Rettet die Wahrscheinlichkeit – in einer nicht zu sehr mathematisierten Form

Manfred BOROVCNIK, Univ. Klagenfurt

Kurzfassung

Es gibt viele Gründe, die Wahrscheinlichkeitsrechnung zu kürzen oder zu streichen, was auch tatsächlich, so sie überhaupt einmal durchgängig unterrichtet worden ist, im Schulalltag passiert. Im Artikel sollen Hinweise gegeben werden, warum aus begrifflichen Gründen solide Kenntnisse aus der Wahrscheinlichkeitsrechnung nicht verzichtbar sind - nicht nur im Hinblick auf ein Verständnis der Beurteilenden Statistik, sondern auch im Hinblick auf ein geeignetes Modellverständnis von Wahrscheinlichkeit und wichtigen Anwendungen. Ferner sollen einige Möglichkeiten aufgezeigt werden, wie Wahrscheinlichkeit besser (hinsichtlich der bestehenden Welt der Intuitionen) und mathematisch einfacher unterrichtet werden kann.

0. Einige Vorbemerkungen

a) Statistische Methoden ohne Wahrscheinlichkeit verstehen

Man kennt den Autor u.a. auch wegen seiner Bemühungen, statistische Bildung und Ausbildung zu fördern. Damit ist insbesondere auch ein "Direkter Zugang zur Beurteilenden Statistik" gemeint, der weitgehend ohne Ausbau der Wahrscheinlichkeitsrechnung erfolgen kann, siehe Borovcnik (1985 und 1992). Ein empirischer Zugang zu Regression und Korrelation in Borovcnik (1989) ist entstanden aus der Not, Biologen als Anwendern, die keine Ahnung von Wahrscheinlichkeit haben (und haben wollen), wichtige statistische Schließweisen nahezubringen.

b) Statistik erleben

In jüngerer Vergangenheit hat der Autor sich auf Schulexperimente im Rahmen des Arbeitskreises "Statistik in Bildung und Ausbildung" der Österreichischen Statistischen Gesellschaft eingelassen, die durch die Leitlinien "einfachste Methoden" und "Anwendungsorientierung", also durch minimalen Bezug zur Wahrscheinlichkeitsrechnung geprägt sind. Mehr dazu findet man in Borovcnik (1999) und auch im Internet unter http://www.osg.or.at; zu den jüngsten Bestrebungen zählen der sogenannte Volksschulkoffer sowie der Posterwettbewerb an allen Schulen.

Es fällt aber auf, daß alle diese Bestrebungen nicht Wahrscheinlichkeit als Begrifflichkeit sondern lediglich den Ausbau des Kalküls zur Wahrscheinlichkeit, insbesondere in hochstilisierter, d.h. mathematisierter Form vermeiden.

c) Stochastisches Denken

Die Beschäftigung mit "Fundamentalen Ideen" eines Faches ist seit dem Irrweg mit der Mengenlehre als alleinigem Unterbau für einen strengen logischen Aufbau der Mathematik in Verruf geraten. Jüngst ist der Ansatz immer mehr hinter der Subjektivität und Personalität der individuellen Konstruktion aller Begriffe verschwunden. Es "gibt", so eine Hauptströmung in der didaktischen Diskussion, keine allgemeinen Begriffe mehr, nur mehr persönliche; wie dabei eine Mathematik immer noch den Anspruch auf Anwendung aufrecht erhalten soll und wie sie immer noch angewendet werden kann und wie sie vielleicht doch noch von wenigen in ihren Auswirkungen (noch immer) im vornhinein verstanden werden kann, wie sie unterrichtet und gar verstanden werden kann, all das bleibt ein Rätsel. Es muß also vielleicht doch irgendwas wie einen allgemeinen Kern eines Begriffes geben, der nicht beliebig subjektiv rekonstruierbar ist.

Oder ist der Ansatz der Fundamentalen Ideen einfach an den Alltagssorgen der Beschaffung des nötigen und sich ständig ändernden know hows im Computerbereich vergessen worden. Für den Autor bleibt es - trotz Mängel, welcher Ansatz hätte keine - einer der vielversprechendsten. Eine Analyse der großen, fundamentalen Ideen der Wahrscheinlichkeit (siehe Borovcnik 1997) zeigt auch, worin diese Ideen, ohne Bindung an einen Kalkül, bestehen und wie bedeutsam sie für ein Verständnis des Wahrscheinlichkeitsbegriffs einschließlich aller Anwendungen statistischer Methoden sind. Die Analyse zeigt auch Ansätze auf, wie man diese fundamentalen Ideen ohne allzu mathematisierten Ausbau des Kalküls lehren kann.

Als Fazit dieser Analysen ergibt sich: Wahrscheinlichkeit in seinen ideellen Ausprägungen ist nicht nur für das Verständnis statistischer Anwendungen unverzichtbar.

d) Einige allgemeine Gründe für den Abstieg der Wahrscheinlichkeitsrechnung

Die Theorie der Wahrscheinlichkeit ist relativ formal und scheint wenig mit den intuitiven Vorstellungen von Personen über den Zufall zu tun zu haben, wie empirische Interviews immer wieder belegen, siehe etwa Traar (1989). Üblicher Unterricht war daher relativ erfolglos im Aufbau eines gesunden Verständnisses auch schon der Grundbegriffe.

Der Druck der Anwendungen kommt scheinbar aus der Ecke der Datenanalyse. Dort kommt man lange mit einfachen deskriptiven Methoden aus; darüber hinaus werden die Methoden der statistischen Beurteilung mit den modernen Rechenhilfsmitteln leicht zugänglich, was nicht heißt, daß sie im Anwendungsfall auch wirklich verstanden werden.

Der Trend, die Modellierung im Mathematik-Unterricht an sich als wichtig zu bewerten (was durchaus zielführend und richtig ist), führt weiters dazu, daß man sich im Unterricht auf einfache Begriffe zurück zieht, damit man ein geeignetes Verständnis der Modellierungsaspekte nicht durch die Komplexität der Mathematik behindert.

Die Veränderung der Stellung der allgemeinbildenden und berufsbildenden höheren Schulen weg von Eliteschulen und hin zur „verbindlichen" Ausbildung für alle führt zu einer Veränderung der Stellung der Mathematik im Curriculum und zu einer Änderung der Haltung zur schulischen

Leistung; das führt zu einer Elementarisierung des Mathematik-Unterrichts.

Neue Medien machen Teile der komplexen Mathematik ohnehin obsolet, komplizierte Rechenmethoden werden durch einen Knopfdruck ersetzt, der wenig mit den Eigenschaften der Begriffe zu tun hat. Man spart Zeit und Aufwand, es verändert / behindert aber auch teilweise begriffliches Verständnis. Für solche Trends findet man weitere Details in Borovcnik (1994 und 1996).

1. Die Rolle von Wahrscheinlichkeit beim Unterrichten

Wahrscheinlichkeit als Begriff blieb in seinen Aspekten und seiner Bedeutung nicht nur von so vielen Anwendern sondern auch in der didaktischen Diskussion unverstanden. Das hatte zur Folge, daß im "gelebten" Lehrplan Wahrscheinlichkeit weitgehend fehlt und von der Statistik, so überhaupt, triviale Methoden an (durchaus wichtigen) realen Beispielen ausgewalzt werden. Daß Wahrscheinlichkeit so weitgehend unverstanden blieb, heißt aber bei weitem nicht, daß der Begriff nicht zugänglich ist. Der bis dato fehlgeschlagene Versuch soll also nicht zu voreiligen Konsequenzen führen.

Leicht wird es nicht und durch eine Beschäftigung nebenbei, wie der Autor immer wieder erlebt hat, wird man nichts erreichen; man muß sich schon ein wenig mehr mit den schillernden Aspekten des Begriffs einlassen. Die Mathematik als Hürde möchte der Autor allerdings mildern; daher lautet auch der Titel dieses Aufsatzes "Rettet die Wahrscheinlichkeit - in einer nicht zu sehr mathematisierten Form".

a) Theoretische Positionen zum Unterricht von Wahrscheinlichkeit

Dazu kann der geneigte Leser, die geneigte Leserin sich in Borovcnik und Peard (1996) ausführlichst kundig machen. Hier seien nur ganz wenige wesentliche Aspekte zusammengefaßt:

Wahrscheinlichkeit einschließlich der Wahrscheinlichkeitsrechnung kann *als zentraler Kern* der Stochastik aufgefaßt werden. Danach hat man begrifflich sauber die Wahrscheinlichkeitstheorie zu entwickeln und erst dann philosophisch - logisch (logisch allein geht nicht) den statistischen Schluß von einer Stichprobe auf die fiktive, zugrunde liegende Population zu begründen und die dafür geeigneten Methoden aufzubauen. Alle einschlägigen Methoden haben extreme Schwächen und Einschränkungen, die man aber erst verstehen kann, wenn man genug von Wahrscheinlichkeit und dem Kalkül dazu beherrscht. Diese Position verstrickt sich allzubald in den bekannten mathematischen Schwierigkeiten mit dem Kalkül, egal ob man den kombinatorischen oder den axiomatischen (oder gar den maßtheoretischen) Weg wählt.

Als Gegenpol dazu hat man eine didaktische Position entwickelt, die *Wahrscheinlichkeit* als absolut *dienend für* ein geeignetes *Verständnis statistischer Methoden* ansieht. Immer noch wird ein umfassendes Verständnis von Wahrscheinlichkeit angestrebt, aber es wird unterstellt, daß man ohne statistische Fragestellungen Wahrscheinlichkeit gar nicht verstehen kann. Also: Wahrscheinlichkeit und der statistische Schluß sind wechselweise begrifflich voneinander abhängig. Aus "taktischen" Gründen hat man jedoch den dienenden Charakter überbetonen wollen und auf den Ausbau des Wahrscheinlichkeitskalküls verzichtet. Als Ansätze hierzu sind zu nennen: (i) der naive Zugang ohne Wahrscheinlichkeit, (ii) der nicht-parametrische Ansatz sowie (iii) der Resampling-Ansatz. (Die beiden letzteren waren aus Gründen der besseren Modellierung sowie der erweiterten Anwendung aus dem Fachgebiet erstanden.)

Über den Fragen, *wie* die Begriffe eigentlich zu verstehen sind, hat man jedoch einen ganz wichtigen Aspekt des *wozu* diese Begriffe sind, übersehen. Wahrscheinlichkeit ist als Werkzeug nicht nur geeignet, diverse Methoden für den statistischen Schluß zu begründen. *Wahrscheinlichkeit* ist in einem von Unsicherheit geprägtem Leben des Menschen ein grundlegendes *Werkzeug, über "Realität"*, was immer darunter zu verstehen ist, *nachzudenken* und seine Position in der Lösung des anstehenden Problems zu verbessern. Das betrifft keineswegs nur Fälle der Anwendung eines statistischen Schlusses, nein, es beginnt bei den - auch im Alltag - kaum wegzudenkenden Begriffen

(i) des Risikos, insbesondere des Restrisikos und

(ii) des Begriffes der Zuverlässigkeit, der Sicherheit, daß "irgend etwas" denn doch funktioniert, wofür es vorgesehen ist. Beim Begriff der Zuverlässigkeit kommen auch mittlere, "zu erwartende" Lebensdauern ins Spiel. Hier liegt man schon sehr nahe zum

(iii) Begriff der Wartezeit, der nicht nur in formalen Warteschlangen-Systemen eine wesentliche Rolle spielt.

Dieser Werkzeug-Charakter zur Modellierung, ja zur Konstruktion einer künstlichen Welt, ist wesentlicher Bestandteil von Wahrscheinlichkeit und macht seine Unterweisung, auch unabhängig von Beurteilender Statistik, erforderlich.

b) Die Rolle von Wahrscheinlichkeit in wirklichem Unterricht

Eine internationale Untersuchung von Nemetz und Borovcnik (in Nemetz, 1997) bestätigt leider, was auch schon so informell bekannt war, nämlich daß Wahrscheinlichkeit in allen Curricula für den Sekundarbereich zurückgeht.

Als Gründe geben Didaktiker und Curriculum-Designer weltweit an, daß Wahrscheinlichkeit

(i) zu sehr mathematisch ausgerichtet sei,

(ii) zu eng mit Spielen, insbesondere Glücksspielen, verbunden sei,

(iii) sowieso nur benötigt werde, um Methoden der statistischen Beurteilung zu begründen.

Zu (i): Die obigen Ausführungen deuten hier ein Mißverständnis von gängiger Unterrichtspraxis und Wesenszug des Begriffs an sich an.

Zu (ii): Hier zeigt sich ein öffentlich fortschreitender Puritanismus, wonach zwar immer mehr Leute sich im privaten Bereich mit Glücksspielen beschäftigen, eine sachliche, fachliche "Aufklärung" darüber aber nicht mehr im öffentlichen, sprich schulischen Interesse steht, zumal das Glücksspiel anrüchig ist und daher aus moralischen Gründen ohnehin zu unterlassen sei. Dieser Zynismus hindert den Staat jedoch nicht, über öffentliche Lotterien, das Lotto oder die Casinos gewaltig zu verdienen. Außerdem wird eine mögliche Begriffsillustration mit Reduktion des Begriffsspektrums verwechselt und man beraubt sich damit unnötigerweise einer hervorragenden Illustration der Begriffe, die sich ja wirklich aus dem Glücksspiel und dem Versicherungsgeschäft heraus entwickelt und so ihre Eigenart erhalten haben.

Zu (iii): Hier zeigt sich das Versäumnis, den Werkzeug-Charakter von Wahrscheinlichkeit an sich zur Erforschung und Strukturierung von "Realität" zu erkennen. Trotz der Allgegenwart des Begriffs des "Restrisikos", besonders in der politischen Diskussion, mußte der Autor in die völlig überraschten Augen von Didaktikern auf internationalen Tagungen blicken, wenn er auf diese grundlegend stochastischen Denkweisen und Anwendungen hinwies.

Wahrscheinlichkeit wird in dieser Studie unter Didaktikern als viel zu eng und zu mathematisch ausgerichtet beschrieben, weshalb man die großen Schwierigkeiten, die Lernende damit haben, gar nicht rechtfertigen könne. Und dann darf man sich nicht wundern, wenn Wahrscheinlichkeit im praktischen Unterricht auf der ständigen Absenzenliste geführt wird.

c) Die Flucht in die EDA

Der Autor selbst hat diesen jungen Zweig der Angewandten Statistik schon früh für unterrichtliche Zwecke genutzt (siehe Borovcnik und Ossimitz 1987). Noch immer ist er der Überzeugung, daß hier ein didaktisch reizvolles Gebiet vorliegt: die Ausrichtung an wirklichen Beispielen aus dem Alltag und die sachliche Auseinandersetzung mit der Fragestellung, erleichtert durch einfachste Techniken und Begriffe, deren Sinn sofort einzusehen ist, gibt ein märchenhaftes Feld für den Mathematisierungsgedanken im Unterricht ab.

Wie dominant im internationalen Trend die EDA alle anderen Zweige der Stochastik überrollt (siehe die neuen Curricula in den USA die data handling von der Grundschule aus bis zur Universität als alleinige Leitlinie verfolgen) kann der Autor nur als Flucht in die EDA bezeichnen. Während die Meriten dieses Ansatzes durchaus zu würdigen sind, darf man jedoch nicht übersehen, daß damit eine vollständige Umgehung der statistischen Beurteilung sowie von stochastischem Modellieren und Stochastischem Denken verbunden ist. Also: EDA ist attraktiv, insbesondere für den Mathematisierungsaspekt, geht jedoch am Zufall und seinen Phänomenen vorbei, erfaßt insbesondere den Begriff des Restrisikos überhaupt nicht.

d) Gründe für eine "starke Wahrscheinlichkeit"

In der Diskussion der Rolle von Wahrscheinlichkeit wurden folgende Punkte genannt:

(i) Wahrscheinlichkeit ist unverzichtbar für ein Verständnis statistischer Beurteilung.

(ii) Wahrscheinlichkeit bietet einen Typ von Werkzeug zum Modellieren und Erschaffen von Realität.

(iii) Wahrscheinlichkeit bietet einen Typ von Denkweise an, wie man über Realität nachdenken kann.

Daraus ergibt sich, daß auf Wahrscheinlichkeit zu verzichten eine wesentliche Einschränkung intellektueller Möglichkeiten bedeutet, woraus der Autor für die Schule die Aufforderung "Rettet die Wahrscheinlichkeit als Denkform" ableitet.

2. Wahrscheinlichkeit zum Verallgemeinern partieller Information

Die Situation im statistischen Schluß ist dadurch gekennzeichnet, daß man über endliche Grundgesamtheiten nur teilweise Information im Wege einer Teilmenge, genannt Stichprobe, hat; oder daß man einen Prozeß des Entstehens von Ereignissen (ist gleich Grundgesamtheit) nur eine Zeitlang beobachten kann (Daten bilden die Stichprobe). In beiden Fällen hat man die unvollständige Information von der Stichprobe auf die größere Einheit der Grundgesamtheit zu verallgemeinern.

Die klassische statistische Inferenz bietet durch einen szenario-ähnlichen Rahmen - was könnte denn eigentlich passieren - die Möglichkeit der Bewertung, ob die beobachtete Stichprobe mit der Grundgesamtheit vereinbar ist oder nicht. Die daraus abgeleitete Entscheidung birgt immer ein sogenanntes Restrisiko für Fehler in sich.

Der EDA-Ansatz versucht eine Zerlegung der Daten in ein allgemeines Muster und individuelle Abweichungen. Wären die Daten aus einer zufälligen Stichprobe, so könnte man das erkannte "Muster" einer klassischen statistischen Analyse unterziehen. Das Hauptaugenmerk liegt jedoch in einer durch Kenntnis über den Sachzusammenhang geleiteten Interpretation der Abweichungen, wodurch man sich weitere, neue Einsichten erhofft.

Auch der für die Anwendungen so wichtige ANOVA-Ansatz versucht eine Zerlegung der Daten in ein allgemeines Mittel, besondere (durch spezifische Ursachen bedingte) Einflüsse sowie zufällige Abweichungen (Fehler). Hier werden die ursachenbedingten Abweichungen durch den Kontext erklärt bzw. geben über den Kontext Aufschluß (etwa, daß verschiedene Behandlungen unterschiedlichen Erfolg zeitigen). Die Fehler werden einer rein statistischen Untersuchung unterzogen, insbesondere erfolgt die Trennung der Daten in ihre Komponenten nach einer gefinkelten statistischen Prozedur.

Umgeht der EDA-Ansatz Wahrscheinlichkeit gänzlich, so ist bei den beiden anderen Ansätzen Wahrscheinlichkeit der Schlüssel zum Verständnis der darauf aufbauenden Verfahren.

a) Klassische Inferenz

Die klassischen statistischen Methoden von Vertrauensintervall und statistischen Tests basieren auf einem Vergleich der gegenwärtig vorhandenen Stichprobe mit allen möglichen Stichproben.

Dabei werden dem datenerzeugenden Prozeß eine Reihe stochastischer Annahmen auferlegt bis einschließlich seiner Wahrscheinlichkeitsverteilung. Das Merkmal bzw. dessen Größe, die in der Grundgesamtheit geschätzt werden soll, hat dabei keinen stochastischen Charakter. Diese Annahme würde auf Bayesianische Sichtweisen und Methoden führen, auf welche hier nicht extra eingegangen sei. Der Bayesianische Ansatz bietet Vor- und Nachteile (siehe Borovcnik 1986); hier sei nur erwähnt, daß eine Modellierung danach ganz stark in die Richtung "Denken in Wahrscheinlichkeiten" geht (Borovcnik 1992).

Es ergeben sich je nach Fragestellung folgende Problemtypen der Übertragung der partiell vorhandenen Information:

(i) Beurteilen des Gewichts von Ergebnissen - Rückschluß auf die Grundgesamtheit in Form von Schätzen von Wahrscheinlichkeiten

(ii) Voraussagen für Ergebnisse zukünftiger Stichproben - Rückschluß auf die Grundgesamtheit und dann auf weitere Stichproben

(iii) Verallgemeinern der Ergebnisse in der gegenwärtigen Stichprobe auf die zu bestimmenden Größen einer Grundgesamtheit

Bei all diesen im klassischen Rahmen dazu entwickelten "Ritualen" spielt die Hintergrundhypothese einer bestimmten Verteilung auf der Grundgesamtheit sowie die Unabhängigkeit der einzelnen Schritte der Datengewinnung eine wesentliche Rolle; der sogenannte datenerzeugende Prozeß - kurz DEP- muß "zufällig", ausgestattet mit bestimmten (strukturell bekannten) Wahrscheinlichkeiten, sein. Dieser DEP, auch als Hyperpopulation apostrophiert, wird üblicherweise im Unterricht durch Simulation erschlossen. Kennzeichnend für alle abgeleiteten Schlüsse nach allen üblichen Methoden ist ein sogenanntes Restrisiko, daß die getroffene Entscheidung doch falsch ist.

Zwei kurze Beispiele sollen die Begrifflichkeit und die Schwierigkeiten erläutern: Eine Wahlumfrage ergibt bei 1000 Befragten einen Anteil von 0,42 für Kandidat XX. Wir wollen davon ausgehen, daß wirklich eine zufällige Stichprobe vorliegt (was Meinungsforschungsinstitute nicht tun, sie können es besser) und daß keine Meßfehler vorliegen, also etwa daß alle Befragten bereits fest entschlossen sind und ihre Wahlabsicht auch wahrheitsgemäß kund tun (was sie nicht sind bzw. was sie in immer höherem Ausmaß nicht tun). Dann leitet der Statistiker ein Vertrauensintervall von (0,42 +/-0,03) zur "Sicherheit" 0,95 ab. Er kann mit seiner Aussage, der Kandidat wird bei der Wahl unter 45% der Stimmen bleiben, auch falsch liegen, das bleibt immer ein Risiko. Er kann das Risiko verringern, indem er mehr Leute befragt oder das Vertrauensintervall vergrößert (er kann das neue Risiko zahlenmäßig angeben).

Was es mit diesem Risiko auf sich hat, kann man vielleicht an folgendem Simulationsbeispiel einsehen. Zu prüfen ist, ob zwei gewöhnliche Würfel stochastisch unabhängig sind oder nicht. Man könnte so vorgehen: Man würfelt 100 Mal und berechnet aufgrund der Stichprobe einen Korrelationskoeffizienten und liegt dieser über einem sogenannten Schwellenwert (auch kritischer Wert genannt), so verwirft man die Hypothese, die Würfel seien unabhängig. Wie gewinnt man nun diesen Schwellenwert? Ganz einfach: Mit Hilfe einer Computersimulation für zwei

Würfel erhält man z.B. 1000 Mal eine 100er-Serie von Wurfergebnissen und berechnet so 1000 Korrelationswerte, die man in einem Histogramm auswertet. Von dieser empirischen Vergleichsbasis aus bestimmt man einen oberen 5%-Punkt, das ist der gesuchte Schwellenwert. Er wird bei tatsächlicher Unabhängigkeit der Würfel nur in 5% der Fälle überschritten. Klar ist, daß die Entscheidung, die Unabhängigkeit bei Überschreiten des Schwellenwerts abzulehnen, ein Restrisiko in sich birgt.

Über die intuitiven Probleme mit Stichproben gibt es eine Reihe von Ergebnissen aus der psychologischen Literatur sowie aus der empirisch-didaktischen Forschung. Man kann sie nicht dadurch vermeiden, indem man einen möglichst wahrscheinlichkeitsfreien Zugang zur statistischen Beurteilung verfolgt. Folgende Typen, die mit der Standard-Interpretation interferieren oder sie überhaupt in Zweifel ziehen, kann man anführen:

(i) Repräsentativitätsfehlschlüsse und Verfügbarkeitsfehlschlüsse (siehe Kahneman, Slovic und Tversky 1982 oder Walter 1983) - wonach der Vergleich zwischen Stichprobe und Grundgesamtheit nicht quantitativ über eine so fiktive Gesamtheit aller Stichproben geführt wird, sondern qualitativ die "Glaubwürdigkeit" einer Stichprobe durch "Ähnlichkeiten" mit der Grundgesamtheit beurteilt wird. Klar, daß dies nicht nur die Akzeptanz sondern auch das Verständnis klassischer Prozeduren behindert.

(ii) Die "outcome"-Orientierung nach Konold (1989), wonach (auch erst gut eingeschätzte) Wahrscheinlichkeiten durch den inneren Zwang, eine Voraussage treffen zu wollen (bzw. zu müssen) völlig verzerrt werden. Etwa wird die Wahrscheinlichkeit ½ keinen Grund liefern, eines von zwei verschiedenen Ergebnissen vorauszusagen. Obwohl in einer konkreten Situation die Person korrekt schätzt/ bewertet, schaut sie sich nach ganz anderen Möglichkeiten um, die eine "bessere" Vorhersagemöglichkeit bieten. Oder: Eine Wahrscheinlichkeit von 0,7 wird von so manchen als "das Ereignis muß jetzt eintreten" interpretiert - ansonsten wäre die Aussage ja wertlos.

(iii) Die Gründe für ein "Übervertrauen" in bestimmte Prognosen können sehr subjektiv und kaum nachvollziehbar sein (astrologische sowie psychodynamische Aspekte), können völlig dem theoretischen Rahmen widersprechen, können aber sogar dynamische, nicht im Modell erfaßbare Aspekte in Bewegung setzen und sich sogar bestätigen. Personen mit solchen Erfolgserlebnissen wird man von der allgemeinen Gültigkeit ihrer Strategien wohl kaum abbringen können.

b) Der EDA-Ansatz zur Inferenz

In der EDA werden beobachtete Daten in ein allgemeines Muster plus davon abweichende Residuen zerlegt.

$$\textit{Daten} = \textit{Muster} + \textit{Residuen}$$

Während nun das Muster einer statistischen Analyse zugänglich wäre, falls die Daten aus einer zufälligen Stichprobe stammten, ist eine geeignete Analyse der Residuen nur individuell und über Kenntnisse aus dem Kontext möglich. Der Anwendungsbereich der EDA ist viel größer als derjenige der klassischen Inferenz, weil der DEP keinerlei Einschränkungen unterliegt. Allerdings gibt es keine ritualisierten Ergebnisse der Art, wenn ein Schwellenwert überschritten wird, muß man zu dem und dem Schluß (Ablehnung der Nullhypothese) kommen. Es bleibt allein in der Verantwortung und Kenntnis des Sachkontexts des Modelleurs, was als Einsicht beurteilt wird.

Hilfen bekommt man durch gewisse Eigenarten der EDA-Techniken. Etwa werden zur Zusammenfassung des Musters (des Glatten, des Modell-Werts) robuste Techniken verwendet, etwa der Median anstelle des Mittelwerts. Das soll verhindern, daß im Muster irgendwelche individuelle Abweichungen repräsentiert sind; das macht gleichzeitig auch die Residuen besonders groß, wie unter der Lupe, so daß sie nicht übersehen werden können. Zur kontextlichen Interpretation - wie auch zur Trennung überhaupt, die ja i.a. nicht in einem Schritt, sondern interaktiv, erfolgt - wird die Individualität der Daten immer wieder herstellbar. So etwa werden besonders weit außen liegende Daten im Boxplot individuell beschriftet, damit man Gemeinsamkeiten der "Ausreißer" oder trennende Merkmale von dem allgemeinen Muster (hier die zentrale 50%-Box) "erkennen" kann.

Als ganz allgemeine Probleme des Ansatzes können angesprochen werden:

(i) Ist das ein Spielen ohne Ziel? Wie kann man verhindern, daß vorgefaßte Ziele durch geschickte Manipulation verschleiert und vorab gewollte Ergebnisse als Resultat einer Analyse vorgegeben werden?

(ii) Wie kann man ein Ergebnis subjektiver Manipulation anderen als verbindlich vorschreiben?

(iii) Wie geht man didaktisch und praktisch mit der Vielzahl gleichberechtigter "Sichtweisen" statt eines wie immer als optimal eingestuften Ergebnisses um?

(iv) Wie verschafft man sich unterrichtlich sowie praktisch die erforderlichen Kenntnisse aus dem Kontext? Die Rolle des Kontexts zur Steuerung der Analyse sowie zur Interpretation der Ergebnisse wird in didaktischen Zugängen weitgehend und weit unterschätzt oder ausgeblendet.

(v) Schwerwiegendster Einwand jedoch ist, daß der EDA-Ansatz am zufälligen Charakter der Daten völlig vorbei modelliert und diesem auch nicht gerecht werden kann. Da hilft auch nichts, daß er auch auf zufällige Daten angewandt werden kann; er nutzt deren spezifische "Struktur" überhaupt nicht aus.

In den Anwendungen hat sich demgemäß der EDA-Ansatz in der Modellfindungsphase durchgesetzt, nicht aber in der "rituellen" Verallgemeinerung von Wissen aus Stichproben auf Grundgesamtheiten unter Beiziehung von entsprechendem Hintergrundwissen und Zufälligkeit des datenerzeugenden Prozesses.

c) Der ANOVA-Ansatz zur Inferenz

Im ANOVA-Ansatz werden beobachtete Daten zerlegt in

Daten = allgemeines Mittel + kausale Einflüsse + zufällige Abweichungen

Der DEP ist hier durch additive Überlagerung von konstant wirkenden "Ursachen" und zufälligen Abweichungen gekennzeichnet. Während die zufälligen Abweichungen einer allgemeinen statistischen Analyse zugänglich sind, müssen die "kausalen" Komponenten vom Kontext her interpretiert werden. Zerlegung in die Komponenten erfolgt nicht nach Kriterien der Sachkenntnis sondern nach einem ritualisierbaren statistischen Verfahren. Das macht auch die weite Verbreitung des Ansatzes und dessen Beliebtheit aus.

Es sollen nicht die technischen Details erläutert werden, da kann man jedes bessere Methodenbuch zur Angewandten Statistik heranziehen. Für interessante Anwendungsbeispiele sei auf Borovcnik (1998) verwiesen. Hier sei nur ein kleines Beispiel angedeutet: Es sollen drei Unterrichtsmethoden miteinander hinsichtlich eines nach bestimmten Kriterien festgelegten Erfolgs beurteilt werden. Der kausale Einfluß soll durch Versuchsanordnung lediglich der Unterrichtsmethode zuschreibbar sein. Die Unterschiede sollen also :

(i) nicht dem Lehrer an sich zuordenbar sein; auch nicht der Wechselwirkung des Lehrers zu Methode, etwa der Lehrer mag Methode 2 nicht - das hat Auswirkungen auf seinen Unterricht;

(ii) nicht den Schülern, so etwa soll ihr Potential in den drei Gruppen so vergleichbar sein, daß sie eigentlich aus Drillingen bestehen, auf die nun einzig unterschiedlich der Stimulus der Unterrichtsmethode wirkt; nicht der Wechselwirkung zwischen bestimmten Schülertypen mit den Lernmethoden - die Typen sind ja durch die Drillings-Idee gleichgeschaltet und damit ausgeschaltet;

(iii) nicht durch die Art der Abhaltung des Unterrichts nach den diversen Methoden etc.

(iv) Etc.

Die Untersuchung nach eventuellen Störfaktoren macht die sogenannte Systemanalyse aus, auf deren Bedeutung der verstorbene angewandte Statistiker Josef Gölles immer wieder - und oft leider ungehört - hingewiesen hat. Wird diese nicht mit der nötigen Sorgfalt betrieben, so kann auch die beste statistische Auswertung der Daten nichts mehr bringen, weil die "kausalen" Einflüsse nun nicht mehr der Unterrichtsmethode zuordenbar wären, siehe auch Borovcnik u. Gölles (1994).

An Problemen der Methode seien genannt:

(i) Die technischen Details behindern ein Verständnis der Vorgangsweise, was zu unterschiedlichen Reaktionen führen kann - Ablehnung ebenso wie Überakzeptanz.

(ii) Die Mischung von Zufall und Kausalität interferiert mit der grundlegenden Schwierigkeit im

Wahrscheinlichkeitsverständnis, das häufig durch eine unselige Überlagerung dieser beiden Komponenten gekennzeichnet ist, wie viele empirische Untersuchungen zeigen.

(iii) In der Phase der Systemanalyse ist eine enge Zusammenarbeit von Statistiker und Kontext-Kenner notwendig. Versäumnisse hier sind, wie gesagt, auch durch noch so gute weitere Schritte nicht mehr korrigierbar.

Zusammenfassend ist festzuhalten, daß einerseits die stark vorhandenen Überlagerungen von zufälligen und kausalen Vorstellungen für viele Personen eine Barriere für Verständnis und Akzeptanz der ANOVA-Methode darstellen. Auch hier ist der Wahrscheinlichkeitsbegriff Basis einer technisch recht anspruchsvollen Prozedur zum Verallgemeinern von Ergebnissen aus Stichproben auf Grundgesamtheiten.

3. Wahrscheinlichkeit zum Strukturieren von Denken

Wahrscheinlichkeit ist keine wie immer geartete Eigenschaft von Objekten oder Prozessen. Wahrscheinlichkeit ist vielmehr Bestandteil einer Modellierweise. So etwa ist der Prozeß der Ankünfte von Kunden in einem System nicht zufällig oder nicht-zufällig, der Zufall tritt erst als Betrachtungsweise des Modelleurs in Erscheinung. So könnte man etwa gleichberechtigt einen Ansatz mit Wahrscheinlichkeiten, etwa mit Abhängigkeiten in Form von Markoff-Ketten versuchen, wie man auch Rekursionsgleichungen unterstellen könnte. Genauso gut könnte man statt der Zufallsmodellierung beim Roulette einen physikalischen Ansatz verfolgen und Croupier und Kugel gut studieren, was ja auch viele Spieler tun. Jüngst wurde dieser kausale Ansatz mit einem Gerät verfolgt, das die Geschwindigkeit der Kugel messen sollte; er war auch ziemlich gut, so gut, daß die Casinos dessen Verwendung verboten.

Da man Wahrscheinlichkeit nicht so einfach als eine Eigenschaft von Objekten oder Prozessen ansehen kann, da es vielmehr von Grund auf ein fiktiver Begriff zur indirekten "Ordnung" entsprechender Phänomene ist, gibt es auch wenige Lernmöglichkeiten und noch weniger Korrekturmöglichkeiten im Lernprozeß. Diese Form von Denken hat sich entsprechend wissenschaftsgeschichtlich sehr spät entwickelt. Entsprechend war es auch früh schon an die mathematische Repräsentation eng gebunden, so daß die Bedeutsamkeit eines intuitiven Zugangs zu dieser Form von Denken groß wird. Ein intuitiver Zugang könnte, ohne die Einschränkung auf mathematisch außerordentliche Fähigkeiten den Phasenraum des Denkens vergrößern und das bei einer breiteren Bevölkerungsschicht als heute.

a) Wahrscheinlichkeit ist keine materielle Eigenschaft

Hierzu ein kleines Beispiel: Man wählt eines der beiden Glücksräder in der Figur unten und dreht danach den Zeiger. Bleibt er im Sektor 1 stehen, so erhält man einen bestimmten Betrag; im Sektor 0 geht man leer aus. Welches Glücksrad sollte man wählen?

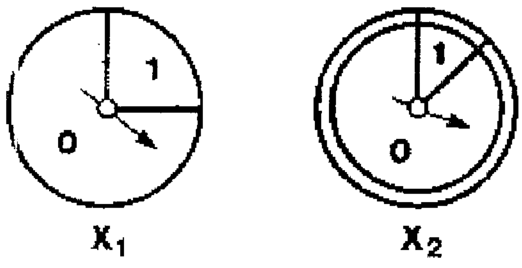

Fig. 1

Schon fünfjährige Probanden sind durchaus in der Lage, die Größe des Sektors als Maß der Erwartung für den Gewinnfall heranzuziehen und diese Größe, auch in komplizierteren Fällen zu vergleichen. Danach hätte man das linke Rad zu nehmen. Merkwürdigerweise hatte ein fünfjähriges Kind mit dieser Aufgabe Probleme; nicht bei gewöhnlichem Layout, jedoch bei besonderer Umrahmung des "schlechteren" Glücksrades. Da änderte es nämlich seine Strategie und wählte nunmehr das Rad mit dem kleineren Gewinnsektor. "Da müsse man dem Glück gegenüber das verheimlichen, worauf man abzielt."

Das belegt, wie widersprüchlich und wie instabil Verhaltensweisen in stochastischen Situationen sind und daß mathematische Kenntnisse allein keineswegs ausreichen. Darüber hinaus jedoch ist ein solcher Proband für die gewöhnliche Wahrscheinlichkeitsrechnung für einige Zeit verloren, wenn er mit seiner Wahl auch noch gewinnt, was ja vorkommen kann. Er verliert ja mit dem besseren Rad auch sehr häufig, so daß die richtige Wahl auch nicht gerade zu einem verstärkenden Erfolgserlebnis durch Gewinn führt. Wir ersehen daran, daß empirische Versuche nur ein indirektes Feedback mit vielen Störmöglichkeiten darstellen (Borovcnik 1994)

Häufigkeiten aus einer zufälligen Serie sind also kein direktes Feedback und bestätigen auch eine begrifflich korrekte Durchdringung eines Problems nicht. Man denke nur an die vielen Möglichkeiten, im zeitlichen Ablauf der Serie ein Muster für Abhängigkeiten der Ergebnisse zu entdekken. Man vergleiche auch hier die outcome-Orientierung, wonach eine abstrakte Wahrscheinlichkeit ja sowieso keine hilfreiche Größe darstellt, um den Ausgang des nächsten "Spiels" vorauszusagen.

Relative Häufigkeiten sind demnach nur eine *Metapher*, sie dienen dazu, irgendwelche "Konsequenzen" des Begriffs Wahrscheinlichkeit zu illustrieren. Dennoch sind sie nützlich, sie illustrieren auch mathematische Zusammenhänge oder verkürzen bzw. ersetzen mathematische Überlegungen, die ohnehin nicht einfach sind. Reduziert man den Wahrscheinlichkeitsbegriff aber allzusehr auf diese Metapher, so darf man sich nicht wundern, wenn man ein verzerrtes Konzept beim Lernenden anlegt.

b) Abwägen von Information

Der Aspekt des Abwägens, eines persönlichen quantifizierten Urteils für das Maß der Erwartung dafür, daß irgendein Ereignis eintritt, ist im klassischen Rahmen

(i) nach kombinatorischer Vielfalt bzw.

(ii) nach vergangener Häufigkeit

möglich. Beide Aspekte sind objekt- bzw. prozeßgebunden, also materiell ausgerichtet. Schon von Beginn der Wahrscheinlichkeitsrechnung aber war der Aspekt des Austausches zwischen Geld und Ungewißheit auch in Situationen (siehe Huygens 1657), die sich einem Urteil nach kombinatorischer Vielfalt bzw. vergangener Häufigkeit entzogen, ein tragender. Nur konnte man hierbei über die Prozedur der Quantifizierung nichts allgemeines entwickeln. Nur konnte man den Kalkül der Wahrscheinlichkeit an solchen Situation weder entwickeln noch illustrieren. Also hat man in Büchern zur Wahrscheinlichkeit eingangs alles entlang von Glücksspielen entwickelt, um dann aber hurtig zu solchen Anwendungen überzugehen.

Hier ist also abzuwägen zwischen Materialisierung eines Begriffes und der Idee hinter dem Begriff. Wenn sich die Idee schwer materialisieren läßt, so bekommt man leicht einen reduzierten Begriff im Denken darüber. Eine didaktisch reizvolle Aufgabe besteht nun darin, der Idee trotz mangelnder Materialisierungsfähigkeit dennoch zum Durchbruch zu verhelfen. Wenn man schon zur Quantifizierung von Wahrscheinlichkeiten wenig Brauchbares sagen kann, jedenfalls nicht viel mehr als daß sich in solchen Fällen Experten sehr bemühen alle vergleichbare Information zu sichten und bewerten, so kann man doch im Kalkül der Wahrscheinlichkeit manches Mal mit dem "Abwägen von Information" besseres Verständnis erreichen (z.B. die outcome-Orientierung als fehlgeleitet entlarven).

c) Zu Mis-Conceptions

Viele in der Literatur belegten Fehlschlüsse - das englische mis-conceptions ist hier als Bezeichnung viel aufschlußreicher - beziehen sich nicht erst auf statistische Schlußweisen (da gibt es auch jede Menge Fehleinschätzungen), sondern auf den Wahrscheinlichkeitsbegriff. Man kann diese Fehlschlüsse wie folgt unterteilen:

(i) Wahrscheinlichkeit - falsche Einschätzung des Werts und / oder falsche Einschätzung der Konsequenzen aus einem bestimmten Wert.

(ii) Bedingte Wahrscheinlichkeit - falsche Einschätzung, insbesondere durch Überlagerung mit nicht-stochastischen Denkweisen wie astrologischen, kausalen und logischen Überlegungen.

(iii) Bayes-Formel - falsche "Kurzschlüsse" unter Umgehung des formalen Kalküls, die weit daneben liegen. Wenn Wahrscheinlichkeit in die öffentliche Diskussion gelangt, sind es häufig Bayes-Probleme wie vor nun schon einigen Jahren das Drei-Türen-Problem, an dem sich gerade die gescheitesten Köpfe fürchterlich reiben und viel Unsinn verbreiten (mehr dazu in Borovcnik 1991 sowie Klemisch 1993).

Dazu kann man in der einschlägigen Literatur mehr nachlesen, auch in Borovcnik (1992) findet man einiges. Viele der Mis-conceptions sind durch die materielle Re-präsentation von Wahrscheinlichkeit bedingt, damit ist die Überbetonung von Wahrscheinlichkeit als einer objektiven Eigenschaft von Dingen und Prozessen bei gleichzeitiger Vernachlässigung des Aspektes des Abwägens von Information gemeint. Die outcome - Orientierung läßt gegebene Wahrscheinlichkeiten völlig falsch einsetzen aber auch falsch aufzeichnen bzw. aus dem Gedächtnis zurückrufen; es gibt viele andere Mechanismen, die zu einer falschen Registrierung vergangener Häufigkeiten führen, u.a. die Wünschbarkeit eines Ereignisses, die Ähnlichkeit von Ereignissen, falsche Vergleichsbasen etc. Auch hier liest man am besten in der Literatur oder bei Borovcnik (1992) nach.

Gleichzeitig deutet sich in der Literatur immer wieder eine Überlagerung verschiedener Denkformen an, wobei sich hinsichtlich der "Mächtigkeit" eine Hierarchie von logisch zu kausal zu probabilistisch ergibt. Hier gilt es, durch Aufbau einer ordentlichen Begrifflichkeit von Wahrscheinlichkeit Denken zu ordnen und "falsche" Ansätze zu klären.

d) Stochastisches Denken

Zu einzelnen Aspekten von stochastischem Denken ist ausreichend bei Borovcnik (1992 und 1997) nachzulesen, weshalb hier wenige Hinweise ausreichen sollten. Drei wichtige Bereiche sind hier zu nennen:

(i) Die *Einzelentscheidung* unter Unsicherheit - schon bei Huygens (1657) im Vordergrund; man vergleiche das Spiel mit den beiden Glücksrädern aus der Figur weiter oben oder den Abschluß irgendeiner Versicherung, die zu einer bestimmten Prämie gewisse, in ihrer Wahrscheinlichkeit unbekannte (Schadensfälle) Risiken abdeckt.

Erst mit J. Bernoulli (1713) und seinem "goldenen Theorem", dem schwachen Gesetz der großen Zahlen, tritt die relative Häufigkeit und die wiederholte Zufallssituation in den Vordergrund. In der Einzelentscheidung jedoch dominiert der Vergleich von erwartetem "Nutzen" für verschiedene Entscheidungen; hierzu findet sich auch einiges in Bentz (1983). Eigentlich ist auch jede Anwendung eines statistischen Schlusses eine solche Einzelentscheidung und entzieht sich einer long run - Deutung, welche immer wieder - gedankenlos - als Gedankenkrücke für ein "Verständnis" des eingesetzten Verfahrens verwendet wird (siehe Borovcnik 1992).

Klar, daß die Erschließung der Einzelsituation nach dem Begriff "Erwartungswert" zu ganz anderen intuitiven Vorstellungen führt als die mit Wahrscheinlichkeiten. Speziell ist der Erwartungswert auch einer nicht ausschließlichen "long run"- Deutung zugänglich, wie man bei Huygens, Freudenthal oder Bentz (1983) nachlesen kann. Diese Deutung entspricht viel besser der Problemstellung der Einzelentscheidung.

(ii) Entscheidung nach Indizien - Formel von Bayes. Aus bereits vorhandenem Wahrscheinlichkeitsurteil über einen Sachverhalt soll, bei bekannten Wahrscheinlichkeiten für das Auftreten von verschiedenen empirischen Befunden, auf ein neues, revidiertes Wahrscheinlichkeitsurteil geschlossen werden, welches das Vorliegen eines bestimmten empirischen Befundes mitberücksichtigt.

Klassisches Beispiel ist die Diagnose einer Krankheit unter Vorliegen bestimmter medizinischer Befunde. Erst wenn man die Wahrscheinlichkeiten verschiedener Befunde (genannt "positiv" und "negativ") unter Krankheit bzw. Fehlen der Krankheit kennt ***und*** wenn man ferner ein Wahrscheinlichkeitsurteil für die Krankheit bei der zu untersuchenden Person heranzieht, kann man aus der Bayes-Formel eine neue (aktuelle) Wahrscheinlichkeit berechnen.

Es wird immer wieder subjektives Urteil von Personen angeführt, das weit von den Ergebnissen der Bayes-Formel abweicht. Das wird in der Literatur als Nachweis gedeutet, daß

- Personen keine passenden Intuitionen, auch nach entsprechender Unterweisung in Wahrscheinlichkeitsrechnung entwickeln können, ferner daß
- der Kalkül keiner intuitiven Deutung zugänglich ist, daß man "nur" auf formale Aspekte Rücksicht nehmen kann/ muß, und ferner, daß
- ein Bayesianischer Zugang mit seinen Deutungen von Wahrscheinlichkeiten (in welchem diese Formel zentraler Bestandteil ist) nicht zielführend ist.

Dazu ist zu sagen, daß eine andere Repräsentation der Bayes-Formel in der Gestalt mit Chancenverhältnissen (sogenannten odds) einer intuitiven Deutung sehr wohl zugänglich ist und daß viele mathematische Zusammenhänge der Bayes-Formel "vernünftigem" Denken direkt entsprechen (siehe Borovcnik 1992): Erhöht sich die a priori-Wahrscheinlichkeit (z.B. der Krankheit), so auch die a posteriori-Wahrscheinlichkeit (die aktuelle Wahrscheinlichkeit); ist der empirische Befund sehr wahrscheinlich unter der Krankheit und sehr unwahrscheinlich unter Fehlen der Krankheit, so differenziert er gut, ist also ein gutes Indiz für die Krankheit; Vorliegen bzw. Nicht-Vorliegen des Indizes macht dann den Unterschied im Wahrscheinlichkeitsurteil über das Vorliegen oder Fehlen der Krankheit besonders groß. Man kann auch intuitiv verstehen, warum man seinerzeit in der AIDS-Debatte von einer verpflichtenden Reihenuntersuchung der Bevölkerung doch wieder abgesehen hat etc.

Es ist nicht so sehr die Frage, wie gut die Bayes-Formel empirisch das wirkliche Denken von Personen beschreibt (was sie nicht tut), sondern wie sehr man unter bestimmten Umständen ein unter anderen Bedingungen abgeleitetes Denkergebnis am Ergebnis der Bayes-Formel (mit bestimmten Eingangsgrößen) vergleichen und verbessern kann und ferner, wie sehr man diese Eingangsgrößen für eine gegebene Entscheidungssituation verbessern kann (mehrere, verschiedenartige Befunde, bessere Diagnoseverfahren usw.).

Probleme macht die Bayes-Formel immer auch wegen ihres Charakters als Einzelentscheidung, was in den üblichen Rahmen mit den long run - Deutungen von Wahrscheinlichkeiten nicht so ganz hinein paßt.

(iii) Der dritte Bereich sei mit dem Namen "Wurzel 1 durch n - Gesetz" umschrieben (siehe Riemer 1991); er handelt von der natürlichen Variation des Zufalls. Was ist bei Vorliegen einer bestimmten Verteilung, die durch irgendeine klar umrissene Hypothese festgelegt ist - für die wiederholten Messungen des zu untersuchenden Phänomens zu "erwarten" (hier nicht im Sinne des Erwartungswertes sondern im Sinne der empirischen Verteilung). Klar, daß hier die Simulation dieser Verteilung leicht zugängliche Ergebnisse liefert. Das ist wohl auch der Grund, daß die Deutung von Wahrscheinlichkeit als relative Häufigkeit sich so stark gegenüber anderen Deutungen durchgesetzt hat; sie ist eben materiell so viel leichter zu erschließen. Es darf jedoch nicht übersehen werden, daß die Basis der Simulation eine Verteilungsannahme aufgrund einer Hypothese ist.

Als Beispiel könnte man hier die in 2 a) genannte Prüfung der Unabhängigkeit zweier Würfel anführen. Die simulierte Verteilung bietet dann Vergleichsbasis für die Bewertung der einzelnen, bereits vorhandenen Stichprobe. Sie erklärt die Entscheidungsprozedur und macht den Begriff des statistischen Risikos zugänglich. Allerdings mit den üblichen, auch in der Literatur berichteten Mißverständnissen (siehe auch Borovcnik 1992), die nur mit allerhand Tricks intuitiv auszumerzen sind. Hierunter ist auch die Methode der Analogien oder Metaphern zu nennen.

Stochastisches Denken ist nicht nur das Anwenden stochastischer Modelle, sondern vielmehr das intuitive Erfassen formaler Begriffe und Ergebnisse mathematischer Berechnungen. Sozusagen ein Kurzschluß, ein Schluß, der wohl nicht dieselbe Strenge hat wie mathematische Überlegungen, aber viel mehr an Überzeugungskraft ausstrahlt, weil er wesentliche Bestandteile der Situation erfaßt. An didaktischen Strategien zur Vertiefung Stochastischen Denkens seien hier genannt:

(i) Analogien - der Kurzschluß anhand einer intuitiv leicht zugänglichen "analogen" Situation, die jedoch wesentliche Bestandteile der den Begriff repräsentierenden artifiziellen Situation oder der Mathematik erfaßt, siehe Brewer (1989) oder Borovcnik (1992).

Als ein Beispiel sei hier die Analogie des Fischens (von zwei Sorten Fischen, die den "richtigen" sowie "falschen" Nullhypothesen entsprechen) mit einem mehr oder weniger feinmaschigen Netz im Vergleich zum Testen von Hypothesen genannt. So etwa ist der Anteil der "richtigen" Fische nach dem Abfischen nicht gleich dem Anteil der "richtigen" Fische zu setzen, die durchs Netz schlüpfen.

(ii) Simulation - die materielle Illustration der Konsequenzen eines Begriffs, hier lediglich hinsichtlich seiner Auswirkungen auf lange Sicht. Das ganze hat trotz seiner einseitigen Erfassung des Begriffs durchaus seine Meriten, weil zu berücksichtigen ist, daß ein voll abgedeckter Begriff einige seiner Eigenschaften immer in bestimmten Deutungen leichter zugänglich macht.

Man muß nur darauf achten, daß der Begriff nicht auf "bequeme" Deutungen reduziert wird, was in der gängigen didaktischen Diskussion häufig der Fall ist. Besonders seien hier die zu einem Torso entarteten Interpretationen der klassischen Prozeduren zur statistischen Inferenz genannt: die Irrtumswahrscheinlichkeit (schon die Bezeichnung ist irreführend; besser Wahrscheinlichkeit eines Fehlers 1. Art) eines Tests ist keineswegs eine (globale) Irrtumswahrscheinlichkeit, sonst müßte ein Test umso besser werden, je kleiner seine Irrtumswahrscheinlichkeit ist, was im "besten" Fall dann zur Entscheidungsregel "lehne die Nullhypothese nie ab" führt (!).

(iii) Reformulierung - In einer anderen Formulierung des Kalküls ist vielleicht ein bestimmter

Aspekt eines Begriffs oder einer Situation viel besser einzusehen. So etwa kann die Einzelentscheidung mit dem Erwartungswert viel besser verstanden werden als mit Wahrscheinlichkeiten (siehe oben), die Bayes-Formel mit Chancenverhältnissen viel besser verstanden werden als mit Wahrscheinlichkeiten (siehe Borovcnik 1990). Das Chancenverhältnis ist W(E) zu W(non E); aus odds (wie man auch sagt) von z.B. 1 : 5 rechnet man eine Wahrscheinlichkeit von 1/6 zurück.

Wenn die a priori odds für eine Krankheit 5 : 10000 (eine seltene Krankheit) und die Indizienwahrscheinlichkeiten unter Krankheit bzw. Nicht-Krankheit wie 100 : 1 stehen (der Befund könnte nicht besser zwischen den beiden Möglichkeiten differenzieren!), dann ergibt sich mit der Bayes-Formel mit Chancenverhältnissen eine a posteriori Wahrscheinlichkeit für Krankheit von 5x100 : 10000x1, also 500 : 10000, oder 5 : 100, was einer Wahrscheinlichkeit von etwas weniger als 0,05 entspricht. Hier sieht man z.B. deutlich, daß die "überraschend" kleine Wahrscheinlichkeit für Krankheit unter Einbeziehung des Befundes von ca. 0,05 einzig und allein durch die geringe a priori-Wahrscheinlichkeit der Krankheit verursacht ist und keineswegs durch die Anwendung eines unbrauchbaren diagnostischen Verfahrens.

Ein ähnliches Problem tritt immer bei seltenen Krankheiten auf. In der Darstellung der Bayes-Formel mit Chancen jedoch kann man wenigstens intuitiv die Schwierigkeit verstehen und den "Stellenwert" der Ergebnisse einstufen.

4. Wahrscheinlichkeit zum Strukturieren und Erschaffen von "Realität"

Vielfach wird in der Wahrscheinlichkeitsrechnung wie auch in der Angewandten Mathematik an sich ein naiver Modellbegriff zugrunde gelegt. Es wird von einer als gegeben vorausgesetzten Situation ausgegangen, derer man sich, allenfalls durch iterativ verbessertes Modellieren immer mehr annähert. Die Probleme daraus hat der Autor ausführlich in Borovcnik (1984) beschrieben. Wahrscheinlichkeitsmodelle entziehen sich jedoch prinzipiell einer quantitativen Beurteilung, inwiefern sie auf die Realität passen. Jedenfalls ist z.B. die Interpretation des Tests auf Vorliegen einer bestimmten Verteilung für ein Merkmal ganz anderer Natur. Dies aufgrund der prinzipiell fehlenden Alternativhypothese bzw. der globalen Alternativhypothese jeder beliebigen anderen Verteilung als in der Nullhypothese festgehaltenen. Das hat zur Folge, daß man prinzipiell keinen Fehler 2. Art berechnen kann, was, wenn überhaupt, den Test einigermaßen rational vertretbar macht (siehe Borovcnik 1984).

Es tritt demgemäß viel stärker ein sogenannter Szenario-Gedanke beim Modellieren in der Stochastik in den Vordergrund: Man geht verschiedene, plausible Modelle durch und vergleicht die Berechnungen im Kontext des Problems. Die einzelnen Modelle repräsentieren ganz unterschiedliche Qualitäten von Annahmen, die allesamt durchaus in Frage zu stellen sind. Im Unterschied der Annahmen und der Resultate jedoch ergeben sich jedoch durchaus vernünftige

Hinweise für kritische Punkte der Modellierung sowie Stützen für zu treffende Entscheidungen.

Dieser Szenario-Gedanke geht also von vornherein davon ab, eine (nicht vorhandene) Wirklichkeit möglichst getreu im Modell abzubilden, das Ganze wird zu einem Planspiel, die verwendeten Modelle werden zu einem Spielzeug/ Werkzeug, um über die Wirklichkeit nachzudenken. Jedenfalls erweitern sie den Phasenraum des Nachdenkens und Problemlösens mächtig. Nur wer die Modelle in ihrem Stellenwert innerhalb der Theorie gut überschauen und verstehen kann (einschließlich aller Annahmen), kann die Szenarien gut vergleichen. Da man aber die ganze Mathematik nicht so leicht überschauen kann, wird wieder einmal klar, wie bedeutsam deren intuitive Erschließung eigentlich ist.

a) Verallgemeinern aus zufälligen Stichproben

Hier wurde schon einiges in Abschnitt 2 a) festgehalten, weitere Details findet man z.B. in Borovcnik (1992 oder 1985). Die Technik der zufälligen Stichproben tritt in der Praxis immer in Konkurrenz zu direkt repräsentativen Stichproben wie etwa der Quotenstichprobe bei Meinungsforschungsinstituten. Als Vorzüge einer zufälligen Stichprobe können summarisch festgehalten werden: Die Technik der zufälligen Stichproben garantiert

(i) die Repräsentativität der gezogenen Stichprobe - ein scheinbarer Widerspruch, denn die Technik garantiert auch, daß jede einzelne, mögliche Teilmenge von beobachteten Objekten dieselbe Auswahlwahrscheinlichkeit besitzt,

(ii) die Kontrolle des sogenannten statistischen "Restrisikos", daß die gezogene Stichprobe doch nicht repräsentativ ist.

Es wird also eine "durchschnittliche" Repräsentativität garantiert, eine Eigenschaft, die dann auf die einzelne Stichprobe übertragen wird. So wird der oben angesprochene Widerspruch aufgelöst. Mit dem Restrisiko kann man (etwa über Vertrauensintervalle) die Ergebnisse aus der Stichprobe auf die Grundgesamtheit übertragen. Es wird nicht nur die Ähnlichkeit zwischen Grundgesamtheit und Stichprobe garantiert - wie auch im Quotenverfahren - sondern es werden auch Genauigkeitsangaben möglich. Durch Schichtung der Grundgesamtheit kann man auch die Vorzüge des Quotenverfahrens nützen.

Für die Prüfung empirischer Hypothesen (etwa: besteht ein Unterschied zwischen verschiedenen Lehrmethoden hinsichtlich eines bestimmten Erfolgskriteriums) benötigt man keineswegs zufällige Stichproben, sondern lediglich die zufällige Aufteilung einer Gruppe von Objekten auf die "Behandlungsgruppen", siehe den ANOVA-Ansatz in 2 c). Allerdings hat man eine generelle, qualitative Repräsentativität der zugrunde gelegten Gruppe doch irgendwie herzustellen. So etwa werden sich die Ergebnisse einer Untersuchung, die sich nur auf hochbegabte Personen bezieht, sicherlich nicht auf die allgemeine Bevölkerung so ohne weiteres übertragen lassen.

b) Beurteilung von Risiken

Es gibt eine Reihe von Risiken, deren Höhe aus unterschiedlichem Interesse (privat bzw. öffentlich) vernünftig eingeschätzt werden muß. Eine Einzelperson muß ihre Risiken bezüglich des Autofahrens einschätzen, wenn sie eine Vollkasko-Versicherung erwägt, sie muß hinsichtlich des zukünftigen Gesundheitszustandes bzw. erreichbaren Lebensalters Einschätzungen vornehmen, wenn sie eine Zusatz-Krankenversicherung bzw. eine Lebensversicherung abschließt. Behörden müssen Risiken eines Bergabbaus (Vergiftung der Gewässer, Verschiebung des Grundwasserspiegels, tektonische Verschiebungen, Sicherheit des Personals etc.) abschätzen. Dieselben Behörden geben umfängliche Studien in Auftrag, um die Sicherheit von Kernkraftwerken oder die Folgen von bestehenden Kernkraftwerken auf die Gesundheit der umliegenden Bevölkerung zu bewerten.

Immer geht es um die Beurteilung von Risiken als Wahrscheinlichkeiten oder Erwartungswerte oder gar in Form einer Verteilung (einer Belastungsgröße). Bei vielen dieser Beurteilungen hat man gar nicht die Möglichkeit einer stichprobenmäßigen Erfassung, man ist daher auf Urteile, auf Einschätzungen, insbesondere von Experten, angewiesen. Diese Expertenurteile bezüglich einiger, einfacherer Komponenten des Problems werden in umfängliche Modellierungen eingebaut und einer weiteren mathematischen Analyse unterzogen. So erhält man konkurrierende Szenarios, immer wieder hört man erstaunt von einander so widersprechenden Ergebnissen.

Für die Beurteilung der Sicherheit von Kernkraftwerken z.B. werden Expertenurteile über die Sicherheit von einzelnen Bauteilen (wie der Sicherheit von Pumpen, Ventilen, Schaltern etc.) in ein umfängliches Netzwerk von gegenseitigen (modellmäßigen) Abhängigkeiten eingebaut und man erhält daraus rechnerische Sicherheit des gesamten Kühlsystems, Notaggregats usw. und daraus schließlich die Wahrscheinlichkeiten eines GAUs oder Super-GAUs (siehe Bungartz 1988). Die Expertenurteile beziehen sich auf die Sicherheit im Labortest und nicht auf die Praxis-Sicherheit, die Abhängigkeiten sind modellmäßig und beziehen Praxisbedingungen nicht ein, menschliche Reaktionen, Alterung der Systeme usw. werden kaum erfaßbar.

Aus der Berechnung erhält man qualitative Ergebnisse, die erst im Vergleich richtig an Wert gewinnen: Redundanz (etwa Einbau doppelter Systeme) gewisser Bauteile erhöht die Sicherheit um so viele Zehnerpotenzen, der Typ A ist um so viel sicherer als Typ B. Vergessen wird, daß hier Wahrscheinlichkeit überhaupt nicht als absolute Größe interpretiert werden kann, schon gar nicht in einer Häufigkeitsdeutung wie, das wird in 100 Jahren nur ein halbes Mal passieren.

Fazit: Bei der Beurteilung von Risiken hat man es vorwiegend mit Wahrscheinlichkeit in einer Einzelsituation und mit Experten-Urteilen (oder subjektiven Einschätzungen) zu tun. Eine long run - Deutung der Ergebnisse geht grundsätzlich fehl. Die Ergebnisse sind als Szenarien für die Wirklichkeit zu verstehen und können durchaus zur qualitativen Verbesserung von Systemen dienen, sie sind aber zur absoluten Beurteilung der Risiken nur bedingt geeignet.

c) Lebensdauer und Zuverlässigkeit

In technischen Anwendungen geht es auch häufig um die Schätzung einer Wahrscheinlichkeit, nämlich um die Überlebenswahrscheinlichkeit technischer Bauteile oder Ausrüstungen, wobei die Lebensdauer durch eine Wahrscheinlichkeitsverteilung modelliert wird. Die Wahrscheinlichkeit, einen bestimmten Zeitraum zu überleben (funktionieren), wird auch Zuverlässigkeit genannt (mehr Details, Beispiele und Fragestellungen findet man in Borovcnik 1997).

In biometrischen Anwendungen geht es um den Nachweis der Wirksamkeit von Medikamenten und oder Operationsmethoden, oder auch um den Vergleich von neuen Medikamenten mit solchen, die gegenwärtig als Standard gelten. Auch hier geht es oft um die Beurteilung der Überlebensdauer (z.B. der Dauer bis zum erneuten Wachstum eines Tumors). Hier wird die mittlere Überlebensdauer unter verschiedenen Behandlungsmethoden miteinander verglichen.

In beiden Anwendungsfällen entzieht sich die Wahrscheinlichkeit weitgehend einer long run - Deutung. Die Unabhängigkeit des Ausfallverhaltens von Bauteilen in technischen Anwendungen ist ein reiner Modellbegriff, mangels vernünftiger Daten über Zusammenhänge eine Art Joker.

Die Lebensdauer wie sie in empirischen Tests im Labor ermittelt wird, ist rein fiktiv, weil sie meist unter wesentlich erhöhtem Stress (Belastung) gemessen und dann auf geringe Belastung in der Praxis übertragen wird. Grund dafür ist, daß bei gefordertem geringen Ausfallverhalten, also bei hoher Lebensdauer (von mehreren Jahren), Ausfälle bei den beobachteten technischen Einheiten während relativ kurzer Beobachtungszeit im Labor (von unter einem Monat, oft noch weniger) überhaupt nicht zu beobachten wären. Man verwendet also Beschleunigungsfunktionen bzw. umgekehrt Entschleunigungsfunktionen, um zwischen hohem Stress im Labor und praktischen Nutzungsbedingungen hin und her zu rechnen. Diese Funktionen entsprechen durchaus technischem know how, müssen aber auch gewissen Kriterien der Einfachheit und der einfachen Interpretation genügen.

Jedenfalls ist eine Deutung von (Überlebens-)Wahrscheinlichkeit als relative Häufigkeit eine reine Fiktion. Dennoch kann man wichtige Kriterien aus den verwendeten Modellen ableiten, etwa Garantiezeit, Garantiekosten, Wartungsintervalle etc. Allenfalls müssen in der Praxis Korrekturen angebracht werden. Keine Korrekturen sind wohl dann möglich, wenn es sich tatsächlich um Einmaleinsätze, wie etwa der Verwendung eines Materials, Bauteils oder Instruments für einen Weltraumflug handelt.

In biometrischen Anwendungen entbehrt die Modellierung meist auch einer Häufigkeitsdeutung von Wahrscheinlichkeit (auch der Unabhängigkeit, die ja jeden Anwendungsfall als gleichwertig typischen Fall hinstellt), sie stellt jedoch die einzige Ratio dar, um neue Medikamente kontrolliert an Menschen auszutesten und die Ergebnisse sinnvoll zu bewerten.

Insgesamt ergibt sich die Anwendung von Wahrscheinlichkeitsmodellen im Szenario-Gedanken. Szenarien, welche die Realität nur sehr indirekt abbilden können, die auch gar nicht passen, bilden die Basis für Transparenz und Vergleichbarkeit von Entscheidungen, die ansonsten nur willkürlich wären. Der Erfolg in technischen Anwendungen sowie die Notwendigkeit, sich in medizinischen Anwendungen besser und rascher zu orientieren, zeigt auch, wie wenig man in der Schule den Gedanken des Szenarios übergehen kann.

c) Warteschlangensysteme

Nur andeutungsweise sei ein weiterer Anwendungsbereich von Wahrscheinlichkeit angeführt: die Verbesserung von "Systemen" zur Betreuung von "Kunden" hinsichtlich des Kriteriums der "Wartezeit". Die Ankünfte werden nach irgendeiner Annahme als zufällig modelliert, z.B. als poisson-verteilt (was nach entsprechenden mathematischen Eigenschaften des Poisson-Prozesses als rein zufällig angesprochen werden kann). Ebenso werden die Bedienzeiten nach einer entsprechenden Verteilung modelliert. Im Modell können dann mathematisch oder durch Simulation (die Mathematik wird schnell unhandlich) viele Eigenschaften dieser Systeme abgeleitet werden: die mittlere Bedienzeit, mittlere Wartezeit, Verteilung der Anzahl der Wartenden (für die Bemessung des Warteraums) etc. Mehr dazu, für den Unterricht aufbereitet, findet man auch in Lindenau (1988).

Auch hier gilt: die verwendeten Modelle können keineswegs den Gedanken der bestmöglichen Abbildung der Verhältnisse in der Realität verfolgen (das wäre wirklich aussichtslos), man muß sich auf den Szenario-Gedanken besinnen und die Ergebnisse verschiedener Szenarien miteinander vergleichen. Dennoch können durchaus praktisch verwertbare qualitative Lösungen und Einsichten gewonnen werden. Je künstlicher die Welt, also je besser man in der Realität die Annahmen der Szenarien herstellen kann, umso besser werden die Resultate. In einer zunehmend künstlichen Welt werden diese Modelle durchaus zur Vorausschau.

Insgesamt sollte man hier festhalten, daß Wahrscheinlichkeit Modelle bereit stellt, die ein Denken in Szenarios begünstigen. Damit erhält man neue Werkzeuge, um über die Realität nachzudenken. Insbesondere gilt es, Szenarien zu bewerten, Szenarien miteinander zu vergleichen und dadurch daraus qualitative Schlüsse für die Realität zu ziehen.

5. Wahrscheinlichkeit: Heuristik und Szenario

Die Ausführungen seien wie folgt zusammen gefaßt:

Wahrscheinlichkeit soll Realität strukturieren:

- Wahrscheinlichkeit wird häufig allzu eng an das Argument zufälliger Stichproben gebunden
- die Modelle werden zu eng als deskriptive Modelle, die eine Realität möglichst gut abbilden, verstanden
- der Szenario-Charakter von Wahrscheinlichkeitsmodellen wird nur ungenügend berücksichtigt

Wahrscheinlichkeit soll Denken strukturieren:

- Üblicherweise wird Wahrscheinlichkeit eingeschränkt auf die Interpretation und Messung durch kombinatorische Vielheit und oder vergangene relative Häufigkeit
- Begrifflich entsteht dadurch eine zu enge Bindung an materielle Re-präsentationen - auch viele Bilder und Sprechweisen drücken das aus
- der heuristische Charakter von Wahrscheinlichkeit zur „spielerischen" Untersuchung realer Phänomene und dem Lernen daraus wird kaum berücksichtigt.

Aus den Schwierigkeiten, Wahrscheinlichkeit in der Vergangenheit erfolgreich zu unterrichten, zieht man nach Ansicht des Autors die falschen Schlüsse. Die Ausführungen in diesem Artikel sollten aufweisen, daß es möglich, wertvoll und erstrebenswert, ja notwendig ist, Wahrscheinlichkeit als Heuristik, die den Denkraum des Menschen erweitert, und als Szenario-Werkzeug, die den Problembearbeitungsraum des Menschen vergrößert, zu unterweisen. In diesem Sinne: „Rettet die Wahrscheinlichkeit!"

Literatur

Bentz, H.J.: Zum Wahrscheinlichkeitsbegriff bei Chr. Huygens. In: *Didaktik der Mathematik* **11**(1983), 76-83.

Bernoulli, J.: *Ars conjectandi.* Basel: 1713.

Borovcnik, M.: *Was bedeuten statistische Aussagen.* Wien: Hölder-Pichler-Tempsky 1984.

Borovcnik, M.: Zum Anwendungsproblem in der Statistik. In: *mathematica didactica* **7**(1984), Teil 1: 21-35; Teil 2: 121-135.

Borovcnik, M.: Ein direkter Zugang zur Beurteilenden Statistik. In: *Didaktik der Mathematik* **13**(1985), 251-271.

Borovcnik, M: Klassische und Bayessche statistische Methoden - Ein kritischer Vergleich. In: *Österreichische Zeitschrift für Statistik und Informatik* **16**(1986), 3-28.

Borovcnik, M.: Ein intuitiver Zugang zur bedingten Wahrscheinlichkeit und zur Bayes-Formel. In: *Stochastik in der Schule* **10**(1990) 3, 22-35.

Borovcnik, M.: Problemecke. In: *Stochastik Schule* **11**(1991)3, 46-51.

Borovcnik, M.: Analogien zum besseren Verständnis von Stochastik. In: *Didaktik der Mathematik* **20**(1992), 39-57.

Borovcnik, M.: *Stochastik im Wechselspiel von Intuitionen und Mathematik.* Mannheim: Bibliographisches Institut 1992.

Borovcnik, M.: Der Einfluß des Computers auf die Statistik-Ausbildung. 1994. In: H. Friedl (Hsg.): *Was ist Angewandte Statistik*, Grazer Mathematische Berichte Nr. 34. Graz: TU Graz, 9-20.

Borovcnik, M.: Statistische Analyse von Zusammenhängen - Regression und Korrelation. In:

Schriftenreihe zur Didaktik der Mathematik der Höheren Schulen der ÖMG **17**(1989), 1-20.

Borovcnik, M.: Intuitive strategies for teaching statistics. In: Brunelli, L. u. Cicchitelli, G.: *Proceedings of the First Scientific Meeting of the IASE.* Perugia: Unversity of Perugia.

Borovcnik, M.: Fundamentale Ideen als Organisationsprinzip in der Mathematik-Didaktik. In: *Schriftenreihe zur Didaktik der Mathematik der Höheren Schulen der ÖMG* **27**(1997), 1-13.

Borovcnik, M.: Trends und Perspektiven in der Stochastik-Didaktik. In: Kadunz, G. e.a. (Hsg.): *Trends und Perspektiven.* Wien: Hölder-Pichler-Tempsky 1996, 39-60.

Borovcnik, M.: Fallstudien in Angewandter Statistik - die Rolle von Systemanalyse und Computern in der Ausbildung in Statistik. Klagenfurt: Manuskript 1998.

Borovcnik, M.: Probleme und Beispiele aus der Zuverlässigkeitstheorie. In: *Schriftenreihe zur Didaktik der Mathematik der Höheren Schulen der ÖMG* **26**(1997), 1-20.

Borovcnik, M.: Bestrebungen zur Förderung Unterricht in Statistik. In: *Schriftenreihe zur Didaktik der Mathematik der Höheren Schulen der ÖMG* **30**(1999), 10-29.

Borovcnik, M. u. Gölles, J. Prinzipien und Probleme der Angewandten Statistik. Klagenfurt: Manuskript 1994.

Borovcnik, M. u. Ossimitz, G.: *Materialien zur Beschreibenden Statistik und Explorativen Datenanalyse.* Wien: Hölder-Pichler-Tempsky 1987.

Borovcnik, M. u. Peard, R.: Probability. In: Kilpatrick, J. e.a. (Hsg.) : *International Handbook of Mathematics Education.* Dordrecht: Kluwer 1996.

Brewer, J.K.: Analogien und Parabeln im Statistik-Unterricht. In: *Stochastik in der Schule* **9**(1989)3, 47-49.

Bungartz: Das Risiko von Kernkraftwerken. In: *Mathematik Lehren* **29**(1988).

Huygens, Chr.: De ratiociniis in ludo aleae. In: F. v. Schooten: *Exercitationes matematicae.* Leyden: 1657.

Kahneman, D., Slovic, P. U. Tversky, A.: *Judgement under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press 1982.

Klemisch, I.: Ein Einstieg über das Drei-Türen-Problem. In: *Stochastik Schule* **13**(1993)1, 9-15.

Konold, C.: Informal conceptions of probability. In: *Cognition and Instruction* **6**(1989)1, 59-68.

Lindenau, V.:, Einfache Simulationsmodelle für Warteschlangen. In: *Stochastik in der Schule* 8(1988)3, 40-57.

Nemetz, T.: An overview of the teaching of probability in secondary schools. In: Phillips, B. (Hsg.): *Papers on Statistical Education.* Hawthorn: Swinburne University of Technology 1997, 75-86.

Riemer, W.: Das ‚Eins durch Wurzel n' - Gesetz. Einführung in statistisches Denken auf der Sekundarstufe I. In: *Stochastik in der Schule* **11**(1991)3, 24-36.

Traar, S.: *Intuitive Vorstellungen zum Wahrscheinlichkeitsbegriff.* Klagenfurt: Univ. Klagenfurt (Diplomarbeit) 1989.

Walter, H.: Heuristische Strategien und Fehlvorstellungen in stochastischen Situationen. In: *Der Mathematik-Unterricht* **29**(1983), 11-23.