Tagungsband zum Wissenschaftsforum 1990

Defizite der Umweltforschung

Wissenschaftszentrum Bonn

8. Februar 1990

Gemeinschaftsveranstaltung
von
GSF und Bild der Wissenschaft
unterstützt vom
Bundesministerium für Forschung und Technologie

Organisation: Dr. E. Hofner, H.-J. Haury
Redaktion des Berichts: Dr. E. Hofner

Inhaltsverzeichnis

Seite

Die Sicht der Umweltorganisationen:

Schlußwort:

Prof. Bodo Pareigis (Universität München)

Fehlinformation der Öffentlichkeit durch statistische Daten

Im Amerikanischen gibt es die Bezeichnung "illiterate", d.h. des Lesens und Schreibens unkundig - wir sprechen dann von Analphabeten. Daneben gibt es dann den Begriff "innumerate", also des Rechnens unkundig. Dieser Begriff macht zur Zeit in den USA Furore. Schon seit einigen Monaten findet man das Buch Innumerancy des Autors John Allen Paulos auf den ersten Plätzen der Bestsellerliste. Sicher ist nicht jedem, der lesen und schreiben kann, gegeben, Gedichte wie Goethe zu schreiben, ja nicht einmal in ihrer Tiefe, in ihren Gefühlen und in ihrer Bedeutung ganz zu verstehen, - so ist es auch nicht jedem, der rechnen kann, möglich, über die einfachen Rechenvorschriften (wie sie uns von dem bekannten mittelalterlichen Rechenmeister Adam Riese überkommen sind) hinaus komplizierte mathematisch formulierte Gesetzmäßigkeiten der Umwelt zu verstehen.
Ich möchte Ihnen heute an einigen Beispielen zeigen, daß auch wir in Deutschland Probleme haben, die unter den Begriff "innumerate" fallen, darüber hinaus aber auch auf viele scheinbar richtige Berechnungen, die wir wegen ihrer Einfachheit sogar überprüfen zu können glauben, hereinfallen können. Ganz besonders viel Gefährliches und Unsinniges findet man bei Anwendungen der Statistik und Wahrscheinlichkeitstheorie in alltäglichen Problemen. Mein Anliegen heute ist es, Sie auf Fallstricke aufmerkasam zu machen, über die häufig sogar Fachleute stolpern, und Sie darauf hinzuweisen, daß eine Überprüfung statistischer Aussagen durch Fachmathematiker immer angebracht ist. Ich will Ihnen zeigen, daß viele Ergebnisse gegen die Intuition verstoßen und daß schon bei einfachen statistischen Aussagen große Fehlermöglichkeiten bestehen.
Lassen Sie mich an den Anfang eine Meldung der Süddeutschen Zeitung vom 23. Dezember 1989 stellen. Sie lautet:
Elfmal mehr Leukämie in Sittensen als üblich
stg. Bremen (Eigener Bericht) - Das Mainzer Universitätsinstitut für Medizinische Statistik und Dokumentation hat jetzt anhand eines von Mainz geführten bundesweiten Kinderkrebsregisters errechnet, daß in den vergangenen fünf Jahren in Sittensen (Kreis Rotenburg/Wümme) elfmal mehr Kinder an Leukämie erkrankt sind, als nach dem Bundesdurchschnitt zu erwarten wäre. Von anderen bösartigen Tumoren würden die Kinder in Sittensen viermal häufiger befallen. In der 9000 Einwohner-Gemeinde wurden jüngst fünf Fälle von Leukämie bei Kindern festgestellt. Eines der Kinder ist inzwischen gestorben. Der Mainzer Institutsdirektor jörg Michaelis wollte auf Anfrage der SZ nicht ausschließen, daß es sich um eine zufällige Häufung handelt. Es sei aber auch denkbar, daß zum Beispiel Umweltbelastungen die Erkrankungen verursacht hätten. Die Gesundheitsbehörden forschen schon seit Tagen vergeblich nach möglichen Ursachen.

Ich werde etwas später auf die Bedeutung solcher Feststellungen eingehen.

Die Mathematik hat unter anderem die Aufgabe übernommen, dem Menschen zu helfen seine umwelt durch Vorausberechnungen zu verstehen und damit auch zu beherrschen. Dadurch werden wir Mathematiker aber auch mit all den Problemen konfrontiert, die moderne Technologie nach heutiger Sicht mit sich bringt. In einem berühmten Report, dem Davis Report, erstellt für die amerikanische Regierung, steht das in knappster Form: "Hochtechnologie ist Mathematik". Die Amerikaner haben das genau erkannt und haben in den letzten Jahren große Anstrengungen gemacht, um diese mathematischen Grundlagen für die zukünftigen Entwicklungen kräftig zu stärken. Ebenso wird in Frankreich und in besonderem Maße auch in Japan die Mathematik in großem Umfang gefördert. Bei uns ist die bisherige Förderung auf mathematischem Gebiet eher mager mit all den vorhersehbaren Konsequenzen für die wirtschaftliche Entwick-lung in den kommenden Jahrzehnten.
Bevor ich auf die Anwendungen der modernen Statistik - zum größten Teil erst in diesem Jahrhundert entwickelte mathematische Gesetze - eingehe, lassen Sie mich auf einige einfache Probleme und Fallstricke eingehen.

I) Prozentrechnung, die Statistik des kleinen Mannes
Die meisten als statistisch ausgegebenen Werte, die man in der Tagespresse lesen kann, sind lediglich Prozentzahlen über gewisse Auszählungen: "Haushalte mit einem Netto-Monatseinkommen von ca. 6000,- haben eine Sparqote von 14%." Schon bei solch einfachen Prozentangaben haben viele von uns Schwierigkeiten. Ich denke hier z.B. an die ehrenwerten Damen und Herren des Bundesfinanzhofes. Ich möchte Ihnen nämlich einen Auszug aus einem Steuerberatungsbuch zeigen, das ich kürzlich einsehen konnte.
Urteil v. 18.10.1983, BStBl. 1984 II S.112
Bei der Ermittlung des Anteils der Arbeitszimmerfläche an der Wohnfläche ist das Arbeitszimmer nicht in die Wohnfläche einzubeziehen.
Das heißt dann

(prozentueller) Anteil des Arbeitszimmers= Fläche des Arbeitszimmers : Wohnfläche (ohne Arbeitszimmer)

Die Finanzverwaltung (Fin. Ger. Rheinland-Pfalz: Urteil v. 13.8.1985, EFG 1986 S. 174) hält inzwischen dagegen, daß man doch wohl rechnen müßte

(prozentueller) Anteil des Arbeitszimmers= Fläche des Arbeitszimmers : Wohnfläche (einschl. Arbeitszimmer)

Da hat offenbar der Bundesfinanzhof den Berechnungsmodus zu anteiligen Berechnung der steuerlich absetzbaren Unkosten für ein Arbeitszimmer auf eine mathematisch sehr fragwürdige Weise festgelegt. Um den Prozentsatz dessen, was man von den Gesamtausgaben

für seine Wohnung oder sein Haus für das Arbeitszimmer absetzen kann, soll man die Fläche des Arbeitszimmers durch die Fläche der restlichen Wohnräume teilen. Man macht sich an folgendem Beispiel klar, daß dieser Berechnungsmodus völlig unsinnig, aber für den Steuerzahler sehr vorteilhaft ist. Bei einem Arbeitszimmer, etwa einem Atelier, von 60 qm und sonstigen Wohnräumen ergäbe sich eine Absetzungsrate von 150%, d.h. man kann das 1 1/2-fache seiner Ausgaben für die Gesamtwohnung steuerlich absetzen, bei 90 und 30 qm sogar das Dreifache der echt entstandenen Auslagen. Eine neuerliche Entscheidung zu dieser "Formel" wird vom obersten deutschen Finanzgericht demnächst erwartet.
Dieses ist ein Beispiel dafür, daß man sich die Unsinnigkeit von Zahlen oder Berechnungen häufig durch Extremfälle klarmachen kann. Die Rechnung selbst mag wohl stimmen, nur hat das, was bei der Rechnung herauskommt, nichts mit der Realität zu tun. Prozentangaben ohne Angabe der Bezugsgröße sind sinnlos, mit Angabe der Bezugsgröße können sie völlig irreführend sein, wenn die Bezugsgröße falsch gewählt ist. Falsche Wahl der Bezugsgröße ist meines Erachtens ein deutliches Zeichen von <u>Innumeracy</u>. Es gibt Hunderte und Tausende von Beispielen dafür. Lassen Sie mich nur die bekannte Scherzfrage vom Kaufmann nennen, der erst seine Ware um 30% verteuert und dann wegen des massiven Umsatzrückgangs die neuen Preise wieder um 30% verringert. Tatsächlich hat er nicht das alte Preisniveau erreicht, sondern ist um 9% billiger als zuvor.
Oder "15% der Bevölkerung glauben, daß Krankenpfleger und Krankenschwestern ihren Beruf ergriffen, weil sie in keinen besseren Beruf hineinkommen konnten." Originalton 3. Fernsehprogramm Bayern. Heißt das, daß 85% meinen, die Krankenpfleger und Krankenschwestern hätten den besten aller Berufe gewählt? Wer verdummt hier wen?
Wenn nach Meldungen im letzten Monat die Arbeitslosenrate von 7.1% auf 6.9% gefallen ist und die Zuverlässigkeit - oder das Konfidenz-Intervall, wie die Statistiker sagen - plus oder minus 1% ist, dann kann der gute Eindruck total täuschen. Statt einer fallenden Rate könnte dann tatsächlich eine steigende Rate bestehen. Aber die Mitteilung der Zuverlässigkeit einer solchen Aussage ist (vielleicht wegen mangelnden Verständnisses?) für die Medien meist zu aufwendig.
Frauen haben nur 59% des Einkommens von Männern (USA 1980). Heißt das, daß sie im Beruf für dieselben Aufgaben auf vergleichbaren Positionen so bezahlt werden? Berücksichtigt diese Zahl den steigenden Anteil der weiblichen Bevölkerung an allen Berufstätigen und die dadurch bedingte geringere Berufserfahrung oder die geänderte Altersverteilung? Wird berücksichtigt, daß Frauen, aus welchen Gründen auch immer, in niedriger bezahlten Berufen tätig sind? Ist eingeschlossen die Tatsache, daß durch den Beruf des Ehemannes der Wohnort und damit das Arbeitsplatzangebot bestimmt werden? Arbeitet ein Teil der Frauen nur für kurzfristig gesetzte Ziele, für einmalige Anschaffungen im Haus? Sicherlich sind diese Punkte nicht berücksichtigt worden. Was ist eine solche Aussage dann noch wert? Das Durchschnittseinkommen der Frauen war 1980 59% des Durchschnittseinkommens der Männer. Eine magere Aussage.
Daraus allein kann man überhaupt keine soziologischen Schlußfolgerungen ziehen. Solche Zahlen sollten so nackt wie hier mitgeteilt niemals der Öffentlichkeit vorgesetzt werden, es sei denn,

man verfolgt damit bestimmte politische Ziele. Dann ist es aber eine Irreführung der Öffentlichkeit und ein Mißbrauch von statistischen Daten.
Sind 50% Chancen für Regen am Samstag und 50% Chancen für Regen am Sonntag 100% Chancen für Regen am Wochenende, oder aber 50% Chancen für Regen am Wochenende oder noch etwas anderes?

II) Würfeln und andere Glücksspiele
Von jeher war der Mensch dem Glücksspiel zugeneigt, etwas Irrationalem, wie wir es empfinden, etwas, das keinen nachvollziehbaren Gesetzen gehorcht. Aber auch Geburtsraten, Todesfälle, Unfälle, die Wirtschaft, das politische Leben, ja mit der Quantenmechanik auch alle weiteren Bereiche unseres Lebens hängen vom Zufall ab, vom berechenbaren Zufall zwar und von Statistiken. Sie sind zwar nicht wirklich deterministisch. Statistische Aussagen werden mit höchst kompliziertem, ausgeklügeltem mathematischen Werkzeug erstellt. Vieles davon wurde erst in den letzten Jahrzehnten entwickelt. Die Entwicklung ist weiterhin stürmisch. Die Erhebung numerischer Daten selbst ist dabei häufig einfach, wenn man einmal von Volkszählungen absieht. Aber sowohl die Auswahl der Daten als auch ihre Auswertung, die ja zu gewissen Vorhersagen führen soll, ist voll von Fallstricken, die nur der Fachmann mit einigermaßen Aussicht auf Erfolg vermeiden kann. Einige solche Fallstricke, die schon in den einfachsten alltäglichen Beispielen vorkommen, möchte ich Ihnen jetzt vorführen.

Bis ins 18. Jahrhundert verstand man die Zusammenhänge kaum. Und erst in unserem Jahrhundert hat die Wahrscheinlichkeitsrechnung und die Statistik einen rasanten Aufschwung genommen. Betrachten Sie z.B. das Würfelspiel. Spielen wir es wie Roulette und nehmen wir an, wir hätten einen guten Würfel ohne Unwucht, ohne Zinken und einen fairen Bankier. Wir setzen auf eine Zahl zwischen Eins und Sechs, die Bank würfelt, und wir bekommen den 6-fachen Einsatz, wenn unsere Zahl geworfen wird. Spricht etwas dagegen, auf die 3 zu setzen? Wohl nicht. Wie entscheiden sie jedoch, wenn Sie gesehen haben, daß bei den letzten vier Spielen jedesmal eine 3 gefallen war. Können wir dann noch eine 3 setzen oder besser nicht? Hat sich die Wahrscheinlichkeit von 1/6 für den nächsten Wurf verändert? Verschlechtert? Nein, sie bleibt weiterhin 1/6 und wir können weiter auf die 3 setzen. Der Würfel hat kein Gedächtnis. Insbesondere - eines der am häufigsten anzutreffenden Mißverständnisse - ist die Auffassung unhaltbar, daß die gewissermaßen zuviel gewürfelten 3-er später wieder ausgeglichen werden dadurch, daß etwas weniger 3-er im Durchschnitt fallen. Das sogenannte Gesetz der großen Zahl sagt soetwas ähnliches voraus, aber eben nicht genau das. Soetwas nennt man auch die Unabhängigkeit von Ereignissen. Zunächst einmal ist das nur eine Grundannahme für das mathematische Modell. Es hat sich aber bei allen Untersuchungen herausgestellt, daß diese Annahme mit den Ereignissen, die wir erfassen können, immer übereinstimmt.

Insofern ist auch eine viel zitierte Äußerung des Sprechers Bob Phillips des britischen Atomkraftwerks Chapel Cross völliger Unsinn. Er sagte im Dezember 1988 nach dem Sprengstoffattentat auf den PAN AM Jumbo Jet über England: "Es mag eine etwas zynische Betrachtungsweise sein, aber die statistische Wahrscheinlichkeit, daß ein Flugzeug auf Chapel Cross abstürzt, ist noch geringer als vorher." Die Wahrscheinlichkeit würde sich wohl ändern, wenn man die Flugzeuge besser überwachen und auf andere Routen senden würde. Aber auf eine solche Änderung nahm Phillips gar nicht Bezug. Hier handelt es sich schlicht und einfach um eine fehlerhafte statistische Aussage: "Wenn im statistischen Durchschnitt alle 10 000 Jahre ein Flugzeug im Radius von 1km um ein Atomkraftwerk herunterfällt, und wenn jetzt gerade eines runtergefallen ist, dann haben wir weitere 10 000 Jahre Zeit, bis das nächste herunterkommt." Oder: "Wenn die 3 im statistischen Durchschnitt jeden 6. Wurf fällt, und gerade die 3 geworfen wurde, dann können wir jetzt erstmal 5 Würfe abwarten, bis wieder eine 3 fällt." Ganz im Gegenteil sind die nächsten Ereignisse unabhängig von dem vorhergehenden, die Aussage selbst ist total unsinnig. Eine Änderung der Wahrscheinlichkeit tritt nur ein, wenn in das Gesamtsystem eingriffen wird, z.B. durch Austauschen des Würfels gegen einen, auf dem es keine 3 gibt, oder durch Umlegen der Flugrouten, so daß niemals ein Flugzeug näher als 10 km an das Kraftwerk herankommt, z.B. durch Fesselballons gesichert oder durch andere Maßnahmen.

Hier habe ich besondere Würfel, die mit den Punktzahlen

Würfel 1:	2,2,2,2,6,6
Würfel 2:	1,1,1,5,5,5
Würfel 3:	0,0,4,4,4,4
Würfel 4:	3,3,3,3,3,3

besetzt sind. Der Würfel 1 hat gegen den Würfel 2 die Gewinnchance 2:1 beim nächsten Wurf die höhere Punktzahl zu erreichen, ebenso der Würfel 2 gegen den Würfel 3 und Würfel 3 gegen den Würfel 4. Aber auch der Würfel 4 hat 2:1 Chancen nunmehr gegen den Würfel 1. Jeder ist eingeladen, das auszuprobieren. Eine gegen die Intuition verstoßende Gesetzmäßigkeit in der Wahrscheinlichkeitsrechnung. Es gibt viele solche.
Ich will in der verbleibenden Zeit zwei der ganz wichtigen Fallstricke darstellen. Der eine basiert auf statistischen Erhebungen mit zu kleinen Zahlen. Darunter fällt das eingangs erwähnte Beispiel.

Einfacher als das Würfeln ist das Werfen einer Münze. Kopf oder Zahl haben je die Chance von 50% zu fallen. Zweimal Kopf oder zweimal Zahl hat dann die Wahrscheinlichkeit von 25%. Damit ist es sehr wahrscheinlich, daß der Kopf auch zweimal, ja auch dreimal, viermal, usf, fällt. Serien müssen also sogar auftreten, natürlich seltener als Einzelwürfe, aber doch so häufig, daß man sich nach Auftreten einer Serie Gedanken macht über den Ausgang der nächsten Würfe. Solche Häufigkeiten treten auch bei Ereignissen im täglichen Leben auf. So sind z.B. bei der Statistik von Geburten in Krankenhäusern an manchen Tagen Häufungen von Buben-Geburten oder von Mädchen-Geburten zu beobachten, obwohl die Wahrscheinlichkeit sagt, daß im Mittel etwa genauso viele Buben wie Mädchen geboren werden.

Eine höchst interessante Frage ist es nun, ob es etwa in kleinen Krankenhäusern oder in großen öfter vorkommt, daß an einem Tage sagen wir mit mehr als 75% der Geburten Buben sind. Zunächst mag man keinen Zusammenhang mit der Größe des Krankenhauses sehen. Wenn man dann aber an Extrem-Fälle denkt, wie etwa Krankenhäuser, in denen es jeden Tag etwa zwei Geburten gibt, dann wird klar, was da passieren kann. In einem Viertel aller Tage kommen dann dort nämlich 2 Buben zur Welt und keine Mädchen. In einem weiteren Viertel aller Tage sind es zwei Mädchen. Bei großen Krankenhäusern mit sagen wir 50 Geburten täglich ist es sehr unwahrscheinlich, daß alle Geburten oder auch nur 75% Buben sind. Solche Häufungen treten also vor allem bei kleinen Proben, wie man sagt, auf. Hier ist eine Liste für solche Geburtsraten.

Es werden mehr als 75% Buben geboren

Geburtenrate	jeden...Tag	%
1	2	50.000
2	4	25.000
3	8	12.500
4	3	31.250
5	5	18.750
6	9	10.938
7	16	6.250
8	6	14.453
9	11	8.984
10	18	5.469
11	30	3.271
12	13	7.300
13	21	4.614
14	34	2.869
15	56	1.758
16	26	3.841
17	40	2.452
18	64	1.544
19	104	0.961
20	48	2.069
21	75	1.330
22	118	0.845

Und solche kleinen Proben liegen oft bei Untersuchungen z.B, über die Anzahl der Krebserkrankungen in der Nähe von Kernkraftwerken vor. Die Auswertung solcher Untersuchungen führt sogar unter den statistisch versierten Wissenschaftlern immer wieder zu heftigen Kontroversen.
Unser Beispiel kann in ähnlicher Form bei statistischen Untersuchungen der Krebshäufigkeit in der Nähe von Kernkraftwerken auftreten. Nehmen wir als Beispiel die Leukämiefälle bei Kindern unter 6 Jahren. Setzen wir voraus, daß spontan, d.h. ohne äußer Einflüsse etwa 2% Leukämiefälle auftreten. Uns interessiert nun, wie häufig es ohne äußere Einflüsse vorkommen kann, daß in einem Ort mehr als 10% Leukämiefälle bei den Kindern unter 6 Jahren beobachtet werden. Das wären 500% der üblichen Rate.

Hier ist unsere Tabelle dafür.

Orte mit über 10% Leukämie-Fällen*

Orte mit n Kindern	betroffene Orte 1 von	%
1	50	2.0000
2	25	3.9600
3	17	5.8808
5	10	9.6079
7	8	13.1874
9	6	16.6252
10	62	1.6178
15	28	3.5338
19	18	5.4616
20	141	0.7069
30	346	0.2893
40	851	0.1175
50	2091	0.0478
60	5119	0.0195
80	30312	0.0033
100	177111	0.0006
120	1024033	0.0001
150	13990568	0.0000

*Angenommene Leukämie-Rate 2%

Ganz in der Nähe von Kernkraftwerken liegen häufiger kleine Orte mit unter 20 Kindern bis 6 Jahre. Weiterweg liegen Orte mit 50 bis 150 Kindern dieser Alterstufe. In der Nähe gibt es nun erschreckend viele Orte (ca. 5%), bei denen die Leukämierate von 10% überschritten wird, weiter weg tritt das nur noch in 0,01% aller Orte auf. Ein Zeichen für radioktive Belastung der Bevölkerung und Einfluß auf Leukämie? Nein, eine statistische Notwendigkeit. Sie sehen, wie dramatisch erhöht die Schadensrate sein muß, z.B. daß 30% aller Orte Leukämieraten über 10% haben, um nur sagen zu können, daß das Ergebnis ungewöhnlich ist. Dann aber ist immer noch kein ursächlicher Zusammenhang hergestellt. All das ist nur die Folge, daß kleine Proben betrachtet werden.

Hierunter dürfte sicherlich auch die anfangs zitierte Meldung der Süddeutschen Zeitung fallen. Ich will hier nicht dafür plädieren, in solchen Fällen garnichts zu tun. Eine Suche nach möglichen anderen Ursachen einer solchen Beobachtung ist mit angemessenem Aufwand vertretbar. Wenn man aber aus der Statistik das notwendige Auftreten solcher Häufungen in kleinen Gemeinden kennt, dann darf man doch mit einer solchen Meldung die Bevölkerung nicht grundlos alarmieren. Ich überlasse den Vertretern der Medien zu entscheiden, ob man in solchen, immer wieder auftretenden Fällen die Bevölkerung mit einer sorgfältigen Darstellung dieser statistischen Gesetzmäßigkeit informiert, oder aber; beschließen sollte, daß ein solcher absolut normaler und natürlicher Vorgang keine Erwähnung in den Medien finden sollte. Hier ist nicht die Verantwortung der Techniker, der Wissenschaftler, der Ärzte gefragt, sondern die Verantwortung der Medien.

Ähnlich sind Tests auf Krankheiten, z.B. Aids, zu bewerten. Wenn ein HIV-Test (angenommene Zahlen) mit 97% Sicherheit die richtige Antwort gibt und wir etwa 300.000 HIV-Infizierte in der Bundesrepublik haben, so gäbe dieser Test nur für 9.000 die falsche Antwort "nicht infiziert". Bei der nicht-infizierten Bevölkerung von 60.000.000 Bundesbürgern jedoch müßte man bei diesem Test etwa 1.800.000 fälschlich als infiziert bezeichnete Personen finden, vielmehr als es überhaupt an HIV-Infizierten gibt.

III) Bedingte Wahrscheinlichkeiten und ihre Fallstricke

In den oben besprochenen Beispielen treten die Fehler noch kontrollierbar und wegen kleiner Testmengen auf. Schwieriger ist die Bewertung von Fehlern der folgenden Art. Die New York Times berichtete vor einigen Jahren - und das ging dann auch noch über einige Presseagenturen - über eine Untersuchung auf die Lese- und Schreibfähigkeit der New Yorker. Man hatte dabei einen verblüffenden und sehr aussagekräftigen Zusammenhang - eine Korrelation - zwischen der Schuhgröße und der Rechtschreibfähigkeit entdeckt: je größer die Schuhgröße der getesteten Personen war, desto besser und fehlerfreier konnten sie lesen und schreiben. Die Nachricht hat damals großes Rätselraten und große Zweifel ausgelöst. Die Wissenschaftler wurden mit allerlei Nachfragen behelligt. Was war passiert? Der betreffende Reporter hatte die ursprüngliche Nachricht gekürzt. In ihr war vermerkt gewesen, daß die Untersuchung an den New Yorker Schulen bei den Schulkindern erfolgt war. Jetzt ist wohl jeder von Ihnen bereit, das merkwürdige Ergebnis zu akzeptieren, denn beides, die Schuhgröße und die Rechtschreibfähigkeit, hängt natürlich von Alter der Kinder ab.

So kann man immer wieder statistische Erhebungen finden, die Zusammenhänge zwischen zwei erhobenen Größen feststellen und damit einen ursächlichen Zusammenhang zwischen den Größen suggerieren. Häufig liegt aber nur eine bisher oder in der Untersuchung unbekannte gemeinsame Ursache für beides vor. Deshalb war oder ist man sich in der Medizin auch immer wieder unsicher über sogenannte Risikofaktoren, z.B. Cholesterin für Herzinfarkte, Rauchen für Lungenkrebs. Ein Mittel wie Aspirin, das - so heißt es - bei regelmäßiger Einnahme erhebliche Schutzwirkungen gegen den Herzinfarkt und andere Kreislaufkrankheiten hat, muß z.B. notgedrungen "Krebs erzeugen", so würde man aus der Statistik ersehen. Warum? Weil die Rate der Krebstoten automatisch steigen würde, wenn man den Prozentsatz der Herz-Kreislauftoten deutlich senken könnte. Die Menschen würden etwas länger leben und dann einer anderen Todesursache, unter anderen also auch dem Krebs, zum Opfer fallen.

Arbeitslosen-Zahlen*

	Akademiker		Nicht-Akademiker	
männlich	120/3.000	4%	200/10.000	2%
weiblich	40/200	20%	2.400/16.000	15%
total	160/3.200	5%	2.600/26.000	10%

*Angaben in 1000 bzw. Prozent, erste Angabe ist Zahl der Arbeitslosen, zweite Angabe ist Zahl der Beschäftigten.

Ein beeindruckendes Beispiel für eine solche Verfälschung von statistischen Aussagen ist eine (fiktive) Arbeitslosenstatistik. Bei Gesamtbewertung der Statistik (letzte Zeile) muß man zu dem Schluß kommen, daß ein Studium das Risiko der Arbeitslosigkeit erheblich verringert. Ist die Statistik jedoch aufgeteilt in männliche und weibliche Arbeitnehmer, so erhöht ein Studium das Risiko der Arbeitslosigkeit in beiden Gruppen. Man könnte also sagen, daß ein Studium für denjenigen besonders empfehlenswert ist, der nicht weiß, ob er Männlein oder Weiblein ist. Ein zugegebenermaßen auch für den Fachmann immer wieder verblüffender Widerspruch. Nehmen sie sich ruhig die Zeit, die Zahlen zu vergleichen und evtl. auch nachzurechnen. Sie werden wohl schnell erkennen, daß eine solche Statistik, mehr noch als die zuvor gezeigten, alle Möglichkeiten zur Manipulation enthält. Je nach politischem oder soziologischem Standort wird man sich die eine oder die andere Aussage für seine Argumentation auswählen. Aber was ist hier tatsächlich geschehen? Die Rechnung ist ganz offensichtlich richtig. Es handelt sich um die Interpretation dieser Daten, also um die Übereinstimmung des mathematischen Modells mit den Daten der Realität. Man hat die Erhebungen in Zahlen gemacht und diese dann umgerechnet in Prozentsätze. Soweit ist alles in ordnung. Die Frage ist nur, was man aus diesen Prozentsätzen schließen soll. Nun, der Schluß eines guten Statistikers ist, daß das Vorkommen von Akademikern und Nicht Akademikern stark von dem Geschlecht abhängt, daß also die beiden Unterscheidungsmerkmale Akademiker/Nicht-Akademiker und männlich/weiblich abhängige statistische Größen sind. Bei genauerem Hinsehen bemerkt man dann auch sehr schnell, daß es bei den weiblichen Beschäftigten nur einen verschwindend kleinen Anteil an Akademikern gibt. Hingegen sind sie bei den Nicht-Akademikern deutlich in der Überzahl. Solche in sich widerspruchsvollen Statistiken ergeben sich häufig, wenn man abhängige Unterscheidungsmerkmale verwendet.

Ein weiteres Beispiel, das ich von einem englischen Kollegen habe, ist eine Lungenkrebs-Statistik, die zur Kontrolle verschiedener Behandlungsmethoden in verschiedenen Kliniken erstellt wurde. Für einen Lungenkrebskranken ist ganz offensichtlich die Behandlung A weitaus erfolgversprechender, als die Behandlung B. Schaut man sich jedoch die Heilungs-Chancen aufgeteilt nach Rauchern und Nichtrauchern an, so ist in beiden Fällen die Behandlung B erfolgreicher. Auch hier sieht man wieder beim zweiten Hinsehen, daß die Behandlung B hauptsächlich an Rauchern und die Behandlung A hauptsächlich an Nichtrauchern angewendet wurde. Wegen dieser Abhängigkeit der Unterscheidungsmerkmale ist der Widerspruch möglich geworden. Es bedarf schon eines ausgebildeten Fachmanns, um bei weniger deutlichen statistischen Daten festzustellen, ob sie für eine Voraussage der weiteren Entwicklung (oder Wirkung) brauchbar oder signifikant sind. Die Aussagen über die Wirkungen der Behandlungen A bzw. B bei den einzelnen Gruppen sind weitgehend unbrauchbar. Die totale Aussage ist dadurch auch betroffen, aber etwas weniger belastet von dem Fehler, der bei der Erhebung gemacht wurde.

Lungenkrebsstatistik*

	Behandlung A		Behandlung B	
Raucher	5/100	5%	101/900	11%
Nichtraucher	400/800	50%	196/200	98%
total	405/900	45%	297/1100	27%

*Erste Angabe ist Zahl der geheilten Fälle, zweite Angabe ist die Gesamtzahl der erfaßten Lungenkrebserkrankungen

Ebenso ist die Aussage der Arbeitslosenstatistik zumindest für die weiblichen Arbeitnehmer weitgehend unbrauchbar. Wieder ist die totale Aussage zwar verzerrt, aber doch in gewissem Umfang und mit gehöriger Vorsicht noch brauchbar. Der Fachmann hat hier weitere Methoden, um die Tragkraft, das Gewicht einer Aussage der Form: "Akademiker haben nur eine halb so große Arbeitslosenrate, wie Nicht-Akademiker" abzuschätzen. Er kann die Wahrscheinlichkeit einer solchen Aussage berechnen. Darauf will ich aber nicht näher eingegehen. Mir scheint wichtig zu sein festzuhalten, daß die Zahlen in Statistiken im allgemeinen stimmen, daß ihre Interpretation jedoch ungemein schwierig, wenn nicht gar in vielen Fällen unmöglich ist und dann nur vom Experten vorgenommen werden kann. Verfälschungen können einerseits dadurch auftreten, daß man nur eine ganz kleine Zahl von Fällen zur Verfügung hat, andererseits, daß man voneinander deutlich abhängige Unterscheidungsmerkmale für die statistische Erfassung vorliegen hat. Häufig bleibt die Abhängigkeit unerkannt. Man wundert sich dann nur, daß die Vorhersagen nicht mit den früheren Messungen übereinstimmen. Solche abhängigen Größen haben wir auch schon bei der Diskussion der Schuhgröße und Rechtschreibfähigkeit der New Yorker Schüler kennengelernt. Beide Größen bedingen sich nämlich nicht etwa gegenseitig, sondern sind von einer dritten mehr zufällig erkannten Größe, dem Alter, abhängig. Ich kann hier nur vor der selbständigen Interpretation von Prozentzahlen warnen. Wenn wirklich wesentliche Entscheidungen davon abhängen, dann sollte man auf jeden Fall einen Fachmann befragen.

Das ist auch das Dilemma, in dem wir heute immer wieder bei Expertenbefragungen stehen. Da geht es um mathematische Modelle, um Auslegung von rechnerischen Ergebnissen durch Experten und selbst ernannte Experten. Spitzfindigkeiten, wie die obigen Beispiele, werden von den Politikern und Richtern z.T. nicht erfaßt, es können sich Interpretationen des Zahlenmaterials durchsetzen, die z.B. von entsprechend ausgebildeten Personen leicht als irrig erkannt werden. Aber der Nachweis dieses Irrtums ist kompliziert, technisch, aufwenidig, und wird nur selten akzeptiert, wegen mangelnder Vorbildung.