitsbegriff ist die mathematische Axiomatisierung. Diese reiht einerseits den Begriff in die pure Mathematik ein, beraubt ihn aber *zunächst* jeden realen Hintergrundes. Auf dieser Basis kann anschließend das „Glasperlenspiel“der Mathematik gespielt werden. Andererseits erfolgt aber sowohl die Setzung der Axiome als auch die Interpretation der daraus abgeleiteten Sätze im Hinblick auf eine Realität im Rahmen der jeweiligen erkenntnistheoretische Position.

Im wesentlichen können zwei Lager unterschieden werden, nämlich Objektivisten und Subjektivisten. Zwischen diesen Lagern gab es - vornehmlich in der ersten Hälfte des letzten Jahrhunderts - vehemente Auseinandersetzungen. Der Objektivist läßt nur Ereignisse zu, deren Wahrscheinlichkeiten als objektiv existent postuliert werden und im Prinzip durch Messungen oder Beobachtungen approximativ bestimmt werden können - allerdings stets mit einer gewissen Unsicherheit. So schließt er von vornherein ein Fülle von Gegebenheiten aus, mit denen wir im täglichen Leben konfrontiert sind, z.B. alle Ereignisse, die nur einmalig auftreten. Der Subjektivist dagegen denkt über die Einschätzung von Unsicherheit durch Individuen nach. Im Idealfall beinhaltet die subjektivistische die objektivistische Theorie. Beachtet man, daß dem Objektivismus der (unbeweisbare) Glaube an real existierende Wahrscheinlichkeiten zugrunde liegt, so scheint der Subjektivismus objektiver zu sein, da er ohne diesen Glaubenssatz auskommt.

Wir zäumen nun das Pferd vom Schwanz her auf, beginnen mit der allgemein akzeptierten Axiomatik und diskutieren diese anschließend.

7.1 Die Axiomatik von Kolmogoroff

Dies ist die noch heute allgemein akzeptierte Axiomatik der Wahrscheinlichkeitstheorie. Sie wurde in A. KOLMOGOROFF (1933,1973) vorgestellt. Ähnlich wie in Geometrie und Algebra wird nicht erklärt, was die betrachteten Objekte denn seien. Es werden nur die Beziehungen zwischen ihnen festgelegt, also die mathematischen „Spielregeln“.

Die Monographie ‘Grundbegriffe der Wahrscheinlichkeitsrechnung’ von Andrei Nikolajewitsch Kolmogorov (1903-1987) [21] enthält die fundamentalen Axiome der Stochastik, wie sie heute noch der Wahrscheinlichkeitsrechnung zugrunde gelegt werden und damit den Kern der heutigen Wahrscheinlichkeitstheorie. Seine Beiträge über ‘Analytische Methoden in der Wahrscheinlichkeitstheorie’ (1938), sind Grundlage für die Theorie der Markov Prozesse.

Im Hintergrund steht jeweils ein Experiment mit einer wohldefinierten Präparierung, d.h. eine bekannte Ausgangssituation. Das Experiment habe mögliche Ausgänge ω, deren Eintreten aber jeweils ungewiß sei. Zunächst wird also die Menge Ω der möglichen *Ergebnisse* ω festgelegt. Zum zweiten wählt man ein System $\mathcal{F}$ von Teilmengen in Ω; dieses wird unseren Wissenstand widerspiegeln. Jedes $F \in \mathcal{F}$ repräsentiert ein *Ereignis*. Ein Ereignis F *tritt ein*, wenn das tatsächliche Ergebnis $\hat{\omega}$ in F liegt. Dabei ist zu *fordern*, daß feststellbar ist, ob $\hat{\omega}$ tatsächlich in F liegt oder nicht.

Unter dieser Interpretation und wenn man die klassische zweiwertige Logik akzeptiert, muß $\mathcal{F}$ die Struktur einer *Mengenalgebra* tragen. Insbesondere muß das Ereignis „irgendetwas geschieht“, d.h. die Menge Ω in $\mathcal{F}$ enthalten sein. Mit einem Ereignis F muß seine Negation, also das Komplement F^c ebenfalls enthalten sein. Eine weitere Forderung entspricht dem logischen „oder“, d.h. mit F_1 und F_2 muß auch die Vereinigung $F_1 \cup F_2$ in $\mathcal{F}$ liegen. Daraus folgt leicht, daß auch die anderen logischen Operationen wie das

„und“ ihre mengentheoretische Entsprechung finden. Ist mit $F_1, F_2, \ldots$ auch die Vereinigung in $\mathcal{F}$, so nennt man dieses eine *(Mengen-)* σ*-Algebra*.

Es sei betont, daß dies ausschließlich der klassischen Logik entspricht. Zum Beispiel ist das *tertium non datur* wesentlich. Für die Quanten- oder auch die weibliche Logik braucht man andere Regeln. Die Forderung nach Verifizierbarkeit oder Falsifizierbarkeit durch eine Beobachtung schränkt die Wahl der Algebra ein. Bepunktet man zum Beispiel die zwei Seiten zweier Münze jeweils mit den Zahlen 0 oder 1 und beobachtet nur die Augensumme, so kann man die Menge $A = \{(0,1),(1,0)\}$ durch Beobachtung nicht auflösen. Sie bildet ein nicht teilbares *Atom* der zugehörigen Algebra. Es ist also nicht statthaft, die Potenzmenge von $\Omega = \{(0,1),(0,0),(1,0),(1,1)\}$ als Algebra zu wählen. Die zugehörige Algebra muß also sein:

$$\mathcal{F} = \{\emptyset, \Omega, \{(0,0)\}, \{(1,1)\}, A, A \cup \{(0,0)\}, A \cup \{(1,1)\}\}\}$$

Entsprechend der Annnahme, daß die Ereignisse unsicher sind, müssen sie nun entsprechend bewertet werden. Dies soll durch eine nichtnegative reellwertige Funktion $\mathbb{P}$ auf $\mathcal{F}$ geschehen. $\mathbb{P}$ bewertet die Ereignisse zwischen absoluter Unmöglichkeit und Sicherheit. Da das Experiment durgeführt werden soll, muß zwingend ein $\omega \in \Omega$ eintreten. Wir vereinbaren für dieses definitiv sichere Ereignis den Wert $\mathbb{P}(\Omega) = 1$. Für das logisch unmögliche Ereignis $\emptyset$ sollte natürlich $\mathbb{P}(\emptyset) = 0$ sein (das Experiment liefert ja irgendeinen Ausgang ω). Dies wird allerdings folgen und muß deshalb nicht gefordert werden.

Die letzte Regel sei zunächst nur für Algebren $\mathcal{F}$ formuliert:

$$\mathbb{P}(F_1 \cup F_2) = \mathbb{P}(F_1) + \mathbb{P}(F_2), \quad \text{für disjunkte Ereignisse } F_1, F_2 \in \mathcal{F}. \tag{10}$$

Aus dieser Regel folgt etwa sofort, daß $\mathbb{P}(\emptyset) = 0$. Während die schwächere Forderung „$F_1 \subset F_2 \implies \mathbb{P}(F_1) \leq \mathbb{P}(F_2)$“ unmittelbar einleuchtet, muß (10) gesondert begründet werden. Dies wird in Abschnitt 7.2 versucht werden. Es wären ja auch andere Funktionen $f(\mathbb{P}(F_1), \mathbb{P}(F_2))$ als die Addition denkbar, z.B.

$$\mathbb{P}(F_1 \cup F_2) = \sqrt{\mathbb{P}(F_1)^2 + \mathbb{P}(F_2)^2}.$$

Auch hieraus würde $\mathbb{P}(\emptyset) = 0$ folgen.

Da man z.B. Aussagen über die Konvergenz der relativen Häufigkeiten eines Ereignisses gegen seine Wahrscheinlichkeit von hohem Interesse sind, steht man Mengen des Typs $\xi_n \longrightarrow \xi$, $n \to \infty$, gegenüber, wobei die ξ_i und ξ Funktionen auf Ω sind. Es ist

$$\{\omega \in \Omega : \xi_n(\omega) \longrightarrow \xi(\omega)\} = \bigcap_{k=1}^{\infty} \bigcup_{n_0 \geq 1} \bigcap_{n \geq n_0} |\xi_n(\omega) - \xi(\omega)| \leq \frac{1}{k}.$$

Hier werden abzählbare Mengenoperationen benötigt, die natürlich in den Wahrscheilichkeitskalkül hinein wirken. Wir formulieren also die folgende Axiomatik:

Seien Ω eine Menge versehen mit einer σ-Algebra $\mathcal{F}$. Sei ferner $\mathbb{P}$ eine reellwertige Abbildung auf $\mathcal{F}$. Es gelten:

Kolmogorov-Axiom 1 (Konvention) $\mathbb{P}$ *ist nichtnegativ,*

Kolmogorov-Axiom 2 (Normalisierung) $\mathbb{P}(\Omega) = 1$,

Kolmogorov-Axiom 3 (σ-Additivität) *für paarweise disjunkte Ereignisse $F_1.F_2, \ldots$ aus $\mathcal{F}$ gilt*

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(F_k).$$

Man nennt das Tripel $(\Omega, \mathcal{F}, \mathbb{P})$ einen *Wahrscheinlichkeitsraum*, und $\mathbb{P}$ ein *Wahrscheinlichkeitsmaß*. Früher war die Bezeichnung Wahrscheinlichkeitsfeld gebräuchlich. Dieses Axiomensystem ist widerspruchsfrei (weil es Beispiele gibt) und unvollständig (weil es viele Beispiele gibt). Es bleibt nur noch die Erweiterung des Additionssatzes auf die σ-Additivität zu diskutieren. Dazu zitieren wir A. Kolmogoroff (1933,1973):

> Unendliche Wahrscheinlichkeitsfelder erscheinen nur als idealisierte Schemata reeller zufälliger Prozesse. *Wir beschränken uns dabei willkürlich auf solche Schemata, welche dem Stetigkeitsaxiom genügen.* Diese Beschränkung erwies sich bis jetzt bei den verschiedensten Untersuchungen als zweckmäßig.

Unter „Stetigkeitsaxiom“ versteht Kolmogoroff die Aussage

Kolmogorov-Axiom 3 (σ-Stetigkeit) *Für eine abnehmende Folge $F_1 \supset F_2 \supset \cdots \supset F_n \supset \cdots$ von Ereignissen in $\mathcal{F}$ mit $\bigcap_k F_k = \emptyset$ gilt $\mathbb{P}(F_k) \to 0,\ k \to \infty$.*

Austausch von (c) und (d) führt zu äquivalenten Axiomensystemen.

Bezüglich der Sigma- Additivität (σA) nimmt Kolmogoroff einen pragmatischen Standpunkt ein, der durch die einleitenden Bemerkungen gestützt wird. Allerdings hat die Erweiterung auf abzählbare Mengenoperationen dramatische Folgen für die abgeleiteten Sätze. Die entsprechende Diskussion sprengt den vorliegenden Rahmen. Es sei hier lediglich de Finetti (siehe Abschnitt 7.3 erwähnt, der die Sigma-Additivität ablehnte, weil sie seiner Auffassung nach zu paradoxen Konsequenzen führt. Zusammen mit Alfréd

Rényi versuchte er, eine alternative Axiomatisierung der Wahrscheinlichkeitsrechnung zu entwickeln. Erfahrungsgemäß führt dieser Ansatz zu erheblichen Schwierigkeiten, was KOLMOGOROFF'S Meinung bestätigt.

Ein Problem, das in diesem Zusammenhang entsteht, sei doch angesprochen, da es mit der Bemerkung oben über die Entscheidbarkeit des Eintretens von Ereignissen nicht kompatibel ist. Betrachtet man z.B. $\Omega = [0,1]$, versieht dies mit der kleinsten σ-Algebra, welche alle Intervalle enthält, der *Borel-σ-Algebra* $\mathcal{B}([0,1])$, so ist das natürliche Längenmaß λ, das *Lebesguemaß*, ein Wahrscheinlichkeitsmaß auf $\mathcal{B}([0,1])$. Nun liegt aber jede einpunktige Menge $\{x\}$, $x \in \Omega$, in $\mathcal{B}([0,1])$ und es gilt $\lambda(\{x\}) = 0$. Ein solches Ereignis ist zwar logisch möglich, aber im Allgemeinen praktisch nicht identifizierbar. Solche Effekte sind der Pragmatik geschuldet.

7.2 Die objektivistische Auffassung

Wir greifen noch einmal die Frage, was Zufall denn sei, auf. Obiges System ist mathematisch einwandfrei, daraus gezogene Schlüsse können nicht angefochten werden. Die Verbindung zur Realität überlassen Mathematiker gerne anderen Wissenschaften, wie Philosophie, Wissenschaftstheorie oder Statistik. Aber gerade in der Verbindung zur Realität liegt ja der Wert der Wahrscheinlichkeitstheorie.

In der objektivistischen Interpretation der Wahrscheinlichkeitstheorie stellt man sich auf folgenden Standpunkt:

Jedem Ereignis wohnt eine Größe - nämlich seine Wahrscheinlichkeit - inne, ähnlich wie ein Körper Ausdehnung oder Masse hat. Diese Größe kann man im Prinzip durch wiederholte unabhängige Messungen beliebig genau bestimmen. In den Worten von HANS RICHTER (1952)(1912-1978) liest sich das so:

> Unter den denkbaren Belegungssystemen mit der in der Wahrscheinlichkeitsrechnung festgelegten Struktur gibt es ein 'absolutes' System', das von der realen Welt gefordert wird; d.h. der numerische Wert der Wahrscheinlichkeit für jedes $E|H$[9] ist wie eine physikalische Größe festgelegt.

Im sowjetisch/stalinistischen Sozialismus/Kommunismus war im Sinne der materialistischen Weltsicht diese objektivistische Interpretation die einzig erlaubte [10]. Das liest man z.B. deutlich aus der Einleitung zu dem nach wie vor empfehlenswerten einführenden Werk B.W. GNEDENKO (1968) heraus.

[9] Das meint ein Ergebnis E eines Versuches H

[10] Deshalb war im Stalinismus auch die Quantenmechanik verboten. Man versuchte den Determinismus durch sogenannte 'verborgene Parameter' zu retten, vgl. Abschnitt 4.2.

Im objektivistischen Rahmen sind nur Ereignisse A für Experimente sinnvoll, die sich - wenigstens im Prinzip - beliebig of unabhängig voneinander realisieren lassen. Man will, daß die relativen Häufigkeiten n_A/n bei n Wiederholungen und der Trefferzahl n_A die - als objektiv existierend postulierte - Wahrscheinlichkeit $\mathbb{P}(A)$ approximieren. Welcher Art die Konvergenz ist, bleibt zunächst offen.

In diesem Rahmen wird die Additionsvorschrift (10) plausibel. Jedenfalls gilt für disjunkte Ereignisse A und B, daß sich die Anzahlen der Treffer addieren, d.h. daß $n_A + n_B = n_{A\cup B}$ gilt; somit folgt für die relativen Häufigkeiten

$$\frac{n_A}{n} + \frac{n_B}{n} = \frac{n_{A\cup B}}{n}.$$

Unterstellen wir die Konvergenz, so folgt (10).

Die Konvergenz findet sich in *Gesetzen der großen Zahlen*. Unter obigen Axiomen beweist man: Sei A ein Ereignis und sei für jedes $k \in \mathbb{N}$ das Ereignis, daß A bei der k-ten Wiederholung des Experimentes eintritt, mit A_k bezeichnet; ferner sei $\chi_{A_k} = 1$ wenn A_k eintritt und $\chi_{A_k} = 0$, wenn A_k nicht eintritt. Die relativen Häufigkeiten des Auftretens von A bei n Wiederholungen sind dann $(\sum_{k=1}^{n} \chi_{A_k})/n$. Sind die Experimente unabhängig, so gilt

$$\mathbb{P}\left(\omega \in \Omega : \frac{1}{n}\sum_{k=1}^{n} \chi_{A_k}(\omega) \;\longrightarrow\; \mathbb{P}(A), \quad n \to \infty\right) = 1.$$

Somit ist die Art der erhofften Konvergenz präzisiert. Allerdings gilt die analytische Konvergenz *nicht für alle* Versuchsausgänge $\omega \in \Omega$. Die Ausnahmen bilden ein Ereignis N mit $\mathbb{P}(N) = 0$, eine $\mathbb{P}$-*Nullmenge*. Diese enthält z.B. alle konstanten Folgen $\omega = (a, a, a, \ldots)$ mit $a \neq \mathbb{P}(A)$ und viele andere mehr. All diese Folgen sind theoretisch denkbare Ergebnisse.

Um diese rein mathematische Aussage mit der Realität in Verbindung zu bringen, wird ein Postulat benötigt, nämlich das Cournotsche Prinzip, Antoine Augustin Cournot (1843, 1973-1984):

> Sei $\varepsilon > 0$ gewählt. Gilt für ein Ereignis $A \in \mathcal{F}$, daß $\mathbb{P}(A) \geq 1 - \varepsilon$, so sollen wir so handeln, als ob das Eintreten von A gewiß wäre.

Ausführliche Quellenangaben und eine Diskusion finden sich in G. Shafer (2006).

Das subjektive Element der Wahl von ε scheint praktisch nicht vermeidbar. Dieses Phänomen scheint insbesondere in Schätz- und Testtheorie auf, wo man $1-\varepsilon$ etwa in Gestalt von Signifikanzniveaus oder bei Konfidenzintervallen begegnet. In der Tat bedienen wir uns dieses Prinzips - ohne darüber nachzudenken - oftmals bei Alltagsentscheidungen. All dies entspricht dem rechten absteigenden Ast in Diagramm 1.

Abbildung 24: Bruno De Finetti, 1906-1985, H. Richter, 1912-1978 (Bild 1976 von Konrad Jacobs, MFO Oberwolfach)

7.3 Die subjektivistische Auffassung

Die objektivistische Theorie läßt nur wenige Ereignisse der täglichen Erfahrung zu, da sie meist nur einmalig stattfinden. Ferner ist die Wiederholbarkeit der Ereignisse - wenn sie denn überhaupt gegeben ist - ein idealisierter Begriff; Experimente können nur gedanklich beliebig oft wiederholt werden. Eine Aussage wie „Caesar war in Großbritannien“ ist z.B. einmal schon deshalb unzulässig, weil nicht entschieden werden kann, ob er dort war oder nicht, und zum anderen auch, weil es sich um ein einmaliges Ereignis handelt.

Ferner gibt es erhebliche *prinzipielle Einwände* gegen das Postulat objektiv existierender Wahrscheinlichkeiten. So sagt Bruno de Finetti (1906-1985), vgl. Abb. 24:

> Es existiert keine objektive Wahrscheinlichkeit.

Seine Versuche einer subjektivistischen Axiomatik liefen parallel zu denen von F. Ramsey (1903-1930); auch Wissenschaftler wie J.M. Keynes (1883-1946) arbeiteten an diesem Gegenstand, letzterer im Rahmen der Ökonomie.

Sie bezieht sich auf die Einschätzung eines Individuums. *Eine* Art, sich das vorzustellen, beruht auf dem Wettbegriff: Zwei „vernünftige Individuen“schließen eine „faire Wette“ auf das Eintreten eines Ereignisses ab; das heißt, daß sie die Rollen tauschen würden. Der so entstehende Wettquotient ist dann die Wahrscheinlichkeit des Ereignisses. Hans Richter (1952) sagt:

> Objektiv (im Sinne von allgemeinverbindlich) an der Wahrschein-

lichkeitsbelegung der $E|H$ ist nur die Struktur dieser Belegung und die Vorschrift zur *Änderung* der Belegungswerte; dagegen sind die numerischen Werte selbst abhängig vom Wissensstande (andere insbesondere psychologische Einflüsse gelten meist als ausgeschaltet).

Zur Grundlegung dieser Begriffe lese man bei Hans Richter (1974a,b, 1975, 1972, 1952, 1953a,b,c, 1954, 1956) nach.

H. Richter argumentiert in [41]: Der Objektivist unterscheidet sich vom Subjektivisten dadurch, daß er an objektiv außerhalb von uns existierende Wahrscheinlichkeiten *glaubt*; er darf aber nicht behaupten, aus seiner Erfahrung heraus sichere Schlüsse auf deren Wert ziehen zu können. Dies zeigt sich besonders in der Statistik, wo meist eine Familie von Wahrscheinlichkeiten (das Modell) vorab gewählt wird. Gewisse dieser werden dann, z.B. durch einen Signifikanzest mit hoher Wahrscheinlichkeit verworfen. Das bedeutet aber keineswegs, daß auf die „wahre Verteilung“geschlossen werden kann.

H. Richter geht vom Begriff der Wette aus, ebenso wie B. De Finetti (1937), I.R Savage (1968), L.J Savage (1954) oder J.G. Kemeney (1955).

Als Diskussionsgrundlage geben wir die Axiomatik an, wie sie von H. Richter in [41] formuliert wurde. Zugrunde liegt die *einfache Wette* auf ein Ereignis E. Sie ist charakterisiert durch ihre *Auszahlungsfunktion*

$$a(\omega) = \alpha\chi_\Omega(\omega) + \beta\chi_E(\omega),\ \omega \in \Omega,$$

wobei α und β reelle Zahlen sind. β ist die Auszahlung im Gewinnfalle, α vertritt den Einsatz und ist im Allgemeinen von der Form $\alpha = -e$, $e \geq 0$, mit dem positiven Einsatz e; dieser ist stets verloren und man erhält im Gewinnfalle den Betrag $\beta - e$ und verliert sonst e. Es sind auch mehrere bzw. abzählbar viele Wetten mit den Auszahlungsfunktionen a_k simultan zugelassen, die gesamte Auszahlung ist dann $a = a_1 + a_2 + \cdots$, sofern die Summe endlich ist; wir nennen dies eine *σ-Wette*.

Die Menge dieser Wetten wird nun eingeteilt in die Gesamtheit $\mathfrak{F}$ der *fairen* und $\mathfrak{U}$ der *unfairen Wetten*. Einige Wetten sind für den Spiele nicht entscheidbar; sie bilden die Klasse $\mathfrak{N}$.

Im täglichen Leben kommen überwiegend unfaire Wetten vor. Ein Standardbeipiel ist das deutsche Lotto, bei dem die Finanzministerien vorweg 50% des Einsatzes einziehen.

Was fair bzw. unfair ist, wird nun in drei Axiomen formalisiert.

Natürlich soll eine Wette mit konstanter Auszahlung 0 fair sein:

Wett-Axiom 1 *Ist $a(\omega) = 0$, für alle $\omega \in \Omega$ so ist die Wette fair.*

Umgekehrt muß gelten:

Wett-Axiom 2 *Ist $a(\omega) \geq 0$, für alle $\omega \in \Omega$ aber nicht $a \equiv 0$, so ist die Wette unfair.*

Dies berücksichtigt, daß die Wette auch unentscheidbar sein kann.

Nun kommen hinterrücks die Wahrscheinlichkeiten ins Spiel. Ein Spieler habe für jedes Ereignis E eine Menge $\mathcal{P}(E)$ von zugelassenen *subjektiven Wahrscheinlichkeiten* festgelegt. Daß dies eine Menge und keine Zahl ist, berücksichtigt, daß sich der Spieler eventuell nicht auf genau eine Zahl festlegen kann. Entscheidend ist

Wett-Axiom 3 *Sei S eine reelle Zahl. Für jedes $p(E) \in \mathcal{P}(E)$ ist die Wette mit Auszahlungsfunktion*

$$a(\omega) = -p(E)S\chi_\Omega + S\chi_E$$

fair.

Dieses Axiom leuchtet sofort ein, wenn man vom Ende her denkt: Wären wir bereits im Besitz der Wahrscheinlichkeiten $\mathbb{P}$ und der Erwartungen $\mathbb{E}$, so wäre in diesem Fall

$$\mathbb{E}(a) = -S \cdot \mathbb{P}(E) \cdot 1 + S \cdot \mathbb{P}(E) = 0.$$

Eine Wette, bei der man weder etwas gewinnt noch verliert, muß natürlich fair sein.

Nun wird die Menge der subjektiven Wahrscheinlichkeiten strukturiert.

Wett-Axiom 4 *Der Spieler verfügt über eine Gesamtheit Θ von subjektiven Wahrscheinlichkeitssystemen. Es ist*

$$\{p_\vartheta(E) : \vartheta \in \Theta\} = \mathcal{P}(E) \qquad \textit{für jedes } E.$$

Dies steht im Gegensatz zu I.R Savage (1968), der meint, daß durch angestrengtes Nachdenken genau eine Zahl $\mathbb{P}(E)$ als subjektive Wahrscheinlichkeit von E gefunden werden könne. Richter argumentiert dagegen, daß man am Schreibtisch selbst bei sehr angestrengtem Nachdenken keine Naturgesetze finden könne, Hans Richter (1972), Seite 67.

Schließlich kommt noch eine Konsistenzforderung:

Wett-Axiom 5 *Hat eine σ-Wette die einzelnen Auszahlungsfunktionen*

$$a_k(\omega) = -S_k p_\vartheta(E_k)\chi_\Omega + S_k\chi_\Omega$$

bei übereinstimmendem $\vartheta \in \Theta$, so ist sie nicht unfair.

Die entscheidende Konsequenz aus diesen Axiomen ist nun der erwartete Satz:

Satz 2 *Für jede $\vartheta \in \Theta$ ist die Wahrscheinlichkeitsbelegung $\{p_\vartheta(E) : E \in \mathcal{F}\}$ ein Wahrscheinlichkeitsmaß auf $\mathcal{F}$.*

Man ist also wieder bei der Axiomatik von Kolmogoroff angelangt. Der Unterschied besteht nicht in den Axiomatiken, sondern in der objektivistischen bzw. subjektivistischen Auffassung.

Bemerkung Obiges ist eine vereinfachte Version der Axiomatik, die aber für das Verständnis vollkommen genügt. In der Originalversion geht H. Richter nur von Algebren anstatt von σ-Algebren aus. Ferner formuliert er im Hinblick auf die Entscheidungtheorie die Axiomatik für bedingte Wetten. Der interessierte Leser sei hierzu auf die Originalarbeiten verwiesen.

7.4 Komplexität

Ein dritter Aspekt betrifft die Komplexität konkreter Folge von Zahlen . Diese läßt sich z.B. durch die Länge eines Programmes auf einer Turing-Maschine messen, welches die Folge erzeugt, der „Kolmogorov-Komplexität“. Sie ist von Bedeutung im Hinblick auf Pseudozufallzahlen. Diese Idee geht auf A.N. Kolmogoroff und einem seiner vielen Schüler Per Martin-Löf zurück. Sie wurde in C.P. Schnorr (1971) schlüssig formuliert.

Eine schlüssige Einordnung in die vorliegenden Gedankengänge sprengt den Rahmen dieses Aufsatzes.

8 Der Zufall in der Kunst

Wir beschränken uns auf drei typische Beispiele aus Literatur und Film.

8.1 Der Zufall bei Karl Valentin

Karl Valentin (1882-1948), der wahrscheinlich bedeutendste Münchener „Komiker“(so komisch war er eigentlich gar nicht) hat den Begriff des Zufalles ideal erfaßt und in seiner skurilen Art in zwei Stücken dargestellt. Seine kongeniale Partnerin war Liesl Karlstadt (1892-1960). Wir erwähnen nur zwei Beispiele.

In seinem Stück (1940) „Der Zufall, Orchesterprobe, Komödie in zwei Akten Alternativtitel: Tingeltangel, So ein Theater Theater in der Vorstadt Die komische Kapelle“Karl Valentin, 1933, schaue Karl Valentin (2009) geht es um den Zufallsbegriff und seine subjektivistische Deutung ganz allgemein.

Valentin ist Mitglied eines Orchesters und kommt wie üblich zu spät zur Probe. Er erzählt, daß er mit einem Bekannten in der Kaufingerstraße gewesen sei, über Radfahrer gesprochen habe, und daß just in diesem Moment ein Radfahrer daher gekommen sei. Dies sei ein unglaublicher Zufall. Der Dirigent entgegnet, daß dies kein Zufall sei, da in der Kaufingerstraße „alle Meter - alle Sekunden" einer vorbeikäme. Valentin antwortet: *Sie haben halt eine eigene Weltanschauung.* Der Dirigent meint, wenn dasselbe mit einem Flugzeug geschehen wäre, so hätte es sich um einen Zufall gehandelt.

In einem weiteren Sketch will Valentin sein Haus gegen ein Bergwerk tauschen. Der Hauskäufer interessiert sich für den Grund. Wir zitieren aus Karl Valentin (1940) [11]:

KARLSTADT Kaufen Sie sich dann wieder ein neues Haus?
VALENTIN Niemals mehr! Ich suche ein altes tausend Meter tiefes Bergwerk zu mieten.
KARLSTADT Und das wollen Sie dann bewohnen?
VALENTIN Selbstverständlich!
KARLSTADT Das ist ja unheimlich!
VALENTIN Schon – aber sicher!
KARLSTADT Vor wem?
VALENTIN Vor Meteorsteinen.
KARLSTADT Aber Meteorsteine sind doch ganz selten.
VALENTIN Schon, aber bei mir geht die Sicherheit über die Seltenheit.

Damit hat Valentin die Grundidee der Extremwertstatistik, wie sie in der Finanzmathematik grundlegend ist, konzise charakterisiert.

8.2 Der Zufall im Film

„Der Zufall möglicherweise"(Originaltitel: Przypadek) ist ein 1981 entstandener Spielfilm des polnischen Regisseurs Krzysztof Kieślowski, der jedoch erst 1987 veröffentlicht wurde. Er wurde mit zahlreichen Auszeichnungen bedacht.

Vor dem politischen Hintergrund des damaligen kommunistischen Polen und der typischen dualen Auseinandersetzung zwischen der Partei und den

[11] Ein ausführlicheres Zitat ist leider aufgrund einer erheblichen Strafandrohung durch Herrn Rechtsanwalt Fette, München, nicht möglich

im Untergrund tätigen Oppositionellen zeigt er drei mögliche Versionen des Lebenslaufes des Protagonisten Witek Długosz. Der Regisseur untersucht die Rolle des Zufalls und des Determinismus als Faktoren des menschlichen Schicksals.

Witek will nach Warschau fahren und begibt sich zum Bahnhof. Von nun an werden drei verschiedene Versionen durchgespielt, die davon abhängen, ob er den Zug noch erreicht oder nicht. In der ersten Version gelingt ihm das und er beginnt in Warschau eine Karriere als kommunistischer Funktionär und Politiker, in der zweiten Version gerät er im Bahnhof mit einem Bahnpolizisten aneinander, und er beginnt eine Karriere als Oppositioneller. In der dritten Version verpaßt er den Zug, trifft aber eine Studienkollegin, in die er sich verliebt und sie heiratet.

8.3 Der Zufall in der Literatur

In seiner Erzählung „Der Schnupfen“ untersucht der polnische Schriftsteller Stanisłav Lem zufällige Ereignisse im Hinblick auf das Gesetz der großen Zahlen. Sein Held heißt Ijon Tichy, nach dem altgriechischen Wort Týche ($\tau\upsilon\chi\eta$) für die Göttin des Schicksals, der glücklichen oder bösen Fügung und des Zufalls, später auch für Schicksal, Zufall, Versehen oder Fehler. Die römische Entsprechung ist Fortuna.

Eine rätselhafte Häufung von Suiziden in Neapel, der zunächst wie ein Kriminalfall aussieht, wird analysiert. Lem konstruiert eine komplizierte Wirkungskette und argumentiert, daß die Folgen, obwohl unwahrscheinlich, aufgrund des Gesetzes der großen Zahlen eintreten müssen, wenn nur genügend viele Wiederholungen vorliegen.

8.4 Der Zufall in Action

Der Salzburger Aktionskünstler Jürgen Fuchs meint: *Kunst ist Zufall.*

9 Zum guten Schluß

Man kann also trefflich über den Begriff ‘Zufall’ diskutieren. Der elsässische Arzt, Theologe und Philosoph Albert Schweitzer (1875-1965) hat eine zugleich geistreiche und witzige Formulierung gefunden:

> Der Zufall ist ein Pseudonym, das der liebe Gott wählt, wenn er anonym bleiben will.

Im Gegensatz dazu siedeln viele Authoren „Gott“ genau innerhalb des Zufalls an, da sie im Rahmen einer deterministischen Weltsicht keinen anderen Platz für ihn finden. Ähnlich verhält es sich mit der Willensfreiheit. Hier bleibt Raum für Zweifel.

In der populärwissenschaftlichen Literatur erfreut sich der Zufall wachsender Beliebtheit. Dies zieht eine Flut von Büchern teils populärwissenschaftlicher, teils (semi-) philosophischer oder (semi-) psychologischer oder soziologischer Ausrichtung nach sich. Neuere Erscheinungen sind etwa W. Krämer (2002), M. Hampe (2006), K. Mainzer (2007) oder N.N. Taleb (2008). Man bilde sich selbst eine Meinung über die jeweiligen Botschaften.

Das vielleicht wichtigste Thema, nämlich die Statistik, wurde in diesem Aufsatz ausgespart. In der Praxis stellt sich heraus, daß die Natur ein ziemlich hinterlistiger Gegner des Statistikers ist. Möglicherweise hat Gott auch einfach Spaß daran, uns hereinzulegen.

Abschließend sei bemerkt, daß dieser Aufsatz der Versuch der Diskussion eines Falles mathemathischer Modellierung ist. Mit dieser wird in der Praxis viel Schindluder getrieben. Zumindestens die Entscheidung, ob ein Modell mit den Daten kompatibel ist, sollte möglich sein.

Literaturverzeichnis

[1] Louis Bachelier. *Théorie de la spéculation.* Gauthier-Villars, Paris, 1900.

[2] Jakob Bernoulli. *Ars conjectandi (Wahrscheinlichkeitsrechnung), I., II., III. und IV. Theil (1713)*, Band 107 von *Ostwalds Klassiker der exakten Wissenschaften.* Verlag Harri Deutsch, Frankfurt am Main, 2002. Nachdruck der Ausgabe 1713.

[3] Girolamo Cardano. *Liber de Ludo Alea.* Piper Verlag, 1663.

[4] Antoine Augustin Cournot. *Exposition de la théorie des chances et des probabilités.* Hachette, Paris, 1843.

[5] Antoine Augustin Cournot. Exposition de la théorie des chances et des probabilités. In *OEuvre complètes.* Vrin, Paris, 1973-1984. 10 Bände.

[6] R.H. Daw and E.S.. Pearson. Abraham De Moivre's 1733 derivation of the normal curve: A bibliographical note. *Biometrika*, 59(3):677–680, 1972.

[7] B. De Finetti. La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. Poincaré*, 7:1–68, 1937.

[8] P. Eichelsbacher. Eine Diskussion der Faustregel von Laplace. *Stochastik in der Schule*, 22–27, 2001.

[9] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme gefordete Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 17:549–560, 1905.

[10] W. Feller. *The foundations of statistics*, Volume 1. John Wiley & Sons, New York, 1968.

[11] P. Gänßler and W. Stute. *Wahrscheinlichkeitstheorie.* Hochschultext. Springer-Verlag, Berlin-Heidelberg-New York, 1977.

[12] B.W. Gnedenko. *Lehrbuch der Wahrscheinlichkeitsrechnung.* Akademie-Verlag, Berlin, 5. Auflage, 1968.

[13] M. Hampe. *Die Macht des Zufalls.* wjs verlag, Wolf Jobst Siedler jr., Berlin, erste Auflage, 2006. ISBN 3-937989-23-4.

[14] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel. *Robust Statistics.* Wiley Series in Probability and Mathematical Statistics. Wiley & Sons, New York, 1986.

[15] HC. Körper and N.A. Whitney-Desaults. Astragalus bones: Artefacts or ecofacts? *Pacific Coast Archeological Society Quarterly*, 35(2-3):69–80, 1999. URL `http://www.pcas.org/Vol35N23/3523Koerper.pdf`.

[16] W. Heisenberg. *Quantentheorie und Philosopie.* Reclam, Stuttgart, 2000.

[17] P. Huber. Projection pursuit. *Ann. Statist.*, 13:435–475, 1985.

[18] P.J. Huber. *Robust Statistics.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York etc., 1981.

[19] B. Inhelder. Die Entwicklung der Grundbegriffe von Zufall und Wahrscheinlichkeit bei Kindern. In B. Inhelder und H.H. Chipman, Editoren, *Von der Kinderwelt zur Erkenntnis der Welt*, 98–114. Akademische Verlagsgesellschaft, Wiesbaden, 1978.

[20] J.G. Kemeney. Fair bets and inductive probabilities. *Journal of Symb. Logic*, 20:263–273, 1955.

[21] A. Kolmogoroff. *Grundbegriffe der Wahrscheinlichkeitsrechnung. Reprint der Erstaufl. Berlin 1933.*, Band 2 von *Ergebnisse der Mathematik und ihrer Grenzgebiete.* Springer-Verlag, Berlin-Heidelberg-New York, 1933,1973.

[22] W. Krämer. *Denkste! Trugschlüsse aus der Welt der Zahlen und des Zufalls.* Piper, München, Zürich, zweite Auflage, 2002. ISBN 3-492-22443-1.

[23] M. Le Comte Laplace. *Théorie analytique des probabilités.* Nummer 57. Courcier, Imprimeur-Libraire pour les Mathemathique, 1812.

[24] K. Mainzer. *Wie das Neue in die Welt kommt.* Verlag C.H. Beck, München, 2007. ISBN 978 3 406 55428 5.

[25] B.B. Mandelbrot and R.L. Hudson. *Fraktale und Finanzen, Märkte zwischen Risiko, Rendite und Ruin.* Piper Verlag, München, 2005.

[26] G. Marsaglia. Random numbers fall mainly in the planes. *Proc. Nat. Acad. Sci.*, 25–28, 1968.

[27] G. Marsaglia. The structure of linear congruential sequences. In Zaremba S.K., Editor, *Applications of Number Theory to Numerical Analysis*, 249–285. Academic Press, 1972.

[28] Abraham De Moivre. Approximiatio ad Summam Terminorum Binomii in Seriem expansi. 1733.

[29] S.K. Park and K.W. Miller. Random number generators: good ones are hard to find. *Comm. Assoc. Comput. Mach.*, 31:1192–1201, 1988.

[30] J. Piaget and B. Inhelder. *Die Psychologie des Kindes.* München, 1978. 4. Auflage 1991.

[31] Platon. Philebos. In W.F. Otto, E. Grassi, and G. Plamböck, editors, *Sämtliche Werke*, Band 5. Rowohlt Verlag, Hamburg, 1957-1959. In der Übersetzung von Friedrich Schleiermacher mit der Stephanus-Numerierung.

[32] Hans Richter. Zur Grundlegung der Wahrscheinlichkeitstheorie. I. *Math. Ann.*, 125:129–139, 1952.

[33] Hans Richter. Zur Grundlegung der Wahrscheinlichkeitstheorie. II. *Math. Ann.*, 125:223–234, 1953a.

[34] Hans Richter. Zur Grundlegung der Wahrscheinlichkeitstheorie. III: Die Begründung des Additions- und des Multiplikationssatzes. *Math. Ann.*, 125:335–343, 1953b.

[35] Hans Richter. Zur Grundlegung der Wahrscheinlichkeitstheorie. IV: Wahrscheinlichkeitsrechnung. *Math. Ann.*, 126:362–374, 1953c.

[36] Hans Richter. Zur Grundlegung der Wahrscheinlichkeitstheorie. V: Indirekte Wahrscheinlichkeitstheorie. *Math. Ann.*, 128:305–339, 1954.

[37] Hans Richter. *Wahrscheinlichkeitstheorie.* (Die Grundlehren der Mathematischen Wissenschaften, Bd. 86.) Berlin-Göttingen-Heidelberg: Springer-Verlag,1956.

[38] Hans Richter. Eine einfache Axiomatik der subjektiven Wahrscheinlichkeit. In *Sympos. math. 9, Calcolo Probab., Teor. Turbolenza 1971, 59-77.* 1972.

[39] Hans Richter. Die historische und logische Verbindung zwischen Wettbegriff und Wahrscheinlichkeitsbegriff. I. *Bl., Dtsch. Ges. Versicherungsmath.*, 11:311–318, 1974a.

[40] Hans Richter. Die historische und logische Verbindung zwischen Wettbegriff und Wahrscheinlichkeitsbegriff. II. *Bl., Dtsch. Ges. Versicherungsmath.*, 11:481–490, 1974b.

[41] Hans Richter. Die historische und logische Verbindung zwischen Wettbegriff und Wahrscheinlichkeitsbegriff. III. *Bl., Dtsch. Ges. Versicherungsmath.*, 12:3–14, 1975.

[42] B.D. Ripley. *Stochastic Simulation.* Wiley Interscience, New York, 1987.

[43] B.D. Ripley. *Statistical Inference for Spatial Processes.* Cambridge University Press, Cambridge New York New Rochelle Melbourne Sidney, 1988.

[44] I.R Savage. *Statistics: Uncertainty and behavior.* Mass., Boston, 1968.

[45] L.J Savage. *The foundations of statistics.* New York, 1954.

[46] I. Schneider. Der Mathematiker Abraham de Moivre (1667-17549. *Archive for history of exact sciences*, 5:177–317, 1968.

[47] I. Schneider. Die Rückführung des Allgemeinen auf den Sonderfall - eine Neubetrachtung des Grenzwertsatzes für binomiale Verteilungen von Abraham de Moivre. In Joseph W. DAUBEN et alii, Editoren, *History of Mathematics: States of the Art, Flores quadrivii, - Studies in Honor of Christoph J. Scriba*, 263–275. Academic Press San Diego, 1996.

[48] C.P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie. (Randomness and probability. An algorithmic foundation of probability theory).* Number 218 in Lecture Notes in Mathematics. Springer-Verlag, Berlin-Heidelberg-New York, 1971.

[49] G. Shafer. From Cournot's principle to market efficiency. Technical report, The Game-Theoretic Probability and Finance Project, 2006. URL `http://www.probabilityandfinance.com`.

[50] N.N. Taleb. *Der Schwarze Schwan.* Hanser, 2008.

[51] Karl Valentin. Orchesterprobe. In *Karl Valentin & Liesl Karlstadt: Die beliebtesten Kurzfilme.* Film 101 München, 2009. DVD.

[52] G. Winkler. Eine Anwendung des Zwischenwertsatzes auf das periodische Verhalten dynamischer Systeme: der Satz von Šarkovskii. *DdM*, 4: 304–320, 1986.

[53] H. Ziezold. Die Faustregel $np(1-p) \overset{>}{\approx} 9)$. *Stochastik in der Schule*, 22–27, 2001.