Anmerkungen zum Testen von Hypothesen

MANFRED BUTH, Hamburg

Zusammenfassung

Signifikanztests, so die im Folgenden vertretene These, werden durchgeführt, wenn man fast keine Information über die vorliegende Situation hat, mit der Folge, dass man sich mit Antworten auf falsch gestellte Fragen zufrieden geben muss, aber auf die eigentlich interessierenden Fragen keine Antwort erhält. Der Schlüssel zur Vermeidung der üblichen Fehler im Unterricht könnte in dem Hinweis liegen, dass Schülerinnen und Schüler dazu neigen, die richtigen Fragen zu stellen.

1 Zur Einführung

Bei der didaktischen Analyse eines möglichen Unterrichtsinhalts sollte die Erörterung der methodischen Einzelheiten erst dann beginnen, wenn die Sache selbst hinreichend klar ist. Diese allgemeine Regel erscheint beim Thema 'Signifikanztests' besonders angebracht. Denn trotz der umfangreichen Literatur dürfte noch immer nicht hinreichend geklärt sein, was beim Testen von Hypothesen eigentlich gespielt wird. Nach Ansicht des Autors besteht ein wesentliches Merkmal der Signifikanztests darin, dass man zwar Antworten auf falsch gestellte Fragen erhält, aber auf die eigentlich interessierenden Fragen keine Antwort zu erwarten hat. Das geschieht jedoch nicht aus bösem Willen, sondern eher aus Notwehr, weil aufgrund der vorliegenden Information nicht mehr zu holen ist. Sollte diese Ansicht zutreffen, dann ergibt sich daraus eine Folgerung über angebliche Schülerfehler: Die Schwierigkeiten mit dem Testen von Hypothesen im Unterricht kommen dadurch zustande, dass Schülerinnen und Schüler dazu neigen, die richtigen Fragen zu stellen, und darin durch die Bezeichnung ‚Testen von Hypothesen' noch bestärkt werden. Man muss ihnen deshalb erklären, warum die Welt nicht so einfach ist, wie sie glauben, und dass sie sich folglich mit viel bescheideneren Erkenntnissen zufrieden geben müssen.

Der nächste Abschnitt enthält eine Erläuterung der These an einem Beispiel. Sie ist so ausführlich gehalten, dass sie als exemplarische Begründung der These gelten kann. Im dritten Abschnitt wird unter Verweis auf schon vorliegende Analysen daran erinnert, dass zwar auch der Satz von Bayes keinen Ausweg aus dem Dilemma bietet, wohl aber der Begriff der bedingten Wahrscheinlichkeit ein geeignetes Mittel bildet, das Problem zu analysieren. Erst im vierten Abschnitt werden didaktische Folgerungen gezogen und Anregungen für die methodische Gestaltung des Unterrichts gegeben.

2 Erläuterung der These an einem Beispiel

In der Einführung war behauptet worden, dass man beim Testen von Hypothesen richtige Antworten auf falsch gestellte Fragen bekommt, aber auf die eigentlich interessierenden Fragen keine Antwort erhält. Um das an einem Beispiel zu erläutern, sei eine endliche Anzahl von Urnen gegeben, die mit schwarzen und weißen Kugeln gefüllt sind. Außer der Farbe sollen an den Kugeln keine unterscheidenden Merkmale vorhanden sein.

Wenn man eine dieser Urnen herausgreift und wissen möchte, wie groß der Anteil der weißen Kugeln ist, dann besteht die einfachste Lösung des Problems darin, die Urne zu öffnen und auszuzählen, wie viele Kugeln von jeder Farbe darin enthalten sind. Falls dieser Weg nicht gangbar ist, aber wenigstens noch die Verteilungsfunktion vorliegt, die aussagt, welche Mischungsverhältnisse überhaupt vorkommen und wie häufig sie auftreten, dann lassen sich immerhin noch Wahrscheinlichkeiten angeben. Sollte auch dieser Fall nicht mehr vorliegen, dann hat man praktisch keine Information in der Hand und auch ein Signifikanztest kann sie nicht herbei zaubern. Das scheint mir beim Testen von Hypothesen der entscheidende Punkt zu sein.

Will man in der gegebenen Situation doch noch etwas über den Inhalt der Urne in Erfahrung bringen, dann muss man sich mit sehr viel bescheideneren Fragen zufrieden geben, z. B.: Liegt das Mischungsverhältnis 1:1 vor oder nicht?

Nach dieser Reduktion der Fragestellung führen drei Maßnahmen weiter. Erstens sollte man die Frage als Nullhypothese H_0 formulieren: Der Anteil der weißen Kugeln beträgt 50%. Zweitens kann man unter der ausdrücklichen Voraussetzung, dass die Hypothese zutrifft, die Wahrscheinlichkeit w(i, k) dafür berechnen, dass beim k-maligen Ziehen einer Kugel mit Zurücklegen genau i weiße Kugeln erscheinen. Daraus wiederum lässt sich für jede Zahl α zwi-

Stochastik in der Schule 22 (2002) 2 S. 27-29

schen 0 und 1, dem Signifikanzniveau, ein Intervall I(α) konstruieren, das im hier diskutierten Fall des zweiseitigen Tests die folgende Eigenschaft besitzt: Die Wahrscheinlichkeit dafür, dass der Anteil der weißen Kugeln beim Ziehen kleiner ausfällt als der linke Rand l(α) des Intervalls, soll genau $\alpha/2$ sein und ebenso die Wahrscheinlichkeit dafür, dass der Anteil der weißen Kugeln jenseits des rechten Randes r(α) liegt. Schließlich kann man noch als dritte Maßnahme tatsächlich k Kugeln mit Zurücklegen ziehen und den Anteil d(k) der weißen Kugeln bestimmen. Die Information, die in der Kenntnis von d(k) liegt, werde mit D bezeichnet.

Der meines Erachtens völlig normale Denkprozess besteht nun darin, den Schluss von der empirischen Kenntnis D auf die Gültigkeit oder Ungültigkeit der Hypothese H_0 zu versuchen, d. h. nach der bedingten Wahrscheinlichkeit $P(H_0 / D)$ zu fragen. Dieser Weg ist aber nicht gangbar. Es bleibt nur übrig, umgekehrt von der bedingten Wahrscheinlichkeit $P(D / H_0)$ auszugehen, das Intervall I(α) zu konstruieren, d(k) zu bestimmen, um dann in bekannter Manier über die Gültigkeit von H_0 auf dem Signifikanzniveau α zu entscheiden.

Üblicherweise wird die Nullhypothese auf dem Niveau α angenommen, wenn

$l(\alpha) \leq d(k) \leq r(\alpha)$

und anderenfalls verworfen. Nach Konstruktion ist α das Risiko, die Nullhypothese irrtümlich zu verwerfen, obwohl sie in Wirklichkeit zutrifft. Denn α gibt, unter der ausdrücklichen Voraussetzung, dass die Nullhypothese zutrifft, die Wahrscheinlichkeit dafür an, dass d(k) außerhalb des Intervalls I(α) liegt. Über den Fehler zweiter Art kann in diesem Beispiel nichts ausgesagt werden, so lange die Alternativhypothese nur feststellt, dass der Anteil der weißen Kugeln von 50% verschieden ist. Denn daraus lassen sich keine Schlüsse ziehen.

3 Die Rolle des Satzes von Bayes beim Testen von Hypothesen.

Der Satz von Bayes kann ganz allgemein als ein Hilfsmittel verstanden werden, um bedingte Wahrscheinlichkeiten umzukehren, d.h. von $P(A / B)$ und $P(A / \overline{B})$ auf $P(B / A)$ zu schließen. Er scheint daher das ideale Werkzeug zu sein, um von $P(D / H_0)$ und $P(D / \overline{H_0})$ zu $P(H_0 / D)$ zu kom- men. Das aber ist eine Illusion. Denn die Anwendung der Gleichung

$$P(H_0 / D) = \frac{P(D / H_0) P(H_0)}{P(D / H_0) P(H_0) + P(D / \overline{H_0}) P(\overline{H_0})} \quad (1)$$

erfordert die Kenntnis der absoluten Wahrscheinlichkeiten $P(H_0)$ und $P(\overline{H_0})$. Wenn man diese jedoch hätte, dann bräuchte man den ganzen Aufwand mit dem Signifikanztest nicht mehr. Denn $P(H_0)$ gibt die Wahrscheinlichkeit dafür an, dass die Nullhypothese zutrifft.

Nun gibt es allerdings noch ein Verfahren, das die Kenntnis der absoluten Wahrscheinlichkeiten trickreich umgeht. Dabei werden diese zunächst willkürlich festgesetzt. Nach dem Ziehen einer Kugel lässt sich der Satz von Bayes anwenden. Daraus ergeben sich bedingte Wahrscheinlichkeiten, die als neue Schätzwerte für die absoluten Wahrscheinlichkeiten festgesetzt werden und so fort. Nach einer Anzahl von Wiederholungen des Verfahrens kommt man damit zu einer recht guten Einschätzung der Nullhypothese. Einzelheiten finden sich bei Riemer (1985). Analysiert man das Verfahren aber im Einzelnen (vgl. Buth (1996)), dann stellt sich heraus, dass die Verbesserung der Schätzwerte erst bei mehrfachem Ziehen einer Kugel eintritt und dass man in diesem Fall eben so gut die relativen Häufigkeiten heran ziehen kann, um die Gültigkeit der Nullhypothese zu beurteilen.

Insgesamt bietet also der Satz von Bayes bei gleichen Voraussetzungen keine brauchbare Alternative zum Signifikanztest.

4 Didaktische Erwägungen

Aus der vorgelegten Sachanalyse ergeben sich unmittelbar drei Folgerungen für die Behandlung von Signifikanztests im Stochastikunterricht.

Schüler machen keinen Fehler, wenn sie die richtigen Fragen stellen und sich bei der Erörterung einer Hypothese dafür interessieren, ob die Vermutung zutrifft oder nicht. Deshalb muss man den Schülern klar machen, dass sie zwar richtig fragen, aber bei Signifikanztests nicht genügend Information zur Verfügung haben, um Antworten auf vernünftige Fragen erwarten zu können, und dass sie sich deshalb mit Antworten auf weniger interessante Fragen begnügen müssen.

Der Begriff der bedingten Wahrscheinlichkeit ist das geeignete Mittel, um die besondere Situation beim Testen von Hypothesen zutreffend zu beschreiben (vgl. hierzu Krauss & Wassner (2001)) und um durch die Gegenüberstellung von

$P(D / H_0)$ und $P(H_0 / D)$ deutlich zu machen, welche Kenntnis man gern hätte und womit man sich begnügen muss. Auch unabhängig davon gibt es gute Gründe für eine starke Betonung der bedingten Wahrscheinlichkeit im Unterricht. In der Mittelstufe oder in Grundkursen bietet sich das Werfen mit zwei Würfeln an, um den Begriff der bedingten Wahrscheinlichkeit einzuführen, seine Anwendung an Beispielen zu üben und – wenn es denn sein muss – den Satz von Bayes zu behandeln.

Die Stärke des Begriffs 'bedingte Wahrscheinlichkeit' beruht übrigens nicht auf der definierenden Gleichung

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \qquad (2)$$

sondern darauf, dass man für jeden Wahrscheinlichkeitsraum W =(Ω, P, σ) und jedes Ereignis B aus σ mit P(B)>0 einen neuen Wahrscheinlichkeitsraum W' = (Ω', P', σ') mit

$$\Omega' = B, \qquad P' = P_B, \quad \sigma' = \{A' | A' = A \cap B \wedge A \in \sigma\}$$

herstellen kann, der eine Einschränkung von Ω auf B darstellt. Wesentlich ist dabei, dass die durch (2) definierte Funktion tatsächlich eine Wahrscheinlichkeit ist. – Diese Konstruktion wäre ein angemessenes Thema für den Leistungskurs.

Den Satz von Bayes sollte man bei den Signifikanztests aus dem Spiel lassen und ihm auch sonst nicht die Aufmerksamkeit zuwenden, die er weitgehend erfährt (vgl. Buth (1999)).

Lehrer haben mir übrigens in Diskussionen gesagt, dass man im Unterricht meistens nicht mehr bis zum Testen Hypothesen kommt und sich das Thema deshalb von allein erledige.

Quellenangaben

Buth, M. (1996). Schwierigkeiten im Umgang mit dem Zufall - eine didaktisch orientierte Sachanalyse, mathematica didactica 19, Bd. 2, 3

Buth, M. (1999). Ist der Satz von Bayes nützlich oder überflüssig?, in: Beiträge zum Mathematikunterricht S.125

Krauss, S. & Wassner, C. (2001). Wie man das Testen von Hypothesen einführen sollte, Stochastik in der Schule, 21, S. 29

Riemer, W. (1985). Neue Ideen zur Stochastik. Lehrbücher und Monographien zur Didaktik der Mathematik. Band 3. BI Wissenschaftsverlag

Anschrift des Verfassers

Manfred Buth

Von-Melle-Park 8

20146 Hamburg

Vorbemerkungen zum folgenden Aufsatz von Kay Rottmann

HEINZ ALTHOFF, BIELEFELD

Zu den Themen der Wiederholung der Stochastik für die Abiturprüfung meines LK 2001 gehörte auch das *Geburtstagsproblem.* In diesem Rahmen bat ich K. R., seinen Mitschülern den Inhalt des Aufsatzes 'Ein Geburtstagsproblem" von G. SCHRAGE (aus Stochastik in der Schule, Heft 2/1992) darzustellen, in dem es um die Wahrscheinlichkeit für das Zusammentreffen dreier Geburtstage bei 15 Personen geht. K. R. überraschte mich dann damit, dass er im zweiten Teil seines Vortrags das Problem auf "mindestens k gemeinsame Geburtstage bei n Personen" erweiterte. Ich ermunterte ihn daraufhin, den folgenden Aufsatz für unsere Zeitschrift auszuarbeiten. Lehrerinnen und Lehrer sollten den Aufsatz vor allem als anregendes Beispiel dafür betrachten, hinreichend begabte Schüler zu motivieren, eine (nicht zu umfangreiche) mathematische Abhandlung anderen verständlich darzustellen, ja sie vielleicht sogar (wie hier geschehen) zu einer eigenen wissenschaftlichen Auseinandersetzung mit einem mathematischen Problem zu "verführen". Kann der Aufsatz von K. R. eventuell eine Kettenreaktion für so etwas auslösen?