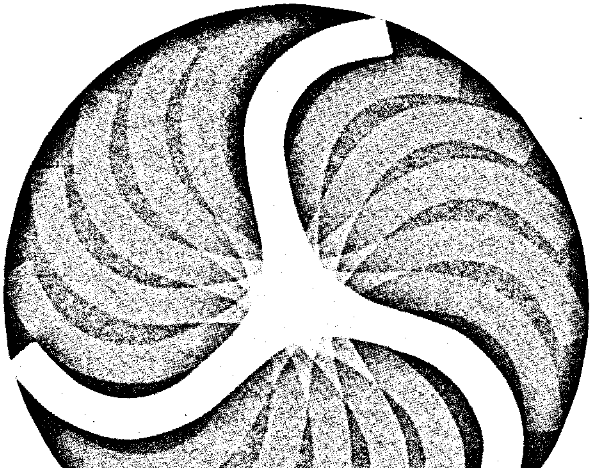
ZUR ENTWICKLUNG DES WAHRSCHEINLICHKEITSBEGRIFFS — DAS ANWENDUNGSPROBLEM IN DER WAHRSCHEINLICHKEITSTHEORIE AUS DIDAKTISCHER SICHT



Institut für Didaktik der Mathematik
der Universität Bielefeld

ZUR ENTWICKLUNG DES WAHRSCHEINLICHKEITSBEGRIFFS — DAS ANWENDUNGSPROBLEM IN DER WAHRSCHEINLICHKEITSTHEORIE AUS DIDAKTISCHER SICHT

Heinz Steinbring

Materialien und Studien Band 18

Institut für Didaktik der Mathematik
der Universität Bielefeld

Fotovorlage: Maria Otte
Gesamtherstellung: Robert Bechauf, Bielefeld

INHALT

Seite

Seite

Seite

Seite

Seite

VORWORT

In der vorliegenden Arbeit geht es darum, anhand der Entwicklung des Wahrscheinlichkeitsbegriffs zum Verständnis der Probleme der Herausbildung und Aneignung mathematischer Begriffe insgesamt beizutragen.

Die grundlegende Schwierigkeit eines jeden Lernprozesses könnte man kurz so formulieren: Wie ist es möglich, sich <u>neues</u> Wissen anzueignen, da letztlich die Neuartigkeit eines jeden Wissensbestandes unvergleichbar mit bisherigen Erkenntnissen sein wird? Das Neue kann nicht vollständig aus schon Bekanntem deduziert werden, es lebt vielmehr aus seiner Eigenständigkeit.

Gerade dem Wahrscheinlichkeitsbegriff ist meist ausschließlich eine Hilfsfunktion auf der Grundlage einer kausaldeterministischen Position zugesprochen worden. Anstatt das Neue und die diesem Begriff eigenen Besonderheiten zum Ausgangspunkt der Untersuchung zu machen, wurde seine Bedeutung oft aus "bekannten" Auffassungen abgeleitet, versuchte man, ihn als ein subjektives Hilfsmittel einer ansonsten prinzipiell strikt kausal bestimmten Erkenntnis zu erklären.
Man muß jedoch gewissermaßen genau umgekehrt vorgehen und sich auf die "Ursprünglichkeit" der Wahrscheinlichkeit einlassen. "I think chance is a more fundamental conception than causality; for whether in a concrete case a cause-effect relation holds or not can only be judged by applying the laws of chance to the observation." (Born, 1964, S. 47)

Hiermit wird ein Grundproblem dieser Arbeit angesprochen. Die Eigenständigkeit der Wahrscheinlichkeit zur Richtschnur der Analyse zu machen, hat zur Folge, sich die Entwicklungen und Veränderungen der inhaltlichen Bedeutung der Wahrscheinlichkeitstheorie, wie sie in den Anwendungen zum Ausdruck kommt, genauer anzusehen. Dazu werden im II. Kapitel ausführlich drei historische Fallstudien bearbeitet. Im III. Kapitel wird eine systematische Darstellung der Entwicklung des wahrscheinlichkeitstheoretischen Anwendungsbezugs gegeben. Auf der Grundlage der historischen und systematischen Diskussion werden dann im IV. Kapitel Schlußfolgerungen für den Stochastikunterricht gezogen.

Die vorliegende Dissertation entstand im Rahmen der Arbeit der Arbeitsgruppe "Mathematiklehreraus- und -weiterbildung" am Institut für Didaktik der Mathematik der Universität Bielefeld. Sie gehört in den Diskussionskontext dieser Gruppe und ist ohne diesen schwerlich denkbar. Für kritische Ratschläge und Unterstützung möchte ich meinen Kollegen Rolf Biehler, Rainer Bromme, Gerd v. Harten, Niels Jahnke, Thomas Mormann und Matthias Paul meinen Dank sagen. Zu besonderem Dank bin ich Herrn Prof. Michael Otte verpflichtet, der über Hilfe und Förderung in der Betreuung dieser Dissertation hinaus, mich in Probleme und Aufgaben der Mathematikdidaktik eingeführt hat.

I. EINFÜHRUNG: WAHRSCHEINLICHKEITSTHEORIE UND ENTWICKLUNGSSTANDPUNKT

Jede Untersuchung, die sich über den Rahmen eng fachwissenschaftlicher Grenzen hinaus mit Fragen zum Status des theoretischen Wissens und der Rolle theoretischer Begriffe, z.B. in sozialen, philosophischen, erkenntnistheoretischen oder auch pädagogischen Kontexten, befaßt, sieht sich zusätzlich im Vergleich zur betreffenden Fachwissenschaft grundsätzlich neuartigen Problemen gegenüber. Vor dieser Schwierigkeit steht letztlich auch die Didaktik der Mathematik. Als eine interdisziplinäre Wissenschaft hat sie sich bei der Analyse des mathematischen Wissens im Hinblick auf die Vermittlungsproblematik um vielfältige unterschiedliche Dimensionen des Wissens, um soziale, pädagogisch-psychologische, fachliche, wissenschaftstheoretische und auch historische Aspekte und deren Koordination zu bemühen.

Gerade der Versuch, eine didaktische Analyse des Wahrscheinlichkeitsbegriffs vorzunehmen, macht den Aspektereichtum zu berücksichtigender Fragestellungen deutlich. Selbst eine erste Eingrenzung der Untersuchung auf primär historische, fachmathematische und erkenntnistheoretische Dimensionen, wie sie in dieser Arbeit angestrebt wird, trifft auf mannigfaltige Probleme.

So zeigt ein erster flüchtiger Gesamtüberblick, daß es wohl kaum einen anderen mathematischen Begriff gibt, der eine ähnliche Vielfalt verschiedenartiger Definitionen und gegensätzlicher Interpretationen aufweist, wie der Begriff der Wahrscheinlichkeit. Da ist z.B. die Bernoullische Gleichwahrscheinlichkeit und die klassische Laplacesche Definition, die Häufigkeitstheorie von Richard von Mises, die logische Wahrscheinlichkeit, etwa von John M. Keynes, und schließlich die bekannte axiomatische Begründung der Wahrscheinlichkeit durch Kolmogoroff.

Berücksichtigt man zudem die schon lang andauernde heftige Diskussion um die subjektive und objektive Wahrscheinlichkeitsauffassung, so kann man insgesamt Kendall nur zustimmen, der sagte: "Few branches of scientific method have been subject to so much difference of opinion as the theory of probability." (Kendall, 1949, S.101) Dies macht den Wahrscheinlichkeitsbegriff zwar zu einem sehr interessanten, aber gleichzeitig wohl zu einem der problematischsten Begriffe für die Schule überhaupt. Denn die didaktische Diskussion kann sich nicht grundsätzlich den Problemen der Bedeutung und der Interpretation von mathematischen Begriffen entziehen. Bezogen auf den Wahrscheinlichkeitsbegriff und seine unterschiedlichen Definitionen kann dies jedoch nicht heißen, daß die Didaktik etwa eine "Lösung" für die eben dargelegte Vielfalt erkenntnistheoretischer Fragen liefern sollte oder gar könnte. Sie muß sich aber um einen angemessenen Standpunkt gegenüber diesen Problemen in Hinsicht auf die Vermittlungsproblematik der Wahrscheinlichkeitstheorie bemühen. Ja, die Einbeziehung einer didaktischen Perspektive in den Rahmen erkenntnistheoretischer Fragestellungen liefert gewissermaßen eine zusätzliche Ebene, an welcher sich ganz allgemein Bearbeitungen dieser Fragen in neuer Weise orientieren können.

Wie wird nun in der wissenschaftstheoretischen Diskussion mit der Divergenz und Vielfalt des Wahrscheinlichkeitsbegriffs umgegangen, und welche vorläufigen Hinweise für eine didaktisch ausgerichtete Untersuchung lassen sich evtl. hieraus gewinnen? Ohne dabei zu weit ins Detail gehen zu können, was weder in aller Ausführlichkeit möglich, noch zentraler Zweck unserer Arbeit ist, läßt sich zu dieser Frage relativ kurz die folgende Einschätzung geben.

Einerseits wird eher der Unterschied, ja Gegensatz zwischen den vielen Wahrscheinlichkeitstheorien, vor allem entlang der Trennungslinie zwischen subjektiven bzw. personalistischen und objektiven Interpretationen betont. Welche dieser Theorien man für die eigentlich angemessene hält, wird letzt-

lich zu einer Art "Glaubensbekenntnis" des jeweiligen Vertreters. Zum anderen gibt es auch vereinzelte Bestrebungen, diese strikte Gegensätzlichkeit insofern zu relativieren, als eine vergleichende Klassifizierung aller dieser Wahrscheinlichkeitstheorien versucht wird. Zu nennen ist in dieser Hinsicht einmal die Arbeit von Good (1950), in der u.a. eine Klassifikation beispielhaft für insgesamt sieben Theorien (Venn, v.Mises, Gleichwahrscheinlichkeit, Jeffrey, Kolmogoroff, Ramsey und Koopmann) entlang den folgenden vier Kriterien durchgeführt wird: "axiomatisch versus nicht-axiomatisch", "objektiv versus subjektiv", "Grad des vernünftigen Dafürhaltens (degrees of belief) versus relative Häufigkeit (frequency bzw. statistical)" und "numerisch versus nicht-numerisch".
Eine abgekürzte Klassifikation läßt sich nun anhand eines zweidimensionalen Schemas darstellen, in dem "ja" - "nein" Markierungen an entsprechenden Stellen erscheinen. (vgl. auch für die ausführlichere Darstellung Good, 1950, S.6-12)
Des weiteren sei auch auf das Buch von Fine (1973) hingewiesen. Fine unternimmt einen Klassifikationsversuch von insgesamt elf Wahrscheinlichkeitstheorien entsprechend fünf Hauptkriterien (mit jeweils bis zu fünf Unterkriterien). (vgl. Fine, 1973, S.1-10)

Ohne nun diese Klassifikationsversuche näher untersuchen zu wollen, kann man hierzu relativ allgemein anmerken, daß neben der Gefahr, aufgrund relativ äußerlicher Kriterien keine differenzierte Gesamtperspektive und inhaltlichen Vergleich für die Vielfalt möglicher Wahrscheinlichkeitstheorien zu erhalten, sich vor allem herausstellt, daß in erster Linie das Verhältnis der jeweiligen Theorie zur Realität bzw. die Rolle und Funktion, welche den Anwendungen zugesprochen wird, in dem Versuch einer "einheitlichen" Beschreibung der Wahrscheinlichkeitstheorien eine besondere Bedeutung zukommt. Denn rein "oberflächlich", d.h. etwa nur den Kalkül oder auch den axiomatischen Aufbau der Theorie betreffend, sind unter den vielen Wahrscheinlichkeitstheorien kaum relevante Unterschiede festzustellen. So bemerkt etwa Gillies im Anschluß an seinen Vergleich der Theorie von Keynes mit der von v.Mises: "It is worth noting that this complete disagreement on all the philosophical issues is accompanied by complete agreement on the mathematical side. Both authors in their books derive the binomial theorem, the Laws of Large Numbers, the Central Limit theorem; that is to say all the main results of the mathematical calculus." (Gillies, 1973, S.14/5)

Und zum Vergleich von Bayesianern und Nicht-Bayesianern schreibt etwa Weber beispielhaft: "The axioms of probability theory and the algebraic rules for manipulating probabilities that follow from them are generally accepted by both classical and Bayesian statisticians - even Bayes's theorem is not questioned with respect to its algebraic validity. The controversy concerns only the definition and interpretation of probabilities, not their algebraic manipulation." (Weber, 1973, S.1)

Die Vielfalt der Wahrscheinlichkeitstheorien besser verstehen und entlang gemeinsam beziehbarer Kriterien vergleichen zu wollen, bedeutet also, sich genauer das Verhältnis der jeweiligen Wahrscheinlichkeitstheorie zur Realität anzusehen. Die je verschiedenen Aspekte wahrscheinlichkeitstheoretischer Interpretationen sind in den unterschiedlichen Auffassungen zu den Anwendungen zu suchen. Hieraus ergibt sich ganz allgemein die große Bedeutung des Anwendungsproblems für die Analyse des Wahrscheinlichkeitsbegriffs.

Die Diskussion des Anwendungsproblems bringt nun neuartige zusätzliche Schwierigkeiten mit sich; und zwar zum einen relativ allgemeine Fragen des Verhältnisses von theoretischen zu empirischen Aspekten, die zum anderen mit spezifisch wahrscheinlichkeitstheoretischen bzw. statistischen Komplikationen behaftet sind. Den Anwendungsbezug, also das Verhältnis von Theorie zu Empirie, etwa rein statisch, zu interpretieren, bedeutet hinsichtlich des Klassifikationsproblems bzw. eines vergleichbaren Standpunktes gegenüber der Vielfalt der Wahrscheinlichkeitstheorien letztlich keine

Verbesserung. Einen solchermaßen statisch-fixierten Erklärungsversuch des wahrscheinlichkeitstheoretischen Anwendungsbezuges hat etwa Gillies in seiner Arbeit (1973) unternommen und in der Tat werden hierüber keine möglichen Zusammenhänge unter verschiedenen Auffassungen und keine ansatzweise "einheitliche" Perspektive sichtbar. Die Anwendungen anderer Theorien bleiben letztlich gegensätzlich und nicht inhaltlich vergleichbar.

Darüber hinaus erlauben es die spezifisch wahrscheinlichkeitstheoretischen Besonderheiten in diesem Problemzusammenhang auch gar nicht, den Anwendungsbezug als absolut und fixiert anzusehen. Wie sich im Verlaufe dieser Arbeit noch deutlicher herausstellen wird, muß man grundlegender Schwierigkeiten wegen, aufgrund von "Zirkularitäten" in den Begründungen und den Anwendungen des Wahrscheinlichkeitsbegriffs, von einer <u>dynamischen Auffassung</u> des Theorie-Empirie-Verhältnisses ausgehen.

Teilweise entspricht diese Notwendigkeit einer dynamischen Sichtweise ganz allgemein der Erkenntnis, daß Theorien nicht statisch, etwa als Systeme von Aussagen, sondern nur in ihrer dynamischen Entwicklung verstanden werden können. Dies ergibt sich aus den bestehenden Differenzen von theoretischen zu empirischen Momenten, aus Paradoxien, die in einem statischen Erklärungsversuch von Theorien auftreten. Ausdruck hierfür sind in der Wahrscheinlichkeitstheorie die immer wieder als zirkulär empfundenen Definitionsversuche des Wahrscheinlichkeitsbegriffs in Form der Gleichwahrscheinlichkeit etwa oder auch der frequentistischen Interpretation. Darüber hinaus haftet auch den Anwendungen der Wahrscheinlichkeitstheorie ein zirkulärer Charakter an; wie wir noch ausführlich analysieren werden, unterliegen etwa theoretische Aussagen zum Anwendungsbezug <u>letztlich</u> selbst wieder einer statistischen Beurteilung.

In seiner Arbeit: "Zum Verhältnis von Wissensentwicklung und Begründung in der Mathematik - Beweisen als didaktisches

Problem" (1979) hat Jahnke im Zusammenhang dieser Problematik die Notwendigkeit einer dynamischen Auffassung von Theorieentwicklung herausgearbeitet, und er faßt die zentrale Konsequenz folgendermaßen zusammen:"Eine der grundlegenden Einsichten ... bestand darin, daß die mit dem Auftreten theoretischer Terme in empirischen Theorien verbundene Paradoxie bzw. Zirkularität, daß nämlich die theoretischen Terme die Theorie bestimmen und umgekehrt die Theorie die Bedeutung der theoretischen Terme festlegt, nur dann auflösbar wird, wenn man Theorien als sich entwickelnde versteht." (Jahnke, 1979, S.118) Wie angemerkt, kommt in der Wahrscheinlichkeitstheorie zusätzlich ein "Zirkel" im Anwendungsbezug hinzu.

Soweit zu unserer eher vorläufigen und skizzenhaften Beschreibung und Herleitung eines Grundproblems für die zentrale Entwicklung eines relativ "einheitlichen" Verständnisses der Wahrscheinlichkeitstheorie. Dabei hat sich die Perspektive der Analyse quasi von einer mehr "äußerlichen" Klassifikation der Theorien verschoben zur Untersuchung eines der wichtigsten Vergleichskriterien, dem Anwendungsbezug der Wahrscheinlichkeitstheorie.
Grundlegender Schwierigkeiten im Verhältnis von Wahrscheinlichkeitstheorie zu den empirischen Anwendungen wegen ist es jedoch notwendig, und dies ist das wesentliche Resumée unserer Vorüberlegungen, <u>eine dynamische Sichtweise auf den wahrscheinlichkeitstheoretischen Anwendungsbezug</u> einzunehmen.

Mit dieser Kennzeichnung einer dynamischen Betrachtungsweise des wahrscheinlichkeitstheoretischen Anwendungsbezuges ist natürlich im Grunde ein äußerst umfangreiches Programm angesprochen. Denkt man etwa vor allem an die ungeheure Vielfalt und Divergenz der verschiedensten Anwendungsverfahren, -regeln und -rezepte der Statistik, oder auch an die so zahlreichen und unterschiedlichen Anwendungsbereiche der Wahrscheinlichkeitstheorie in der Physik, der Chemie, der Biologie, den Sozialwissenschaften usw., so wird deutlich, daß man in bezug auf die Entwicklung des Anwendungsproblems der Wahrscheinlichkeitstheorie bestimmte Fragestellungen und Teilprobleme ausgrenzen muß und nur diese zunächst wird bearbeiten können.

In welcher Hinsicht wollen wir nun versuchen, Teilprobleme des allgemeinen wahrscheinlichkeitstheoretischen Anwendungsproblems auszusondern? Um dem wesentlichen Anliegen einer dynamischen Perspektive auf die Anwendungen gerecht zu werden, d.h. etwa die in der Entwicklung sichtbar werdenden Veränderungen, Erweiterungen und Verallgemeinerungen dieser Beziehung besser zu verstehen, wollen wir uns zentral als der wichtigsten Materialgrundlage der gesamten Arbeit der Geschichte der Wahrscheinlichkeitstheorie unter dieser Fragestellung nach der Rolle der Anwendungen zuwenden. Auch dies wäre sicherlich ohne zusätzliche Einschränkungen ein zu umfangreiches Unternehmen. Deshalb erscheint es als sinnvoll, sich auf insgesamt drei wichtige historische Perioden zu konzentrieren, in denen sich gewisse zentrale Veränderungen und Umbrüche in dem Verhältnis von Wahrscheinlichkeitstheorie und Realität zeigen, wie sie in den Erweiterungen des Anwendungsbereichs zum Ausdruck kommen. Diese Erweiterungen in den Anwendungen bewirken nun auch auf der Ebene der Wahrscheinlichkeitstheorie selbst Entwicklungen und bringen Verbesserungen hinsichtlich vorhandener Methoden und Konzepte mit sich. Zudem gewinnt das Verhältnis der Wahrscheinlichkeitstheorie zu den realen Anwendungen im Verlaufe dieser Veränderungen selbst eine immer differenziertere Struktur, was verbesserte Einsichten in die Bedeutung und den Status des sich entwickelnden Wahrscheinlichkeitsbegriffs erlaubt.

Die drei historischen Perioden, die wir in dieser Arbeit untersuchen wollen, betreffen erstens Bernoulli und den Beginn der Überlegungen und Beweise zum (schwachen) Gesetz der großen Zahlen, dann zweitens die Anfänge der Fehlertheorie und eng damit verbunden die Aufstellung der Normalverteilung als einer grundlegenden wahrscheinlichkeitstheoretischen Gesetzmäßigkeit durch Gauß, und drittens die ersten Anwendungen der Wahrscheinlichkeitstheorie in der Physik, wie sie vor allem in den Arbeiten Boltzmanns zur kinetischen Gastheorie und statistischen Mechanik behandelt werden.

Die Auswahl dieser drei historischen Etappen zeigt, daß wir uns auf der Grundlage der historischen Einzelstudien primär in allgemein erkenntnistheoretischer Weise mit dem Verhältnis von Wahrscheinlichkeitstheorie zur Realität und den Veränderungen dieses Verhältnisses auseinandersetzen wollen. Es geht uns darum, das theoretische Objektverständnis der Wahrscheinlichkeit und seine besondere Problematik besser zu verstehen. Demgegenüber sind gerade die in der Vielfalt statistischer Anwendungsverfahren zum Ausdruck kommenden Besonderheiten und Schwierigkeiten ausgeklammert; eine Bearbeitung dieser Probleme würde etwa, was die historischen Aspekte betrifft, vor allem die Analyse der Anwendungen der Wahrscheinlichkeitstheorie bzw. Statistik in den Sozialwissenschaften samt ihrer Entwicklungen voraussetzen.

Im Rahmen dieser relativ grundlegenden Theorie-Empirie-Beziehung werden wir gleichermaßen auch das Subjekt-Objekt-Verhältnis der Wahrscheinlichkeitstheorie relativ allgemein und grundsätzlich angehen. Das heißt etwa, genauso wenig, wie wir die Fragen und Probleme der (schließenden) Statistik und statistischer Entscheidungstheorien in ihren vielfältigen Details untersuchen werden, ohne jedoch prinzipiell die sich hierin zeigende allgemeine Struktur des Anwendungsbezuges zu vernachlässigen, genauso werden wir zwar die Kontroverse um die subjektive bzw. objektive Auffassung der Wahrscheinlichkeitstheorie in ihren technischen Feinheiten nicht diskutieren können und wollen, jedoch dabei in grundlegender Weise die Subjekt-Objekt-Dialektik einer jeden wissenschaftlichen Tätigkeit als wichtiges Analysekriterium heranziehen. Dementsprechend wird im folgenden in prinzipieller Weise auf das Anwendungsproblem der Wahrscheinlichkeitstheorie Bezug genommen und sowohl die statistischen Schwierigkeiten als auch die subjektive versus objektive Wahrscheinlichkeitskontroverse aufgegriffen, ohne daß wir diese jedoch in ihren besonderen Aspekten verfolgen werden.

Der Schwerpunkt unserer Arbeit wird also primär auf dem Verhältnis der Wahrscheinlichkeits<u>theorie</u> zur Realität liegen, also die Frage nach den Entwicklungen und Veränderungen des wahrscheinlichkeitstheoretischen Anwendungsgegenstandes betreffen. Dies wird zwar letztlich nicht ohne grundlegende statistische Betrachtungsweisen möglich sein, diese würden jedoch für sich genommen darüber hinaus eine eigenständige und umfassende Analyse im Hinblick auf die vielen Besonderheiten des Anwendungsproblems der Statistik erforderlich machen. Wir nehmen also insgesamt eine allgemeinere erkenntnistheoretische Position zum Anwendungsproblem in der Wahrscheinlichkeitstheorie ein, was es ermöglicht, die hier auftretenden spezifischen Probleme grundlegender zu beurteilen.

Die Einnahme eines solchermaßen stärker grundsätzlichen Standpunktes gegenüber dem Verhältnis von Wahrscheinlichkeitstheorie zu den Anwendungen scheint uns gerade auch im Hinblick auf die didaktischen Probleme zunächst sinnvoll zu sein, denn im Unterricht der Stochastik kann es u.E. nicht in erster Linie darum gehen, den Schülern einfach eine Vielfalt von unterschiedlichen statistischen Anwendungsverfahren zu vermitteln, sondern darüber hinaus ist auf grundlegende Vorstellungen zum Verhältnis von Mathematik zur Realität zu achten, was exemplarisch in besonders prägnanter Weise anhand der Wahrscheinlichkeitstheorie erreicht werden könnte.

Des weiteren kommt eine dynamische Betrachtungsweise der Begriffsentwicklung insgesamt dem prozessualen Charakter des Begriffslernens entgegen; zudem ist diese Herangehensweise auf ein verbessertes Verständnis dessen ausgerichtet, welche inhaltliche Bedeutung der Wahrscheinlichkeit zukommt, was einer didaktischen Untersuchung nur zugute kommen kann.

Auch läßt sich erwarten, und mit dieser Bemerkung möchten wir unsere einführenden Überlegungen abschließen, daß in einem so grundlegenden Rahmen die Diskussion der Dynamik wahrscheinlichkeitstheoretischer Anwendungen die Besonderheiten der Wahrscheinlichkeitstheorie verstärkt mit ähnlichen Fragestellungen anderer Theorien vergleichbar macht, und daß auf diese Weise auch Anstöße für die wissenschaftstheoretische Analyse des Wahrscheinlichkeitsbegriffs vielfältig fruchtbar werden können.

II. AUS DER GESCHICHTE DER WAHRSCHEINLICHKEITSTHEORIE - DIE ROLLE DER ANWENDUNGEN FÜR DIE ENTWICKLUNG DES WAHRSCHEINLICHKEITSBEGRIFFS

Mit dem Vorhaben, historische Untersuchungen für erkenntnistheoretische und didaktische Zwecke zu nutzen, steht man vor der Entscheidung, sich unter dieser Perspektive auf wichtige Fragen an die Geschichte zu beschränken und darauf zu konzentrieren. Uns geht es in erster Linie darum, genauer die Dynamik der Wahrscheinlichkeitstheorie zu verstehen, und wir wollen entsprechend den Ausführungen im ersten Kapitel dazu die Entwicklung des Verhältnisses der Theorie zu ihren Anwendungen analysieren. Den roten Faden unserer historischen Betrachtung stellt somit der Anwendungsbezug dar. Uns kann es dementsprechend nicht auf Vollständigkeit in allen historischen Details ankommen; vielmehr wollen wir speziell aus der Analyse des sich historisch entwickelnden Anwendungsproblems Nutzen für didaktische Fragestellungen ziehen, und wir hoffen, auch umgekehrt dadurch gewisse Aspekte in der Geschichte der Wahrscheinlichkeitstheorie besser einschätzen zu können.

II.1. Symmetrie und Zufall

Wir wollen den Zusammenhang zwischen praktischen und theoretischen Problemstellungen und der Entwicklung des Wahrscheinlichkeitsbegriffs zunächst zur Zeit dessen Herausbildung betrachten. Der Philosoph Ian Hacking äußert in seinem Buch "The Emergence of Probability" zu diesem Problem die folgende Ansicht: "There are two ways in which a science develops: in response to problems which itself creates, and in response to problems that are forced on it from the outside." Und bezogen auf die Entstehung und Entwicklung des Wahrscheinlichkeitsbegriffs fährt er fort: "Only very recently has probability theory been hardy enough to create its own problems and to generate its own programme of research. The stimulus used to come from other disciplines." (Hacking, Ähnlich hebt der Wahrscheinlichkeitstheoretiker Michel Loève in seinem kurzen Abriß zur Geschichte der Wahrscheinlichkeitstheorie die lang andauernde und enge Verknüpfung zwischen der Entwicklung wahrscheinlichkeitstheoretischer Methoden und außermathematischen Problemen hervor:"Durant près de trois siècles, la motivation des concepts et des problèmes du Calcul des probabilités vint surtout de l'extérieur. ... Ce n'est que récemment durant le vingtième siècle, que le Calcul des probabilités s'est libéré de son rôle d'instrument, et est devenu un membre de plein droit de la famille mathématique." (Loève, 1978, S.279)

Die hier ganz allgemein beschriebene Spannung zwischen den praktischen Anforderungen und Problemen einerseits und der Herausarbeitung theoretischer Methoden andererseits als einer der wichtigsten Triebfedern für die Entwicklung der Wahrscheinlichkeitstheorie, spielt gerade für die Geburt des mathematischen Wahrscheinlichkeitsbegriffs in der Mitte des 17. Jhdts. eine überaus bedeutsame Rolle.

Umfassend analysieren und erklären läßt sich die Herausbildung dieses Begriffs wohl nur auf dem Hintergrund der Entstehung der neuzeitlichen Wissenschaften insgesamt während dieser Zeit und im Zusammenhang der in dieser Periode stattfindenden gesellschaftlichen, ökonomischen und wissenschaftlichen Veränderungen und Umwälzungen.

So versucht etwa Dirk Struik in seinem Aufsatz: "Soziologische Aspekte des Ursprungs der mathematischen Wahrscheinlichkeitstheorie" auf die von ihm gestellte Frage:"Wie kommt es, daß kurz nach 1650 fast gleichzeitig sowohl die mathematische Wahrscheinlichkeitstheorie als auch die mathematische Statistik entstehen, obgleich Glücksspiele, Versicherungen, Bevölkerungsstatistiken schon lange bekannt waren?" (Struik, 1977, S.1) durch soziologische Betrachtungen Hinweise für eine Antwort zu geben. Er sieht eine Erklärung für die Entstehung der Wahrscheinlichkeitstheorie in diesem Zeitpunkt und "fast nur in den Ländern nahe der Nordsee, genau in den Ländern, wo das Bürgertum eine hervorragende Stellung einnahm" (Struik, 1977, S.1), vor allem in dem "Übergang von der Rechenhaftigkeit des Bürgertums des 16. Jhdts. zu dem, was ich gern als Mathematisierung oder besser Mathematisierungsmethode oder Mathematisierungsphilosophie bezeichne, die typisch war für die Führungsschicht des Bürgertums des 17. Jhdts., besonders der letzten Hälfte." (Struik, 1977, S.8). Hervorgerufen wurde dieser Übergang durch eine vielfältig wechselwirkende Beziehung zwischen den sich stellenden neuartigen praktischen Problemen und den zu ihrer Bearbeitung schon vorhandenen und noch zu entwickelnden Methoden und technischen Geräten handwerklichen Ursprungs; diese Problematik erforderte dann auch insgesamt die Weiterentwicklung und Nutzung der Wissenschaften für die tägliche Praxis.

Und Ian Hacking stellt sich in seiner Arbeit zur Entstehung der Wahrscheinlichkeitstheorie vor allem die Aufgabe, die zur Herausbildung des neuartigen wahrscheinlichkeitstheoretischen Denkens notwendigen Veränderungen und Voraussetzungen auf der wissenschaftlichen und philosophisch-erkenntnistheoretischen Ebene zu analysieren, wie sie sich im Wandel von der im Mittelalter vorherrschenden antiken, kontemplativen Wissenschaftsauffassung zu einer aktiven operativen Auffassung der neuzeitlichen Wissenschaften zeigen.

Diese sehr knappen Anmerkungen weisen auf die Vielfalt und Unterschiedlichkeit der zur Analyse der Entstehung des Wahrscheinlichkeitsbegriffs notwendigen Aspekte hin; als wichtig für jede Untersuchungsebene, für jede Herangehensweise an dieses Problem erweist sich jedoch ganz allgemein die Beachtung des Verhältnisses der Theorie zur Praxis, der wissenschaftlichen Erkenntnisse zu den konkreten praktischen Anforderungen.

Diese Theorie-Praxis-Spannung bringt in ihrer aspektreichen Form auch den Wahrscheinlichkeitsbegriff hervor. Bezogen auf die Analyse der in diesem Entstehungszusammenhang wichtigen mathematischen Methoden und Konzepte samt ihrer Struktur, auf deren Untersuchung wir uns beschränken wollen, läßt sich diese Spannung zugespitzt an dem Wechselverhältnis von Glücksspiel und relativer Häufigkeit darstellen.

II.1.1. Glücksspiel und relative Häufigkeit

Welch' überaus große Bedeutung das Glücksspiel für die erste Phase der Wahrscheinlichkeitstheorie hat, wird allein schon aus der Tatsache sichtbar, daß alle frühen Arbeiten zu diesem Thema fast ausschließlich Glücksspielprobleme behandelten. Die bekanntesten sind die im Briefwechsel zwischen Pascal und Fermat (1654) diskutierten Probleme des Chevalier de Mérê: Bei wieviel Würfen ist es sinnvoll, auf einen Sechserpasch zu wetten? (Problème des dés") und, Wie ist bei der Unterbrechung eines Glücksspiels der Gewinn aufzuteilen? ("Problème des points"). Mit der Lösung dieser beiden Aufgaben wird ganz allgemein der Beginn der mathematischen Wahrscheinlichkeitstheorie angesetzt. (Der Briefwechsel zwischen Pascal und Fermat ist vollständig abgedruckt in F.N. David, 1962; vgl. weiter hierzu Todhunter, 1865, Ore, 1960,und Ian Hacking, 1975, Kapitel 7.)

In dem ersten Büchlein zur Wahrscheinlichkeitstheorie von Christian Huygens:"Van Rekeningh in Spelen van Geluck", das 1657 als Anhang in van Schootens "Exercitionum Mathematicorum" unter dem Titel "Ratiociniis in aleae ludo" auf Lateinisch erschien, wird erstmals versucht, auf der Grundlage des Begriffs der Erwartung ansatzweise eine Theorie der Wahrscheinlichkeit im Rahmen der Glücksspiele zu entwickeln.

Mit Hilfe einfacher Sätze, die in dieser Theorie formuliert und bewiesen werden, behandelt Huygens dann einfache Glücksspielprobleme; am Ende seines Traktats finden sich zudem weitere fünf neue Probleme.

Als zusätzliche Belege für die Bedeutung des Glücksspiels seien noch Jakob Bernoullis "Ars Conjectandi" (1713), Rémond de Montmorts "Essai d'analyse sur les jeux de hasards" (1708/1713) und Abraham de Moivres' "The Doctrine of chances" (1718) angeführt. In allen genannten Büchern werden wahrscheinlichkeitstheoretische Probleme fast immer in die Form von Glücksspielen eingekleidet; allein in de Moivres Werk werden so beispielsweise rund achtzig Probleme und Aufgaben dargestellt und gelöst.

Doch aus dieser Tatsache zu schließen, die Wahrscheinlichkeitstheorie sei zu dieser Zeit ausschließlich eine Theorie der Glücksspiele gewesen, wäre völlig falsch; dies würde nicht den Aufschwung, den diese Theorie genommen hat, erklären können. Zu Recht merkt Hacking in Übereinstimmung mit Maistrov an, "that the emergence of probability had little to do with gambling." (Maistrov, 1974, S.viii)

Um die Bedeutung der Glücksspiele für die Wahrscheinlichkeitstheorie besser einschätzen zu können, muß berücksichtigt werden, daß man gleichzeitig mit der Entwicklung des Wahrscheinlichkeitsbegriffs vor vielfältigen praktischen Aufgaben und Problemen stand und auch diesen neuen Begriff in möglichst vielen Bereichen anzuwenden versuchte. So sind vor allem die Anwendungen in den Sterblichkeits- und Rentenberechnungen (vgl. etwa John Graunt: Natural and Political Observations made upon the Bills of Mortality, London, 1663 und De Witt: Waerdye van lyf-renten naer proportie van los-renten, S'Gravenhage, 1671) und im Bereich der Rechtsprechung und des Gesetzes (Leibniz, 1665) zu nennen. Ja, man kann ohne zu übertreiben sagen, daß die Absicht bestand, die neue Wahrscheinlichkeitstheorie in jedem Bereich des täglichen Lebens benutzen zu wollen, sie auf "bürgerliche, sittliche und wirtschaftliche Verhältnisse" (Jakob Bernoulli) anzuwenden. Die Wahrscheinlichkeit wurde als ein "Führer des Lebens", wie man später sagte, aufgefaßt, als eine universelle Methode zur Behandlung verschiedenartigster praktischer Probleme.

Beispielhaft wird an der Aussage Pierre Rêmond de Montmorts, mit welcher er die Einleitung zu seinem Werk beginnt, die große Hoffnung, die man an die neue Wissenschaft knüpfte, sehr schön deutlich: "Il y a long-temps que les Géomètres se vantent de pouvoir par leurs méthodes découvrir dans les Sciences naturelles toutes les verités qui sont à la portée de l'esprit humain; et il est certain que par le merveilleux alliage qu'ils ont fait depuis cinquante ans de la Géométrie avec la Physique, ils ont forcé les hommes à reconnoître que ce qu'ils disent à l'avantage de la Géomètrie n'est pas sans fondement. Quelle gloire seroit-ce pour cette Science si elle pouvoit encore servir à regler les jugements et la conduite des hommes dans la pratique des choses de la vie!

L'aîné de Messieurs Bernoulli si connus l'un et l'autre dans le monde sçavant, n'a pas cru qu'il fût impossible de porter la Géométrie jusqu'à ce point." (Montmort, 1708, S.1)

In welchem Zusammenhang steht nun einerseits das Glücksspiel und andererseits diese vermeintlich universelle Anwendbarkeit des Wahrscheinlichkeitsbegriffs?

Die Zuversicht, die Wahrscheinlichkeitstheorie so ausgiebig benutzen zu können, entstammte letztlich der im Laufe der Entwicklung der neuzeitlichen Wissenschaften gewonnenen, relativ neuen Erkenntnis, daß sich in vielen konkreten praktischen Situationen gewisse stabile und reguläre Verhältnisse von Häufigkeiten zeigen. Aus der relativen Häufigkeit der beobachtbaren Anzeichen konnten zwar keine absolut gewissen, jedoch "wahrscheinliche" Prognosen über zukünftige Erscheinungen getroffen werden. So diagnostizierte etwa der Arzt auf Grund bestimmter Krankheitszeichen beim Patienten mit einer gewissen Wahrscheinlichkeit vorliegende Krankheitsursachen, und der Bergmann konnte beispielsweise in Anbetracht der Beschaffenheit des Gesteins anhand relevanter Anzeichen, etwa auf Metalle im Berg, mit mehr oder weniger großer Wahrscheinlichkeit schließen.

Die Rolle des "Zeichenbegriffs" der mittelalterlichen Wissenschaften und die Bedeutung seiner Wandlung für die Entstehung des neuen Wahrscheinlichkeitsbegriffs untersucht Hacking ausführlich in seinem Buch (Hacking, 1975, Kap.5).

Indem das "Zeichen" nicht mehr einfach mit seinem Gegenstand identifiziert wurde, sondern diesem gegenüber eine gewisse Variabilität erlangte und so die Universalität der Bedeutung des Zeichens aufgehoben wurde, konnten nun mit Hilfe des Zeichens, das nun auch als ein An-Zeichen für gewisse Ursachen angesehen wurde, wahrscheinliche Urteile und Prognosen vorgenommen werden. Die in der jeweiligen praktischen Situation gewonnenen Erfahrungen darüber, wie oft das betrachtete Zeichen zutrifft oder nicht, erlaubte eine erste Quantifizierung des Wahrscheinlichkeitsbegriffs. Diese Veränderung des Zeichenbegriffs beurteilt Hacking als eine der wesentlichen Aspekte bei der Entstehung der mathematischen Wahrscheinlichkeit aus dem alten Begriff der "Probabilitas":"Probability was communicated by what we should now call law-like regularities and frequencies. Thus the connection of probability namely testimony with law-like frequencies is a result of the way in which the new concept of internal evidence came into being." (Hacking, 1975, S.44). Und weiter unten belegt Hacking anhand des Zitats eines Textes von Hobbes die Bedeutsamkeit des Zeichenbegriffs und damit des intuitiven Begriffs der relativen Häufigkeit für den Begriff der Wahrscheinlichkeit. Hobbes schreibt 1640: "This taking of signs by experience, is that wherein men do ordinarily think, the difference stands between man and man in wisdom, by which they commonly understand a man's whole ability or power cognitive; but this is an error; for the signs are but conjectural; and according as they have often or seldom failed, so their assurance is more or less; but never full and evident: for though a man have always seen the day and night to follow one another hitherto, yet can he not thence conclude they shall do so, or that they have done so eternally: experience concludeth nothing universally. If the signs hit twenty times for one missing, a man lay a wager of twenty to one of the event; but may not conclude it for a truth." (Hobbes, 1650, IV.10, S.17/18)

Hierzu merkt Hacking an: "Here, in a text published in 1650, probability has emerged in all but name." (Hacking, 1975, S.48)

In den letzten Äußerungen Hobbes' wird nun sehr schön der Zusammenhang, den man zwischen der relativen Häufigkeit und dem idealen Glücksspiel herstellte, deutlich: Das als ideal angenommene stabile Verhältnis der Häufigkeiten wird mit einer Wette, einem Spiel mit denselben Chancenverhältnissen verglichen. Das Glücksspiel erklärt sich also nicht einfach selbst, es dient der Erläuterung anderer Situationen.

Das Glücksspiel gibt die Möglichkeit, die beobachteten empirischen Häufigkeiten zu modellieren, sie zu erklären und ihnen eine Struktur aufzuprägen; und auch umgekehrt kann die Kenntnis einer möglichst genau beobachteten relativen Häufigkeit dazu beitragen, die bezüglich des Glücksspiels formulierten theoretischen Verfahrensweisen zu verbessern. So läßt sich beispielsweise feststellen, daß ein Teil der vielen Glücksspielprobleme geradezu aus Widersprüchlichkeiten zwischen empirischen Erfahrungen und theoretischen Berechnungen entstand, oder solche selbst beinhaltete. Hacking interpretiert etwa ein Problem, welches von Galilei bearbeitet wurde "as a conflict between experimental result and a particular theory." (Hacking, 1975, S.53) Hierbei ging es darum, ob man mit drei Würfeln eine Summe von 9 bzw. 12 Augen genauso oft werfen kann, wie eine Summe von 10 bzw. 11 Augen. Legt man, modern gesagt, eine Bose-Einstein-Statistik für Würfel zugrunde, was fälschlicherweise oft gemacht wurde,und nicht die Maxwell-Boltzmann-Statistik, dann ist mit drei Würfeln tatsächlich 9 bzw. 12 genauso wahrscheinlich wie 10 bzw. 11. Galilei war jedoch vorsichtiger; er befand "it is known from long observation that dice players consider 10 and 11 more advantageous than 9 and 12." (Zitiert nach Hacking, 1975, S.52) Und dieser offensichtliche Widerspruch führte ihn letztlich zur richtigen Lösung mittels der angemessenen Statistik: Man hat nur 25 Möglichkeiten, die 9 bzw. 12, jedoch 27 Möglichkeiten, die 10 bzw. 11 zu werfen.

Dieses Beispiel verdeutlicht den Wechselbezug zwischen theoretischen Berechnungen und empirischen Beobachtungen und zeigt, wie anhand einer recht präzisen Kenntnis der relativen Häufigkeit die "Theorie des Glücksspiels" entwickelt wurde.

Dem idealen Glücksspiel kommt in gewisser Weise die Funktion eines natürlichen Erklärungsrahmens für den Wahrscheinlichkeitsbegriff zu, wie er sich in den stabilen Häufigkeitsverhältnissen zeigte; die ideale Symmetrie des Würfels oder des Kartenspiels macht das zugrundeliegende Verhältnis der Häufigkeiten und die Art und Weise ihres Zustandekommens deutlich. Zudem erlaubt dieser vorgegebene Rahmen des idealen Glücksspiels in einfacher Form die wichtigsten Grundregeln der Wahrscheinlichkeitsrechnung, wie etwa das Additions- und Multiplikationstheorem abzuleiten und auch weitere Grundbegriffe, wie den der bedingten Wahrscheinlichkeit und Unabhängigkeit ansatzweise in exemplarischen Situationen einzuführen.

Kurz, das ideale Glücksspiel gestattet, die beobachtbaren stabilen Häufigkeitsverhältnisse, also die Wahrscheinlichkeiten in einem Modell zu fassen, in einer ersten und einfachen Weise. Zudem kann unter dieser Perspektive des Spiels eine ungeheure Vielzahl von Problemen, wie wir angemerkt haben, "mathematisch" formuliert und bearbeitet werden, und so die Entwicklung der Wahrscheinlichkeitsrechnung vorantreiben. Hier ist für die frühe Phase der Wahrscheinlichkeitstheorie vor allem die systematische Einbeziehung und auch Erarbeitung kombinatorischer Methoden, wie beispielsweise dem heute sogenannten "Pascalschen Dreieck" anzuführen.

Die Wahrscheinlichkeitsrechnung dieser Zeit kann wohl zu Recht als eine Theorie der Chancenverhältnisse interpretiert werden, die so zu einem Teilgebiet der damaligen Algebra wurde (vgl. Ivo Schneider, 1972, Kap. 1). Dies bringt deutlich zum Ausdruck, daß die beobachtbaren Häufigkeitsverhältnisse (in unzulässiger Weise, wie sich später zeigen wird) zunächst ideal als absolut bestimmte Verhältnisse, ja Zahlen aufgefaßt wurden; die ideale Symmetrie als Ursache dieser Verhältnisse determinierte die Wahrscheinlichkeit, der Zufall spielte demgegenüber bisher keine grundlegende Rolle. Jedoch darf nicht die so vielfältig und breit intendierte Anwendbarkeit dieser Wahrscheinlichkeitsrechnung bloß als Beiwerk abgetan werden; sie verhinderte, daß diese neue Rechnung ausschließlich eine Glücksspielrechnung blieb.

Gerade die Autoren, die auf der Grundlage der Glücksspiele die Wahrscheinlichkeit darstellen, betonen die weitreichende Anwendbarkeit und die damit verbundene Bedeutung dieser Rechnung über das Glücksspiel hinaus. So findet sich etwa in dem berühmten Werk "Die Logik oder Die Kunst des Denkens" (Arnauld, 1972, erstmals 1662 veröffentlicht) im Kapitel XVI des letzten Teils zunächst eine Darstellung von Wahrscheinlichkeit mit Hilfe des Glücksspiels:"Es gibt Spiele, in denen von zehn Personen jeder einen Taler einsetzt, und es nur einen gibt, der das Ganze gewinnt, während alle anderen verlieren. In diesem Fall riskiert jeder Spieler nur einen Taler und er gewinnt möglicherweise neun. Wenn man nur den Gewinn und Verlust für sich genommen in Erwägung ziehen würde, würde es den Anschein haben, daß alle Spieler in diesem Fall einen Vorteil haben. Man muß aber außerdem erwägen, daß es, wenn jeder neun Taler gewinnen kann und dabei nur einen riskiert, in bezug auf jeden einzelnen neunmal wahrscheinlicher ist, daß er seinen Taler verlieren und die neun Taler nicht gewinnen wird. Auf diese Weise hat jeder für sich neun Taler zu erhoffen und einen zu verlieren, neun Grade der Wahrscheinlichkeit, einen Taler zu verlieren und einen einzigen, die neun Taler zu gewinnen: was die Sache vollkommen ausgleicht." (Arnauld, 1972, S.346) Nach weiteren Anmerkungen und Erklärungen zum Glücksspiel fährt der Autor dann jedoch fort:"Diese Überlegungen scheinen geringfügig zu sein, und sie sind es in der Tat, wenn man dabei stehen bleibt. Man kann sie aber in den Dienst viel wichtigerer Dinge stellen. Der hauptsächliche Nutzen, den man daraus ziehen muß, besteht darin, daß wir in unseren Hoffnungen und in unseren Befürchtungen vernünftiger werden. Es gibt zum Beispiel viele Leute, die höchst erschreckt sind, wenn sie Donner hören. ... Ist es ... die Angst, durch den Donner zu sterben, die ihre außergewöhnliche Furcht verursacht, so ist es leicht zu zeigen, daß sie nicht vernünftig ist, denn es ist viel, wenn einer unter zwei Millionen Menschen auf diese Weise stirbt; und man kann selbst sagen, daß es keinen gewaltsamen Tod gibt, der weniger häufig vorkommt als dieser. Da also die Furcht vor dem Übel nicht nur der Größe des Übels angemessen sein soll, sondern auch der Wahrscheinlichkeit des Ereignisses, und da es kaum eine seltenere Todesart gibt, als durch den Donner zu sterben, gibt es auch kaum eine Todesart, die uns weniger Furcht verursachen sollte, vor allem in Anbetracht dessen, daß die Furcht keineswegs dazu beiträgt, daß man dem Sterben durch den Donner entgeht." (Arnauld, 1972, S.347/48)

Diese ausführlichen Passagen machen den wechselseitigen Zusammenhang zwischen Glücksspiel und relativer Häufigkeit sehr schön anschaulich deutlich; zudem wird beispielhaft sichtbar, was unter praktischer Benutzbarkeit des Wahrscheinlichkeitsbegriffs verstanden wurde.

Auch Christian Huygens, der als erster Ansätze einer <u>Spieltheorie</u> ausarbeitete, verweist in seinem Büchlein auf die eigentliche Bedeutung dieser Rechnungen: "Toutefois je veux croire qu'en considérant ces choses plus attentivement, le lecteur apercevra bientôt qu'il ne s'agit pas ici d'un simple jeu d'esprit mais qu'on y jette les fondements d'une spéculation fort intéressante et profonde." (Huygens, Werke Bd. 14, 1920, S.56).

Die wechselseitige Spannung zwischen der Herausbildung des Wahrscheinlichkeitsbegriffs und den praktischen Anforderungen für die Entwicklung dieser neuen Theorie zeigt sich exemplarisch an dem Verhältnis zwischen Glücksspiel und relativer Häufigkeit, als der mathematischen Modellierung und Simulation relevanter praktischer Beobachtungsgrößen. Dieser anfängliche Zusammenhang zwischen Modell und Realität war eine einfache "Übersetzung", eine Isomorphie. Um der allgemein beabsichtigten und geforderten breiten Anwendbarkeit auch in komplizierten Fällen gerecht zu werden, bedurfte es einer Weiterentwicklung und Präzisierung dieser Beziehung zwischen der mathematischen Modellierung und den Anwendungsgegenständen. Schon für die allererste Phase war das Theorie-Praxis-Verhältnis jedoch zentral, und es gelang, eine spezifisch wahrscheinlichkeitstheoretische Ausprägung dieser Beziehung in einer ersten einfachen Form durch "Glücksspiel" und "Häufigkeitsverhältnis" herauszuarbeiten; die Differenzierung dieses wahrscheinlichkeitstheoretischen Anwendungsbezuges, welche wir im weiteren analysieren wollen, machte dann immer offensichtlicher, was Huygens prophezeit hat, nämlich daß hier "der Grundstein einer sehr interessanten und tiefgreifenden Theorie gelegt wird."

Eine erste und sehr wichtige Etappe in der weiteren Bearbeitung und Verbesserung dieses wahrscheinlichkeitstheoretischen Anwendungsbezuges stellt das berühmte Theorem von Jakob Bernoulli dar, das wir im nächsten Abschnitt ausführlich untersuchen wollen.

II.1.2. Bernoullis Theorem

Jakob Bernoullis Werk über die Wahrscheinlichkeitsrechnung "Ars conjectandi" (oder: "Die Kunst des Mutmaßens", erstmals 1713 veröffentlicht; wir benutzen im folgenden die 1899 erschienene deutsche Übersetzung in Ostwalds Klassiker, Nr. 107/108) stellt einen entscheidenden konzeptionellen Fortschritt in der frühen Phase der Entwicklung der Wahrscheinlichkeitstheorie dar, so die einhellige Meinug vieler Mathematiker und Historiker. (vgl. etwa Hacking, 1975, Todhunter, 1965 und Kolmogoroff, 1947). Das Neue und überaus Wichtige war natürlich das heute nach ihm benannte "Bernoullische Theorem", welches im vierten und letzten Teil des Werkes bewiesen wird. Es ist das erste, wichtige Theorem der Wahrscheinlichkeitstheorie überhaupt. Michel Loève hebt besonders die Bedeutung dieses Theorems für die Entwicklung der Wahrscheinlichkeitstheorie heraus:"C'est seulement, avec la publication posthume de l'ouvrage de Jacques Bernoulli, et plus exactement avec ce qu'il appelle son 'théorème d'or' (la loi des grands nombres) que le Calcul des probabilités a commencé sa lente évolution vers une discipline mathématique pure." (Loève, 1978, S.282) Dem Anwendungsbezug der in Bernoullis Theorem in spezieller Weise zum Ausdruck kommt, wird somit eine wichtige und zentrale Rolle für die Entstehungsgeschichte der gesamten Wahrscheinlichkeitstheorie zuerkannt.

II.1.2.1 Zur Bedeutung von Bernoullis "Ars conjectandi"

Wie schon der Titel des Bernoullischen Werkes "Kunst des Mutmaßens" ankündigt, ging es dem Autor hiermit um weit mehr, als sich in erster Linie, wie es die meisten Autoren vor ihm taten, auf die Glücksspiele zu beschränken; Jakob Bernoulli strebte an, explizit Anwendungen der Wahrscheinlichkeitstheorie in allen Bereichen des täglichen Lebens vorzunehmen. "Das Würfelspiel ist ein Spiel. Man kann es auch lassen, wenn man nicht vom Spielteufel besessen ist. Aber Mutmaßen, das ist eine unentbehrliche Tätigkeit. Jeder Richterspruch, jede wissenschaftliche Theorie, jede militärische Strategie und jede wirtschaftliche Unternehmung beruht auf Vermutungen. Bei jeder Tat gibt man sich Rechenschaft von den wahrscheinlichen Folgen, oder zumindest sollte man sich darüber Rechenschaft geben, und die Wahrscheinlichkeitsrechnung kann einem helfen, die Wahrscheinlichkeiten der angenehmen und unangenehmen Folgen abzuschätzen. Heute ist uns das alles geläufig. Versicherungsgesellschaften schätzen Sterbewahrscheinlichkeiten auf Grund von Statistiken und richten ihre Tarife danach ein. Es gibt eine mathematische Statistik, die uns lehrt, die Irrtumswahrscheinlichkeiten von Erfahrungsschlüssen abzuschätzen. Es gibt eine Entscheidungstheorie und eine Unternehmungsforschung, die immerfort mit Wahrscheinlichkeiten rechnen. Aber vor Jakob Bernoulli gab es das alles nicht. Er war es, der die Menschheit zuerst auf die theoretische und praktische Wichtigkeit der'Mutmaßungskunst' aufmerksam gemacht hat." (Bernoulli, Werke Bd. 3, 1975, S.12) Als so breit angelegt, beschreibt B.L. van der Waerden, in der historischen Einleitung zu Bernoullis wahrscheinlichkeitstheoretischen Arbeiten, das intendierte Vorhaben.

Mehr als 20 Jahre hat Bernoulli sich mit der Wahrscheinlichkeitstheorie beschäftigt. Vor allem aus seinen "Meditationes" (abgedruckt in seinen Werken Bd. 3, 1975), einem chronologisch geführten Tagebuch mit Notizen und kurzen Abhandlungen für spätere Veröffentlichungen, kann man recht genau ersehen, wann er sich mit welchen Problemen und Fragestellungen bezüglich der Wahrscheinlichkeitsrechnung befaßt, und wie er sich in dieses Gebiet eingearbeitet hat. Am Anfang stehen gerade die fünf Glücksspielprobleme, die Huygens ohne Lösung am Ende seines Traktats publiziert hat. Zwischen 1684 und 1685 löst Bernoulli diese und fügt eigene Probleme an. In einer zweiten Phase von 1685 bis 1686 wendet er dann die neue Kunst auf Probleme des Alltags, so z.B. Erbangelegenheiten an und stößt schließlich während der dritten Arbeitsphase 1687 bis 1689 auf das Gesetz der großen Zahlen und gibt auch einen Beweis hierfür.

Dies zeigt, wie früh Bernoulli im Besitz seines Theorems einschließlich des gültigen Beweises war; das in den sich anschließenden Jahren ausgearbeitete Werk, die "Ars conjectandi" konnte Bernoulli, vor allem was den letzten Teil betraf, nicht recht befriedigen, so daß er sich wegen der Unvollständigkeiten nie zu einer Veröffentlichung durchringen konnte. Erst 8 Jahre nach seinem Tode (1705) veröffentlichte der Neffe Nikolaus Bernoulli die "Ars conjectandi" 1713 in Basel.

Der Aufbau der "Kunst des Mutmaßens" ist teilweise eng an den Vorüberlegungen in den "Meditationes" ausgerichtet. So behandelt der erste der vier Teile dieses Werkes wiederum unter dem Titel "Abhandlungen über die bei Glücksspielen möglichen Berechnungen" ausführlich die Arbeit von Christian Huygens samt den Lösungen von Glücksspielproblemen. Der zweite Teil "Permutations- und Kombinationslehre" stellt, wie sein Titel anzeigt, einen systematischen Aufbau der Kombinatorik dar, er behandelt u.a. Eigenschaften der Binomialkoeffizienten und gibt erstmals die für die allgemeine Berechnung von Potenzsummen notwendigen "Bernoullischen Zahlen". Im dritten Teil "Anwendungen der Kombinationslehre auf verschiedene Glücks- und Würfelspiele" wird die erarbeitete Theorie zur Berechnung von Erwartungswerten für Spielgewinne benutzt. Der vierte und weitaus wichtigste Teil, die "Anwendung der vorhergehenden Lehre auf bürgerliche, sittliche und wirtschaftliche Verhältnisse", mit dem wir uns im folgenden ausführlich befassen wollen, beinhaltet gerade Bernoullis Theorem samt Beweis.

Die hier in gedrängter Form gegebene Übersicht zu Bernoullis "Ars conjectandi" läßt jedoch schon den Charakter dieser Arbeit deutlich werden: Zum einen stellt sie eine systematische Zusammenfassung der bis dahin in der Wahrscheinlichkeitstheorie erzielten Resultate dar und zum anderen führt der vierte Teil mit dem Grenztheorem einen neuen Gesichtspunkt in die theoretische Diskussion ein, ja mit ihm beginnt die theoretische Entwicklung des Wahrscheinlichkeitsbegriffs. Ian Hacking faßt die Bedeutung der "Ars conjectandi" folgendermaßen zusammen und weist zudem auf einen weiteren Punkt hin:"Once a research programme is under way the occasional masterpiece of permanent value often has three distinct characteristics. It does something almost completely new which, although much in the air at the time, has never before crystallized, but, once written down, sets the direction for all future enquiry. Secondly, it epitomizes what everyone has known for a long time but has been unable to state succinctly. Thirdly, and much less often noticed, it ends certain possible lines of development which, until that node in history, were perfectly open but now become closed. The first two features of Bernoulli's work are evident. The third feature is this: until 1713 it was not in the least determined that the addition law for probability would be accepted. Bernoulli was the last master to contemplate non-additive probabilities." (Hacking, 1975, S.144) Wir wollen uns jedoch gerade auf den ersten Punkt, auf das "Neue", wie es im vierten Teil in Form des Grenztheorems vorgestellt wird, konzentrieren. "It is Part IV that revolutionizes probability theory. The revolution is twofold. For the first time a 'subjective' conception of probability is explicitly avowed, and the first limit theorem is proven." (Hacking, 1975, S.145) Damit bezieht Hacking sich neben Bernoullis Theorem auf die umstrittene erkenntnsitheoretische Interpretation des Wahrscheinlichkeitsbegriffs: Handelt es sich hierbei um einen objektiven, "in der Wirklichkeit vorkommenden" oder um einen subjektiven, vom Menschen willkürlich konstruierten Begriff? Gleichzeitig mit seiner Entstehung, so Hacking, hatte der Wahrscheinlichkeitsbegriff als Nachfolger des alten, eher subjektiv zu kennzeichnenden und für alltägliche Entscheidungen gebrauchten Begriffs der "probabilitas" diesen dualen Charakter, diesen Januskopf aleatorischer und epistemologischer Bedeutung. "... the probability emerging in the time of Pascal is essentially dual. It has to do both with stable frequencies and with degrees of belief. It is, as I shall put it, both aleatory and epistemological. This quite specific character of probability is one of the clues to its emergence." (Hacking, 1975, S.10) Der erste, welcher nun explizit von "subjektiver" und "objektiver" Gewißheit - bezogen auf die Wahrscheinlichkeitstheorie - spricht, ist Jakob Bernoulli. Und auch nur, soweit es für das Verständnis seiner Wahrscheinlichkeitskonzeption und der Interpretation seines Theorems wichtig ist, wollen wir uns in diese problematische Diskussion einlassen. Auch in den folgenden Teilen werden wir, entsprechend unserem Vorhaben, den sich entwickelnden Anwendungsbezug der Wahrscheinlichkeitstheorie zu analysieren, nur unter dieser Perspektive das Problem von "subjektiver und objektiver Wahrscheinlichkeit" diskutieren. Wir erhoffen jedoch, daß gerade das Anwendungsproblem der auf der allgemeinen philosophischen bzw. erkenntnistheoretischen Ebene so strikt alternativ geführten Diskussion zusätzliche erklärende Aspekte gibt, und es ermöglicht, auf dem Hintergrund des Theorie-Praxis-Bezugs das Problem der subjektiven versus objektiven Wahrscheinlichkeit als Teil der Subjekt-Objekt-Problematik der modernen Naturwissenschaften insgesamt besser verstehen und einschätzen zu können.

Kommen wir nun jedoch zur Analyse des Bernoullischen Theorems. In welchem konzeptionellen Rahmen hat Jakob Bernoulli seinen Wahrscheinlichkeitsbegriff formuliert, um anschließend sein Theorem vorstellen und beweisen zu können?

II.1.2.2 Bernoullis Überlegungen zum Wahrscheinlichkeitsbegriff

Zu Beginn des vierten Teils der "Ars conjectandi", im ersten Kapitel:"Einleitende Bemerkungen über Gewißheit, Wahrscheinlichkeit, Notwendigkeit und Zufälligkeit der Dinge" erklärt Bernoulli den Unterschied von objektiver und subjektiver Gewißheit:"Die Gewißheit irgendeines Dinges läßt sich entweder objektiv, d.h. an sich betrachten, und bezeichnet in diesem Falle nichts anderes als das wirkliche gegenwärtige oder zukünftige Vorhandensein jenes Dinges, oder subjektiv, d.h. in bezug auf uns und besteht dann in dem Maße unserer

Erkenntnis hinsichtlich dieser Wirklichkeit." (Bernoulli, 1899, S.71) Jedoch "aufgrund göttlicher Voraussicht und Vorherbestimmung" (S.72) so Bernoulli, hat "alles, was unter der Sonne existiert oder entsteht, das Vergangene, das Gegenwärtige und das Zukünftige ... an sich die höchste Gewißheit." (S.71)

Bezogen auf uns Menschen "(variiert jedoch), so Bernoulli, die betrachtete Gewißheit der Dinge ... vielfach nach oben und unten." Und er fährt fort:"Jene Dinge, von welchen es uns durch Offenbarung, Überlegung, sinnliche Wahrnehmung, Erfahrung, Autopsie oder irgendwie anderes gewiß ist, daß wir an ihrer gegenwärtigen oder zukünftigen Existenz nicht Zweifel haben dürfen, besitzen für uns die höchste und absolute Gewißheit. Alle übrigen Dinge erhalten ein, gemäß unserer Erkenntnis, unvollkommeneres Maß der Gewißheit, welches größer oder kleiner ist, je nachdem mehr oder weniger Wahrscheinlichkeiten dafür vorhanden sind, daß irgendein Ding ist, sein wird oder gewesen ist." (S.72)

Damit kann Bernoulli nun seinen Wahrscheinlichkeitsbegriff folgendermaßen definieren:"Die Wahrscheinlichkeit ist nämlich ein Grad der Gewißheit und unterscheidet sich von ihr wie ein Teil vom Ganzen. Wenn z.B. die volle und absolute Gewißheit, welche wir mit a oder 1 bezeichnen, aus fünf Wahrscheinlichkeiten oder Teilen bestehend angenommen wird, von denen drei für das gegenwärtige oder zukünftige Eintreten irgendeines Ereignisses, und die übrigen beiden dagegen sprechen, so soll das Ereignis $\frac{3}{5} \cdot a$ oder $\frac{3}{5}$ der Gewißheit besitzen." (S.72)

Auf der Grundlage seiner allgemeinen Vorstellung von Wahrscheinlichkeit als einem Teil-Ganze-Verhältnis, geht Bernoulli letztlich von einem kombinatorischen Gleichwahrscheinlichkeitsbegriff aus, der sich <u>a priori</u> aus den idealen Eigenschaften (beispielsweise der Symmetrie) der betrachteten Ereignisse bestimmen läßt. Damit wird der Zusammenhang seines Konzeptes zum Wahrscheinlichkeitsbegriff der Glücksspiele sichtbar. Weiterhin machen die zitierten Aussagen Bernoullis seinen "deistischen Determinismus" deutlich: Die Wahrscheinlichkeit ist als ein Grad der Gewißheit geradezu Ausdruck der Unsicherheit und Ungewißheit des unvollkommenen und unwissenden Menschen über die "auf Grund göttlicher Voraussicht und Vorherbestimmung" unterliegenden prinzipiell absolut gewissen Gesetzmäßigkeiten.

An einigen Beispielen verdeutlicht Bernoulli seine Auffassung, indem er im Zusammenhang mit der "Zufälligkeit" die prinzipielle Determiniertheit jeglichen Naturgeschehens unterstreicht. "Ganz gewiß ist es, daß ein Würfel bei gegebener Lage, Geschwindigkeit und Entfernung vom Würfelbrette, von dem Augenblick an, in welchem er die Hand verläßt, nicht anders fallen kann, als er tatsächlich auch fällt. Ebenso kann das Wetter bei einer bestimmten gegenwärtigen Beschaffenheit der Atmosphäre, bei bestimmter Menge, Lagerung, Bewegung, Richtung, Geschwindigkeit der Winde, Dünste und Wolken und bestimmten mechanischen Gesetzen, nach welchen sich diese sämtlich untereinander bewegen, morgen nicht anders sein, als es wirklich sein wird. Diese Wirkungen folgen aus ihren nächsten Ursachen nicht weniger notwendig, als die Erscheinungen der Finsternisse aus der Bewegung der Gestirne." (S.73/74) Welche Rolle spielt dann jedoch die Zufälligkeit? "Und dennoch hält man an der Gewohnheit fest, nur die Finsternisse zu den notwendigen Ereignissen, die Fälle des Würfels und die zukünftige Gestaltung des Wetters aber zu den zufälligen Ereignissen zu rechnen. Der Grund hiervon liegt ausschließlich darin, daß das, was zur Bestimmung späterer Geschehnisse als gegeben angenommen wird und in Wirklichkeit auch gegeben ist, uns noch nicht hinreichend bekannt ist; wäre es uns aber hinreichend bekannt, so ist das Studium der Mathematik und Physik genügend weit ausgebildet, damit wir aus gegebenen Ursachen die späteren Wirkungen ebenso berechnen könnten, wie wir z.B. aus den bekannten astronomischen Gesetzen die Finsternisse berechnen und voraussagen können. Bevor aber die Astronomie zu solcher Vollkommenheit gelangt war, mußte man die Finsternisse deshalb genau so wie die beiden anderen oben angeführten Ereignisse zu den künftig zufällig eintretenden zählen. Daraus folgt, daß einem Menschen und zu einer bestimmten Zeit etwas als zufällig erscheinen kann, was einem anderen Menschen (ja sogar auch demselben) zu einer anderen Zeit, nachdem die Ursachen davon erkannt sind, als notwendig erscheint. Daher hängt die Zufälligkeit vornehmlich auch von unserer Erkenntnis ab, insofern als wir keinen Grund wahrnehmen können, welcher dagegenspricht, daß etwas nicht ist oder nicht sein wird, trotzdem es auf Grund der nächsten, uns aber noch unbekannten Ursachen notwendig ist oder sein wird." (Bernoulli, 1899, S.74)

Der Wahrscheinlichkeit kommt nach Bernoullis Ansicht somit nur eine vorläufige Hilfsfunktion bei der Erkenntnis zu; sind "Mathematik und Physik genügend weit ausgebildet", dann kann im Grunde auf die Wahrscheinlichkeit verzichtet werden.

Was macht es nun für Bernoulli erforderlich, diesen doch relativ geschlossenen und einheitlichen Begriffskontext aufzubrechen und zu erweitern?

Der wichtigste Grund hierfür war gerade das zentrale Anliegen Bernoullis, ja sein Hauptinteresse an der Wahrscheinlichkeitstheorie überhaupt, nämlich die Erweiterung des Anwendungsbereichs dieser neuen Rechnung, die Anwendung der Glücksspiellehre auf "bürgerliche, sittliche und wirtschaftliche Verhältnisse".

Um dies zu erreichen, entwickelt Bernoulli zunächst das "Prinzip der moralischen Gewißheit", welches wichtig für den alltäglichen Umgang mit Wahrscheinlichkeiten ist und dem Menschen in gewissen Entscheidungssituationen nützlich sein soll. Dazu erklärt Bernoulli zuerst, was er unter der "moralischen Gewißheit" versteht:" <u>Moralisch gewiß</u> ist etwas, dessen Wahrscheinlichkeit nahezu der vollen Gewißheit gleichkommt, so daß ein Unterschied nicht wahrgenommen werden kann. <u>Moralisch unmöglich</u> dagegen ist das, was nur so viel Wahrscheinlichkeit besitzt, als dem moralisch Gewissen an der vollen

Gewißheit mangelt. Wenn man also das, was $\frac{999}{1000}$ der Gewißheit für sich hat als moralisch gewiß betrachtet, so ist das, was nur $\frac{1}{1000}$ der Gewißheit für sich hat, moralisch unmöglich." (S.73)

Im zweiten Kaptel (des vierten Teils) erläutert Bernoulli dann die Aufgabe der Vermutungskunde:"Irgendein Ding vermuten, heißt so viel, als seine Wahrscheinlichkeit messen. Deshalb bezeichnen wir als Vermutungs- oder Mutmaßungskunst (ars conjectandi sive stochastice) die Kunst, so genau als möglich die Wahrscheinlichkeiten der Dinge zu messen, und zwar zu dem Zwecke, daß wir bei unseren Urteilen und Handlungen stets das auswählen und befolgen können, was uns besser, trefflicher, sicherer oder ratsamer erscheint. Darin allein beruht die ganze Weisheit des Philosophen und die ganze Klugheit des Staatsmannes." (S.75)

Nun ist Bernoulli in der Lage, das Prinzip der moralischen Gewißheit in einer Reihe von Regeln zu formulieren, "welche die bloße Vernunft jedem Menschen mit gesundem Verstande diktiert." Dieses Prinzip sagt aus, wie man sich bei dem in praktischen Situationen notwendig ungenauen Vermuten, also unpräzisen Messen von Wahrscheinlichkeiten verhalten sollte. "9. Weil aber doch nur selten volle Gewißheit erlangt werden kann, so wollen es die Notwendigkeit und das Herankommen, daß das, was nur moralisch gewiß ist, für unbedingt gewiß gehalten wird. Es würde also nützlich sein, wenn auf Veranlassung der Obrigkeit bestimmte Grenzen für die moralische Gewißheit festgesetzt würden, wenn z.B. entschieden würde, ob zur Erzielung dieser $\frac{99}{100}$ oder $\frac{999}{1000}$ der Gewißheit verlangt werden müssen, damit ein Richter nicht parteiisch sein kann, sondern einen festen Gesichtspunkt hat, welchen er beim Fällen des Urteils beständig im Auge behält." (S.80)

Mit Hilfe dieser Vorüberlegungen kann Jakob Bernoulli nun den eigentlichen und wichtigen Teil seiner Untersuchung in Angriff nehmen. Im Kapitel IV:"Über die zwei Arten, die Anzahl der Fälle zu ermitteln. Was von der Art, sie durch

Beobachtung zu ermitteln, zu halten ist. Hauptproblem hierbei und anderes." hebt er hervor, daß es nur in den wenigsten Fällen möglich ist, die Wahrscheinlichkeit, a priori aus den idealen Eigenschaften des betrachteten Gegenstandes zu ermitteln, in den allermeisten Anwendungen der Wahrscheinlichkeitstheorie auf zufällige Ereignisse ist dagegen diese Berechnungsweise undurchführbar. Wir möchten hier eine etwas längere Passage zitieren, in welcher Bernoullis Überlegungen sehr schön zum Ausdruck kommen, und man erkennt, daß der wichtigste Impuls für die Verallgemeinerung und Entwicklung des Gleichwahrscheinlichkeitsbegriffs von den praktischen Anforderungen herrührt: Man möchte die Wahrscheinlichkeitstheorie auf eine sehr viel größere Klasse von Gegenständen anwenden können. Zunächst stellt Bernoulli fest: "Wir sind also dahin gelangt, daß zur richtigen Bildung von Vermutungen über irgendeine Sache nichts anderes zu tun erforderlich ist, als daß wir zuerst die Zahl dieser Fälle genau ermitteln und dann bestimmen, um wieviel die einen Fälle leichter als die anderen eintreten können." (S.88) Damit ist offensichtlich die kombinatorische Gleichwahrscheinlichkeit gemeint. Sofort kommt Bernoulli dann auf die Problematik dieser Wahrscheinlichkeitsdefinition zu sprechen: "Und hier scheint uns gerade die Schwierigkeit zu liegen, da nur für die wenigsten Erscheinungen und fast nirgends anders als in Glücksspielen dies möglich ist; die Glücksspiele wurden aber von den ursprünglichen Erfindern, damit die Spielteilnehmer gleiche Gewinnaussichten haben sollten, so eingerichtet, daß die Zahlen der Fälle, in welchen sich Gewinn oder Verlust ergeben muß, im voraus bestimmt und bekannt sind, und daß alle Fälle mit gleicher Leichtigkeit eintreten können. Bei den weitaus meisten anderen Erscheinungen aber, welche von dem Walten der Natur oder von der Willkür der Menschen abhängen, ist dies keineswegs der Fall. So sind z.B. bei Würfeln die Zahlen der Fälle bekannt, denn es gibt für jeden einzelnen Würfel ebenso viele Fälle als er Flächen hat; alle diese Fälle sind auch gleich leicht möglich, da wegen der gleichen Gestalt aller Flächen und wegen des gleichmäßig verteilten Gewichts des Würfels kein Grund dafür vorhanden ist, daß eine Würfelfläche leichter als eine andere fallen sollte, was der Fall sein würde, wenn die Würfelflächen verschiedene Gestalt besäßen und ein Teil des Würfels aus schwererem Material angefertigt wäre als der andere Teil. ... Welcher Sterbliche könnte aber je die Anzahl der Krankheiten (das ist ebenso vieler Fälle), welche den menschlichen Körper an allen seinen Teilen und in jedem Alter befallen und den Tod herbeiführen können, ermitteln und angeben, um wieviel leichter diese als jene Krankheit, die Pest als die Wassersucht, die Wassersucht als Fieber den Menschen zugrunde richtet, um daraus eine Vermutung über das Verhältnis von Leben und Sterben künftiger Geschlechter abzuleiten? ... Da diese und ähnliche Dinge von ganz verborgenen Ursachen abhängen, welche über dies noch durch die unendliche Mannigfaltigkeit ihres Zusammenwirkens unsere Erkenntnis beständig täuschen, so würde es völlig sinnlos sein, auf diese Weise etwas erforschen zu wollen." (S.88/89) Wie aber kann man sonst in diesen Fällen die Wahrscheinlichkeit bestimmen? Bernoulli fährt fort:"Aber ein anderer Weg steht uns hier offen, um das Gesuchte zu finden und das, was wir a priori nicht bestimmen können, wenigstens a posteriori, d.h. aus dem Erfolge, welcher bei ähnlichen Beispielen in zahlreichen Fällen beobachtet wurde, zu ermitteln." (S.89)

Man soll also mit Hilfe der beobachteten relativen Häufigkeit die Wahrscheinlichkeit berechnen. "Diese empirische Art, die Zahl der Fälle durch Beobachtungen zu bestimmen, ist weder neu noch ungewöhnlich; denn schon der berühmte Verfasser des Werkes 'L'Art de penser', ein scharfsinniger und talentvoller Mann, hat ... ein ganz ähnliches Verfahren beschrieben, und alle Menschen beobachten im täglichen Leben dasselbe Verfahren. Auch leuchtet jedem Menschen ein, daß es nicht genügt, nur eine oder die andere Beobachtung anzustellen, um auf diese Weise über irgendein Ereignis zu urteilen, sondern daß eine große Anzahl von Beobachtungen erforderlich sind. Zuweilen hat auch schon ein recht einfältiger Mensch in Folge irgendeines natürlichen Instinkts von sich aus und ohne jede vorangegangene Unterweisung die Erfahrung gemacht (was wirklich wunderbar ist), daß man, je mehr diesbezügliche Beobachtungen vorliegen, umso weniger Gefahr läuft, von der Wahrheit abzuirren." Man muß jedoch achtgeben; dieses, heute würde man sagen, empirische Gesetz der großen Zahlen bedarf nämlich einer genauen Analyse. Bernoulli merkt kritisch an:"Obgleich nun dies aus der Natur der Sache heraus von jedem eingesehen wird, so liegt doch der auf wissenschaftlichen Prinzipien gegründete Beweis durchaus nicht auf der Hand, und es liegt mir daher ob, ihn an dieser Stelle zu erbringen." (S.90) Der mathematisch exakte Beweis erfordert die Berücksichtigung <u>neuartiger</u> Überlegungen, welche Bernoullis spezifische Herangehensweise an die Bearbeitung der Konvergenz ausdrückt: Die Messung der Wahrscheinlichkeit des betrachteten Ereignisses mittels immer mehr beobachteten relativen Häufigkeiten wird von Bernoulli selbst als ein der Wahrscheinlichkeit unterliegendes Ereignis angesehen. Dementsprechend stellen sich für die Lösung des Problems erst unter dieser neuen Perspektive sinnvolle und theoretisch bearbeitbare Fragen; Bernoulli zählt diese auf:"Man muß vielmehr noch weiteres in Betracht ziehen, woran vielleicht niemand bisher auch nur gedacht hat. Es bleibt nämlich noch zu untersuchen, ob durch Vermehrung der Beobachtungen ständig auch die Wahrscheinlichkeit dafür wächst, daß die Zahl der günstigen zu der Zahl der ungünstigen Beobachtungen das wahre Verhältnis erreicht, und zwar in dem Maße, daß diese Wahrscheinlichkeit schließlich jeden beliebigen Grad der Gewißheit übertrifft, oder ob das Problem vielmehr, sozusagen, seine Asymptote hat, d.h. ob ein bestimmter Grad der Gewißheit, das wahre Verhältnis der Fälle gefunden zu haben, vorhanden ist, welcher auch bei beliebiger Vermehrung der Beobachtungen niemals überschritten werden kann, ... Ist dies nicht der Fall, so gestehe ich, daß es um unseren Versuch, die Zahl der Fälle durch Beobachtungen zu ermitteln, schlecht bestellt ist. Wenn es aber der Fall ist, und man schließlich auf diese Weise moralische Gewißheit erhält (daß dies wirklich so ist, will ich in dem folgenden Kapitel zeigen), so können wir die Zahlen der Fälle a posteriori fast ebenso genau finden, als wenn sie uns a priori bekannt wären. Und dies ist für das bürgerliche Leben, wo das moralisch Gewisse als absolut gewiß angesehen wird, nach Axiom 9 des Kapitels II hinreichend, um unsere Vermutung in jedem beliebigen Zufallsgebiet nicht weniger wissenschaftlich zu leiten als bei den Glücksspielen." (S.90ff)

Dies macht ansatzweise deutlich, wie Bernoulli im Gegensatz zur absolut bestimmten Wahrscheinlichkeit des idealen Glücksspiels den im täglichen Leben gebrauchten Begriff verstanden wissen will; nie ist dieser Begriff durch das Verhältnis der günstigen zu allen Beobachtungen absolut genau zu ermitteln, immer wird er nur bis auf "moralische Gewißheit" meßbar sein. Der, man könnte sagen, "statistische Charakter", den Bernoulli der Wahrscheinlichkeit beigibt, wird in seiner Darstellung deutlich, die das Häufigkeitsverhältnis im vorgegebenen Exaktheitsintervall einschließt:"Damit ... dies nicht unrichtig verstanden werde, ist noch zu bemerken, daß wir das Verhältnis zwischen den Zahlen der Fälle, welches durch wir Beobachtungen zu bestimmen unternehmen, nicht absolut genau (denn so würde ganz das Gegenteil herauskommen und desto unwahrscheinlicher werden, daß das richtige Verhältnis gefunden sei, je mehr Beobachtungen gemacht werden), sondern nur mit einer bestimmten Annäherung erhalten, d.h. zwischen zwei Grenzen einschließen wollen, welche aber beliebig nahe beieinander angenommen werden können." (S.92)

Äußerst wichtig ist für den nun folgenden Beweis des Theorems u.E. eine fast nebenbei gemachte Anmerkung Bernoullis, die verrät, auf welches mathematische Modell er sich bezieht, um vorgegebenen konkreten Situationen einen entsprechenden theoretischen Untersuchungsrahmen zu geben. Im Gegensatz zum bisherigen, von vornherein durch seine Symmetrie, also a priori bekannten "Würfelmodell" wählt Bernoulli nun das Modell einer Urne; er sagt:"Damit durch ein Beispiel deutlich werde, was ich meine, nehme ich an, es seien in einer Urne ohne dein Vorwissen 3000 weiße und 2000 schwarze Steinchen und du wollest durch Versuche das Verhältnis derselben bestimmen, indem du ein Steinchen nach dem anderen herausnimmst (jedoch so, daß du jedes gezogene Steinchen wieder zurücklegst, ehe du ein neues herausnimmst, damit die Zahl der Steinchen in der Urne nicht kleiner wird) und beobachtest, wie oft ein weißes, wie oft ein schwarzes Steinchen herauskommt." (S.91) Dieses Urnenmodell bezieht in der von Bernoulli beschriebenen Weise systematisch die Zufälligkeit der Ereignisse mit ein; es ist gewissermaßen das erste Simulationsmodell realer Zufallserscheinungen, und es stellt somit einen ersten Übergang von der Symmetrie zum Zufall dar.

Aus diesen hier ausführlich wiedergegebenen Vorüberlegungen Bernoullis werden zusammenfassend mehrere Aspekte deutlich sichtbar:

Von dem konkret praktischen Interesse geleitet, den Anwendungsbereich der Wahrscheinlichkeitstheorie von idealsymmetrischen Glücksspielen auf echte Zufallsereignisse zu erweitern, sieht Bernoulli als erster die Notwendigkeit, einen mathematischen Zusammenhang zwischen der Wahrscheinlichkeit eines Ereignisses und seiner beobachtbaren relativen Häufigkeit herzustellen. Zudem betont er, daß dieser herzustellende Zusammenhang selbst von wahrscheinlichkeitstheoretischer Natur ist, und er kann so Anforderungen nach mathematischer Exaktheit stellen; diesen "Vorsichtsmaßnahmen" Bernoullis kann man nur beipflichten, wenn man die doch recht komplizierte Struktur des infrage stehenden Zusammenhangs kennt.

Insgesamt wird sich zeigen, daß mit Bernoullis Theorem nicht einfach nur dem alten kombinatorischen Wahrscheinlichkeitsbegriff eine größere Anwendungsbreite verschaffen wurde, sondern daß hiermit eine Entwicklung ihren Anfang nahm, die auf eine Veränderung des Wahrscheinlichkeitsbegriffs selbst hinauslief.

II.1.2.3 Zum Verhältnis von Wahrscheinlichkeit und relativer Häufigkeit - der Beweisgang des Bernoullischen Theorems

Im V. und letzten Kapitel "Lösung des vorigen Problems" wird der Beweis dieses wichtigen Theorems mit "vollkommener analytischer Strenge", so das Urteil Kolmogoroffs (vgl. Maistrov, 1974, S.75) gegeben. In Bernoullis Sprechweise lautet der Satz:"Es möge sich die Zahl der günstigen Fälle zu der Zahl der ungünstigen Fälle genau oder näherungsweise wie $\frac{r}{s}$, also zu der Zahl aller Fälle wie $\frac{r}{r+s} = \frac{r}{t}$ - wenn $r + s = t$ gesetzt wird - verhalten, welches letztere Verhältnis zwischen den Grenzen $\frac{r+1}{t}$ und $\frac{r-1}{t}$ enthalten ist. Nun können, wie zu beweisen ist, so viele Beobachtungen gemacht werden, daß es beliebig oft (z.B. c-mal) wahrscheinlicher wird, daß das Verhältnis der günstigen zu allen angestellten Beobachtungen innerhalb dieser Grenzen liegt als außerhalb derselben, also weder größer als $\frac{r+1}{t}$ noch kleiner als $\frac{r-1}{t}$ ist." (S.104)

Dieser Satz, so könnte man etwas vereinfacht sagen, ist die mathematische Präzisierung der Leitidee, mit welcher Bernoulli um 1689 in seinen "Meditationes" den Satz samt Beweis ankündigt:"Minùs à verâ proportione aberrare possum, si saepiùs quam si variùs observem." (Bernoulli, 1975, S.75) ("Ich kann weniger vom wahren Verhältnis abweichen, wenn ich öfter als wenn ich seltener beobachte.")

Die Entwicklung der Beweisidee vollzieht Bernoulli in den Meditationes anhand einfacher Beispiele; so schreibt er etwa konkret für eine Reihe von Beobachtungen jeweils die Anzahl der Fälle hin, die er in einfachen Situationen mit Hilfe des "Pascalschen Dreiecks" bestimmen kann. Aus solchen Zahlenbeispielen wird der allgemeine Beweis erarbeitet, der sich dann auch allein in der "Ars conjectandi" wiederfindet.

Der einfacheren Rechnung wegen setzt Bernoulli die Anzahl der (unabhängigen) Versuche gleich $n \cdot t = n(r+s)$ und studiert die Binomialentwicklung von

$$(r+s)^{n(r+s)};$$

er bestimmt hiervon das größte Glied als

$$M = \binom{nt}{nr} r^{nr} s^{ns} = \binom{nt}{ns} r^{nr} s^{ns} .$$

Dann weist er nach, daß man n so groß wählen kann, "daß die Summe aller Glieder von dem größten M an nach beiden Seiten bis zu den Gliedern

$$L_n \left(= \binom{nt}{n(s-1)} r^{n(r+1)} \cdot s^{n(s-1)}\right) \qquad \text{und}$$

$$R_n \left(= \binom{nt}{n(r-1)} r^{n(r-1)} \cdot s^{n(s+1)}\right)$$

(einschließlich) zur Summe aller übrigen Glieder, welche nach beiden Seiten außerhalb dieser Grenzen L_n und R_n liegen, ein Verhältnis von größerem Werte als irgendein gegebenes bildet." (S.99)

Während Bernoulli also das <u>Verhältnis</u> der Summen (grob) abschätzt, ließe sich der "direkte" Beweisgang vielleicht so skizzieren:

Die Wahrscheinlichkeit, daß die relative Häufigkeit h_n in das Intervall

$$\left[\frac{r-1}{t} , \frac{r+1}{t} \right]$$

fällt, ist

$$P = \sum_{k=(r-1)\cdot n}^{(r+1)\cdot n} \binom{nt}{k} p^k (1-p)^{nt-k}$$

(mit $p = \frac{r}{t}$). Anschließend wird gezeigt, daß für genügend große n diese Wahrscheinlichkeit nahe bei Eins liegt:

$$P > 1 - \varepsilon .$$

Nun, wie wir noch genauer sehen werden, ist es einfacher das Verhältnis der Summen als die Summen selbst zu bestimmen, und zwar auch in endlichen konkret benutzbaren Berechnungen und nicht nur in den von Bernoulli zunächst durchgeführten "Unendlichkeitsbetrachtungen". Die Hauptschwierigkeit liegt in der Bestimmung bzw. der möglichst genauen Abschätzung der Binomialkoeffizienten für große Zahlen n und k. War somit die Vorgehensweise Bernoullis, gerade auch für praktische Anwendungen sinnvoll, so führte die Betrachtung des Verhältnisses jedoch zu sehr groben Abschätzungen und veranlaßte etwa Karl Pearson dazu, dem Bernoullischen Theorem seine Bedeutung überhaupt abzusprechen (vgl. Pearson, 1925).

Auf der Grundlage von fünf Hilfssätzen, in welchen die vorhin skizzierten Schritte mathematisch exakt nachgewiesen wurden, kann Bernoulli sodann den Beweis seines Hauptsatzes schnell durchführen. Sieht man sich die Aussage des Theorems und den Beweisgang nochmals genau an, so stellt man fest, daß zwar die Wahrscheinlichkeit des betrachteten Ereignisses

$$p = \frac{r}{t} = \frac{r}{r+s} \,,$$

die als konstant vorausgesetzt wurde, auch während der Argumentationen in den Beweisgängen selbst unveränderlich blieb, daß man jedoch jeweils das Exaktheitsintervall

$$\left[\frac{r-1}{t} \,,\ \frac{r+1}{t}\right] \,,$$

in welches die realtive Häufigkeit "fällt", eigentlich beliebig "fein" vorgeben kann. D.h., Bernoulli betrachtet die durch das Verhältnis der Beobachtungsanzahlen gegebene Wahrscheinlichkeit $p = \frac{r}{t}$ als konstant, jedoch diese Anzahlen selbst als variabel. Mit anderen Worten, das von Bernoulli vorgeschlagene Urnenmodell zur Behandlung des

Problems ist nicht einfach eine konkrete, wirkliche Urne, vielmehr handelt es sich hierbei jeweils um eine ganze Klasse solcher Urnen, die ein gleiches Verhältnis von Steinchen zweier verschiedener Arten beinhalten.

Als ersten wichtigen Punkt kann man also festhalten, daß Bernoulli das Exaktheitsintervall variabel hält während des Beweisganges und daß gerade diese Variabilität dem Urnenmodell seine Wirksamkeit zur Analyse des praktischen Problems verleiht.

Führt man diese Variable explizit ein, so läßt sich das bewiesene Theorem (auch mathematisch formaler) folgendermaßen "übersetzen":
"Sei die Wahrscheinlichkeit des betrachteten Ereignisses $p = \frac{r}{t}$. Sei abhängig von $k \in \mathbb{N}$ das Exaktheitsintervall $\left[\frac{kr-1}{kt}, \frac{kr+1}{kt}\right]$. Dann können, wie zu beweisen ist, so viele Beobachtungen gemacht werden, daß für die relative Häufigkeit h_n gilt:

$$P\left(h_n \in \left[\frac{kr-1}{kt}, \frac{kr+1}{kt}\right]\right) \longrightarrow 1$$

(bzw. $P > \frac{c-1}{c}$ für jedes beliebige c)."

In den modernen Darstellungen wird dieser Aussage oft die kurze Form gegeben:

$$\bigwedge_{\varepsilon > 0} \lim_{n \to \infty} P(|h_n - p| < \varepsilon) = 1$$

Hierin drückt sich prägnant ausschließlich das Faktum aus, daß letztendlich mit"absoluter"Sicherheit die beobachtete relative Häufigkeit mit der Wahrscheinlichkeit des Ereignisses beliebig genau übereinstimmt. Die Häufigkeiten, so sagt man, "stabilisieren" sich um die Wahrscheinlichkeit. Es war tatsächlich wohl auch das hauptsächliche Anliegen Bernoullis, diese"Konvergenz"der relativen Häufigkeiten nachzuweisen.

Man kann jedoch dies hier zunächst nur unter dem Aspekt seiner Konvergenz betrachtete Theorem auch in anderer Weise interpretieren, wenn man genauer die spezifische Art und Weise der hier anzutreffenden Konvergenz (etwa ihre "Schnelligkeit" etc.) untersucht und nach einer Abschätzung für die Sicherheit P nach einer endlichen Anzahl von Beobachtungen fragt, was gerade für praktische Anwendungen des Grenztheorems von zentraler Bedeutung ist. Dabei tritt dann deutlicher die dem Theorem eigene mathematische Struktur samt der in ihm steckenden technischen und begrifflichen Probleme zutage. In moderner Schreibweise hätte man beispielsweise die folgende Form, deren strukturelle Ähnlichkeit mit dem "übersetzten" Bernoulli-Theorem sofort ins Auge fällt:

$$\bigwedge_{\varepsilon > 0} \quad \bigwedge_{\eta > 0} \quad \bigvee_{n_0} \quad \bigwedge_{n \geq n_0} \quad P(|h_n - p| < \varepsilon) > 1-\eta$$

Insgesamt hat man also drei variable Größen: einmal die Exaktheit der betrachteten Aussage, welche durch ε gemessen wird, dann die Sicherheit, mit welcher die Aussage gilt, gemessen durch η und schließlich die Anzahl der gemachten Versuche, gegeben durch n. Diese drei Parameter bedingen sich gegenseitig, man kann jeweils zwei festlegen und anschließend versuchen, den dritten abzuschätzen.

Es ist nun interessant zu sehen, daß auch Bernoulli sich schon diesem Problem widmet. Im Verlaufe der Beweisführung macht er eine Anmerkung zu seinen Hilfssätzen:"Gegen den vierten und fünften Hilfssatz könnte von denen, welche sich nicht mit Unendlichkeitsbetrachtungen befreundet haben" (S.100), ein Einwand erhoben werden. Hierbei handelt es sich um die mathematisch unpräzise Aussage, daß "die Summe aller Glieder ... unendlich oft größer ist als die Summe ebenso vieler Glieder, welche nach links ... folgen." (S.99)

"Diesen Bedenken, so fährt Bernoulli fort, kann ich nicht besser entgegentreten, als daß ich jetzt die Berechnung für einen endlichen Wert von n wirklich durchführe; ich werde zeigen, daß auch in einer endlich hohen Potenz des Binoms (das ist $(r+s)^{nt}$) die Summe der innerhalb der Grenzglieder L_n und R_n (einschließlich) stehenden Glieder zur Summe aller übrigen Glieder ein Verhältnis hat, welches jedes beliebig groß gegebene Verhältnis c an Wert übertrifft. Ist dies aber gezeigt, so muß der Einwand notwendigerweise in sich zusammenfallen." (S.100)

Der darauf durchgeführte "endliche" Nachweis ergibt eine Abschätzung für die drei Größen Exaktheit ε bzw. k, Sicherheit η bzw. c und der Anzahl n der Beobachtungen.

Bernoullis Abschätzung ist auf Grund der von ihm benutzten relativ elementaren Techniken recht grob, was ihren praktischen Wert doch sehr einschränkt.

Den von Bernoulli gefundenen Abschätzungsformeln kann man die folgende Form geben:

Bezeichne r die Zahl der günstigen, s die der ungünstigen und $t = r+s$ die aller Fälle. Sei abhängig von $k \in \mathbb{N}$ das folgende Exaktheitsintervall gewählt,

$$\left[\frac{kr-1}{kt}, \frac{kr+1}{kt}\right] = \left[\frac{r'-1}{t'}, \frac{r'+1}{t'}\right]$$

und sei die Sicherheit der Aussage mittels c festgesetzt. Dann wähle für die Anzahl der durchzuführenden Beobachtungen den größeren der beiden folgenden Werte:

$$n_1t' = mt' + \frac{s't'(m-1)}{r'+1} \quad , \text{ wobei } \quad m \geq \frac{\log c(s'-1)}{\log (r'+1) - \log r'}$$

und

$$n_2t' = mt' + \frac{r't'(m-1)}{s'+1} \quad , \text{ wobei } \quad m \geq \frac{\log c(r'-1)}{\log (s'+1) - \log s'}$$

$(t' = kt,\ r' = kr,\ s' = ks)$.

Mittels dieser so komplizierten Formeln lassen sich nun tatsächlich aus je zwei vorgegebenen Größen grobe Abschätzungen für die restliche ermitteln. So kommt Bernoulli in einem von ihm durchgerechneten Beispiel zu dem Resultat, daß bei einem Ereignis mit Wahrscheinlichkeit $\frac{3}{5}$ "nach dem oben bewiesenen Satze (es) mehr als 1000-mal wahrscheinlicher ist, daß bei 25 550 angestellten Beobachtungen das Verhältnis der günstigen zu allen Beobachtungen innerhalb der Grenzen $\frac{31}{50}$ und $\frac{29}{50}$ (diese einschließlich) liegt als außerhalb derselben. Setzt man $c = 10\,000$ oder $c = 100\,000$, so findet man auf gleiche Weise, daß 31 258 Beobachtungen notwendig sind, damit es 10 000-mal wahrscheinlicher ist, daß das angegebene Verhältnis innerhalb der genannten Grenzen liegt als außerhalb derselben, und daß, damit es 100 000-mal wahrscheinlicher wird 36 966 Beobachtungen nötig sind, und so fort in das Unendliche, indem man immer ein Vielfaches von 5 708 Beobachtungen zu 25 550 hinzuaddiert." (S.107)

II.1.2.4 Die Weiterentwicklung des Wahrscheinlichkeitsbegriffs auf der Grundlage des Bernoullischen Theorems

Bernoulli war in erster Linie an einer Zurückführung der relativen Häufigkeit auf die Gleichwahrscheinlichkeit interessiert, so wie er es auch in den letzten Sätzen seines Werkes betont:"Wenn also alle Ereignisse durch alle Ewigkeit hindurch fortgesetzt beobachtet würden (wodurch schließlich die Wahrscheinlichkeit in volle Gewißheit übergehen müßte), so würde man finden, daß alles in der Welt aus bestimmten Gründen und in bestimmter Gesetzmäßigkeit eintritt, daß wir also gewzungen werden, auch bei noch so zufällig erscheinenden Dingen eine gewisse Notwendigkeit, und sozusagen ein Fatum anzunehmen." (S.107)

In dem zusätzlich, der Präzisierung wegen durchgeführten "endlichen" Beweis des Theorems zeigt sich nun im Keim ein grundsätzliches strukturelles Problem des Wahrscheinlichkeitsbegriffs. Geht man von den idealsymmetrischen Gegenständen zu zufälligen Ereignissen über, nimmt man anstelle des Würfel-Modells das der (verdeckten) Urne, um dem Zufall wirklich Rechnung zu tragen, d.h. ihn zur Grundlage des (unbekannten) Untersuchungsgegenstandes zu machen, so läßt sich die Wahrscheinlichkeit nun nicht mehr a priori aus der Symmetrie in quasi "deterministischer" Weise bestimmen. Die Wahrscheinlichkeit eines zufälligen Ereignisses wird selbst zu einer "Zufallsgröße". Dies wird aus dem komplizierten Zusammenhang zwischen der relativen Häufigkeit und der Wahrscheinlichkeit, den Bernoullis Theorem gerade zum Ausdruck bringt, deutlich: Die relative Häufigkeit ist ein "statistisches" Maß für die Wahrscheinlichkeit.

Hier wird für den Wahrscheinlichkeitsbegriff zum ersten Mal in seiner Entwicklung die im ersten Kapitel angesprochene "Zirkularität" ansatzweise sichtbar; will man die Wahrscheinlichkeit eines bestimmten zufälligen Ereignisses durch die beobachtete relative Häufigkeit erklären, so muß man eigentlich immer schon vorher wissen, was Wahrscheinlichkeit "ist"; denn der Erklärungszusammenhang, der die Wahrscheinlichkeit mittels der relativen Häufigkeit bestimmen soll, unterliegt, wie wir vor allem im "endlichen" Fall gesehen haben, wiederum selbst einer Wahrscheinlichkeit, die man somit voraussetzen muß.

Man könnte sagen, Bernoulli versuchte, sich den Schwierigkeiten dieser Zirkularität dadurch zu entziehen, daß er zum idealen Grenzfall übergeht; bei unendlich vielen Versuchen erhält man scheinbar eine einfache Definitionsgleichung $h_n = p$. Jedoch auch hier wird, wie es sich im starken Gesetz der großen Zahlen zeigt: $P(\lim_{n \to \infty} h_n = \frac{r}{t}) = 1$, wiederum von Wahrscheinlichkeit Gebrauch gemacht. In der statistischen Praxis, in der man es immer mit einer endlichen Versuchszahl zu tun hat, benutzt man das Konzept der "kleinen Wahrscheinlichkeiten". Man vernachlässigt kleine Unsicherheiten und nimmt die entsprechende relative Häufigkeit als die Wahrscheinlichkeit. Hierbei muß man jedoch vorher zumindest schon wissen, was kleine bzw. große Wahrscheinlichkeiten sind.
Die hier angesprochene Zirkularität des Wahrscheinlichkeitsbegriffs wird in den verschiedensten Formen sichtbar. So steckt letztlich bereits in der Annahme der Gleichwahrscheinlichkeit, von einem entwickelteren Standpunkt her gesehen, dieser Aspekt; auch hier wird Wahrscheinlichkeit schon vorausgesetzt. Des weiteren hängt die in Bernoullis Theorem vorkommende "Sicherheits"-Wahrscheinlichkeit P von der Wahrscheinlichkeit $\frac{r}{t}$ des Ereignisses ab.

All' diese verschiedenen Aspekte der Zirkularität in den Definitionen und Anwendungen des Wahrscheinlichkeitsbegriffs, insbesondere erstmals in Bernoullis Theorem, weisen auf den theoretischen Charakter dieses Begriffs hin: Er kann nicht auf andere reduziert werden, sein Inhalt erschöpft sich nicht in Eigenschaften oder Strukturen der von ihm bearbeiteten Gegenstände. Der Zirkel bringt die relative Eigenständigkeit des Begriffs zum Ausdruck.

Ließ sich etwa aus der idealen Symmetrie des Würfels nach damaliger Auffassung die Bedeutung des Wahrscheinlichkeitsbegriffs ableiten, so ist dies auf der Grundlage des Zufalls nicht mehr möglich. Denn hier muß auch umgekehrt der unbekannte und noch unentwickelte Wahrscheinlichkeitsbegriff zu Hilfe genommen werden, um zu erklären, was "zufällig" bedeutet, um die spezifischen mit dem Konzept der Zufälligkeit verbundenen wichtigen Charakteristika (wie etwa den Begriff der Unabhängigkeit) der speziell von der Wahrscheinlichkeitstheorie bearbeiteten Gegenstände herauszupräparieren.

Auch diese wechselseitig abhängende Entwicklung und Herausarbeitung des "wahrscheinlichkeitstheoretisch-zufälligen" Gegenstandes einerseits und der Wahrscheinlichkeitstheorie bzw. der spezifisch-wahrscheinlichkeitstheoretischen Operationen mit diesem Gegenstand andererseits sind nichts anderes als Aspekte dieser von uns angesprochenen Zirkularität. Und dies verdeutlicht, daß in der "zirkelhaften" Struktur des Bernoullischen Theorems nicht nur das Problem der Begründung, sondern zusätzlich ein Problem der Anwendung theoretischer Begriffe ansatzweise zum Ausdruck kommt. Das Theorem von Bernoulli behandelt in erster und elementarer Form den Anwendungsbezug des Wahrscheinlichkeitsbegriffs: Die Anwendung dieses Begriffs, soviel zeigt sich an dieser Stelle, kann nicht einfach "naiv" oder unter rein pragmatischen Gesichtspunkten vorgenommen werden, sie ist selbst theorieabhängig;der Zusammenhang von Wahrscheinlichkeit und relativer Häufigkeit als einem einfachen Anwendungstyp wird von einer theoretischen Struktur, wie sie im Bernoulli-Theorem vorkommt, geprägt.

Bernoulli hat einen wichtigen Schritt in der Entwicklung der Wahrscheinlichkeitstheorie vorangetan, indem er die zu stark idealisierte Modellierung der Anwendungsgegenstände der Wahrscheinlichkeitstheorie mit Hilfe des Würfels durch das realere Urnenmodel ersetzte. Zudem, und das hat die Analyse des Beweises zum Bernoulli-Theorem gezeigt, benutzte er nicht einfach dieses Modell als Ersatz in simulativer Weise für konkrete praktische Situationen, vielmehr wurde es als Klasse von Modellen zu einem wichtigen theoretischen Untersuchungsinstrument. Gerade das Einführen von Variablen in sein Modell in Form der Größen der Exaktheit k, der Sicherheit c und der Anzahl der Beobachtungen n erlaubte es, sich dem gestellten Problem gegenüber flexibel zu verhalten und dann auch zum ersten Mal einen theoretischen Strukturzusammenhang zwischen den drei genannten Parametern in Form des Theorems aufzustellen, um so den "Zufall" zu behandeln bzw. die Wahrscheinlichkeit zufälliger Ereignisse zu messen. Darüber hinaus hat Bernoulli mit seinem "Axiom der moralischen Gewißheit" schon so etwas wie ein "Prinzip der kleinen bzw. großen Wahrscheinlichkeiten" ähnlich der später sogenannten "Cournotschen Brücke" zur Verfügung gehabt, als einer praktisch bedingten Regel, die später gerade zur "Überbrückung" der sich im Zirkel des Grenztheorems zeigenden Differenz zwischen den theoretisch bestimmten und praktisch brauchbaren Werten der Wahrscheinlichkeit diente.

Bernoulli gebührt das Verdienst, als erster Grundzüge der zur Bearbeitung des Phänomens "Zufall" notwendigen theoretischen Struktur herausgearbeitet zu haben; und nur in diesem Sinne kann man die anfangs zitierte Aussage Michel Loèves und ihre Bedeutsamkeit einschätzen, daß mit Bernoullis "Theorème d'or" die langsame Entwicklung der Wahrscheinlichkeitstheorie zu einer mathematischen Disziplin in Gang gesetzt wird. In dieser weiteren Entwicklung werden dann die noch vorhandenen technischen Mängel beseitigt und der Charakter der der Wahrscheinlichkeitstheorie spezifischen Anwendungsgegenstände Schritt für Schritt herausgearbeitet. Einmal sind hier konkret die sehr groben Abschätzungen zu nennen, zum anderen ist beispielhaft auf die unbedeutende Rolle des "Unabhängigkeitsbegriffs" für die Untersuchungen Bernoullis hinzuweisen: War bezüglich des von ihm benutzten Urnenmodells mehr oder weniger "empirisch" klar, was unabhängige Ziehungen sind, so wird gerade dieser Begriff für die Charakterisierung der "Zufällgikeit" von Ereignissen im weiteren Verlaufe immer wichtiger und seine theoretische Natur immer deutlicher.

Daß Bernoulli diese weitreichende Bedeutung des von ihm aufgestellten Zusammenhangs hat sehen können, ist wohl mit Bestimmtheit zu bestreiten; ihm genügte es, mit Hilfe der entwickelten Theorie den Nachweis zu erbringen, daß letztlich die Wahrscheinlichkeit mit der relativen Häufigkeit "übereinstimmt". Man kann sagen, daß Bernoulli gar nicht explizit die mit dem von uns angesprochenen Zirkel verbundenen Schwierigkeiten bemerkt hat, denn in dem für ihn in erster Linie interessanten idealen Grenzfall "löst" sich der Zirkel in gewisser Weise auf; letztendlich bleibt für Bernoulli der Wahrscheinlichkeitsbegriff seiner Definition gemäß ein kombinatorischer.

Trotz dieser Einschränkungen hinsichtlich der Allgemeinheit von Bernoullis wahrscheinlichkeitstheoretischer Konzeption, bleibt der grundsätzlich neuartige und richtungsweisende Charakter seiner Untersuchungen für die folgenden Arbeiten auf diesem Gebiet unbestritten. Daß Bernoulli tatsächlich mit seiner Konzeption auf den zentralen Kern der Wahrscheinlichkeitstheorie zielte und mit seinem Theorem eine erste Grundstruktur samt ihren kontroversen Aspekten zur weiteren Bearbeitung aufstellte, wird unseres Erachtens sehr schön in einer Aussage Hackings sichtbar, daß Vertreter fast einer "jeden wahrscheinlichkeitstheoretischen Schule" sich auf Bernoulli berufen: "He (Bernoulli) has been fathered with the first subjective conception of probability. Yet Richard von Mises could cast him as a stalwart frequentist. More recent statisticians such as A.P. Dempster say he anticipated Jerzy Neyman's approach to inference via confidence intervals. P.M. Boudot has argued that Bernoulli was a good inductivist and anticipated the theories of Rudolf Carnap. Since each of these approaches to probability is customarily deemed inconsistent with every other, each school claims Bernoulli as its own." (Hacking, 1975, S.143)

II.1.3. Erste Präzisierungen des Bernoullischen Theorems

Ein wichtiges Problem hatte Bernoullis Beweis des Grenztheorems mit sich gebracht: Gerade für praktische Zwecke mußte man eine bessere als die von Bernoulli gegebene Abschätzung der Sicherheit

$$P(|h_n - p| < \varepsilon)$$

finden.

Abraham de Moivres (1667-1754) Beiträge zur Wahrscheinlichkeitstheorie beinhalten einen ersten, wichtigen Schritt zur Lösung dieser Aufgabe.

II.1.3.1 Die Grundzüge der Wahrscheinlichkeitstheorie in de Moivres "Doctrine of Chances"

Insgesamt hat de Moivre zur Entwicklung der Mathematik in vielen Bereichen beigetragen (vgl.die ausgezeichnete Biographie de Moivres von Ivo Schneider, 1968/69). Seine wahrscheinlichkeitstheoretischen Arbeiten faßt Schneider folgendermaßen zusammen:"The demands of the calculus of probability led de Moivre to an approximation for the binomial coefficients $\binom{n}{m}$ for large values of n.

The interaction between de Moivre and James Stirling, particularly in regard to the asymptotic series for log (n!), is treated at length. This work supplied the foundation for de Moivre's limit theorem for the binomial distribution.

The calculus of probability, which occupied him from 1708 onward, became in time ever more the center of de Moivre's inquiries. Proceeding from contemporary collections of gambling exercises, de Moivre, by introducing an explicit measure of probability for the so-called Laplace experiments, found the beginnings of a theory of probability. De Moivre expanded the classic application of probability calculus to games of chance by addressing himself to the problem of annuities and by adopting Halley's work with its conception of "Probabi-

lity of life". De Moivre was the first to publish a mathematically formulated law for the decrements of life derived from mortality tables." (Schneider, 1968/69, S.177)
Dies macht schon die verschiedenartigen Fragestellungen und Probleme bezüglich der Wahrscheinlichkeitstheorie deutlich, welche de Moivre interessierten. Vor allem sind die von ihm entwickelten analytischen Methoden zur Behandlung von Glücksspielproblemen bedeutsam. So wird ihm zumindest die Idee der "erzeugenden Funktionen" (bei der Behandlung rekurrenter Folgen) zur Lösung des berühmten "Problems der Spieldauer oder des Ruins" zugeschrieben (vgl. Freudenthal/ Steiner, 1966, S.158-166; H.L. Seal, 1949, S.67ff).
Das Interesse de Moivres an der Rentenrechnung bezeugt seine Auffassung darüber, daß über die Glücksspiele hinaus dieser Lehre praktische Bedeutung zukommt.

In der zweiten und dritten Auflage seines Wahrscheinlichkeitsbuches:"The Doctrine of Chances" (1738, 1756) ist sein "Treatise of Annuities on Lives" abgedruckt; in der zu Beginn des Buches gegebenen Widmung an den Lord Carpenter merkt de Moivre an, nachdem er ausdrücklich hervorgehoben hat, "... that this Doctrine is so far from encouraging Play, it is rather a Guard against it ...", wie praktisch fruchtbar diese Rechnung ist:"... one Branch of this useful Knowledge extends to the valuation of Annuities founded on Contingencies of Life, and that I have made it my particular Care to facilitate and improve the Rules I have formerly given on that Subject; ... ."

Wir wollen unsere Analyse neben einigen Anmerkungen zu den bei de Moivre anzutreffenden Systematisierungstendenzen für einige Grundbegriffe und -regeln vor allem auf die Verbesserung der Abschätzung im Bernoulli-Theorem konzentrieren.

Die "Doctrine of Chances" (sowohl schon die 1. Ausgabe von 1718, umso mehr auch die stark erweiterte 3. Auflage von 1756, auf die wir uns im folgenden beziehen werden) wird häufig als das erste Lehrbuch der Wahrscheinlichkeitsrechnung überhaupt angesehen. (vgl. Schneider, 1972, S.61) Dieses Buch gliedert sich in ein Vorwort, die die Grundlagen der Wahrscheinlichkeitstheorie darstellende Einleitung, einen daran anschließenden Teil mit dem Titel:"Solutions of several sorts of Problems, deduced from the Rules laid down in the Introduction", dann der sogenanhten "Approximatio", in welcher Bernoullis Theorem behandelt wird und abschließend dem "Treatise of Annuities on Lives". Im Anhang befinden sich noch sieben Appendices zu den verschiedensten wahrscheinlichkeitstheoretischen Fragen.

"Für die Entwicklung des Wahrscheinlichkeitsbegriffs ist (die) ... 'Doctrine' in zweierlei Hinsicht zu beachten. Erstens stellt de Moivre jetzt in der auf das Vorwort folgenden 'Introduction', sicherlich auch angeregt durch die inzwischen erschiene und am Ende des Vorworts ausdrücklich erwähnte 'Ars conjectandi', als ersten und zentralen Begriff seiner Theorie die mathematische Wahrscheinlichkeit an den Anfang eines theoretischen Aufbaus. Zweitens enthält die Widmung und das Vorwort dieser Arbeit von 1718 bereits Ansätze für eine erkenntnistheoretische Einbettung des Begriffs 'chance' in das von Newton geprägte Weltbild." (Schneider, 1972, S.62)

Die hier erwähnte erkenntnistheoretische Interpretation der Wahrscheinlichkeit und Zufälligkeit bei de Moivre wird besonders ausgeprägt deutlich in den beiden "Remarks" im Anschluß an seinen Grenzwertsatz in der dritten Auflage der "Doctrine"; wir werden nach der Analyse dieses Theorems kurz auf de Moivres "objektiven" Wahrscheinlichkeitsbegriff eingehen.

In der "Introduction" der Doctrine findet sich gleich zu Beginn die Definition des Wahrscheinlichkeitsbegriffs:"The Probability of an Event is greater or less, according to the number of Chances by which it may happen, compared with the whole number of Chances by which it may either happen or fail. ... Thus if an Event has 3 Chances to happen and 2 to fail, the Fraction $\frac{3}{5}$ will fitly represent the Probability of its happening, and may be taken to be the measure of it." (de Moivre, 1756, 1967, S.1/2) Hieran schließen sich weitere Definitionen von Grundbegriffen wie beispielsweise dem der Erwartung, des Verlustes, des Gewinns usw. und von Grundregeln an. Der Summensatz erscheint nur in der speziellen Form, daß sich die Wahrscheinlichkeit zweier komplementärer Ereignisse zu 1 aufaddieren.

Demgegenüber, und das interessiert uns besonders im Hinblick auf den Grenzwertsatz, wird der Produktsatz sowohl für unabhängige als auch abhängige Ereignisse formuliert. Die Definition der Unabhängigkeit bzw. Abhängigkeit lautet: "Two Events are independent, when they have no connexion one with the other, and the happening of one neither forwards nor obstructs the happening of the other. Two Events are dependent, when they are so connected together as that the Probability of either's happening is altered by the happening of the other." (de Moivre, 1756, S.6)

Die "Unabhängigkeit" wird bei de Moivre also "absolut" definiert, allein aufgrund der vorliegenden Ereignisse ergibt sich "empirisch" die Unabhängigkeit; demgegenüber wird "Abhängigkeit" gerade bezüglich der sich dann ändernden Wahrscheinlichkeit definiert, wie dies modern durch den Produktsatz geschieht. Dies wird zudem aus den von de Moivre vorgestellten Beispielen sichtbar. Hieran anschließend findet sich dann der Produktsatz für abhängige Ereignisse formuliert:"... the Probability of the happening of two Events dependent, is the product of the Probability of the happening of one of them, by the Probability which the other will have of happening, when the first is considered as having happened; and the same Rule will extend to the happening of as many Events as may be assigned." (de Moivre, 1756, S.7) Hierin steckt bereits der Begriff der bedingten Wahrscheinlichkeit. Mit seiner Hilfe gibt de Moivre dann eine ganz allgemeine Version des Multiplikationssatzes: "But to determine, in the easiest manner possible, the Probability of the happening of several Events dependent, it will be convenient to distinguish by thought the order of those Events, and to suppose one of them to be the first, another to be the second, and so on: which being done, the Probability of the happening of the first may be looked upon as independent, the Probability of the happening of the second, is to be determined from the supposition of the first's having happened, the Probability of the third's happening, is to be determined from the supposition of the first and second having happened, and so on: then the Probability of the happening of them all will be the product of the Multiplication of the several Probabilities which have been determined in the manner prescribed." (de Moivre, 1756, S.7/8)

De Moivre hat als einer der ersten die Notwendigkeit gesehen, grundlegende Begriffe klar zu fixieren, um Ungenauigkeiten zu vermeiden. "... for fear that, in what is to be said afterwards, the terms independent or dependent might occasion some obscurity, it will be necessary, before I proceed any farther, to settle intirely the notion of those terms." (de Moivre, 1756, S.6) Benutzt werden diese beiden Begriffe von de Moivre jedoch in einer a priori entscheidbaren Weise; abhängig bzw. unabhängig sind gewisse Ereignisse von vornherein allein aufgrund der Versuchsbedingungen. Dies ist so selbstverständlich, daß im Zusammenhang mit Bernoullis Theorem, für dessen Beweis die Voraussetzung der Unabhängigkeit der einzelnen Versuche zentral ist, de Moivre jedoch mit keinem Wort auf diese wichtigen Voraussetzungen zu sprechen kommt. Aus der Unabhängigkeit, die als empirisch gegeben angesehen wird, ergibt sich deduktiv die Behauptung des Satzes; die Probleme, die mit dem "Nachweis" der Unabhängigkeit für konkrete Situationen bzw. mit der Frage nach dem methodologischen Status der Unabhängigkeit verbunden sind, spielen bei de Moivre noch überhaupt keine Rolle.

II.1.3.2 De Moivres Entwicklung des Theorems von Bernoulli

Kommen wir nun zu der von de Moivre gegebenen "Verschärfung" des Theorems von Bernoulli. Erstmals hat de Moivre seine "Approximatio ad Summam Terminorum Binomii $(a+b)^n$ in Seriem expansi" (vgl. R.C. Archibald, 1926, der im Anschluß an seinen Aufsatz de Moivres Original nochmals abdruckt) in lateinischer Sprache 1733 für einige Freunde drucken lassen; 1738 wird die "Approximatio" in englischer Sprache in die 2. erweiterte Auflage der "Doctrine", erweitert durch den ersten Remark, aufgenommen; auch in der 3. Auflage erscheint die "Approximatio", diesmal um den zweiten "Remark" erweitert. (Zur Entstehungsgeschichte der Approximatio vgl. R.C. Archibald, 1926, K. Pearson, 1924 und I. Schneider, 1968/69 und 1972)

Auch für de Moivre läßt sich konstatieren, daß ganz allgemein gesehen seine Arbeit am Bernoullischen Theorem unter der Intention der praktischen Anwendbarkeit der Wahrscheinlichkeitstheorie stand (vgl. O.B. Sheynin, 1970, S.233). Dies wird aus den beiden Problemen LXXII und LXXIII deutlich, die direkt vor der "Approximatio" in der "Doctrine" stehen.

Sofort hieran schließt die Arbeit "A Method of approximating the Sum of the Terms of the Binomial $(a+b)^n$ expanded into a Series, from whence are deduced some practical Rules to estimate the Degree of Assent which is to be given to Experiments" an. Und hierin kommt de Moivre gleich zu Beginn auf das zentrale Problem, welches schon Bernoulli bearbeitet hatte, zu sprechen:"Altho'the Solution of Problems of Chance often requires that several Terms of the Binomial $(a+b)^n$ be added together, nevertheless in very high Powers the thing appears so laborious, and of so great difficulty, that few people have undertaken that Task; for besides James and Nicolas Bernoulli, two great Mathematicians, I know of no body that has attempted it; in which, tho' they have shewn very great skill, and have the praise which is due to their

Industry, yet some things were farther required; for what they have done is not so much an Approximation as the determining very wide limits, within which they demonstrated that the Sum of the Terms was contained." (de Moivre, 1756, S.243)

Zur Bearbeitung dieser Aufgabe hatte de Moivre in seinen "Miscellanea Analytica" (1730) einige wichtige Methoden entwickelt und Probleme gelöst. Zum besseren Verständnis des Beweisganges bei de Moivre stellen wir nun explizit das von ihm behandelte Problem in einer uns geläufigen Form vor:

Sei $p = q = \frac{1}{2}$ die Wahrscheinlichkeit des betrachteten Ereignisses. Sei h_n die relative Häufigkeit bei n (unabhängigen) Versuchen und sei $\ell = \frac{1}{2}\sqrt{n}$ die Genauigkeitsgrenze. Abzuschätzen ist dann die Wahrscheinlichkeit P, daß h_n in das Intervall $\left[\frac{1}{2}n-\ell,\ \frac{1}{2}n+\ell\right]$ fällt:

$$P\left(|h_n - np| \leq \ell\right) =$$

$$P\left(|h_n - \tfrac{n}{2}| \leq \tfrac{\sqrt{n}}{2}\right) = \sum_{k=\frac{n}{2}-\ell}^{\frac{n}{2}+\ell} \binom{n}{k} \frac{1}{2^n} \, .$$

Die weitere Untersuchung befaßt sich nun mit einer möglichst genauen Berechnung der Summe:

$$\sum_{K=\frac{n}{2}-\ell}^{\frac{n}{2}+\ell} \binom{n}{k} \frac{1}{2^n} \, .$$

Diese Summe kann umgeschrieben werden in

$$\sum_{k=\frac{n}{2}-\ell}^{\frac{n}{2}+\ell} \frac{\binom{n}{\frac{n}{2}}}{2^n} \cdot \frac{\binom{n}{k}}{\binom{n}{\frac{n}{2}}}$$

Und von einer soweit vorangetriebenen Problemformulierung geht de Moivre eigentlich implizit aus. Die beiden in dieser Summe auftretenden Verhältnisse, das des mittleren Terms $T_o = \binom{n}{\frac{n}{2}}$ zur Summe 2^n aller Terme der Entwicklung des Binoms $(1+1)^n$ und des Terms $T_k = \binom{n}{k}$ zum mittleren T_o, hatte de Moivre nämlich schon in seiner "Miscellanea Analytica" berechnet. Diese Ergebnisse benutzt er im folgenden;

es gilt:

$$\frac{T_o}{2^n} \cong \frac{2}{B\sqrt{n}}$$

mit

$$\ln B = 1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} \quad \ldots$$

Die Zahl B erweist sich gerade als $\sqrt{2\pi}$; Stirling hatte diesen Wert vorgeschlagen und de Moivre gelang der mathematische Nachweis. (vgl. Schneider 1968/69, Kap.5.3; Adams, 1974, S.21 und K. Pearson, 1924) Dies erfolgte im Zusammenhang mit de Moivres Abschätzungen von $m!$, wobei er auf die Formel:

$$\ln m! = \frac{1}{2}\ln B + (m+\tfrac{1}{2})\ln m - m + \frac{1}{12m} - \frac{1}{360m^3} + \ldots$$

stieß, aus der sich dann die heute sogenannte Stirlingsche Formel:

$$m! \sim \sqrt{2\pi}\, m^{m+\frac{1}{2}}\, e^{-m}$$

ergibt (vgl. Czuber, 1897, S.72; und für einen modernen Beweis, Feller, 1968, S.52/53).

In de Moivres Beweisgang erhält man somit:

$$\frac{T_o}{2^n} \cong \frac{2}{\sqrt{2\pi n}}$$

Anschließend gibt de Moivre, unter Bezugnahme auf schon vorliegende Resultate, das Verhältnis eines beliebigen Gliedes zum mittleren so an:

$$\ln \frac{T_k}{T_o} = - \frac{2k^2}{n}$$

(vgl. Corollary I, S.245)

In moderner Form könnte man diese Formel so aufschreiben:

$$\frac{T_k}{T_o} = e^{\frac{-2k^2}{n}} .$$

Anzumerken ist jedoch hier, und dies ist wichtig für das weitere Verständnis, daß de Moivre für die Funktion e^x ausschließlich die Potenzreihendarstellung:

$$e^x = \sum_{\nu=o}^{\infty} \frac{x^\nu}{\nu!}$$

benutzt. Dies weist auf die bei de Moivre anzutreffende ausschließlich <u>instrumentelle</u> Natur des Funktionsbegriffs für die Wahrscheinlichkeitsrechnung hin.
Im "Corollary 2" wird die gesuchte Abschätzung nun gegeben. Aufgrund der bisher gemachten Überlegungen hat man folgendes erhalten:

$$\sum_{k=\frac{n}{2}-\ell}^{\frac{n}{2}+\ell} \frac{\binom{n}{\frac{n}{2}}}{2^n} \cdot \frac{\binom{n}{k}}{\binom{n}{\frac{n}{2}}} \sim \frac{2}{\sqrt{2\pi n}} \sum_{k=\frac{n}{2}-\ell}^{\frac{n}{2}+\ell} e^{-\frac{2k^2}{n}}$$

Bei de Moivre findet sich nun die betrachtete e-Funktion in ihrer Reihendarstellung wiedergegeben; zudem begnügt er sich der vorhandenen Symmetrie wegen, nur die Hälfte der Terme aufzusummieren und dann das erzielte Resultat am Ende zu verdoppeln.

$$\frac{2}{\sqrt{2\pi n}} \sum_{k=\frac{n}{2}-\ell}^{\frac{n}{2}} e^{-\frac{2k^2}{n}} = \frac{2}{\sqrt{2\pi n}} \sum_{k=\frac{n}{2}-\ell}^{\frac{n}{2}} \left(\sum_{\mu=0}^{\infty} \frac{1}{\mu!} (-2)^{\mu} \frac{k^{2\mu}}{n^{\mu}} \right)$$

$$= \frac{2}{\sqrt{2\pi n}} \sum_{k=\frac{n}{2}-\ell}^{\frac{n}{2}} \left(1 - \frac{2k^2}{n} + \frac{4k^4}{2n^2} - \frac{8k^6}{6n^3} + \dots \right)$$

Genau genommen steht nur diese letzte Form bei de Moivre. Die hier gesuchte Summe erhält er, indem er sie im Grunde durch das Integral

$$\int_0^{\ell} e^{-\frac{2k^2}{n}}$$

ersetzt. Konkret integriert de Moivre die unter der Summe stehenden Terme einfach gliedweise von 0 bis ℓ und erhält somit:

$$\frac{2}{\sqrt{2\pi n}} \left(\ell - \frac{2\ell^3}{3n} + \frac{4\ell^5}{5 \cdot 2n^2} - \frac{8\ell^7}{6 \cdot 7n^3} + \dots \right)$$

In dieser Formel wird $\ell = \frac{1}{2} . \sqrt{n}$ eingesetzt und es ergibt sich:

$$\frac{2}{\sqrt{2\pi}} \left(\frac{1}{2} - \frac{1}{3 \cdot 4} + \frac{1}{2 \cdot 5 \cdot 8} - \frac{1}{6 \cdot 7 \cdot 10} + \dots \right),$$

"which converges so fast, that by help of no more than seven or eight Terms, the Sum required may be carried to six or seven places of Decimals." (de Moivre , 1756, S.245).

Dieser Wert wird von de Moivre als 0.341 344 berechnet, also der gesuchte doppelte Wert als 0.682 688. Damit hat de Moivre seine im 3. Korollar formulierte Behauptung nachgewiesen:"... if it was possible to take an infinite number of Experiments, the Probability that an Event which has an equal number of Chances to happen or fail, shall neither appear more frequently than $\frac{1}{2}$ n + $\frac{1}{2}\sqrt{n}$ times, nor more rarely than $\frac{1}{2}$n - $\frac{1}{2}\sqrt{n}$ times, will be expressed by the double Sum of the number exhibited in the second Corollary, that is by 0.682 688, and consequently the Probability of the contrary, which is that of happening more frequently or more rarely than in the proportion above assigned will be 0.317 312, those two Probabilities together compleating Unity, which is the measure of Certainty: Now the Ratio of those Probabilities is in small Terms 28 to 13 very near." (de Moivre, 1756, S.246)

In einer heutzutage verständlichen Form könnte man deutlicher diesen Sachverhalt durch die folgende Formel ausdrücken:

$$\lim_{n \to \infty} P\left(\left|h_n - \frac{n}{2}\right| \leq \frac{\sqrt{n}}{2}\right) = 2\,\frac{2}{\sqrt{2\pi n}} \int_0^{\frac{\sqrt{n}}{2}} e^{-\frac{2x^2}{n}}\,dx \approx 0.682688.$$

In Korollar 4 stellt de Moivre dann beispielhaft die Anwendung seines Theorems auf <u>endliche</u> Versuchsanzahlen vor:"But altho' the taking an infinite number of Experiments be not practicable, yet the preceding Conclusions may very well be applied to finite numbers, provided they be great: for Instance, if 3600 Experiments be taken, make n = 3600, hence $\frac{1}{2}$ n will be = 1800, and $\frac{1}{2}\sqrt{n}$ = 30, then the Probability of the Event's neither appearing oftner than 1830 times, nor more rarely than 1770, will be 0.682 688." (de Moivre, 1756, S.246)

Im Korollar 5 befaßt de Moivre sich mit der Genauigkeitsgrenze $\ell = \frac{\sqrt{n}}{2}$ und beschreibt einige wichtige Aspekte in diesem Zusammenhang; später ist die wichtige Bedeutung der durch $\sqrt{n}$ ausgedrückten Streuung für die "statistische" Konvergenz erkannt worden. Die weiteren Korollare dienen de Moivre dazu, das bis dahin erreichte spezielle Resultat für $p = q = \frac{1}{2}$ auf den allgemeinen Fall zu erweitern. Dazu wird die Binomialentwicklung von $(a+b)^n$ für große n untersucht und es wird entsprechend wie oben vorgegangen. De Moivre erhält anschließend wiederum das Verhältnis eines beliebigen Gliedes der Binomialentwicklung zum mittleren Glied; in moderner Bezeichnung könnte man dies so schreiben:

$$\frac{T_k}{T_o} = e^{-\frac{(a+b)^2 k^2}{2abn}}$$

(vgl. Korollar 9, S.250). Abschließend macht de Moivre dazu folgende Anmerkung:"If the Probabilities of happening and failing be in any given Ratio of inequality, the Problems relating to the Sum of the Terms of the Binomial $(a+b)^n$ will be solved with the same facility as those in which the Probabilities of happening and failing are in a Ratio of Equality." (de Moivre, 1756, S.250)

Damit ist der wichtige mathematische Teil der "Approximatio" beendet. Zum Beweisgang bei de Moivre gibt nun Czuber die folgenden Erläuterungen:"Die zur vollen Lösung des Problems, das er sich gestellt, erforderliche Summierung hat de Moivre unrichtig ausgeführt. Er gibt als Wert für die Summe

$T_\ell + T_{\ell-1} + \ldots + T_1 + T_o$

den Ausdruck

$$\frac{2}{\sqrt{2\pi s}} \left\{1 - \frac{2\ell^3}{3s} + \frac{4\ell^5}{2\cdot 5s^2} - \ldots\right\};$$

dies aber ist der in eine Reihe entwickelte Wert des Integrals

$$\frac{2}{\sqrt{2\pi s}} \int_o^\ell e^{-\frac{2\ell^2}{s}} d\ell.$$

Das doppelte dieses Betrages faßt er schließlich als Wahrscheinlichkeit dafür auf, daß in s Beobachtungen jedes der Ereignisse A, B in einer Wiederholungszahl erscheine, die zwischen $\frac{s}{2} - \ell$ und $\frac{s}{2} + \ell$ enthalten ist. Der eine Fehler liegt in der Ersetzung der endlichen Summe

$$\sum_{x=o}^{\ell} T_x$$

durch das Integral $\int_o^{\ell} T_x dx$;

der andere darin, daß vermöge der Verdoppelung das maximale Glied T_o zweimal gezählt wird, statt wie jedes andere nur einfach genommen zu werden; dieser zweite Fehler würde nur in dem Falle, wo A, B gleichwahrscheinlich sind und n eine ungerade Zahl ist, entfallen.

Das Mittel, auch diesen Teil der Aufgabe korrekt durchzuführen ist später durch Maclaurin (1742) und Euler (1755) gefunden worden." (Czuber, 1897, S.73/74)

Hierbei handelt es sich um die heute sogenannte Euler-Maclaurinsche Summenformel (vgl. zur Geschichte dieser Formel, Dieudonné, 1978, Bd. I, S.26).

"(Es) blieb ... Laplace vorbehalten (ab 1781), mit den von Moivre, Stirling, Maclaurin und Euler geschaffenen Hilfsmitteln eine exakte, in gewissem Sinne abschließende analytische Formulierung des Bernoullischen Theorems zu geben. In der von uns gewählten Betrachtung lautet dieselbe dahin, es sei

$$P = \frac{2}{\sqrt{\pi}} \int_o^{\gamma} e^{-t^2} dt + \frac{e^{-\gamma^2}}{\sqrt{2\pi s\, pq}}$$

die Wahrscheinlichkeit dafür, daß die Wiederholungszahl des Ereignisses A, dessen a priorische Wahrscheinlichkeit p ist, in einer sehr großen Anzahl s eingeschlossen sei zwischen den Grenzen

$$ps \pm \gamma \sqrt{2pqs}\ ,$$

das Verhältnis jener Wiederholungszahl zur Gesamtzahl der Versuche, also zwischen den Grenzen

$$p \pm \gamma \sqrt{\frac{2pq}{s}}\ ."$$

(Czuber, 1897, S.74/75; vgl. zur Herleitung des de Moivreschen Theorem bei Laplace Todhunter, 1865, S. 548ff.)

Auch bei Sheynin findet sich die hier gegebene Einschätzung, daß Laplace, wenngleich mathematisch exakter, in derselben Art und Weise wie de Moivre die Verschärfung des Bernoulli-Theorems durchführte:"... and Laplace goes on to find the sum of such sums by means of the Maclaurin-Euler summation formula. Exactly the use of this formula distinguishes Laplace's deduction from that of de Moivre." (Sheynin, 1976, S.148)

II.1.3.3 Zu einigen wesentlichen Aspekten in der wahrscheinlichkeitstheoretischen Arbeit de Moivres

Wir haben uns so ausführlich die mathematische Herleitung bei de Moivre angesehen, weil sie uns exemplarisch dafür zu sein scheint, wie in den Anfängen des Grenzwertsatzes bis zu den ersten wahrscheinlichkeitstheoretischen Arbeiten von Laplace die neue Infinitesimalrechnung benutzt wurde. Spricht man de Moivre berechtigterweise zu, als erster systematisch die Methoden dieser neuen Infinitesimalrechnung für die Berechnung wahrscheinlichkeitstheoretischer Probleme nutzbar gemacht zu haben, so muß man sich jedoch gleichzeitig nach der damals herrschenden Auffassung von der Analysis fragen.

Anhand des Funktionsbegriffs charakterisiert Verley kurz und übersichtlich die Entwicklung und die Auffassungen in der Analysis von 1650 bis ungefähr 1800:"Dans la seconde moitié du dix-septième siècle, apparaissent les premiers développements en série des fonctions élémentaires: développement en série de log (1+x) donné indépendamment par Newton (1665) et Mercator (1668), développements de sin x, cos x, Arc sin x, e^x (Newton dans De Analysi, 1669; Leibniz 1673), tandis que Taylor utilise la formule d'interpolation de Gregory-Newton pour obtenir, en 1712, la formule portant son nom qui donne le développement en série d'une fonction 'quelconque'. Nous renvoyons ..., pour l'analyse de la validité de ces divers résultats et insisterons ici essentiellement sur l'aspect algébrique (dans l'anneau des séries formelles): les propriétés des fonctions élémentaires s'obtiennent par des procédés algébriques, et, en langage actuel, les techniques de différentiation et d'intégration du calcul infinitésimal ont un caractère formel dans l'anneau des séries formelles considérées. Connue sous le nom d'analyse algébrique, cette étude des algorithmes illimités de nombres réels ou complexes et celle des méthodes spéciales permettant de représenter à l'aide de tels algorithmes les fonctions élémentaires sera le coeur des recherches des analystes du dix-huitième siècle. On sait que Lagrange essayera, dans sa Théorie des fonctions analytiques, de réduire précisément à une approche purement algébrique l'étude des fonctions, mais ce n'est pas avant Weierstrass ... que ce programme sera réalisé." (Verley, 1978, S.131; vgl. hierzu auch die damit übereinstimmenden Analysen von Grattan-Guinness, 1970, und Juschkewitz, 1976 und 1972).

Die hier beispielhaft in der Aussage Verleys skizzierte Auffassung der Analysis als einer formal-algebraischen Manipulation von Potenzreihen zeigte sich auch deutlich in den von uns wiedergegebenen Überlegungen de Moivres zum Bernoullischen Theorem. Den Methoden der Infinitesimalrechnung kommt für die Wahrscheinlichkeitstheorie dabei ausschließlich eine instrumentelle Hilfsfunktion zu, und zwar dienen sie ausschließlich der mathematisch-präziseren Approximation. Auch Adams bestreitet, daß de Moivre die e-Funktion etwa schon als eine Wahrscheinlichkeitsverteilung, also einen speziellen Funktionsbegriff als einen wahrscheinlichkeitstheoretischen Begriff an sich betrachtet hätte, wie dies beispielsweise bei Karl Pearson anklingt (vgl. Pearson, 1924 und 1925).

"It is in de Moivre's Approximation that we find the first use of integrals of e^{-x^2} to approximate sought after probabilities, and it is in this sense that de Moivre's theorems are prototypes of the Central Limit Theorem. It is perhaps natural, in a spasm of generosity, to attribute perspectives to de Moivre which developed later, as some writers have done. I can find no evidence to support the view that de Moivre thought of what we now call the normal law as itself being a probability distribution or as being a probability law of errors." (Adams, 1974, S.27)

Und des weiteren wird aus den anschließenden Arbeiten zum frühen zentralen Grenzwertsatz, wie sie von Daniel Bernoulli (1778) und dann vor allem von Laplace (ab 1781) vorgenommen wurden, der Hilfscharakter, die Form eines "approximation tool" (Adams) der Analysis für die Wahrscheinlichkeitstheorie deutlich. Vor allem ging es während dieser Phase darum, möglichst gute Approximationsmethoden für Integrale der Form e^{-x^2} zu finden, einmal durch Potenzreichen und dann auch durch Kettenbruchentwicklungen. Ähnlich Adams betont auch Sheynin, daß weder de Moivre noch Daniel Bernoulli der Verteilungskurve selbst Aufmerksamkeit geschenkt hätten. (vgl. O.B. Sheynin, 1977 Bd.II, S.1o1-1o4) Erst im Zusammenhang mit der Theorie der Beobachtungsfehler, zu der dann auch Laplace Beiträge leistete, entwickelt sich der Begriff der Verteilung und hiermit kommt dem Funktionsbegriff in Form der Wahrscheinlichkeitsverteilung erstmals grundlegende Bedeutung für die Wahrscheinlichkeitstheorie zu. Zudem bezieht man sich dabei dann nicht mehr bloß auf eine Zufallsvariable mit zwei Ausfällen, sondern auf ein ganzes System solcher Zufallsvariablen.

Doch die vor allem durch Jakob Bernoulli und Abraham de Moivre gelegten mathematischen Grundlagen für die Anwendungsbeziehung der Wahrscheinlichkeitstheorie in Form dieses Grenzwertsatzes, hat den Rahmen der weiteren Entwicklung umrissen.
"En 1733, de Moivre, en approfondissant le 'theorême d'or' à l'aide de la formule asymptotique pour n! obtenue par lui même et par Stirling, obtient la loi normale (ainsi nommée par Poincaré). Le leit-motiv de l'évolution mathématique du Calcul des probabilités est créé." (Michel Loève, 1978, Bd. II, S.282)

Zum Abschluß der Analyse von de Moivres wahrscheinlichkeitstheoretischer Arbeit möchten wir einige ergänzende Anmerkungen zu seiner Auffassung vom Status des Wahrscheinlichkeitsbegriffs machen, wie sie sich in den beiden "Remarks" zu Ende der "Approximatio" findet. Auf der Grundlage einer ausführlichen Analyse dieser "Remarks" kommt Ivo Schneider (vgl.Schneider,1972, S.75-79) zu dem folgenden Urteil:"Eine erkenntnistheoretische Deutung der in der 'Doctrine' eingeführten mathematischen Wahrscheinlichkeit bringt de Moivre erst mit der Interpretation seines 1733 entdeckten Grenzwertsatzes, der das Bernoullische Theorem verschärfte. Diesen Grenzwertsatz, mit dem de Moivre ähnlich wie Jakob Bernoulli das Umkehrproblem einer Bestimmung von Wahrscheinlichkeiten aus relativen Häufigkeiten leisten zu können glaubte, sah de Moivre als die mathematische Bestätigung einer planvollen Ordnung in der Welt durch 'determinate laws' und damit als Basis eines Theismus an. Diese 'determinate laws' sind für de Moivre die durch den Grenzübergang zu unendlich vielen Beobachtungen erreichten vom Informationsstand eines Subjekts völlig unabhängigen und damit 'objektiven' Wahrscheinlichkeiten.

De Moivres mit der Bestätigung eines Theismus verknüpfter objektiver Wahrscheinlichkeitsbegriff ist historisch bedingt auch durch die philosophisch-theologische Auseinandersetzung der Kirche in England mit dem naturwissenschaftlichen Weltbild Newtons." (Schneider, 1972, S.196) Diese resümierende Analyse Schneiders macht deutlich, wie gegensätzlich die Interpretationen des Wahrscheinlichkeitsbegriffs waren. Zudem wird klar, wie stark diese strikt alternative Auffassung von einerseits subjektiver und andererseits objektiver Wahrscheinlichkeit abhängig war von der dem jeweiligen Autor eigenen philosophischen, erkenntnistheoretischen oder sogar theologischen Auffassung und nicht primär die dem Wahrscheinlichkeitskonzept spezifischen Aspekte der Erkenntnis betraf. Denn mit einer fast gleichartigen Begründung hat Jakob Bernoulli seinen Wahrscheinlichkeitsbegriff "subjektiv" interpretiert:"Aufgrund göttlicher Voraussicht und Vorherbestimmung" (Bernoulli 1899, S.71) hat alles die höchste Gewißheit; der Wahrscheinlichkeit kommt ausschließlich die "subjektive" Rolle zu, Mittel zur Annäherung an diese absolute Gewißheit zu sein.

Wir glauben, diese kontroverse Diskussion um die Interpretation des Wahrscheinlichkeitsbegriffs nur in der Analyse der weiteren Entwicklung der Wahrscheinlichkeitstheorie sinnvoll, d.h. konkreter auf den Wahrscheinlichkeitsbegriff selbst und seine Bedingungen bezogen, aufnehmen zu können; vor allem die weitere Präzisierung der <u>Anwendungsbeziehung</u> wird deutlicher den der Wahrscheinlichkeitstheorie spezifischen "Gegenstand" herausarbeiten und somit zur Frage der objektiven bzw. real-wirksamen Funktionen dieses Begriffs beitragen. Erst auf diesem Hintergrund läßt sich u.E. sinnvoll die Frage nach dem Verhältnis subjektiver und objektiver Momente des Wahrscheinlichkeitsbegriffs wie eines jeden anderen mathematischen Begriffs stellen.

II.1.3.4 Anmerkungen zum Theorem von Bayes

Gerade im Zusammenhang des Theorems von Bernoulli und seiner Verschärfung durch de Moivre erscheint es uns notwendig, kurz auf die Arbeit von Thomas Bayes (1702-1761) einzugehen; denn oft wird gesagt, daß erst Bayes das vollbracht hat, was schon Bernoulli und de Moivre im Grunde intendierten: nämlich die Abschätzung einer "wirklich" unbekannten Wahrscheinlichkeit p aufgrund der beobachteten relativen Häufigkeiten h_n. Sowohl Bernoulli als auch de Moivre mußten zur Herleitung und zum Nachweis des Grenzwertsatzes relativ viel voraussetzen; neben der stillschweigenden Annahme, daß die einzelnen Versuche voneinander unabhängig sind und daß die Wahrscheinlichkeit p des betrachteten Ereignisses konstant bleibt, war es sogar notwendig, wollte man bei endlich vielen Versuchen, eine Abschätzung vornehmen, die Wahrscheinlichkeit p selbst zu kennen.

Die Arbeit von Bayes: "An essay towards solving a problem in the doctrine of chances" (wiederabgedruckt in Pearson/ Kendall, 1970, S.131-154) wurde 1763 posthum von seinem Freund Richard Price veröffentlicht. Sie beinhaltet eine von Price verfaßte Einführung, 2 Teile, von denen der erste systematisch Grundbegriffe und -regeln darstellt und im zweiten das eigentliche Problem bearbeitet wird, und abschließend einen wiederum von Price gegebenen Appendix mit Anwendungen.

Bemerkenswert ist in dem systematischen Teil, daß Bayes eine recht "moderne" Definition der Unabhängigkeit gibt, die nicht "direkt" an die betrachteten Ereignisse selbst, sondern an ihre Wahrscheinlichkeiten anknüpft: "Events are independent when the happening of any one of them does neither increase nor abate the probability of the rest." (Bayes, 1970, S.137) Diese Definition wird im folgenden benutzt, um den Produktsatz (Prop. 6, S.139) herzuleiten; auch im Nachweis der Bernoullischen Formel (Prop.7, S.140) findet sich ein Hinweis auf die unterliegende Unabhängigkeit der Versuche: "For the happening or failing of an event in different trials are so many independent events." (Bayes, 1970, S.140) In der Form dieser Bernoullischen Formel geht dann die Unabhängigkeit mehr oder weniger explizit als eine wichtige Voraussetzung in die Behandlung des eigentlichen Problems ein. Für den Bayesschen Essay läßt sich also im Vergleich zu Bernoulli und dann zu de Moivre sagen, daß immer systematischer die Grundbegriffe, die zunächst implizit vorausgesetzt und benutzt wurden, eingeführt, exakt definiert werden und ihre Benutzung im Verlaufe der Beweisführung deutlicher gekennzeichnet wird. Jedoch trotz der wachsenden Aufmerksamkeit, die den Grundbegriffen und -regeln geschenkt wird, bleibt festzuhalten, daß natürlich auch bei Bayes dem Unabhängigkeitsbegriff ausschließlich empirische Bedeutung zukommt; d.h. ob Ereignisse unabhängig sind oder nicht, ist a priori aus den Bedingungen des Versuchsaufbaus etwa entscheidbar. Man sieht noch nicht die Notwendigkeit und hat dementsprechend auch keine Methoden, die Unabhängigkeit von Ereignissen theoretisch zu überprüfen.

Die Fragestellung des Bayesschen Problems steht in der Tradition von Bernoulli und de Moivre. Price weist in seiner Einführung darauf hin:"... it is also a problem that has never before been solved. Mr. de Moivre, indeed, the great improver of this part of mathematics, has in his Laws of Chance after Bernoulli, and to a greater degree of exactness, given rules to find the probability there is, that after a very great number of trials be made concerning any event, the proportion of the number of times it will happen, to the number of times it will fail in those trials, should differ less than by small assigned limits from the proportion of the probability of its happening to the probability of its failing in one single trial. But I know of no person who has shewn how to deduce the solution of the converse problem to this; namely, 'the number of times an unknown event has happened and failed being given, to find the chance that the probability of its happening should lie somewhere between any, two named degrees of probability'. What Mr. de Moivre has done therefore cannot be thought sufficient to make the consideration of this point unnecessary: especially, as the rules he has given are not pretended to be rigorously exact, except on supposition that the number of trials made are infinite; from whence it is not obvious how large the number of trials must be in order to make them exact enough to be depended on in practice." (Bayes, 1970, S.135).

Bayes beginnt dann seinen Essay mit der Formulierung des Hauptproblems:"Given the number of times in which an unkown event has happened and failed: Required the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named." (Bayes, 1970, S.136)

Wir wollen nun im folgenden nicht exakt die von Bayes durchgeführte Analyse seines sogenannten "Umkehrproblems" verfolgen, sondern versuchen, anhand einer Herleitung der Bayesschen Formel, wie sie später erwa auch von Laplace gegeben wurde, die grundsätzlichen Aspekte seiner Überlegungen im Unterschied zu dem Vorgehen Bernoullis und de Moivres zu skizzieren. (vgl. hierzu u.a. Czuber, 1897, S.91-110)

Die erste wichtige Überlegung von Bayes besteht darin, die unbekannte Wahrscheinlichkeit p des Zufallsereignisses selbst zu einer Zufallsvariablen zu machen und sich dafür ein geeignetes Modell zu konstruieren. Er nimmt dazu eine quadratische Platte und läßt auf diese Kugeln fallen. Dabei geht er davon aus, daß die Kugeln mit gleicher Wahrscheinlichkeit auf jeden Punkt zwischen der linken und rechten Begrenzung der Platte auftreffen können. Die erste gefallene Kugel teilt nun die Platte in ein unbekanntes Verhältnis p:q (p+q=1), indem parallel zu den Begrenzungen der Platte links und rechts eine Linie durch den Auftreffpunkt dieser Kugel gezogen wird. Im folgenden wird versucht, aus der Kenntnis der Anzahl der Experimentkugeln, die links oder rechts vom Strich auf die Platte gefallen sind, das unbekannte Verhältnis p:q zu ermitteln. Die Wahl dieses geometrischen Modells erlaubt es Bayes zudem, eine Behandlung des Problems mit kontinuierlichen

Methoden vorzunehmen.

Üblicherweise wird dieses von Bayes eingeführte Modell in das folgende Urnenmodell übertragen: Man nimmt unendlich viele Urnen mit jeweils unendlich vielen Kugeln; jede Urne beinhaltet ein spezielles Verhältnis $p:q$ $(p+q=1)$ von schwarzen zu weißen Kugeln. Dabei ist die Dichte der Urnen mit dem Verhältnis $p:q$ für jedes p dieselbe; man sagt, die Zufallsvariable p ist a priori gleichverteilt. Zunächst wird nun eine beliebige Urne mit unbekanntem Kugelverhältnis gewählt; anhand der Anzahl der aus dieser Urne gezogenen weißen Kugeln etwa will man nun auf die Urne, d.h. auf das durch sie repräsentierte Verhältnis $p:q$ zurückschließen. Hat man nun eine Urne ausgewählt und ihr m weiße Kugeln bei n Ziehungen (mit Zurücklegen) entnommen, so ist nach der Bernoullischen Formel die Wahrscheinlichkeit, dies gerade bei einer Urne mit dem Verhältnis $p:q$ gemacht zu haben:

$$\binom{n}{m} p^m \, (1-p)^{n-m}$$

für jedes $p \in [0,1]$.

Die Funktion

$$p \longrightarrow \frac{p^m (1-p)^{n-m}}{\int_0^1 p^m (1-p)^{n-m} \, dp}$$

$$p \in [0,1]$$

ist die Wahrscheinlichkeitsverteilung der Zufallsvariablen p. Der Erwartungswert dieser Zufallsvariablen bezüglich ihrer Verteilung ist nun die Wahrscheinlichkeit dafür, mit dem nächsten, also $(n+1)$-ten Zuge wiederum eine weiße Kugel zu erhalten:

$$E\left(p \middle/ \begin{array}{l}\text{bei } n \text{ Versuchen} \\ m \text{ weiße Kugeln} \\ \text{gezogen}\end{array}\right) = \frac{\int_0^1 p \, p^m (1-p)^{n-m} \, dp}{\int_0^1 p^m (1-p)^{n-m} \, dp}$$

$$= \frac{(m+1)! \quad (n-m)!}{(n+2)!} \cdot \frac{(n+1)!}{m! \; (n-m)!} = \frac{m+1}{n+2} .$$

Diese Wahrscheinlichkeit liegt also nahe bei der beobachteten relativen Häufigkeit. Es läßt sich nun eine in ihrer Struktur ähnlich dem Satz von Bernoulli lautende Aussage formulieren:

Die Wahrscheinlichkeit P, daß das zufällige (unbekannte) p nahe bei der beobachteten relativen Häufigkeit $h_n = \frac{m}{n}$ liegt, etwa innerhalb vorgegebener Grenzen $\pm \ell$, ist:

$$P = \frac{\int_{h_n-\ell}^{h_n+\ell} p^m (1-p)^{n-m}\, dp}{\int_0^1 p^m (1-p)^{n-m}\, dp} \; .$$

Diese Aussage stellt im Prinzip nun die von Bayes vorgelegte Formel dar.

Die zur Herleitung dieser Formel notwendigerweise getroffene Voraussetzung der a priori Gleichverteilung konnte nun später verallgemeinert werden auf andere a priori Verteilungen, ohne dieses Resultat im wesentlichen zu verändern.

Freudenthal und Steiner beurteilen die Formel von Bayes und ihre weitere Auswirkung folgendermaßen:"Zeitgenossen und Nachwelt gaukelte sie ein Paradies vor. Bayes' Formel war auch zu schön, um nicht mißbraucht zu werden. Laplace, Condorcet, Poisson und viele andere betrieben das dann auch gründlich. Sie geriet in Verruf und davon hat sie sich noch nicht erholt, trotz des Ernstes, mit dem Abraham Wald (1950) sie wieder aufgenommen hat.
Der schwache Punkt war natürlich die a priori Verteilung. Man rechtfertigte Bayes' Gleichverteilung mit der Maxime, daß bei vollständiger Unwissenheit alles gleichwahrscheinlich sei.
Das mag schon richtig sein. Wenn einer in vollständiger Unwissenheit Kopf und Adler spielt, sollte er nicht einmal den Unterschied zwischen Kopf und Adler kennen, und dann sind sie in der Tat für ihn gleichwahrscheinlich. Aber normalerweise ist die vollständige Unwissenheit eine Fiktion.
Auch sonst wandte man die Formel ganz kritiklos an. Ein Beispiel: Man erlaubte sich, so die Wahrscheinlichkeit zu berechnen, daß die Sonne morgen wieder aufgeht, wenn bekannt ist, daß sie

- seit der Schöpfung - jeden Tag aufgegangen ist." (Freudenthal/Steiner, 1966, S.184)

Zwar ist die Kenntnis der Wahrscheinlichkeit p des betrachteten Ereignisses für die Benutzung der Bayesschen Formel nicht notwendig, jedoch muß auch hier eine Voraussetzung gemacht werden, eine Voraussetzung, deren große Gewichtigkeit sich erst im Verlaufe weiterer Anwendungen und Entwicklungen zeigte. Im folgenden wurde dann auch deutlich, daß es nicht darum geht, ob bei Anwendungen der Wahrscheinlichkeitstheorie überhaupt Annahmen gemacht werden müssen oder nicht; bei jeder wahrscheinlichkeitstheoretischen Anwendung, so stellte sich heraus, müssen zunächst Annahmen irgendwelcher Art gemacht werden, etwa in bezug auf die Unabhängigkeit von Zufallsvariablen oder in bezug auf Schätzwerte der gesuchten Wahrscheinlichkeit. Wichtig ist es jedoch, ob mit Hilfe der jeweils entwickelten Anwendungsverfahren bzw. der wahrscheinlichkeitstheoretischen Meßvorschrift es dann möglich ist, nach der Messung zumindest Hinweise für das Erfülltsein der gemachten Annahmen zu erhalten. Die erzielten Resultate müssen in gewisser Weise die zunächst vorläufigen Voraussetzungen als zweckmäßig rechtfertigen.

Bei dem Bayesschen Verfahren bleibt nun die Unabhängigkeit bzw. die a-priori-Verteilung immer eine "subjektiv" gefällte Voraussetzung, deren Gültigkeit oder Zutreffen nicht durch das Verfahren von Bayes selbst "überprüft" werden könnte; die Voraussetzung der a-priori-Verteilung ist in gewisser Weise"zu stark", da in ihr schon "alles steckt": Man paßt nur entsprechend den gemachten Beobachtungen seine Verteilung sofort wieder an, man nimmt immer schon den so durch "Anpassung" erhaltenen Wert von p als den gesuchten wahren Wert. Man geht also hierbei nicht so vor, daß man aus gemachten Voraussetzungen wahrscheinlichkeitstheoretische Konsequenzen zieht und diese mit einem empirisch beobachteten Wert entsprechend einer Meßvorschrift vergleicht, um so Hinweise für das Zutreffen der gemachten Annahmen und der Zulässigkeit der Anwendung der Wahrscheinlichkeitstheorie zu erhalten. Bei dem Bayesschen "Meßverfahren" wird jede Beobachtung sofort adaptiert, und es gibt somit keinen Unterschied zwischen einem theoretisch-prognostizierten und einem empirisch-beobachtbaren Wert, aus welchem man Rückschlüsse auf die Richtigkeit der Anwendung ziehen könnte. Auf die hiermit verbundenen Probleme wollen wir in der systematischen Analyse im Kapitel III nochmals zurückkommen.

Aus diesen Schwierigkeiten mit dem Bayesschen Theorem ergab sich eine heftige Kontroverse um die subjektive und objektive Wahrscheinlichkeit. "This paper has been the focus of what may be the most heated controversy in the history of probability and statistics, a controversy that extends to the present time." (Neyman/Le Cam, 1965, S.III)

Dabei ging es vor allem um den subjektiven Charakter der a-priori-Verteilung; ungerechtfertigte Annahmen etwa von Gleichverteilungen führten zu den von Freudenthal und Steiner erwähnten Kuriositäten, ja sie trugen mit dazu bei, daß in der Zeit nach Laplace viele die Wahrscheinlichkeitstheorie als "mathematischen Zirkus" betrachteten.

Zugegeben, auch Bernoulli und de Moivre mußten starke Voraussetzungen machen, um ihre Ergebnisse herleiten zu können und zudem waren auch ihre Methoden noch nicht geeignet, die Gültigkeit gemachter Voraussetzungen nachzuweisen. In der sich an Bernoulli und de Moivre anschließenden Weiterentwicklung des zentralen Grenzwertsatzes konnte man jedoch diese Voraussetzungen weiter "abschwächen" und mit der Herausbildung des Begriffs "Wahrscheinlichkeitsverteilung" auch Methoden zur Überprüfung dieser Voraussetzungen verstärkt entwickeln.

Die zentrale Entwicklung des wahrscheinlichkeitstheoretischen Anwendungsbezuges vollzieht sich historisch in der Nachfolge der von Bernoulli und de Moivre gelegten Grundlagen. Aber auch die kurze Analyse des Bayesschen Essays, in dem ja die gleiche Problematik behandelt werden sollte, schien uns als kontrastierender Vergleich zu dem Theorem von Bernoulli und de Moivre notwendig; historisch betrachtet gehört er zu den frühen Bemühungen, mathematische Verfahren für den wahrscheinlichkeitstheoretischen Anwendungsbezug zu entwickeln.