Skripte zur Wissenschaftstheorie, II

Epistemische und aleatorische Wahrscheinlichkeit

G. Rohwer

Version 1 **Februar 2000**

Vorbemerkung. Bei den „Skripten zur Wissenschaftstheorie“ handelt es sich um Texte, die als Leitfäden für Seminare zur Wissenschaftstheorie der Sozialwissenschaften dienen sollen. Der vorliegende Text beschäftigt sich mit Grundbegriffen der Wahrscheinlichkeitstheorie.

Götz Rohwer
Fakultät für Sozialwissenschaft
Ruhr-Universität Bochum, GB 1
44780 Bochum

Email: `goetz.rohwer @ ruhr-uni-bochum.de`

wtb.tex Februar 2000
Homepage: `http://www.stat.ruhr-uni-bochum.de [lecture notes]`

Inhaltsverzeichnis

1 Einleitung **1**

2 Möglichkeiten und Hypothesen **5**
2.1 Reden über Möglichkeiten . . . 5
2.2 Vermutungen und Hypothesen . . . 9
2.3 Hypothesen über Sachverhalte . . . 10
2.4 Situationen und Ereignisse . . . 13
2.5 Vergangenheit und Zukunft . . . 15
2.6 Statische und dynamische Hypothesen . . . 18
2.7 Hypothesen über Dispositionen . . . 19
2.8 Situationen und Situationstypen . . . 21
2.9 Definite und indefinite Hypothesen . . . 24

3 Epistemische Wahrscheinlichkeit **27**
3.1 Bildung von Hypothesen . . . 27
3.2 Algebra für Hypothesen . . . 30
3.3 Epistemische Wahrscheinlichkeit . . . 32
3.4 Abhängigkeit vom Wissensstand . . . 37
3.5 Komparative Ordnung von Hypothesen . . . 41
3.6 Elementare Konsistenzbedingungen . . . 44
3.7 Bewertungen und Wahrscheinlichkeiten . . . 46
3.7.1 Additive Bewertungsfunktionen . . . 46
3.7.2 Bewertungen von Hypothesen . . . 48
3.8 Ergänzungen und Exkurse . . . 51
3.8.1 Begründung durch Verfahren? . . . 52
3.8.2 Wahrscheinlichkeit als Relation? . . . 53
3.8.3 Logisch vs. empirisch? . . . 60

4 Aleatorische Wahrscheinlichkeit **63**
4.1 Einleitende Bemerkungen . . . 63
4.2 Zufallsgeneratoren . . . 67
4.2.1 Verfahren zur Ereigniserzeugung . . . 67
4.2.2 Elementare Zufallsgeneratoren . . . 71
4.3 Aleatorische Wahrscheinlichkeitsmaße . . . 73
4.3.1 Motivation der Definition . . . 73
4.3.2 Aleatorische Wahrscheinlichkeitsaussagen . . . 74
4.3.3 Einwände und Mißverständnisse . . . 78
4.4 Zufallsvariablen . . . 82
4.4.1 Einfache Zufallsvariablen . . . 82
4.4.2 Mehrdimensionale Zufallsvariablen . . . 87
4.4.3 Funktionen von Zufallsvariablen . . . 88

4.5 Allgemeine Zufallsgeneratoren 89
4.5.1 Verwendung von Zufallsvariablen 89
4.5.2 Unabhängige Verbindungen 90
4.5.3 Mischungen 93
4.6 Bedingte Wahrscheinlichkeiten 94

5 Erwartungswerte und Wiederholungen 97
5.1 Erwartungswerte und Varianzen 98
5.1.1 Erwartungswerte 98
5.1.2 Varianzen und Kovarianzen 101
5.2 Unabhängige Wiederholungen 105
5.2.1 Repetitive Zufallsgeneratoren 105
5.2.2 Die Binomialverteilung 106
5.3 Aussagen über Wiederholungen 108
5.3.1 Reformulierung der Fragestellung 108
5.3.2 Bernoullis Gesetz der großen Zahlen 109
5.4 Wahrscheinlichkeiten und Erwartungen 111

6 Frequentistische Definitionsansätze 115
6.1 Einleitende Bemerkungen 115
6.2 Grenzwerte von Häufigkeiten? 121
6.3 Zufällige Folgen 129
6.3.1 Was heißt 'zufällig'? 129
6.3.2 v. Mises' Regellosigkeit 134
6.4 Sinngrenzen für Anwendungen 137

Literatur 141

Kapitel 1

Einleitung

1. Ein Studium der Sozialwissenschaften beinhaltet fast immer eine Einführung in die Statistik, und in deren Kontext wird meistens auch eine Einführung in die Wahrscheinlichkeitsrechnung vermittelt. Allerdings fehlt es dann oft an Zeit, um sich mit ihren Grundbegriffen in der erforderlichen Ausführlichkeit zu beschäftigen. Der vorliegende Text setzt hier an. Das Ziel besteht darin, eine Reihe von Überlegungen zur Diskussion zu stellen, die die Frage betreffen, wie Wahrscheinlichkeitsaussagen verstanden werden können.

2. Im Hinblick auf dieses Ziel ist es sinnvoll, nicht sogleich mit Begriffen der Wahrscheinlichkeits*rechnung* anzufangen. Denn wie schon der Name sagt, wird in der Wahrscheinlichkeitsrechnung von vornherein vorausgesetzt, daß 'Wahrscheinlichkeit' ein metrisierbarer Größenbegriff ist. Ob und ggf. in welchen Kontexten eine solche Voraussetzung begründet werden kann, ist jedoch nicht selbstverständlich. Tatsächlich wird das Wort 'wahrscheinlich' zunächst außerhalb des engen Umkreises der Wahrscheinlichkeitsrechnung in *epistemischen* Kontexten verwendet, nämlich zur Qualifizierung von Vermutungen, ohne dafür vorauszusetzen, daß auch numerische Wahrscheinlichkeitsgrade angegeben werden können. In erster Näherung kann man sich an dann einem Begriffsverständnis orientieren, das von I. Kant folgendermaßen formuliert worden ist:

„Unter Wahrscheinlichkeit ist ein Fürwahrhalten aus unzureichenden Gründen zu verstehen, die aber zu den zureichenden ein größeres Verhältnis haben, als die Gründe des Gegenteils.“ (Kant 1800, S. 512)

3. In den Kapiteln 2 und 3 wird versucht, ein etwas genaueres Verständnis epistemischer Wahrscheinlichkeitsaussagen zu gewinnen. Der Leitgedanke ist, daß mit solchen Aussagen angestrebt wird, die Gründe, die für oder gegen Hypothesen sprechen, einschätzbar zu machen. Ob und wie das gelingen kann, hängt natürlich insbesondere von der Art der Hypothesen ab, auf die man sich beziehen möchte; und wir beginnen deshalb in Kapitel 2 mit einer ausführlichen Diskussion unterschiedlicher Arten von Hypothesen. Erst im Anschluß daran, in Kapitel 3, folgen Überlegungen, wie man sich mit Wahrscheinlichkeitsaussagen auf Hypothesen beziehen kann. Dabei wird zu zeigen versucht, daß dies bestenfalls in einer komparativen Weise gelingen kann. Allerdings soll auch deutlich werden, daß ein komparatives Verständnis epistemischer Wahrscheinlichkeitsaussagen es nicht ausschließt, auch ggf. verfügbare Wahrscheinlichkeitsmaße für ihre Begründung zu verwenden.

4. Eine Beschäftigung mit Wahrscheinlichkeitsrechnung im engeren Sinne beginnt in Kapitel 4. Um ein Verständnis zu gewinnen, folgen wir ihrer historischen Entwicklung aus einer Beschäftigung mit Glücksspielen bzw. – in einer abstrakteren Formulierung – mit Zufallsgeneratoren. Es wird also der Versuch gemacht, die Wahrscheinlichkeitsrechnung als eine mathematische Theorie der Zufallsgeneratoren zu verstehen. Dann läßt sich, wie ich zeigen möchte, auch ein quantitativer Wahrscheinlichkeitsbegriff begründen, und man kann einen Anschluß an die üblichen Begriffsbildungen der Wahrscheinlichkeitsrechnung finden, insbesondere ein Verständnis des Begriffs einer „Zufallsvariablen“ (im Unterschied zu statistischen Variablen) gewinnen.

5. Die Überlegungen führen also zu zwei unterschiedlichen Arten von Wahrscheinlichkeitsaussagen: einerseits zu epistemischen Wahrscheinlichkeitsaussagen, durch die Hypothesen einschätzbar gemacht werden sollen; andererseits zu aleatorischen Wahrscheinlichkeitsaussagen, die sich auf modale Eigenschaften von Zufallsgeneratoren beziehen. Daraus resultieren zwei Fragen. Zunächst stellt sich die Frage, in welchem Verhältnis die beiden Arten von Wahrscheinlichkeitsaussagen zueinander stehen. Dieser Frage wird in Abschnitt 4.3.2 und dann in Kapitel 5 nachgegangen. Dafür wird ein bereits in Abschnitt 3.7.2 entwickelter Gedanke verwendet: daß zwar epistemische Wahrscheinlichkeiten nicht quantifiziert, dennoch zur Begründung epistemischer Wahrscheinlichkeitsaussagen auch ggf. verfügbare Wahrscheinlichkeitsmaße verwendet werden können. Dies setzt natürlich voraus, daß sich die Hypothesen auf Situationen, Ereignisse oder Sachverhalte beziehen, für deren Zustandekommen ein Wahrscheinlichkeitsmaß begründet werden kann.

6. Versteht man die Wahrscheinlichkeitsrechnung als eine Theorie der Zufallsgeneratoren, setzt dies voraus, daß es sich um Situationen, Ereignisse oder Sachverhalte handelt, für deren Zustandekommen Zufallsgeneratoren unterstellt werden können. Somit stellt sich die Frage, unter welchen Bedingungen eine solche Unterstellung begründet werden kann. Das ist sicherlich dann möglich, wenn es sich um Sachverhalte oder Ereignisse handelt, die durch *artifizielle* Zufallsgeneratoren entstehen, wie z.B. bei Glücksspielen. Einen weiteren sinnvoll konzipierbaren Anwendungsbereich gibt es vielleicht dort, wo – ebenfalls artifizielle – Zufallsgeneratoren zur Auswahl von Beobachtungen, gewissermaßen als

„Beobachtungsgeneratoren“, verwendet werden können. Problematisch wird es indessen, wenn bei Anwendungen der Wahrscheinlichkeitsrechnung unterstellt wird, daß die Sachverhalte oder Ereignisse, auf die sich die Wahrscheinlichkeitsaussagen beziehen, durch *hypothetische* Zufallsgeneratoren erzeugt werden. Dieser Gedanke durchzieht insbesondere die historische Entwicklung der „theoretischen Sozialstatistik“, seit sie – etwa durch Quetelet, Fechner und Lexis – mit Begriffsbildungen der Wahrscheinlichkeitsrechnung verbunden worden ist. Solchen problematischen Anwendungen der Wahrscheinlichkeitsrechnung wird jedoch in diesem Skript nicht näher nachgegangen.[1]

7. Ein weiteres Problem stellt sich indessen im Hinblick auf die begrifflichen Grundlagen der Wahrscheinlichkeitsrechnung. Wenn man sie im Anschluß an ihre historische Entwicklung als eine Theorie der Zufallsgeneratoren konzipiert, bekommen Wahrscheinlichkeitsaussagen den Charakter von Aussagen über Möglichkeiten im Sinne von dispositionalen Eigenschaften von Zufallsgeneratoren. Von Aussagen dieser Art führt jedoch kein direkter (deduktiver) Gedankengang zu Aussagen über empirisch realisierte Sachverhalte oder Ereignisse, insbesondere nicht zu relativen Häufigkeiten in empirisch realisierten Gesamtheiten von Sachverhalten oder Ereignissen. Dieser „logische Bruch“ kann, wie ich in Kapitel 5 zu zeigen versuche, nur wiederum durch *epistemische* Wahrscheinlichkeitsaussagen überbrückt werden. Allerdings gibt es konkurrierende Konzeptionen, bei denen versucht wird, den „logischen Bruch“ dadurch zu vermeiden, daß Wahrscheinlichkeiten nicht modal, sondern durch relative Häufigkeiten definiert werden. Zwei Varianten solcher „frequentistischen“ Konzeptionen werden in Kapitel 6 besprochen.

8. Schließlich noch eine Vorbemerkung. In diesem Text wird oft von umgangssprachlichen Redeweisen ausgegangen, die man sich an einfachen Beispielen überlegen kann. Das erscheint sinnvoll, weil ein Verständnis unserer Erfahrungswelt zunächst an ein Verständnis unserer Umgangssprache gebunden ist. Damit ist zwar nicht gemeint, daß umgangssprachliche Formulierungen stets ausreichend sind. Dennoch erscheint mir richtig, was J. L. Austin (1979, S. 185) einmal so gesagt hat:

> “Certainly, then, ordinary language is *not* the last word: in principle it can everywhere be supplemented and improved upon and superseded. Only remember, it *is* the *first* word.”

[1] Es ist geplant, dazu ein gesondertes Skript mit dem Arbeitstitel „Metaphysik der Sozialstatistik“ zu erstellen und in einem separaten Seminar zur Diskussion zu stellen.

Vorschläge zur sprachlichen Differenzierung und Präzisierung sollten deshalb aus einer Reflexion umgangssprachlicher Redeweisen gewonnen und durch sie motiviert werden. Dies betrifft auch die Verwendung symbolischer Notationen, die in diesem Text des öfteren verwendet werden, um Redeweisen zu vereinfachen und eindeutiger zu machen und um einen Anschluß an mathematische Notationen der Wahrscheinlichkeitsrechnung herzustellen. Schließlich besteht jedoch die Aufgabe auch dieser symbolischen Notationen nur darin, Hilfsmittel zur Reflexion unseres gewöhnlichen Redens bereitzustellen.

Kapitel 2

Möglichkeiten und Hypothesen

Wahrscheinlichkeitsaussagen treten in unterschiedlichen Formen auf, und es gibt unterschiedliche Verständnisse ihrer Bedeutungen. Ein Zugang zum Verständnis knüpft an unser Reden über Möglichkeiten an. Damit beschäftigen wir uns in diesem und im folgenden Kapitel.

2.1 Reden über Möglichkeiten

1. Eine elementare Unterscheidung betrifft Tatsachen und Möglichkeiten. Gleichgültig, in welcher Situation man sich befindet, ist man überhaupt in der Lage, auf die Situation zu achten, kann man stets auch mindestens einige Merkmale der Situation *als Tatsachen* feststellen. Angenommen, wir sind Teilnehmer eines Seminars; dann können wir zum Beispiel die in der gegebenen Situation anwesenden Teilnehmer zählen und dadurch ihre Anzahl feststellen und die Seminarsituation charakterisieren. Aber das Seminar soll in einer Woche noch einmal stattfinden. Die Teilnehmer an dem Seminar, das erst in einer Woche stattfinden wird, können wir heute noch nicht zählen. Es ist unbestimmt, wieviele Personen an dem Seminar, das in der nächsten Woche stattfinden soll, teilnehmen werden. Es ist möglich, daß die Anzahl der Teilnehmer gleich bleibt, aber auch, daß sie kleiner oder größer wird. Das sind Möglichkeiten, über die wir heute bestenfalls Vermutungen anstellen können. Nicht einmal ist sicher, daß das Seminar in einer Woche tatsächlich stattfinden wird. Immerhin ist angekündigt worden, daß es in einer Woche wieder stattfinden soll. Insofern gibt es einen Grund für die Annahme, daß es tatsächlich stattfinden wird; und stattdessen könnten wir auch sagen: es ist wahrscheinlich, daß es in einer Woche wieder stattfinden wird.

2. Das Beispiel zeigt, daß man sich auf Möglichkeiten in der sprachlichen Form von Vermutungen beziehen kann. Das Wort ‘Vermutung’ soll zum Ausdruck bringen, daß man etwas nicht sicher oder nicht genau weiß, und es impliziert, daß sich eine Vermutung als falsch herausstellen kann. Daraus bezieht auch das Wort ‘möglich’ in diesem Zusammenhang seinen Sinn: es könnte anders sein.

3. Aber warum könnte irgendetwas anders sein – anders als man vermutet hat? Eine Möglichkeit ist unmittelbar einsichtig: es kann uns an

Wissen oder Fähigkeiten mangeln, so daß man infolgedessen keine zutreffenden Feststellungen machen, sondern nur Vermutungen äußern kann, die sich als falsch herausstellen können. Hierzu gehört auch, daß man Fehler machen kann. Es ist zum Beispiel möglich, daß man beim Zählen der Seminarteilnehmer einen Fehler gemacht hat und infolgedessen die ermittelte Anzahl nicht zutreffend ist. Bei dieser Betrachtungsweise wird vorausgesetzt, daß es einen Sachverhalt gibt und das Problem nur darin besteht, daß wir nicht oder nur mangelhaft in der Lage sind, den Sachverhalt in einer geeigneten sprachlichen Form festzustellen. Wir nehmen zum Beispiel an, daß die Anzahl der Personen, die sich zu einem gewissen Zeitpunkt in einem Fußballstadion aufhalten, als eine Eigenschaft der Situation bestimmt ist; aber es könnte sein, daß wir nicht in der Lage sind, diese Anzahl genau zu ermitteln.

4. Die eben angedeutete Betrachtungsweise erscheint sinnvoll, wenn und insoweit bestimmte Sachverhalte als gegeben angenommen werden können. Infolge dieser Annahme sind unzutreffende Vermutungen eine Folge oder ein Ausdruck mangelhaften Wissens oder – in gewisser Weise allgemeiner – mangelhafter Fähigkeiten (zur Wissensbildung), wozu auch gehört, daß man Fehler machen kann. Die Annahme impliziert, daß es zutreffende Feststellungen geben könnte.

5. Aber ist diese Annahme stets sinnvoll? Denken wir noch einmal an das Seminar, das – vielleicht, vermutlich, wahrscheinlich – in der nächsten Woche stattfinden wird. Diesem Seminar entspricht noch kein in unserer Erfahrungswelt realisierter Sachverhalt, es existiert – wenn dieser Ausdruck hier überhaupt angewendet werden kann – nur als eine zukünftige Möglichkeit. Sicherlich können wir gleichwohl Vermutungen anstellen. Wir können vermuten, daß es stattfinden wird; und wir können auch weitergehende Vermutungen anstellen, z.B. über die Anzahl der Teilnehmer, das Thema, den Ablauf des Seminars. Diese Vermutungen beziehen sich aber nicht auf bereits realisierte Sachverhalte, sondern auf Vorstellungen, die wir uns von einem zukünftigen Geschehen heute schon machen können. Nehmen wir jetzt an, daß wir einige solche Vermutungen angestellt haben. Sicherlich können sie sich als falsch erweisen. Es ist möglich, daß das Seminar nicht stattfindet, obwohl wir vermutet haben, daß es stattfinden wird. Wenn wir dann nach dem Grund fragen, warum unsere Vermutung unzutreffend gewesen ist, gibt es zwei verschiedene Möglichkeiten, nach einer Antwort zu suchen.

6. Die erste Möglichkeit entspricht der oben skizzierten Auffassung: wir könnten sagen, daß unsere Vermutung falsch gewesen ist, weil wir – in der Situation, in der wir die Vermutung gebildet haben – nicht gewußt haben, daß das Seminar nicht stattfinden wird. Diese Betrachtungsweise erscheint sinnvoll, wenn und insoweit wir hätten wissen können, daß das Seminar nicht stattfinden wird. So oder so, wir suchen dann den Grund für die Falschheit unserer Vermutung in einem Mangel an Wissen, das uns zur Formulierung der Vermutung zur Verfügung stand, oder in fehlerhaften Überlegungen bei der Verwendung des verfügbaren Wissens. Um auf diese Betrachtungsweise zu verweisen, kann man von einer *epistemischen Unbestimmtheit* sprechen; die Unbestimmtheit – die sich darin äußert, daß sich eine Vermutung als falsch herausstellen kann – verdankt sich bei dieser Betrachtungsweise einem Mangel an Wissen oder Fähigkeiten zur Wissensbildung.

7. Der Frage, warum das Seminar nicht stattgefunden hat, obwohl wir vermutet haben, daß es stattfinden wird, kann indessen noch eine andere Bedeutung gegeben werden. Wir können uns einen Prozeß vorstellen, in dessen Ablauf darüber „entschieden“ wird, ob das Seminar stattfinden wird oder nicht; und wir können fragen, wie dieser Prozeß abgelaufen ist und schließlich dazu geführt hat, daß das Seminar nicht stattgefunden hat.[1] Wir beziehen dann unser Reden von Gründen (oder Ursachen) nicht auf den epistemischen Prozeß, der zur Bildung unserer Vermutung geführt hat, sondern auf einen Prozeß, von dem wir annehmen, daß von seinem Ablauf abhängt, ob und ggf. wie ein Ereignis – in diesem Beispiel das fragliche Seminar – realisiert wird. Wir könnten dann vielleicht herausfinden, daß das Seminar nicht stattgefunden hat, weil der Seminarleiter tags zuvor krank geworden ist. Dann hätten wir einen Grund gefunden, mit dem erklärt werden könnte, warum sich unsere Vermutung, daß das Seminar stattfinden wird, als falsch herausgestellt hat.

8. Aber welcher Art ist dieser Grund? Er bezieht sich auf ein Ereignis, das eingetreten ist, nachdem wir unsere Vermutung gebildet haben. Hätten wir es voraussehen können? Vielleicht ja, vielleicht nein. Sicherlich treten oft Ereignisse ein, die wir nicht voraussehen und bei denen die Annahme, daß man sie hätte voraussehen können, nicht begründbar erscheint. Diese Überlegung betrifft indessen zunächst wieder nur eine epistemische Unbestimmtheit: es mangelt uns an Wissen, die Zukunft verläßlich vorauszusagen. Ein neuer Gesichtspunkt entsteht jedoch durch die Überlegung, daß Menschen auf zukünftiges Geschehen Einfluß

[1] Ich setze hier das Wort ‘entschieden’ in Anführungszeichen, weil es sich um einen metaphorischen Sprachgebrauch handelt. Es ist indessen an dieser Stelle nicht erforderlich, sich ein genaues Bild solcher Prozesse zu machen, durch die mögliche Seminare real werden oder unrealisiert bleiben.

nehmen können. Angenommen, wir begeben uns in der nächsten Woche in den Raum, in dem das Seminar stattfinden soll, und stellen dort fest, daß der Seminarleiter nicht anwesend ist, weil er kurz zuvor erkrankt ist. Gleichwohl könnten wir vorschlagen, daß das Seminar trotzdem – wenn auch ohne Anwesenheit des Seminarleiters – stattfinden soll. Ob das Seminar tatsächlich stattfindet, hängt dann davon ab, ob der Vorschlag von den übrigen Seminarteilnehmern angenommen wird.

9. Das Beispiel kann verallgemeinert werden. In vielen Fällen hängt das, was geschehen wird, auch davon ab, wie sich Akteure verhalten, die am Geschehen beteiligt sind. Aus der Sicht dieser Akteure stellt sich dann nicht nur die Frage, was geschehen wird, sondern auch: was geschehen sollte und wie sie vielleicht Einfluß nehmen können. Tatsächlich lassen sich in den meisten praktischen Handlungssituationen beide Fragestellungen nicht klar unterscheiden.[2]

10. Aber auch ohne selbst eine Akteursperspektive einzunehmen, kann ein Unterschied festgestellt werden. Wenn an einer Situation Akteure beteiligt sind, hängt die Entwicklung der Situation auch davon ab, wie sich die Akteure verhalten werden. Akteure können jedoch Entscheidungen treffen; und insoweit ihre Entscheidungen erst im Verlauf einer Situation getroffen werden, ist auch über die Entwicklung der Situation „noch nicht entschieden“ worden. Es erscheint deshalb sinnvoll, außer einer epistemischen Unbestimmtheit, die sich einem Mangel an Wissen verdankt, noch eine andere Art von Unbestimmtheit in Betracht zu ziehen, die daraus resultiert, daß Akteure Entscheidungen treffen und dadurch auf die Entwicklung einer Situation Einfluß nehmen können. Zur Unterscheidung von epistemischer Unbestimmtheit kann man von *ontologischer Unbestimmtheit* sprechen; denn in diesem Fall ist „noch nicht entschieden“, wie die Situation sein wird, auf die sich eine Vermutung ihrer Intention nach bezieht.

11. Ob bzw. in welcher Weise die Unterscheidung epistemischer und ontologischer Unbestimmtheit für die Reflexion von Vermutungen, die sich auf zukünftige Situationen beziehen, wichtig ist, ist eine schwierige Frage, und sie wird in der Literatur zur Wahrscheinlichkeitstheorie

[2] In einigen Fällen, wenn man sich unmittelbar auf eigene Handlungsmöglichkeiten bezieht, erscheint sogar die Frage, was geschehen *wird*, überhaupt nicht sinnvoll. Zum Beispiel bin ich in der Lage, mithilfe meiner Hände und eines Bleistifts Sätze auf ein Stück Papier zu schreiben, also die entsprechenden Ereignisse hervorzubringen. Aber es erscheint nicht sinnvoll, mich jetzt zurückzulehnen und Vermutungen darüber anzustellen, welche Sätze meine Hände in den nächsten 5 Minuten zu Papier bringen werden.

fast immer ausgeblendet. Vielmehr wird von der Annahme ausgegangen, daß die Situationen, auf die sich Vermutungen richten, als ontologisch *bestimmte* Situationen vorausgesetzt werden können und daß es infolgedessen nur eine epistemische Unbestimmtheit gibt, die in einem mangelhaften Wissen darüber besteht, wie die Situationen tatsächlich beschaffen sind.[3] In der weiteren Diskussion werden wir diese Annahme als eine Sinnvoraussetzung wahrscheinlichkeitstheoretischer Überlegungen zugrunde legen und nur gelegentlich auf daraus resultierende Sinngrenzen hinweisen.

2.2 Vermutungen und Hypothesen

1. Im allgemeinen Sprachgebrauch erscheint es nicht erforderlich, zwischen den Worten ‘Vermutung’ und ‘Hypothese’ zu unterscheiden. Man kann sie weitgehend synonym verwenden und ggf. durch zusätzliche Erläuterungen kenntlich machen, auf welche Art von Vermutungen oder Hypothesen man sich beziehen möchte und wie man dies tun möchte. Wir werden hier diesem allgemeinen Sprachgebrauch folgen und Hypothesen als Vermutungen auffassen. Durch einige ergänzende Bemerkungen soll indessen unser Sprachgebrauch etwas präzisiert werden.

2. Die erste Bemerkung betrifft die Formulierung von Hypothesen. Wenn wir von Hypothesen sprechen, meinen wir *explizit formulierte* Vermutungen. Dies ist erforderlich, weil bei wahrscheinlichkeitstheoretischen Überlegungen Hypothesen *als Gegenstand* der Theoriebildung auftreten. Wir reden in diesem Kontext nicht unmittelbar über unsere Erfahrungswelt, sondern *über Hypothesen*. Also muß explizit gemacht werden, worüber wir sprechen wollen und wie wir dies tun wollen.

3. Damit hängt ein zweiter Punkt unmittelbar zusammen. Eine Hypothese (Vermutung) kann falsch sein. Oder anders formuliert: es könnte sein, daß eine Hypothese falsch und stattdessen eine andere Hypothese richtig ist. Wenn man in der Wahrscheinlichkeitstheorie über Hypothe-

[3] Eine oft verwendete Formulierung für diese Sinnvoraussetzung wurde von Keynes in seinem *Treatise on Probability* (1921, S. 3) so ausgedrückt: „All propositions are true or false, but the knowledge we have of them depends on our circumstances.“ Ob eine solche Annahme ohne Einschränkungen auch für zukünftiges Geschehen sinnvoll ist, ist indessen seit Aristoteles’ berühmter Diskussion „zukünftiger Kontingenz“ fragwürdig. Das Gegenargument ist auch leicht einsehbar: Wenn heute noch nicht darüber „entschieden“ ist, ob das Seminar in der nächsten Woche stattfinden wird oder nicht, kann die Aussage, daß es stattfinden wird, nicht bereits *heute* wahr oder falsch sein.

sen spricht, ist es deshalb zweckmäßig, stets ein *vollständiges System von Hypothesen* in Betracht zu ziehen. Damit ist gemeint: eine Menge von zwei oder mehr Hypothesen, die sich gegenseitig ausschließen und zusammengenommen alle – zur Zeit der Hypothesenbildung – denkbaren Möglichkeiten ausschöpfen. Wenn man zunächst nur an eine Hypothese denkt, kann das dadurch erreicht werden, daß man außerdem ihre einfache Negation in Betracht zieht. Wird zum Beispiel als Hypothese formuliert, daß das oben erwähnte Seminar in der nächsten Woche stattfinden wird, wäre die zugehörige Alternativhypothese, daß es nicht stattfinden wird. Ein vollständiges System von Hypothesen kann indessen auch mehr als zwei Hypothesen umfassen. Bezieht man sich zum Beispiel auf die mit einem Würfel erzielbaren Augenzahlen, kann man mindestens sechs unterschiedliche Hypothesen bilden.

4. Wie Hypothesen gebildet werden, ist natürlich in gewissen Grenzen freigestellt. Auch bei der Bezugnahme auf die Ergebnisses des Wurfs mit einem Würfel könnte man nur zwei Hypothesen unterscheiden: daß eine gerade oder eine ungerade Augenzahl erscheint; und noch viele andere Möglichkeiten der Hypothesenbildung sind vorstellbar. In Abschnitt 3.2 wird dies durch die Konzeption einer „Algebra für Hypothsen“ präzisiert. Zunächst soll nur festgehalten werden, daß für das Reden von Hypothesen stets vorausgesetzt werden soll, daß man sich auf ein vollständiges System von Hypothesen beziehen kann.

5. Schließlich ist wichtig, daß deutlich gemacht wird, von welcher Art die Hypothesen sind, über die gesprochen werden soll. Wenn in der sozialwissenschaftlichen Methodenliteratur von Hypothesen gesprochen wird, sind meistens Vermutungen über „Zusammenhänge zwischen Variablen“ gemeint. Jedoch ist es nicht nur eine offene Frage, wie einem solchen Verständnis von Hypothesen ein nachvollziehbarer Sinn gegeben werden kann; viele sinnvoll formulierbare Hypothesen haben mit „Zusammenhängen zwischen Variablen“ überhaupt nichts zu tun. Es ist also in jedem Fall wichtig, sich zu überlegen, *worauf* sich Hypothesen beziehen sollen und *was* mit ihnen vermutet werden soll.

2.3 Hypothesen über Sachverhalte

1. In vielen Fällen beziehen sich Hypothesen auf Sachverhalte. Zum Beispiel kann sich eine Hypothese auf die Anzahl der Teilnehmer an einem Seminar richten, das in der letzten Woche stattgefunden hat. Wenn ich dabeigewesen wäre und die Anzahl der Teilnehmer festgestellt und

mir gemerkt hätte, könnte ich jetzt in der Form einer Feststellung sagen: es waren so ... viele. Aber wenn ich nicht dabeigewesen wäre oder nicht darauf geachtet hätte, wieviele Teilnehmer anwesend waren, oder einfach vergessen hätte, wieviele es gewesen sind, dann könnte ich jetzt bestenfalls eine Vermutung äußern. — Wenn sich Hypothesen auf Sachverhalte beziehen, kann man ein vollständiges System von Hypothesen im allgemeinen so formulieren: es könnte so oder so oder so ... (gewesen) sein.

2. Aber worüber reden wir, wenn wir Hypothesen über Sachverhalte formulieren? Die naheliegende Antwort, daß wir dann über den Sachverhalt reden, auf den wir uns mit unserer Hypothese beziehen, hat offenbar einen Mangel. Denn die Hypothese könnte ja falsch sein, und dann hat es den Sachverhalt nicht wirklich, sondern nur in unserer Einbildung gegeben. Sollte man infolgedessen sagen, daß sich Hypothesen stets nur auf Einbildungen beziehen? Damit würde man jedoch eine wesentliche Intention preisgeben: daß wir uns in vielen Fällen mit der Formulierung von Hypothesen darauf beziehen wollen, wie gewisse Aspekte unserer Erfahrungswelt tatsächlich beschaffen sind. Wir wollen dann nicht über Einbildungen, sondern über Aspekte unserer Erfahrungswelt sprechen; und wir sind nur deshalb gezwungen, dies in der sprachlichen Form von Vermutungen zu tun, weil das vorhandene Wissen nicht ausreicht, um Feststellungen treffen zu können.

3. Es gibt jedoch eine Möglichkeit, das angedeutete Dilemma zu vermeiden. Denken wir noch einmal an die Vermutung über die Anzahl der Teilnehmer des Seminars, das in der letzten Woche stattgefunden hat. Die Vermutung bezieht sich in diesem Fall auf eine Eigenschaft des Seminars. Daß das Seminar stattgefunden hat und daß es eine bestimmte Anzahl von Teilnehmern gegeben hat, wird für die Formulierung der Hypothese vorausgesetzt. Die Hypothese bezieht sich nicht darauf, ob das Seminar überhaupt stattgefunden hat, sondern auf die Anzahl der Teilnehmer. Wir reden also in diesem Fall über eine Situation, von der wir bei der Formulierung der Hypothese annehmen, daß es sie tatsächlich gegeben hat. Fraglich sind nur gewisse Aspekte oder Eigenschaften der Situation – und darauf beziehen sich die Hypothesen.

4. Dieses Beispiel kann verwendet werden, um eine für Hypothesen über Sachverhalte geeignete allgemeine sprachliche Form zu finden. Solche Hypothesen beziehen sich auf Situationen, von denen wir annehmen und bei der Formulierung der Hypothesen voraussetzen, daß es sie in unserer Erfahrungswelt gibt oder gegeben hat. Fraglich sind jedoch bestimmte Aspekte oder Eigenschaften der Situation: es könnte so oder so oder so (gewesen) sein. Oder anders formuliert: fraglich ist, welche aus einer Reihe möglicher Kennzeichnungen für die Situation zutreffend ist. Eine Hypothese besteht bei dieser Betrachtungsweise in der Vermutung, daß für die Situation, auf die man sich bezieht, eine gewisse Kennzeichnung zutreffend ist. Diese Betrachtungsweise liefert auch eine sinnvolle Explikation für unser Reden von Sachverhalten, nämlich: ein Sachverhalt besteht in der *zutreffenden* Kennzeichnung einer Situation. Zum Beispiel: An dem Seminar haben ... Personen teilgenommen. Dies ist eine Aussage über einen Sachverhalt. Es ist auch eine Aussage über eine Situation. Aber die Situation – das Seminar, auf das wir uns beziehen – kann natürlich noch auf viele andere Weisen beschrieben werden. Insofern können Sachverhalte und Situationen nicht umstandslos identifiziert werden. Man kann aber sinnvoll davon sprechen, daß Sachverhalte in jeweils spezifischen Aspekten einer Situation bestehen.

5. Diese Überlegungen erlauben es, die sprachliche Form von Hypothesen über Sachverhalte auch symbolisch zu vergegenwärtigen. Es gibt zunächst eine Situation s; durch sie wird angegeben, worauf sich die Hypothesen beziehen. Fraglich ist, welche aus einer Reihe möglicher Kennzeichnungen für die Situation zutrifft. Hypothesenbildung besteht also zunächst darin, eine Reihe möglicher Kennzeichnungen für die Situation s zu fixieren. Wir verwenden dafür die Formulierung[4]

$$\tilde{K} := \{\tilde{k}_1, \ldots, \tilde{k}_m\}$$

Jedes Element von $\tilde{K}$ kann verwendet werden, um – in der Form einer Vermutung – die Situation s zu charakterisieren. Jedem $\tilde{k} \in \tilde{K}$ entspricht somit eine Vermutung

$$\langle\, s, \tilde{k}\,\rangle$$

mit der Bedeutung: die Kennzeichnung von s durch $\tilde{k}$ ist ein *möglicher* Sachverhalt; und es wäre tatsächlich ein Sachverhalt, wenn die Vermutung zutreffen würde. Auch wenn wir nicht oder nicht genau wissen, ob ein Sachverhalt zutrifft, sprechen wir also nicht über beliebige Einbildungen, sondern über eine Situation, von der wir wissen oder annehmen, daß es sie gibt oder gegeben hat.

6. Schließlich kann durch die Verwendung dieser Schreibweisen auch leicht ein vollständiges System von Hypothesen angegeben werden. Man

[4]Das Zeichen ':=' wird verwendet, um zum Ausdruck zu bringen, daß es sich um eine Definition handelt; der Ausdruck auf der linken Seite des Zeichens wird durch den Ausdruck auf der rechten Seite des Zeichens definiert.

muß nur bei der Bildung von $\tilde{K}$ darauf achten, daß zwei Bedingungen erfüllt sind. Erstens, daß sich die Kennzeichnungen der Situation s durch die Elemente von $\tilde{K}$ wechselseitig ausschließen; und zweitens, daß die in $\tilde{K}$ enthaltenen Kennzeichnungen vollständig sind, so daß genau ein Element eine zutreffende Kennzeichnung der Situation liefern kann. Ein vollständiges System von Hypothesen kann dann durch die Menge

$$\mathcal{H} := \{\, \langle\, s, \tilde{k}\,\rangle \mid \tilde{k} \in \tilde{K}\,\}$$

angegeben werden. Mit dem Symbol $\mathcal{H}$ soll auch im weiteren stets auf ein vollständiges System von Hypothesen verwiesen werden.

2.4 Situationen und Ereignisse

1. Bisher wurde das Wort 'Sachverhalt' verwendet, um eine Variante von Hypothesen – Hypothesen über Sachverhalte – zu erläutern. Sowohl in der Umgangssprache als auch in Texten zur Wahrscheinlichkeitstheorie wird dagegen oft von Ereignissen gesprochen. In der Literatur zur Wahrscheinlichkeitstheorie wird das Wort 'Ereignis' meistens in einer weitestgehend unspezifischen Weise verwendet, oft auch austauschbar mit dem Wort 'Sachverhalt'.[5] In der Umgangssprache verwenden wir dagegen die Worte 'Sachverhalt' und 'Ereignis' in unterschiedlichen Bedeutungen. Mit dem Wort 'Sachverhalt' verweisen wir auf Eigenschaften von Situationen (oder Objekten); mit dem Wort 'Ereignis' verbinden wir dagegen meistens die Vorstellung eines „Geschehens" oder eines „Vorkommnisses".[6] Zum Beispiel: daß an unserem Seminar heute 15 Personen teilnehmen, ist ein Sachverhalt, der einen Aspekt der Seminarsituation fixiert; wir würden in diesem Fall nicht von einem Ereignis sprechen. Andererseits kann man sagen, daß das heutige Stattfinden des Seminars ein Ereignis ist; und man kann sich auch auf eine Vielzahl von Ereignissen beziehen, die während des Seminars auftreten.

[5]Hier ist ein Beispiel (Applebaum 1996, S. 43): „Suppose that you are going to buy a second-hand car. You find one that you like and after a brief inspection you convince yourself that it would have a high probability of being roadworthy. Being a cautious individual you then inspect the car again in the company of a friend, who is a skilled mechanic and who discovers a large number of major problems. Consequently, you decrease your personal probability of the car begin roadworthy. In the above example we have seen an important general principle at work: namely, *Gain in information about an event leads to a change in its probability.*" Folgt man der Umgangssprache, ist allerdings der Zustand eines Gebrauchtwagens kein Ereignis.

[6]Zu diesem Verständnis des Ereignisbegriffs vgl. man Hacker, 1982.

2. Einige Aspekte des Redens von Ereignissen können jedoch auch mit der im vorangegangenen Abschnitt entwickelten sprachlichen Form für Hypothesen über Sachverhalte erfaßt werden. Denn in vielen Fällen erscheint die Vorstellung sinnvoll, daß Ereignisse in Situationen stattfinden. Wir befinden uns in einem Seminar, und dann meldet sich ein Teilnehmer zu Wort oder ein anderer verläßt den Raum. Oder wir unternehmen einen Spaziergang, und dann beginnt es zu regnen. Diese und ähnliche Beispiele legen es nahe, daß man Situationen sowohl durch Sachverhalte (Zustände) als auch durch Vorkommnisse (Ereignisse) charakterisieren kann und daß es sich in gewisser Weise um komplementäre Arten der Charakterisierung von Situationen handelt.[7]

3. Soweit dies komplementäre Verständnis von Sachverhalten und Ereignissen sinnvoll erscheint, kann man das im vorangegangenen Abschnitt entwickelte Schema für Hypothesen über Sachverhalte auch zur Formulierung von Hypothesen über Ereignisse verwenden. Es ist nur erforderlich, die Kennzeichnungsmenge $\tilde{K}$ als eine Menge möglicher Ereignisse zu konzipieren, die in der Situation s vorgefallen sein können und die mithin verwendet werden können, um Vermutungen über die Situation zum Ausdruck zu bringen. Wir werden deshalb im weiteren bei der Bezugnahme auf Hypothesen nicht unterscheiden, ob sich die Elemente von $\tilde{K}$ auf mögliche Sachverhalte oder auf mögliche Ereignisse beziehen.

4. Das Reden von möglichen Ereignissen bedarf jedoch noch einer Erläuterung. Angenommen, wir beziehen uns auf eine Situation s und fixieren mit einer Kennzeichnungsmenge $\tilde{K}$ mögliche Ereignisse, die in der Situation s stattgefunden haben können. Man sollte sich klarmachen, daß es sich bei den Elementen von $\tilde{K}$ selbst nicht um mögliche Ereignisse handelt, vielmehr um sprachliche Formulierungen, mit denen Ereignisse, die in der Situation s aufgetreten sein könnten, charakterisiert werden sollen. Um auf die möglichen Ereignisse zu verweisen, muß dagegen auf Vorkommnisse in der Situation s und mithin zunächst auf die Situation selbst bezug genommen werden. Für die möglichen Ereignisse selbst verwenden wir deshalb die sprachliche Form $\langle s, \tilde{k} \rangle$, also unsere Sprachform für Hypothesen. Als Hypothese verweist diese sprachliche Form auf ein mögliches Ereignis. Ist die Hypothese zutreffend, kann man von einem tatsächlich realisierten Ereignis sprechen. Insofern gibt es auch eine Parallele zum Reden von möglichen und realisierten Sachverhalten.

[7] Zu dieser Komplementarität des Redens von Zuständen und Ereignissen vgl. man Galton 1984, insb. Kap. 2.

2.5 Vergangenheit und Zukunft

1. Bei den bisherigen Überlegungen zur Formulierung von Hypothesen sind wir davon ausgegangen, daß sich Hypothesen auf Situationen beziehen und daß angenommen werden kann, daß es diese Situationen in unserer Erfahrungswelt gibt oder gegeben hat. Somit konnte auch angenommen werden, daß diese Situationen in einer jeweils bestimmten Weise gegeben sind – sei es durch eigene Erfahrungen oder durch Nachrichten – und daß man nur deshalb Hypothesen bilden muß, weil das Wissen über die Situationen unvollständig ist. Andererseits kann man sich mit Hypothesen auch auf zukünftiges Geschehen beziehen.

2. Ein wichtiger Unterschied ist unmittelbar einsichtig. Von gegenwärtigen oder vergangenen Situationen kann man sagen, daß es sie gibt oder gegeben hat. Beim Reden über zukünftiges Geschehen kann man sich dagegen nicht auf Situationen beziehen, die es bereits gibt oder gegeben hat. In gewisser Weise gibt es die zukünftigen Situationen noch nicht; insbesondere können zukünftige Situationen nicht beschrieben werden, und es ist auch schwer vorstellbar, in welcher Weise es heute bereits Nachrichten oder Berichte über zukünftige Situationen geben könnte. Dennoch kann man sich auch über zukünftiges Geschehen Gedanken machen. Man kann sich *Vorstellungen* bilden und sich wünschen oder glauben oder annehmen oder vermuten oder erwarten oder davon überzeugt sein, daß es so kommen wird, wie man es sich vorstellbar gemacht hat.[8]

3. An dieser Stelle erscheint es sinnvoll, zwei Reflexionsformen zu unterscheiden, in denen wir uns auf zukünftiges Geschehen beziehen können. Die eine Reflexionsform orientiert sich an Fragestellungen der Art: Was wird geschehen? Oder auch konditional: Angenommen, es ist eine Situation s mit den Eigenschaften ... entstanden, wird dann auch eine Situation s' mit den Eigenschaften ... entstehen? Ich nenne dies eine *prozessuale Reflexionsform* über zukünftiges Geschehen. Es ist einsichtig, daß man in dieser Reflexionsform auch über gegenwärtiges und vergangenes Geschehen nachdenken kann. Zwar sind gegenwärtige und vergangene Situationen bereits in jeweils bestimmter Weise real geworden und können nachträglich nicht mehr verändert werden. Trotzdem

[8] Es sei indessen angemerkt, daß das Wort ‘Vorstellungen’ die Unterscheidung zwischen Vergangenheit und Zukunft, wie sie sich aus menschlicher Perspektive darstellt, nicht wirklich trifft. Denn auch über vergangene Situationen kann man sich heute bestenfalls noch Vorstellungen machen; und auch das, was gerade jetzt im Nebenzimmer vor sich geht, kann ich mir nur vorstellen.

kann man in kontrafaktischen Reflexionsformen auch noch über vergangenes und gegenwärtiges Geschehen nachdenken; und solche kontrafaktischen Überlegungen spielen auch für unser prozessuales Nachdenken über zukünftiges Geschehen eine wichtige Rolle.[9]

4. Man kann sich indessen auch in einer nicht-prozessualen Reflexionsform auf zukünftiges Geschehen beziehen. Als Beispiel kann das Seminar dienen, daß in der nächsten Woche stattfinden wird. Ob es stattfinden wird, kann man heute noch nicht sicher wissen, denn es ist „noch nicht entschieden"; insofern könnte man in einer prozessualen Reflexionsform darüber nachdenken, ob es stattfinden wird oder nicht; d.h. man kann über einen Prozeß nachdenken, in dessen Verlauf darüber „entschieden" wird, ob das Seminar stattfinden wird oder nicht. Andererseits können wir vermuten, daß es stattfinden wird, und uns mit einer anderen Frage beschäftigen: wieviele Teilnehmer es beim Seminar in der nächsten Woche geben wird. Offenbar unterscheiden sich die beiden Fragestellungen. Im ersten Fall bezieht sich die Frage darauf, ob ein gewisses Ereignis in der nächsten Woche stattfinden wird. Im zweiten Fall setzen wir voraus, daß eine Situation – in diesem Beispiel das Seminar – gegeben ist, daß wir jedoch bestimmte Aspekte oder Eigenschaften der Situation nicht kennen und (auch) darüber nur Vermutungen anstellen können. Bei der Voraussetzung handelt es sich um eine Annahme, die der Bildung von Hypothesen vorausgesetzt wird. Natürlich entsteht durch diese Annahme nicht die zukünftige Situation, auf die sich die Hypothesen beziehen. Wenn wir zur Formulierung der Frage nach der Anzahl der Teilnehmer am Seminar, das in der nächsten Woche stattfinden soll, voraussetzen, daß das Seminar stattfinden wird, impliziert dies nicht, daß das Seminar in der nächsten Woche tatsächlich stattfinden wird. Das Seminar, das in der nächsten Woche *vielleicht* stattfinden wird, können wir uns heute nur in Form einer vorstellbaren Situation vergegenwärtigen. Der wichtige Punkt ist jedoch, daß wir Hypothesen über zukünftiges Geschehen auch in einer sprachlichen Form bilden können, bei der sich die Hypothesen auf zukünftige Situationen beziehen; und zwar so, daß in der sprachlichen Formulierung die ontologische Differenz zwischen Vergangenheit und Zukunft unsichtbar bleiben kann.

5. Die Unterscheidung zwischen prozessualen und nicht-prozessualen Reflexionsformen zukünftigen Geschehens ist hier wichtig, weil sich dementsprechend die sprachlichen Formulierungen für Hypothesen unterscheiden. Mit dem bisher für Hypothesen über Sachverhalte und Ereignisse entwickelten sprachlichen Schema können Vermutungen, die im Rahmen einer prozessualen Reflexionsform sinnvoll erscheinen, nicht angemessen formuliert werden. Es kann aber durchaus verwendet werden, um nicht-prozessuale Hypothesen auszudrücken. Es ist dafür nur erforderlich, sich die Situation s, auf die sich die Hypothesen beziehen, als eine zukünftige Situation vorzustellen. Um bei unserem Beispiel zu bleiben, würde dann s auf die Situation verweisen, die durch das Seminar entsteht, wenn es in der nächsten Woche stattfindet; und die Kennzeichnungsmenge $\tilde{K}$ könnte als eine Menge von Zahlen konzipiert werden, mit denen Hypothesen über die Anzahl der Teilnehmer formuliert werden können.

[9]Eine allgemeine Einführung in diese Reflexionsform findet sich bei Rescher, 1996.

6. Wenn man sich mit den begrifflichen Hilfsmitteln der Wahrscheinlichkeitstheorie auf zukünftiges Geschehen bezieht, wird stets eine nicht-prozessuale Betrachtungsweise angenommen. (Dies entspricht der in Abschnitt 2.1 besprochenen „Bestimmtheitsannahme" für Situationen.) Zugrunde liegt die Vorstellung, daß man sich auf Zukünftiges in der gleichen sprachlichen Form beziehen kann wie auf Gegenwärtiges und Vergangenes. Es gibt dann nur unterschiedliche Grade epistemischer Unbestimmtheit; oder etwas krasser formuliert: aus dieser Sicht besteht der einzige für die Theoriebildung relevante Unterschied darin, daß man über die Zukunft meistens schlechter informiert ist als über die Vergangenheit.[10] Ich betone diesen Punkt, weil die in der Literatur zur Wahrscheinlichkeitstheorie verwendete Rhetorik oft suggeriert, daß man sich mit dem Eintreten von Ereignissen beschäftigt; zum Beispiel wird von „Eintrittswahrscheinlichkeiten für Ereignissse" gesprochen. Dementsprechend werden oft Formulierungen verwendet, die eine prozessuale Reflexionsform nahelegen. Sobald man jedoch die Begriffsbildungen genauer analysiert, stellt man fest, daß es sich um eine irreführende Rhetorik handelt.

7. Um das Argument zu verdeutlichen, kann ein Standardbeispiel aus der Wahrscheinlichkeitstheorie dienen: würfeln. Hat man einen Würfel, kann man ihn verwenden, um Ereignisse zu erzeugen. Ein solches Ereignis besteht darin: jemand nimmt einen Würfel und wirft ihn und wartet dann, bis der Würfel in einer bestimmten Lage zur Ruhe gekommen ist. Von Ereignissen *dieser Art* wird jedoch in der Wahrscheinlichkeitstheorie

[10]Zur Begründung wird manchmal auf physikalische Theorien verwiesen, in denen eine Unterscheidung zwischen Vergangenheit und Zukunft nicht vorkommt; vgl. zum Beispiel die Ausführungen von Smart 1963, Kap. 7. Daß eine solche Unterscheidung auch grundsätzlich nicht erforderlich sei, wird in der Philosophie seit einiger Zeit unter der Überschrift „the new theory of time" diskutiert; eine gute Einführung geben die Diskussionsbeiträge in Oaklander und Smith, 1994.

gar nicht gesprochen. Wenn man sich dort auf die „möglichen Ereignisse“ bezieht, die beim Werfen eines Würfels auftreten können, meint man vielmehr unterschiedliche Kennzeichnungen einer Situation, die zuvor – durch das Werfen des Würfels – erzeugt worden ist. Man bezieht sich auf Hypothesen der Form $\langle s, \tilde{k} \rangle$, deren Formulierung bereits voraussetzt, daß es eine Situation s gibt, sei es als eine reale oder als eine vorstellbare Situation. Ob und ggf. wie es in unserer Erfahrungswelt zu einer solchen Situation gekommen ist oder kommen wird, tritt als Fragestellung in der (elementaren) Wahrscheinlichkeitstheorie nicht auf.

2.6 Statische und dynamische Hypothesen

1. Man kann einwenden, daß es in der Wahrscheinlichkeitstheorie und Statistik durchaus Begriffsbildungen und Überlegungen gibt, um auch über *Prozesse* sprechen zu können. Es liegt also nahe, eine Unterscheidung zwischen statischen und dynamischen Hypothesen einzuführen, etwa so: *statische* Hypothesen beziehen sich auf die Beschaffenheit einer Situation, dagegen beziehen sich *dynamische* Hypothesen auf den Ablauf eines Prozesses.

2. Orientiert man sich an einer solchen Unterscheidung, handelt es sich bei den in den vorangegangenen Abschnitten besprochenen Hypothesen um statische Hypothesen. Denn sie beziehen sich auf die Frage, wie eine als gegeben voraussetzbare Situation durch Sachverhalte oder Ereignisse (die in der Situation stattgefunden haben oder stattfinden könnten), charakterisiert werden kann. Hypothesenbildung dieser Art setzt voraus, daß von den prozeßhaften Aspekten von Situationen abstrahiert wird. Es handelt sich um eine Abstraktion, denn alle Situationen, über die wir in unserer Erfahrungswelt sprechen können, verändern sich mehr oder weniger schnell. Diese Fähigkeit, im historischen Ablauf des Geschehens Situationen fixieren und durch Sachverhalte und Vorkommnisse charakterisieren zu können, ist jedoch für unser Nachdenken und Kommunizieren konstitutiv; insofern handelt es sich nicht um eine beliebige Abstraktion.

3. Andererseits ist es auch eine schwierige Frage, wie ein geeigneter Begriffsrahmen für das Reden von Prozessen begründet werden kann. Zwar stellt uns unsere Sprache eine Fülle von Möglichkeiten zur Verfügung, um auf Prozesse hinzuweisen und sie zu charakterisieren. Diese sprachlichen Möglichkeiten sind jedoch durchgängig mit einer „prozessualen Reflexionsform“ vergangenen, gegenwärtigen und zukünftigen Geschehens verbunden; und eine solche Reflexionsform kann nicht bruchlos mit einer Wissenschaftssprache verknüpft werden, die sich zunächst und prototypisch an Sachverhalten orientiert. Eine weitere Schwierigkeit resultiert daraus, daß das umgangssprachliche Reden über Prozesse stets auch mit Vorstellungen über Akteure verbunden ist, die gewissermaßen als „Generatoren“ für die sich in Prozessen ausdrückenden Veränderungen und Ereignisabfolgen angenommen werden. Die Entwicklung der neuzeitlichen (Natur-) Wissenschaft hat sich demgegenüber vor allem darum bemüht, alle Anklänge an ein „animistisches“ Weltbild, wie es unserer Lebenswelt und ihrer Umgangssprache zugrunde liegt, so weit wie möglich aus dem Blickfeld geraten zu lassen.

4. Die angedeuteten Schwierigkeiten beim Reden über Prozesse sollen an dieser Stelle nicht näher besprochen werden. In der weiteren Diskussion beziehen wir uns nur auf statische Hypothesen, die sich Fragestellungen nach der Beschaffenheit von Situationen verdanken. Mit der schwierigen Frage, wie man sich sprachlich auf Prozesse beziehen kann, werden wir uns in diesem Skript nicht näher beschäftigen.

2.7 Hypothesen über Dispositionen

1. Statische Hypothesen über die Beschaffenheit von Situationen sollten von Hypothesen über Dispositionen unterschieden werden, die sich darauf beziehen, wie sich Objekte verhalten können oder sich unter gewissen Umständen verhalten werden. Denn bei Dispositionen handelt es sich nicht um Sachverhalte, die als Eigenschaften von Situationen oder Objekten expliziert werden können. Zwar bezieht man sich beim Reden über Dispositionen oft auf Situationen, aber nur um dadurch auf einen Kontext zu verweisen, in dem sich Dispositionen in der Gestalt von Verhaltensweisen von Objekten zeigen können. Dispositionale Eigenschaften dienen zur Charakterisierung von Objekten, die sich in Situationen unterschiedlich verhalten könnnen. Demgegenüber erscheint es als ein metaphorischer Sprachgebrauch, wenn auch Situationen selbst gelegentlich mit dispositionalen Formulierungen charakterisiert werden; zum Beispiel, wenn davon gesprochen wird, daß ein Entwicklungsprozeß in einer Situation „angelegt“ ist.

2. Der Wunsch, Wissen über Dispositionen von Objekten zu gewinnen, ist sicherlich ein zentrales Motiv der neuzeitlichen Wissenschaft. Dem entspricht die Bedeutung, die Experimenten – im Unterschied zu „bloßen Beobachtungen“ – als Instrument zur Ermittlung von Einsichten in dispositionale Eigenschaften gegeben wird. d'Alembert hat in der zusammen mit Diderot herausgegebenen „Enzyklopädie“ (1756, S. 434) den Unterschied so formuliert:

„Die Beobachtung, die weniger originell und tiefgründig ist, beschränkt sich auf die Tatsachen, die man vor Augen hat, das heißt darauf, Erscheinungen aller Art, die uns das Schauspiel der Natur darbietet, gut zu betrachten und ausführlich zu beschreiben; dagegen sucht das Experiment die Natur tiefer zu erforschen, ihr das zu entreißen, was sie verbirgt, und durch mannigfache Kombination der Körper neue Erscheinungen hervorzubringen, um diese wiederum zu studieren – kurz, es beschränkt sich nicht darauf, die Natur zu belauschen, sondern es befragt sie und zwingt sie zur Auskunft.“

Das Motiv ist verständlich: erst Wissen über dispositionale Eigenschaften von Objekten erlaubt es dem Menschen, sie zu nutzen und ggf. für seine Zwecke zu verändern. So betrachtet – und dies war auch die zentrale Botschaft Francis Bacons – besteht ein enger Zusammenhang zwischen der wissenschaftlichen Konzeption von Dispositionen und der Frage, was man mit Objekten aller Art machen kann.[11]

3. Diese Bemerkungen dienen hier zunächst nur dazu, um Hypothesen über die Beschaffenheit von Situationen von Hypothesen über Dispositionen abzugrenzen. Im Vorgriff auf spätere Überlegungen kann jedoch an dieser Stelle schon bemerkt werden, daß man viele Begriffs- und Modellbildungen der Wahrscheinlichkeitstheorie und Statistik nur verstehen kann, wenn man sie auf die Absicht bezieht, Wissen über dispositionale Eigenschaften zu gewinnen. Einen Hinweis darauf liefert bereits die Tatsache, daß die elementaren Begriffsbildungen oft ausgehend von einer primären Konzeption von (Zufalls-) Experimenten eingeführt werden. Z.B. heißt es in einem Lehrbuch von McPherson (1990, S. 3):

“Those scientific investigations in which Statistics might be employed are characterized by the fact that (i) they generate *data*, and (ii) there is an *experimental set-up* which is the source of the data. Scientists construct *models* of the experimental set-up and use the data to accept or modify those models.”

[11] Fragestellungen, die sich auf Dispositionen beziehen, kann man im Hinblick auf Objekte aller Art formulieren. Insbesondere für unser Verständnis von Akteuren erscheinen sie grundlegend. Allerdings gibt es in diesem Fall eine Besonderheit, die der Logik technisch konzipierter Experimente entgegen steht: daß sich Akteure auf jedenfalls einige ihrer dispositionalen Eigenschaften reflexiv (nicht unbedingt bewußt) beziehen können und demzufolge auch lernfähig sind. Dispositionale Eigenschaften dieser Art könnten „Fähigkeiten“ genannt werden; sie entziehen sich einem technisch konzipierten Stimulus-Response-Schema.

Allerdings wird oft kein klarer Unterschied zwischen Beobachtungen und Experimenten gemacht. Die Ausführungen McPhersons (1990, S. 3) gehen so weiter:

“For statistical purposes there are two descriptions of experimental set-ups which have common application.

1. *Populations.* The first, which is most useful in areas of biological and social sciences, views the object of investigation as a *population* or collection of objects from which a *sample* of population members has been drawn for examination in the investigation. This is the view which is useful, for example, in surveys which seek information on the collective opinion of people in a society.

2. *Physical Processes.* The second useful description of an experimental set-up is of a physical process which successively generates outcomes. A favorite academic example is the tossing of a coin where the outcome is the appearance of a head or a tail showing uppermost on the coin. In economic terms, the daily rates of exchange of a currency may be viewed as the outcomes from the complex process which is determined by a country's economic and political management.”

Ob und ggf. in welcher Weise Daten, die aus Umfragen gewonnen werden, als Ergebnisse von Experimenten verstanden werden können, ist sicherlich nicht evident; und gleichermaßen fragwürdig erscheint der Vorschlag, soziale Sachverhalte oder Prozesse als durch ein „experimental set-up“ erzeugt zu konzipieren.

2.8 Situationen und Situationstypen

1. In den bisherigen Überlegungen haben sich Hypothesen auf jeweils bestimmte Situationen bezogen, die sich in unserer Erfahrungswelt identifizieren lassen oder die man sich als bestimmte Situationen vorstellen kann. In den Sozialwissenschaften und insbesondere in ihren statistisch orientierten Varianten beziehen sich Hypothesen dagegen oft auf eine Mehrzahl von Situationen. Wie kann man das genauer verstehen?

2. Eine Möglichkeit besteht darin, eine Menge von Situationen zu vergegenwärtigen, deren Existenz durch Verweise auf unsere vergangene oder gegenwärtige Erfahrungswelt empirisch expliziert werden kann. Mengen dieser Art sind infolgedessen stets endlich, und man kann die Notation

$$\mathrm{S} := \{s_1, \ldots, s_N\}$$

verwenden, um auf eine solche Menge zu verweisen. Zum Beispiel kann es sich um die Gesamtheit der Statistik-Seminare handeln, die in einem bestimmten Zeitraum durchgeführt worden sind; oder um die Lebenssituationen einer Menge von Menschen, die in einem bestimmten Zeitraum arbeitslos geworden sind.

3. Andererseits erlaubt uns unsere Sprache, auch von *Situationstypen* zu sprechen; zum Beispiel von Situationen des Typs „Seminarsituation", des Typs „Fußballspiel", des Typs „Arbeitslosigkeit", usw. Diese sprachliche Möglichkeit wird auch meistens genutzt, um bestimmte Mengen von Situationen zu konzipieren, so daß sie als Beispiele eines Situationstyps verglichen werden können. Zur Notation wird in diesem Text die Schreibweise

$$s \models \mathcal{S}$$

verwendet, womit gemeint sein soll, daß die Situation s vom Typ $\mathcal{S}$ ist; oder anders formuliert: s ist ein Beispiel für Situationen des Typs $\mathcal{S}$. Wichtig ist, daß zwar s auf eine bestimmte (reale oder vorstellbare) Situation verweist, nicht jedoch $\mathcal{S}$. Situationstypen sind keine Situationen, sondern sprachliche Konstrukte, um Situationen als Beispiele eines Typs zu charakterisieren.

4. Situationstypen müssen auch von Mengen von Situationen unterschieden werden, und wir verwenden deshalb unterschiedliche Symbole. Das Symbol S wird verwendet, um auf eine jeweils bestimmte endliche Menge von Situationen zu verweisen. Dagegen wird das Symbol $\mathcal{S}$ verwendet, um auf einen Situationstyp zu verweisen.

5. Die Konzeption von Situationstypen scheint nun noch eine andere Möglichkeit zu liefern, um sich auf eine Mehrzahl von Situationen zu beziehen, nämlich: auf *alle Situationen eines Typs* $\mathcal{S}$. Zum Beispiel könnte man sagen: für alle Situationen des Typs „Fußballspiel" gilt, daß es zwei Mannschaften und einen Ball gibt. Diese Aussage erscheint durchaus sinnvoll, aber es handelt sich nur scheinbar um eine Aussage über alle Fußballspiele. Vielmehr betrifft die Aussage unseren Sprachgebrauch. Sie sagt etwas über unser Verständnis des Wortes 'Fußballspiel'. Finden wir zum Beispiel eine Situation, die in den meisten Aspekten als ein Fußballspiel erscheint, in der es jedoch zwei Bälle gibt, würde dadurch die Aussage, daß es in Fußballspielen nur einen Ball gibt, nicht falsch. Die Schlußfolgerung wäre vielmehr, daß es sich bei der von uns beobachteten Situation nicht um ein Fußballspiel im gewöhnlichen Sinn des Wortes handelt; aber vielleicht um eine Situation, die einem Fußballspiel sehr ähnlich ist.

6. Wie Aussagen über alle Situationen eines gewissen Typs formuliert werden können, die nicht nur darin bestehen, einen Sprachgebrauch

zu fixieren, ist sehr unklar. Die Schwierigkeiten beginnen bereits bei der Suche nach einer verständlichen Sprachform. In der Literatur findet man oft Formulierungen, die unterstellen, daß sinnvoll von einer *Menge* (oder *Klasse*) aller Situationen eines Typs $\mathcal{S}$ gesprochen werden kann. Solchen Redeweisen kann aber keine empirisch explizierbare Bedeutung gegeben werden. Nicht nur gibt es – zum Beispiel – die Menge aller Fußballspiele nicht, man kann sich eine solche Menge auch nicht in irgendeiner bestimmten Weise vorstellen. Man denke zum Beispiel an die Frage: Gehören die Fußballspiele, die vielleicht in der Zukunft noch stattfinden werden, zu dieser Menge? Natürlich kann man sich *einige* Fußballspiele vorstellen, die in der Zukunft stattfinden werden. Aber was ist mit denjenigen, die man sich heute noch nicht vorstellen kann? Und in welcher Weise könnte man sagen, daß es alle in der Zukunft noch möglichen Fußballspiele schon heute als Mitglieder einer Menge aller Fußballspiele gibt?

7. In der Mathematik kann man sprachliche Probleme dieser Art dadurch vermeiden, daß man unendliche Mengen durch Konstruktionsverfahren einführt; zum Beispiel die Menge aller natürlichen Zahlen. Aussagen über alle natürlichen Zahlen können dann als Aussagen darüber verstanden werden, was mit dem Konstruktionsverfahren erzeugt werden kann. Durch Rückgriff auf das Konstruktionsverfahren kann ihnen nicht nur ein bestimmter Sinn gegeben werden, sondern man kann sie auch beweisen, in diesem Fall mit der Methode der vollständigen Induktion. Aber um von „allen Situationen eines bestimmten Typs" zu sprechen, steht uns kein (Konstruktions-) Verfahren zur Verfügung. Es gibt zum Beispiel kein Verfahren, mit dem sich alle möglichen Fußballspiele – zumindest als Vorstellungen – erzeugen ließen. Die Methode der Mathematiker, Mengen durch Konstruktionsverfahren einzuführen, kann für das Reden über unsere Erfahrungswelt nicht sinnvoll nachgeahmt werden.

8. In diesem Text wird deshalb die Vorstellung, daß Situationstypen als Mengen von Situationen aufgefaßt werden können, nicht verwendet. Dementsprechend wird auch nicht die Notation „$s \in \mathcal{S}$", sondern stattdessen „$s \models \mathcal{S}$" verwendet, um zum Ausdurck zu bringen, daß die Situation s ein Beispiel für den Situationstyp $\mathcal{S}$ ist oder sein soll.

2.9 Definite und indefinite Hypothesen

1. Die Bildung von Hypothesen kann sich zunächst auf eine bestimmte endliche Menge von Situationen

$$\mathrm{S} := \{s_1, \dots, s_N\}$$

beziehen. Ich spreche dann von *definiten* Hypothesen, weil man in diesem Fall die Situationen, auf die sich die Hypothesen beziehen sollen, angeben kann, sei es als eine bestimmte Menge realer oder als eine bestimmte Menge vorstellbarer Situationen. Insbesondere sind natürlich singuläre Hypothesen, die sich auf *eine* bestimmte Situation beziehen, auch definite Hypothesen.

2. Wenn es in der Menge S zwei oder mehr Situationen gibt, kann man zwei unterschiedliche Arten der Hypothesenbildung in Betracht ziehen. Einerseits kann man an Hypothesen denken, die sich als eine Konjunktion von singulären Hypothesen über die jeweils einzelnen Situationen in S formulieren lassen. Andererseits gibt es *statistische* Hypothesen, die sich nicht auf singuläre Hypothesen reduzieren lassen, sondern nur als Aussagen *über die Gesamtheit* der einzelnen Situationen verstanden werden können. Dennoch kann man auch in diesem Fall von definiten Hypothesen sprechen, weil bzw. wenn sie sich auf eine bestimmte Menge von Situationen beziehen.

3. Statistische Hypothesen beziehen sich auf Merkmalsverteilungen bei einer Menge von Situationen (oder Objekten). Zunächst werden statistische Variablen als Abbildungen

$$X : \mathrm{S} \longrightarrow \tilde{X}$$

definiert, die jeder Situation $s \in \mathrm{S}$ einen Merkmalswert $X(s) \in \tilde{X}$ zuordnen; $\tilde{X}$ ist der Merkmalsraum der statistischen Variablen X. Für jede Menge von Merkmalswerten $\tilde{x} \subseteq \tilde{X}$ kann dann eine relative Häufigkeit definiert werden, die sich auf den Anteil von Situationen in S bezieht, die einen Merkmalswert in $\tilde{x}$ aufweisen. Die statistische oder auch Häufigkeitsverteilung der statistischen Variablen X ist dementsprechend eine Abbildung

$$\mathrm{P_S} : \mathcal{P}(\tilde{X}) \longrightarrow \mathbf{R}$$

die jeder Menge von Merkmalswerten $\tilde{x} \in \mathcal{P}(\tilde{X})$ ihre relative Häufigkeit $\mathrm{P_S}(\tilde{x})$ zuordnet.[12]

[12]Mit dem Symbol $\mathcal{P}(\tilde{X})$ soll auf die Potenzmenge von $\tilde{X}$ verwiesen werden. In der elementaren Statistik wird gezeigt, daß $\mathrm{P_S}$ eine additive Bewertungsfunktion ist, das heißt, für alle $\tilde{x}, \tilde{x}' \subseteq \tilde{X}$ gilt:

a) $0 \le \mathrm{P_S}(\tilde{x}) \le 1$

b) $\mathrm{P_S}(\mathrm{S}) = 1$

c) $\mathrm{P_S}(\tilde{x}^c) = 1 - \mathrm{P_S}(\tilde{x})$

d) $\tilde{x} \cap \tilde{x}' = \emptyset \implies \mathrm{P_S}(\tilde{x} \cup \tilde{x}') = \mathrm{P_S}(\tilde{x}) + \mathrm{P_S}(\tilde{x}')$

Zur Notation sei angemerkt, daß gelegentlich $\mathrm{P_S}(X \in \tilde{x})$ anstelle von $\mathrm{P_S}(\tilde{x})$ verwendet wird, wenn $\tilde{x}$ eine Teilmenge von $\tilde{X}$ ist; und $\mathrm{P_S}(X = \tilde{x})$ anstelle von $\mathrm{P_S}(\{\tilde{x}\})$, wenn $\tilde{x}$ ein Element aus $\tilde{X}$ ist.

4. In der sozialwissenschaftlichen Statistik handelt es sich zunächst immer um definite statistische Hypothesen; dann nämlich, wenn man deskriptives Wissen über in der sozialen Realität bestimmbare (und infolgedessen endliche) Gesamtheiten gewinnen möchte. Exemplarisch kann man daran denken, daß man Informationen über die Verdienstverhältnisse in einem Betrieb oder einer Branche oder Kenntnisse über die Dauer der Arbeitslosigkeit in einem gewissen Zeitraum gewinnen möchte; oder daß man sich dafür interessiert, wie sich die Berufsbildungsabschlüsse während eines gewissen Zeitraums verändert haben. Einem verbreiteten Verständnis zufolge geht es in den Sozialwissenschaften jedoch in erster Linie nicht um Fragestellungen dieser Art, sondern um indefinite Hypothesen: man möchte Aussagen „über alle Situationen eines gewissen Typs“ machen. In der sozialwissenschaftlichen Methodenliteratur wird tatsächlich von vielen Autoren die Auffassung vertreten, daß solche indefiniten Hypothesen mit Begriffsbildungen der Statistik formuliert und mit statistischen Methoden begründet werden können. Diese Auffassung wird meistens von Autoren vertreten, die glauben, daß in den Sozialwissenschaften „Gesetzmäßigkeiten“ des gesellschaftlichen Lebens und seiner historischen Entwicklung gefunden und begründet werden können.

5. Allerdings ist bereits problematisch, wie indefinite Hypothesen formuliert und verständlich gemacht werden können. Denn worüber wird geredet, wenn – vorgeblich – „über alle Situationen eines gewissen Typs“ geredet wird? Eine Möglichkeit, solchen Redeweisen einen Sinn zu geben, besteht darin, sie zu reformulieren. Man könnte sagen: man redet dann nicht über eine indefinite Menge von Situationen, sondern über Situationstypen. Als Sinnvoraussetzung wird dann unterstellt, daß eine Aussage über einen Situationstyp $\mathcal{S}$ auch für jede bestimmte Situation s zutrifft, die als Beispiel für $\mathcal{S}$ vorstellbar gemacht werden kann. Aber es stellt sich natürlich die Frage, wie Aussagen über Situationstypen, deren Verständnis diese Sinnvoraussetzung beinhaltet, nicht nur formuliert, sondern auch begründet werden können.

6. Eine Möglichkeit, Aussagen über Situationstypen zu machen, wurde schon im vorangegangenen Abschnitt am Beispiel des Situationstyps „Fußballspiel" erläutert. In diesem Fall handelte es sich jedoch um Aussagen über einen Sprachgebrauch, durch den das Verständnis eines Situationstyps fixiert werden soll. Daß die dadurch fixierten Eigenschaften eines Situationstyps auch allen Situationen zukommen, die als Beispiele des Situationstyps vorstellbar sind, ist dann kein empirischer Sachverhalt, sondern eine Implikation des (intendierten) Sprachgebrauchs. Oder anders formuliert: Eine Situation, die diese Eigenschaften nicht aufweist, ist dann *deshalb* kein Beispiel für den vorab fixierten Situationstyp. Offenbar gibt es hier eine Parallele zur Methode der Mathematiker, eine Menge vorstellbarer Gegenstände durch Konstruktionsverfahren einzuführen. Dennoch bleibt ein wesentlicher Unterschied. Durch die Fixierung von Eigenschaften für Situationstypen durch sprachliche Regeln wird zwar der durch „$s \models \mathcal{S}$" symbolisierten Vorstellung, daß s ein Beispiel für den Situationstyp $\mathcal{S}$ ist, ein bestimmter Sinn gegeben. Aber diese Regeln liefern uns keine Menge aller Situationen, die als Beispiele für $\mathcal{S}$ möglich sind.

7. Eine schwierige Frage ist, ob auch Aussagen über Situationstypen gemacht werden können, bei denen es sich nicht um Festlegungen eines Sprachgebrauchs handelt. Diese Frage steht im Mittelpunkt vieler wissenschaftstheoretischer Überlegungen und Kontroversen, und sie kann hier deshalb nicht angemessen besprochen werden. Jedoch soll wenigstens angedeutet werden, von welcher Auffassung sich dieser Text leiten läßt. Diese Auffassung besteht kurz gesagt darin, daß indefinite Hypothesen nur sinnvoll sind, wenn und insoweit sie sich als Hypothesen über Verfahren auffassen lassen, mit denen zumindest hypothetisch beliebig viele Situationen erzeugt werden können. Wenn ein solches Verständnis einer indefiniten Hypothese möglich ist, kann man nämlich Aussagen über „alle Situationen eines gewissen Typs" als Aussagen über Verfahren (nicht unbedingt „Mechanismen") verstehen, durch die die Situationen, über die man sprechen möchte, erzeugt werden oder erzeugt werden könnten.

8. Diese Auffassung dient in diesem Text zunächst als ein Leitfaden, um die aleatorische Wahrscheinlichkeitsrechnung zu verstehen, die sich mit „zufällig entstandenen" Situationen beschäftigt. Beginnend in Kapitel 4 wird versucht, sie als eine Theorie zu verstehen, die sich mit Verfahren (Zufallsgeneratoren) beschäftigt, mit denen Situationen (Ereignisse, Sachverhalte) erzeugt werden können.

Kapitel 3

Epistemische Wahrscheinlichkeit

Im vorangegangenen Kapitel wurde besprochen, wie man einige Aspekte unseres Redens über Möglichkeiten durch die Formulierung von Hypothesen explizit machen kann. Insbesondere haben wir uns mit Hypothesen beschäftigt, die sich darauf beziehen, wie Situationen durch Sachverhalte oder Ereignisse charakterisiert werden können. Da man sich bei Hypothesen dieser Art stets auf eine bestimmte Situation oder auf eine bestimmte endliche Menge von Situationen beziehen kann, wird von *definiten Hypothesen* gesprochen. In diesem Kapitel soll die Diskussion solcher Hypothesen fortgesetzt und überlegt werden, wie man von einer Wahrscheinlichkeit von Hypothesen sprechen kann.

3.1 Bildung von Hypothesen

1. Es scheint ein Gemeinplatz zu sein, daß man sich im praktischen Leben fortwährend in Unsicherheitssituationen befindet. Tatsächlich wird jedoch dieser Aspekt selten zum Gegenstand expliziter Überlegungen. Meistens verläßt man sich auf Selbstverständlichkeiten und Gewohnheiten; und die Aussage, daß man sich gleichwohl – ohne es zu wissen? – in Unsicherheitssituationen befindet, erscheint insoweit wenig sinnvoll. Es muß vielmehr in irgendeiner Form einen Anlaß geben, der eine *reflexive* Haltung motiviert, in der Möglichkeiten und Unsicherheiten einer Situation ins Blickfeld geraten. Erst dadurch *entsteht* eine Unsicherheitssituation.

2. Aber auch wenn eine Unsicherheitssituation entstanden ist, kommt es selten zu einer expliziten Hypothesenbildung. Vielmehr erwägen wir unterschiedliche Vermutungen, die zum Beispiel in Kommunikationsprozessen ausgetauscht werden. Sicherlich begleiten solche Reflexionsprozesse auch unser praktisches Handeln; aber es ist ziemlich unklar, ob und ggf. wie man davon sprechen kann, daß sie unserem Handeln „zugrunde liegen". Glücklicherweise braucht diese Frage hier nicht diskutiert zu werden. Es genügt, darauf aufmerksam zu werden, daß auch die Reflexion von Vermutungen meistens weitgehend selbstverständlich und gewohnheitsmäßig abläuft. Hält man sich an den in Abschnitt 2.2 erläuterten Sprachgebrauch, kann man insoweit auch noch nicht von einer Hypothesenbildung sprechen. Hypothesenbildung setzt vielmehr voraus, daß Vermutungen *explizit* gemacht und zu einem *Gegenstand* des Nachdenkens und der Kommunikation gemacht werden.

3. Es ist also keineswegs selbstverständlich, Hypothesen zu bilden und über ihre Wahrscheinlichkeit nachzudenken. Man kann es vielleicht so sagen: eine explizite Formulierung von Hypothesen mit dem Ziel, ihre Wahrscheinlichkeit einschätzbar zu machen, ist nur erforderlich, wenn (a) der gewohnheitsmäßige und selbstverständliche Umgang mit Möglichkeiten und Unsicherheiten unzureichend erscheint – also ein „Problem“ entstanden ist –, und (b) wenn und insoweit es wichtig erscheint, nicht einfach nur unterschiedliche Möglichkeiten zu bedenken und sich mehr oder weniger selbstverständlich von Plausibilitäten leiten zu lassen, sondern explizit Gründe zu finden, um Wahrscheinlichkeiten von Hypothesen einschätzbar zu machen.

4. Dafür gibt es hauptsächlich zwei Motive. Einerseits ein wissenschaftliches Erkenntnisinteresse, in dessen Verfolgung man sich nicht mit mehr oder weniger beliebigen Vermutungen zufriedengeben möchte. Andererseits praktische Interessen, die ein Motiv dafür liefern können, über die Folgen von Handlungsalternativen nachzudenken. Insoweit diese Folgen auch von Umständen abhängen, über die nur Vermutungen angestellt werden können, kann es sinnvoll erscheinen, explizit Hypothesen über Handlungsbedingungen und Handlungsfolgen zu bilden und durch Argumentationen einschätzbar zu machen.

5. Vermutungen, die bei wissenschaftlichen Gedankengängen oder beim Nachdenken über Handlungsmöglichkeiten wichtig erscheinen können, sind unterschiedlicher Art. Es sollte deshalb betont werden, daß wir uns hier zunächst nur mit einer bestimmten Sorte von Vermutungen beschäftigen, nämlich Vermutungen, die in der Form definiter Hypothesen formuliert werden können. Den Ausgangspunkt bildet die sprachliche Fixierung einer Situation, auf deren Beschaffenheit sich die Hypothesen beziehen.[1] Weiterhin erforderlich ist die Bildung einer Menge von Kennzeichnungen, die zur hypothetischen Charakterisierung der Situation verwendet werden können.

6. Wie eine Kennzeichnungsmenge gebildet werden kann, hängt davon ab, was wir bereits über die Situation wissen, wozu auch ggf. vorhandenes Wissen darüber gehört, wie die Situation entstanden sein könnte

[1] Stattdessen kann man sich auch auf eine bestimmte endliche Menge von Situationen beziehen und z.B. statistische Hypothesen formulieren. In diesem Kapitel wird zur Illustration stets auf singuläre Hypothesen bezug genommen.

oder entstehen wird. Hypothesenbildung bezieht sich auf Aspekte einer Situation, die wir nicht oder nicht genau kennen; aber bereits um sagen zu können, was wir nicht oder nicht genau wissen, benötigen wir einiges Wissen über die Beschaffenheit einer Situation. Ohne ein solches Wissen könnten wir uns gar nicht reflexiv auf eine Situation beziehen.

7. Es sollte auch betont werden, daß es im allgemeinen keinen unmittelbaren Zusammenhang zwischen Hypothesen und Entscheidungen gibt. Sicherlich kann die Bildung und Bewertung von Hypothesen im Rahmen von Entscheidungsprozessen eine Rolle spielen. Wenn es dagegen um wissenschaftliche Hypothesen geht, ist es weder erforderlich noch sinnvoll, Entscheidungen zu treffen. Die Aufgabe besteht dann vielmehr darin, die Gründe darzulegen, die für bzw. gegen Hypothesen sprechen, um dadurch ihre Wahrscheinlichkeit einschätzbar zu machen. Dann endet ein Gedankengang vielleicht vorläufig damit, daß eine der in Betracht gezogenen Hypothesen am relativ wahrscheinlichsten erscheint. Aber es gibt dennoch keinen Grund, sie deshalb „für wahr zu halten“ oder „zu glauben“ oder sich „für sie zu entscheiden“.

8. Im übrigen ist es auch im praktischen Leben keineswegs immer sinnvoll, die relativ wahrscheinlichste Hypothese für zutreffend zu halten. Denken wir noch einmal an das Seminar, das in der nächsten Woche wahrscheinlich stattfinden wird. Vielleicht haben wir einen Grund für die Annahme, daß die Hypothese, daß das Seminar nicht stattfinden wird, etwas wahrscheinlicher ist als die entgegengesetzte Hypothese. Sollte man deshalb dem Seminar in der nächsten Woche fernbleiben? Das ist offenbar eine ganz andere Frage. Bei Entscheidungen geht es darum, was getan werden *sollte*. Bei der Einschätzung von Hypothesen geht es darum, wie reale oder vorstellbare Situationen durch Sachverhalte oder Ereignisse charakterisiert werden können. Zwischen beiden Arten von Fragestellungen gibt es keinen unmittelbaren Zusammenhang.[2]

[2] Entsprechend hat sich auch Hacking (1965, S. 28) geäußert: „If one hypothesis is better supported than another, it would usually be, I believe, right to call it the more reasonable. But of course it need not be reasonable positively to believe the best supported hypothesis, nor the most reasonable one. Nor need it be reasonable to act as if one knew the best supported hypothesis were true.“

3.2 Algebra für Hypothesen

1. Für die weiteren Überlegungen soll vorausgesetzt werden, daß man sich auf ein System definiter (singulärer) Hypothesen

$$\mathcal{H} := \{ \langle s, \tilde{k} \rangle \mid \tilde{k} \in \tilde{K} \}$$

beziehen kann. In diesem Abschnitt werden zwei Möglichkeiten besprochen, wie man aus einem solchen System von Hypothesen neue Hypothesen bilden kann.

2. Die erste Idee besteht darin, Hypothesen zu vergröbern. Um zu verdeutlichen, was gemeint ist, denke man an ein Beispiel: würfeln. Die Hypothesen beziehen sich auf eine Situation, die als Ergebnis des Wurfs eines Würfels entsteht oder bereits entstanden, aber nicht bekannt ist. Ein gewöhnlicher Würfel erlaubt es, sechs Möglichkeiten (Augenzahlen) zu unterscheiden; und es liegt nahe, zur Bildung von Hypothesen die Kennzeichnungsmenge

$$\tilde{K} := \{1, 2, 3, 4, 5, 6\}$$

zu verwenden. Vielleicht soll es indessen nur darum gehen, ob eine gerade oder eine ungerade Augenzahl geworfen wird. Dann kann man die folgende gröbere Kennzeichnungsmenge verwenden:

$$\tilde{K}^* := \{\tilde{k}_1^*, \tilde{k}_2^*\}$$

wobei $\tilde{k}_1^* := \{2, 4, 6\}$ und $\tilde{k}_2^* := \{1, 3, 5\}$ ist. Bildet man auf diese Weise aus einer Kennzeichnungsmenge $\tilde{K}$ eine gröbere Kennzeichnungsmenge $\tilde{K}^*$, kann man mit ihr ein neues System von Hypothesen bilden:

$$\mathcal{H}^* := \{ \langle s, \tilde{k}^* \rangle \mid \tilde{k}^* \in \tilde{K}^* \}$$

Wenn $\mathcal{H}$ ein System von Hypothesen ist, dann ist offenbar auch $\mathcal{H}^*$ ein System von Hypothesen.

3. Der zweiten Überlegung liegt die Idee zugrunde, daß man die Hypothesen aus $\mathcal{H}$ verwenden kann, um aus ihnen neue Hypothesen zu bilden. Dafür bieten sich die elementaren Operatoren der Aussagenlogik an: *und* ($\wedge$), *oder* ($\vee$), *nicht* ($\neg$). Wenn also $\langle s, \tilde{k} \rangle$ eine Hypothese ist, kann man zum Beispiel die neue Hypothese

$$\neg\langle s, \tilde{k} \rangle$$

bilden, womit gemeint ist, daß die Kennzeichnung $\tilde{k}$ für die Situation s *nicht* zutrifft. Ganz entsprechend kann man mit den Operatoren $\vee$ und $\wedge$ jeweils zwei Hypothesen zu einer neuen Hypothese verknüpfen. Zum Beispiel bedeutet

$$\langle s, \tilde{k} \rangle \vee \langle s, \tilde{k}' \rangle$$

daß die Kennzeichnung $\tilde{k}$ *oder* die Kennzeichnung $\tilde{k}'$ für die Situation s zutrifft. Und

$$\langle s, \tilde{k} \rangle \wedge \langle s, \tilde{k}' \rangle$$

bedeutet, daß sowohl die Kennzeichnung $\tilde{k}$ als auch die Kennzeichnung $\tilde{k}'$ für die Situation s zutrifft. Da in diesem Beispiel $\tilde{k}$ und $\tilde{k}'$ Elemente der Kennzeichnungsmenge $\tilde{K}$ sind und sich infolgedessen wechselseitig ausschließen, kann allerdings die zuletztgenannte Hypothese niemals richtig sein.

4. Dennoch kann auch der $\wedge$-Operator durchaus nützlich sein, denn der Prozeß der Bildung neuer Hypothesen kann fortgesetzt werden. In den bisher angegebenen Beispielen sind wir von Hypothesen der Form $\langle s, \tilde{k} \rangle$ ausgegangen, wobei $\tilde{k}$ ein Element der vorgegebenen Kennzeichnungsmenge $\tilde{K}$ ist. Wir sprechen in diesen Fällen von *elementaren* Hypothesen, denn ein System von Hypothesen stellt uns zunächst diese Elementarhypothesen zur Verfügung. Wenn man jedoch aus den Elementarhypothesen neue Hypothesen gebildet hat, kann man mithilfe der logischen Operatoren auch aus diesen neuen Hypothesen wiederum andere Hypothesen konstruieren, usw. Den Sinn des $\wedge$-Operators kann man sich auch dadurch verdeutlichen, daß er durch die beiden anderen Operatoren definiert werden kann. Es gilt nämlich für beliebige Aussagen A und B folgende (aussagenlogische) Äquivalenz:

$$A \wedge B \equiv \neg(\neg A \vee \neg B) \tag{3.2.1}$$

Die Klammern sind erforderlich, um die Reihenfolge festzulegen, in der die Operatoren angewendet werden sollen. Als Konvention wird angenommen, daß $\neg$ die höchste, $\wedge$ die zweithöchste und $\vee$ die niedrigste Priorität hat.

5. In der Wahrscheinlichkeitstheorie ist es üblich, nicht nur Elementarhypothesen zu betrachten, sondern außerdem alle Hypothesen, die sich aus ihnen mithilfe der drei genannten Operatoren bilden lassen.

Die Gesamtheit dieser Hypothesen wird eine *Algebra für Hypothesen* genannt.[3] Ihre Definition hängt natürlich davon ab, daß zunächst ein System von Elementarhypothesen eingeführt wird. Um auf eine Algebra für Hypothesen zu verweisen, verwenden wir deshalb den Ausdruck $\mathcal{A}(\mathcal{H})$, wobei $\mathcal{H}$ ein System von Elementarhypothesen ist. Die Regeln, die für eine Algebra für Hypothesen gelten, kann man so formulieren:

a) $H \in \mathcal{A}(\mathcal{H}) \implies \neg H \in \mathcal{A}(\mathcal{H})$

b) $H \in \mathcal{A}(\mathcal{H}),\ H' \in \mathcal{A}(\mathcal{H}) \implies H \vee H' \in \mathcal{A}(\mathcal{H})$

c) $H \in \mathcal{A}(\mathcal{H}),\ H' \in \mathcal{A}(\mathcal{H}) \implies H \wedge H' \in \mathcal{A}(\mathcal{H})$

Weitere Regeln können aus diesen Regeln abgeleitet werden. Man kann mithilfe von (3.2.1) auch bereits die dritte Regel aus den ersten beiden Regeln ableiten.

3.3 Epistemische Wahrscheinlichkeit

1. Hypothesen können mehr oder weniger wahrscheinlich sein. Jedenfalls ist uns eine solche Redeweise aus der umgangssprachlichen Reflexion und Kommunikation von Vermutungen vertraut. Zum Beispiel kann man sagen: sehr wahrscheinlich wird das Seminar in der nächsten Woche wieder stattfinden. Fragt man indessen, „wie groß“ die Wahrscheinlichkeit einer Hypothese ist, gerät man sofort in Schwierigkeiten.

2. Bei den meisten Hypothesen erscheint es nicht sinnvoll, ihnen einen numerisch bestimmten Wahrscheinlichkeitsgrad zuzuschreiben. Zur Illustration sei auf einige Beispiele aus einem Buch zur Wahrscheinlichkeitslehre von H. Reichenbach (1949, S. 1) hingewiesen:

„Übrigens wäre es irrtümlich zu glauben, daß der Wahrscheinlichkeitsbegriff nur bei Zukunftsaussagen auftritt. Es gibt eine ganze Reihe von Vergangenheitsaussagen, in denen wir offensichtlich von dem Wahrscheinlichkeitsbegriff Gebrauch machen. Daß Nero den Brand Roms veranlaßt hat, hält der Historiker für sehr wahrscheinlich, daß Henriette von England, die als Herzogin von Orléans am Hofe Ludwigs XIV lebte, ermordet worden ist, hält er für weniger wahrscheinlich, daß Bacon der Autor der Shakespeareschen Dramen ist, ist

[3] Man spricht auch von einer *Booleschen Algebra*, um an George Boole (1815 – 1864) zu erinnern, der die Idee einer aussagenlogischen Algebra zuerst explizit entwickelt hat. Eine auch unter historischen Gesichtspunkten interessante Einführung in seine Gedankenwelt vermittelt die von Grattan-Guinness und Bornet herausgegebene Textzusammenstellung in Boole, 1997.

unwahrscheinlich. Auch diese Tatbestände der Vergangenheit lassen sich also nur mit Wahrscheinlichkeitsanspruch behaupten.“

Es erscheint nicht nur unmöglich, sondern hier wird auch nicht der Versuch gemacht, den Hypothesen numerisch bestimmte Wahrscheinlichkeiten zu geben. Vielmehr bedient sich der Autor – so wie man es in der Umgangssprache macht – nur vager Formulierungen: daß eine Hypothese „unwahrscheinlich“, „sehr wahrscheinlich“ oder „weniger wahrscheinlich“ ist.

3. Dennoch wird in der Literatur zur Wahrscheinlichkeitsrechnung meistens davon ausgegangen, daß ‘Wahrscheinlichkeit’ als ein metrisierbarer Größenbegriff konzipiert werden kann, so daß „Grade der Wahrscheinlichkeit“ numerisch angegeben werden können. Es wird also angenommen, daß man eine Funktion

$$\mathrm{P_e} : \ \mathcal{A}(\mathcal{H}) \longrightarrow \mathbf{R}$$

konzipieren kann, mit der jeder Hypothese H aus einem System von Hypothesen $\mathcal{A}(\mathcal{H})$ eine numerisch bestimmte Wahrscheinlichkeit $\mathrm{P_e}(H)$ zugeordnet werden kann. Ob und ggf. wie quantitative Wahrscheinlichkeitsmaße dieser Art begründet werden könnten, ist allerdings fragwürdig. In diesem Text wird die Auffassung vertreten, daß quantitative Wahrscheinlichkeitsaussagen nur unter sehr spezifischen Voraussetzungen möglich sind (wenn man sich auf Zufallsgeneratoren beziehen kann) und daß die Annahme eines quantitativen Wahrscheinlichkeitsmaßes deshalb nicht von vornherein und in allen Anwendungskontexten vorausgesetzt werden kann.[4] Stattdessen wird in diesem Text von einem *komparativen* Verständnis von Wahrscheinlichkeitsaussagen ausgegangen.[5] Die sprachliche Form besteht dann darin: daß eine Hypothese H „wahrscheinlicher

[4] Dies wird auch in der Literatur zur Wahrscheinlichkeitstheorie gelegentlich betont. Zum Beispiel heißt es bei Czuber (1908, S. 76): „Das tägliche Leben veranlaßt uns sehr häufig wenn auch nicht zu Wahrscheinlichkeitsbestimmungen, so doch zu Wahrscheinlichkeitsschätzungen. Wir urteilen über das Eintreffen oder Nichteintreffen eines Ereignisses auf Grund allgemeiner Erwägungen, bei denen wir uns durch erworbene Erfahrungen leiten, vielfach aber auch unseren Seelenzustand mit zu Worte kommen lassen; das Ergebnis wird nur ausnahmsweise eine numerische Wahrscheinlichkeit, sondern in der Regel lediglich eine allgemeine Aussage über die Größenbeziehung der Wahrscheinlichkeiten für das Eintreffen und das Nichteintreffen sein. Der subjektive Charakter eines so gewonnenen Resultates schließt einen erkenntnistheoretischen Wert desselben aus.“ Abgesehen von dem letzten Satz erscheint diese Bemerkung sehr plausibel.

[5] Dieses Verständnis wurde insbesondere von Keynes (1921) ausgearbeitet, dem wir insoweit folgen. Ein weiterer Autor, der ein komparatives Verständnis von Wahrscheinlichkeitsaussagen vorgeschlagen hat, ist Hacking, 1965.

als“ oder „nicht weniger wahrscheinlich als“ eine andere Hypothese H' ist.

4. Aber was ist mit Aussagen dieser Art gemeint? Man kann sich hier nicht auf einen wohldefinierten Wahrscheinlichkeitsbegriff berufen; denn die Worte ‘wahrscheinlich’ oder ‘Wahrscheinlichkeit’ werden in unterschiedlichen Kontexten und mit unterschiedlichen Bedeutungen verwendet. Um dennoch einen Ausgangspunkt für weitere Überlegungen zu gewinnen, gehen wir von folgendem Verständnis aus: daß mit der Aussage „die Hypothese H ist wahrscheinlicher als die Hypothese H'“ gemeint ist, daß die Gründe, die für H sprechen, besser sind als die Gründe, die für H' sprechen. Wahrscheinlichkeitsaussagen haben es bei diesem Verständnis mit Gründen zu tun, die für oder gegen Hypothesen angeführt werden können.

5. Um Wahrscheinlichkeitsaussagen als komparative Aussagen zu formulieren, wird die Schreibweise

$$H' \prec_e H$$

verwendet, wenn gemeint ist, daß H wahrscheinlicher ist als H'; und die Schreibweise

$$H' \preceq_e H$$

wenn gemeint ist, daß H nicht weniger wahrscheinlich ist als H'. Aussagen dieser Art nennen wir *epistemische Wahrscheinlichkeitsaussagen*; denn es handelt sich um Aussagen über Hypothesen, mit denen eine Bewertung von Gründen erfolgt, die für oder gegen die Hypothesen angeführt werden können.

6. Um ein Verständnis epistemischer Wahrscheinlichkeitsaussagen zu gewinnen, ist es wichtig, auf die sprachliche Form zu achten. Denn man kann das Wort ‘wahrscheinlich’ auf zwei unterschiedliche Weisen verwenden. Einerseits kann man es verwenden, um *modale* Aussagen zu formulieren. Die Verwendung des Wortes ‘wahrscheinlich’ für modale Aussagen wurde von White (1975, S. 59) so charakterisiert:

> “For both the possibility that something should be so, that is, the possibility of ‘can’, and the possibility that something is so, that is, the possibility of ‘may’, circumstances may be such that *either* they exclude the possibility of something *or* they exclude the possibility of its opposite *or* they allow both its possibility and the possibility of its opposite. Where they exclude the possibility of something, they make it *impossible* – though the exclusion of the possibility of ‘may’ is usually expressed by ‘not possible’. Where, however, they exclude the possibility of something's opposite they make the thing itself either *necessary* or *certain*, depending [...] on whether the possibility that is excluded is the possibility that it should be so (‘can’) or the possibility that it is so (‘may’). *Probability* enters at this stage which is intermediate between the exclusion of the possibility of something and the exclusion of the possibility of its opposite, that is, the stage where there exist both the possibility of something and the possibility of its opposite. To say that something is ‘probable’, or ‘likely’, is to indicate the degree to which circumstances favour its being so in comparison with the existence of its possible alternatives.”

Unsere Sprache vermittelt uns zunächst ein solches modales Verständnis, bei dem sich Wahrscheinlichkeitsaussagen unmittelbar auf Möglichkeiten beziehen, die dem Geschehen in unserer Erfahrungswelt gewissermaßen innewohnen. Wenn das gemeint ist, spreche ich von *modalen* Wahrscheinlichkeitsaussagen. — Davon zu unterscheiden ist ein *epistemisches* Verständnis von Wahrscheinlichkeitsaussagen.[6]

7. Bei epistemischen Wahrscheinlichkeitsaussagen sollte zunächst auf die sprachliche Form geachtet werden: es sind *Aussagen über Hypothesen*; zum Beispiel: H ist wahrscheinlicher als H'. Im Unterschied zu modalen Wahrscheinlichkeitsaussagen sind epistemische Wahrscheinlichkeitsaussagen nicht unmittelbar Aussagen über unsere Erfahrungswelt, sondern Aussagen über die Wahrscheinlichkeit von Hypothesen. Eine Bezugnahme auf die Erfahrungswelt erfolgt nur indirekt durch die Hypothesen, durch die auf mögliche Sachverhalte oder Ereignisse bezug genommen wird. Epistemische Wahrscheinlichkeitsaussagen dienen dann – gewissermaßen auf einer zweiten Stufe – dazu, diese Hypothesen durch einen Hinweis auf Gründe, die für oder gegen die Hypothesen sprechen, einschätzbar zu machen.[7]

---

[6] In der philosophischen Literatur wird eine solche Unterscheidung manchmal mit den Ausdrücken *de re* und *de dicto* angedeutet. Aussagen *de re* beziehen sich auf unsere Erfahrungswelt, Aussagen *de dicto* dagegen auf unser Reden über ... Verwendet man diese Ausdrücke, sind modale Wahrscheinlichkeitsaussagen *de re* und epistemische Wahrscheinlichkeitsaussagen *de dicto*. Wie von Hacking (1971) dargestellt worden ist, sind in der Geschichte der Wahrscheinlichkeitstheorie beide Aspekte oft vermischt worden.

[7] Ich betone hier die Unterscheidung zwischen Aussagen über unsere Erfahrungswelt und Aussagen über Hypothesen, weil sich daraus – wie wir später noch genauer sehen werden – einige wichtige Implikationen für das Verständnis von Wahrscheinlichkeitsaussagen ergeben. Die Unterscheidung wendet sich zugleich gegen eine in der Literatur oft anzutreffende Neigung, gar nicht ernsthaft darüber nachzudenken, *worüber* Aussagen gemacht werden sollen. Exemplarisch sei auf Lenzen (1980, S. 81f) hingewiesen: „Der durch ‘$X \leq_a Y$’ zu symbolisierende Grundbegriff subjektiver Wahrscheinlichkeitstheorie beinhaltet, daß eine Person a das Ereignis X für höchstens so

8. Das Bemühen, Wahrscheinlichkeitstheorie auf der Ebene epistemischer Wahrscheinlichkeitsaussagen zu formulieren, kann man als einen Versuch verstehen, die Theoriebildung von den Schwierigkeiten, die mit modalen Sprech- und Denkweisen verbunden sind, zu befreien. Es ist allerdings eine offene Frage, wieweit ein solcher Versuch gelingen kann. Man sieht das schon daran, daß auch epistemische Wahrscheinlichkeitsaussagen meistens nur als Vermutungen formuliert werden können,[8] und daß zu ihrer Begründung ein Rückgriff auf modale Formulierungen im allgemeinen nicht vermieden werden kann. Dadurch bleiben auch epistemische Wahrscheinlichkeitsaussagen an ein elementares Verständnis von Modalitäten gebunden. Die Unterscheidung ist dennoch wichtig. Erst eine sprachliche Ablösung von modalen Wahrscheinlichkeitsaussagen, die sich unmittelbar auf unsere Erfahrungswelt beziehen, erlaubt es, Wahrscheinlichkeitsaussagen zum Gegenstand einer eigenständigen Theoriebildung zu machen, die sich nicht unmittelbar mit Aspekten unserer Erfahrungswelt, sondern mit epistemischen Wahrscheinlichkeitsaussagen beschäftigt.[9]

9. Allerdings bezieht sich diese Überlegung zunächst nur auf epistemische Verwendungen des Wahrscheinlichkeitsbegriffs. Konzipiert man die Wahrscheinlichkeitsrechnung als eine Theorie der Zufallsgeneratoren, wie ich später vorschlagen werde, bekommen dadurch *ihre* Wahrscheinlichkeitsaussagen einen modalen Charakter: sie beziehen sich auf Möglichkeiten, die durch Zufallsgeneratoren definiert werden.

---

wahrscheinlich hält wie das Ereignis Y. Dabei interessiert den Wahrscheinlichkeitstheoretiker nicht, über welchen Gegenstandsbereich die Variablen $X, Y, \ldots$ laufen, d.h. welche Entitäten 'Ereignisse' sind. Je nach Kontext könnten Ereignisse als Mengen von Sachverhalten, von Zuständen, oder gar von 'möglichen Welten' interpretiert werden. Der Wahrscheinlichkeitstheoretiker abstrahiert von diesen möglichen Deutungen, indem er jede beliebige Menge als 'Ereignis' zuläßt.“ Daß bereits das Reden von einem „subjektiven“ Wahrscheinlichkeitsbegriff fragwürdig ist, werden wir im nächsten Abschnitt besprechen.

[8] Infolgedessen ist es möglich, auch Wahrscheinlichkeitsaussagen als Hypothesen zu formulieren. Unabhängig davon, ob und ggf. wie dies geschieht, sollten jedoch Wahrscheinlichkeitsaussagen von den Hypothesen unterschieden werden, *über* die sie eine Aussage (Vermutung, Hypothese) formulieren.

[9] Hier sollte darauf geachtet werden, daß man bei einer metasprachlichen Betrachtungsweise auch epistemische Wahrscheinlichkeitsaussagen „modal“ nennen kann, etwa so wie in der folgenden Bemerkung von Lucas (1970, S. 4): „'Probably' is therefore a modal word. It affects not the content of what is said, but the way in which it is said, the force with which it is said, and the degree to which it is to be relied on.“ Dieses Verständnis von „modal“ muß jedenfalls von einer Konzeption modaler Aussagen unterschieden werden, wie sie in den oben angeführten Bemerkungen von White zum Ausdruck kommt.

3.4 Abhängigkeit vom Wissensstand

1. Aussagen müssen von ihren Begründungen unterschieden werden. Mit einer epistemischen Wahrscheinlichkeitsaussage „$H' \prec_e H$“ wird zum Ausdruck gebracht, daß die Hypothese H wahrscheinlicher ist als die Hypothese H'. Wissenswert wird eine solche Aussage aber nur dadurch, daß eine Begründung angegeben wird. In dieser Hinsicht unterscheiden sich Wahrscheinlichkeitsaussagen nicht von anderen Aussagen.[10]

2. Will man die Aussage „$H' \prec_e H$“ begründen, muß man Gründe angeben, die sich auf die Hypothesen H und H' beziehen. Die Möglichkeit, solche Gründe angeben zu können, hängt wiederum von unserem Wissen über die Situation ab, auf die sich die Hypothesen beziehen. Denken wir an unser Seminarbeispiel. Unter normalen Umständen ist es viel wahrscheinlicher, daß das Seminar stattfinden wird, verglichen mit der Alternative, daß es ausfällt. Wenn wir dann jedoch erfahren, daß der Seminarleiter erkrankt ist, erscheint es wahrscheinlicher, daß das Seminar ausfallen wird.

3. Die jeweils vorhandenen Möglichkeiten zur Formulierung und Begründung von Wahrscheinlichkeitsaussagen hängen also von unserem Wissen ab. So formuliert handelt es sich allerdings nicht um eine Besonderheit von Wahrscheinlichkeitsaussagen, sondern gilt generell für die Formulierbarkeit und Begründbarkeit von Aussagen aller Art. Bei epistemischen Wahrscheinlichkeitsaussagen gibt es jedoch eine Besonderheit, die daraus resultiert, daß es sich nicht unmittelbar um Aussagen über unsere Erfahrungswelt handelt, sondern um Aussagen über Hypothesen. Die Aussage „$H' \prec_e H$“ *bedeutet*, daß es für H bessere Gründe gibt als für H'. In dieser Betrachtungsweise beziehen sich epistemische Wahrscheinlichkeitsaussagen *in ihrer Bedeutung* auf Begründungen und mithin auch auf das Wissen, das für Begründungen zur Verfügung steht. Man kann es vielleicht so sagen: für gewöhnliche Aussagen (und auch für modale Wahrscheinlichkeitsaussagen) ist Wissen als eine Ressource für ihre Formulierung und Begründung erforderlich; aber der gemeinte Sinn läßt sich von ihrer Begründung unterscheiden. Dagegen wird mit epistemischen Wahrscheinlichkeitsaussagen in komparativer Weise über Begründungen von Hypothesen gesprochen.

4. Daß epistemische Wahrscheinlichkeitsaussagen auf spezifische Weise von unserem Wissen abhängig sind, hat einige Autoren veranlaßt,

---

[10] Wenn hier und im folgenden ohne Zusatz von Wahrscheinlichkeitsaussagen gesprochen wird, sind stets epistemische Wahrscheinlichkeitsaussagen gemeint.

von „subjektiven“ Wahrscheinlichkeitsaussagen bzw. von einem „subjektiven“ Wahrscheinlichkeitsbegriff zu sprechen. Demgegenüber sollte zunächst betont werden, daß – in einem allgemeinen Sinn – die Abhängigkeit von unserem Wissensstand keine Besonderheit von Wahrscheinlichkeitsaussagen ist, sondern für alle Aussagen zutrifft. Infolgedessen liefert das Attribut ‘subjektiv’ keine sinnvolle Charakterisierung.

5. Wenn in der Literatur von einem „subjektiven“ Wahrscheinlichkeitsbegriff gesprochen wird, ist indessen oft etwas anderes gemeint, nämlich eine Auffassung, die zum Beispiel von De Morgan (1838, S. 6f) folgendermaßen zum Ausdruck gebracht wurde:

> “The quantities which we propose to compare are the forces of the different impressions produced by different circumstances. The phraseology of mechanics is here extended: by force, we merely mean cause of action, considered with reference to its magnitude, so that it is more or less according as it produces greater or smaller effect. It is one of the most essential points of the subject to draw the distinction we now explain. Probability is the feeling of the mind, not the inherent property of a set of circumstances. It is frequently referred to external objects, as if it accompanied them independently of ourselves, in the same manner as we imagine colour, form, &c. to abide by them. Thus we hold it just to say, that a white ball may be shut up in a box, and whether we allow light to shine on it or not, it is still a white ball. And if we were to translate the common notion, we should also say that in a lottery of balls shut up in a box, each ball has its probability of being drawn inseparably connected with it, just as much as form, size, or colour. But this is evidently not the case: Two spectators, who stand by the drawner, may be very differently affected with the notion of likelihood in respect to any ball being drawn. Say that the question is, whether a red or a green ball shall be drawn, and suppose that A feels certain that all the balls are red, B, that all are green, while C knows nothing whatever about the matter. We have here, then, in reference to the drawing of a red ball, absolute certainty for or against, with absolute indifference, in three different persons, coming under different previous impressions. And thus we see that the real probabilities may be different to different persons.”

Diese Passage ist interessant, weil sich in ihr eine richtige und eine fragwürdige Überlegung vermischen.

6. Es ist einsichtig, daß Wahrscheinlichkeiten nicht als Eigenschaften von Objekten oder Situationen verstanden werden können. Dies ist jedenfalls dann einsichtig, wenn man sich auf epistemische Wahrscheinlichkeitsaussagen bezieht.[11] In diesem Kontext kann man indessen Wahrscheinlichkeiten auch nicht als Eigenschaften von Hypothesen verstehen; bei der Formulierung epistemischer Wahrscheinlichkeitsaussagen dienen Formulierungen wie „ist wahrscheinlicher als“ vielmehr dazu, etwas über die Gründe zu sagen, die für oder gegen die Hypothesen, auf die sich die Aussage bezieht, angeführt werden können.

[11] Wir werden uns jedoch in Kapitel 4 mit einer Konzeption von Wahrscheinlichkeiten beschäftigen, die es in gewisser Weise erlaubt, sie zur Charakterisierung von Verfahren (Zufallsgeneratoren) zu verwenden.

7. Aus der richtigen Einsicht, daß Wahrscheinlichkeiten keine Eigenschaften von Objekten oder Situationen sind, zieht De Morgan jedoch eine fragwürdige Schlußfolgerung: daß Wahrscheinlichkeit als ein „feeling of the mind“ verstanden werden sollte. Diese Auffassung ist fragwürdig, weil sie weder dem Verständnis noch der Verwendung epistemischer Wahrscheinlichkeitsaussagen gerecht wird. Wenn wir eine Wahrscheinlichkeitsaussage formulieren, geben wir dadurch nicht einen Bericht über einen Gefühlszustand. Dies ist im allgemeinen weder der gemeinte Sinn einer Wahrscheinlichkeitsaussage, noch verhält es sich so, daß wir Wahrscheinlichkeitsaussagen durch Hinweise auf Gefühle begründen. Weder reden wir mit Wahrscheinlichkeitsaussagen über Gefühle (oder wie auch immer konzipierte „mentale Zustände“), noch können Hinweise auf Gefühle Begründungen liefern. Vielmehr muß man sich zur Begründung einer epistemischen Wahrscheinlichkeitsaussage auf die Hypothesen beziehen, die durch ein Wahrscheinlichkeitsurteil vergleichbar gemacht werden sollen. Es geht also um Gründe, die für oder gegen die Hypothesen angeführt werden können.

8. Daß solche Gründe nicht „objektiv festgestellt“ werden können, ist allerdings selbstverständlich, denn Gründe gibt es nicht als wahrnehmbare Objekte oder Sachverhalte in unserer Erfahrungswelt. Wenn überhaupt, gibt es sie nur als Argumente im Kontext reflexiver und kommunikativer Begründungen von Behauptungen oder Vermutungen. Insofern ist es auch nicht ausgemacht, daß man bei der Formulierung epistemischer Wahrscheinlichkeitsaussagen stets eine gemeinsame Auffassung erzielen kann. Aber auch das ist keine Besonderheit von Wahrscheinlichkeitsaussagen.

9. Es sei angemerkt, daß nicht alle Autoren, die von einem „subjektiven“ Wahrscheinlichkeitsbegriff sprechen, damit die Auffassung verbinden, daß Wahrscheinlichkeitsaussagen Aussagen über „mentale Zustände“ sind. Allerdings werden oft Formulierungen verwendet, die ein Verständnis sehr erschweren. Verbreitet ist insbesondere die Redeweise, daß Wahrscheinlichkeiten als „degrees of belief“ verstanden werden können. Die Verwendung dieses Ausdrucks verbindet sich dann oft mit einer psychologisierenden Rhetorik. Zum Beispiel heißt es bei Ramsey (1926, S. 166):

"The subject of our inquiry is the logic of partial belief, and I do not think we can carry it far unless we have at least an approximate notion of what partial belief is, and how, if at all, it can be measured. [...] We must therefore try to develop a purely psychological method of measuring belief."

Diese und ähnliche Formulierungen erwecken den Eindruck, daß Wahrscheinlichkeitsaussagen als Aussagen *über* „mentale Zustände" konzipiert werden sollen.[12]

10. Es ist jedoch bemerkenswert, daß Ramsey in dem gleichen Essay die Auffassung kritisiert, daß Wahrscheinlichkeitsaussagen als Aussagen über Gefühlszustände verstanden werden können. Er sagt zum Beispiel (1926, S. 170):

"But when we seek to know what is the difference between believing more firmly and believing less firmly, we can no longer regard it as consisting in having more or less of certain observable feelings; at least I personally cannot recognize any such feelings."

Der Vorschlag, den Ramsey schließlich entwickelt, um „degrees of belief" quantifizierbar zu machen, hat auch tatsächlich mit einer psychologischen Theorie (mit der vielleicht Aussagen über „mentale Zustände" formuliert werden können) nichts zu tun.[13] Er schlägt nämlich vor, „degrees of belief" von Menschen dadurch zu „messen", daß man ihnen Wetten vorschlägt. Dann geht es aber nicht darum, Gefühlszustände zu ermitteln, sondern man möchte herausfinden, welche Einschätzung von Hypothesen Menschen für „vernünftig" oder „vertretbar" halten. Ob das durch Wetten auf eine vernünftige Weise erreicht werden kann, ist natürlich eine ganz andere Frage. Die Idee, sich auf Wetten zu verlassen, ist allerdings von zahlreichen Autoren, die einen „subjektiven" Wahrscheinlichkeitsbegriff vertreten, aufgegriffen und weiter verfolgt worden. Es sollte deshalb überlegt werden, ob wir diese Auffassung später noch etwas genauer diskutieren sollten.[14]

[12] Zum Beispiel hat auch de Finetti (1972, S. xiii) von Wahrscheinlichkeiten als „expressions of subjective feeling" gesprochen.

[13] In einer späteren Notiz bemerkte Ramsey (1929, S. 256) auch ausdrücklich: „The defect of my paper on probability was that it took partial belief as a psychological phenomenon to be defined and measured by a psychologist."

[14] Ein kritisches Argument von Benenson (1984, S. 46) kann schon an dieser Stelle erwähnt werden: „... by explaining judgments of probability as propensities to wager, subjective theories of probability put the horse before the cart: all but the most naive gamblers try to establish (by reference to the evidence available) what the correct probability of an event is – to what degree it is reasonable to expect it to occur – in order to decide how to wager; they do not blindly wager and then conclude what degree of belief they must have had in its occurrence."

11. Für unser Verständnis epistemischer Wahrscheinlichkeitsaussagen halten wir fest, daß sie sich auf Begründungen für Hypothesen beziehen und mithin selbst einer Begründung bedürfen. Von entscheidender Bedeutung ist deshalb das Wissen, das man über den Gegenstandsbereich hat, auf den sich die Hypothesen beziehen. Infolgedessen ist es auch angebracht, skeptisch gegenüber all denjenigen Versuchen zu sein, die eine allgemeine Theorie epistemischer Wahrscheinlichkeit anstreben, ohne auf einen jeweils bestimmten Gegenstandsbereich bezug zu nehmen. Daß auch in diesem Kapitel in sehr unspezifischer Weise von Hypothesen gesprochen wird, ohne auf einen bestimmten Gegenstandsbereich bezug zu nehmen, soll zunächst auch nur dazu dienen, ein allgemeines Vorverständnis für die Problematik epistemischer Wahrscheinlichkeitsaussagen zu gewinnen. In späteren Kapiteln wird es sich als wichtig erweisen, jeweils genau auf die Art der Hypothesen zu achten, die mit Wahrscheinlichkeitsaussagen einschätzbar gemacht werden sollen.

3.5 Komparative Ordnung von Hypothesen

1. Zwei Fragen sollten unterschieden werden: wie man Wahrscheinlichkeitsaussagen *verstehen* und wie man sie *begründen* kann. Eine solche Unterscheidung erleichtert es auch, ein Vorurteil zu vermeiden, daß in der Literatur zur Wahrscheinlichkeitstheorie häufig auftritt; nämlich die Auffassung, daß eine Definition von 'wahrscheinlich' oder 'Wahrscheinlichkeit' durch die Angabe eines Verfahrens erfolgen sollte, mit dem Wahrscheinlichkeitsaussagen begründet werden können. Meistens wird diese Auffassung zugleich mit dem Postulat verknüpft, daß 'Wahrscheinlichkeit' von vornherein und in allen Anwendungsgebieten als ein (metrisierbarer) Größenbegriff aufgefaßt werden kann. Die Frage, wie Wahrscheinlichkeitsaussagen verstanden werden können, wird dann mit einer ganz anderen Frage vermischt: wie – durch eine Verfahren? – (numerisch bestimmte) Wahrscheinlichkeiten ermittelt werden können.

2. Wie schon gesagt worden ist, wird in diesem Text die Auffassung vertreten, daß numerische Wahrscheinlichkeitsaussagen nur in einigen spezifischen Kontexten sinnvoll sind (mit denen wir uns später beschäftigen werden) und daß stattdessen zunächst – im allgemeinen – von einem komparativen Verständnis von Wahrscheinlichkeitsaussagen ausgegangen werden sollte. Die Frage, wie Wahrscheinlichkeitsaussagen zu verstehen sind, kann dann so beantwortet werden, wie bereits in Abschnitt 3.3

angedeutet worden ist: Die Aussage

$$H \preceq_e H'$$

soll bedeuten, daß die Gründe, die für die Hypothese H' sprechen, mindestens so gut und überzeugend sind, wie die Gründe, die für die Hypothese H sprechen. Das ist ersichtlich keine operationale Definition. Diese Erläuterung des Sinns von Wahrscheinlichkeitsaussagen liefert uns keine Methode, um bei zwei beliebig vorgegebenen Hypothesen entscheiden zu können, ob die eine oder die andere wahrscheinlicher ist. Dennoch liefert sie ein Verständnis dessen, was mit einer Wahrscheinlichkeitsaussage gemeint ist. Diese Erläuterung impliziert im übrigen, daß sich Wahrscheinlichkeitsaussagen nicht auf „mentale Zustände", sondern auf Begründungen für Hypothesen beziehen.

3. Redet man in sehr allgemeiner Weise von Wahrscheinlichkeitsaussagen, erscheint es möglich, daß man ihre Sprachform zum Vergleich beliebiger Hypothesen verwenden kann. Dann wird jedoch ziemlich rätselhaft, wie eine komparative Einschätzung von Gründen, die für oder gegen die Hypothesen angeführt werden können, sinnvoll erfolgen kann. Wir werden deshalb unsere weiteren Überlegungen stets auf ein vorgegebenes System von Hypothesen beziehen, so daß eine Situation (oder eine bestimmte Menge von Situationen) fixiert werden kann, auf die sich die Hypothesen konkurrierend beziehen. Dafür wird die in Abschnitt 2.3 eingeführte Notation:

$$\mathcal{H} := \{ \langle s, \tilde{k} \rangle \mid \tilde{k} \in \tilde{K} \}$$

verwendet. Soweit es sich nur um endlich viele Hypothesen handelt, wird zur Abkürzung von Schreibweisen auch die Notation

$$\mathcal{H} := \{ H_1, \ldots, H_m \}$$

verwendet.

4. Hat man als Ausgangspunkt ein System von Hypothesen $\mathcal{H}$, geht es im weiteren um die Formulierung und Begründung von Wahrscheinlichkeitsaussagen. Unter *elementaren* Wahrscheinlichkeitsaussagen verstehen wir Aussagen der Form $H \preceq_e H'$, wobei H und H' Hypothesen aus $\mathcal{H}$ sind. Daran anschließend können zwei weitere Sprachformen für elementare Wahrscheinlichkeitsaussagen eingeführt werden. Zunächst:

$$H \equiv_e H' \iff H \preceq_e H' \text{ und } H' \preceq_e H$$

was auch so verstanden werden kann, daß es (zur Zeit) keine Gründe gibt, eine der beiden Hypothesen für besser begründet zu halten als die andere. Weiterhin kann definiert werden:

$$H \prec_e H' \iff H \preceq_e H' \text{ und nicht } H' \equiv_e H$$

womit gemeint ist, daß die verfügbaren Gründe die Hypothese H' wahrscheinlicher machen als die Hypothese H.

5. Bezieht man sich in einer Diskussion auf ein System von Hypothesen $\mathcal{H}$, besteht ein sinnvolles Ziel darin, die Hypothesen mithilfe der Relationen $\preceq_e$, $\prec_e$ und $\equiv_e$ in eine Rangordnung zu bringen. Es kann jedoch nicht vorausgesetzt werden, daß dies stets gelingt. Es kann unklar sein, ob bei zwei Hypothesen ein sinnvoller Wahrscheinlichkeitsvergleich vorgenommen werden kann; außerdem können bei einer Diskussion unterschiedliche Auffassungen vertreten werden. Weiterhin ist es auch nicht unter allen Umständen sinnvoll, nach einer definitiven Wahrscheinlichkeitsaussage zu suchen, wie Keynes (1921, S. 30) mit folgendem Beispiel deutlich gemacht hat:

> "Is our expectation of rain, when we start out for a walk, always *more* likely than not, or *less* likely than not, or *as* likely as not? I am prepared to argue that on some occasions *none* of these alternatives hold, and that it will be an arbitrary matter to decide for or against the umbrella. If the barometer is high, but the clouds are black, it is not always rational that one should prevail over the other in our minds, or even that we should balance them, – though it will be rational to allow caprice to determine us and to waste no time on the debate."

Somit sollte im allgemeinen davon ausgegangen werden, daß uns durch eine Menge *begründbarer* Wahrscheinlichkeitsaussagen nur eine partielle Ordnung für die Hypothesen aus $\mathcal{H}$ verfügbar gemacht wird.[15]

6. Aber auch wenn es gelingt, alle Hypothesen aus $\mathcal{H}$ im Hinblick auf ihre Wahrscheinlichkeit in eine begründbare Ordnung zu bringen, erhält man dadurch nicht zugleich auch eine vollständige Ordnung für eine aus $\mathcal{H}$ konstruierbare Algebra für Hypothesen $\mathcal{A}(\mathcal{H})$. Zwar ist dies trivialerweise der Fall, wenn $\mathcal{H}$ nur zwei Hypothesen enthält. Sobald es indessen mehr als zwei Hypothesen gibt, erhält man aus einer Rangordnung für $\mathcal{H}$ nur eine partielle Ordnung für $\mathcal{A}(\mathcal{H})$. Man denke zum

[15] Es sei aber angemerkt, daß in der Literatur zur Wahrscheinlichkeitstheorie – soweit überhaupt von einem komparativen Verständnis von Wahrscheinlichkeitsaussagen ausgegangen wird – zumeist ohne Begründung unterstellt wird, daß es eine vollständige Ordnung für Wahrscheinlichkeitsvergleiche beliebiger Aussagen gibt. Vgl. z.B. Lenzen 1980, S. 82f.

Beispiel an die Hypothesen $H \vee H'$ und $H' \vee H''$. Wie kann über sie eine Wahrscheinlichkeitsaussage begründet werden? Auch eine Kenntnis aller elementaren Wahrscheinlichkeitsaussagen würde dafür keine hinreichende Grundlage liefern.

3.6 Elementare Konsistenzbedingungen

1. Die zuvor angeführte Überlegung führt zu einer allgemeineren Fragestellung: Welche Regeln können für das Argumentieren mit komparativen Wahrscheinlichkeiten vorausgesetzt werden? Einige dieser Regeln, die als Konsistenzbedingungen sinnvollen Argumentierens verstanden werden können, sollen im folgenden besprochen werden.

2. Zunächst kann sicherlich die Transitivität der Relationen $\preceq_e$, $\prec_e$ und $\equiv_e$ vorausgesetzt werden. Verwendet man exemplarisch die Relation $\preceq_e$, lautet diese Regel:

$$\text{Für alle } H, H', H'' \in \mathcal{A}(\mathcal{H}): \quad H \preceq_e H',\ H' \preceq_e H'' \implies H \preceq_e H'' \tag{3.6.1}$$

Es sollte indessen betont werden, daß es sich hierbei nicht um eine irgendwie gegebene Eigenschaft der Relation $\preceq_e$ handelt, sondern um eine Regel für sinnvolles Argumentieren.

3. Eine zweite Regel betrifft den Vergleich einer durch $\vee$ zusammengesetzten Hypothese mit ihren Komponenten:

$$\text{Für alle } H, H' \in \mathcal{A}(\mathcal{H}): \quad H \preceq_e (H \vee H') \quad \text{und} \quad H' \preceq_e (H \vee H') \tag{3.6.2}$$

Denn alle Gründe, die für H sprechen, sprechen auch für $(H \vee H')$. Und entsprechend: alle Gründe, die für H' sprechen, sprechen ebenfalls für $(H \vee H')$.

4. Mit einer analogen Überlegung kann man sich klarmachen, daß folgende Regel sinnvoll ist:

$$\text{Für alle } H, H' \in \mathcal{A}(\mathcal{H}): \quad (H \wedge H') \preceq_e H \quad \text{und} \quad (H \wedge H') \preceq_e H' \tag{3.6.3}$$

Denn jeder Grund, der für $(H \wedge H')$ angeführt werden, kann sicherlich auch für H und ebenfalls für H' angeführt werden.

5. Schließlich gibt es in einer Algebra für Hypothesen noch zwei Sonderfälle, die sich für eine Formulierung von Regeln verwenden lassen. Man kann einerseits die Hypothese

$$H^s := H_1 \vee \cdots \vee H_m$$

bilden. Da $\mathcal{H}$ als ein vollständiges System von Hypothesen vorausgesetzt worden ist, ist diese Hypothese stets zutreffend; und sie wird deshalb gelegentlich als „sichere Hypothese“ bezeichnet. Somit folgt aus Regel (3.6.2) auch die folgende Regel:

$$\text{Für alle } H \in \mathcal{A}(\mathcal{H}): \ H \preceq_e H^s \tag{3.6.4}$$

Andererseits kann man auch eine Hypothese

$$H^u := \neg H^s$$

bilden, die offenbar niemals zutreffend sein kann.[16] Und man gewinnt mithin die Regel

$$\text{Für alle } H \in \mathcal{A}(\mathcal{H}): \ H^u \preceq_e H \tag{3.6.5}$$

Man kann den Sinn dieser beiden Regeln auch so ausdrücken: keine Hypothese kann weniger wahrscheinlich sein als eine Hypothese, die niemals zutreffen kann; und keine Hypothese kann wahrscheinlicher sein als eine Hypothese, die immer zutreffend ist.

6. Es sei angemerkt, daß in der Literatur noch weitere Regeln für Wahrscheinlichkeitsrelationen vorgeschlagen worden sind; meistens mit dem Ziel, hinreichend viele Regeln zu haben, um mit ihrer Hilfe einen *quantitativen* Wahrscheinlichkeitsbegriff definieren zu können.[17] Im Unterschied zu den oben angeführten Regeln haben jedoch die für einen quantitativen Wahrscheinlichkeitsbegriff zusätzlich erforderlichen Regeln keine intuitive Plausibilität; und wir werden uns hier deshalb mit Versuchen, eine „axiomatische Begründung“ für einen quantitativen Wahrscheinlichkeitsbegriff zu finden, nicht näher beschäftigen.

[16] Hypothesen dieser Art können auf unterschiedliche Weisen formuliert werden. Zum Beispiel ist auch $H \wedge H'$ eine Hypothese, die niemals zutreffend sein kann, wenn H und H' elementare Hypothesen aus $\mathcal{H}$ sind. Der erforderliche formale Aufwand, um alle Hypothesen dieser Art zu identifizieren, lohnt hier jedoch nicht.

[17] Der erste Versuch dieser Art wurde von Koopman (1940) ausgedacht; weitere Hinweise finden sich bei Lenzen 1980, Kap. 4.

3.7 Bewertungen und Wahrscheinlichkeiten

Es wurde schon darauf hingewiesen, daß sich ein großer Teil der in der Literatur zur Wahrscheinlichkeitstheorie und mathematischen Statistik unternommenen Anstrengungen darauf richtet, quantitative Wahrscheinlichkeitsmaße zu konstruieren. In diesem Abschnitt soll überlegt werden, wie sich Bemühungen dieser Art mit einem qualitativen und komparativen Verständnis von Wahrscheinlichkeitsaussagen verbinden lassen.

3.7.1 Additive Bewertungsfunktionen

1. Im Mittelpunkt vieler Bestrebungen in der Wahrscheinlichkeitstheorie steht folgende Aufgabe: es soll eine Funktion

$$\mathrm{P_e} : \mathcal{A}(\mathcal{H}) \longrightarrow \mathbf{R}$$

konstruiert werden, mit der jeder Hypothese H aus einem System von Hypothesen $\mathcal{A}(\mathcal{H})$ eine numerisch bestimmte Wahrscheinlichkeit $\mathrm{P_e}(H)$ zugeordnet werden kann. Dabei wird fast immer vorausgesetzt, daß diese Funktion einigen elementaren Regeln genügen soll; und zwar soll für alle Hypothesen $H, H' \in \mathcal{A}(\mathcal{H})$ gelten:[18]

a) $0 \leq \mathrm{P_e}(H) \leq 1$

b) $\mathrm{P_e}(H^s) = 1$

c) $\mathrm{P_e}(\neg H) = 1 - \mathrm{P_e}(H)$

d) $H \wedge H' = H^u \implies \mathrm{P_e}(H \vee H') = \mathrm{P_e}(H) + \mathrm{P_e}(H')$

$\mathrm{P_e}$ wird dann ein *Wahrscheinlichkeitsmaß* für das System von Hypothesen $\mathcal{A}(\mathcal{H})$ genannt.

2. Diese Regeln für ein Wahrscheinlichkeitsmaß entsprechen offenbar denjenigen, die für statistische Häufigkeitsfunktionen gelten (vgl. Abschnitt 2.9). Zwar bezieht sich eine statistische Häufigkeitsfunktion $\mathrm{P_S}$ auf die Häufigkeit von Merkmalen in einer Menge von Situationen S, hingegen ein Wahrscheinlichkeitsmaß $\mathrm{P_e}$ auf die Wahrscheinlichkeit von Hypothesen;[19] bei einer rein formalen Betrachtung kann man sich jedoch in beiden Fällen auf eine Algebra für Mengen beziehen und dann

[18] Zur Definition von H^s und H^u vgl. Abschnitt 3.6. Mit der Formulierung $H \wedge H' = H^u$ ist gemeint, daß sich die Hypothesen H und H' ausschließen.

[19] Im hier vorliegenden Kontext; im nächsten Kapitel werden Wahrscheinlichkeitsmaße für Zufallsgeneratoren besprochen und durch $\mathrm{P_a}$ symbolisiert.

annehmen, daß es eine Mengenfunktion gibt, die den oben angegebenen Regeln genügt. Mengenfunktionen dieser Art werden in diesem Text *additive Bewertungsfunktionen* genannt. Diese Bezeichnung erscheint sinnvoll, um nicht von vornherein den Anspruch zu unterstellen, daß mit der Konstruktion einer solchen Funktion eine Quantifizierung epistemischer Wahrscheinlichkeiten angestrebt wird.

3. Der Vorschlag, eine mathematische Wahrscheinlichkeitsrechnung als eine Theorie additiver Bewertungsfunktionen zu konzipieren, stammt (in der Form, wie er inzwischen weitgehend akzeptiert worden ist) von A. Kolmogoroff (1933). Sein Ziel bestand darin, eine Wahrscheinlichkeitsrechnung axiomatisch und rein formal einzuführen, um sie dadurch von den inhaltlichen Kontroversen um das Verständnis von Wahrscheinlichkeitsaussagen ablösen und als ein Teilgebiet der Mathematik begründen zu können. In der Einleitung zu seiner Arbeit schrieb Kolmogoroff:

„Zweck des vorliegenden Heftes ist eine axiomatische Begründung der Wahrscheinlichkeitsrechnung. Der leitende Gedanke des Verfassers war dabei, die Grundbegriffe der Wahrscheinlichkeitsrechnung, welche noch unlängst für ganz eigenartig galten, natürlicherweise in die Reihe der allgemeinen Begriffsbildungen der modernen Mathematik einzuordnen.“ (S. III)

„Die Wahrscheinlichkeitstheorie als mathematische Disziplin soll und kann genau in demselben Sinne axiomatisiert werden wie die Geometrie oder die Algebra. Das bedeutet, daß, nachdem die Namen der zu untersuchenden Gegenstände und ihrer Grundbeziehungen sowie die Axiome, denen diese Grundbeziehungen zu gehorchen haben, angegeben sind, die ganze weitere Darstellung sich ausschließlich auf diese Axiome gründen soll und keine Rücksicht auf die jeweilige konkrete Bedeutung dieser Gegenstände und Beziehungen nehmen darf.“ (S. 1)

Nur ein kurzer Abschnitt in Kolmogoroffs Arbeit beschäftigt sich mit einer inhaltlichen Motivation des Vorschlags, von der Konzeption einer additiven Bewertungsfunktion auszugehen; und zur Begründung wird im wesentlichen nur auf die formale Analogie zu einer relativen Häufigkeitsfunktion hingewiesen. In dem Abschnitt mit der Überschrift „Das Verhältnis zur Erfahrungswelt“, in dem diese Überlegung ausgeführt wird, heißt es gleich zu Beginn (S. 3):

„Ein Leser, der sich nur für die rein mathematische Entwicklung der Theorie interessiert, braucht diesen Paragraphen nicht zu lesen – die weitere Darstellung beruht auf den Axiomen des § 1 und benutzt nicht die Überlegungen des gegenwärtigen Paragraphen. In diesem wollen wir uns mit dem bloßen Hinweis auf die empirische Entstehung der Axiome der Wahrscheinlichkeitsrechnung begnügen und lassen deshalb eine eingehende philosophische Untersuchung des Wahrscheinlichkeitsbegriffes in der Erfahrungswelt bewußt beiseite.“

Im weiteren werden dann über das „Verhältnis zur Erfahrungswelt“ die folgenden Bemerkungen gemacht:

„1. Es wird ein gewisser Komplex $\mathfrak{S}$ von Bedingungen vorausgesetzt, welcher unbeschränkter Wiederholung fähig ist.

2. Man untersucht einen bestimmten Kreis von Ereignissen, welche infolge der Realisation der Bedingungen $\mathfrak{S}$ entstehen können. In den einzelnen Fällen der Realisation der Bedingungen $\mathfrak{S}$ verlaufen die erwähnten Ereignisse im allgemeinen auf verschiedene Weisen. Es sei E die Menge der verschiedenen möglichen Varianten $\xi_1, \xi_2, \ldots$ des Verlaufes der besagten Ereignisse. Einige unter diesen Varianten brauchen dabei überhaupt nicht zur Realisation zu gelangen. Wir nehmen in die Menge E alle Varianten auf, die wir a priori für möglich erachten.

3. Wenn die nach der Realisation der Bedingungen $\mathfrak{S}$ praktisch aufgetretene Variante unserer Ereignisse zu der (durch irgendwelche Bedingungen definierten) Menge A gehört, so sagen wir, daß das Ereignis A stattgefunden hat. [Als Beispiel wird dann das Werfen einer Münze erwähnt.]

4. Unter gewissen Bedingungen, auf die wir hier nicht näher eingehen wollen, kann man voraussetzen, daß einem Ereignis A, welches infolge der Bedingungen $\mathfrak{S}$ auftritt oder nicht, eine gewisse reelle Zahl $\mathrm{P}(A)$ zugeordnet ist, welche folgende Eigenschaften besitzt:

A. Man kann praktisch sicher sein, daß, wenn man den Komplex der Bedingungen $\mathfrak{S}$ eine große Anzahl von n Malen wiederholt und dabei durch m die Anzahl der Fälle bezeichnet, bei denen das Ereignis A stattgefunden hat, das Verhältnis m/n sich von $\mathrm{P}(A)$ nur wenig unterscheidet.

B. Ist $\mathrm{P}(A)$ sehr klein, so kann man praktisch sicher sein, daß bei einer einmaligen Realisation der Bedingungen $\mathfrak{S}$ das Ereignis A nicht stattfindet.“

Der auf diese Weise hergestellt Zusammenhang zwischen $\mathrm{P}(A)$ und relativen Häufigkeiten wird schließlich verwendet, um zu motivieren, daß P axiomatisch als eine additive Bewertungsfunktion konzipiert werden sollte, d.h. daß für P die gleichen Regeln wie für relative Häufigkeiten gelten sollten.

3.7.2 Bewertungen von Hypothesen

1. Den Sinn einer axiomatischen Begründung kann man darin sehen, daß dadurch die Entwicklung einer *mathematischen* Wahrscheinlichkeitsrechnung von den grundlegenden und kontroversen Fragestellungen der

Wahrscheinlichkeits*theorie* abgekoppelt werden kann. Ein hervorragender Vertreter der mathematischen Wahrscheinlichkeitsrechnung hat das so gesagt:

“Now that probability theory has become an acceptable part of mathematics, words like ‘occurrence’, ‘event’, ‘urn’, ‘die’, and so on can be dispensed with. These words and the ideas they represent, however, still add intuitive significance to the subject and suggest analytical methods and new investigations. It is for this reason that probability language is still used even in purely theoretical investigations.” (Doob 1953, S. 1)

Wenn jedoch Überlegungen der Wahrscheinlichkeitstheorie zur Reflexion von Hypothesen verwendet werden sollen, die sich auf unsere Erfahrungswelt beziehen, ist der axiomatische Standpunkt der Mathematiker unzweckmäßig. Dann kommt es vielmehr darauf an, den möglichen Sinn von Wahrscheinlichkeitsaussagen zu verstehen.[20]

2. Noch ein weiterer Grund spricht dafür, bei unseren Überlegungen zum Verständnis und zur Begründbarkeit von Wahrscheinlichkeitsaussagen nicht vorauszusetzen, daß eine als Wahrscheinlichkeitsmaß verwendbare additive Bewertungsfunktion immer schon zur Verfügung steht. Folgt man nämlich der von Kolmogoroff angegebenen Motivation für sein Axiomensystem, wie es im vorangegangenen Abschnitt kurz angedeutet worden ist, erkennt man, daß es nur innerhalb eines sehr spezifischen Verwendungszusammenhangs eine sinnvolle Interpretation erlaubt: wenn man es mit Experimenten zu tun hat, die unter gleichen Bedingungen beliebig oft wiederholbar sind. Kolmogoroff setzt also ein sehr spezifisches „Verhältnis zur Erfahrungswelt“ voraus, um zu zeigen, daß mit den begrifflichen Hilfsmitteln seines Axiomensystems auch Wahrscheinlichkeitsaussagen (und nicht nur mathematische Sätze über additive Bewertungsfunktionen) formuliert werden können. Dieser Verwendungszusammenhang kann jedoch bei den meisten Wahrscheinlichkeitsaussagen nicht vorausgesetzt werden.

3. Man kann sich infolgedessen die Frage stellen, ob die in der Wahrscheinlichkeitsrechnung und mathematischen Statistik betriebenen Bemühungen um eine Konstrukton von Bewertungsfunktionen, die sich als

[20]Dazu sei folgende Bemerkung von R. von Mises (1990, S. 226) angeführt: „Es läßt sich natürlich nichts dagegen sagen, daß sich jemand mit dem rein tautologischen Teil der Theorie eines Sachgebietes befaßt und dafür eine zusammenfassende Darstellung gibt, ohne auf den empirischen Ausgangspunkt zurückzugreifen. Nur darf er dann nicht die Resultate der Rechnung in einer Form aussprechen, die den Eindruck erweckt, als würden diese Überlegungen Aussagen über bestimmte wirkliche Vorgänge gewonnen.“

Wahrscheinlichkeitsmaße interpretieren lassen oder aus ihnen abgeleitet werden können, zum Verständnis und zur Begründung *epistemischer* Wahrscheinlichkeitsaussagen überhaupt relevant sind. Aus der in diesem Text vertretenen Auffassung, daß epistemische Wahrscheinlichkeiten nicht quantifizierbar sind, sondern bestenfalls komparative Vergleiche erlauben, scheint eine negative Antwort zu folgen. Diese negative Antwort bezieht sich jedoch nur darauf, daß die Idee epistemischer Wahrscheinlichkeitsaussagen durch ein quantitatives Wahrscheinlichkeitsmaß nicht erschöpfend erfaßt werden kann. Die Möglichkeit, daß quantitative Bewertungsfunktionen gleichwohl *zur Beurteilung* von Wahrscheinlichkeitsaussagen dienen können, sollte jedenfalls nicht von vornherein ausgeschlossen werden.

4. Angenommen, man kann sich auf eine Algebra für Hypothesen $\mathcal{A}(\mathcal{H})$ beziehen. Nach der in diesem Text vertretenen Auffassung erscheint es in den meisten Fällen aussichtslos, ein Wahrscheinlichkeitsmaß zu finden, mit dem folgende Behauptung begründet werden könnte:

$$\mathrm{P_e}(H) \leq \mathrm{P_e}(H') \iff H \preceq_e H' \tag{3.7.1}$$

Dies folgt aus unserem Verständnis von Wahrscheinlichkeitsaussagen der Form „$H \preceq_e H'$“: daß mit ihnen gemeint ist, daß es für H' bessere Gründe gibt als für H. Denn ob dies der Fall ist, hängt stets von einem explizit darzulegenden Begründungszusammenhang ab, der nicht situationsunabhängig fixiert und reguliert werden kann.

5. Dennoch ist es nicht ausgeschlossen, daß Bewertungsfunktionen konstruiert werden können, die sich mit Wahrscheinlichkeitsaussagen auf folgende Weise in einen Zusammenhang bringen lassen:

$$\mathrm{P_e}(H) \leq \mathrm{P_e}(H') \rightsquigarrow H \preceq_e H' \tag{3.7.2}$$

womit gemeint sein soll: die Aussage „$\mathrm{P_e}(H) \leq \mathrm{P_e}(H')$“ liefert *ein Argument* zur Begründung der Wahrscheinlichkeitsaussage „$H \preceq_e H'$“. Ein solches Verständnis muß offenbar von einer Auffassung, wie sie in (3.7.1) zum Ausdruck kommt, unterschieden werden. Denn (3.7.2) impliziert nicht den Anspruch, daß durch $\mathrm{P_e}$ eine erschöpfende Explikation epistemischer Wahrscheinlichkeit gegeben wird (ein Anspruch, der auch stets nur zu Kontroversen geführt hat); der Anspruch ist viel bescheidener: daß Aussagen, die mithilfe von $\mathrm{P_e}$ über ein System von Hypothesen gemacht werden können, bei der Einschätzung ihrer epistemsichen Wahrscheinlichkeiten *als ein Argument* verwendet werden können. Damit wird insbesondere nicht die Möglickeit ausgeschlossen, daß es zur

Beurteilung der fraglichen Hypothesen noch andere Gesichtspunkte und Argumente geben kann, die in der Konzeption von $\mathrm{P_e}$ nicht berücksichtigt worden sind. Weiterhin soll durch die Formulierung deutlich werden, daß es bei der Verwendung einer Bewertungsfunktion im Kontext einer Begründung epistemischer Wahrscheinlichkeitsaussagen darauf ankommt, welche substantiellen Argumente in der Form einer Bewertungsfunktion zum Ausdruck gebracht werden können.

6. Wenn wir uns in späteren Kapiteln auf Bewertungsfunktionen beziehen, die in der Literatur für Wahrscheinlichkeitsaussagen vorgeschlagen worden sind, werden wir uns an der durch (3.7.2) angedeuteten Auffassung orientieren. Dann wird sich auch zeigen, daß es sich dabei nicht immer um additive Bewertungsfunktionen im Sinne des Kolmogoroffschen Axiomensystems handelt. Wir werden deshalb den Ausdruck ‘Bewertungsfunktion’ in einem allgemeineren Sinne verwenden und jede für eine Algebra von Hypothesen definierte Funktion so nennen, wenn sie im Sinne des Argumentationsschemas (3.7.2) verwendet werden soll. Der Ausdruck ‘additive Bewertungsfunktion’ wird dagegen auch weiterhin so verwendet, wie er im vorangegangenen Abschnitt erläutert worden ist, d.h. als eine rein formale Konstruktion, mit der nicht eo ipso eine bestimmte Interpretation verbunden ist. Nur wenn eine solche Funktion explizit mit dem Anspruch formuliert wird, daß sie im Sinne von (3.7.2) verwendet werden kann, sprechen wir von einem epistemischen Wahrscheinlichkeitsmaß. Davon zu unterscheiden sind schließlich noch aleatorische Wahrscheinlichkeitsmaße, die den Ausgangspunkt für eine aleatorische Wahrscheinlichkeitsrechnung bilden; sie werden in Kapitel 4 besprochen.

3.8 Ergänzungen und Exkurse

Wie bereits die Bemerkungen in Abschnitt 3.4 gezeigt haben, gibt es unterschiedliche Verständnisse epistemischer Wahrscheinlichkeitsaussagen. In den folgenden Abschnitten werden deshalb noch einige ergänzende Überlegungen zur Diskussion gestellt, die vielleicht helfen können, die in diesem Text vertretene Auffassung – daß sich Wahrscheinlichkeitsaussagen auf *Begründungen* von Hypothesen beziehen – verständlicher zu machen.

3.8.1 Begründung durch Verfahren?

1. In Abschnitt 3.4 haben wir unser Verständnis epistemischer Wahrscheinlichkeitsaussagen von „subjektiven“ – im Sinne von psychologischen – Charakterisierungen von ‘Wahrscheinlichkeit’ abgegrenzt. Demgegenüber wurde betont, daß mit Wahrscheinlichkeitsaussagen ein komparatives Urteil über Begründungen von Hypothesen formuliert wird. Aus diesem Verständnis erhält man auch einen Hinweis, wie Wahrscheinlichkeitsaussagen zu begründen sind: Zur Begründung einer Wahrscheinlichkeitsaussage, mit der behauptet wird, daß eine Hypothese H besser begründet ist als eine Hypothese H', muß *dieser Anspruch* einsichtig gemacht werden.[21] Dies geschieht praktisch dadurch, daß Gründe, die für oder gegen die Hypothesen sprechen, erwogen und verglichen werden. Wie stets, wenn eine Behauptung begründet werden soll, ist eine Argumentation erforderlich, die die Behauptung einsichtig machen soll.

2. In der Geschichte der Wahrscheinlichkeitstheorie ist immer wieder der Versuch gemacht worden, für Argumentationen dieser Art Prinzipien anzugeben; oft in der Form von Regeln oder Verfahren, mit denen sich – nach der Intention ihrer Urheber – entscheiden läßt, ob eine Hypothese wahrscheinlicher ist als eine andere Hypothese. Mit einigen Prinzipien dieser Art werden wir uns in späteren Kapiteln genauer beschäftigen. Bereits an dieser Stelle ist es jedoch sinnvoll, auf einige Schwierigkeiten hinzuweisen.

3. Eine erste Schwierigkeit ist unmittelbar einsichtig, wenn man von dem hier vorgeschlagenen Verständnis von Wahrscheinlichkeitsaussagen ausgeht. Sie beziehen sich dann nämlich auf *Begründungen*, die in komparativer Weise für oder gegen Hypothesen angeführt werden können. Begründungen, die dem Zweck dienen sollen, eine Behauptung einsichtig zu machen, können jedoch nicht durch ein formal handhabbares Entscheidungsverfahren ersetzt werden.[22] Dies ist schon deshalb nicht

[21] Wenn eine Aussage sich unmittelbar auf unsere Erfahrungswelt bezieht, könnte man die Aufgabe, sie zu begründen, so formulieren: es muß gezeigt werden, daß es sich *tatsächlich* so verhält, wie mit der Aussage behauptet wird. Wahrscheinlichkeitsaussagen beziehen sich jedoch nicht unmittelbar auf die Erfahrungswelt, sondern auf Begründungen für Hypothesen. Dann aber macht es keinen Sinn, davon zu sprechen, daß eine Hypothese *tatsächlich* besser begründet ist als eine andere Hypothese; denn Begründungen sind keine Tatsachen.

[22] Dies gilt jedenfalls dann, wenn sich die zu begründende Behauptung direkt oder – wie bei Wahrscheinlichkeitsaussagen – indirekt auf unsere Erfahrungswelt bezieht. Ob bzw. in welcher Weise es sich in der Mathematik anders verhält, braucht hier nicht erörtert zu werden.

möglich, weil Begründungen kontextabhängig sind und sich auf ein jeweils verfügbares Wissen beziehen müssen, das nur begrenzt situationsunabhängig und schon gar nicht rein formal charakterisiert werden kann.

4. Diese Feststellung ist allerdings nur dann unmittelbar einsichtig, wenn man sich an dem hier vertretenen Verständnis orientiert, daß sich Wahrscheinlichkeitsaussagen auf Gründe beziehen, die für oder gegen Hypothesen angeführt werden können. In Abschnitt 3.5 wurde bereits darauf hingewiesen, daß in der Wahrscheinlichkeitstheorie oft eine andere Auffassung vertreten wird: daß der Sinn von Wahrscheinlichkeitsaussagen durch Verfahren definiert werden sollte, mit denen entschieden werden kann, ob eine Wahrscheinlichkeitsaussage richtig (oder „am besten begründet“) ist. Dieser Auffassung entspricht, daß man auch die Prinzipien oder Verfahren, die zur Begründung von Wahrscheinlichkeitsaussagen vorgeschlagen worden sind, als jeweils spezifische Definitionen der Bedeutung oder des Sinns von Wahrscheinlichkeitsaussagen auffassen kann. Aber die Konsequenzen dieser Auffassung sind ersichtlich unbefriedigend; denn für jedes unterschiedliche Begründungsprinzip gäbe es dann einen jeweils spezifischen Wahrscheinlichkeitsbegriff.

5. Deshalb erscheint es mir sinnvoller, daran festzuhalten, daß der Sinn von Wahrscheinlichkeitsaussagen vorgängig und unabhängig von Verfahren expliziert werden sollte, die – möglicherweise – helfen können, Wahrscheinlichkeitsaussagen zu begründen. Aus der oben vertretenen Auffassung, daß die Aufgabe, Wahrscheinlichkeitsaussagen zu begründen, nicht an formal handhabbare Entscheidungsverfahren delegiert werden kann, folgt natürlich nicht, daß Prinzipien oder Regeln, die sich auf Argumentationen beziehen, überflüssig oder sinnlos sind. Es kann jedenfalls nicht von vornherein ausgeschlossen werden, daß solche Prinzipien oder Regeln *helfen* können, durchaus sinnvolle Argumentationsmöglichkeiten aufzuzeigen. Es handelt sich dann jedoch nicht um Entscheidungsverfahren, sondern um Ressourcen für Argumentationsmöglichkeiten, die *im Kontext* von Begründungen für Wahrscheinlichkeitsaussagen sinnvoll werden können. Das sollte auch durch das Argumentationsschema (3.7.2) zum Ausdruck gebracht werden.

3.8.2 Wahrscheinlichkeit als Relation?

1. Ein Verständnis der Aufgabe, Wahrscheinlichkeitsaussagen zu begründen, hängt davon ab, wie Wahrscheinlichkeitsaussagen (was durch sie behauptet werden soll) verstanden werden können. Ich möchte hier, auch zur Unterstützung der Überlegungen des vorangegangenen Abschnitts, noch einmal den Vorschlag verteidigen, daß epistemische Wahrscheinlichkeitsaussagen als komparative Aussagen über Gründe für Hypothesen aufgefaßt werden sollten. Denn es gibt eine alternative Konzeption, die den Gedanken nahelegt, daß Begründungen für Wahrscheinlichkeitsaussagen in die Form eines Kalküls gebracht werden könnten.

2. Der Grundgedanke dieser alternativen Konzeption besteht darin, Wahrscheinlichkeitsaussagen als Aussagen über eine „epistemische Relation" zwischen einer Hypothese und einem für ihre Begründung verfügbaren Wissen zu konzipieren. Dieser Gedanke wurde zum Beispiel von Keynes (1921, S. 3f) so formuliert:

> "All propositions are true or false, but the knowledge we have of them depends on our circumstances; and while it is often convenient to speak of propositions as certain or probable, this expresses strictly a relationship in which they stand to a *corpus* of knowledge, actual or hypothetical, and not a characteristic of the propositions in themselves. A proposition is capable at the same time of varying degrees of this relationship, depending upon the knowledge to which it is related, so that it is without significance to call a proposition probable unless we specify the knowledge to which we are relating it."

Keynes hat deshalb vorgeschlagen, Wahrscheinlichkeitsaussagen in der Form H/W zu schreiben, wobei H auf eine (hypothetische) Aussage und W auf eine als (hypothetisches) Wissen voraussetzbare Menge von Aussagen verweist. Die Idee besteht insoweit darin, Wahrscheinlichkeitsaussagen als Aussagen über eine „epistemische Relation" zwischen W und H aufzufassen; in den Worten von Keynes (1921, S. 6):

> "Between two sets of propositions, therefore, there exists a relation, in virtue of which, if we know the first, we can attach to the latter some degree of rational belief. This relation is the subject-matter of the logic of probability."

3. Eine sehr ähnliche Konzeption wurde auch von Reichenbach (1949, S. 41f) vorgeschlagen:

> „Betrachten wir irgendeine Wahrscheinlichkeitsaussage, z.B. die Behauptung, daß beim Würfeln das Auftreten der Seite 1 mit der Wahrscheinlichkeit 1/6 zu erwarten ist, so stellen wir fest, daß diese Behauptung die logische Form einer *Relation* hat. Es wird nämlich das Auftreten der Seite 1 keineswegs schlechthin mit der Wahrscheinlichkeit 1/6 behauptet, sondern diese Behauptung wird an eine Bedingung geknüpft, die Bedingung nämlich, daß der Würfel geworfen wird. Wenn man den Würfel wirft, so ist mit der Wahrscheinlichkeit 1/6 das Auftreten der Seite 1 zu erwarten – dies ist die Form, in der wir die Wahrscheinlichkeitsaussage aussprechen; und niemand würde behaupten, daß die Wahrscheinlichkeit, einen Würfel mit obenliegender Seite 1 auf dem Tisch zu finden,

> 1/6 auch dann beträgt, wenn gar nicht geworfen wird. Die Wahrscheinlichkeit ist deshalb vom Charakter einer Implikation; sie enthält ein Vorderglied und ein Hinterglied, und zwischen diesen wird die Relation der Wahrscheinlichkeit behauptet. Wir nennen diese Relation *Wahrscheinlichkeitsimplikation* und stellen sie dar durch das Zeichen
>
> $$\underset{p}{\mathrel{-\!\!\!\supset}}$$
>
> Dies ist das einzige neue Zeichen, welches die Wahrscheinlichkeitsrechnung zu den Zeichen des Logikkalküls hinzunimmt. Seine Verwandtschaft zu der strengen Implikation der Logik soll durch die Form angedeutet werden, indem wir das Zeichen der strengen Implikation mit einem Strich kreuzen. Während die strenge Implikation Aussagen von der Form macht: 'wenn a wahr ist, dann ist b wahr' macht die Wahrscheinlichkeitsimplikation Aussagen der Form 'wenn a wahr ist, ist b wahrscheinlich vom Grade p'."

Zwar schränkt Reichenbach den Bereich möglicher Aussagen, die durch eine Wahrscheinlichkeitsimplikation verknüpft werden können, im weiteren ein; zugelassen sind bei ihm nur „Aussagen über die Klassenzugehörigkeit von Elementen gewisser gegebener Folgen" (ebda., S. 43). Dennoch ist der Grundgedanke mit dem von Keynes vergleichbar: beide Autoren konzipieren Wahrscheinlichkeitsaussagen in einer Analogie zur logischen Implikation.[23] Dies ist auch deshalb bemerkenswert, weil Reichenbach schließlich eine frequentistische Deutung von Wahrscheinlichkeitsaussagen vorschlägt. (Damit werden wir uns jedoch erst in Kapitel 6 etwas genauer beschäftigen.)

4. Auch die Notation, die wir in Abschnitt 3.3 eingeführt haben, nämlich

$$H \preceq_e H'$$

[23] Es gibt noch eine Reihe weiterer Varianten. Etwa kann auf Carnaps (1962) Versuche hingewiesen werden, eine „induktive Logik" zu entwickeln, um in quantifizierender Weise davon sprechen zu können, in welchem Ausmaß eine Menge von Aussagen durch eine andere Menge von Aussagen wahrscheinlich gemacht wird; Carnap spricht auch von einem „degree of confirmation" und von „induktiver Wahrscheinlichkeit"; eine Einführung findet sich bei Carnap und Stegmüller, 1959. Verbreitet ist der Gedanke, daß Wahrscheinlichkeit als eine Relation zwischen Aussagen konzipiert werden kann, auch in einer Reihe von „subjektiven" Wahrscheinlichkeitsdeutungen, die ihren Basisbegriff – „degree of belief" – relational deuten; exemplarisch kann auf Jeffreys (1948) hingewiesen werden: „Our fundamental idea will not be simply the probability of a proposition p, but the probability of p on data q." (S. 15) „Our primitive notion, then, is that of the relation 'given p, q is more probable than r', where p, q, and r are three propositions." (S. 16)

verweist auf eine Relation. Es ist jedoch wichtig, die in dieser Notation gemeinte Relation von derjenigen zu unterscheiden, die durch eine relationale Konzeption des Wahrscheinlichkeitsbegriffs vorstellbar gemacht werden soll. In unserer Notation bezieht sich eine Wahrscheinlichkeitsaussage auf zwei Hypothesen; ein Wahrscheinlichkeitsurteil soll zum Ausdruck bringen, welche der beiden Hypothesen besser begründet ist. Das jeweils verfügbare Wissen, das für die Begründung einer solchen Wahrscheinlichkeitsaussage zur Verfügung steht, wird in dieser Notation nicht explizit erwähnt. Bei der oben exemplarisch anhand von Keynes und Reichenbach erläuterten Konzeption geht es dagegen darum, in welchem Ausmaß eine Aussage durch eine andere Aussage wahrscheinlich gemacht wird. In diesem Fall wird also versucht, den *Sinn* von Wahrscheinlichkeitsaussagen durch eine „epistemische Relation“ zwischen Aussagen zu explizieren.

5. Die Unterscheidung wird dadurch etwas erschwert, daß Keynes (wie auch Jeffreys in der oben angeführten Fußnote) Wahrscheinlichkeitsaussagen zunächst in einer komparativen Weise formuliert; im Anschluß an die Keynes'sche Notation: $H/W < H'/W$. Hierbei wird jedoch angenommen, daß sowohl H/W als auch H'/W bereits sinnvolle sprachliche Ausdrücke sind: die Wahrscheinlichkeit von H in Relation zu W und die Wahrscheinlichkeit von H' in Relation zu W. Der Gedanke, daß man Wahrscheinlichkeit als eine „epistemische Relation“ auffassen könne, soll gerade dazu dienen, diese separaten Ausdrücke zu explizieren. Es handelt sich ersichtlich um einen entscheidenden Argumentationsschritt, wenn man schließlich eine Quantifizierung von Graden der Wahrscheinlichkeit einführen möchte. Denn ein numerisches Wahrscheinlichkeitsmaß kann nicht an unserer Relation $\preceq_e$ ansetzen, sondern nur an Ausdrücken der Form H/W. Die hier wichtige Frage besteht deshalb darin, ob sich bei Ausdrücken der Form H/W – ohne auf alternative Hypothesen bezug zu nehmen – sinnvoll von einer Wahrscheinlichkeit sprechen läßt.

6. Eine Diskussion dieser Frage ist natürlich sinnlos, wenn eine Theoriebildung damit beginnt, einen sprachlichen Ausdruck $P(H/W)$ als subjektiv quantifizierbaren „degree of belief“ zum Ausgangspunkt weiterer Begriffsbildungen zu machen. Der Vorschlag, daß man einen Ausdruck der Form $P(H/W)$ durch eine „epistemische Relation“ zwischen W und H explizierbar machen kann, ist dagegen gerade deshalb interessant, weil durch ihn auf das Problem aufmerksam gemacht wird: ob epistemische Wahrscheinlichkeiten quantifizierbar sind. Denn diesem Zweck dient schließlich die Behauptung, daß man von einer „epistemischen Relation“ zwischen W und H sprechen könne. Die Behauptung zielt darauf, daß

man für Wahrscheinlichkeitsaussagen die Existenz einer „epistemischen Relation“ zwischen W und H *als etwas durch W und H Gegebenes* voraussetzen könne.[24] Das war auch die Auffassung von Keynes, wie zum Beispiel folgende Bemerkung (1921, S. 4) deutlich macht:

> “A proposition [in unserer Notation H] is capable at the same time of varying degrees of this relationship, depending upon the knowledge [in unserer Notation W] to which it is related [...]. To this extent, therefore, probability may be called subjective. But in the sense important to logic, probability is not subjective. It is not, that is to say, subject to human caprice. A proposition is not probable because we think it so. When once the facts are given which determine our knowledge, what is probable or improbable in these circumstances has been fixed objectively, and is independent of our opinion.”

Die Annahme besteht also darin, daß man die „epistemische Relation“ zwischen W und H in Analogie zu einem objektiven Sachverhalt auffassen könne.[25]

7. Aber für die Annahme, daß eine „epistemische Relation“ zwischen W und H als ein (erkennbarer) *Sachverhalt* vorausgesetzt werden kann, gibt es keinerlei Begründung.[26] In dieser Formulierung ist die Annahme sogar dann fragwürdig, wenn H als eine logische Implikation aus W folgt, denn auch dann handelt es sich nicht um einen relationalen Sachverhalt, sondern um eine argumentativ herstellbare Beziehung zwischen Aussagen. In diesem Fall wären allerdings Wahrscheinlichkeitsaussagen überflüssig, denn H könnte durch W begründet werden. Wenn jedoch H keine logische Implikation von W ist, kann W bestenfalls ein unvollständiges Argument liefern, das für die Einschätzung der Wahrscheinlichkeit von H eine Rolle spielen kann. Aber auch diese Formulierung ist potentiell irreführend, insofern sie unterstellt, daß man eine einzelne

[24] Diese Annahme ist zwar nicht hinreichend, um schließlich auch noch eine Metrik einzuführen; aber sie bildet die entscheidende Voraussetzung, um epistemische Wahrscheinlichkeit als einen Größenbegriff konzipieren zu können.

[25] Das Motiv für diese Annahme wird daraus deutlich, wie das Zitat weiter geht: „The Theory of Probability is logical, therefore, because it is concerned with the degree of belief which it is *rational* to entertain in given conditions, and not merely with the actual beliefs of particular individuals, which may or may not be rational.“ Nicht nur Keynes, auch viele andere Autoren haben sich bemüht, ihre Überlegungen zur Bedeutung und Begründbarkeit von Wahrscheinlichkeitsaussagen mit einer jeweils spezifischen normativen Konzeption von „Rationalität“ zu verknüpfen. Aber die Wahrscheinlichkeitstheorie ist sicherlich nicht dafür zuständig, normative Standards für rationales Denken und Handeln zu entwickeln.

[26] Dies ist auch das Kernargument in Ramseys (1926, S. 160ff) ausführlicher Kritik des Ansatzes von Keynes.

Hypothese mehr oder weniger wahrscheinlich nennen kann. Sieht man ein, daß es sich bei der „epistemischen Relation" zwischen W und H bestenfalls um ein Argument handeln kann, muß die Fragestellung selbst reformuliert werden, etwa so: Welches argumentative Gewicht kann dem durch W vorausgesetzten Wissen für einen Vergleich von zwei oder mehr konkurrierenden Hypothesen gegeben werden? Dann wird jedoch selbst ein metaphorisches Reden von einer „epistemischen Relation" zwischen W und H hinfällig. Man ist vielmehr zu einem komparativen Verständnis epistemischer Wahrscheinlichkeitsaussagen zurückgekehrt.

8. Aber könnte es nicht auch bei einem komparativen Verständnis von Wahrscheinlichkeitsaussagen sinnvoll sein, das zu ihrer Begründung (zur Einschätzung der konkurrierenden Hypothesen) verfügbare Wissen explizit anzugeben? Denn darin bestand ja zunächst der prima facie plausible Grundgedanke von Keynes und vieler anderer Autoren: daß es darum gehen sollte, die Abhängigkeit von Wahrscheinlichkeitsaussagen von einem jeweils verfügbaren Wissen explizit zu machen. Dieser Gedanke kann jedoch auf zwei unterschiedliche Weisen verstanden werden. Ein aus meiner Sicht sinnvolles Verständnis kann etwa so formuliert werden: Bei der Begründung einer epistemischen Wahrscheinlichkeitsaussage sollten die dafür verwendeten Fakten, Annahmen und Überlegungen möglichst explizit gemacht werden. Die Theoriebildungen von Keynes, Jeffreys, Carnap u.a. folgen jedoch einem anderen Verständnis, indem sie von der Annahme ausgehen, daß das zur Begründung einer Wahrscheinlichkeitsaussage verfügbare Wissen vor und unabhängig von einem tatsächlichen Bemühen um eine Begründung fixiert werden kann. Erst diese Annahme erlaubt es, von einem verfügbaren Wissen so zu sprechen, als ob es sich um einen Sachverhalt handelt (der dann durch W symbolisch repräsentiert werden kann). Und erst dadurch kann man schließlich auf den Gedanken kommen, von einer „epistemischen Relation" zwischen W und H zu sprechen.

9. Man kann sich jedoch überlegen, daß diese Vorstellung eines „verfügbaren Wissens" problematisch ist. Bestenfalls kann man vorab – d.h. bevor man sich tatsächlich um die Begründung einer Wahrscheinlichkeitsaussage bemüht – einige der Voraussetzungen fixieren, die für eine Begründung relevant sein könnten. Dementsprechend spricht zum Beispiel Jeffreys davon, daß das jeweils verfügbare Wissen in einer Menge von „Daten" besteht.[27] Daß jedoch eine Charakterisierung des verfügbaren Wissen durch Daten ganz unzureichend ist, kann man sich anhand eines Beispiels verdeutlichen, das Jeffreys (1948, S. 15) selbst angeführt hat:

> "Suppose that I know that Smith is an Englishman, but otherwise know nothing particular about him. He is very likely, on that evidence, to have a blue right eye. But suppose that I am informed that his left eye is brown – the probability is changed completely."

Daraus, daß das linke Auge von Mr. Smith braun ist, folgt natürlich nicht, daß auch sein rechtes Auge braun ist. Das Wissen, das Jeffreys in diesem Beispiel verwendet, besteht auch keineswegs nur in der Information über die Farbe des linken Auges. Diese Information könnte gar nicht genutzt werden, wenn nicht ein weiteres Wissen vorhanden wäre: daß sehr oft (oder nur bei Engländern?) beide Augen die gleiche Farbe haben. Aber bei diesem zusätzlich erforderlichen Wissen handelt es sich nicht um „Daten". Man kann vielleicht sagen, daß in diesem Beispiel die Information über die Farbe des linken Auges von Mr. Smith ein „Datum" ist. Aber die Annahme, daß in vielen Fällen beide Augen die gleiche Farbe haben, ist kein „Datum", sondern eine „Regel", die nicht als etwas „Gegebenes" zur Verfügung steht.[28]

10. Das Beispiel sollte zeigen, daß es im allgemeinen nicht möglich ist, das zur Begründung von Wahrscheinlichkeitsaussagen *verwendete* Wissen vollständig vorab zu fixieren. Man kann es auch anders formulieren: Wissen läßt sich nicht auf Information (Daten) reduzieren; wir müssen auch wissen, wie mit Daten bei Überlegungen, die sich auf Hypothesen beziehen, sinnvoll umgegangen werden kann. So wird auch noch einmal die Auffassung einsichtig, die bereits im vorangegangenen Abschnitt dargelegt wurde: daß sich Begründungen epistemischer Wahrscheinlichkeitsaussagen nicht durch einen Kalkül oder eine „Methode der Inferenz" ersetzen lassen. — Dennoch ist es in vielen Fällen sicherlich sinnvoll, die Daten, auf die man sich bei der Begründung einer Wahrscheinlichkeitsaussage berufen möchte, explizit zu machen.

[27] Vgl. das oben angeführte Zitat. Auch andere Autoren sprechen in diesem Zusammenhang oft von „Daten". Zum Beispiel heißt es bei Carnap und Stegmüller (1959, S. 6): „Eine Aussage über *induktive Wahrscheinlichkeit* hat eine Relation zwischen einer Hypothese und einer Gesamtheit von Erfahrungsdaten (z.B. bestimmten Beobachtungsergebnissen) zum Inhalt."

[28] Übrigens eine Regel, von der man wohl nicht sagen kann, daß es sich um ein „Naturgesetz" handelt. Jedenfalls kann man sich vorstellen, daß Menschen die Fähigkeit haben oder erlangen könnten, sich unterschiedliche Augenfarben zu machen, auch wenn sie „von Natur aus" gleich sein sollten.

3.8.3 Logisch vs. empirisch?

1. Hier soll noch auf eine weitere fragwürdige Implikation hingewiesen werden. Wenn epistemische Wahrscheinlichkeitsaussagen als Aussagen über eine „epistemische Relation“ zwischen einer Hypothese und „Daten“ konzipiert werden, wird unklar, in welcher Weise sie überhaupt noch einen empirischen Sinn haben können. Einige Autoren, die diese „epistemische Relation“ als eine „logische Relation“ zwischen Aussagen auffassen, sind infolgedessen auch zu der Auffassung gelangt, daß epistemische Wahrscheinlichkeitsaussagen – vergleichbar mit logischen Aussagen – überhaupt keinen empirischen Gehalt haben. Diese Auffassung haben zum Beispiel Carnap und Stegmüller (1959, S. 7) vertreten:

> „Für diesen Begriff der induktiven Wahrscheinlichkeit oder des Bestätigungsgrades ist der relationale Charakter wesentlich: Eine Aussage, die vom Bestätigungsgrad einer Hypothese handelt, muß stets auch auf ein Erfahrungsdatum Bezug nehmen. Dies ist nicht mit dem für alle naturwissenschaftlichen Aussagen gültigen Prinzip zu verwechseln, daß die Aussage auf Beobachtungen beruhen muß. Die Aussage ‘die Wahrscheinlichkeit, daß es morgen regnen wird, ist 1/6’ ist ein unvollständiger Satz, solange nicht der Zusatz hinzugefügt wird ‘in bezug auf diese und diese Erfahrungsdaten’. Die Erfahrungsdaten werden z.B. im vorliegenden Falle gewisse meteorologische Beobachtungen sein. Ist der Satz durch diese Einbeziehung einmal vollständig geworden, dann spielt es für seine Gültigkeit weder eine Rolle, ob die erwähnten Erfahrungsdaten wahr sind oder nicht, noch ob sie dem Sprecher bekannt sind. [...] Eine induktive Wahrscheinlichkeitsaussage hat also die Gestalt ‘die Wahrscheinlichkeit (der Bestätigungsgrad) der Hypothese h auf Grund der Erfahrungsdaten e ist so und so groß’. Alles für die Wahrscheinlichkeitsbeurteilung relevante Wissen ist bereits innerhalb dieses Satzes selbst ausgedrückt, nämlich durch die Einbeziehung von e. Daraus ergibt sich, daß der Grund für die Gültigkeit einer solchen Wahrscheinlichkeitsaussage nicht selbst wiederum empirischer Natur sein kann.“

Offenbar beruht das Argument auf der Annahme, daß es möglich sei, das gesamte für die Beurteilung einer Hypothese relevante empirische Wissen im Konditionalteil einer Wahrscheinlichkeitsaussage explizit zu machen. Wenn dies jedoch nicht richtig ist, wie das zuvor besprochene Beispiel zeigen sollte, entfällt nicht nur die Möglichkeit, den Zusammenhang zwischen Hypothese und Daten als eine „logische Relation“ zu deuten, sondern die Wahrscheinlichkeitsaussagen selbst können dann nicht als „logische“, im Sinne von nicht-empirischen Aussagen verstanden werden.

2. Aber handelt es sich infolgedessen um empirische Aussagen? In der empiristischen Wissenschaftstheorie, wie sie von Carnap und Stegmüller vertreten wird, gäbe es dann nur diese Alternative. Aber auch in dieser Frage zeigt sich, wie ich glaube, unsere Konzeption epistemischer Wahrscheinlichkeitsaussagen als sinnvoller. Wie in Abschnitt 3.3 betont worden ist, sind epistemische Wahrscheinlichkeitsaussagen in unserem Verständnis Aussagen über Begründungen von Hypothesen. Sie beziehen sich also nicht unmittelbar auf unsere Erfahrungswelt; der unmittelbare Bezug auf die Erfahrungswelt wird vielmehr durch die Hypothesen hergestellt. Aber indem sich Wahrscheinlichkeitsaussagen auf *Begründungen* von Hypothesen beziehen, betreffen sie gleichwohl unsere Wissensbildung über die Erfahrungswelt. Es soll ja durch eine Einschätzung der Gründe für und gegen konkurrierende Hypothesen Wissen gebildet werden, wenn auch nur in der Form eines mehr oder weniger gut begründeten hypothetischen Wissens. Aber zur Begründung dieses Wissens spielen natürlich empirische Argumente eine entscheidende Rolle. Es geht nicht um eine „logische Beziehung“ zwischen einer Hypothese H und einer Menge von Daten W, sondern W soll *als ein Argument* dienen, um einschätzbar zu machen, wie gut H (im Vergleich zu einer anderen Hypothese H') begründet ist. Infolgedessen ist es – ganz im Gegensatz zu der oben zitierten Auffassung von Carnap und Stegmüller – gerade wichtig, „ob die erwähnten Erfahrungsdaten wahr sind oder nicht“ und „ob sie dem Sprecher bekannt sind“.

Kapitel 4

Aleatorische Wahrscheinlichkeit

Im vorangegangenen Kapitel wurden epistemische Wahrscheinlichkeitsaussagen besprochen, deren Sinn darin liegt, zum Ausdruck zu bringen, wie eine Menge alternativer Hypothesen auf der Grundlage eines gegebenen Wissens zu beurteilen ist. Dort wurde auch festgestellt, daß für derartige Wahrscheinlichkeitsaussagen im allgemeinen kein quantitatives Wahrscheinlichkeitsmaß zur Verfügung steht. Die Entwicklung der Wahrscheinlichkeitsrechnung beruht demgegenüber darauf, daß es unter spezifischen Voraussetzungen möglich erscheint, auch zu quantitativen Wahrscheinlichkeitsaussagen zu gelangen. Welche Voraussetzungen dafür erforderlich sind, ist umstritten. Wir folgen in diesem Kapitel der historischen Entwicklung der Wahrscheinlichkeitsrechnung, die von einer Beschäftigung mit Glücksspielen ausgegangen ist. Um daran zu erinnern, sprechen wir von einem *aleatorischen* Wahrscheinlichkeitsbegriff.

4.1 Einleitende Bemerkungen

1. Beginnen wir mit einem kurzen Rückblick. In Kapitel 2 ging es darum, wie man über Möglichkeiten sprechen und Vermutungen über Möglichkeiten in der Form von Hypothesen sprachlich fixieren kann. Dabei wurde davon ausgegangen, daß sich Hypothesen stets auf bestimmte Situationen oder Mengen von Situationen beziehen. Solche Hypothesen wurden „definit" genannt. In Kapitel 3 wurde die Diskussion definiter Hypothesen fortgesetzt und überlegt, wie Wahrscheinlichkeitsaussagen für Hypothesen (die eine Algebra bilden) formuliert werden können. Wir haben in diesem Zusammenhang von epistemischen Wahrscheinlichkeitsaussagen gesprochen, da ihr Sinn darin liegt, zum Ausdruck zu bringen, wie eine Menge alternativer Hypothesen auf der Grundlage eines gegebenen Wissens zu beurteilen ist.

2. Weiterhin wurde besprochen, daß eine Quantifizierung epistemischer Wahrscheinlichkeiten in den meisten Anwendungsfällen aussichtslos erscheint. Epistemische Wahrscheinlichkeiten wurden deshalb als ein komparatives Konzept eingeführt, mit dem komparative Aussagen über zwei oder mehr Hypothesen formuliert werden können. Dagegen beginnt die Geschichte der modernen Wahrscheinlichkeitstheorie mit dem Grundgedanken, daß es möglich sei, auch einen quantitativen Wahrscheinlichkeitsbegriff zu begründen. In diesem Kapitel soll es darum gehen, ein Verständnis für diesen Grundgedanken zu finden.

3. Zum Verständnis ist zunächst wichtig, daß die moderne Wahrscheinlichkeitstheorie aus einer Beschäftigung mit Glücksspielen entstanden ist.[1] Als ihr historischer Ausgangspunkt wird oft auf einen Briefwechsel aus dem Jahr 1664 hingewiesen, in dem sich die Mathematiker Blaise Pascal und Pierre de Fermat mit einer Berechnung von Gewinnchancen in Würfelspielen beschäftigt haben.[2] Das erste größere Werk zur Wahrscheinlichkeitsrechnung stammt von dem Mathematiker Jakob Bernoulli. Es trägt den Titel *Ars conjectandi* („Kunst der Vermutungen“) und wurde zuerst posthum im Jahr 1713 veröffentlicht.[3] Das Buch besteht aus vier Teilen.[4] Die ersten drei Teile beschäftigen sich in erster Linie mit Glücksspielen (Münzwurf, Würfel- und Kartenspiele);[5] dies waren die in der frühen Literatur bevorzugt behandelten Beispiele zur Erläuterung wahrscheinlichkeitstheoretischer Überlegungen.[6] Dennoch ging es Bernoulli wie auch den meisten anderen Autoren, die an der Entwicklung der Wahrscheinlichkeitstheorie beteiligt waren, nicht in erster Linie um Glücksspiele. Vielmehr ging es ihnen darum, eine *allgemeine* Theorie vernünftiger Urteilsbildungen für Situationen der Unsicherheit zu entwickeln;[7] und dies ist auch das Thema des vierten Teils der *Ars conjectandi*, der den Titel trägt: *Anwendung der vorhergehenden Lehre auf bürgerliche, sittliche und wirtschaftliche Verhältnisse*.

[1]Eine Einführung in den historischen Kontext findet sich bei David, 1962.

[2]In einer berühmten Geschichte der Wahrscheinlichkeitsrechnung von Todhunter (1865, S. 20) heißt es: „Neglecting the trifling hints which may be found in preceding writers we may say that the Theory of Probability really commenced with Pascal and Fermat.“ Inzwischen sind jedoch eine Reihe von Arbeiten erschienen, die zeigen, daß es eine viel weiter zurückreichende Geschichte der Wahrscheinlichkeitsrechnung gibt; u.a. Ineichen 1996, Rabinovitch 1973. Dt. Übersetzungen von Auszügen aus dem Briefwechsel zwischen Pascal und Fermat finden sich bei Schneider 1988, S. 25ff. Weitere Hinweise zum historischen Kontext dieses Briefwechsels finden sich auch bei Ore, 1960.

[3]Bernoulli lebte 1654–1705; eine gute Darstellung des historischen Kontextes seiner Arbeiten zur Wahrscheinlichkeitstheorie findet sich bei van der Wærden, 1975.

[4]Ich beziehe mich auf die deutsche Übersetzung von R. Haussner, vgl Bernoulli 1899a, b.

[5]Der erste Teil des Buches besteht im wesentlichen aus einem früheren Text von Christian Huygens, *De ratiociniis in ludo aleae*, der zuerst 1657 veröffentlicht wurde. Bernoulli reproduzierte diesen Text und versah ihn mit Anmerkungen und Lösungen der von Huygens gestellten Aufgaben.

[6]Ein anderes berühmtes Buch, das sich fast ausschließlich mit Glücksspielen beschäftigt, ist Abraham de Moivre's *The Doctrine of Chances; or A Method of Calculating the Probabilities of Events in Play*. Dies Buch erschien in drei Ausgaben: 1718, 1738 und 1756.

4. Natürlich stellt sich die Frage, ob eine allgemeine „Theorie rationaler Vermutungen“ sinnvoll formuliert werden kann und ob dafür eine Theorie quantitativer Wahrscheinlichkeiten hilfreich sein kann. Daß hier Zweifel angebracht sind, folgt schon daraus, daß es im allgemeinen keine sinnvolle Methode gibt, epistemische Wahrscheinlichkeiten zu quantifizieren. Aber auch wenn man an dieser Stelle den Allgemeinheitsanspruch aufgibt, bleibt die Frage, ob überhaupt eine Theoriebildung, die ihre leitenden Begriffe aus einer Beschäftigung mit Glücksspielen oder, etwas allgemeiner, mit Zufallsgeneratoren gewinnt, auch zur Reflexion von Fragestellungen genutzt werden kann, die sich auf mögliche Sachverhalte in unserer Erfahrungswelt beziehen.

5. Diese Frage wird uns in späteren Kapiteln noch mehrfach beschäftigen. Bereits an dieser Stelle kann jedoch darauf hingewiesen werden, daß sich auch Bernoulli des Problems durchaus bewußt war, wie seine folgenden Ausführungen zeigen:

„Wir sind also dahin gelangt, dass zur richtigen Bildung von Vermuthungen über irgend eine Sache nichts anderes zu thun 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. 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 Gücksspiele wurden aber von den ursprünglichen Erfindern, damit die Spieltheilnehmer 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 andern 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 giebt für jeden einzelnen Würfel ebensoviele Fälle als er Flächen hat; alle diese Fälle sind auch gleich leicht möglich, da wegen der gleichen Gestalt aller Flächen und

[7]Eine ausführliche Diskussion dieses Erkenntnisinteresses findet sich bei Daston, 1988. Hier mag als Hinweis folgende Bemerkung von Bernoulli (1899b, S. 75) genügen: „Irgend ein Ding *vermuten* heisst soviel als seine Wahrscheinlichkeit messen. Deshalb bezeichnen wir als *Vermuthungs-* oder *Muthmaassungskunst* (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 Urtheilen und Handlungen stets das auswählen und befolgen können, was uns besser, trefflicher, sicherer oder rathsamer erscheint. Darin allein beruht die ganze Weisheit des Philosophen und die ganze Klugheit des Staatsmannes.“

wegen des gleichmässig vertheilten Gewichtes 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 Theil des Würfels aus schwererem Materiale angefertigt wäre als der andere Theil. So sind auch die Zahlen der Fälle für das Ziehen eines weissen oder eines schwarzen Steinchens aus einer Urne bekannt und können alle Steinchen auch gleich leicht gezogen werden, weil bekannt ist, wieviele Steinchen von jeder Art in der Urne vorhanden sind, und weil sich kein Grund angeben lässt, warum dieses oder jenes Steinchen leichter als irgend ein anderes gezogen werden sollte. Welcher Sterbliche könnte aber je die Anzahl der Krankheiten (d.i. ebensovieler Fälle), welche den menschlichen Körper an allen seinen Theilen 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 Vermuthung über das Verhältniss von Leben und Sterben künftiger Geschlechter abzuleiten? Oder wer könnte die unzähligen Fälle von Veränderungen aufzählen, welchen die Luft täglich unterworfen ist, um daraus schon heute vermuthen zu wollen, welche Beschaffenheit sie nach einem Monate oder gar nach einem Jahre hat? [...] Da diese und ähnliche Dinge von ganz verborgenen Ursachen abhängen, welche überdies noch durch die unendliche Mannigfaltigkeit ihres Zusammenwirkens unsere Erkenntnis beständig täuschen, so würde es völlig sinnlos sein, auf diese Weise etwas erforschen zu wollen. 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.“ (Bernoulli 1899b, S. 88f.)

6. Die Unterscheidung ist deutlich. Wenn man sich mit Glücksspielen oder, allgemeiner, mit Zufallsgeneratoren beschäftigt, gewinnt man Kenntnisse über Wahrscheinlichkeiten aus Einsichten in ihre Konstruktion. Es handelt sich um ein Wissen *a priori* in dem Sinne, daß man es aus einer Untersuchung von Konstruktionsverfahren gewinnen kann und nicht darauf angewiesen ist, es aus Beobachtungen realisierter Spiele zu bilden.[8] Wenn es sich jedoch andererseits um Vorkommnisse handelt, „welche von dem Walten der Natur oder von der Willkür der Menschen abhängen“, kann unser Wissen nur aus einer Beobachtung der Vorkommnisse selbst und – wenn möglich – der Prozesse, durch die sie entstehen, gewonnen werden. Zur Begründung von Erwartungen, die sich auf *mögliche* Vorkommnisse beziehen, kann man sich dann bestenfalls auf in der Vergangenheit gemachte Beobachtungen berufen, und ggf. auf eine Kenntnis von Regeln für das Verhalten von Dingen und Akteuren. Und somit stellt sich die Frage, ob dafür eine Wahrscheinlichkeitstheorie nützlich sein kann, die ihre grundlegenden Begriffe und Vorstellungen aus einer Bezugnahme auf Zufallsgeneratoren gewinnt; insbesondere dann, wenn es sich um Erwartungen für mögliche Vorkommnisse handelt, die nicht durch Zufallsgeneratoren, sondern durch Interaktionsprozesse von Akteuren hervorgebracht werden.

[8] Man sollte darauf achten, daß mit der Charakterisierung ‘a priori’ hier nicht gemeint ist, daß es sich um ein Wissen „vor aller Erfahrung“ handelt; vielmehr soll der Ausdruck in diesem Kontext darauf hinweisen, daß man dieses Wissen aus Einsichten in die Beschaffenheit bzw. Konstruktion eines Verfahrens gewinnt.

4.2 Zufallsgeneratoren

Als prototypische Beispiele zur Motivation von Begriffsbildungen der Wahrscheinlichkeitsrechnung dienen – auch heute noch – Glücksspiele. In diesem Abschnitt orientieren wir uns an solchen Beispielen, um daraus die allgemeine Konzeption eines Zufallsgenerators zu gewinnen.[9]

4.2.1 Verfahren zur Ereigniserzeugung

1. Denken wir an ein Würfelspiel. Man nimmt einen Würfel und kann, indem man ihn wirft, ein Ereignis erzeugen. Das kann man beliebig oft machen. Also kann man das Würfelspiel als ein Verfahren zur Erzeugung von Ereignissen auffassen. Weiterhin kann man sich überlegen, welche Ereignisse möglich sind. Entsprechend den Regeln des Würfelspiels kann beim Wurf eines Würfels genau eines von sechs möglichen Ereignissen auftreten, unterschieden durch die jeweils erzielte Augenzahl. Hier muß man allerdings etwas aufpassen. Einerseits kann man mit einem Würfel beliebig viele Ereignisse erzeugen, andererseits gibt es nur sechs mögliche Ereignisse? Offenbar benötigt man eine sprachliche Unterscheidung. Wir orientieren uns an dem in Abschnitt 2.4 diskutierten Sprachgebrauch und verwenden das Wort ‘Ereignis’ für jeweils einmalige Vorkommnisse im historischen Ablauf des Geschehens. Dementsprechend sagen wir, daß man mit einem Würfel beliebig viele Ereignisse erzeugen kann. Andererseits trifft jedoch für jedes Ereignis genau eine von sechs möglichen Kennzeichnungen zu, nämlich eine Kennzeichnung durch die jeweils erzielte Augenzahl.

[9] Die Darstellung orientiert sich an Überlegungen von P. Lorenzen 1974, S. 209–218, und 1978. Dies betrifft insbesondere den in Abschnitt 4.5 weiter verfolgten Gedanken, elementare Zufallsgeneratoren zur Konstruktion allgemeiner Zufallsgeneratoren zu verwenden.

2. Jetzt kann man sich überlegen, wie Hypothesen gebildet werden können. Dabei ist eine Unterscheidung wichtig, die bereits in Abschnitt 2.8 besprochen worden ist. Wir können uns einerseits auf Situationen beziehen, in denen Würfelspiele durchgeführt werden. In solchen Situationen können offenbar sehr viele unterschiedliche Ereignisse auftreten. Zum Beispiel kann eine der anwesenden Personen einen Würfel nehmen und werfen, eine andere Person kann über den Ausgang des Wurfes eine Vermutung anstellen, eine dritte Person kann vorschlagen, das Spiel abzubrechen, usw. All diese in der Realität möglichen Ereignisse bleiben jedoch außer Betracht, wenn Würfelspiele zur Motivation von Begriffsbildungen der Wahrscheinlichkeitsrechnung verwendet werden. Die Hypothesen beziehen sich insbesondere nicht darauf, *ob* mit einem Würfel ein Ereignis erzeugt wird. Sie beziehen sich überhaupt nicht unmittelbar auf Situationen, in denen – zum Beispiel – Würfelspiele durchgeführt werden, sondern – in diesem Fall – auf Situationen, die durch Werfen eines Würfels erzeugt werden können. Um das sprachlich zu fassen, hilft die in Kapitel 2 eingeführte Notation für Hypothesen. Es wird unterstellt, daß durch das Werfen eines Würfels eine Situation s erzeugt werden kann, die sich durch eine von sechs Möglichkeiten kennzeichnen läßt. Es gibt also eine Kennzeichnungsmenge, die in diesem Beispiel durch

$$\tilde{K} := \{1,2,3,4,5,6\}$$

vergegenwärtigt werden kann; und es gibt dementsprechend sechs Hypothesen der Form $\langle s, \tilde{k} \rangle$, wobei $\tilde{k} \in \tilde{K}$ eine der möglichen Kennzeichnungen ist. Hypothesen beziehen sich also auf Aspekte von Situationen, die mit einem Würfel erzeugt werden können, nicht darauf, ob solche Situationen eintreten.

3. Bei dieser Betrachtungsweise ist s eine jeweils bestimmte Situation im historischen Ablauf des Geschehens. Für Würfelspiele ist jedoch charakteristisch, daß beliebig viele Situationen der gleichen Art erzeugt werden können. Damit ist zweierlei gemeint. Erstens, daß die Situationen – wenn sie erzeugt werden – stets auf die gleiche Art erzeugt werden, wobei die Art der Erzeugung durch ein Verfahren (eine Menge von Regeln) festgelegt wird. In unserem Beispiel bezieht sich das Verfahren darauf, wie der Würfel beschaffen sein soll und wie mit ihm geworfen werden soll. Zweitens ist gemeint, daß die Situationen ausschließlich unter dem Gesichtspunkt betrachtet werden sollen, daß sie sich durch jeweils ein Element der Kennzeichnungsmenge $\tilde{K}$ charakterisieren lassen. Diese Bedingungen legen nicht nur fest, was unter „Situationen der gleichen Art“ verstanden werden soll; sie legen auch fest, wann zwei Situationen als gleich gelten sollen: nämlich genau dann, wenn es sich um Situationen der gleichen Art handelt und wenn sie durch das gleiche Element der Kennzeichnungsmenge $\tilde{K}$ charakterisiert werden können.

4. Schließlich ist eine weitere Bedingung wesentlich: das Verfahren, mit dem Situationen der gleichen Art erzeugt werden können, soll so beschaffen sein, daß man vor einer Anwendung des Verfahrens nicht wissen kann, welche Situation resultieren wird. Diese Bedingung motiviert, daß man von „Zufallsgeneratoren“ spricht. Die Resultate sollen „zufällig“ entstehen, womit indessen hier nur gemeint ist, daß sich vor einer Anwendung des Verfahrens mit dem gegenwärtig verfügbaren Wissen keine sicheren Voraussagen darüber machen lassen, welches Element der vorab festgelegten Kennzeichnungsmenge $\tilde{K}$ in der resultierenden Situation realisiert wird.[10]

5. Jetzt kann der allgemeine Begriff eines Zufallsgenerators definiert werden. Wir verwenden dafür die Notation

$$\mathcal{G} := (\mathcal{V}, (\mathcal{S}, \tilde{K}))$$

Der Zufallsgenerator $\mathcal{G}$ besteht aus einem Verfahren $\mathcal{V}$, mit dem beliebig viele Situationen des Typs $(\mathcal{S}, \tilde{K})$ erzeugt werden können, die ihrerseits durch jeweils genau ein Element einer Kennzeichnungsmenge $\tilde{K}$ charakterisiert werden können. Außerdem soll gelten, daß vor einer Anwendung des Verfahrens mit dem gegenwärtig verfügbaren Wissen keine sicheren Voraussagen darüber gemacht werden können, welches Element der Kennzeichnungsmenge $\tilde{K}$ realisiert wird.

6. Die Schreibweise $(\mathcal{S}, \tilde{K})$ wird verwendet, um auf den Typ der Situationen zu verweisen, die mit dem Zufallsgenerator erzeugt werden können. Die Schreibweise soll es erlauben, daß man sich beim Reden über einen Zufallsgenerator auf eine oder mehrere Situationen dieses Typs beziehen kann. Es sollte jedoch beachtet werden, daß $(\mathcal{S}, \tilde{K})$ selbst weder eine Situation ist noch auf eine Situation verweist. Situationen, für die wir die Symbole $s, s_1, s_2, \ldots$ verwenden, sind stets bestimmte (reale oder vorgestellte) Situationen. Der Ausdruck $(\mathcal{S}, \tilde{K})$ soll vielmehr dafür dienen, um über solche jeweils bestimmten Situationen sprechen zu können.

[10] Es sei darauf hingewiesen, daß diese Bedingung bei den in der statistischen Praxis verwendeten Zufallsgeneratoren meistens nicht erfüllt ist. Sie ist insbesondere dann nicht erfüllt, wenn mit spezifischen Computerprogrammen sogenannte Zufallszahlen erzeugt werden. Man spricht dann gelegentlich von Pseudo-Zufallsgeneratoren. Wie von „zufällig“ gesprochen werden kann, wird etwas genauer in Abschnitt 6.3.1 genauer besprochen.

Wir verwenden die Schreibweise

$$s \models (\mathcal{S}, \tilde{K})$$

um damit zum Ausdruck zu bringen, daß s eine bestimmte (reale oder vorgestellte) Situation des Typs $(\mathcal{S}, \tilde{K})$ ist.[11]

7. Hat man verstanden, wie $(\mathcal{S}, \tilde{K})$ als ein Situationstyp zur Charakterisierung jeweils bestimmter Situationen verwendet werden kann, kann man in analoger Weise auch von Situationen des Typs $(\mathcal{S}, \tilde{k})$ sprechen, wobei $\tilde{k}$ ein Element aus der Kennzeichnungsmenge $\tilde{K}$ ist. Zu sagen, daß s eine Situation vom Typ $(\mathcal{S}, \tilde{k})$ ist, soll bedeuten, daß s eine Situation des Typs $(\mathcal{S}, \tilde{K})$ ist, die durch $\tilde{k} \in \tilde{K}$ charakterisiert werden kann. Es sei darauf hingewiesen, daß es sich bei $\tilde{k}$ auch um eine beliebige Teilmenge von $\tilde{K}$ handeln kann. Eine Situation des Typs $(\mathcal{S}, \tilde{k})$ ist dann eine Situation des Typs $(\mathcal{S}, \tilde{K})$, die durch irgendein Element aus $\tilde{k}$ charakterisiert werden kann. Somit kann man auch von einer Algebra für Situationstypen sprechen.

8. Man sollte sich schließlich noch einmal klarmachen, daß Zufallsgeneratoren Artefakte sind, von Menschen ausgedachte Verfahren, um auf wiederholbare Weise Situationen zu erzeugen, deren mögliche Eigenschaften – soweit sie durch eine Kennzeichnungsmenge vorab fixiert worden sind – nicht sicher vorausgesagt werden können. Die Aussage, daß bei einem Zufallsgenerator nur Situationen eines vorab fixierten Typs $(\mathcal{S}, \tilde{K})$ auftreten können, ist keine empirische Aussage. Die Aussage bezieht sich vielmehr auf den Sinn des Verfahrens, durch das ein Zufallsgenerator definiert wird. Zum Beispiel können bei einem Würfelspiel zahlreiche Ereignisse auftreten, die durch die Spielregeln nicht vorgesehen sind. Niemand kann garantieren, daß sich Menschen, wenn sie einen Zufallsgenerator zur Erzeugung von Ereignissen verwenden, an die Regeln halten. Und auch wenn ein Zufallsgenerator durch eine Maschine realisiert wird, zum Beispiel durch einen Computer (etwa in der Form eines Glücksspielautomaten), können Ereignisse eintreten, die dem vorab fixierten Typ $(\mathcal{S}, \tilde{K})$ nicht entsprechen; zum Beispiel kann ein Computer defekt werden oder es kann ein Stromausfall eintreten.

[11] Wie schon in Abschnitt 2.8 besprochen worden ist, wird in diesem Text das Symbol '$\models$' anstelle der mengentheoretischen Notation '$\in$' verwendet, um zu betonen, daß ein Situationstyp nicht sinnvoll als eine Menge möglicher Situationen konzipiert werden kann. Mit einem Zufallsgenerator $(\mathcal{V}, (\mathcal{S}, \tilde{K}))$ können zwar beliebig viele Situationen des Typs $(\mathcal{S}, \tilde{K})$ hergestellt werden; aber es gibt kein Verfahren, mit dem sinnvoll eine „Menge aller möglichen Situationen dieses Typs“ definiert werden könnte.

4.2.2 Elementare Zufallsgeneratoren

1. Eine wichtige Frage ist noch nicht angesprochen worden: *wie* bei einem Zufallsgenerator der Prozeß beschaffen ist, durch den Situationen erzeugt werden können. Zur Diskussion dieser Frage beziehen wir uns auf einen Zufallsgenerator $\mathcal{G} := (\mathcal{V}, (\mathcal{S}, \tilde{K}))$ mit der Kennzeichnungsmenge

$$\tilde{K} := \{\tilde{k}_1, \ldots, \tilde{k}_m\}$$

Man kann dann sagen, daß der Zufallsgenerator ein Verfahren ist, um Situationen zu erzeugen, die sich durch jeweils genau ein Element $\tilde{k} \in \tilde{K}$ charakterisieren lassen. Es kommt also gar nicht auf weitere Aspekte solcher Situationen an, sondern nur darauf, durch welches Element aus der Kennzeichnungsmenge $\tilde{K}$ sie charakterisiert werden können. Insofern erscheint es möglich, auch bei der Charakterisierung des Verfahrens $\mathcal{V}$, das durch den Zufallsgenerator $\mathcal{G}$ definiert wird, von der Vorstellung auszugehen, daß es sich um einen *Prozeß der Auswahl* eines Elementes aus der Kennzeichnungsmenge $\tilde{K}$ handelt.

2. Diese Betrachtungsweise erlaubt es, die Frage nach der Funktionsweise von Zufallsgeneratoren so zu formulieren: Wie ist das Verfahren beschaffen, mit dem bei der Betätigung eines Zufallsgenerators jeweils ein Element aus $\tilde{K}$ ausgewählt wird? Bei *elementaren* Zufallsgeneratoren besteht die leitende Idee darin, das Verfahren so zu gestalten, daß es für jedes Element von $\tilde{K}$ die „gleiche Chance“ gibt, ausgewählt zu werden. Diese Idee wird zum Beispiel verfolgt, wenn man das Werfen von Münzen oder Würfeln als Zufallsgeneratoren verwendet. Für analytische Zwecke am besten geeignet sind Urnen, die mit vollständig symmetrischen Objekten (meistens Kugeln) gefüllt sind, die sich nur durch einen Namen (z.B. ihre Farbe oder eine Nummer) unterscheiden. Soll die Kennzeichnungsmenge eines Zufallsgenerators m verschiedene Elemente enthalten, wird die Urne mit m unterscheidbaren Objekten gefüllt. Das praktische Verfahren besteht dann darin, die Objekte in der Urne zunächst zu mischen und dann ein Objekt aus der Urne blind herauszuziehen; der Name des herausgezogenen Objekts liefert die Kennzeichnung der erzeugten Situation. Schließlich wird das herausgezogene Objekt wieder in die Urne zurückgelegt, um gleiche Ausgangsbedingungen für eine Wiederholung herzustellen.

3. Hat bei diesem Verfahren jedes Objekt in der Urne die gleiche Chance, gezogen zu werden? Die Frage kann in dieser Form nicht sinnvoll beantwortet werden, denn es steht keine unabhängige Definition für „gleiche Chancen“ zur Verfügung. Die Frage ist vielmehr, wie eine solche

Definition gefunden werden kann. Die grundlegende Idee besteht darin, daß man elementare Zufallsgeneratoren verwenden kann, *um zu explizieren*, was mit „gleichen Chancen" gemeint sein soll. Bei dem Verfahren, das eine Urne verwendet, kann man so argumentieren: keine der Kugeln wird durch das Auswahlverfahren bevorzugt; und *das* soll damit gemeint sein, daß jede Kugel die gleiche Chance hat, gezogen zu werden.

4. Man sollte sich klarmachen, daß es hier nicht um eine empirische Frage geht, die durch Beobachtungen der mit einem Zufallsgenerator erzeugten Situationen entschieden werden könnte. Sicherlich kann man solche Beobachtungen machen. Man kann einen Zufallsgenerator n-mal betätigen und dann für jedes $\tilde{k} \in \tilde{K}$ die relative Häufigkeit feststellen, mit der Situationen des Typs $(\mathcal{S}, \tilde{k})$ eingetreten sind. So erhält man die relativen Häufigkeiten

$$h_n(\tilde{k}_j) \qquad (\text{für } j = 1, \ldots, m)$$

Aber wollte man versuchen, die Bedeutung von „gleiche Chancen" durch die Forderung

$$h_n(\tilde{k}_1) = \ldots = h_n(\tilde{k}_m)$$

zu definieren, würde dies sogleich in eine Sackgasse führen; denn diese Forderung kann gar nicht für beliebige n erfüllt werden.

5. Es sei indessen betont: das Argument zielt darauf, daß Beobachtungen der mit einem Zufallsgenerator erzeugten Situationen keine Grundlage liefern, um den Sinn von „gleiche Chancen" zu definieren. Empirische Argumente spielen gleichwohl eine wichtige Rolle. Aber sie müssen sich auf die Konstruktion des Zufallsgenerators bzw. auf die zu seiner Verwendung festgelegten Regeln beziehen. Wird zum Beispiel eine Urne verwendet, ist es wichtig, daß die Kugeln vorher *gemischt* werden und daß dann eine der Kugeln *blind* herausgezogen wird. Es sind solche elementaren Symmetrieforderungen, die unser Verständnis von „gleiche Chancen" begründen; und es ist eine empirische Frage, ob ihnen bei der Konstruktion bzw. Verwendung eines Zufallsgenerators entsprochen worden ist.

6. Man kann es auch so sagen: ein elementarer Zufallsgenerator sollte so konstruiert werden, daß sich aus Hinweisen auf seine Konstruktion (einschließlich der Regeln für seinen Gebrauch) keine Argumente gewinnen lassen, mit denen gezeigt werden könnte, daß der Prozeß, durch den Situationen des Typs $(\mathcal{S}, \tilde{K})$ erzeugt werden, von Eigenschaften abhängt, die nicht allen Elementen $\tilde{k} \in \tilde{K}$ gleichermaßen zukommen.

7. Die Idee, daß bei einem elementaren Zufallsgenerator Situationstypen $(\mathcal{S}, \tilde{k})$ gleiche Realisierungschancen haben, hat eine wichtige Implikation: der Zufallsgenerator muß so konstruiert werden, daß das Verfahren $\mathcal{V}$ unabhängig davon funktioniert, welche Situationen bei früheren Anwendungen des Verfahrens realisiert worden sind. Bei der Verwendung von Urnen wird dies zum Beispiel dadurch erreicht, daß die Kugeln gemischt werden, bevor eine von ihnen herausgezogen wird. Man kann auch sagen, daß das Verfahren keine Erinnerung an frühere Anwendungen haben darf. In dieser Betrachtungsweise sind elementare Zufallsgeneratoren zugleich Verfahren, um auf möglichst perfekte Weise die Idee eines gedächtnislosen Prozesses zu realisieren.

4.3 Aleatorische Wahrscheinlichkeitsmaße

In diesem Abschnitt wird besprochen, wie Zufallsgeneratoren verwendet werden können, um Wahrscheinlichkeitsmaße zu definieren.

4.3.1 Motivation der Definition

1. Für die Überlegungen dieses Abschnitts wird ein elementarer Zufallsgenerator $\mathcal{G} := (\mathcal{V}, (\mathcal{S}, \tilde{K}))$ mit der Kennzeichnungsmenge

$$\tilde{K} := \{\tilde{k}_1, \ldots, \tilde{k}_m\}$$

vorausgesetzt. Der Zufallsgenerator hat also die Eigenschaft, daß bei den Situationen, die mit ihm erzeugt werden können, alle Werte aus $\tilde{K}$ gleich möglich sind.

2. Es soll jetzt ein Wahrscheinlichkeitsmaß

$$\mathrm{P_a} : \mathcal{P}(\tilde{K}) \longrightarrow \mathbf{R}$$

konstruiert werden, das jeder Menge $\tilde{k} \subseteq \tilde{K}$ eine Zahl $\mathrm{P_a}(\tilde{k})$ zuordnet, die als „Grad der Möglichkeit" für den Situationstyp $(\mathcal{S}, \tilde{k})$ interpretiert werden kann. Dafür steht zunächst der Gedanke zur Verfügung, daß alle m Situationstypen $(\mathcal{S}, \tilde{k}_j)$ gleich möglich sind. Also muß das Wahrscheinlichkeitsmaß folgende Bedingung erfüllen:

$$\mathrm{P_a}(\{\tilde{k}_1\}) = \ldots = \mathrm{P_a}(\{\tilde{k}_m\}) \tag{4.3.1}$$

Qua Konstruktion des Zufallsgenerators steht auch noch eine weitere Überlegung zur Verfügung: daß der Grad der Möglichkeit eines Situationstyps $(\mathcal{S}, \tilde{k})$ nur von der Anzahl der Merkmalswerte in $\tilde{k}$ abhängt. Also muß das Wahrscheinlichkeitsmaß auch folgende Bedingung erfüllen:

$$\mathrm{P_a}(\tilde{k}) = \mathrm{P_a}(\tilde{k}') \iff |\,\tilde{k}\,| = |\,\tilde{k}'\,| \tag{4.3.2}$$

wobei $\tilde{k}$ und $\tilde{k}'$ beliebige Teilmengen in $\tilde{K}$ sind. Zusammengenommen liefern die beiden Bedingungen eine Begründung, um das Wahrscheinlichkeitsmaß durch folgende Definition festzulegen:

$$\text{Für alle } \tilde{k} \subseteq \tilde{K}: \; \mathrm{P_a}(\tilde{k}) = \frac{|\,\tilde{k}\,|}{m} \tag{4.3.3}$$

Willkürlich ist dabei nur die Normierung $\mathrm{P_a}(\tilde{K}) = 1$.

3. Infolge dieser Definition ist $\mathrm{P_a}$ eine additive Bewertungsfunktion im Sinne der in Abschnitt 3.7.1 gegebenen Definition, d.h. es gelten die folgenden Regeln:

a) $0 \leq \mathrm{P_a}(\tilde{k}) \leq 1$

b) $\mathrm{P_a}(\tilde{K}) = 1$

c) $\mathrm{P_a}(\tilde{k}^c) = 1 - \mathrm{P_a}(\tilde{k})$

d) $\tilde{k} \cap \tilde{k}' = \emptyset \implies \mathrm{P_a}(\tilde{k} \cup \tilde{k}') = \mathrm{P_a}(\tilde{k}) + \mathrm{P_a}(\tilde{k}')$

4. Die durch (4.3.3) definierte Funktion wird ein *aleatorisches Wahrscheinlichkeitsmaß* genannt. Es handelt sich um ein spezielles Wahrscheinlichkeitsmaß, das auf der Voraussetzung eines elementaren Zufallsgenerators beruht. In Abschnitt 4.5 wird besprochen, wie aleatorische Wahrscheinlichkeitsmaße konstruiert werden können, die nicht zu einer Gleichverteilung führen.

4.3.2 Aleatorische Wahrscheinlichkeitsaussagen

1. Das im vorangegangenen Abschnitt eingeführte aleatorische Wahrscheinlichkeitsmaß knüpft unmittelbar an eine Idee gleicher Möglichkeiten an, wie sie durch einen elementaren Zufallsgenerator expliziert wird. Insoweit handelt es sich nur darum, einer solchen Idee gleicher Möglichkeiten eine mathematische Ausdrucksform zu geben. Aleatorische Wahrscheinlichkeitsaussagen sind also Aussagen über Möglichkeiten – oder auch „Realisierungschancen" – von Situationen, die mit einem Zufallsgenerator erzeugt werden können.

2. Es sind infolgedessen *keine* Aussagen über relative Häufigkeiten, mit denen Situationen des einen oder anderen Typs eintreten. Es gibt keinen unmittelbaren Zusammenhang zwischen aleatorischen Wahrscheinlichkeiten und relativen Häufigkeiten. „Der Satz, daß das, was gleich möglich ist, auch gleich oft eintreten werde, läßt sich nicht aus der bloßen Voraussetzung herleiten, daß gleich mögliche Fälle vorliegen." (Bruns 1906, S. 14) Er läßt sich tatsächlich überhaupt nicht herleiten;[12] und wollte man ihn als ein „Prinzip" annehmen, wäre es falsch.[13]

3. Die Aussagen, die mit einem aleatorischen Wahrscheinlichkeitsmaß gemacht werden können, sind also keine Aussagen über relative Häufigkeiten. Aber es handelt sich auch nicht um epistemische Wahrscheinlichkeitsaussagen, die sich auf Hypothesen beziehen. Es sind vielmehr Aussagen über Eigenschaften eines Zufallsgenerators. Es ist also wichtig, aleatorische und epistemische Wahrscheinlichkeitsaussagen zu unterscheiden.[14]

---

[12] Bruns beruft sich stattdessen auf „Erfahrung"; seine Bemerkung geht so weiter: „Unter solchen Umständen bleibt nur übrig, die Begründung des Satzes auf die Aussagen der Erfahrung zu stützen und ihn als eine durch *Induktion* gewonnene Verallgemeinerung *beobachteter* Tatsachen anzusehen." Auf diesem Gedanken beruht die sog. frequentistische Wahrscheinlichkeitskonzeption, mit der wir uns in Kapitel 6 beschäftigen werden.

[13] Folgende Bemerkung von Fry (1928, S. 82) weist darauf hin: „The reader who has dealt at all with statistics will already have remarked that nothing whatever has been said about the frequency with wich an event will happen. It has been said that in tossing a penny heads is 'equally likely' to appear as tails; never that 'heads will appear as often as tails'. The reason is, that the latter statement is not true. Try it and see. It is not even true 'in the long run', unless that phrase is given a rather unusual shade of meaning which is dangerously near begging the question. Neither is it true that 'in a large number of independent runs, heads will as often exceed tails as tails will exceed heads'; for that is merely shifting attention from the event 'a single throw' to the event 'a run', and weak logic is never made stronger by obfuscation. Again the answer is, 'Try it and see'."

[14] Ähnliche Unterscheidungen sind von vielen Autoren vorgeschlagen worden, wobei allerdings ganz unterschiedliche Auffassungen über die beiden Arten von Wahrscheinlichkeit vertreten worden sind. Insbesondere wird oft anstelle eines aleatorischen Wahrscheinlichkeitsbegriffs eine Deutung durch beobachtbare relative Häufigkeiten vorgeschlagen (damit beschäftigen wir uns in Kapitel 6); auf einige unterschiedliche Auffassungen epistemischer Wahrscheinlichkeiten wurde bereits in Abschnitt 3.8 hingewiesen. Folgende Bemerkungen von Braithwaite (1955, S. 120) zeigen die Vielfalt des Sprachgebrauchs: „Many of the philosophers who have recently emphasized the distinction between the two concepts of probability (and these philosophers have held quite different views as to the analyses of the two concepts) have proposed to use different names for the two concepts. Probability in the first sense – probability within a science – has been renamed by Rudolf Carnap *relative frequency*, by Bertrand Russell *mathematical probability* and by William Kneale *chance*. It has also

4. Ein Zusammenhang mit epistemischen Wahrscheinlichkeitsaussagen kann jedoch auf einfache Weise hergestellt werden. Denn ein Zufallsgenerator ist ein Verfahren, mit dem sich Situationen erzeugen lassen; und über diese Situationen können Hypothesen gebildet werden. Solche Hypothesen haben die Form

$$\langle\, s, \tilde{k}\,\rangle$$

wobei $\tilde{k}$ eine Teilmenge der Kennzeichnungsmenge $\tilde{K}$ des Zufallsgenerators ist. Also kann man durch die Definition

$$\mathrm{P_e}(\langle\, s, \tilde{k}\,\rangle) \;:=\; \mathrm{P_a}(\tilde{k}) \tag{4.3.4}$$

auch ein epistemisches Wahrscheinlichkeitsmaß einführen.

5. Durch diese Definition werden die beiden Wahrscheinlichkeitsmaße formal identifiziert; sie haben aber dennoch eine unterschiedliche Bedeutung.[15] Mit dem aleatorischen Wahrscheinlichkeitsmaß $\mathrm{P_a}$ können Aussagen über einen Zufallsgenerator gemacht werden. Das epistemische Wahrscheinlichkeitsmaß $\mathrm{P_e}$ soll dagegen dazu dienen, um epistemische Wahrscheinlichkeitsaussagen zu begründen, die sich auf Situationen beziehen, die mit dem Zufallsgenerator erzeugt werden können. Das in Abschnitt 3.7.2 erläuterte Argumentationsschema verdeutlicht die intendierte Verwendung:

$$\mathrm{P_e}(\langle\, s, \tilde{k}\,\rangle) \leq \mathrm{P_e}(\langle\, s, \tilde{k}'\,\rangle) \;\rightsquigarrow\; \langle\, s, \tilde{k}\,\rangle \preceq_e \langle\, s, \tilde{k}'\,\rangle \tag{4.3.5}$$

Orientiert man sich an diesem Schema, soll das epistemische Wahrscheinlichkeitsmaß $\mathrm{P_e}$ Argumentationsmöglichkeiten verschaffen, um die Wahrscheinlichkeit von Hypothesen über Situationen, die mit einem Zufallsgenerator erzeugt werden, einschätzbar zu machen.

---

been called *statistical probability* and *empirical probability*. There seem to me to be objections to all these alternative names; so I shall use *probability* without any qualification for probability within a scientific hypothesis. The other sort of probability, which Carnap calls *confirmation*, Russell *credibility* and Kneale *acceptability*, I shall straighforwardly call *reasonableness*, and shall neither call it probability nor suppose that it is a quantity measurable by a number.“ Auch Braithwaite ist hier offenbar der Auffassung, daß epistemische Wahrscheinlichkeit nicht quantifiziert werden kann. Sein terminologischer Vorschlag, von *reasonableness* zu sprechen, paßt auch gut zu unserem Verständnis epistemischer Wahrscheinlichkeit.

[15] Es sei erwähnt, daß in der Literatur zur Wahrscheinlichkeitstheorie selten eine klare Unterscheidung gemacht wird. Zum Beispiel bezieht sich Bernoulli in seinen Rechnungen stets auf aleatorische Wahrscheinlichkeiten, sagt aber zu Beginn des 4. Teils seiner *Ars conjectandi*: „Die Wahrscheinlichkeit ist nämlich ein Grad der Gewissheit und unterscheidet sich von ihr wie ein Theil vom Ganzen.“

6. Wie in Abschnitt 3.7.2 überlegt worden ist, liefert ein epistemisches Wahrscheinlichkeitsmaß $\mathrm{P_e}$ im Kontext des Schemas (4.3.5) stets nur Argumentationsmöglichkeiten, ohne daß ausgeschlossen wird, daß für epistemische Wahrscheinlichkeitsaussagen noch weiteres Wissen über eine Situation relevant werden kann, das in $\mathrm{P_e}$ nicht berücksichtigt worden ist. Wenn man sich auf Situationen bezieht, die durch Zufallsgeneratoren erzeugt werden, liegt jedoch in gewisser Weise ein besonderer Fall vor. Denn durch ihre Konstruktion und durch die Regeln für ihren Gebrauch soll sichergestellt werden, daß für Wahrscheinlichkeitsaussagen *ausschließlich* das Wissen über die Konstruktion des Zufallsgenerators, nicht jedoch irgendein zusätzliches Wissen über die Situationen, die mit ihm erzeugt worden sind oder erzeugt werden können, verwendet werden kann. Insofern erscheint in diesem speziellen Fall die Auffassung begründbar, daß für Wahrscheinlichkeitsaussagen ausschließlich das durch (4.3.4) definierte epistemische Wahrscheinlichkeitsmaß $\mathrm{P_e}$ verwendet werden sollte.

7. Aber so plausibel diese Auffassung auch erscheinen mag, man sollte sich klarmachen, daß mit ihr der Kontext einer aleatorischen Wahrscheinlichkeitstheorie verlassen und vielmehr eine Rationalitätsnorm formuliert wird. Es handelt sich um einen Vorschlag, wie man vernünftigerweise Wahrscheinlichkeitsaussagen über Hypothesen machen sollte, die sich auf Situationen beziehen, die durch einen Zufallsgenerator erzeugt werden. Wie schon in Abschnitt 4.1 exemplarisch durch einen Verweis auf J. Bernoulli bemerkt worden ist, richtete sich die Intention vieler Autoren, die an der Entwicklung der Wahrscheinlichkeitstheorie mitgewirkt haben, oft darauf, Rationalitätsnormen zu begründen.[16] Tatsächlich vertreten auch heute noch viele Autoren die Auffassung, daß die Wahrscheinlichkeitstheorie Rationalitätsnormen für eine Bewertung von Hypothesen begründen könne; einige Autoren sind sogar der Auffassung, daß Normen für „rationale Entscheidungen“ begründet werden können. Aber der Glaube, daß mit den begrifflichen Hilfsmitteln

---

[16] Zur Illustration sei hier eine Bemerkung von Laplace (1814, S. 82f) angeführt: „Setzen wir den unvermeidlichen Wirkungen des Fortschrittes der Aufklärung keinen nutzlosen und oft gefährlichen Widerstand entgegen; aber wechseln wir nur mit außerordentlicher Behutsamkeit unsere Einrichtungen und Gebräuche, denen wir uns seit langer Zeit angepaßt haben. Wir kennen zwar aus der Erfahrung der Vergangenheit die Übelstände, die sie mit sich bringen, aber wir kennen das Ausmaß der Übel nicht, die eine Abänderung derselben verursachen kann. Bei dieser Unwissenheit schreibt die Wahrscheinlichkeitstheorie vor, jede Veränderung zu unterlassen, namentlich aber plötzliche Änderungen zu vermeiden, die in der moralischen wie physischen Ordnung sich niemals ohne bedeutenden Verlust an lebendiger Kraft vollziehen.“

der Wahrscheinlichkeitstheorie Rationalitätsnormen formuliert oder sogar begründet werden könnten, ist fragwürdig.[17] Die Frage, was möglicherweise „rational" genannt werden kann, ist sicherlich wichtig, würde aber eine philosophische Diskussion erfordern, die sich nicht auf Überlegungen zur Wahrscheinlichkeitstheorie einschränken ließe. Fragen, die Rationalitätsnormen betreffen, werden deshalb in diesem Text ausgeklammert; und nur gelegentlich wird darauf hingewiesen, wenn sie in problematischer Weise in Gedankengänge der Wahrscheinlichkeitstheorie einfließen.

4.3.3 Einwände und Mißverständnisse

1. Die Idee, aleatorische Wahrscheinlichkeitsmaße durch Rückgriff auf Zufallsgeneratoren einzuführen, die eine vorgängige Explikation gleicher Möglichkeiten erlauben, erschien in den Anfängen der Wahrscheinlichkeitstheorie – wie in Abschnitt 4.1 exemplarisch am Beispiel J. Bernoullis gezeigt wurde – weitgehend selbstverständlich. Insofern kann man diesen Gedankengang auch als „klassische" Definition eines quantitativen Wahrscheinlichkeitsbegriffs bezeichnen. Im folgenden sollen einige Einwände und Mißverständnisse besprochen werden, die im Hinblick auf diese Definition aufgetreten sind.[18]

2. Ein erster Einwand richtet sich dagegen, daß man sich mit einem aleatorischen Wahrscheinlichkeitsmaß P_a nur auf Situationen beziehen

[17] Hier sei noch ein Beispiel erwähnt, an dem man sich die Problematik überlegen kann. Bei Carnap und Stegmüller (1959, S. 10) heißt es: Die induktive Logik „hat nicht nur Verfahren zu konstruieren, um den Bestätigungsgrad von Hypothesen zu beurteilen, sondern auch Regeln für die Vornahme von Schätzungen aufzustellen. [...] Die Gewinnung einer allgemeinen Schätzungsmethode ist nicht nur in theoretischer Hinsicht bedeutsam, sondern vor allem auch für die Frage der Vornahmer rationaler Entschlüsse von großer Wichtigkeit. Angenommen, eine im Wirtschaftsleben stehende Person habe zwischen verschiedenen möglichen Handlungen zu wählen. Für jede der Handlungen wird sie versuchen, den vermutlichen Geldgewinn zu schätzen. Falls sie imstande ist, die Bestätigungsgrade aller möglichen Resultate einer in Erwägung gezogenen Handlung zu bestimmen, so kann sie die Summe der Gewinne berechnen, wobei jeder Gewinn mit dem dazugehörigen Bestätigungsgrad zu multiplizieren ist. Das Ergebnis dieser Kalkulation wird dann auf Grund der erwähnten Definition gerade die Schätzung des Gewinnes aus dieser Handlung darstellen. In analoger Weise kann die Person die Schätzungen des Gewinnes für die anderen Handlungen ermitteln. Wenn sie rational vorgehen will, muß sie jene Handlung wählen, für welche der geschätzte Gewinn am größten ist. An diesem Beispiel zeigt sich zugleich die praktisch-ökonomische Funktion der induktiven Logik."

[18] Eine übersichtliche Zusammenstellung der Einwände findet sich bei Gottinger 1980, S. 63f.

könne, bei denen sinnvoll von „gleichen Möglichkeiten" gesprochen werden kann. Zum Beispiel heißt es bei R. v. Mises (1981, S. 80):[19]

> "Equally possible cases do not always exist, e.g., they are not present in the game with a biased die, or in life insurance. Strictly speaking, the propositions of the classical theory are therefore not applicable to these cases."

Daran ist richtig, daß sich die Definition von P_a zunächst auf elementare Zufallsgeneratoren bezieht, die gerade dem Zweck dienen sollen, die Idee *gleicher* Möglichkeiten zu explizieren. Die Konzeption von Zufallsgeneratoren kann jedoch verallgemeinert werden, so daß mit ihrer Hilfe weitgehend beliebige Wahrscheinlichkeitsmaße definiert werden können. Damit werden wir uns in Abschnitt 4.5 beschäftigen.

3. Eher ein Mißverständnis kommt in der folgenden Bemerkung von Moore (1991, S. 334f) zum Ausdruck:

> "You may feel that it is obvious from the balance of a coin that the probability of a head is about 1/2. Such opinions are not always correct. Take a penny and, instead of tossing it, hold it on end on a hard surface with the index finger of one hand and snap it with the other index finger. The coin will spin for several seconds and then fall with either heads or tails showing. A long series of trials reveals that the probability of a head in this random experiment is clearly different from 1/2. Moral: We defined probability *empirically*, that is, in terms of observations. Only by observation can we be reasonably sure of the approximate value of the probability of an outcome."

Sicherlich wird man Moore darin zustimmen, daß nicht jede beliebige Methode zur Erzeugung von Ereignissen als ein (elementarer) Zufallsgenerator aufgefaßt werden kann. Das liefert jedoch keine Rechtfertigung für Moores Schlußfolgerung. Der wesentliche Punkt ist vielmehr, daß sich *einige* praktische Verfahren angeben lassen, die als Beispiele elementarer Zufallsgeneratoren gelten und mithin zur Explikation eines aleatorischen Wahrscheinlichkeitsmaßes dienen können.

4. An dieser Stelle wird indessen ein Problem sichtbar, auf das bereits in Abschnitt 4.1 im Anschluß an die Ausführungen Bernoullis hingewiesen worden ist. Ein aleatorisches Wahrscheinlichkeitsmaß kann sinnvoll verwendet werden, *wenn* man sich auf einen Zufallsgenerator beziehen kann. Aber wenn das nicht der Fall ist, wenn man es mit Ereignissen zu tun hat, die – in den Worten Bernoullis – „von dem Walten der Natur oder von der Willkür der Menschen abhängen", ist zunächst ganz unklar,

[19] Dieser Einwand wird hauptsächlich von Autoren vorgetragen, die eine „Häufigkeitsinterpretation" des Wahrscheinlichkeitsbegriffs vorschlagen (Kapitel 6); vgl. auch z.B. Nagel 1939, S. 45.

ob und ggf. wie ein aleatorischer Wahrscheinlichkeitsbegriff sinnvoll angewendet werden könnte. Bernoullis Vorschlag war, in diesen Fällen Wahrscheinlichkeitsaussagen durch empirisch ermittelbare relative Häufigkeiten zu begründen.

5. Die Entwicklung der Wahrscheinlichkeitstheorie ist zunächst nicht dem Vorschlag Bernoullis gefolgt. Stattdessen ist versucht worden, den aleatorischen Wahrscheinlichkeitsbegriff auf spekulative Weise allgemein anwendbar zu machen. Zum Beispiel hat bereits A. de Moivre (1756, S. 1) den Wahrscheinlichkeitsbegriff so definiert:

> "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."

Einerseits nimmt die Definition in unspezifischer Weise auf beliebige Ereignisse bezug, andererseits verwendet sie Begriffsbildungen, die nur durch eine Bezugnahme auf Zufallsgeneratoren sinnvoll expliziert werden können. So bleibt schließlich ganz unklar, was man sich unter einer „number of chances" vorstellen soll.

6. Besonders einflußreich ist eine ganz analog konzipierte Definition geworden, die von P. S. de Laplace in seinem *Essai philosophique sur les probabilités* (1814) gegeben worden ist:

> „Die Theorie des Zufalls ermittelt die gesuchte Wahrscheinlichkeit eines Ereignisses durch Zurückführung aller Ereignisse derselben Art auf eine gewisse Anzahl gleich möglicher Fälle, d. s. solcher, über deren Existenz wir in gleicher Weise unschlüssig sind, und durch Bestimmung der dem Ereignis günstigen Fälle. Das Verhältnis dieser Zahl zu der aller möglichen Fälle ist das Mass dieser Wahrscheinlichkeit, die also nichts anderes als ein Bruch ist, dessen Zähler die Zahl der günstigen Fälle und dessen Nenner die Zahl aller möglichen Fälle ist." (Laplace 1814, S. 4)

Auch in diesen Ausführungen gibt es keinerlei Bezugnahme auf einen Zufallsgenerator, durch den eine Vorstellung „gleicher Möglichkeiten" expliziert werden könnte. Stattdessen wird suggeriert, daß eine solche Vorstellung bei Ereignissen beliebiger Art sinnvoll gebildet werden kann.

7. Infolge der Definitionen von Moivre und Laplace ist der „klassische" aleatorische Wahrscheinlichkeitsbegriff weitgehend aufgelöst und mit einem fragwürdigen Prinzip vermischt worden, das in der Geschichte der Philosophie unter dem Namen „Prinzip des unzureichenden Grundes" diskutiert und kritisiert worden ist.[20] Das Prinzip besagt, daß man

[20] Es wurde bereits in der Antike diskutiert, vgl. Makin 1993. Ausführliche historische Ausführungen finden sich bei Keynes (1921), der von einem „principle of indifference"

berechtigt sei, zwei oder mehr Ereignisse (oder Sachverhalte) schon dann für „gleich möglich" zu halten, wenn man keinen Grund angeben kann, daß eines der Ereignisse „eher möglich" sei als ein anderes, oder in der oben zitierten Formulierung von Laplace: wenn man über die Ereignisse „in gleicher Weise unschlüssig" ist. Man kann allerdings bezweifeln, daß dieses „Prinzip des unzureichenden Grundes" eine tragfähige Grundlage für Wahrscheinlichkeitsaussagen liefern kann. In prägnanter Weise hat das Leslie Ellis bereits 1842 so gesagt: „Mere ignorance is no ground for any inference whatever. *Ex nihilo nihil.*"[21] Diese Kritik wurde dann vor allem von J. von Kries (1886) ausgearbeitet und später von vielen anderen Autoren unterstützt.[22]

8. Durch die Vermengung der „klassischen" Definition eines aleatorischen Wahrscheinlichkeitsmaßes mit dem „Prinzip des unzureichenden Grundes" sind in der Geschichte der Wahrscheinlichkeitstheorie viele Unklarheiten entstanden;[23] noch bedauerlicher ist, daß infolgedessen auch die Begründung des aleatorischen Wahrscheinlichkeitsbegriffs durch eine explizite Konzeption von Zufallsgeneratoren unklar gemacht worden ist. Zwar ist es in der Lehrbuchliteratur bis heute üblich, einen quantitativen Wahrscheinlichkeitsbegriff durch Hinweise auf Glücksspiele und Zufallsgeneratoren zu motivieren. In der neueren Literatur wird aber die „klassische" Definition meistens nicht explizit entwickelt, sondern stattdessen auf Laplace und das „Prinzip des unzureichenden Grundes" verwiesen, um schließlich eine frequentistische Wahrscheinlichkeitsdeutung als eine vergleichweise sinnvolle Alternative darzustellen. Ob es sich um eine sinnvolle Alternative handelt, werden wir in Kapitel 6 besprechen. Hier soll nur noch einmal betont werden, daß der aleatorische Wahrscheinlichkeitsbegriff vor allem deshalb interessant ist, weil er sich nicht auf relative Häufigkeiten, sondern auf Konstruktionsprinzipien von Zufallsgeneratoren stützt. Wie wir sehen werden, wird es auch erst dadurch möglich, einen einsichtigen Zusammenhang zwischen aleatorischen Wahrschein-

spricht.

[21] Zitiert bei Keynes 1921, S. 85.

[22] Eine gute Übersicht über die Diskussion findet sich bei Czuber 1923, Kap. II. Es sei erwähnt, daß das „Prinzip des unzureichenden Grundes" nicht immer nur im Sinne eines Nicht-Wissens interpretiert wird, sondern von einigen Autoren auch mit begründbaren Symmetrieannahmen verknüpft wird. In dieser Weise hat schon Kries (1886) versucht, eine begründbare Variante des Prinzips zu verteidigen. Neuere Überlegungen, die die gleiche Intention verfolgen, finden sich u.a. bei Struik 1934, Lucas 1970, Kap. VII.

[23] Viele Hinweise auf solche Unklarheiten und Zweideutigkeiten finden sich bei Hacking, 1971.

lichkeiten und relativen Häufigkeiten herzustellen.

9. An dieser Stelle setzt indessen ein weiteres Mißverständnis ein, auf das kurz hingewiesen werden soll. Es äußert sich in dem Vorwurf, daß die „klassische" Definition des aleatorischen Wahrscheinlichkeitsbegriffs zirkulär sei, weil sie „gleich wahrscheinlich" durch „gleich möglich" definiere.[24] In gewisser Weise ist diese Feststellung durchaus zutreffend. Der aleatorische Wahrscheinlichkeitsbegriff gewinnt seine Bedeutung durch ein vorgängiges Verständnis der Bedeutung „gleicher Möglichkeiten"; „gleich wahrscheinlich" wird durch „gleich möglich" definiert und hat mithin die gleiche Bedeutung. Dennoch ist es ein Mißverständnis, hierin eine zirkuläre Definition zu sehen. Denn folgt man der „klassischen" Definition, kann ja gerade durch eine Bezugnahme auf operativ herstellbare Zufallsgeneratoren expliziert werden, was mit „gleich möglich" gemeint sein soll. Es ist diese Fundierung in praktisch herstellbaren Zufallsgeneratoren, die den Begriffsbildungen der aleatorischen Wahrscheinlichkeitsrechnung einen ausweisbaren Sinn gibt.

4.4 Zufallsvariablen

In der neueren Literatur zur Wahrscheinlichkeitsrechnung wird selten von Zufallsgeneratoren gesprochen, bei den Begriffsbildungen wird stattdessen von „Zufallsvariablen" ausgegangen. Bei einer rein formalen Betrachtung mag dies sinnvoll sein; denn eine Zufallsvariable kann dann einfach als eine mathematische Variable konzipiert werden, für deren Wertebereich eine Mengenalgebra und eine additive Bewertungsfunktion unterstellt werden kann. Wenn jedoch diese additive Bewertungsfunktion als ein Wahrscheinlichkeitsmaß interpretiert werden soll, wird es erforderlich, den Begriff einer Zufallsvariablen mit inhaltlich ausweisbaren Vorstellungen über Situationen und Ereignisse zu verknüpfen. Dem Ansatz dieses Kapitels entsprechend verwenden wir hierfür Zufallsgeneratoren.

4.4.1 Einfache Zufallsvariablen

1. Wir beziehen uns also im folgenden auf einen Zufallsgenerator $\mathcal{G} := (\mathcal{V}, (\mathcal{S}, \tilde{K}))$. Ist dieser Ausgangspunkt gegeben, kann man Zufalls-

[24] Zum Beispiel spricht R. v. Mises (1931, S. 3) von einer „petitio principii [...], da 'gleichmöglich' doch nur ein Deckwort für gleichwahrscheinlich sein kann." Den gleichen Einwand formuliert auch Nagel 1939, S. 46.

variablen als Funktionen

$$X : \tilde{K} \longrightarrow \tilde{X} \tag{4.4.1}$$

definieren, die jedem Element $\tilde{k}$ der Kennzeichnungsmenge $\tilde{K}$ einen Wert $X(\tilde{k}) \in \tilde{X}$ zuordnen. In dieser Formulierung ist $\tilde{X}$ der Merkmalsraum, innerhalb dessen die Zufallsvariable X ihre Werte annehmen kann; und wie bei statistischen Variablen nehmen wir auch für Zufallsvariablen an, daß uns für den Merkmalsraum stets eine numerische Repräsentation zur Verfügung steht:

$$X : \tilde{K} \longrightarrow \tilde{X} \subseteq \mathbf{R}$$

Wir nehmen also an, daß die Merkmalsräume von Zufallsvariablen stets in die Menge der reellen Zahlen „eingebettet" sind.

2. Insofern gibt es eine formale Analogie zwischen Zufallsvariablen und statistischen Variablen. Es ist allerdings wichtig, die beiden Begriffsbildungen zu unterscheiden. Eine statistische Variable ist eine Funktion

$$X_{\mathrm{S}} : \mathrm{S} \longrightarrow \tilde{X}$$

die sich auf eine Menge S bezieht, die als eine in unserer Erfahrungswelt feststellbare Menge von Situationen (oder Objekten) vorstellbar ist. Jeder Situation $s \in \mathrm{S}$ wird durch eine solche Variable ein bestimmter Merkmalswert $X_{\mathrm{S}}(s)$ in einem Merkmalsraum $\tilde{X}$ zugeordnet. Statistische Variablen beziehen sich infolgedessen in ihrer Bedeutung auf Eigenschaften von Situationen oder Objekten in unserer Erfahrungswelt. Dagegen beziehen sich Zufallsvariablen auf die Elemente einer Kennzeichnungsmenge $\tilde{K}$, und sie haben infolgedessen zunächst überhaupt keinen Bezug zu realen Phänomenen. Ein solcher Bezug kommt nur indirekt dadurch zustande, daß die Elemente von $\tilde{K}$ zur Bildung von Hypothesen über Ereignisse oder Sachverhalte verwendet werden können. Diese Hypothesen haben die Form $\langle s, \tilde{k} \rangle$, wobei s eine reale oder vorgestellte Situation und $\tilde{k}$ ein Element oder eine Teilmenge der Kennzeichnungsmenge $\tilde{K}$ ist. Kann man sich dann auf eine Zufallsvariable X beziehen, die für diese Kennzeichnungsmenge definiert ist, kann man auch sie verwenden, um Hypothesen über Situationen zu formulieren.

3. Werden zur Formulierung von Hypothesen Zufallsvariablen verwendet, bezieht man sich nicht unmittelbar auf die Kennzeichnungsmenge $\tilde{K}$, sondern auf den Merkmalsraum (Wertebereich) der Zufallsvariablen, also in der hier verwendeten Notation auf den Merkmalsraum $\tilde{X}$, wenn X der Name der Zufallsvariablen ist. Mit jeder Teilmenge $\tilde{x} \subseteq \tilde{X}$ kann dann eine entsprechende Hypothese formuliert werden: daß die Zufallsvariable X bei einer Situation s einen Wert in $\tilde{x}$ annimmt. Ob eine solche Hypothese zutrifft oder nicht, hängt natürlich davon ab, welcher Wert in der Kennzeichnungsmenge $\tilde{K}$ realisiert wird. Also kann man die Formulierung von Hypothesen mithilfe von Zufallsvariablen in der folgenden Weise auf die bisher verwendeten Formulierungen für Hypothesen zurückführen:

$$\langle s, \tilde{x} \rangle := \langle s, \{\tilde{k} \in \tilde{K} \mid X(\tilde{k}) \in \tilde{x}\} \rangle \tag{4.4.2}$$

wobei $\tilde{x}$ eine beliebige Teilmenge aus dem Merkmalsraum $\tilde{X}$ ist.

4. Gelegentlich werden auch noch andere Formulierungen verwendet, zum Beispiel:

$$\langle s, X \in \tilde{x} \rangle := \langle s, \tilde{x} \rangle \quad \text{oder} \quad \langle s, X = \tilde{x} \rangle := \langle s, \{\tilde{x}\} \rangle$$

wobei $\tilde{x}$ im ersten Fall eine Teilmenge und im zweiten Fall ein Element aus $\tilde{X}$ ist. Weiterhin wird gelegentlich die Formulierung

$$X^{-1}(\tilde{x}) := \{\tilde{k} \in \tilde{K} \mid X(\tilde{k}) \in \tilde{x}\}$$

verwendet, um sich auf einfache Weise auf das Urbild einer Menge $\tilde{x}$ zu beziehen.

5. Die begriffliche Verknüpfung von Zufallsvariablen mit Zufallsgeneratoren kann schließlich verwendet werden, um auch für Zufallsvariablen aleatorische Wahrscheinlichkeitsmaße zu definieren. Hat man nämlich für den Zufallsgenerator ein aleatorisches Wahrscheinlichkeitsmaß

$$\mathrm{P_a} : \mathcal{P}(\tilde{K}) \longrightarrow \mathbf{R}$$

eingeführt, kann man ganz entsprechend auch eine Funktion

$$\mathrm{P_a}[X] : \mathcal{P}(\tilde{X}) \longrightarrow \mathbf{R} \tag{4.4.3}$$

konzipieren, die durch

$$\mathrm{P_a}[X](\tilde{x}) := \mathrm{P_a}(X^{-1}(\tilde{x})) \tag{4.4.4}$$

definiert ist.[25] Eine aleatorische Wahrscheinlichkeit für Mengen im Merkmalsraum von X wird dadurch auf eine vorgängig definierte aleatorische

[25]Wir verwenden die Schreibweise $\mathrm{P_a}[X]$, um damit anzudeuten, daß X hier als ein Attribut (Index), nicht als ein Argument für das Wahrscheinlichkeitsmaß verstanden werden soll.

Wahrscheinlichkeit für Teilmengen der Kennzeichnungsmenge eines Zufallsgenerators zurückgeführt.

6. Infolge dieser Definition ist $\mathrm{P_a}[X]$ ebenfalls eine additive Bewertungsfunktion und hat also die für ein Wahrscheinlichkeitsmaß erforderlichen formalen Eigenschaften. $\mathrm{P_a}[X]$ wird deshalb als ein *Wahrscheinlichkeitsmaß für die Zufallsvariable X* oder kurz als *Wahrscheinlichkeitsverteilung* von X bezeichnet. Im allgemeinen sind natürlich $\mathrm{P_a}[X]$ und $\mathrm{P_a}$ nicht identisch. In der Wahrscheinlichkeitsrechnung ist es jedoch üblich, dies nicht durch eine entsprechende Indizierung der Wahrscheinlichkeitsmaße, sondern nur durch jeweils geeignete Schreibweisen für die Argumente kenntlich zu machen; das heißt: es wird meistens die Schreibweise

$$\mathrm{P_a}(X \in \tilde{x}) \quad \text{anstelle von} \quad \mathrm{P_a}[X](\tilde{x})$$

verwendet. Im weiteren folgen wir dieser Konvention.

7. Der Sinn dieser Begriffsbildungen wird deutlich, wenn man Zufallsvariablen verwendet, um Hypothesen über Situationen zu formulieren, die mithilfe von Zufallsgeneratoren erzeugt werden können. Die argumentative Verknüpfung erfolgt so, wie dies in Abschnitt 4.3.2 bereits für Hypothesen der Form $\langle s, \tilde{k} \rangle$ besprochen worden ist. Dort wurde die Definition

$$\mathrm{P_e}(\langle s, \tilde{k} \rangle) := \mathrm{P_a}(\tilde{k}) \tag{4.4.5}$$

verwendet, um ein epistemisches Wahrscheinlichkeitsmaß für Hypothesen zu definieren. Ein epistemisches Wahrscheinlichkeitsmaß für Hypothesen, die mithilfe einer Zufallsvariablen formuliert werden, folgt dann aus (4.4.2), nämlich

$$\mathrm{P_e}(\langle s, \tilde{x} \rangle) = \mathrm{P_e}(\langle s, X^{-1}(\tilde{x}) \rangle) \tag{4.4.6}$$

Man findet dann auch einen zu (4.4.5) analogen Zusammenhang für Zufallsvariablen, nämlich

$$\mathrm{P_e}(\langle s, \tilde{x} \rangle) = \mathrm{P_a}(X \in \tilde{x}) \tag{4.4.7}$$

8. Folgt man diesen Überlegungen zum Begriff einer Zufallsvariablen, handelt es sich also nicht um einen Grundbegriff der Wahrscheinlichkeitstheorie, sondern nur um eine Begriffsbildung, die es erlaubt, sich in mathematisch flexibler Weise auf die Elemente der Kennzeichnungsmenge eines Zufallsgenerators zu beziehen. Dieser Punkt sollte betont werden, da (4.4.3) und (4.4.4) eine gewissermaßen „abgekürzte“ Definition von Zufallsvariablen zu erlauben scheinen, etwa so: Eine Zufallsvariable ist eine Variable X mit einem Merkmalsraum $\tilde{X}$, so daß sich für Aussagen der Form „$X \in \tilde{x}$“ ein Wahrscheinlichkeitsmaß unterstellen läßt. So heißt es zum Beispiel bei Tschuprow (1925, S. 20):[26]

„Der Begriff der zufälligen Variablen ordnet sich unter den allgemeinen mathematischen Begriff der variablen Größen als genus proximum ein, wobei als differentia specifica das Vorhandensein des Verteilungsgesetzes erscheint. Eine einziffrige Zahl ist eine diskontinuierliche Variable, welche zehn verschiedene Werte von 0 bis 9 annehmen kann. Sie wird zu einer zufälligen Variablen zehnter Ordnung, falls sie diese verschiedenen Werte mit bestimmten Wahrscheinlichkeiten annimmt.“

Bei Formulierungen dieser Art bleibt jedoch unklar, was es bedeuten soll, daß eine mathematische Variable ihre möglichen Werte mit bestimmten Wahrscheinlichkeiten annimmt. Ein Verständnis wird erst möglich, wenn man Zufallsvariablen begrifflich mit Hypothesen verknüpft, die sich auf Ereignisse oder Sachverhalte beziehen, die mit Zufallsgeneratoren erzeugt werden können.

9. Ist ein solcher Kontext explizit eingeführt worden, kann man natürlich auch bei Zufallsvariablen von einer Wahrscheinlichkeitsverteilung sprechen. Man kann sogar noch einen Schritt weiter gehen und Zufallsvariablen zur *Repräsentation* von Zufallsgeneratoren verwenden, indem man die Kennzeichnungsmenge eines Zufallsgenerators als Wertebereich einer Zufallsvariablen auffaßt. Man verwendet zur Definition der Zufallsvariablen dann einfach eine identische Abbildung auf der Kennzeichnungsmenge des Zufallsgenerators, also eine Abbildung

$$X : \tilde{K} \longrightarrow \tilde{K}$$

die durch $X(\tilde{k}) = \tilde{k}$ definiert ist. Somit kann man Zufallsgeneratoren auch als Verfahren zur Konstruktion von Zufallsvariablen auffassen. Soweit es in den folgenden Abschnitten nur um die Entwicklung eines formalen Begriffsrahmens geht, wird von dieser Korrespondenz zwischen Zufallsvariablen und Zufallsgeneratoren zur Vereinfachung der Redeweisen oft Gebrauch gemacht. Ein Rückgriff auf Zufallsgeneratoren als praktisch realisierbare Verfahren wird jedoch stets dann erforderlich, wenn bestimmte Eigenschaften von Wahrscheinlichkeitsmaßen für Zufallsvariablen begründet werden müssen.

[26] Zitiert bei Anderson (1929, S. 5), der sich dieser Definition anschließt.

4.4.2 Mehrdimensionale Zufallsvariablen

1. So wie bei statistischen Variablen ist es auch bei Zufallsvariablen oft sinnvoll, ihren Merkmalsraum explizit durch eine Kombination mehrerer einzelner Merkmalsräume zu konzipieren. Der Merkmalsraum einer Zufallsvariablen X hat dann die Darstellung

$$\tilde{X} = \tilde{X}_1 \times \cdots \times \tilde{X}_m$$

In dieser Schreibweise setzt sich der Merkmalsraum $\tilde{X}$ aus m Komponenten (Teilmerkmalsräumen) zusammen. Dementsprechend besteht auch die Zufallsvariable X aus m Komponenten und kann in der Form

$$X = (X_1, \ldots, X_m)$$

ausgedrückt werden. Explizit als Abbildung geschrieben hat man schließlich die Darstellung:

$$(X_1, \ldots, X_m) : \tilde{K} \longrightarrow \tilde{X}_1 \times \cdots \times \tilde{X}_m \tag{4.4.8}$$

Man spricht dann von einer *m-dimensionalen* Zufallsvariablen. Jedem Element $\tilde{k}$ aus der Kennzeichnungsmenge $\tilde{K}$ wird ein Wert $X(\tilde{k})$ im Merkmalsraum $\tilde{X}$ zugeordnet, der aus m Einzelwerten besteht:

$$X(\tilde{k}) = (X_1(\tilde{k}), \ldots, X_m(\tilde{k}))$$

Jede Komponente $X_j(\tilde{k})$ nimmt einen Wert im entsprechenden Teilmerkmalsraum $\tilde{X}_j$ an.

2. Denkt man daran, daß man auch die Kennzeichnungsmenge eines Zufallsgenerators als Wertebereich einer Zufallsvariablen auffassen kann, kann man also auch Zufallsgeneratoren mit mehrdimensionalen Kennzeichnungsmengen konzipieren. Exemplarisch kann man an eine Urne denken, deren Kugeln sowohl durch eine Nummer als auch durch eine Farbe unterschieden werden können. Bildet man mit einer solchen Urne einen Zufallsgenerator, kann er durch eine zweidimensionale Zufallsvariable $X = (X_1, X_2)$ mit dem Merkmalsraum $\tilde{X} = \tilde{X}_1 \times \tilde{X}_2$ repräsentiert werden, wobei $\tilde{X}_1$ die möglichen Nummern und $\tilde{X}_2$ die möglichen Farben der Kugeln umfaßt.

3. Die Notation für Wahrscheinlichkeitsaussagen verläuft ganz analog wie bei eindimensionalen Zufallsvariablen. Zur Verdeutlichung beziehen wir uns auf eine Teilmenge

$$\tilde{x} = \tilde{x}_1 \times \cdots \times \tilde{x}_m \subseteq \tilde{X}_1 \times \cdots \times \tilde{X}_m$$

des Merkmalsraums $\tilde{X}$ der m-dimensionalen Zufallsvariablen X. Als Schreibweise für Wahrscheinlichkeitsaussagen hat man dann:

$$\mathrm{P_a}(X \in \tilde{x}) = \mathrm{P_a}(X_1 \in \tilde{x}_1, \ldots, X_m \in \tilde{x}_m)$$

In dem eben angeführten Beispiel könnte man also

$$\mathrm{P_a}(X_1 \in \{3\}, X_2 \in \{\text{‘rot’}\}) \quad \text{oder auch} \quad \mathrm{P_a}(X_1 = 3, X_2 = \text{‘rot’})$$

verwenden, um sich auf die Wahrscheinlichkeit zu beziehen, daß aus der Urne eine rote Kugel mit der Nummer 3 gezogen wird.

4.4.3 Funktionen von Zufallsvariablen

1. Weitere Möglichkeiten zur Begriffsbildung können schließlich dadurch gewonnen werden, daß man Funktionen von Zufallsvariablen betrachtet. Im einfachsten Fall gibt es zum Beispiel eine Zufallsvariable X mit dem Merkmalsraum $\tilde{X}$ und eine Funktion

$$g : \tilde{X} \longrightarrow \tilde{Y}$$

die jedem Merkmalswert in $\tilde{X}$ einen Wert in einem neuen Merkmalsraum $\tilde{Y}$ zuordnet. Dann kann eine neue Zufallsvariable

$$Y : \tilde{K} \longrightarrow \tilde{Y}$$

gebildet werden, deren Werte durch die Festsetzung

$$Y(\tilde{k}) = g(X(\tilde{k}))$$

zustande kommen. Y ist dann eine Funktion von X, wofür zur Abkürzung auch $Y = g(X)$ geschrieben werden kann.

2. Als Funktion von X ist dann auch Y eine Zufallsvariable. Ihre Wahrscheinlichkeitsverteilung wird in der folgenden Weise durch die Wahrscheinlichkeitsverteilung von X induziert:

$$\mathrm{P_a}(Y \in \tilde{y}) \;=\; \mathrm{P_a}(X \in g^{-1}(\tilde{y}))$$

wobei $\tilde{y}$ eine beliebige Teilmenge des Merkmalsraums $\tilde{Y}$ ist.

3. Wichtig sind insbesondere Funktionen mehrdimensionaler Zufallsvariablen. Ausgangspunkt ist dann eine m-dimensionale Zufallsvariable $X = (X_1, \ldots, X_m)$ und eine Funktion

$$g : \tilde{X}_1 \times \cdots \times \tilde{X}_m \longrightarrow \tilde{Y}$$

die jedem Wert aus dem Merkmalsraum $\tilde{X} = \tilde{X}_1 \times \cdots \times \tilde{X}_m$ einen Wert in dem neuen Merkmalsraum $\tilde{Y}$ zuordnet. Dann können die Werte einer neuen Zufallsvariablen Y durch

$$Y(\tilde{k}) = g(X_1(\tilde{k}), \ldots, X_m(\tilde{k}))$$

gebildet werden.

4.5 Allgemeine Zufallsgeneratoren

Bisher wurden nur elementare Zufallsgeneratoren explizit eingeführt und als Ausgangspunkt zur Definition eines aleatorischen Wahrscheinlichkeitsmaßes verwendet. In diesem Abschnitt soll gezeigt werden, wie man auf konstruktive Weise zu weitgehend allgemeinen Zufallsgeneratoren gelangen kann.

4.5.1 Verwendung von Zufallsvariablen

1. Ausgangspunkt ist ein elementarer Zufallsgenerator

$$\mathcal{G} \;:=\; (\mathcal{V}, (\mathcal{S}, \tilde{X}))$$

Die Kennzeichnungsmenge $\tilde{X}$ wird als Wertebereich einer Zufallsvariablen X konzipiert (vgl. Abschnitt 4.4.1). Jede Funktion

$$g : \tilde{X} \longrightarrow \tilde{Y}$$

die die Elemente des Wertebereichs $\tilde{X}$ in einen neuen Merkmalsraum $\tilde{Y}$ abbildet, kann verwendet werden, um einen neuen Zufallsgenerator

$$\mathcal{G}_{\mathrm{y}} \;:=\; (\mathcal{V}, (\mathcal{S}, \tilde{Y}))$$

zu definieren, dessen Kennzeichnungsmenge aus dem neuen Merkmalsraum $\tilde{Y}$ besteht. Dieser neue Zufallsgenerator kann durch eine Zufallsvariable

$$Y : \tilde{X} \longrightarrow \tilde{Y}$$

repräsentiert werden, die durch

$$Y(\tilde{x}) := g(X(\tilde{x}))$$

definiert ist. Ist $\mathrm{P_a}$ das zunächst für den Zufallsgenerator $\mathcal{G}$ definierte Wahrscheinlichkeitsmaß, erhält man durch

$$\mathrm{P_a}(Y \in \tilde{y}) \;=\; \mathrm{P_a}(X \in g^{-1}(\tilde{y})) \qquad (4.5.1)$$

ein Wahrscheinlichkeitsmaß für $\mathcal{G}_{\mathrm{Y}}$.

2. Der Gedankengang zeigt, wie man ausgehend von einem elementaren Zufallsgenerator neue Zufallsgeneratoren konstruieren kann. Faßt man Zufallsgeneratoren als Verfahren zur Konstruktion von Zufallsvariablen auf, zeigt der Gedankengang, wie man Zufallsvariablen mit beliebigen Wahrscheinlichkeitsverteilungen bilden kann. Möchte man eine Zufallsvariable Y mit einer beliebig vorgegebenen Wahrscheinlichkeitsverteilung konstruieren, ist es nur erforderlich, eine Funktion g zu ermitteln, die den Wertebereich eines elementaren Zufallsgenerators in den Merkmalsraum von Y abbildet und so beschaffen ist, daß die Gleichung (4.5.1) erfüllt wird.

3. Wie man das praktisch machen kann, hängt davon ab, welche technischen Hilfsmittel zur Konstruktion eines elementaren Zufallsgenerators zur Verfügung stehen. Hier verwenden wir zur Illustration eine Urne, die mit einer gewissen Anzahl von Kugeln gefüllt ist. Angenommen, man möchte eine Zufallsvariable Y mit dem Merkmalsraum $\tilde{Y} := \{1, \ldots, m\}$ und einer Wahrscheinlichkeitsverteilung

$$\mathrm{P_a}(Y = j) \;=\; \frac{n_j}{n}$$

konstruieren. Aus den Regeln für ein Wahrscheinlichkeitsmaß folgt dann $n = n_1 + \ldots + n_m$, und man kann eine Urne verwenden, die insgesamt n Kugeln enthält, wobei n_1 Kugeln durch die Nummer 1, n_2 Kugeln durch die Nummer 2, usw. kenntlich gemacht werden.

4.5.2 Unabhängige Verbindungen

1. Betrachten wir jetzt zwei elementare Zufallsgeneratoren

$$\mathcal{G}_1 := (\mathcal{V}_1, (\mathcal{S}_1, \tilde{X}_1)) \quad \text{und} \quad \mathcal{G}_2 := (\mathcal{V}_2, (\mathcal{S}_2, \tilde{X}_2))$$

Außerdem nehmen wir an, daß beide Zufallsgeneratoren unabhängig voneinander sind, womit folgendes gemeint sein soll: man kann mit beiden Zufallsgeneratoren jeweils gesondert beliebig viele Situationen (Ereignisse, Sachverhalte) erzeugen; und es gilt wechselseitig: wenn man einen der beiden Zufallsgeneratoren verwendet, verändert das in keiner Weise

die Bedingungen für die Verwendung und die Funktionsweise des jeweils anderen Zufallsgenerators. Man kann dann einen kombinierten Zufallsgenerator

$$\mathcal{G} := (\mathcal{V}, (\mathcal{S}, \tilde{X}_1 \times \tilde{X}_2))$$

bilden, der folgendermaßen funktioniert: es werden (entweder parallel oder seriell) die beiden Zufallsgeneratoren $\mathcal{G}_1$ und $\mathcal{G}_2$ aktiviert, und das Ergebnis wird als *eine* Situation aufgefaßt, die durch den Zufallsgenerator $\mathcal{G}$ entstanden ist. Für den neuen Zufallsgenerator gibt es dann offenbar eine zweidimensionale Kennzeichnungsmenge $\tilde{X}_1 \times \tilde{X}_2$; jede der durch ihn erzeugten Situationen wird durch einen zweidimensionalen Merkmalswert $(\tilde{x}_1, \tilde{x}_2) \in \tilde{X}_1 \times \tilde{X}_2$ charakterisiert. Man kann auch sagen, daß der neue Zufallsgenerator durch eine zweidimensionale Zufallsvariable (X_1, X_2) mit dem Merkmalsraum $\tilde{X}_1 \times \tilde{X}_2$ repräsentiert werden kann.

2. Wegen der für $\mathcal{G}_1$ und $\mathcal{G}_2$ vorausgesetzten Unabhängigkeit kann man leicht ein Wahrscheinlichkeitsmaß für diese zweidimensionale Zufallsvariable finden. Denn infolge dieser Voraussetzung haben alle Ereignisse, die mit dem kombinierten Zufallsgenerator $\mathcal{G}$ erzeugt werden können, die gleiche aleatorische Wahrscheinlichkeit, nämlich $1/(m_1 m_2)$, wobei $m_1 = |\tilde{X}_1|$ und $m_2 = |\tilde{X}_2|$ ist. Also folgt zunächst für beliebige Werte $(\tilde{x}_1, \tilde{x}_2) \in \tilde{X}_1 \times \tilde{X}_2$ die Gleichung

$$\mathrm{P_a}(X = (\tilde{x}_1, \tilde{x}_2)) \;=\; \mathrm{P_a}(X_1 = \tilde{x}_1)\,\mathrm{P_a}(X_2 = \tilde{x}_2) \qquad (4.5.2)$$

Aus den Regeln für Wahrscheinlichkeitsmaße folgt indessen hieraus auch für beliebige Teilmengen $\tilde{x}_1 \times \tilde{x}_2 \subseteq \tilde{X}_1 \times \tilde{X}_2$ eine entsprechende Gleichung:

$$\mathrm{P_a}(X \in \tilde{x}_1 \times \tilde{x}_2) \;=\; \mathrm{P_a}(X_1 \in \tilde{x}_1)\,\mathrm{P_a}(X_1 \in \tilde{x}_1) \qquad (4.5.3)$$

Wenn diese Gleichung für alle möglichen Merkmalsmengen in $\tilde{X}_1 \times \tilde{X}_2$ erfüllt ist, sagt man, daß die Zufallsvariablen X_1 und X_2 *stochastisch unabhängig* voneinander sind.

3. Man sollte darauf achten, daß die Unabhängigkeit von zwei Zufallsvariablen eine Begründung verlangt und daß eine solche Begründung nur dadurch gegeben werden kann, daß man sich auf die Zufallsgeneratoren bezieht, die durch die Zufallsvariablen repräsentiert werden sollen. Wir haben für die eben angestellte Überlegung vorausgesetzt, daß die beiden Zufallsgeneratoren $\mathcal{G}_1$ und $\mathcal{G}_2$ unabhängig voneinander sind. Damit ist gemeint, daß sich die beiden Zufallsgeneratoren wechselseitig in ihrer

Funktionsweise nicht beeinflussen, daß sie kausal unabhängig sind. Nur wenn man dafür eine Begründung angeben kann, kann man infolgedessen die beiden Zufallsvariablen im Sinne der Gleichung (4.5.3) stochastisch unabhängig nennen.

4. Das Konstruktionsverfahren kann in zweierlei Hinsicht verallgemeinert werden. Eine erste Verallgemeinerung ergibt sich daraus, daß es sich bei $\mathcal{G}_1$ und $\mathcal{G}_2$ um beliebige, nicht unbedingt elementare Zufallsgeneratoren handeln kann; also z.B. um Zufallsgeneratoren, die mit den in Abschnitt 4.5.1 besprochenen Methoden konstruiert worden sind. Wiederum kann dann ein kombinierter Zufallsgenerator gebildet werden, der durch eine zweidimensionale Zufallsvariable (X_1, X_2) repräsentiert wird. Kann vorausgesetzt werden, daß sich die beiden Zufallsgeneratoren in ihrer Funktionsweise wechselseitig nicht beeinflussen, sind auch die Komponenten dieser zweidimensionalen Zufallsvariablen im Sinne von Definition (4.5.3) stochastisch unabhängig. Die stochastische Unabhängigkeit von zwei Zufallsvariablen hängt also nicht von der Art ihrer Verteilung ab, sondern nur davon, ob man begründen kann, daß die Zufallsgeneratoren, auf die sich ihre Definition bezieht, unabhängig voneinander funktionieren.

5. Eine zweite Verallgemeinerung ergibt sich, wenn man von einer beliebigen Anzahl von Zufallsgeneratoren ausgeht:

$$\mathcal{G}_j := (\mathcal{V}_j, (\mathcal{S}_j, \tilde{X}_j)) \qquad (j = 1, \ldots, m)$$

Sie können zu einem kombinierten Zufallsgenerator

$$\mathcal{G} := (\mathcal{V}, (\mathcal{S}, \tilde{X}_1 \times \cdots \times \tilde{X}_m))$$

verbunden werden, der durch eine m-dimensionale Zufallsvariable

$$(X_1, \ldots, X_m)$$

repräsentiert werden kann. Kann vorausgesetzt werden, daß die Zufallsgeneratoren sich in ihrer Funktionsweise gegenseitig nicht beeinflussen, sind auch die Komponenten von X stochastisch unabhängig, d.h. es gilt:

$$\mathrm{P_a}(X_1 \in \tilde{x}_1, \ldots, X_m \in \tilde{x}_m) = \prod_{j=1}^{m} \mathrm{P_a}(X_j \in \tilde{x}_j) \qquad (4.5.4)$$

für beliebige $\tilde{x}_1 \times \cdots \times \tilde{x}_m \subseteq \tilde{X}_1 \times \cdots \times \tilde{X}_m$. Der Beweis kann durch vollständige Induktion geführt werden.

4.5.3 Mischungen

1. Schließlich können Zufallsgeneratoren auch durch Mischungen miteinander verbunden werden. Hierfür werden $m+1$ Zufallsgeneratoren

$$\mathcal{G}_j := (\mathcal{V}_j, (\mathcal{S}_j, \tilde{X}_j)) \qquad (j = 0, 1, \ldots, m)$$

vorausgesetzt. Es wird angenommen, daß die Kennzeichnungsmenge von $\mathcal{G}_0$ durch

$$\tilde{X}_0 = \{1, \ldots, m\}$$

charakterisiert werden kann. Dann kann ein neuer Zufallsgenerator $\mathcal{G} := (\mathcal{V}, (\mathcal{S}, \tilde{Y}))$ konstruiert werden, der folgendermaßen funktioniert: zuerst wird $\mathcal{G}_0$ aktiviert, dann wird abhängig von dem realisierten Wert $X_0 = j$ der Zufallsgenerator $\mathcal{G}_j$ verwendet, um einen Wert für die Zufallsvariable Y zu erzeugen, die den Zufallsgenerator $\mathcal{G}$ repräsentiert.

2. Die Aufgabe besteht darin, die Wahrscheinlichkeitsverteilung der Zufallsvariablen Y zu ermitteln. Ihr Wertebereich ist durch

$$\tilde{Y} = \tilde{X}_1 \cup \ldots \cup \tilde{X}_m$$

gegeben. Um die weiteren Überlegungen zu vereinfachen, wird angenommen, daß die Merkmalsräume $\tilde{X}_1, \ldots, \tilde{X}_m$ identisch sind.[27]

3. Um die Überlegung durchzuführen, ist es am einfachsten, sich die Zufallsgeneratoren als Urnen vorzustellen. Es wird also angenommen, daß $\mathcal{G}_j$ durch eine Urne gegeben ist, die

$$n_j = n_{j1} + \ldots + n_{jm_j}$$

Kugeln enthält, davon n_{jk} mit der Nummer k. Dem entsprechen die aleatorischen Wahrscheinlichkeitsmaße

$$\mathrm{P_a}(X_j = k) = \frac{n_{jk}}{n_j} \qquad (4.5.5)$$

Offenbar kann auch angenommen werden, daß in den Urnen für die Zufallsgeneratoren $\mathcal{G}_1, \ldots, \mathcal{G}_m$ jeweils gleich viele Kugeln enthalten sind, also $n_1 = \ldots = n_m$.

[27] Das ist keine Einschränkung, denn der Merkmalsraum einer Zufallsvariablen kann beliebig erweitert werden, indem man für die zusätzlich aufgenommenen Werte die Wahrscheinlichkeit Null annimmt.

4. Wird dies vorausgesetzt, gibt es insgesamt

$$n = \sum_{j=1}^{m} n_{0j} n_j$$

gleichwahrscheinliche Möglichkeiten dafür, wie eine Kugel gezogen werden kann, die einen Wert für die Zufallsvariable Y liefert. Fragt man nun nach $\mathrm{P_a}(Y = k)$, muß man sich überlegen, wieviele dieser Möglichkeiten zu einer Kugel mit der Nummer k führen. Offenbar lautet die Antwort:

$$\mathrm{P_a}(Y = k) \;=\; \frac{\sum_{j=1}^{m} n_{0j} n_{jk}}{\sum_{j=1}^{m} n_{0j} n_j}$$

Unter Verwendung von (4.5.5) findet man schließlich die Darstellung

$$\mathrm{P_a}(Y = k) \;=\; \sum_{j=1}^{m} \mathrm{P_a}(X_0 = j)\, \mathrm{P_a}(X_j = k) \tag{4.5.6}$$

die zeigt, wie die Wahrscheinlichkeitsverteilung von Y als eine durch X_0 vorgenommene Mischung der Wahrscheinlichkeitsverteilungen der Zufallsvariablen $X_1, \ldots, X_m$ gebildet wird.

4.6 Bedingte Wahrscheinlichkeiten

1. Schließlich muß noch der wichtige Begriff eines bedingten Wahrscheinlichkeitsmaßes besprochen werden. Dafür wird angenommen, daß man sich auf einen Zufallsgenerator beziehen kann, der durch eine Zufallsvariable X mit dem Wertebereich $\tilde{X}$ repräsentiert wird und für den es ein aleatorisches Wahrscheinlichkeitsmaß $\mathrm{P_a}$ gibt, um für alle Teilmengen $\tilde{x} \subseteq \tilde{X}$ Aussagen der Form $\mathrm{P_a}(X \in \tilde{x})$ machen zu können. Sei nun $\tilde{x}'$ eine weitere Teilmenge von $\tilde{X}$. Der Ausdruck

$$\mathrm{P_a}(X \in \tilde{x} \cap \tilde{x}')$$

gibt dann an, mit welcher Wahrscheinlichkeit X einen Wert annimmt, der sowohl in $\tilde{x}$ als auch in $\tilde{x}'$ liegt. Kann jetzt vorausgesetzt werden, daß $\mathrm{P_a}(\tilde{x}') > 0$ ist, kann man folgende Definition vornehmen:

$$\mathrm{P_a}(X \in \tilde{x} \,|\, X \in \tilde{x}') \;:=\; \frac{\mathrm{P_a}(X \in \tilde{x} \cap \tilde{x}')}{\mathrm{P_a}(X \in \tilde{x}')} \tag{4.6.1}$$

Dieser Ausdruck wird die *bedingte Wahrscheinlichkeit* von $X \in \tilde{x}$ unter der Bedingung $X \in \tilde{x}'$ genannt.

2. Die Definition (4.6.1) kann verwendet werden, um ein neues Wahrscheinlichkeitsmaß

$$\mathrm{P_a}(\ldots \,|\, X \in \tilde{x}') : \mathcal{P}(\tilde{X}) \;\longrightarrow\; \mathbf{R} \tag{4.6.2}$$

zu bilden, das jeder Teilmenge $\tilde{x} \subseteq \tilde{X}$ seine durch (4.6.1) definierte bedingte Wahrscheinlichkeit zuordnet. Offenbar handelt es sich um ein Wahrscheinlichkeitsmaß, denn es genügt allen dafür erforderlichen Regeln. Es wird ein *bedingtes Wahrscheinlichkeitsmaß* (unter der Bedingung $X \in \tilde{x}'$) genannt.

3. Zur Motivation der Begriffsbildungen kann man daran denken, daß mit einem aleatorischen Wahrscheinlichkeitsmaß epistemische Wahrscheinlichkeitsaussagen über Hypothesen begründet werden sollen, die sich auf Situationen beziehen, die durch den Zufallsgenerator erzeugt werden können. Bezieht man sich auf eine Hypothese $\langle s, \tilde{x} \rangle$, liefert $\mathrm{P_a}$ zunächst die Wahrscheinlichkeit $\mathrm{P_e}(\langle s, \tilde{x} \rangle) = \mathrm{P_a}(\tilde{x})$. Die epistemische Wahrscheinlichkeit hängt jedoch davon ab, welche Informationen über die Situation s verfügbar sind. $\mathrm{P_a}(\tilde{x})$ liefert ein sinnvolles Maß für die epistemische Wahrscheinlichkeit, wenn man nur weiß, daß die Situation s durch den Zufallsgenerator erzeugt worden ist. Sie verändert sich aber, wenn man zum Beispiel weiß, daß die Zufallsvariable X in der Situation s einen Wert in der Merkmalsmenge $\tilde{x}'$ angenommen hat. Dann liefert $\mathrm{P_a}(X \in \tilde{x} \,|\, X \in \tilde{x}')$ ein sinnvolles Maß für die epistemische Wahrscheinlichkeit $\mathrm{P_e}(\langle s, \tilde{x} \rangle)$.

4. Schließlich sollen noch drei Regeln besprochen werden, die für das Rechnen mit bedingten Wahrscheinlichkeiten oft hilfreich sind. Eine erste Regel lautet:

$$\text{Wenn } \tilde{x}_1, \ldots, \tilde{x}_m \subseteq \tilde{X} \text{ und } \mathrm{P_a}(\tilde{x}_1 \cap \ldots \cap \tilde{x}_m) > 0: \tag{4.6.3}$$

$$\mathrm{P_a}(\tilde{x}_1 \cap \ldots \cap \tilde{x}_m) \;=\; \prod_{j=1}^{m-1} \mathrm{P_a}(\tilde{x}_j \,|\, \tilde{x}_{j+1} \cap \ldots \cap \tilde{x}_m)\, \mathrm{P_a}(\tilde{x}_m)$$

Diese Regel kann durch Induktion unmittelbar aus der Definition bedingter Wahrscheinlichkeiten abgeleitet werden.

5. Eine zweite Regel geht von Mengen $\tilde{x}_1, \ldots, \tilde{x}_m \subseteq \tilde{X}$ aus, die paarweise disjunkt sind und für die außerdem $\mathrm{P_a}(\tilde{x}_j) > 0$ gilt. Weiterhin sei eine Menge $\tilde{x} \subseteq \tilde{x}_1 \cup \ldots \cup \tilde{x}_m$ gegeben. Dann gilt offenbar:

$$\tilde{x} = \tilde{x} \cap (\tilde{x}_1 \cup \ldots \cup \tilde{x}_m) = (\tilde{x} \cap \tilde{x}_1) \cup \ldots \cup (\tilde{x} \cap \tilde{x}_m)$$

Daraus folgt dann die Regel

$$\mathrm{P_a}(\tilde{x}) = \sum_{j=1}^{m} \mathrm{P_a}(\tilde{x} \cap \tilde{x}_j) = \sum_{j=1}^{m} \mathrm{P_a}(\tilde{x} \mid \tilde{x}_j)\, \mathrm{P_a}(\tilde{x}_j) \tag{4.6.4}$$

Diese Regel wird gelegentlich *Satz über die vollständige Wahrscheinlichkeit* genannt.

6. Eine dritte Regel verwendet die gleichen Voraussetzungen wie die zweite Regel und zusätzlich $\mathrm{P_a}(\tilde{x}) > 0$. Dann findet man zunächst:

$$\mathrm{P_a}(\tilde{x}_i \mid \tilde{x}) = \frac{\mathrm{P_a}(\tilde{x}_i \cap \tilde{x})}{\mathrm{P_a}(\tilde{x})} = \frac{\mathrm{P_a}(\tilde{x}_i)\, \mathrm{P_a}(\tilde{x} \mid \tilde{x}_i)}{\mathrm{P_a}(\tilde{x})}$$

Verwendet man zur Darstellung des Nenners (4.6.4), erhält man

$$\mathrm{P_a}(\tilde{x}_i \mid \tilde{x}) = \frac{\mathrm{P_a}(\tilde{x} \mid \tilde{x}_i)\, \mathrm{P_a}(\tilde{x}_i)}{\sum_{j=1}^{m} \mathrm{P_a}(\tilde{x} \mid \tilde{x}_j)\, \mathrm{P_a}(\tilde{x}_j)} \tag{4.6.5}$$

Diese Regel wird *Satz von Bayes* genannt und bildet den Ausgangspunkt für eine spezifische Variante statistischer Inferenzverfahren, oft 'Bayesianische Inferenz' genannt.[28]

[28] In diesem Skript beschäftigen wir uns ausschließlich mit einigen Grundbegriffen der Wahrscheinlichkeitstheorie. Eine Diskussion von Varianten und Problemen „statistischer Inferenz" soll in einem anderen Skript erfolgen.

Kapitel 5

Erwartungswerte und Wiederholungen

In Abschnitt 3.3 wurden epistemische Wahrscheinlichkeitsaussagen von modalen Aussagen unterschieden, die sich auf Möglichkeiten im Sinne von Dispositionen beziehen. Der aleatorische Wahrscheinlichkeitsbegriff, der im vorangegangenen Kapitel besprochen wurde, ist dagegen ausdrücklich ein modaler Begriff, zu dessen Explikation auf ein Verständnis von Möglichkeiten – insbesondere gleicher Möglichkeiten – als Dispositionen eines Zufallsgenerators zurückgegriffen wird. Daraus resultiert ein zentrales Problem der Wahrscheinlichkeitstheorie: Modalitäten sind nicht unmittelbar beobachtbar; wie also können modale Wahrscheinlichkeitsaussagen mit beobachtbaren Sachverhalten in einen Zusammenhang gebracht werden? In der Geschichte der (aleatorischen) Wahrscheinlichkeitstheorie sind vor allem zwei Überlegungen diskutiert worden, um zu einer Antwort zu gelangen.

a) Man kann einerseits an der Vorstellung festhalten, daß sich aleatorische Wahrscheinlichkeitsaussagen auf Möglichkeiten beziehen, so daß Aussagen dieser Art keine logisch ableitbaren Implikationen für Hypothesen haben, die sich auf Realisationen solcher Möglichkeiten richten. Infolgedessen erlauben aleatorische Wahrscheinlichkeitsaussagen nur Hypothesen über Situationen, die durch Zufallsgeneratoren erzeugt werden; oder anders formuliert: um einen Zusammenhang zwischen aleatorischen Wahrscheinlichkeiten und Realisationen möglicher Ereignisse herzustellen, sind wiederum epistemische Wahrscheinlichkeitsaussagen erforderlich.

b) Ein alternativer Gedankengang nimmt die „logische Kluft" zwischen klassisch konzipierten aleatorischen Wahrscheinlichkeiten und empirisch ermittelbaren Sachverhalten zum Anlaß, modal konzipierte Wahrscheinlichkeitsbegriffe grundsätzlich zu verwerfen und stattdessen zu versuchen, Wahrscheinlichkeiten durch relative Häufigkeiten zu definieren. Die Idee ist, einen Wahrscheinlichkeitsbegriff so zu konzipieren, daß Wahrscheinlichkeitsaussagen als Aussagen über relative Häufigkeiten verstanden werden können.

Mit dem zweiten Gedankengang, der zu frequentistischen Wahrscheinlichkeitskonzeptionen führt, werden wir uns im nächsten Kapitel beschäftigen. In diesem Kapitel folgen wir dem ersten Gedankengang. Dies erfordert zunächst eine Reihe von Vorüberlegungen, die sich sowohl auf Begriffsbildungen als auch auf die Idee der Wiederholbarkeit von Zufallsgeneratoren beziehen. Das wesentliche Argument wird erst im letzten Abschnitt im Anschluß an Bernoullis „Gesetz der großen Zahlen“ ausgeführt.

5.1 Erwartungswerte und Varianzen

Eine Art Schlüsselrolle für die Frage nach dem Verhältnis zwischen aleatorischen Wahrscheinlichkeiten und empirisch realisierten Situationen, die durch Zufalsgeneratoren entstehen können, kommt dem Begriff eines „Erwartungswertes“ zu. Eine Schlüsselrolle deshalb, weil dieser Begriff eine fundamentale Ambivalenz aufweist. Einerseits handelt es sich um einen Begriff der aleatorischen Wahrscheinlichkeitsrechnung und infolgedessen um einen modalen Begriff, der sich auf mögliche Realisationen von Zufallsgeneratoren bezieht. Andererseits suggeriert das Wort ein epistemisches Verständnis: etwas, das man „erwarten“ sollte. Dementsprechend wird dies Wort oft so verwendet, als ob mit seiner Hilfe ein Übergang von aleatorischen Möglichkeiten zu epistemischen Erwartungen hergestellt werden könne. Dadurch wird allerdings das Problem, auf das in den einleitenden Bemerkungen zu diesem Kapitel hingewiesen worden ist, eher verschleiert als aufgeklärt. Bei den folgenden Begriffserläuterungen sollte deshalb darauf geachtet werden, daß es sich zunächst ausschließlich um Begriffsbildungen der aleatorischen Wahrscheinlichkeitsrechnung handelt.

5.1.1 Erwartungswerte

1. Wenn X eine statistische Variable ist, wenn man sich also auf eine Abbildung $X : \mathrm{S} \longrightarrow \tilde{X}$ beziehen kann, kann man ihren arithmetischen Mittelwert folgendermaßen definieren:

$$\mathrm{M}(X) := \sum_{\tilde{x} \in \tilde{X}} \tilde{x}\, \mathrm{P_S}(X = \tilde{x}) \qquad (5.1.1)$$

Hier bezeichnet $\mathrm{P_S}$ die Häufigkeitsverteilung der statistischen Variablen X in der Menge S (vgl. Abschnitt 2.9). Ganz analog kann man für eine

Zufallsvariable $X : \tilde{K} \longrightarrow \tilde{X}$ den Ausdruck

$$\mathrm{E}(X) := \sum_{\tilde{x} \in \tilde{X}} \tilde{x}\, \mathrm{P_a}(X = \tilde{x}) \qquad (5.1.2)$$

definieren. Er wird *Erwartungswert* der Zufallsvariablen X genannt.

2. Hinter der formalen Analogie verbirgt sich ein wichtiger Bedeutungsunterschied. Mit dem Mittelwert einer statistischen Variablen kann man empirische Aussagen über eine Menge von Merkmalswerten (die als Eigenschaften von Situationen in S eine empirische Bedeutung haben) formulieren. Aber welcher Art sind demgegenüber Aussagen über Erwartungswerte von Zufallsvariablen? Sie beziehen sich nicht auf eine Menge real existierender Situationen, sondern auf Situationen, die mit einem Zufallsgenerator erzeugt werden könnten. Hier muß man jedoch auf die Formulierung achten. Denn sicherlich kann man einen Zufallsgenerator verwenden, um eine bestimmte Menge von Situationen (Ereignissen, Sachverhalten) zu erzeugen, etwa die Menge

$$\mathrm{S} = \{s_1, \dots, s_n\}$$

Dann kann man für jede dieser Situationen feststellen, welcher Wert der Zufallsvariablen X in ihr realisiert worden ist; und schließlich kann man auch den Mittelwert dieser Werte berechnen. Das Resultat ist dann jedoch ein *statistischer* Mittelwert entsprechend der Definition (5.1.1). Der Erwartungswert der Zufallsvariablen X muß von diesem statistischen Mittelwert unterschieden werden; schon deshalb, weil sich der statistische Mittelwert mit jeder neuen Situation, die man der Menge S hinzufügt, verändert.

3. Warum spricht man überhaupt von einem *Erwartungswert*? Die Bezeichnung stammt aus der Frühphase der Wahrscheinlichkeitsrechnung, in der man sich in erster Linie mit Glücksspielen beschäftigt hat und Zufallsvariablen verwendet wurden, um den möglichen Ergebnissen eines Glücksspiels Geldbeträge zuzuordnen, die ein Spieler – wenn das Ergebnis eintritt – gewinnt.[1] Man denke z.B. an ein Spiel mit einem

[1] Hinweise zur Begriffsgeschichte finden sich bei van der Wærden, 1975. In der älteren Literatur zur Wahrscheinlichkeitsrechnung wurde auch oft von einer „mathematischen Hoffnung“ gesprochen; z.B. findet sich bei Meyer (1879, S. 131) folgende Definition: „Die mathematische Hoffnung ist das Product aus dem erwarteten Gewinn mit der Wahrscheinlichkeit ihn zu erlangen. Man nennt sie auch die Erwartung oder den Vortheil des Spielers.“ In gewisser Weise gewinnen Erläuterungen dieser Art ihre Bedeutung dadurch, daß an ein Verständnis epistemischer Kategorien angeknüpft wird. Wie die folgende Passage aus einem älteren „Handbuch der Lebensversicherung“

Würfel, bei dem man mit einer geraden Augenzahl 1 DM und mit einer ungeraden Augenzahl 3 DM gewinnt. Dieses Spiel kann durch eine Zufallsvariable

$$X : \{1, 2, 3, 4, 5, 6\} \longrightarrow \{1, 3\}$$

repräsentiert werden, wobei

$$X(1) = X(3) = X(5) = 1 \quad \text{und} \quad X(2) = X(4) = X(6) = 3$$

ist. Der Erwartungswert von X beträgt dann $\mathrm{E}(X) = 2$ (DM). Aber heißt das, daß man erwarten kann, bei diesem Spiel 2 DM zu gewinnen? Das kann man so offenbar nicht sagen. Denn es gibt zwei Möglichkeiten: man kann 1 DM oder 3 DM gewinnen; beide Möglichkeiten haben die gleiche aleatorische Wahrscheinlichkeit 0.5, aber sicherlich wird man nicht 2 DM gewinnen.

4. Der Begriff 'Erwartungswert', so wie er durch (5.1.2) definiert wird, hat tatsächlich mit „Erwartungen" zunächst überhaupt nichts zu tun, sondern dient zur Formulierung aleatorischer Wahrscheinlichkeitsaussagen. Man kann sich das am besten anhand einer Zufallsvariablen X verdeutlichen, die nur die Werte 0 und 1 annehmen kann. Dann ist

$$\mathrm{E}(X) \;=\; \mathrm{P_a}(X = 1)$$

d.h. der Erwartungswert ist mit einer aleatorischen Wahrscheinlichkeit identisch.

zeigt, gibt es gleichermaßen einen Bezug zur Frage, wie ein Wertbegriff definiert werden kann: „Es ist einleuchtend, dass eine Summe Geldes, welche Jemand *gewiss* zu erhalten hat, einen ganz anderen Werth repräsentirt, als dieselbe Summe Geldes, welche er nur in dem Falle erhält, wenn ein *ungewisses* Ereignis eintritt. Diese letztere Summe hat allemal einen kleineren Werth, als die erstere, und zwar ist diese um so kleiner, je geringer die Wahrscheinlichkeit ist, dass das Ereignis eintreten wird, und umgekehrt. Nehmen wir an, dass in einer Urne 19 schwarze und 1 weisse Kugel sind, und es soll derjenige, welcher in einer Ziehung die weisse Kugel zieht, 40 Thaler erhalten. In dem vorliegenden Fall haben wir nur 1 günstigen Fall, 19 ungünstige und mithin 20 überhaupt mögliche Fälle. Wenn alle Kugeln gezogen würden, so würde die weisse gleichfalls gezogen werden und der Gewinn = 40 Thlr. würde absolut sicher sein. Da sämmtliche Fälle also einen Werth von 40 Thlr. repräsentiren, so ist es klar, dass 1 Fall nur 1/20 von jenem Werth hat; folglich hat der 1 günstige Fall (die weisse Kugel) nur einen Erwartungswerth = 1/20 von 40 Thaler, also 2 Thlr. [...] Der wirkliche Werth solcher ungewissen oder durch das etwaige Eintreffen eines oder mehrerer Ereignisse bedingten Geldsummen wird also gefunden, wenn man die Wahrscheinlichkeit, eine solche zu erhalten, mit dem absoluten Wert multiplicirt. Das Produkt heisst die *mathematische Hoffnung* (Erwartung) oder der *Geldwerth der mathematischen Wahrscheinlichkeit*." (Karup 1885, S. 184f.)

5. Ein weiterführender Gedanke bezieht sich darauf, daß Zufallsgeneratoren, die einen Anlaß zum Reden von Erwartungswerten geben können (z.B. Glücksspiele oder der Verkauf von Lebensversicherungen), meistens oft wiederholt werden. Man kann dann versuchen, den Erwartungswert als denjenigen Betrag aufzufassen, von dem man erwarten kann, daß man ihn bei einer häufigen Wiederholung eines Spiels im Durchschnitt gewinnen (oder verlieren) wird. Dementsprechend wird oft gesagt, daß der Erwartungswert einer Zufallsvariablen der zu erwartende Durchschnittswert einer größeren Anzahl von Realisationen der Zufallsvariablen ist. Diese Formulierung macht allerdings erneut deutlich, wie sich im Begriff eines Erwartungswertes aleatorische und epistemische Überlegungen vermischen. Wir werden uns später überlegen, wie ein Zusammenhang explizit hergestellt werden kann.

5.1.2 Varianzen und Kovarianzen

1. Auch der Begriff der Varianz einer Zufallsvariablen kann in Analogie zum Begriff der Varianz einer statistischen Variablen definiert werden. Ist X eine statistische Variable, kann man die Definition:

$$\mathrm{V}(X) \;:=\; \textstyle\sum_{\tilde{x}\in\tilde{X}} (\tilde{x} - \mathrm{M}(X))^2\, \mathrm{P_S}(X = \tilde{x}) \tag{5.1.3}$$

verwenden. Ganz analog lautet die Definition für die *Varianz einer Zufallsvariablen* X:

$$\mathrm{V}(X) \;:=\; \textstyle\sum_{\tilde{x}\in\tilde{X}} (\tilde{x} - \mathrm{E}(X))^2\, \mathrm{P_a}(X = \tilde{x}) \tag{5.1.4}$$

Der Unterschied kommt wiederum dadurch zustande, daß zur Gewichtung im ersten Fall eine Häufigkeitsfunktion, im zweiten Fall ein aleatorisches Wahrscheinlichkeitsmaß verwendet wird.

2. Auch für die inhaltliche Interpretation stellen sich die gleichen Probleme, wie sie im vorangegangenen Abschnitt für den Begriff des Erwartungswertes erwähnt worden sind. Am einfachsten sieht man das, wenn man die Varianz einer Zufallsvariablen auf folgende Weise ausdrückt:

$$\mathrm{V}(X) \;=\; \mathrm{E}((X - \mathrm{E}(X))^2) \tag{5.1.5}$$

Diese Schreibweise zeigt, daß sich die Varianz auf den Erwartungswert für die quadrierten Abweichungen der Werte einer Zufallsvariablen von ihrem Erwartungswert bezieht.[2]

[2]Um dies ganz explizit zu machen, kann man zunächst eine neue Zufallsvariable $Z := (X - \mathrm{E}(X))^2$ definieren. Dann ist die Varianz von X mit dem Erwartungswert einer Zufalllsvariablen, nämlich Z, identisch.

3. Schließlich sei noch angemerkt, daß sich aus der Definition (5.1.4) unmittelbar die Gleichung

$$\mathrm{V}(X) \;=\; \mathrm{E}(X^2) - \mathrm{E}(X)^2 \tag{5.1.6}$$

ableiten läßt, die zur Berechnung von Varianzen oft hilfreich ist.

4. In Abschnitt 4.4.3 wurde besprochen, wie man durch Funktionen einer oder mehrerer gegebener Zufallsvariablen neue Zufallsvariablen konstruieren kann. Im allgemeinen ist eine m-dimensionale Zufallsvariable $X = (X_1, \ldots, X_m)$ gegeben, die für die Elemente einer Kennzeichnungsmenge $\tilde{K}$ definiert ist, und außerdem eine Funktion

$$g : \tilde{X}_1 \times \cdots \times \tilde{X}_m \longrightarrow \tilde{Y}$$

die einen neuen Wertebereich $\tilde{Y}$ für eine neue Zufallsvariable Y erzeugt, deren Werte durch

$$Y(\tilde{k}) := g(X_1(\tilde{k}), \ldots, X_m(\tilde{k}))$$

definiert sind. Erwartungswert und Varianz von Y sind dann durch

$$\mathrm{E}(Y) \;=\; \textstyle\sum_{\tilde{y}\in\tilde{Y}} \tilde{y}\, \mathrm{P_a}(Y = \tilde{y})$$
$$\mathrm{V}(Y) \;=\; \textstyle\sum_{\tilde{y}\in\tilde{Y}} (\tilde{y} - \mathrm{M}(Y))^2\, \mathrm{P_a}(Y = \tilde{y})$$

berechenbar.

5. Ein wichtiger Spezialfall tritt auf, wenn die neue Zufallsvariable als eine Linearkombination der Zufallsvariablen $X_1, \ldots, X_m$ gebildet wird, wenn also Y durch

$$Y(\tilde{k}) = a_1 X_1(\tilde{k}) + \ldots + a_m X_m(\tilde{k})$$

definiert wird, wobei $a_1, \ldots, a_m$ beliebig vorgegebene Zahlen sind. In diesem Fall folgt nämlich aus der Definition von Erwartungswerten die Gleichung

$$\mathrm{E}(Y) = a_1 \mathrm{E}(X_1) + \ldots + a_m \mathrm{E}(X) \tag{5.1.7}$$

so daß sich der Erwartungswert von Y unmittelbar aus den Erwartungswerten der Zufallsvariablen $X_1, \ldots, X_m$ berechnen läßt.

6. Für die Varianz von Y gilt eine entsprechende Formel *nicht* ohne weiteres. Um dies genauer zu untersuchen, beziehen wir uns zunächst auf den Fall, daß $Y = X_1 + X_2$, also Summe von nur zwei Zufallsvariablen ist. Unter Verwendung von (5.1.6) ist dann

$$\mathrm{V}(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 \tag{5.1.8}$$

Für die beiden Summanden auf der rechten Seite findet man durch Ausrechnen und Anwendung von (5.1.7):

$$\mathrm{E}(Y^2) = \mathrm{E}(X_1^2) + \mathrm{E}(X_2^2) + 2\,\mathrm{E}(X_1 X_2)$$
$$\mathrm{E}(Y)^2 = \mathrm{E}(X_1)^2 + \mathrm{E}(X_2)^2 + 2\,\mathrm{E}(X_1)\,\mathrm{E}(X_2)$$

Durch Einsetzen in (5.1.8) folgt dann

$$\mathrm{V}(Y) = \mathrm{V}(X_1) + \mathrm{V}(X_2) + 2\,(\mathrm{E}(X_1 X_2) - \mathrm{E}(X_1)\,\mathrm{E}(X_2)) \tag{5.1.9}$$

7. Um die Überlegung fortzusetzen, wird ein neuer Begriff eingeführt: die *Kovarianz* von zwei Zufallsvariablen X_1 und X_2, die durch

$$\mathrm{Cov}(X_1, X_2) := \mathrm{E}((X_1 - \mathrm{E}(X_1))(X_2 - \mathrm{E}(X_2))) \tag{5.1.10}$$

definiert wird.[3] Durch Ausrechnen findet man

$$\mathrm{Cov}(X_1, X_2) = \mathrm{E}(X_1 X_2) - \mathrm{E}(X_1)\,\mathrm{E}(X_2) \tag{5.1.11}$$

Setzt man diesen Ausdruck für die Kovarianz in (5.1.9) ein, findet man

$$\mathrm{V}(Y) = \mathrm{V}(X_1) + \mathrm{V}(X_2) + 2\,\mathrm{Cov}(X_1, X_2) \tag{5.1.12}$$

[3]Ein normiertes Maß für die Kovarianz einer zweidimensionalen Zufallsvariablen liefert der *Korrelationskoeffizient*:

$$\mathrm{Corr}(X_1, X_2) := \frac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\mathrm{V}(X_1)\,\mathrm{V}(X_2)}}$$

Es gilt:

$$-1 \;\leq\; \mathrm{Corr}(X_1, X_2) \;\leq\; 1$$

Einen Beweis erhält man, indem man zunächst durch Ausrechnen zeigt:

$$\mathrm{V}\left(\frac{X_1}{\sqrt{\mathrm{V}(X_1)}} \pm \frac{X_2}{\sqrt{\mathrm{V}(X_2)}}\right) \;=\; 2\,(1 \pm \mathrm{Corr}(X_1, X_2))$$

Da diese Varianz stets nicht-negativ ist, folgt dann die zuvor angegebene Ungleichung für den Korrelationskoeffizienten.

In dieser Weise kann man also die Varianz von Y durch die Varianzen und die Kovarianz der Zufallsvariablen X_1 und X_2 berechnen.

8. Die Formel läßt sich weiter vereinfachen, wenn man annehmen kann, daß X_1 und X_2 stochastisch unabhängig sind (vgl. Abschnitt 4.5.2). Zunächst findet man nämlich unter dieser Voraussetzung

$$\begin{aligned} \mathrm{E}(X_1X_2) &= \textstyle\sum_{\tilde{x}_1\in\tilde{X}_1}\sum_{\tilde{x}_2\in\tilde{X}_2}\tilde{x}_1\tilde{x}_2\,\mathrm{P_a}(X_1=\tilde{x}_1,X_2=\tilde{x}_2) \\ &= \textstyle\sum_{\tilde{x}_1\in\tilde{X}_1}\sum_{\tilde{x}_2\in\tilde{X}_2}\tilde{x}_1\tilde{x}_2\,\mathrm{P_a}(X_1=\tilde{x}_1)\,\mathrm{P_a}(X_2=\tilde{x}_2) \\ &= \textstyle\sum_{\tilde{x}_1\in\tilde{X}_1}\tilde{x}_1\,\mathrm{P_a}(X_1=\tilde{x}_1)\sum_{\tilde{x}_2\in\tilde{X}_2}\tilde{x}_2\,\mathrm{P_a}(X_2=\tilde{x}_2) \end{aligned}$$

also

$$\mathrm{E}(X_1X_2) = \mathrm{E}(X_1)\,\mathrm{E}(X_2) \tag{5.1.13}$$

Daraus folgt: wenn X_1 und X_2 stochastisch unabhängig sind, ist ihre Kovarianz Null; und man erhält dann aus (5.1.12) die Formel

$$\mathrm{V}(Y) = \mathrm{V}(X_1) + \mathrm{V}(X_2) \tag{5.1.14}$$

Dieses Ergebnis kann auch für mehr als zwei Zufallsvariablen verallgemeinert werden. Vorausgesetzt werden muß nur ihre stochastische Unabhängigkeit im Sinne von Definition (4.5.4).

9. Wie oben gezeigt wurde, ist die Kovarianz (und infolgedessen auch die Korrelation) Null, wenn X_1 und X_2 stochastisch unabhängig sind. Umgekehrt folgt jedoch aus $\mathrm{Cov}(X_1,X_2)=0$ nicht unbedingt, daß die beiden Zufallsvariablen stochastisch unabhängig sind. Man muß also zwischen unkorrelierten und unabhängigen Zufallsvariablen unterscheiden.

5.2 Unabhängige Wiederholungen

Zur Vorbereitung der Überlegungen des nächsten Abschnitts werden in diesem Abschnitt repetitive Zufallsgeneratoren besprochen und wird für ihre Charakterisierung die Binomialverteilung eingeführt.

5.2.1 Repetitive Zufallsgeneratoren

1. Ausgangspunkt ist ein beliebiger Zufallsgenerator $\mathcal{G} := (\mathcal{V},(\mathcal{S},\tilde{X}))$. Aus der Definition von Zufallsgeneratoren folgt, daß es beliebig viele unabhängige Realisierungen gibt. Damit ist gemeint: wie mit dem Zufallsgenerator eine neue Situation (Ereignis, Sachverhalt) erzeugt wird, hängt nicht davon ab, welche Situationen schon zuvor mit ihm erzeugt worden sind. Also kann man einen neuen Zufallsgenerator

$$\mathcal{G}_n := (\mathcal{V}_n,(\mathcal{S}_n,\tilde{X}^n))$$

konstruieren, der darin besteht, den Zufallsgenerator $\mathcal{G}$ n-mal zu verwenden. Jede Betätigung von $\mathcal{G}_n$ liefert eine Situation, die sich durch einen Wert

$$\tilde{x} = (\tilde{x}_1,\ldots,\tilde{x}_n) \in \tilde{X}^n$$

charakterisieren läßt. Ein Wahrscheinlichkeitsmaß für $\mathcal{G}_n$ erhält man durch

$$\mathrm{P_a}(\{(\tilde{x}_1,\ldots,\tilde{x}_n)\}) = \prod_{j=1}^{n}\mathrm{P_a}(\{\tilde{x}_j\})$$

da für die Konstruktion angenommen wird, daß es sich um unabhängige Wiederholungen des Zufallsgenerators $\mathcal{G}$ handelt.

2. Für die weiteren Überlegungen wird angenommen, daß der Zufallsgenerator $\mathcal{G}$ durch eine Zufallsvariable X mit dem Wertebereich $\tilde{X}$ repräsentiert wird. Da der Zufallsgenerator $\mathcal{G}_n$ dadurch gebildet wird, daß $\mathcal{G}$ n-mal betätigt wird, kann man auch für jede einzelne Betätigung von $\mathcal{G}$ eine eigene Zufallsvariable definieren. (Man kann sich vorstellen, daß n Kopien von $\mathcal{G}$ verwendet werden.) So erhält man die Zufallsvariablen $X_1,\ldots,X_n$. Jede dieser Zufallsvariablen hat die gleiche Wahrscheinlichkeitsverteilung wie X, und aus ihrer Konstruktion folgt außerdem, daß sie stochastisch unabhängig sind. Also kann der Zufallsgenerator $\mathcal{G}_n$ durch eine n-dimensionale Zufallsvariable

$$(X_1,\ldots,X_n)$$

mit dem Wertebereich $\tilde{X}^n$ repräsentiert werden, deren Komponenten stochastisch unabhängig sind und identische (durch $\mathcal{G}$ bestimmte) Wahrscheinlichkeitsverteilungen haben.

3. Sei nun $\tilde{x}$ ein beliebiger Wert in $\tilde{X}$. Man kann dann eine neue Zufallsvariable $Z_n(\tilde{x})$ definieren, mit der erfaßt wird, wie oft der Wert $\tilde{x}$ bei einer Betätigung des Zufallsgenerators $\mathcal{G}_n$ auftritt. Zur Definition dieser Zufallsvariablen ist es zweckmäßig, Indikatorfunktionen zu verwenden:

$$I(X_j = \tilde{x}) := \begin{cases} 1 & \text{wenn } X_j = \tilde{x} \\ 0 & \text{andernfalls} \end{cases}$$

$Z_n(\tilde{x})$ kann dann folgendermaßen definiert werden:

$$Z_n(\tilde{x}) := \sum_{j=1}^{n} I(X_j = \tilde{x})$$

Ihr Wertebereich ist durch

$$\tilde{Z}_n = \{0, 1, \dots, n\}$$

gegeben. Die weitere Frage ist, wie man für $Z_n(\tilde{x})$ eine Wahrscheinlichkeitsverteilung finden kann.

5.2.2 Die Binomialverteilung

1. Zu bestimmen ist für jeden Wert $k \in \tilde{Z}_n$ die aleatorische Wahrscheinlichkeit

$$\mathrm{P_a}(Z_n(\tilde{x}) = k)$$

Offenbar muß man die möglichen Resultate des Zufallsgenerators $\mathcal{G}_n$ betrachten und sich überlegen, bei wievielen von ihnen der Wert $\tilde{x}$ genau k-mal vorkommt. Betrachten wir zunächst eine der möglichen Ergebnisfolgen: $(\tilde{x}_1, \dots, \tilde{x}_n)$. Die Wahrscheinlichkeit, daß der Wert $\tilde{x}$ in dieser Folge genau k-mal vorkommt, ist

$$p_{\tilde{x}}^k (1 - p_{\tilde{x}})^{n-k} \tag{5.2.1}$$

Dabei ist $p_{\tilde{x}} := \mathrm{P_a}(X = \tilde{x})$ die Wahrscheinlichkeit für das Auftreten des Wertes $\tilde{x}$ bei einer einmaligen Betätigung des Zufallsgenerators $\mathcal{G}$.

2. Jetzt muß noch ermittelt werden, wieviele Folgen es gibt, bei denen $\tilde{x}$ genau k-mal vorkommt. Für die dazu erforderliche kombinatorische

Überlegung sei zunächst an zwei Definitionen erinnert. Zunächst an die Definition von

$$n! := 1 \cdot 2 \cdots n$$

(gesprochen: n Fakultät), wobei n eine beliebige natürliche Zahl ist und außerdem vereinbart wird, daß $0! = 1$ sein soll. Weiterhin an die Definition des Binomialkoeffizieten (gesprochen: n über k):

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \qquad \text{für } 0 \le k \le n$$

Mithilfe dieser Definitionen kann man sich von der Richtigkeit der Gleichung

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \qquad \text{für } 0 \le k < n \tag{5.2.2}$$

überzeugen.

3. Diese Gleichung kann man schließlich verwenden, um – mithilfe des Prinzips der vollständigen Induktion – zu zeigen, daß die Anzahl von Folgen der Länge n, bei denen $\tilde{x}$ genau k-mal vorkommt, gerade n über k ist. Also folgt unter Verwendung von (5.2.1) für die gesuchte Wahrscheinlichkeitsverteilung:

$$\mathrm{P_a}(Z_n(x) = k) = \binom{n}{k} p_{\tilde{x}}^k (1 - p_{\tilde{x}})^{n-k} \tag{5.2.3}$$

Sie wird *Binomialverteilung* (mit dem Parameter $p_{\tilde{x}}$) genannt.

4. Um den Erwartungswert der Zufallsvariablen $Z_n(\tilde{x})$ zu bestimmen, erinnern wir uns an ihre Definition als eine Summe von Indikatorvariablen $I(X_j = \tilde{x})$. Jede dieser Indikatorvariablen ist selbst eine Zufallsvariable, die genau zwei Werte annehmen kann: 1 mit der Wahrscheinlichkeit $p_{\tilde{x}}$ und 0 mit der Wahrscheinlichkeit $1 - p_{\tilde{x}}$. Der Erwartungswert ist also

$$\mathrm{E}(I(X_j = \tilde{x})) = p_{\tilde{x}}$$

und somit folgt aus (5.1.7) für den Erwartungswert von $Z_n(\tilde{x})$:

$$\mathrm{E}(Z_n(\tilde{x})) = n\, p_{\tilde{x}} \tag{5.2.4}$$

Ganz analog findet man, daß $\mathrm{V}(I(X_j = \tilde{x})) = p_{\tilde{x}}(1 - p_{\tilde{x}})$ ist, und da die Indikatorvariablen stochastisch unabhängig sind, erhält man

$$\mathrm{V}(Z_n(\tilde{x})) = n\, p_{\tilde{x}}\, (1 - p_{\tilde{x}}) \tag{5.2.5}$$

für die Varianz der Zufallsvariablen $Z_n(\tilde{x})$.

5.3 Aussagen über Wiederholungen

Wie in Abschnitt 5.1.1 angedeutet worden ist, suggeriert der Begriff eines Erwartungswertes, daß es einen Zusammenhang gibt zwischen aleatorischen Wahrscheinlichkeiten und Erwartungen, die sich auf relative Häufigkeiten in Mengen von Realisationen einer Zufallsvariablen beziehen. Es wurde allerdings auch schon darauf hingewiesen, daß ein unmittelbarer begrifflicher Zusammenhang nicht hergestellt werden kann, da sich aleatorische Wahrscheinlichkeiten auf *mögliche*, dagegen relative Häufigkeiten auf tatsächlich realisierte Ereignisse beziehen. In diesem Abschnitt soll besprochen werden, wie gleichwohl ein indirekter Zusammenhang hergestellt werden kann.

5.3.1 Reformulierung der Fragestellung

1. Wie kann die Vermutung eines Zusammenhangs zwischen aleatorischen Wahrscheinlichkeiten und relativen Häufigkeiten formuliert werden? Unter Verwendung der in Abschnitt 5.2.2 eingeführten Notationen, könnte man versuchen, sie so zu formulieren:

$$Z_n(\tilde{x})/n \approx p_{\tilde{x}}$$

In dieser Formulierung ist ein Vergleich allerdings kaum möglich. Auf der rechten Seite steht eine bestimmte Wahrscheinlichkeit, nämlich

$$p_{\tilde{x}} = \mathrm{P_a}(X = \tilde{x})$$

die für einen Zufallsgenerator $\mathcal{G}$ definiert ist. Auf der linken Seite steht dagegen keine bestimmte Zahl, sondern eine Zufallsvariable, die für einen Zufallsgenerator $\mathcal{G}_n$ definiert ist, der durch eine n-malige Wiederholung von $\mathcal{G}$ zustande kommt. — Um dennoch einen Vergleich vornehmen zu können, muß die Fragestellung verändert werden. Dies geschieht in zwei Schritten. In einem ersten Schritt wird die Fragestellung in ein Problem der aleatorischen Wahrscheinlichkeitsrechnung transformiert. In einem zweiten Schritt kann dann ein Übergang zu epistemischen Wahrscheinlichkeitsaussagen vollzogen werden.

2. Eine Fragestellung der aleatorischen Wahrscheinlichkeitsrechnung erhält man durch die Konzeption einer Zufallsvariable

$$|\, Z_n(\tilde{x})/n - p_{\tilde{x}} \,|$$

die sich auf Abweichungen zwischen einer aleatorischen Wahrscheinlichkeit und ihr korrespondierenden relativen Häufigkeiten bezieht. Man muß allerdings darauf achten, daß sich diese Zufallsvariable nicht unmittelbar auf den Zufallsgenerator $\mathcal{G}$, sondern auf den repetitiven Zufallsgenerator $\mathcal{G}_n$ bezieht. Weiterhin ist auch klar, daß über diese Zufallsvariable nur Wahrscheinlichkeitsaussagen gemacht werden können.

3. Der weitere Gedankengang bezieht sich auf aleatorische Wahrscheinlichkeiten der Form

$$\mathrm{P_a}(\,|\, Z_n(\tilde{x})/n - p_{\tilde{x}} \,| < a)$$

wobei a eine beliebig kleine positive Zahl ist. Man kann nämlich zeigen, daß diese Wahrscheinlichkeit für hinreichend großes n beliebig an den Wert 1 angenähert werden kann. Diese Aussage wurde zum erstenmal von J. Bernoulli in seiner *Ars conjectandi* entwickelt; sie wird oft *Bernoullis Gesetz der großen Zahlen* genannt, obwohl es sich natürlich nicht um ein Gesetz, sondern um einen mathematischen Satz handelt.

5.3.2 Bernoullis Gesetz der großen Zahlen

1. Um Bernoullis „Gesetz“ zu verstehen, ist es am einfachsten, sich zunächst eine andere Aussage zu überlegen, nämlich die Ungleichung

$$\text{Für alle } a > 0: \quad \mathrm{P_a}(\,|\, X - \mathrm{E}(X) \,| \geq a) \;\leq\; \frac{\mathrm{V}(X)}{a^2} \tag{5.3.1}$$

wobei X eine beliebige Zufallsvariable ist. Sie wird *Tschebyscheffsche Ungleichung* genannt.[4]

2. Ihre Richtigkeit ist leicht einzusehen. Sei a eine beliebige positive Zahl. Dann ist

$$\begin{aligned} \mathrm{V}(X) &= \textstyle\sum_{\tilde{x} \in \tilde{X}} (\tilde{x} - \mathrm{E}(X))^2\, \mathrm{P_a}(X = \tilde{x}) \\ &\geq \textstyle\sum_{\tilde{x} \in \{\tilde{x}' \,|\, a \leq |\, \tilde{x}' - E(X) \,|\}} (\tilde{x} - \mathrm{E}(X))^2\, \mathrm{P_a}(X = \tilde{x}) \\ &\geq \textstyle\sum_{\tilde{x} \in \{\tilde{x}' \,|\, a \leq |\, \tilde{x}' - E(X) \,|\}} a^2\, \mathrm{P_a}(X = \tilde{x}) \\ &= a^2\, \mathrm{P_a}(\,|\, X - \mathrm{E}(X) \,| \geq a) \end{aligned}$$

woraus (5.3.1) unmittelbar folgt.

[4]Nach dem russischen Mathematiker P. L. Tschebyscheff (1821 – 1894).

3. Unter Zuhilfenahme der Tschebyscheffschen Ungleichung kann die Aussage von Bernoulli auf einfache Weise bewiesen werden. Sei nämlich a eine beliebige positive Zahl. Man kann dann sicherlich eine natürliche Zahl n_0 finden, so daß gilt:

$$\frac{p_x(1-p_x)}{a^3} \leq n_0$$

Für jede beliebige natürliche Zahl $n \geq n_0$ folgt dann aus der Tschebyscheffschen Ungleichung:

$$\mathrm{P_a}(\,|\,Z_n(\tilde{x})/n - p_{\tilde{x}}\,| \geq a) \;=$$

$$\mathrm{P_a}(\,|\,Z_n(\tilde{x}) - np_{\tilde{x}}\,| \geq na) \;\leq\; \frac{n\,p_{\tilde{x}}(1-p_{\tilde{x}})}{(na)^2} \;=\; \frac{p_{\tilde{x}}(1-p_{\tilde{x}})}{na^2}$$

also auch

$$\mathrm{P_a}(\,|\,Z_n(\tilde{x})/n - p_{\tilde{x}}\,| < a) \;=$$

$$1 - \mathrm{P_a}(\,|\,Z_n(\tilde{x})/n - p_{\tilde{x}}\,| \geq a) \;\geq\; 1 - \frac{p_{\tilde{x}}(1-p_{\tilde{x}})}{na^2}$$

Nach Voraussetzung ist jedoch

$$\frac{p_x(1-p_x)}{a^3} \;\leq n_0 \;\leq n$$

also auch

$$1 - a \;\leq\; 1 - \frac{p_x(1-p_x)}{na^2}$$

Es folgt somit, daß man für jedes vorgegebene $a > 0$ eine natürliche Zahl n_0 finden kann, so daß für alle größeren natürlichen Zahlen n gilt:

$$\mathrm{P_a}(\,|\,Z_n(\tilde{x})/n - p_{\tilde{x}}\,| < a) \;\geq\; 1 - a$$

5.4 Wahrscheinlichkeiten und Erwartungen

1. Bernoullis „Gesetz“ hat in der Geschichte der Wahrscheinlichkeitsrechnung eine wichtige, allerdings auch kontrovers diskutierte Rolle gespielt.[5] Von Bernoulli selbst wurde es unmittelbar mit einer metaphysischen Spekulation verknüpft, die zum Beispiel in folgender Bemerkung zum Ausdruck kommt (Bernoulli 1899b, S. 107):

> „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 gezwungen werden, auch bei noch so zufällig erscheinenden Dingen eine gewisse Nothwendigkeit, und sozusagen ein Fatum anzunehmen.“

Dagegen sollte zunächst betont werden, daß es sich bei seinem „Gesetz“ um eine Aussage der aleatorischen Wahrscheinlichkeitsrechnung handelt. Sofern sie einen Realitätsbezug hat, bezieht sie sich auf Zufallsgeneratoren bzw. auf Ereignisse, die durch Prozesse entstehen, die als Zufallsgeneratoren konzipiert werden können.

2. Aber auch wenn man sich zum Verständnis von Bernoullis „Gesetz“ auf Zufallsgeneratoren bezieht, liefert es nicht unmittelbar eine Brücke zwischen Wahrscheinlichkeiten und Häufigkeiten. Es handelt sich vielmehr um eine Wahrscheinlichkeitsaussage über eine Zufallsvariable $Z_n(\tilde{x})/n$, die Wiederholungen eines Zufallsgenerators repräsentiert: wenn die Anzahl der Wiederholungen n sehr groß ist, ist die aleatorische Wahrscheinlichkeit dafür, daß sich Werte von $Z_n(\tilde{x})/n$ nur wenig von $p_{\tilde{x}}$ unterscheiden, nahezu gleich 1. Der Zusammenhang zwischen Wahrscheinlichkeiten und relativen Häufigkeiten wird also nur indirekt – durch eine Wahrscheinlichkeitsaussage – hergestellt. Und zwar durch eine *aleatorische* Wahrscheinlichkeitsaussage, also durch eine Aussage, die selbst modaler Natur ist und sich auf Eigenschaften eines Zufallsgenerators bezieht; und zwar in diesem Fall auf Eigenschaften eines *repetitiven*, durch Wiederholungen definierten Zufallsgenerators.

3. Von Aussagen dieser Art führt jedoch kein unmittelbarer Weg zu Erwartungen. Vielmehr ist noch eine explizite Argumentation erforderlich, um zur Begründung von epistemischen Wahrscheinlichkeitsaussagen

[5] Es ist auch zum Ausgangspunkt für eine große Anzahl ähnlicher Aussagen geworden, die sich auf Wahrscheinlichkeitsverteilungen von Zufallsvariablen beziehen, die durch repetitive Zufallsgeneratoren definiert werden können. Da es in diesem Text nur um konzeptionelle Fragen geht, braucht darauf nicht näher eingegangen zu werden.

zu gelangen, die sich nicht auf Zufallsgeneratoren, sondern auf Hypothesen beziehen, die durch Zufallsgeneratoren realisierte oder realisierbare Situationen betreffen. Um die erforderliche Argumentation zu verdeutlichen, betrachten wir eine Hypothese, die sich auf die relative Häufigkeit von $\tilde{x}$ in einer Menge S_n bezieht, die aus n Situationen besteht, die mit dem Zufallsgenerator $\mathcal{G}$ erzeugt worden sind. Eine solche Hypothese kann folgendermaßen formuliert werden:

$$\langle\, \mathrm{S}_n \,|\, Z_n(\tilde{x})/n - p_{\tilde{x}} \,| < a \,\rangle$$

Diese Hypothese bezieht sich auf *eine bestimmte* Menge S_n von Situationen, die mit dem Zufallsgenerator $\mathcal{G}$ erzeugt worden sind, und deutet sie als *ein* Ergebnis eines repetitiven Zufallsgenerators $\mathcal{G}_n$; und sie besagt, daß in dieser Menge von Situationen ein Wert von $Z_n(\tilde{x})/n$ realisiert wird, der von $p_{\tilde{x}}$ um wenger als a abweicht. Um den Satz von Bernoulli für eine epistemische Wahrscheinlichkeitsaussage über eine solche Hypothese verwenden zu können, muß also zunächst ein Zusammenhang zwischen epistemischen und aleatorischen Wahrscheinlichkeiten hergestellt werden. Es liegt natürlich nahe, dabei so vorzugehen, wie in Abschnitt 4.3.2 besprochen worden ist:

$$\mathrm{P_e}(\langle\, \mathrm{S}_n,\, |\, Z_n(\tilde{x})/n - p_{\tilde{x}} \,| < a \,\rangle) \;:=\; \mathrm{P_a}(\,|\, Z_n(\tilde{x})/n - p_{\tilde{x}} \,| < a)$$

Dann kann man den Wert der aleatorischen Wahrscheinlichkeit als ein Argument verwenden, um eine epistemische Wahrscheinlichkeit zu begründen.

4. Aber auch wenn man auf diese Weise aleatorische und epistemische Wahrscheinlichkeiten verknüpft, ist es noch eine offene Frage, ob und ggf. wie man zu Begründungen für Erwartungen gelangen kann. Das hängt natürlich auch davon ab, wie man das Wort 'Erwartung' verwenden möchte. Der eben angedeutete Gedankengang liefert jedenfalls zunächst nur eine Begründung für epistemische Wahrscheinlichkeitsaussagen, durch die Hypothesen einschätzbar gemacht werden. Daß solche Aussagen keine Entscheidungen implizieren, wurde bereits in Abschnitt 3.1 besprochen. Andererseits scheint klar zu sein, daß man die Wahrscheinlichkeit von Hypothesen oft deshalb einschätzbar machen möchte, um Begründungen für Erwartungen reflektierbar zu machen. Trotzdem erscheint es mir sinnvoll, Erwartungen von Wahrscheinlichkeiten für Hypothesen begrifflich zu unterscheiden. Angenommen z.B., die aleatorische Wahrscheinlichkeit für ein Ereignis ist 0.8; soll man dann erwarten, daß dies Ereignis eintritt? — Einige Wahrscheinlichkeitstheoretiker haben

gleichwohl versucht, auch für Erwartungsbildungen eine Art von Prinzip zu formulieren. Eine Variante haben wir bereits bei der Besprechung von Kolmogoroffs Axiomensystem kennengelernt; man erinnere sich an die dort zitierten Formulierungen der Form „man kann praktisch sicher sein, daß …“ (Abschnitt 3.7.1). Die Idee besteht darin, daß man eine (aleatorische) Wahrscheinlichkeit, die nahe bei ihrem Maximalwert 1 liegt, als Begründung für eine „praktisch sichere Erwartung“ verwenden könne.[6]

[6]Dieser Gedanke wurde insbesondere von dem Wahrscheinlichkeitstheoretiker Emile Borel vertreten, u.a. in seinem Buch *Probabilité et Certitude* (1950), in dessen Vorwort der Grundgedanke so formuliert wird: „In all my writings hitherto – and particularly in my book on chance – I adopted, with reference to those physical phenomena whose probability is extremely small, the language commonly used by physicists. The latter confine themselves to stating that it is 'higly improbable' that such phenomena will occur, but they hasten to add that this is not a matter of certainty. On reflection it seemed to me that such an attitude was not sufficiently realistic and did not take into account the whole range of our knowledge about the universe; and I arrived at the conclusion that we should not be afraid to use the word *certainty* to describe a probability that falls short of unity by a sufficiently small quantity.“

Kapitel 6

Frequentistische Definitionsansätze

Wie in den vorausgegangenen Kapiteln ausgeführt worden ist, gingen die „klassischen“ Begründungen der Wahrscheinlichkeitsrechnung nicht von relativen Häufigkeiten in jeweils bestimmten endlichen Gesamtheiten von Situationen aus, sondern von Zufallsgeneratoren, mit denen Situationen erzeugt werden können und deren Funktionsweise sich auf ein Verständnis „gleicher Möglichkeiten“ zurückführen läßt. Versuche, stattdessen von relativen Häufigkeiten *auszugehen*, sind erst im 19. Jahrhundert explizit unternommen worden; hinzuweisen ist vor allem auf Arbeiten von R. L. Ellis (1849) und die „Logic of Chance“ von John Venn (1888, erste Ausgabe 1866). Um auf solche Bestrebungen zu verweisen, spreche ich von *frequentistischen* Wahrscheinlichkeitskonzeptionen. In diesem Kapitel werden einige Überlegungen zu ihrer Diskussion dargestellt.

6.1 Einleitende Bemerkungen

1. Hauptsächlich drei Motive für die Entwicklung und Verbreitung frequentistischer Auffassungen lassen sich feststellen. (a) Zunächst spielt sicherlich die Erfahrungstatsache eine Rolle, daß man sich in vielen Fällen von einem Zusammenhang zwischen aleatorischen Wahrscheinlichkeiten und relativen Häufigkeiten praktisch überzeugen kann; so kann sich der Gedanke aufdrängen, daß man diese Erfahrungstatsache auch für eine Definition des Wahrscheinlichkeitsbegriffs verwenden könne. (b) Ein zweites Motiv resultiert aus der Kritik des „Prinzips des unzureichenden Grundes“, durch das die „klassische“ Konzeption eines aleatorischen Wahrscheinlichkeitsbegriffs zunehmend ersetzt und mit spekulativen Gedankengängen vermengt worden ist (vgl. Abschnitt 4.3.3). Wird stattdessen von beobachtbaren Häufigkeiten ausgegangen, ist es vielleicht möglich, einen empirischen Ansatz zum Verständnis von Wahrscheinlichkeiten zu finden. (c) Damit hängt ein drittes Motiv zusammen, das sich auf Anwendungen der Wahrscheinlichkeitsrechnung bezieht. Der „klassische“ Ansatz, eine Grundlegung durch Zufallsgeneratoren herzustellen, schränkt mögliche Anwendungen auf Situationen, Ereignisse und Sachverhalte ein, die durch Zufallsgeneratoren zustande kommen. Das „Prinzip des unzureichenden Grundes“ kann insofern auch unter dem

Gesichtspunkt betrachtet werden, daß es den durch Zufallsgeneratoren eng umgrenzten Bereich möglicher sinnvoller Anwendungen der Wahrscheinlichkeitsrechnung beliebig erweiterbar machen sollte; nämlich auf alle Gebiete, wo man sich „gleichermaßen unschlüssig" (Laplace) ist. Frequentistische Ansätze scheinen es demgegenüber zu erlauben, die Anwendungsmöglichkeiten der „klassischen" Wahrscheinlichkeitsrechnung auf nicht-spekulative Weise auf alle Gebiete auszuweiten, in denen sich sinnvoll von „Massenerscheinungen" und relativen Häufigkeiten sprechen läßt.

2. Alle drei Motive finden sich inbesondere bei R. v. Mises, der als wichtigster Vertreter einer frequentistischen Auffassung angesehen werden kann.[1] Seine Kritik an der „klassischen" Konzeption des aleatorischen Wahrscheinlichkeitsbegriffs, die er allerdings im wesentlichen mit der insbesondere von Laplace vertretenen Deutung durch das „Prinzip des unzureichenden Grundes" gleichsetzt, wurde bereits in Abschnitt 4.3.3 erwähnt. — In seiner „Wahrscheinlichkeitsrechnung" (1931, S. 3) knüpft v. Mises stattdessen an Poisson an:

> „Nach Laplace gab es noch einmal eine hervorragende originale Leistung in dem Werke von Siméon Denis Poisson (1781 – 1840), der u.a. eine wichtige Erweiterung des Bernoullischen Satzes fand. In der Einleitung seines Buches schildert Poisson in ausgezeichneter Weise die allgemeine Erfahrungsgrundlage, die den Ausgangspunkt aller wahrscheinlichkeitstheoretischen Untersuchungen bildet: 'Erscheinungen verschiedenster Art sind einem allgemeinen Gesetz unterworfen, das man das Gesetz der großen Zahlen nennen kann. Es besteht darin, daß, wenn man sehr große Anzahlen von gleichartigen Ereignissen beobachtet, die von konstanten Ursachen und von solchen abhängen, die regelmäßig, nach der einen und anderen Richtung veränderlich sind, ohne daß ihre Veränderung in einem bestimmten Sinn fortschreitet, man zwischen diesen Zahlen Verhältnisse finden wird, die nahezu unveränderlich sind. Für jede Art von Erscheinungen haben diese Verhältnisse besondere Werte, denen sie sich um so mehr nähern, je größer die Reihe der beobachteten Erscheinungen ist, und die sie in aller Strenge erreichen würden, wenn es möglich wäre, die Reihe der Beobachtungen ins Unendliche auszudehnen.'"

[1] R. v. Mises lebte 1883 – 1953. Informationen über seine Biographie und wissenschaftlichen Arbeiten finden sich bei Stadler (1990). Ich beziehe mich im weiteren hauptsächlich auf sein Lehrbuch zur Wahrscheinlichkeitsrechnung (1931) und die zweite Auflage seines Buches „Probability, Statistics and Truth" (1981). Diese zweite Auflage erschien zuerst 1957; die erste Auflage erschien in deutscher Sprache 1928. Weiterhin beziehe ich mich auch auf sein „Kleines Lehrbuch des Positivismus", in dem sich ein Kapitel mit dem Wahrscheinlichkeitsbegriff beschäftigt. Das Buch erschien zuerst 1939 in deutscher Sprache, dann, zuerst 1951, in mehreren amerikanischen Ausgaben und 1990 erneut in einer deutschen Ausgabe.

Dieses „empirische Gesetz der großen Zahlen" spielt in allen frequentistischen Ansätzen eine wichtige Rolle und wird auch in Teilen der heutigen Lehrbuchliteratur als eine Grundlage der Wahrscheinlichkeitsrechnung angesehen.

3. Gibt es ein solches „Gesetz"? Zwei Fragestellungen sollten unterschieden werden. Einerseits: Gibt es wenigstens einige Beispiele, bei denen sich die von Poisson intendierte Aussage als eine Erfahrungstatsache behaupten läßt? Daran könnten sich Überlegungen anschließen, wie der Sachverhalt in diesen Beispielen verstanden werden kann. Eine ganz andere Frage bezieht sich darauf, wie man die Arten von Ereignissen eingrenzen kann, bei denen die von Poisson behaupteten Erfahrungsgrundlagen der Wahrscheinlichkeitsrechnung sinnvoll unterstellt werden können. Im weiteren geht es zunächst nur um die erste Frage.

4. Als Beispiele kann man natürlich an Zufallsgeneratoren denken. Auf solche Beispiele beruft sich auch v. Mises:

> „Jeder Knabe weiß, daß beim 'Kopf-oder-Adler'-Spiel mit einer gewöhnlichen Münze, wenn es nur hinreichend lange fortgesetzt wird, ungefähr die Hälfte aller Würfe 'Kopf' ergibt (wofür wir das Zeichen '1' setzen) und ungefähr die Hälfte 'Adler' (wofür wir das Zeichen '0' wählen). Werden also n Würfe ausgeführt, so ist die Zahl n_1 der 1-Würfe annähernd $n/2$ oder die 'relative Häufigkeit' der 1-Würfe $n_1/n \sim 1/2$."

Ist dies eine Erfahrungstatsache? Die Frage muß wohl bejaht werden; jedenfalls in dem Sinne, in dem wir überhaupt von Erfahrungstatsachen sprechen: daß man auf viele Beispiele hinweisen kann, anhand derer man sich überzeugen kann.

5. Dann aber stellt sich sogleich eine weitere Frage: Wie kann man über Erfahrungstatsachen dieser Art sinnvoll sprechen? Es sei betont, daß es bei dieser Frage nicht darum geht, Erklärungen zu finden, sondern um eine vorgängige, aber auch grundlegendere Frage: wie sich geeignete Begriffe finden lassen, um über Erfahrungstatsachen der angedeuteten Art sinnvoll sprechen zu können. Natürlich reicht dafür in gewisser Weise schon unsere Umgangssprache; aber ihre Ausdrucksmittel sind hier auf bemerkenswerte Weise ungenau. Etwa so wie in dem oben angeführten Zitat: die beobachtete relative Häufigkeit der 1-Würfe wird „ungefähr" oder „annähernd" 1/2 sein.

6. Es sollte betont werden, daß hier kein „Meßproblem" vorliegt. Das Problem liegt offenbar nicht darin, daß man relative Häufigkeiten nicht genau ermitteln kann. Wenn man eine Münze 100 Mal wirft, kann man hinterher nicht nur ungefähr, sondern ganz genau sagen, wie oft eine der

Seiten aufgetreten ist; Meßfehler könnten nur daraus resultieren, daß man sich verzählt oder nicht genau erkennen kann, was herausgekommen ist. Das Wort 'ungefähr' bezieht sich vielmehr darauf, daß man die relative Häufigkeit der 1-Würfe nicht genau, sondern nur ungefähr *voraussagen* kann.

7. Diese Überlegung macht nicht nur deutlich, was hier mit dem Wort 'Erfahrungstatsache' gemeint ist: daß man in einigen Fällen – unter gewissen Voraussetzungen – relative Häufigkeiten näherungsweise richtig voraussagen kann. Sie zeigt auch, worin das Begriffsbildungsproblem besteht. Es besteht nicht darin, Begriffe zum Reden über relative Häufigkeiten zu finden; dafür genügen vollständig die Begriffsbildungen einer Statistik, die ganz ohne Wahrscheinlichkeitstheorie auskommt. Es besteht vielmehr – in diesem Beispiel – darin, sprachliche Ausdrucksformen zu entwickeln, mit deren Hilfe Voraussagen über relative Häufigkeiten formuliert und Begründungen reflektiert werden können. Aber auch diese Überlegung macht noch nicht vollständig klar, warum man einen Wahrscheinlichkeitsbegriff benötigt. Denn um Voraussagen *formulieren* zu können, muß man nur in der Lage sein, in geeigneter Weise darüber zu sprechen, was man voraussagen möchte. In unserem Beispiel genügt ein Verständnis des Begriffs 'relative Häufigkeit', um eine entsprechende Voraussage zu machen. Ein neues Problem, das auch neue Begriffsbildungen erfordert, beginnt erst mit der Überlegung, daß man solche Voraussagen *nicht mit Sicherheit* machen kann, sondern nur mit mehr oder weniger großer Wahrscheinlichkeit.

8. Soweit führt der Gedankengang jedoch nur dorthin zurück, wo in Kapitel 3 begonnen wurde: daß man einen Wahrscheinlichkeitsbegriff benötigt, um in der Form *epistemischer* Wahrscheinlichkeitsaussagen zum Ausdruck zu bringen, daß Hypothesen (Voraussagen) nicht sicher, sondern nur mehr oder weniger wahrscheinlich sind. Die Erfahrungstatsache, auf die sich Poisson und v. Mises berufen, hat jedoch zwei Aspekte. Wenn man sich auf das Münzwurfbeispiel bezieht, kann man folgende Formulierung verwenden, um die beiden Aspekte sichtbar zu machen: Wenn man die Münze 100 Mal wirft, kann man *ziemlich sicher* sein, daß die relative Häufigkeit von 1-Würfen *ungefähr* 1/2 beträgt.[2] Die Formulierung 'ziemlich sicher' deutet an, daß es sich um eine epistemische Wahrscheinlichkeitsaussage handelt, die sich auf folgende Hypothese (Voraussage) bezieht: Wenn man die Münze 100 Mal wirft, ist die relative

[2] Man vgl. die ganz ähnliche Formulierung von Kolmogoroff, die in Abschnitt 3.7.1 angeführt worden ist.

Häufigkeit von 1-Würfen *ungefähr* 1/2. Ein zweiter Aspekt kommt in der Verwendung des Wortes 'ungefähr' für die Formulierung der Voraussage zum Ausdruck. Aber warum wird hier dieses Wort benötigt? Es liegt offenbar kein sprachliches Problem vor; denn man könnte die Voraussage auch so formulieren: die relative Häufigkeit von 1-Würfen wird *exakt* 1/2 betragen; und man könnte genau so gut jede andere Zahl zwischen 0 und 1 verwenden, um eine *exakte* Voraussage zu formulieren.[3] Das Problem besteht ersichtlich nicht darin, daß man keine exakten Voraussagen *formulieren* kann, sondern daß sich exakt formulierte Voraussagen fast immer als falsch herausstellen würden.

9. An diesem zweiten Aspekt setzt nun ein wichtiger Gedankengang der frequentistischen Wahrscheinlichkeitskonzeption an: daß man das „ungefähr" der relativen Häufigkeiten in empirisch realisierten Gesamtheiten als Abweichungen von hypothetisch konzipierbaren Größen auffassen könne, die sich den relativen Häufigkeiten als ihre Wahrscheinlichkeiten unterstellen lassen.

10. Was gemeint ist, wird vielleicht durch folgende Ausführungen eines neueren Lehrbuchs von Gnedenko (1991, S. 22) etwas deutlicher:

„Die Wahrscheinlichkeitstheorie [befaßt sich] nur mit Ereignissen, die eine sog. *statistische Stabilität* oder, anders gesagt, eine Stabilität der relativen Häufigkeiten aufweisen. Diese Bedingung müssen wir genau erläutern. Es möge eine Folge von Versuchen angestellt werden, bei denen ein Ereignis A auftreten oder auch nicht auftreten kann. Die Versuche sollen unter gleichbleibenden Umständen durchgeführt werden, und die Ergebnisse bestimmter Verrsuche sollen keinen Einfluß auf die Ergebnisse der anderen Versuche haben (man sagt, die Versuche sollen unabhängig sein). Es möge die Zahl μ angeben, wie oft das Ereignis A in n vorher bestimmten (z.B. in n aufeinanderfolgenden) Versuchen auftrat; dann wird das Verhältnis μ/n die *relative Häufigkeit* von A in den betreffenden Versuchen genannt; statistische Stabilität von A liegt vor, wenn μ/n für $n \gg 1$ in der Nähe einer Konstanten liegt und sich nur wenig ändert, wenn man weitere Versuchsserien der Länge n durchführt und die entsprechenden μ bestimmt."

Es sind diese hypothetischen „Konstanten", die der frequentistische Ansatz als Wahrscheinlichkeitsgrößen konzipieren möchte. Sie müssen natürlich von epistemischen Wahrscheinlichkeiten, die sich auf Hypothesen beziehen, unterschieden werden. Sie unterscheiden sich aber auch von

[3] Anstelle von 'exakt' könnte man auch das Wort 'genau' verwenden. In diesem Text ist jedoch mit 'genau' meistens nur sprachliche, nicht unbedingt numerische Genauigkeit gemeint; dagegen soll das Wort 'exakt' darauf verweisen, daß numerische Genauigkeit angestrebt wird.

aleatorischen Wahrscheinlichkeiten, die beim „klassischen" Ansatz als Eigenschaften von Zufallsgeneratoren expliziert werden. Um auf die hypothetischen Größen des frequentistischen Ansatzes bezug zu nehmen, spreche ich im folgenden von *frequentistischen* oder kurz F-Wahrscheinlichkeiten.[4]

11. Um sich den Gedankengang vorstellbar zu machen, kann man an gewöhnliche Größen denken, deren Werte sich durch Meßverfahren ermitteln lassen. Man kann zum Beispiel an Objekte denken, die ein Gewicht haben. Ihr Gewicht kann durch ein Meßverfahren ermittelt werden, aber nur „ungefähr", es treten Meßfehler auf, die man als Abweichungen von einer theoretisch konzipierten Größe auffassen kann. Wenn ich es richtig verstehe, liegt der frequentistischen Wahrscheinlichkeitskonzeption ein ganz analog konzipierter Gedankengang zugrunde: daß man sich die Abweichungen zwischen empirisch ermittelbaren relativen Häufigkeiten und theoretisch unterstellbaren F-Wahrscheinlichkeiten so vorstellen könne wie die Abweichungen zwischen Meßergebnissen und einer durch sie zu ermittelnden theoretischen Größe.[5] Aber auch wenn sich durch diese Analogie verstehen läßt, was vielleicht gemeint sein könnte, stellt sich natürlich die Frage, ob sich der Gedankengang für die Wahrscheinlichkeitstheorie eignet. Von Autoren, die sich bemüht haben, einen frequentistischen Wahrscheinlichkeitsbegriff zu begründen, ist zur Beantwortung dieser Frage die Idee entwickelt worden, daß man F-Wahrscheinlichkeiten als „Grenzwerte" relativer Häufigkeiten auffassen kann. Damit beschäftigen wir uns im nächsten Abschnitt.

[4] Einige Anhänger des frequentistischen Ansatzes glauben, daß auch epistemische Wahrscheinlichkeiten einer frequentistischen Deutung zugänglich sind. Insbesondere hat H. Reichenbach zu zeigen versucht, „daß die Häufigkeitsinterpretation der Wahrscheinlichkeit für alle Verwendungen des Wortes ‘wahrscheinlich’ durchgeführt werden kann." (Reichenbach 1949, S. xiii) Wir werden jedoch diese Frage zunächst ausklammern und nur diskutieren, ob F-Wahrscheinlichkeiten als eine sinnvolle Alternative zu aleatorischen Wahrscheinlichkeiten verstanden werden können.

[5] So heißt es zum Beispiel bei H. Freudenthal (1963, S. 132f): „Insbesondere können wir versuchen, Wahrscheinlichkeiten zu messen, wie der Physiker Längen, Kräfte, Spannungen mißt. Meßinstrument ist das Ziehen einer Stichprobe." „Die Messung einer Wahrscheinlichkeit ist eine physikalische Messung und wie jede physikalische Messung mit einer Unsicherheit behaftet." Ganz ähnlich sagt F. Sixtl (1993, S. 241): „Die Wahrscheinlichkeit eines Ereignisses kann man, falls sie existiert, auch messen: jede auf einem endlichen Stichprobenumfang basierende relative Häufigkeit p_n ist ein Meßwert der festen, aber unbekannten Wahrscheinlichkeit."

6.2 Grenzwerte von Häufigkeiten?

1. Die Vorstellung, daß F-Wahrscheinlichkeiten als Grenzwerte relativer Häufigkeiten aufgefaßt werden können, kommt bereits in den Ausführungen von Poisson gut zum Ausdruck (man vgl. das zu Beginn des vorangegangenen Abschnitts angeführte Zitat). Ganz ähnlich heißt es bei Venn in seiner „Logic of Chance" (1888, S. 4) folgendermaßen:

> "What, for instance, is the meaning of the statement that two new-born children in three fail to attain the age of sixty-three? It certainly does not declare that in any given batch of, say, thirty, we shall find just twenty that fail: whatever might be the strict meaning of the words, this is not the import of the statement. It rather contemplates our examination of a large number, of a long succession of instances, and states that in such a succession we shall find a numerical proportion, not indeed fixed and accurate at first, but which tends in the long run to become so."

Weitgehend ähnliche Formulierungen finden sich bis heute in vielen Lehrbüchern der Statistik und Wahrscheinlichkeitsrechnung.[6]

2. Ein ausführlicher und durchdachter Versuch, F-Wahrscheinlichkeiten als Grenzwerte relativer Häufigkeiten zu konzipieren, findet sich in den Arbeiten von R. v. Mises. Als Ausgangspunkt dient ihm der Begriff eines „Kollektivs". Dieser Begriff ist zwar verwandt mit unserem Begriff eines Zufallsgenerators; es gibt jedoch einen gravierenden Unterschied. Wir haben Zufallsgeneratoren als *Verfahren* eingeführt, mit denen sich beliebig viele Situationen erzeugen lassen, die sich durch ihre Zugehörigkeit zu einer Reihe von Situationstypen charakterisieren lassen; die entscheidende Voraussetzung besteht darin, daß man durch Einsichten in die Konstruktion des Zufallsgenerators eine Vorstellung von „gleichen Möglichkeiten" gewinnen kann. Dagegen gründet v. Mises seine Überlegungen nicht auf Eigenschaften eines Verfahrens, sondern auf Eigenschaften der mithilfe eines Verfahrens erzeugbaren Folgen von Situationen.

3. Den Unterschied kann man auch so beschreiben. Es sei irgendein Verfahren gegeben, mit dem sich der Reihe nach beliebig viele Situationen

$$s_1, s_2, s_3, \ldots, s_n, s_{n+1}, \ldots$$

erzeugen lassen. Der „klassische" Ansatz, dem wir in Kapitel 4 gefolgt sind, besteht darin, daß man sich auf Eigenschaften des Verfahrens be-

[6] Vgl. z.B. Fahrmeir et al. 1999, S. 190.

ruft. Aleatorische Wahrscheinlichkeitsaussagen sind zunächst Aussagen *über* Zufallsgeneratoren, nicht über die durch sie erzeugbaren Situationen, insbesondere nicht über relative Häufigkeiten in Mengen solcher Situationen; eine Verknüpfung erfordert ergänzende epistemische Wahrscheinlichkeitsaussagen (vgl. Abschnitt 5.4). — Dem Kollektivbegriff v. Mises' liegt dagegen die Vorstellung zugrunde, daß man sich die Gesamtheit der durch das Verfahren erzeugten Situationen als eine „unendliche Folge" vorstellen könne und Wahrscheinlichkeitsaussagen als Aussagen *über solche Folgen* konzipieren könne. Dementsprechend sagt v. Mises (1933, S. 757):

„Nach der neueren Auffassung der Wahrscheinlichkeitstheorie, wie ich sie in verschiedenen Schriften entwickelt habe, ist jeder Satz der Wahrscheinlichkeitsrechnung eine Aussage über eine unendliche Zeichen- oder Zahlenfolge bestimmter Art."

4. Verfolgen wir den Gedankengang etwas genauer. In seinem Lehrbuch (1931, S. 7) beginnt v. Mises mit folgender Bemerkung:

„In der rationellen Wahrscheinlichkeitsrechnung, mit der wir uns in diesem Buch ausschließlich befassen, handelt es sich stets um eine *Massenerscheinung* oder einen *Wiederholungsvorgang*, bestehend aus sehr vielen Einzelelementen; man kann auch sagen, um eine Folge von 'Beobachtungen', deren jedesmaliges Ergebnis oder Merkmal (Nummer des Loses, Farbe der Erbenblüte) unser Interesse besitzt. Für eine derartige Folge werden wir später (unter genauerer Präzisierung) die Bezeichnung '*Kollektiv*' einführen." Beispiele sind: Die beliebig oft wiederholbare Ziehung eines Loses, die in großen Massen ausgeführte Züchtung von Erbsen, die wiederholt vorgenommene Messung einer bestimmten physischen Größe."

Auch v. Mises' Kollektivbegriff setzt also einen wiederholbaren Vorgang voraus, durch den Situationen (Beobachtungen) erzeugt werden.[7] Aber der Begriff bezieht sich nicht auf den Vorgang (im Sinne eines Verfahrens), *sondern* auf die durch den Vorgang entstehenden Folgen von Situationen.[8] Dies wird allerdings durch den frequentistischen Ansatz erforderlich gemacht: man möchte von relativen Häufigkeiten *ausgehen*.

[7] v. Mises (1931, S. 9) betont: „Wir werden daher auch statt Element des Kollektivs den konkreten Ausdruck 'Beobachtung', statt Merkmal 'Beobachtungsergebnis' gebrauchen. [...] Wesentlich ist, daß der Einzelvorgang *in sich abgeschlossen* ist und sich nach einer *dauernd gleichbleibenden Vorschrift* vollzieht, die über die Bestimmung des Merkmals oder die eventuelle Ausscheidung des Versuches eindeutig verfügt." Wir später besprechen, bei welchen „Massenerscheinungen" dies angenommen werden kann.

[8] Man sollte indessen beachten, daß v. Mises' Formulierungen in diesem Punkt eine charakteristische Ambivalenz aufweisen. Z.B. wird in dem oben angeführten Zitat von der „beliebig oft wiederholbaren Ziehung eines Loses" gesprochen, wohingegen sich jedoch der Kollektivbegriff nicht auf diesen Vorgang, sondern auf dessen realisierte Ergebnisse bezieht.

Aber relative Häufigkeiten gibt es nur in endlichen Mengen von bereits realisierten Situationen oder Ereignissen.

5. Im weiteren möchte ich zeigen, daß die Schwierigkeiten der frequentistischen Auffassung im wesentlichen aus dieser „Verkehrung" des Ansatzes resultieren: daß nicht mit Eigenschaften eines Verfahrens, sondern mit Folgen *realisierter* Situationen begonnen wird. Und zwar gibt es hauptsächlich zwei Schwierigkeiten. Die erste Schwierigkeit, mit der wir uns in diesem Abschnitt beschäftigen, betrifft die Konzeption von Grenzwerten. eine zweite Schwierigkeit betrifft die Definition „zufälliger Folgen". — Es gibt auch noch eine dritte Schwierigkeit, die darin besteht, daß man es bei statistischen Anwendungen meistens nicht mit Folgen, sondern mit irgendwie abgegrenzten Mengen (Gesamtheiten) von Situationen zu tun hat, die nicht ohne weiteres als Folgen aufgefaßt werden können. Bei v. Mises bleibt diese Schwierigkeit unsichtbar, weil er sich von vornherein auf Folgen bezieht, also – ohne es in jedem Fall explizit zu machen – auf ein Verfahren, durch das die Situationen, die die Elemente der Folge bilden, sequentiell entstehen oder erzeugt werden können.[9] Wenn man es jedoch, wie insbesondere in der Sozialstatistik, mit beliebig zustandegekommenen Gesamtheiten zu tun hat, werden Begriffsbildungen, die eine Folge von Situationen voraussetzen, fragwürdig. Dieses Problem hat zu Varianten des frequentistischen Ansatzes geführt, die nicht von Folgen, sondern von beliebigen Gesamtheiten (Klassen) ausgehen.

6. Um F-Wahrscheinlichkeiten als Grenzwerte relativer Häufigkeiten zu definieren, knüpft v. Mises an das im vorangegangenen Abschnitt besprochene Münzwurf-Beispiel an und fährt dann folgendermaßen fort:

„Damit ist die *erste Forderung*, die wir an eine einfache Massenerscheinung (Alternative[10]) stellen müssen, wenn sie Gegenstand der Wahrscheinlichkeitsrechnung werden soll, gefunden: Es muß möglich sein, sich die Folge der Einzelerscheinungen oder Beobachtungen (Elemente) *unbeschränkt fortsetzbar zu denken, und es müssen die relativen Häufigkeiten* für das Auftreten der beiden

[9] In Reichenbachs Variante der frequentistischen Konzeption heißt es (1949, S. 136): „Daß sich die Wahrscheinlichkeitsaussage stets auf Folgen bezieht, haben wir in die logische Grundlegung unserer axiomatischen Wahrscheinlichkeitsrechnung aufgenommen; infolgedessen können wir auch Aussagen über die Ordnung dieser Folgen in der axiomatischen Wahrscheinlichkeitsrechnung machen."

[10] v. Mises bezieht sich hier auf Situationen, bei bei denen nur zwei Merkmalswerte, 0 und 1, unterschieden werden, wie in dem „Kopf-oder-Adler"-Spiel, das er als ein Beispiel verwendet.

Merkmale *bestimmte Grenzwerte besitzen.* Diese Grenzwerte sind es, die man als die (mathematischen) Wahrscheinlichkeiten für das Auftreten der beiden Merkmale anzusehen hat.“

Aber kann man annehmen, und in welcher Weise kann man davon sprechen, daß es solche Grenzwerte überhaupt gibt? Die Frage muß gestellt werden, und auch v. Mises (1931, S. 11) bemerkt:

„Empirisch feststellen läßt sich natürlich nicht, ob ein Grenzwert existiert oder nicht, da man immer nur endlich viele Versuche machen kann. Die unendliche Reihe von Nullen und Einsern, die das Ergebnis eines ideellen Kollektivs darstellt, ist uns weder unmittelbar noch formelmäßig gegeben. Man kann daher auch nicht durch Ausführung einer Rechnung nachprüfen, ob die Forderung nach Existenz des Grenzwertes in einem bestimmten Fall erfüllt ist.“

Das heißt jedoch, daß die in der Mathematik übliche Definition eines Grenzwerts in diesem Fall nicht auf eine konstruktive Weise anwendbar ist. Denn dies würde voraussetzen, daß man die Glieder einer Folge numerischer Ausdrücke durch ein *induktives* Verfahren definieren kann. Damit ist gemeint, daß man zur Definition einer Folge

$$s_1, s_2, s_3, \ldots, s_n, s_{n+1}, \ldots$$

eine Regel benötigt, die festlegt, wie jedes Glied s_n durch die vorangegangenen Glieder $s_1, \ldots, s_{n-1}$ gebildet werden kann. Und zwar muß diese Regel unabhängig von ihrer Verwendung zur tatsächlichen Erzeugung von Gliedern der Folge definierbar sein; oder anders formuliert: es muß möglich sein, die Regel „ein für allemal“ festzulegen, also die Folge *durch die Regel zu definieren.*

7. Beziehen wir uns auf die oben angegebene Folge (s_i) und bezeichnen mit $h_n(\tilde{x})$ die relative Häufigkeit des Merkmals $\tilde{x}$ in den ersten n Gliedern der Folge, muß man für die Existenz eines Grenzwertes in der üblichen mathematischen Notation folgendes zeigen:

$$\forall \epsilon > 0 \; \exists m \; \forall n, n' \geq m \; : \; |\, h_n(\tilde{x}) - h_{n'}(\tilde{x}) \,| < \epsilon$$

Aber um das in konstruktiver Weise zeigen zu können, muß es möglich sein, sich auf eine Regel zu beziehen, nach der die Glieder der Folge zustande kommen. Man kann es auch so sagen: wenn man *heute* zeigen will, daß eine Folge einen Grenzwert hat, dann muß man auch heute schon über die gesamte Folge sprechen können. Dagegen können die mathematischen Begriffsbildungen und insbesondere der Begriff eines Grenzwertes

nicht sinnvoll auf Folgen angewendet werden, deren Elemente sich in einem historischen Prozeß realisieren, der heute noch nicht abgeschlossen ist.

8. Ein Kollektiv kann jedoch nicht durch eine induktive Regel definiert werden, wie sie für den mathematischen Grenzwertbegriff erforderlich ist.[11] Und zwar sind hier zwei Überlegungen wichtig. Zunächst kann man an Folgen bereits realisierter Situationen denken. Solche Folgen sind jedoch stets endlich und man weiß nicht, wie sie weitergehen werden. Schon die Vorstellung, sich Folgen dieser Art als beliebig in die Zukunft hinein fortgesetzt zu denken, bereitet Schwierigkeiten. Aber auch wenn man zu einer solchen Vorstellung bereit wäre, könnte natürlich mit Hinweisen auf ein endliches Anfangsstück der Folge niemals eine Aussage über einen Grenzwert begründet werden. Wenn man gleichwohl einfach annimmt oder postuliert, daß es auch für die beliebig fortgesetzt gedachten Folgen einen Grenzwert relativer Häufigkeiten gibt, verwandelt sich die Überlegung in eine metaphysische Spekulation, die auf eine Negation historischer Prozesse hinausläuft. Im übrigen würde eine solche Spekulation auch gar nichts nützen, denn auch wenn es einen Grenzwert gäbe, würden sich daraus keinerlei Implikationen für jedes endliche Anfangsstück der Folge ergeben; d.h. die empirischen Häufigkeiten könnten sich beliebig lange ganz unabhängig von den postulierten F-Wahrscheinlichkeiten entwickeln.[12]

9. v. Mises hat sich deshalb in seinen theoretischen Überlegungen gar nicht auf reale Kollektive bezogen, sondern auf Verfahren, durch die

[11] Ich spreche von einer *induktiven* Regel, um an das mathematische Prinzip der vollständigen Induktion zu erinnern, auf dem insbesondere auch Beweise für Existenzaussagen über Grenzwerte beruhen.

[12] Dazu bemerkte Braithwaite (1955, S. 125): „... since a sequence of numbers can start with any finite set of terms and still tend to any limit whatsoever, not only is the limiting value of the class-ratio not determined by the class-ratio in any finite set of classes of reference, but – and this is more serious – a limiting value does not determine what will be the class-ratio in any finite class of reference. The former lack of connexion implies that the limiting value cannot be regarded as a logical construction out of class-ratios in observable classes; the latter lack of connexion implies that it cannot be regarded as a theoretical concept either, since no conclusion as to class-ratios in finite classes (which are all that are observable) can be drawn from it.“ — Dies wird auch von einigern Anhängern frequentistischer Wahrscheinlichkeitskonzeptionen eingeräumt, z.B. von Kolmogorov (1963, S. 369): „The frequency concept based on the notion of *limiting frequency* as the number of trials increases to infinity, does not contribute anything to substantiate the applicability of the results of probability theory to real practical problems where we have always to deal with a finite number of trials.“

Kollektive beliebiger Länge erzeugt werden können. Infolgedessen kann er sich jedenfalls gedanklich auf eine beliebig lange Folge von Situationen und mithin von relativen Häufigkeiten beziehen. Aber es wird dann erforderlich, diese Folgen genauer zu charakterisieren. Dafür gibt es zwei Möglichkeiten. Man kann bei Eigenschaften des Verfahrens ansetzen, mit dem die Elemente der Folgen erzeugt werden. Dies ist der „klassische" Ansatz, dem wir mit unserer Konzeption von Zufallsgeneratoren gefolgt sind. Aber es ist offensichtlich, daß ein Zufallsgenerator keine induktive Definition für die Folge von Situationen liefert, die durch ihn erzeugt werden; und mithin kann auch die Vorstellung eines Grenzwertes bei Folgen, die durch einen Zufallsgenerator erzeugt werden, nicht sinnvoll angewendet werden.[13] Eine zweite Möglichkeit besteht darin, das Verfahren durch Eigenschaften der Folgen zu charakterisieren, die durch das Verfahren erzeugt werden. Diesen Weg schlägt v. Mises ein. Wie in Abschnitt 6.3.2 genauer diskutiert werden wird, versucht er, Kollektive als in spezifischer Weise „regellose" Folgen zu definieren. Dann gerät man jedoch erneut in einen Konflikt mit der Annahme, daß es für die in den Folgen realisierten relativen Häufigkeiten einen bestimmten Grenzwert gibt. Denn der Versuch, eine Folge als „regellos" zu definieren, impliziert, daß sie nicht durch ein induktives Verfahren definiert werden kann; und infolgedessen ist auch in diesem Fall die Vorstellung eines Grenzwerts nicht durchführbar.

10. Aber auch wenn es gelingen könnte, die Idee einer „regellosen Folge" durch ein mathematisches Konstruktionsverfahren zu definieren (also selbst durch eine Regel[14]), wäre für das Anliegen des frequentistischen Ansatzes nicht viel gewonnen. Denn die eigentlich interessante, aber auch fragwürdige Behauptung betrifft ja nicht ein mathematisches Konstruktionsproblem, sondern die Konzeption eines empirisch verwendbaren Wahrscheinlichkeitsbegriffs. Die Frage ist, mit welcher Berechtigung bzw. unter welchen Bedingungen man empirisch ermittelbare relative Häufigkeiten als „Ausdruck" oder „Manifestationen" hypothetisch unterstellbarer „Konstanten" (F-Wahrscheinlichkeiten) deuten kann.

[13]Ein Zufallsgenerator soll gerade garantieren, daß keine induktive Regel gefunden werden kann, mit der sein Verhalten vorausgesagt werden könnte. Eine andere Frage ist, ob Zufallsgeneratoren, die dieser Forderung genügen, praktisch hergestellt werden können. Bei allen in Verwendung befindlichen Zufallsgeneratoren wird man nicht ausschließen können, daß solche Regeln gefunden werden könnten. Aber das bedeutet nur, daß ein Verfahren nicht länger als ein Zufallsgenerator verwendbar wäre, sobald man eine Regel gefunden hätte, um seine Funktionsweise voraussagbar zu machen. Es ist klar, daß in diesem Sinne die Möglichkeit, Zufallsgeneratoren praktisch herstellen zu können, von dem jeweils verfügbaren Regelwissen abhängig ist.

[14]Dies würde natürlich voraussetzen, daß die Anforderungen an „Regellosigkeit" mehr oder weniger weitgehend abgeschwächt werden.

11. Daß sich hinter dieser Deutungsabsicht ein Problem verbirgt, kann man sich anhand folgender Ausführungen von E. Nagel (1939, S. 19) verdeutlichen:

> "The basic ideas of the frequency conception of probability emerge upon an examination of such a statement as 'The probability that a person of thirty residing in the United States survives his thirty-first birthday is .945.' The meaning of such a statement can be ascertained by examining how it is established. That procedure, greatly simplified, is somewhat as follows. Suppose that during a period of years there is no migration to or from the United States, and that during these years exact counts are made of its inhabitants who fall into definite age groups. Thus, suppose that in 1900 there are 2,000,000 persons who have just reached their thirtieth birthday, and that exactly one year later there are 1,890,000 persons who have just reached their thirty-first birthday; that is, of the thirty-year-olds in 1900 a ratio of .9450 survive at least another year. We imagine that similar figures are obtained for the four succeeding years, and that the ratios of thirty-year-olds who survive their thirty-first birthday are .9452, .9456, .9451, and .9454, respectively. We notice that, although these ratios are not constant, the differences do not appear until the fourth decimal is reached. We may say, therefore, that during these five years approximately 945 out of a thousand thirty-year-old residents of the United States live for at least another year; and we may make the further *assumption* that for an *indefinite number of future years* the corresponding ratios of survivals remain in the neighborhood of .945. Accordingly, the statement 'The probability that a thirty-year-old resident of the United States survives his thirty-first birthday is .945' means that in the long run the relative frequency with which thirty-year-olds in the United States survive for at least one year is approximately .945."

Soweit es sich in diesen Ausführungen um empirische Aussagen handelt, wird offenbar ein Wahrscheinlichkeitsbegriff gar nicht benötigt, sondern man kommt vollständig mit relativen Häufigkeiten aus. Auch um die in vier Jahren beobachteten relativen Häufigkeiten als näherungsweise gleich oder als kleine Abweichungen von ihrem Durchschnittswert zu betrachten, ist ein Wahrscheinlichkeitsbegriff nicht erforderlich. Ein solcher Begriff wird hier von Nagel vielmehr nur deshalb verwendet, um eine metaphysische und in ihrer Begründbarkeit gar nicht reflektierte Hypothese einzuführen: daß es auch in aller Zukunft so bleiben wird, wie man in den vier Jahren beobachtet hat. Aber auch um diese Hypothese zu formulieren, benötigt man tatsächlich gar keinen Wahrscheinlichkeitsbegriff; denn man kann ja einfach sagen, daß man – aus welchen Gründen auch immer – daran glaubt, daß die in einer gewissen Zeitspanne beobachteten relativen Häufigkeiten sich auch in der Zukunft nicht oder nur wenig verändern werden.

12. Dieses Beispiel verdeutlicht ein grundsätzliches Problem, in das sich der frequentistische Ansatz verstrickt. Zunächst möchten seine Anhänger vorschlagen, daß man die *Bedeutung* von Wahrscheinlichkeitsaussagen dadurch verstehen kann, daß man sie in Aussagen über relative Häufigkeiten transformiert. Aber wenn bzw. insoweit dies gelingt, kann man über relative Häufigkeiten sprechen, und ein davon zu unterscheidender Wahrscheinlichkeitsbegriff wird ganz überflüssig. Insbesondere können alle empirischen Aussagen der Sozialstatistik ganz ohne Verwendung eines Wahrscheinlichkeitsbegriffs formuliert werden. Wenn dagegen gleichwohl an der Konzeption eines Wahrscheinlichkeitsbegriffs festgehalten werden soll, dient dies nicht der Explikation eines semantischen Problems – wie Wahrscheinlichkeitsaussagen *verstanden* werden können –, sondern es wird eine fragwürdige metaphysische Annahme eingeführt. Und zwar nicht so, daß man an diese Annahme glauben oder auch nicht glauben kann; sondern ein Bekenntnis zu dieser Annahme wird zu einer *Sinnvoraussetzung* für die Formulierung von Wahrscheinlichkeitsaussagen gemacht. Zwei Fragen werden also in einer wenig durchsichtigen Weise vermengt: wie man Wahrscheinlichkeitsaussagen verstehen und begründen kann; und ob man glauben soll, daß sich auf der Grundlage beobachtbarer relativer Häufigkeiten Gesetzmäßigkeiten für ihre Entwicklung in beliebig in die Zukunft fortsetzbaren historischen Prozessen formulieren und begründen lassen.[15]

[15]Fragen der zweiten Art werden von Philosophen unter dem Namen „Induktionsproblem“ diskutiert. Dabei geht es darum, wie man Regelmäßigkeiten im Ablauf des Geschehens feststellen, formulieren und mithilfe von Beobachtungen und Experimenten begründen kann. Insofern kann man auch sagen, daß der frequentistische Ansatz Fragen, die sich auf das Verständnis und die Begründbarkeit von Wahrscheinlichkeitsaussagen richten, mit problematischen Annahmen zum Induktionsproblem vermischt. Reichenbach hat für seine Variante des frequentistischen Ansatzes einen expliziten Lösungsversuch unternommen; aber er bemerkt auch (1945, S. 512): „That the inductive inference is indispensable for the frequency theory and that it constitutes the most serious difficulty of this theory, is now generally recognized.“

6.3 Zufällige Folgen

Eine zweite Schwierigkeit für den frequentistischen Ansatz entsteht daraus, daß man sich auf relative Häufigkeiten in „zufälligen“ Folgen von Situationen (Sachverhalten, Ereignissen) beziehen möchte. Um die Schwierigkeiten zu verstehen, beginnen die folgenden Überlegungen mit einigen allgemeinen Bemerkungen zum Verständnis des Wortes ‘zufällig’.

6.3.1 Was heißt ‘zufällig’?

1. Das Wort ‘zufällig’ wird in unterschiedlichen Bedeutungen verwendet. Eine erste wichtige Unterscheidung ergibt sich daraus, ob man sich mit dem Wort auf den Entstehungszusammenhang oder auf die Beschaffenheit von etwas bezieht. Im ersten Fall kann man davon sprechen, daß etwas *zufällig entstanden* oder *sich zufällig ereignet* hat. Im zweiten Fall ist meistens gemeint, daß sich in der Beschaffenheit eines Sachverhalts *keine Regelmäßigkeiten* erkennen lassen.

2. In der Umgangssprache wird das Wort ‘zufällig’ oft dann verwendet, wenn für das Entstehen einer Situation (eines Ereignisses, eines Sachverhalts) keine erkennbaren Gründe vorliegen. Hieran hat auch bereits Aristoteles in seiner Physikvorlesung angeknüpft, wo es heißt:[16]

„Unter dem, was geschieht, erfolgen die einen (Ereignisse) *wegen irgendetwas*, die anderen nicht – von den ersten erfolgen die einen gemäß *vorsätzlicher Absicht*, die anderen nicht nach solcher Absicht, beide befinden sich aber unter den Ereignissen wegen etwas [...]. ‘Wegen etwas’ ist alles das, was sowohl durch planende Vernunft hervorgebracht sein könnte oder auch durch Naturanlage. Wenn Derartiges also nun *als nebenbei eintretende Wirkung* erfolgt, dann sagen wir, es sei ‘aus Fügung’.“ (S. 75) „[...] Ihr Unterschied liegt darin, daß ‘Zufall’ sich über einen weiteren Bereich erstreckt. Ereignisse auf Grund von Fügung sind alle auch zufällig; (umgekehrt) diese sind nicht alle auf Grund von Fügung.“ (S. 79) „Schicksalsfügung und (Ereignisse) auf Grund von Fügung können nur für solche (Wesen) zutreffen, für die auch ‘Glückhaben’ gelten kann, und überhaupt ‘Handlung’. Also muß *Fügung* sich notwendig beziehen auf mögliche Handlungen. [...] Der *Zufall* hingegen trifft auch auf Tiere und einen großen Teil des Unbelebten zu; z.B. ‘das Pferd kam zufällig heraus’ sagen wir dann, wenn es zwar durch sein Herauskommen einem Unglück entging, aber nicht herauskam in der Absicht, diesem Unglück zu entgehen. Und: ‘Der Schemel fiel zufällig um’; er stand zwar da, um darauf zu sitzen, aber er ist nicht, um darauf zu sitzen, umgefallen! Es ist also klar: Wenn im Bereich der Geschehnisse, die im strengen Sinn wegen etwas eintreten und deren Ursache außer ihnen

[16]Ich zitiere nach einer Übersetzung von H. G. Zekl, vgl. Aristoteles 1987, S. 75ff.

liegt, etwas geschieht, das mit dem Ergebnis nicht in eine Deswegen-Beziehung zu bringen ist, dann nennen wir das 'zufällig'. 'Auf Grund von Fügung' (sagen wir) von solchen Ereignissen, die im Bereich sinnvoll gewollter Handlungen bei (Wesen), die die Fähigkeit zu planendem Vorsatz haben, zufällig eintreten.“ (S. 79–81)

Diese Erläuterungen machen eine, auch in der heutigen Umgangssprache noch verständliche Bedeutung des Wortes 'zufällig' deutlich.

3. An diesen Sprachgebrauch knüpft auch der in diesem Text verwendete Begriff eines Zufallsgenerators an. Ein Zufallsgenerator ist ein Verfahren, mit dem Situationen (eines gewissen Typs) erzeugt werden können; aber welche Situation jeweils erzeugt wird, ist nicht voraussagbar. Ein Zufallsgenerator ist gewissermaßen die Negation intentionalen Handelns. Man kann mit einem Zufallsgenerator nicht absichtlich bestimmte Situationen erzeugen. Es ist bestenfalls ein metaphorischer Sprachgebrauch, wenn man sagt, daß jemand sich vornehmen könne, mit einem Würfel eine '6' zu erzielen. Der Begriff eines Zufallsgenerators schließt es aus, daß man üben und lernen könnte, mit ihm absichtlich bestimmte Resultate zu erzeugen.

4. Dieser Sprachgebrauch, daß etwas zufällig entsteht oder entstanden ist, impliziert in keiner Weise, daß „der Zufall“ als eine besondere oder auch vielleicht nur unbekannte Art von Ursache aufzufassen sei. Das hat auch Aristoteles mit seinen Erläuterungen nicht gemeint. Wird davon gesprochen, daß etwas zufällig entstanden ist oder entsteht, hat dies weder im positiven noch im negativen Sinn Implikationen für die Frage, ob das Geschehen durch Ursachen vollständig oder teilweise oder gar nicht bestimmt wird. Es ist deshalb auch müßig, darüber zu spekulieren, ob die Situationen, die durch einen Zufallsgenerator erzeugt werden, durch ihn und seine Verwendung „determiniert“ sind. Zufallsgeneratoren sind Artefakte, von Menschen ausgedachte und hergestellte Verfahren, um Situationen zufällig zu erzeugen. Und sie haben weiterhin den Zweck, (Denk-) Modelle zu verschaffen, um über das zufälllige Entstehen von etwas nachdenken zu können.[17] Diese Verwendung impliziert allerdings, daß man über die Beschaffenheit des Prozesses, der die Situationen hervorbringt, keine „substantiellen“ Aussagen machen kann

[17]Sie werden natürlich auch für Glücksspiele aller Art verwendet; und sie werden auch gelegentlich verwendet, um „faire“ Entscheidungen zu erreichen. Beides hat eine sehr lange Geschichte; zur Geschichte der Verwendung von Zufallsverfahren für Entscheidungen vgl. Bennett 1998, Kap. 2.

bzw. möchte.[18]

5. Wenn man sich auf Zufallsgeneratoren bezieht, gewinnt also das Wort 'zufällig' seine Bedeutung daraus, das es auf einen Prozeß verweist, durch den etwas entstanden ist. Daß eine Situation zufällig ist, meint in diesem Fall, daß sie zufällig *entstanden* ist, daß sie durch einen Zufallsgenerator hervorgebracht worden ist; wobei letzteres zweierlei bedeuten kann: daß die Situation tatsächlich durch die Verwendung eines artifiziellen Zufallsgenerators hervorgebracht worden ist oder daß man sich den Prozeß, durch den die Situation entstanden ist, als einen Zufallsgenerator vorstellen möchte.

6. Eine wesentlich andere Verwendung des Wortes 'zufällig' bezieht sich auf die Beschaffenheit realisierter Situationen. Man bezieht sich dann nicht darauf, wie eine Situation entstanden ist, sondern auf Aspekte ihrer Erscheinungsform. Um sich die Unterscheidung zu verdeutlichen, denke man an einen Zufallsgenerator, mit dem Nullen und Einsen erzeugt werden können. Betätigt man diesen Zufallsgenerator 12 Mal, können zum Beispiel folgende Sequenzen entstehen:

$$1,1,1,1,1,1,1,1,1,1,1,1$$
$$0,1,0,1,0,1,0,1,0,1,0,1$$
$$1,0,0,1,1,1,0,1,0,1,1,0$$

Betrachtet man diese Sequenzen als Ergebnisse eines Zufallsgenerators, sind sie gleichermaßen zufällig hervorgebracht worden. Wenn der Zufallsgenerator so beschaffen ist, daß es für das Auftreten von Nullen und Einsen gleiche Möglichkeiten gibt, haben die drei Sequenzen (genauer formuliert: die ihnen entsprechenden Situationstypen) die gleiche aleatorische Wahrscheinlichkeit. Andererseits erscheinen die ersten beiden Sequenzen als sehr „regelmäßig“, die letzte jedoch als „unregelmäßig”. Also kann man die letzte Sequenz auch „zufällig“ nennen, womit dann aber nicht gemeint ist, daß sie durch einen Zufallsgenerator erzeugt worden ist, sondern daß sie „unregelmäßig“ beschaffen ist.

7. In der Geschichte der Wahrscheinlichkeitsrechnung sind einige Unklarheiten daraus entstanden, daß zwischen den beiden Weisen, von „zufällig“ zu sprechen, nicht immer deutlich genug unterschieden wor-

[18]Darauf beruht der unter Statistikern verbreitete Glaube, daß man bei Anwendungen wahrscheinlichkeitstheoretischer Begriffsbildungen von der Natur der Prozesse, auf die man sich beziehen möchte, vollständig abstrahieren könne. Andererseits beruht darauf auch eine eigentümliche Schwäche statistischer Modelle: daß sie keine Einsichten in Prozesse vermitteln können.

den ist. Dazu beigetragen haben auch die Versuche, einen quantitativen Wahrscheinlichkeitsbegriff durch eine „Idealisierung“ relativer Häufigkeiten zu definieren. Ein scheinbares Paradox kann zur Verdeutlichung dienen. Zur Erläuterung beziehe ich mich auf den Zufallsgenerator, der zur Erzeugung der drei oben angeführten Sequenzen verwendet worden ist. Angenommen, man erzeugt mit diesem Zufallsgenerator eine vierte Sequenz. Nachdem man ihn 8 Mal verwendet hat, stellt man fest, daß jedesmal eine ‘0’ herausgekommen ist. Sollte man dann nicht erwarten, daß die Wahrscheinlichkeit, beim nächstenmal eine ‘1’ zu erzielen, größer als 1/2 ist? Andererseits impliziert der Begriff eines Zufallsgenerators, daß die aleatorischen Wahrscheinlichkeiten, die durch ihn definiert werden, unabhängig davon sind, welche Ergebnisse in der Vergangenheit aufgetreten sind; also ganz gleichgültig, ob eine Sequenz mit 8 Nullen oder irgendeine andere Sequenz aufgetreten ist, die Wahrscheinlichkeit, daß bei der nächsten Anwendung eine ‘1’ herauskommt, ist 1/2.

8. Wie kann man dieses scheinbare Paradox verstehen?[19] Man muß sich überlegen, warum es plausibel erscheinen könnte, nach dem Auftreten von 8 Nullen zu erwarten, daß die Wahrscheinlichkeit, beim nächstenmal eine ‘1’ zu erzielen, größer als 1/2 sein sollte. Es scheint so, daß dies durch ein frequentistisches Verständnis von Wahrscheinlichkeiten nahegelegt wird, jedenfalls wenn man sich an die oft zur Erläuterung verwendete Formulierung hält, daß Wahrscheinlichkeit die zu erwartende relative Häufigkeit sei. Wenn also in unserem Beispiel die Wahrscheinlichkeiten für Nullen und Einsen jeweils 1/2 sind, sollte man hiernach erwarten, daß in einer Sequenz der Länge 12 ungefähr 6 Nullen und 6 Einsen vorkommen. Also sollte man doch erwarten, daß, wenn man schon 8 Nullen bekommen hat, danach eher Einsen auftreten werden? Aber diese Erwartung steht ersichtlich in einem Widerspruch zu unserem Wissen über die Beschaffenheit und Funktionsweise des Zufallsgenerators.[20] Der Denkfehler resultiert deshalb aus einem Mißverständnis des aleatorischen Wahrscheinlichkeitsbegriffs; allerdings ein Mißverständnis, dem durch eine frequentistische Rhetorik Vorschub geleistet wird.

[19] Es wird auch als „gambler’s fallacy“ bezeichnet; eine instruktive Diskussion findet sich bei Robertson 1905, Letter VII.

[20] „Es ist also bei der Berechnung der zu erwartenden Gewinne nur auf die Spiele, welche noch gemacht werden müssen, Rücksicht zu nehmen, nicht auf die bereits gemachten. Denn für jedes einzelne folgende Spiel ist die Wahrscheinlichkeit, dass das Glück diejenigen Spieler begünstigt, welche es bisher bevorzugt hat, nicht grösser als die, dass es diejenigen begünstigt, welche es bisher stiefmütterlich behandelt hat.“ (Bernoulli 1899a, S. 13)

9. Das scheinbare Paradox zeigt auch erneut, wie wichtig es ist, daß man sich zur Begründung eines quantitativen Wahrscheinlichkeitsbegriffs auf Zufallsgeneratoren bezieht. Wüßte man in unserem Beispiel nicht, daß die Nullen und Einsen durch einen Zufallsgenerator erzeugt werden (oder erzeugt worden sind), gäbe es zunächst überhaupt keinen Grund für irgendeine Erwartung; oder genauer gesagt: eine Erwartungsbildung könnte sich dann nicht auf Wahrscheinlichkeiten berufen. Denn man kann sich natürlich immer irgendwelche Erwartungen bilden. Hat man z.B. beobachtet, daß 8 Mal eine ‘0’ entstanden ist, kann man erwarten, daß das nächste Mal wieder eine ‘0’ herauskommen wird; oder man kann glauben, daß der erzeugende Prozeß, über den man nichts weiß, zu Abwechslungen neigt, und deshalb erwarten, daß beim nächstenmal eine andere Zahl erscheinen wird; oder wenn man sich ganz unschlüssig ist, kann man ein Orakel befragen. Aber nur wenn man sich auf einen Zufallsgenerator beziehen kann, kann man zur Begründung von Erwartungen auf quantitative Wahrscheinlichkeiten verweisen (im Unterschied zu relativen Häufigkeiten, die durch kontextabhängige Beobachtungen einer Menge realisierter Situationen entstehen).[21]

10. v. Mises war sich dieses Problems durchaus bewußt, und er hat deshalb für seinen Begriff eines Kollektivs (im Sinne einer beliebig langen Folge von Beobachtungen) nicht nur gefordert, daß es einen „Grenzwert“ der relativen Häufigkeiten geben soll; er hat außerdem gefordert, daß die Folge der Beobachtungen „regellos“ bzw. „zufällig“ sein soll. Dadurch entsteht aber für den frequentistischen Ansatz ein weiteres schwieriges Problem. Denn wie bei der Besprechung dieses Ansatzes in Abschnitt 6 dargelegt worden ist, besteht der Grundgedanke darin, zur Definition eines Wahrscheinlichkeitsbegriffs nicht von Zufallsgeneratoren auszugehen, sondern von Folgen von Situationen (Ereignissen, Beobachtungen), die *irgendwie* entstanden sein können. Um zu klären, was es bedeuten könnte, daß es sich um zufällige Folgen handelt, kann sich deshalb der frequentistische Ansatz nicht auf Eigenschaften eines Zufallsgenerators berufen. Es kann also eine Folge von Situationen auch nicht deshalb als zufällig bezeichnet werden, weil sie zufällig (durch einen Zufallsgenerator) *entstanden* ist. Vielmehr muß der frequentistische Ansatz bei der zweiten der oben unterschiedenen Verwendungsweisen des Wortes

[21] Bei v. Mises (1931, S. 4) heißt es in einer scheinbaren Analogie: „Erst muß ein ‘Kollektiv’ da sein, bevor von Wahrscheinlichkeiten gesprochen werden kann.“ Die Formulierung ist allerdings merkwürdig, denn würde man sie wörtlich nehmen, könnte man erst dann von Wahrscheinlichkeiten sprechen, wenn der unendliche Prozeß, der die Elemente eines Kollektivs erzeugt, abgeschlossen ist.

'zufällig' anknüpfen; d.h. es muß der Versuch unternommen werden, den Sinn des Wortes 'zufällig' durch Rückgriff auf Eigenschaften eines Kollektivs, einer Folge von *realisierten* oder als realisiert gedachten Situationen, zu explizieren. Das jedoch führt, wie im nächsten Abschnitt besprochen werden soll, in eine Reihe von Schwierigkeiten.

6.3.2 v. Mises' Regellosigkeit

1. Die Frage ist, ob man im Hinblick auf eine *gegebene* Folge von Situationen sagen kann, daß sie „zufällig beschaffen" ist. Der Versuch, eine solche Charakterisierung zu begründen, kann sich also nicht auf Überlegungen berufen, die sich darauf beziehen, wie die Elemente der Folge zustandegekommen sind, sondern nur auf in der Folge realisierte Eigenschaften. In diesem Abschnitt besprechen wir Überlegungen, die v. Mises angestellt hat, um zu einer Definition zu gelangen. In seinem Lehrbuch (1931, S. 11f) beginnt er mit mit folgendem Beispiel:

„Betrachten wir als Massenerscheinung die längs einer Eisenbahnlinie aufgestellten Kilometersteine und die dazwischen stehenden kleinen 100-Meter-Marken. Den ersteren wollen wir das Merkmal '1', den letzteren '0' zuweisen. Dann ist klar, daß, wenn wir nur lange genug an der Bahnlinie entlang fahren und die Nullen und Einser zählen, annähernd der zehnte Teil aller Steine das Beobachachtungsergebnis 1 aufweisen wird, während 9/10 das Merkmal 0 besitzen. Es läßt sich sehr leicht ausrechnen, daß auf einer unbegrenzt fortlaufenden Bahnstrecke tatsächlich

$$\lim_{n\to\infty} \frac{n_1}{n} = 0.1, \quad \lim_{n\to\infty} \frac{n_0}{n} = 0.9$$

gelten muß.

Der wesentliche Unterschied, der zwischen diesem Fall und den früher angeführten Beispielen [exemplarisch das 'Kopf-oder-Adler'-Spiel] von Massenerscheinungen besteht, kommt in folgender Überlegung zum Ausdruck: Wenn man beim Zählen nicht alle Steine, sondern beispielsweise *nur jeden zweiten* beachtet, so wird die relative Häufigkeit der großen Steine auf keinen Fall mehr annähernd 1/10 sein, sondern entweder 0 oder 1/5, je nachdem, ob man bei dem 1., 3., 5., 7., 9. der auf einen großen Stein folgenden kleinen Steine zu zählen anfängt, oder bei dem 2., 4., 6., 8. kleinen Stein bzw. bei einem großen Stein selbst. Wenn man aber 'Kopf-oder-Adler' spielt, so ändert sich, wie jedermann weiß, nichts an der Aussicht, Kopf zu werfen, dadurch, daß man etwa nur jeden zweiten Wurf beachtet. Es gehört zu den wesentlichsten Vorstellungen, die sich für uns mit dem Begriff des 'Zufalls' oder des Glücksspiels untrennbar verknüpfen, daß man seine Chancen nicht durch eine systematische Auswahl der Spiele beeinflussen kann. Hier kommt eben das zur Geltung, was in der

Einleitung als das 'Prinzip der Unmöglichkeit eines Spielsystems' angedeutet wurde und was gemeint ist, wenn man den Zufall 'gesetzlos' oder 'unvereinbar mit irgendeinem Gesetz' nennt."

Mit dieser Überlegung begründet v. Mises, daß ein Kollektiv, so wie er den Begriff verwenden möchte, nicht nur der Forderung genügen muß, daß es einen „Grenzwert" relativer Häufigkeiten gibt, sondern daß es außerdem „gesetzlos" sein muß.

2. Das oben angeführte Zitat geht folgendermaßen weiter:

„Die Massenerscheinungen, die den Gegenstand der Wahrscheinlichkeitsrechnung bilden, müssen noch einer *zweiten Forderung* genügen, die wir vorläufig wie folgt formulieren: Wird aus der gesamten unendlich gedachten Folge der Elemente eine *unendliche Teilfolge* so gebildet, daß über die Zugehörigkeit eines Elementes zur Teilfolge *ohne Kenntnis seines Merkmals*, nur nach der Stellung, die dieses Element in der Gesamtfolge hat, entschieden wird, so *müssen auch innerhalb der Teilfolge die relativen Häufigkeiten* des Auftretens der beiden Merkmale *Grenzwerte besitzen*, und zwar *dieselben* wie in der Gesamtfolge."

Hier stellt sich natürlich die Frage, wie dieser Gedankengang, der sich auf eine *unendliche* Folge bezieht, genauer gefaßt werden kann. Die Ausführungen v. Mises' gehen so weiter:

„Um dies näher zu präzisieren, führen wir den Begriff der *'Stellenauswahl'* oder 'Positionsauswahl' ein und geben dafür zunächst folgende zwei Beispiele: a) Irgendein arithmetisches Gesetz wird vorgegeben, etwa in der Form: Jede zehnte Beobachtung soll ausgewählt werden; oder: Jedes Spiel, dessen Nummer in der Spielfolge eine Primzahl ist, soll gelten usf. b) Beim 'rouge et noir' soll nur ein solches Spiel gelten, dessen vorvorhergehendes 'rot' ergeben hat; oder jedes, das auf eine ununterbrochene Serie von 4-mal 'schwarz' folgt. – Das Gemeinsame an diesen Auswahlverfahren ist, wie schon vorausgeschickt, daß über die Zugehörigkeit eines Elements zu der ausgewählten Teilfolge *vor* der Feststellung seines Merkmals, also unabhängig davon, ob die Beobachtung '0' oder '1' ergibt, entschieden wird. Wir definieren daher den Begriff 'Stellenauswahl' wie folgt: Aus einer unendlichen Folge von Beobachtungen wird durch 'Stellenauswahl' eine Teilfolge gebildet, indem man eine Vorschrift angibt, durch die über die Zugehörigkeit oder Nichtzugehörigkeit der $n$ten Beobachtung ($n = 1, 2, \ldots$) zur Teilfolge unabhängig von dem Ergebnis dieser $n$ten Beobachtung und höchstens unter Benutzung der Kenntnisse der vorhergegangenen Beobachtungsergebnisse entschieden wird."

3. Schließlich sei zur Verdeutlichung des Gedankengangs noch folgende zusammenfassende Bemerkung (S. 13f) angeführt:

„Ganz allgemein muß es als charakteristische Eigenschaft aller Massenerscheinungen, mit denen man es in der Wahrscheinlichkeitsrechnung zu tun hat, angesehen werden, daß die relativen Häufigkeiten sich durch eine Stellenauswahl nicht beeinflussen lassen. Man mag ein Auswahlgesetz wie immer vorschreiben: wenn man die Beobachtungen dann nur *lange genug fortsetzt*, wird sich immer zeigen, daß innerhalb der ausgewählten Teilfolge die gleichen relativen Häufigkeiten der beiden Ergebnisse bestehen. Auf Glücksspiele angewendet heißt dies, daß ein ‘Spielsystem’ nicht möglich ist. Wir nennen die hier auftretende, durch die Unmöglichkeit des Spielsystems charakterisierte, Art der Zuordnung zwischen den Merkmalen und den Ordnungsnummern der Elemente eine *‘regellose’* oder *‘zufallsartige’*. Für Erscheinungen oder Vorgänge, die – gehörig idealisiert – den beiden Forderungen nach Existenz des Grenzwertes und nach ‘Regellosigkeit’ der Zuordnung genügen, führen wir den Ausdruck *‘Kollektiv’* ein, und speziell *einfachstes Kollektiv* oder *Alternative*, sobald nur von zwei Merkmalen ‘0’ oder ’1’ die Rede ist.“

4. Man kann sich fragen: Gibt es Kollektive, die der Definition durch v. Mises genügen? Diese Frage muß zunächst von einer anderen unterschieden werden: Kann bei einer vorliegenden Menge von Situationen (Merkmalswerten) entschieden werden, ob sie als Teilmenge oder Teilfolge eines Kollektivs aufgefaßt werden kann? Denn diese Frage betrifft nicht die Begriffsbildungen, sondern Fragen ihrer Anwendbarkeit. Jede vorliegende Menge von Merkmalswerten ist endlich, und die Überlegungen, auf denen v. Mises’ Kollektivbegriff beruht, können schon deshalb nicht angewendet werden. Hinzu kommt, daß eine Menge von Merkmalswerten im allgemeinen nicht als eine Folge aufgefaßt werden kann. Der Begriff eines Kollektivs setzt dagegen voraus, daß man von einer *Folge* sprechen kann; und dies setzt wiederum voraus – sobald man eine platonistisch konzipierte Welt der Mathematik verläßt und sich auf Vorgänge in der Erfahrungswelt beziehen möchte –, daß man sich in irgendeiner Weise auf ein Verfahren beziehen kann, durch das Merkmalswerte (Beobachtungen) sequentiell zustande kommen.

5. Die Frage, ob es Kollektive im Sinne v. Mises’ gibt, hat zu einer langen und bis heute nicht abgeschlossenen Diskussion geführt.[22] Wir werden auf diese Diskussion hier nicht näher eingehen, da sie rein mathematischer Natur ist. In ihrem Mittelpunkt steht die Frage, ob sich mathematische Folgen konstruieren lassen, die – in einer geeignet abgeschwächten Weise – den von v. Mises formulierten Regellosigkeitsbedingungen genügen. Die Fragestellung ist sicherlich sinnvoll, wenn man sie folgendermaßen versteht: Kann man ein mathematisches Modell für Zufallsgeneratoren finden? Die Frage ist wichtig, wenn man nach neuen und vielleicht besseren Methoden sucht, um Zufallsgeneratoren zu konstruieren. Es ist aber, wie ich glaube, nicht richtig, daß den Ergebnissen dieser Diskussion eine wichtige, oder sogar entscheidende Bedeutung für die Frage zukommt, ob eine frequentistische Deutung von Wahrscheinlichkeitsaussagen zu einem sinnvollen Verständnis von Anwendungen der Wahrscheinlichkeitsrechnung führen kann.

[22]Eine Darstellung der Diskussion findet sich bei Martin-Löf, 1969.

6.4 Sinngrenzen für Anwendungen

1. Blicken wir kurz zurück. Die Grundidee des frequentistischen Ansatzes besteht in der Annahme, daß man empirisch ermittelbare relative Häufigkeiten so auffassen könne, daß sich in ihnen näherungsweise gewisse „Konstanten“ (F-Wahrscheinlichkeiten) manifestieren; und daß diese „Konstanten“ als Eigenschaften irgendwelcher, aber nicht explizit gemachter Prozesse gedeutet werden können, die die Situationen hervorbringen, auf die sich die relativen Häufigkeiten beziehen. Dies ist natürlich eine metaphysische Annahme, für die es eine Begründung geben sollte. Aber der Versuch, eine solche Begründung dadurch zu geben, daß man F-Wahrscheinlichkeiten als Grenzwerte relativer Häufigkeiten auffassen könne, gelingt nicht; und zwar weder im Sinne einer empirisch gemeinten Definition noch als eine theoretische Konstruktion.[23] Für die Annahme, daß empirisch ermittelbaren relativen Häufigkeiten gewisse „Konstanten“ (F-Wahrscheinlichkeiten) zugrunde liegen, kann durch den frequentistischen Ansatz keine Begründung gefunden werden. Wenn überhaupt eine Begründung gefunden werden kann, muß man sich auf die Verfahren oder Prozesse beziehen, durch die die Situationen entstehen, aus denen die relativen Häufigkeiten ermittelt werden. Einen nachvollziehbaren Gedankengang liefert der „klassische“ Ansatz bei Zufallsgeneratoren; die „Konstanten“ können dann als Eigenschaften eines Zufallsgenerators verstanden werden.

2. Die Frage, ob empirisch ermittelbare relative Häufigkeiten als zufällige Abweichungen von ihnen voraussetzbaren „Konstanten“ aufgefaßt werden können, hängt eng mit Fragen nach Voraussetzungen und Sinngrenzen für Anwendungen der Wahrscheinlichkeitsrechnung zusammen. Folgt man dem „klassischen“ Ansatz, stellt sich die Frage so: Unter wel-

[23]Überlegungen zur Kritik des frequentistischen Ansatzes finden sich bei vielen Autoren. Auf Braithwaite (1955) wurde schon verwiesen; man vgl. auch Fry 1928, S. 88ff, Stegmüller 1973b, S. 32ff.

chen Bedingungen kann sinnvoll angenommen werden, daß Situationen (Ereignisse, Sachverhalte) durch einen Zufallsgenerator erzeugt werden. Genau dann, wenn dies der Fall ist, können Begriffsbildungen der Wahrscheinlichkeitsrechnung verwendet werden. Sinnvolle Anwendungen gibt es infolgedessen jedenfalls dann, wenn man sich auf artifizielle Zufallsgeneratoren beziehen kann. Es bleibt nur die Frage, ob und unter welchen Bedingungen es auch sinnvoll sein könnte, sich nicht artifiziell erzeugte Phänomene in unserer Erfahrungswelt als durch hypothetische Zufallsgeneratoren erzeugt vorzustellen, um Begriffsbildungen der Wahrscheinlichkeitsrechnung anwendbar zu machen.

3. Folgt man der frequentistischen Auffassung, stellt sich die Frage jedoch anders, etwa so: Unter welchen Bedingungen können relative Häufigkeiten als zufällige Abweichungen von ihnen unterstellbaren F-Wahrscheinlichkeiten aufgefaßt werden? Dafür gibt es allerdings keine objektivierbaren Kriterien, wenn man – wie es die frequentistische Auffassung impliziert – eine explizite Bezugnahme auf Zufallsgeneratoren für nicht erforderlich hält. Übrig bleibt dann nur die Frage, ob man relative Häufigkeiten, die sich ja in allen möglichen endlichen Gesamtheiten feststellen lassen, als „Hinweise" oder „Ausdruck" von F-Wahrscheinlichkeiten *deuten* möchte. Begriffsbildungen der Wahrscheinlichkeitsrechnung werden dann zu einem Deutungsschema für empirisch ermittelbare relative Häufigkeiten.[24]

4. Der frequentistische Versuch, F-Wahrscheinlichkeiten nicht durch eine Bezugnahme auf Zufallsgeneratoren einzuführen, sondern sich stattdessen auf Kollektive zu beziehen, macht noch eine weitere Sinnvoraussetzung der Wahrscheinlichkeitsrechnung unklar: daß F-Wahrscheinlichkeiten (genau wie aleatorische Wahrscheinlichkeiten) nur dann sinnvoll angenommen werden können, wenn man sich auf ein Verfahren (oder einen „Mechanismus") beziehen kann, der beliebig viele Situationen des gleichen Typs *unter gleichen Bedingungen* erzeugen kann. Wie schon erwähnt worden ist, hat auch v. Mises diese Sinnvoraussetzung betont: „Wesentlich ist, daß der Einzelvorgang [...] sich nach einer *dauernd gleichbleibenden Vorschrift* vollzieht." (v. Mises 1931, S. 9) Diese Sinnvoraussetzung wird jedoch dadurch verwischt, daß sich die frequentistischen Begriffsbildungen schließlich nicht auf ein Erzeugungsverfahren für Situationen, sondern auf Kollektive als Mengen tatsächlich realisierter Situationen beziehen. Infolgedessen entsteht zunächst ein theoretisches Dilemma. Einerseits benötigt man Kollektive als Mengen *realisierter* Situationen, um sinnvoll von relativen Häufigkeiten sprechen zu können. Andererseits liefern die Situationen, aus denen ein Kollektiv besteht (in der Sprechweise v. Mises' die „Beobachtungen"), keinerlei Aufschluß über das Verfahren oder den Prozeß, durch den sie zustandegekommen sind. Insbesondere weiß man nicht, ob sie nach einer „dauernd gleichbleibenden Vorschrift" entstanden sind. Im „klassischen" Ansatz wird dieses Problem dadurch gelöst, daß man die Begriffsbildungen auf das Verfahren bezieht, auf den Zufallsgenerator, durch den nach einer „dauernd gleichbleibenden Vorschrift" Situationen entstehen oder erzeugt werden können. Wenn man stattdessen von Kollektiven ausgeht und darunter – wie in der Statistik – irgendwie zusammengefaßte Mengen von Situationen oder Objekten versteht, wird ganz unklar, wie festgestellt werden könnte, ob sie tatsächlich nach einer „dauernd gleichbleibenden Vorschrift" zustandegekommen sind.

5. Dies Problem stellt sich insbesondere bei der Verwendung von Begriffen der Wahrscheinlichkeitsrechnung zur Deutung sozialstatistischer Daten. In einem Zitat, das schon angeführt worden ist, spricht v. Mises davon, daß sich Begriffsbildungen der Wahrscheinlichkeitsrechnung sowohl auf „Massenerscheinungen" als auch auf „Wiederholungsvorgänge" beziehen können. Aber die „Massenerscheinungen", die den Gegenstand der Sozialstatistik bilden, entstehen keineswegs nach einer „dauernd gleichbleibenden Vorschrift", sondern ihre Entstehungsbedingungen unterliegen einem historischen Wandel. Dieses Problem stellt sich natürlich gleichermaßen, wenn man im Anschluß an den „klassischen" Ansatz versucht, die Situationen, auf die sich sozialstatistische Daten beziehen, als durch Zufallsgeneratoren erzeugt aufzufassen. Dann wird jedoch sichtbar, daß hier wirklich ein Problem vorliegt. Man muß einen *hy-*

[24] Ich verwende absichtlich diese Formulierung, weil Anhänger einer frequentistischen Position, insbesondere auch v. Mises, mit dem Anspruch auftreten, einen „objektiven" Wahrscheinlichkeitsbegriff begründen zu können, mit dem sich Aussagen über „objektive Eigenschaften" unserer Erfahrungswelt machen lassen. Sie kritisieren dementsprechend „subjektivistische" Auffassungen, die von einem epistemischen Wahrscheinlichkeitsbegriff ausgehen. Tatsächlich begründet die frequentistische Auffassung jedoch keine Alternative, sondern macht die Annahme von F-Wahrscheinlichkeiten beliebig. Zur Illustration sei auf folgende Ausführungen von Whittle (1992, S. 4) verwiesen: „Suppose one conducted two separate sampling surveys, measuring in each the cholesterol levels of n people, say. Then, if pains have been taken to sample representatively, the average cholesterol levels from the two samples will be found to agree very much more closely than would the cholesterol levels of two randomly chosen individuals. Furthermore, the larger n, the better the agreement. There is a feeling, then, that the sample average measures 'something real' and approaches this 'something real' as the sample size increases. This is, of course, the justification of opinion polls and market surveys. It is on this feature of empirical convergence that one founds probability theory; by postulating the existence of an idealized 'long-run proportion' (a *probability*) or 'long-run average' (an *expectation*)."

pothetischen Zufallsgenerator explizit konstruieren und begründen können, daß es Sinn machen könnte, ihn zum Verständnis oder zur Prognose realer Situationen zu verwenden. Das kann indessen nur insoweit gelingen, wie man sich tatsächlich auf Eigenschaften der Prozesse beziehen kann, durch die Situationen *in der sozialen Realität* entstehen; genau so wie auch die Begründung von Wahrscheinlichkeitsaussagen für artifizielle Zufallsgeneratoren nur insoweit gelingen kann, als man über Einsichten in deren Konstruktion verfügt. — Der frequentistische Vorschlag, daß man sich F-Wahrscheinlichkeiten als „Grenzwerte“ oder „Idealisierungen“ empirisch ermittelbarer relativer Häufigkeiten *vorstellen* könne, macht dagegen weder das Problem sichtbar, noch liefert er Ansatzpunkte für eine Lösung.

Literatur

Anderson, O. 1929. Zur Problematik der empirisch-statistischen Konjunkturforschung. Bonn: Schroeder.

Applebaum, D. 1996. Probability and Information. Cambridge: Cambridge University Press.

Aristoteles 1987. Aristoteles' Physik. 1. Halbband, Bücher I – IV. Übers. von H. G. Zekl. Hamburg: Meiner.

Austin, J. L. 1979. Philosophical Papers. 3rd ed. Oxford: Oxford University Press.

Benenson, F. C. 1984. Probability, Objectivity and Evidence. London: Routledge & Kegan Paul.

Bennett, D. J. 1998. Randomness. Cambridge: Harvard University Press.

Bernoulli, J. 1899a. Wahrscheinlichkeitsrechnung (Ars conjectandi). Teil I und II. Dt. Übers. von R. Haussner. Leipzig: W. Engelmann.

Bernoulli, J. 1899b. Wahrscheinlichkeitsrechnung (Ars conjectandi). Teil III und IV. Dt. Übers. von R. Haussner. Leipzig: W. Engelmann.

Boole, G. 1952. Studies in Logic and Probability. La Salle, Ill.: Open Court.

Boole, G. 1997. Selected Manuscripts on Logic and Its Philosophy. Hrsg. von I. Grattan-Guinness, G. Bornet. Basel: Birkhäuser.

Borel, E. 1950. Probabilité et Certitude. Engl. Transl. Probability and Certainty. New York: Walker and Comp. 1963.

Braithwaite, R. B. 1955. Scientific Explanation. A Study of the Function of Theory, Probability and Law in Science. Cambridge: Cambridge University Press.

Bruns, H. 1906. Wahrscheinlichkeitsrechnung und Kollektivmasslehre. Leipzig: Teubner.

Carnap, R. 1962. Logical Foundations of Probability. 2nd ed. Chicago: University of Chicago Press.

Carnap, R., Stegmüller, W. 1959. Induktive Logik und Wahrscheinlichkeit. Wien: Springer-Verlag.

Czuber, E. 1908. Wahrscheinlichkeitsrechnung. Band 1: Wahrscheinlichkeitstheorie, Fehlerausgleichung, Kollektivmasslehre. Leipzig: Teubner.

Czuber, E. 1923. Die philosophischen Grundlagen der Wahrscheinlichkeitsrechnung. Leipzig: Teubner.

d'Alembert, J. 1756. Artikel „Experimentell“. In: Artikel aus der von Diderot und d'Alembert herausgegebenen Enzyklopädie, 433–437. Hrsg. von M. Naumann. Frankfurt: Röderberg 1984.

Daston, L. 1988. Classical Probability in the Enlightenment. Princeton: Princeton University Press.

David, F. N. 1962. Games, Gods and Gambling. London: Charles Griffin.

de Finetti, B. 1972. Probability, Induction and Statistics. New York: Wiley.

De Morgan, A. 1838. An Essay on Probabilities and On Their Application to Life Contingencies and Insurance Offices. London: Longman.

Doob, J. L. 1953. Stochastic Processes. New York: Wiley.

Ellis, R. L. 1849. On the Foundations of the Theory of Probabilities. Transactions of the Cambridge Philosophical Society 8, 1–6.

Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G. 1999. Statistik. Der Weg zur Datenanalyse. Berlin: Springer-Verlag.

Fisher, R. A. 1973. Statistical Methods and Scientific Inference. Third Ed. New York: Hafner Press.

Freudenthal, H. 1963. Wahrscheinlichkeit und Statistik. München: Oldenbourg.

Fry, T. C. 1928. Probability and Its Engineering Uses. New York: van Nostrand.

Galton, A. 1984. The Logic of Aspect. Oxford: Clarendon.

Gnedenko, B. W. 1991. Einführung in die Wahrscheinlichkeitstheorie. Berlin: Akademie-Verlag.

Gottinger, H. W. 1980. Elements of Statistical Analysis. Berlin: de Gruyter.

Hacker, P. M. S. 1982. Events and Objects in Space and Time. Mind 91, 1–19.

Hacking, I. 1965. Logic of Statistical Inference. Cambridge: Cambridge University Press.

Hacking, I. 1971. Equipossibility Theories of Probability. British Journal for the Philosophy of Science 22, 339–355.

Halmos, P. R. 1972. Naive Mengenlehre (3. Auflage). Göttingen: Vandenhoeck & Ruprecht.

Ineichen, R. 1996. Würfel und Wahrscheinlichkeit. Stochastisches Denken in der Antike. Heidelberg: Spektrum.

Jeffreys, H. 1948. Theory of Probability. 2nd ed. Oxford: Clarendon Press.

Kant, I. 1800. Logik. In: Werke in 10 Bänden, hrsg. von W. Weischedel. Band 5, 421–582. Darmstadt: Wissenschaftliche Buchgesellschaft 1968.

Karup, W. 1885. Handbuch der Lebensversicherung. Leipzig: Fritsch.

Kendall, M. G. 1949. On the Reconciliation of Theories of Probability. Biometrika 36, 101–116.

Keynes, J. M. 1921. A Treatise on Probability. London: Macmillan.

Knuth, D. E. 1981. The Art of Computer Programming, Vol. II: Seminumerical Algorithms. Reading: Addison-Wesley.

Kolmogoroff, A. 1933. Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlin: Springer-Verlag.

Kolmogorov, A. N. 1963. On Tables of Random Numbers. Sankhya A 25, 369–376.

Koopman, B. O. 1940. The Axioms and Algebra of Intuitive Probability. Annals of Mathematics 41, 269–292.

Kries, J. von 1886. Die Principien der Wahrscheinlichkeitsrechnung. Freiburg: Mohr.

Laplace, P. S. de 1814. Philosophischer Versuch über die Wahrscheinlichkeit. Dt. Übers. hrsg. von R. v. Mises. Leipzig: Akademische Verlagsgesellschaft 1932.

Lenzen, W. 1980. Glauben, Wissen und Wahrscheinlichkeit. Systeme der epistemischen Logik. Wien: Springer-Verlag.

Lorenzen, P. 1974. Konstruktive Wissenschaftstheorie. Frankfurt: Suhrkamp.

Lorenzen, P. 1978. Konstruktive Deutung des Dualismus in der Wahrscheinlichkeitstheorie. Zeitschrift für allgemeine Wissenschaftstheorie 9, 256–275.

Lucas, J. R. 1970. The Concept of Probability. Oxford: Clarendon Press.

Makin, S. 1993. Indifference Arguments. Oxford: Blackwell.

Martin-Löf, P. 1969. The Literature on von Mises' Kollektivs Revisited. Theoria 35, 12–37.

McPherson, G. 1990. Statistics in Scientific Investigation. Its Basis, Application, and Interpretation. New York: Springer-Verlag.

Meinong, A. 1915. Über Möglichkeit und Wahrscheinlichkeit. Nachdruck in: Gesammelte Werke, Bd. 6, hrsg. von R. Haller, R. Kindinger, R. M. Chisholm. Graz: Akademische Druck- und Verlagsanstalt 1972.

Meyer, A. 1879. Vorlesungen über Wahrscheinlichkeitsrechnung. Dt. Übers. von E. Czuber. Leipzig: Teubner.

Mises, R. v. 1931. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Leipzig-Wien: Deuticke.

Mises, R. v. 1933. Über Zahlenfolgen, die ein kollektiv-ähnliches Verhalten zeigen. Mathematische Annalen 108, 757–772.

Mises, R. v. 1951. Positivism. A Study in Human Understanding. Transl. from first German edition, 1939. Cambridge: Harvard University Press. Neue dt. Ausgabe: Kleines Lehrbuch des Positivismus. Frankfurt: Suhrkamp 1990.

Mises, R. von 1981. Probability, Statistics and Truth. Reprint of 2nd ed. 1957. New York: Dover.

Moivre, A. de 1756. The Doctrine of Chances: or, A Method of Calculating the Probabilities of Events in Play (3rd ed). London: A. Millar.

Moore, D. S. 1991. Statistics. Concepts and Controversies (3rd ed). New York: Freeman & Comp.

Nagel, E. 1939. Principles of the Theory of Probability. International Encyclopedia of Unified Science, Vol. I, No. 6. Chicago: University of Chicago Press.

Neyman, J. 1967. A Selection of Early Statistical Papers. Berkeley: University of California Press.

Oaklander, L. N., Smith, Q. (eds.) 1994. The New Theory of Time. New Haven: Yale University Press.

Ore, O. 1960. Pascal and the Invention of Probability Theory. American Mathematical Monthly 67, 409–419.

Rabinovitch, N. L. 1973. Probability and Statistical Inference in Ancient and Medieval Jewish Literature. Toronto: University of Toronto Press.

Ramsey, F. P. 1926. Truth and Probability. In: ders., The Foundations of Mathematics and other Logical Essays, 156–198. London: Routledge & Kegan Paul 1931.

Ramsey, F. P. 1929. Probability and Partial Belief. In: ders., The Foundations of Mathematics and other Logical Essays, 256–257. London: Routledge & Kegan Paul 1931.

Reichenbach, H. 1945. Reply to Donald C. Williams' Criticism of the Frequency Theorie of Probability. Philosophy and Phenomenological Research 5, 508–512.

Reichenbach, H. 1949. Wahrscheinlichkeitslehre. Reprint in: Gesammelte Werke Bd. 7, hrsg. von A. Kamlah, M. Reichenbach. Braunschweig: Vieweg 1994.

Rescher, N. 1996. Process Metaphysics. An Introduction to Process Philosophy. New York: State University of New York Press.

Robertson, J. M. 1905. Letters on Reasoning, 2nd ed. London: Watts & Co.

Schneider, I. (ed.) 1988: Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933. Einführungen und Texte. Darmstadt: Wissenschaftliche Buchgesellschaft.

Sixtl, F. 1993. Der Mythos des Mittelwerts. Neue Methodenlehre der Statistik. München: Oldenbourg.

Smart, J. J. C. 1963. Philosophy and Scientific Realism. London: Routledge & Kegan Paul.

Stadler, F. 1990. Richard von Mises (1883–1953) – Wissenschaft im Exil. In: R. von Mises, Kleines Lehrbuch des Positivismus, 7 –51. Frankfurt: Suhrkamp.

Stegmüller, W. 1973a. Personelle und statistische Wahrscheinlichkeit. 1. Halbband: Personelle Wahrscheinlichkeit und Rationale Entscheidung. Berlin: Springer-Verlag.

Stegmüller, W. 1973b. Personelle und statistische Wahrscheinlichkeit. 2. Halbband: Statistisches Schließen, statistische Begründung, statistische Analyse. Berlin: Springer-Verlag.

Sterzinger, O. 1911. Zur Logik und Naturphilosophie der Wahrscheinlichkeitslehre. Leipzig: Xenien-Verlag.

Struik, D. J. 1934. On the Foundations of the Theory of Probability. Philosophy of Science 1, 50–70.

Todhunter, I. 1865. A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace. Reprint New York: Chelsea Publ. 1949.

Tschuprow, A. A. 1925. Grundbegriffe und Grundprobleme der Korrelationstheorie. Leipzig.

Venn, J. 1888. The Logic of Chance. An Essay on the Foundations and Province of the Theory of Probability. 3rd ed. London: Macmillan.

Wærden, B. L. van der 1975. Historische Einleitung. In: Jakob Bernoulli, Werke, Band III, 1–18. Basel: Birkhäuser.

White, A. R. 1975. Modal Thinking. Oxford: Basil Blackwell.

Whittle, P. 1992. Probability via Expectation. New York: Springer-Verlag.