Problemecke

von Manfred Borovcnik, Klagenfurt

Eigenartiges Denken

In der amerikanischen Fernsehshow 'Let's make a deal' darf der Kandidat, der die Endrunde erreicht, aus drei Boxen mit den Aufschriften A, B und C wählen. Eine enthält den Schlüssel für einen Lincoln Continental, die zwei anderen sind leer. Der Kandidat wählt Box B; bevor er sie öffnet, bietet ihm der Moderator Geld für die Box an. Es gibt ja zwei weitere Boxen, in der Schlüssel sein kann, die Chancen für den Kandidaten stehen nur 1:2 für ihn. Der Moderator erhöht sein Angebot bis auf $ 500, der Kandidat zögert. Der Moderator erinnert an die beiden anderen Boxen. Die Wahrscheinlichkeit, daß der Kandidat schon den Schlüssel zum Auto hat, beträgt danach nur 1/3. Die Zuhörer rufen in Aufregung "Nein!". Der Kandidat behält die Box. Nun ändert der Moderator seine Taktik und öffnet eine der verbleibenden Boxen, sagen wir die Box A auf dem Tisch. Sie ist leer! Das Publikum applaudiert. Der Moderator: "Es sind nun nur mehr zwei Boxen übrig, die Wahrscheinlichkeit, daß Sie das Auto schon haben, beträgt nun 1/2. Ich gebe Ihnen $ 1000 für Ihre Box." Zur Überraschung aller bietet der Kandidat jedoch seine Box gegen die am Tisch verbliebene an.

Diese überraschende Wende führte schon 1975 zu einer heftigen Diskussion. Wer hat recht, der Moderator oder der Kandidat? Heuer wurde das Thema aufgewärmt als derselbe Moderator in seiner Show das Spiel wieder einführte. Nun verbirgt sich hinter einer von drei Holztüren ein Auto, hinter den beiden anderen eine lebende Ziege. Nach der ersten Wahl öffnet der Moderator heute eine der verbleibenden Türen direkt (natürlich meckert eine Ziege dahinter) und bietet dem Kandidaten an, seine Wahl zurückzunehmen und die andere, noch verbliebene Tür zu nehmen. In den USA nahm Marylin vos Savant, die wegen ihres außerordentlich hohen Intelligenzquotienten berühmt ist, in ihrer Zeitungskolumne dazu Stellung: "Wechseln hilft." Daraufhin ist

Stochastik in der Schule 11 (1991), 46-51.

die Diskussion erneut voll aufgeflammt und in diesem Sommer bis nach Deutschland und Österreich getragen worden. Ein Artikel im *Spiegel* (Nr.34, 212-214) stellt zwar die Dinge klar, hat jedoch regen Widerspruch gefunden.

Was macht Beispiele aus der elementaren Wahrscheinlichkeitsrechnung so schwierig, daß selbst nach Darstellung der formal richtigen Lösung so viele Zweifel bestehen bleiben? Oft überzeugt nicht einmal ein Simulationsexperiment. Die kombinatorische Lösung gibt einen Überblick über alle Möglichkeiten: Abzählen liefert 2/3 für den Erfolg beim Wechseln.

Schlüssel sind in	Kandidat wählt	Moderator öffnet	Kandidat wechselt zu	Entscheidung
A	A	B oder C	B oder C	falsch
A	B	C	A	richtig
A	C	B	A	richtig
B	A	C	B	richtig
B	B	A oder C	A oder C	falsch
B	C	A	B	richtig
C	A	B	C	richtig
C	B	A	C	richtig
C	C	A oder B	A oder B	falsch

Tab.1

Man kann auch Baumdiagramme zur Lösung heranziehen, die Boxen seien mit S für Schlüssel bzw. L für leer bezeichnet. Die erste Stufe ist die Wahl des Kandidaten, die zweite beschreibt, welche Box der Quizmaster öffnet. Das Ergebnis ist dasselbe, wenn der Moderator die andere leere Box L_2 öffnet, was der Kandidat aber nicht unterscheiden kann.

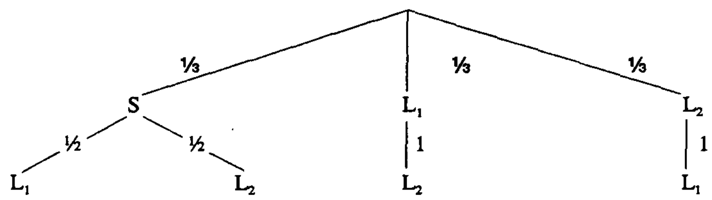

Fig.1

$$W(A|L_1) = \frac{\frac{1}{3}\cdot\frac{1}{2}}{\frac{1}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot 1} = \frac{1}{1+2} = \frac{1}{3}$$

Die Wahrscheinlichkeit für den Kandidaten, die richtige Wahl getroffen zu haben, ist daher 1/3 wie zuvor, daran hat sich durch die geöffnete Box nichts verändert! Allerdings bleibt die Restwahrscheinlichkeit von 2/3 an der noch nicht geöffneten Tür hängen, es ist also tatsächlich vorteilhaft, wenn der Kandidat tauschen kann. Einerseits läßt die Information über die geöffnete Box die Wahrscheinlichkeit, daß die bereits gewählte den Schlüssel enthält, unverändert; sie ist also irrelevant. Andererseits aber läßt diese Information die Wahrscheinlichkeit, daß die noch nicht geöffnete Türe die richtige ist, von 1/3 auf 2/3 ansteigen?! Und warum ist die Lösung 1/2 für jede der beiden noch nicht geöffneten Boxen falsch? Die formale Lösung glaubt man lange nicht, weil sie keine Einsicht bietet. Sie erscheint uns so paradox, weil wir nicht gelernt haben, mit Informationen und wie diese unseren Kenntnisstand in Form von Wahrscheinlichkeiten verändern umzugehen.

Dabei treffen wir recht häufig auf solche Situationen, etwa wenn es um die Beurteilung geht, ob der eine oder andere von zwei Zuständen zutrifft (krank oder gesund, schuldig oder unschuldig). Weitere Informationen sollen dies klären (Blutuntersuchungen, Indizien). Wenn ein Beobachtungsbefund wesentlich stärker unter der einen Möglichkeit vorkommt, so läßt er entsprechend die Wahrscheinlichkeit davon größer werden. (Ein Blutbefund, der bei einer gewissen Krankheit stark gehäuft vorkommt, läßt diese entsprechend wahrscheinlicher werden.)

Dieser natürliche Umgang mit dem stochastischen Wert einer Information kann auch in diesem Beispiel klären, was eigentlich vor sich geht. Der Zustand einer Box ist S oder L, sie hat den Schlüssel oder ist leer. Über jede der drei Boxen hat man anfangs die Wahrscheinlichkeit 1/3 für S. Als weiteren Befund bekommt man über die *Box des Kandidaten*, bezeichnen wir sie mit B, nichts herein, denn der Moderator konnte sie weder öffnen noch konnte er sie bei dieser Auswahl unberücksichtigt lassen. Die geöffnete Box, be-

zeichnen wir sie mit A, sagt nichts über die Wahl des Kandidaten aus. Man hat folgendes Diagramm:

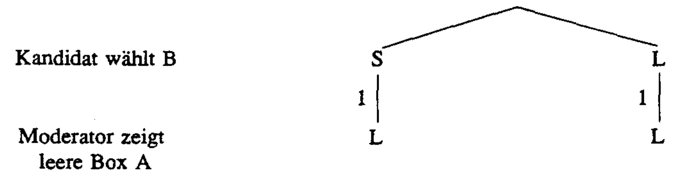

Fig.2

Geht es jedoch um die *am Tisch verbliebene Box* C, so konnte die bei der Auswahl des Moderators berücksichtigt werden oder nicht. Über sie sagt der Umstand, daß sie nicht gewählt wurde, sehr wohl etwas aus. Sie könnte aus dem Grund nicht gewählt worden sein, weil sie tatsächlich den Schlüssel enthält. Das Diagramm sieht so aus:

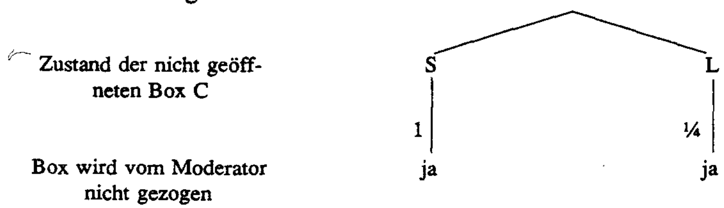

Fig.3

Wenn Box C leer ist, hat man folgende zwei Fälle: Box B des Kandidaten enthält den Schlüssel, dann hat der Moderator freie Wahl zwischen A und C; mit ½ nimmt er Box C *nicht*. Ist jedoch der Schlüssel in der anderen Box, nämlich in A, so muß er C wählen; er 'nimmt mit Wahrscheinlichkeit 0 Box C nicht'. Insgesamt ergibt das ½.½ + 0 = ¼ dafür, daß der Moderator Box C nicht nimmt. Eine weitere Schwierigkeit mag darin liegen, daß man hier auf den Zeitpunkt zurückgehen muß, bevor der Moderator Box C bei seiner Wahl nicht berücksichtigt hat. Nun zeigen die bedingten Wahrscheinlichkeiten, daß der Umstand 'wurde nicht gezogen' ein deutliches Indiz dafür ist, daß Box C den Schlüssel enthält. Es ging hier darum, das Ergebnis der formalen Rechnung einsichtig zu machen.

Gibt es eine Variante des Beispiels, in der es vorteilhafter ist, bei der ersten Wahl zu verbleiben. Verwirrend? Ein bißchen Verweilen und Knobeln hilft.

Überraschungen und Koinzidenzen

Wenn sich zwei zunächst verschiedene Dinge zur selben Zeit ereignen, ist die Überraschung groß. Mit solchen Koinzidenzen sind viele Fehleinschätzungen verbunden, diese bilden eine große Hürde bei der Akzeptanz von Wahrscheinlichkeit. Zumeist wird einfach übersehen, wie oft ein entsprechendes 'Experiment' stattfinden konnte, welches dann mit 'Koinzidenz' oder 'Unauffälliges' endet. Mit einer kleinen Wahrscheinlichkeit und einer großen Anzahl von Versuchen gibt es aber doch eine entsprechende Zahl von erwarteten Koinzidenzen. In Heft 3 (1990) wurde über eine solche Übereinstimmung von Autokennzeichen berichtet: "Snap! New car is a 'twin'". Im englischen System ist das Format *XYZ 123 A*; *X* ist ein frei wählbarer Buchstabe, *123* eine dreistellige Zahl; *A* bezeichnet das Zulassungsjahr, *YZ* kennzeichnet die Region der Registrierung.

Ohne diese Einschränkungen zu beachten, wurde in einem Zeitungsartikel die Wahrscheinlichkeit von 1 durch 1 Million für das Übereinstimmen von Kennzeichen berechnet. Es war die Rede davon, daß jemand beim Kauf eines gebrauchten Zweitwagens eine solche Übereinstimmung erlebte. Im genannten Artikel wurden diese Einschränkungen berücksichtigt, was eine Wahrscheinlichkeit von 1 durch 250000 ergab.

Es gibt jedoch weitere realistische Einschränkungen, die diese Koinzidenz noch wesentlich wahrscheinlicher machen. Die Buchstabenkombination *YZ* gibt genauer den Bezirk oder die Stadt der Registrierung an. Verkauf und Kauf eines Gebrauchtwagens erfolgt vielleicht eher in derselben Region als auswärts; dafür gibt es gute Gründe (Vertrauen z.B.). Man könnte davon ausgehen, daß wenigstens 60% der angebotenen PKWs aus demselben Bezirk stammen, also dieselbe Kombination *YZ* haben.

Für diese 60% ist aber folgendes plausibel: Nehmen wir an, daß im Jahr 5000 Autos registriert werden und die Kennzeichen nicht durch Zufall erzeugt sondern in Serie vergeben werden. Dies möge so erfolgen, daß alle Ziffern aber nur 5 der möglichen Buchstaben ausgenützt werden. Für Autos aus demselben Bezirk hat man eine Wahrscheinlichkeit von 1/5000 für das

gleiche Kennzeichen (ausgenommen das Registrierungsjahr). Insgesamt bläst sich die Koinzidenz auf zu

$$0.60 \cdot \frac{1}{5000} + 0.4 \cdot \frac{1}{250000} = 0.00012$$

Der Einfachheit halber wurde über die Vergabe von Nummern in anderen Bezirken nichts angenommen, sie möge zufällig sein. Die Gewichte von 0.6 und 0.4 erklären sich aus dem Anteil 'heimischer' und 'fremder' Autos beim Gebrauchtwagenhändler. In etwa einem von 10000 Fällen ist daher eine Nummerngleichheit zu erwarten. Die Übereinstimmung ist also um eine ganze Zehnerpotenz kleiner. Es sollte hin und wieder möglich sein, über solche Fälle zu berichten. Wenn Sie ein Beispiel zu Koinzidenzen haben, das sich bei näherer Analyse auch 'normalisiert', so berichten Sie uns doch darüber.

Paradoxe Mittelwerte

Die Reihe der Puzzles soll mit einer Aufgabe (von O. Anderson aus *Teach. Stat.* 13, 1991) an Sie, liebe Leserin, lieber Leser, abgerundet werden. An der Universität von Western Ontario werden einige Klassen in Statistik für das erste Studienjahr angeboten. Einige dieser Kurse haben mehr als 100 Hörer. Die durchschnittliche Größe der Klassen, so wie die Studenten sie erleben, ist 152. Die mittlere Zahl der Studenten, die Lehrenden vor sich haben, ist jedoch nur 32 (mit einer Standardabweichung von 64; bei einem Puzzle wird wie bei Zauberern das Wichtigste so nebenbei gesagt).

Ehrenwort, alle Mittelwerte sind arithmetische Mittelwerte, nicht Mediane oder so. Das Beispiel geht auf, auch wenn es auf den ersten Blick noch so verrückt klingt. Wie viele Klassen gibt es an dieser Universität? Schicken Sie Ihre Lösung an uns. Die ersten zwei richtigen Antworten werden mit einem Jahresabonnement *Stochastik in der Schule* belohnt.