Vom Rückwärtsschließen im Baumdiagramm zum Testen von Hypothesen

HELMUT WIRTHS, OLDENBURG

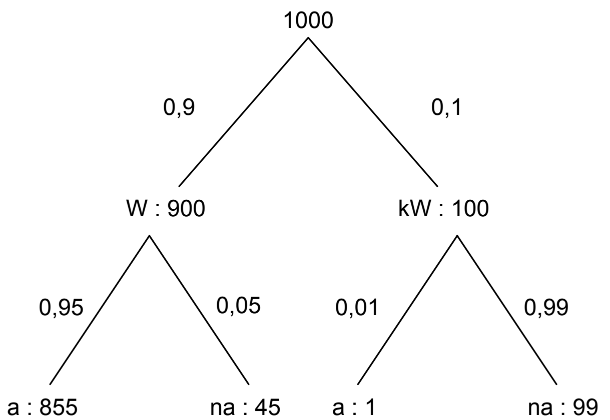
Zusammenfassung: In diesem Beitrag werden mehrfach erprobte Ideen für eine Unterrichtseinheit vorgestellt, die die Bausteine „Rückwärtiges Schließen im Baumdiagramm“ und den Teilbaustein „Alternativtests“ der neuen niedersächsischen Richtlinien für die Klassen 9 und 10 des Gymnasiums miteinander verbindet.

1. Ein Spamfilter

In den neuen niedersächsischen Richtlinien für den Mathematikunterricht am Gymnasium für die Klassen 7 bis 10 (2003) werden Bausteine mit entsprechenden Unterrichtsinhalten angegeben. Hier soll über Erfahrungen mit dem Baustein „Rückwärtsschließen im Baumdiagramm“ für Klasse 9 und dem Teilbaustein „Alternativtests“ für Klasse 10 berichtet werden. Die Unterrichtseinheit zum Baustein „Rückwärtiges Schließen im Baumdiagramm“ könnte mit folgender **Aufgabe** beginnen:

„Till will Werbemüll (Spam) von seinem e-mail-Konto aussperren. Er installiert den von einer Computerzeitschrift ermittelten Testsieger, der 95% aller Werbemails ausfiltert. Leider sortiert das Programm auch 1% aller privaten e-mails und von Till bestellten Infobriefe als Spam aus. Beurteile die Qualität dieses Spamfilters. Würdest Du ihn benutzen?

Wir schließen mit dieser Aufgabe unmittelbar an die mit dem Baustein „Baumdiagramme“ in Klasse 8 gemachten Erfahrungen an.



In obigem Baumdiagramm bedeuten:

W: Werbe-mail (Spam)
kW: keine Werbe-mail
a: aussortiert
na: nicht aussortiert

Wenn Lernende ein Baumdiagramm erstellen, stellen sie fest, dass eine Information über den Anteil der Werbe-mails an allen e-mails noch fehlt. Im Baumdiagramm links unten gehen wir von 1000 in einer nicht näher festgelegten Zeiteinheit eingehenden e-mails und einem Anteil von 90% an Werbemails aus. Es kommen also 900 Werbe-mails und 100 andere e-mails an. Dies wird in der ersten Stufe des Baumdiagramms dargestellt. In der zweiten Stufe wird die Aufteilung nach den in der Aufgabe genannten Prozentzahlen vollzogen. Diese Situation kann man gut und übersichtlich in einer mit Erläuterungen erweiterten Vierfeldertafel darstellen, wobei die eigentliche Vierfeldertafel in der Mitte besonders hervorgehoben wird:

	aussortiert	nicht aussortiert	Summe
Werbe-mail	855	45	900
keine Werbung	1	99	100
Summe	856	144	1 000

Weitere Vierfeldertafeln zu anderen Anteilen der Werbe-mails an der Gesamtzahl an e-mails ergänzen die Betrachtungen in den Lerngruppen. Selbstverständlich wird auch parallel zum Baumdiagramm mit den natürlichen Anzahlen das übliche Baumdiagramm betrachtet, bei dem die Gesamtwahrscheinlichkeit 1 gemäß den Pfadwahrscheinlichkeiten auf alle vier Ergebnisse verteilt wird.

Grietje formuliert: „Ich bekomme zwei bis drei e-mails pro Tag. Ich kann das Löschen unerwünschter mails per Hand erledigen. Daher brauche ich keinen Spamfilter. Aber anders sieht es bei einer Firma aus, die 1000 e-mails an einem Tag erhält. Hier sortiert der Filter die einen in einen Ordner „Posteingang“, die anderen in „Spamverdacht“. Aus dem Ordner „Posteingang“ muss ich die nicht erkannten Werbe-mails entfernen, aus dem Ordner „Spamverdacht“ die erwünschten in den Ordner „Posteingang“ verschieben. Danach kann ich im

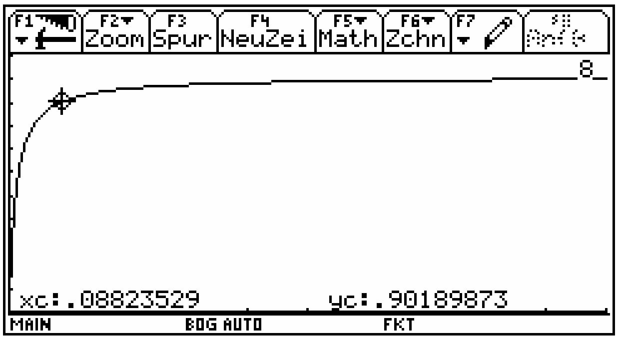
Ordner „Spamverdacht“ alle verbliebenen Werbemails markieren und mit einem Klick löschen. Darin sehe ich den einzigen Vorteil des Vorsortierens durch den Spamfilter bei großem Posteingang.“

Bis jetzt war noch keine Rede von Wahrscheinlichkeiten, auch nicht in Grietjes Beitrag. Ich habe bewusst die Vorzüge des Baumdiagramms mit den natürlichen Anzahlen ausgenutzt. Und das soll auch im nächsten Schritt mit der **Aufgabe** erfolgen:

„Wie groß ist die Wahrscheinlichkeit, dass die als Werbemüll aussortierten e-mails tatsächlich Werbemüll sind ?“

Wir betrachten das auf der vorigen Seite abgebildete Baumdiagramm. Von den 1000 eingegangenen e-mails wurden 856 in den Ordner „Spamverdacht“ aussortiert. 855 davon sind tatsächlich Spam. Also beträgt die gesuchte Wahrscheinlichkeit $\frac{855}{856} \approx 0{,}9988$. Unter diesen Bedingungen erweist sich der Filter als sehr wirksam.

Mit dieser Aufgabe gehen wir über die Fragestellungen aus Klasse 8 hinaus. Damals haben wir im Baumdiagramm von oben nach unten geschlossen, von den Teilwahrscheinlichkeiten auf die Wahrscheinlichkeiten von Ergebnissen und Ereignissen. Bei dieser neuen Aufgabe kommt eine weitere Information hinzu: Wir wissen, wie viele e-mails unter „Spamverdacht“ abgespeichert sind. Damit wird ein Ereignis E definiert. Nun schauen wir von unten (von den Ergebnissen/Ereignissen) im Baumdiagramm zurück und schließen auf die Anteile, die die einzelnen Wege, die zu E führen, an E haben.

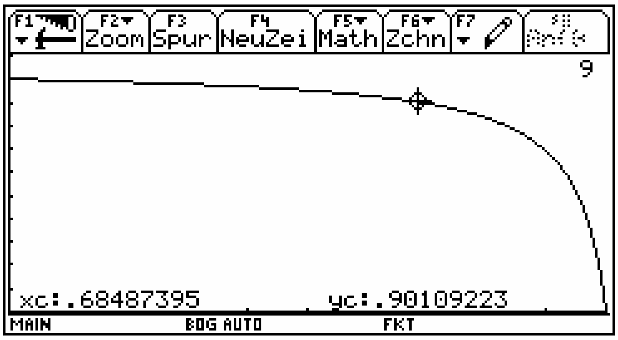


Wenn wir allgemein für den Anteil Spam an allen e-mails die Variable p benutzen, erhalten wir den Term: $\frac{0{,}95\cdot p}{0{,}95\cdot p+(1-p)\cdot 0{,}01} = \frac{95\cdot p}{94\cdot p+1}$. Diesen Term setzen wir in das „Y=Menü“ unseres graphikfähigen Taschenrechners ein. Allerdings müssen wir anstelle der Variablen p bei der Eingabe des Terms in den Rechner x wählen. Dann können wir uns sowohl den Verlauf für alle $p \in [0;1]$ graphisch darstellen lassen als auch im „Table“-Menü die Wertetafel anschauen. Dies wird im obigen Bild dargestellt. Es ist nicht schwer, auch in Klasse 9 schon zu begründen, dass der Graph streng monoton wächst; denn Zähler und Nenner wachsen, der Zähler schneller als der Nenner. Da der Nenner immer größer als der Zähler ist und beide positiv sind, ist der Quotient positiv und kleiner als 1. Wir lesen von Graphen oder aus der Wertetafel ab, dass von einem Anteil p = 0,087 an der Anteil Spam an den aussortierten e-mails größer als 90% ist und mit steigendem p wächst. Dies wird als eine gute Eigenschaft des Filters angesehen.

„Die Software löst Probleme, die man ohne Computer nicht hat.“, sagt Malte. „Die Software schafft aber auch Probleme, die man ohne diese Software nicht hat.“, ergänzt Katharina. Beide haben entdeckt, dass der Filter unter einem anderen Gesichtspunkt nicht sehr wirksam sein kann, und formulieren für alle in der Lerngruppe die **Aufgabe**:

„Wie groß ist die Wahrscheinlichkeit, dass die nicht als Werbe-mails (Spam) aussortierten e-mails tatsächlich kein Spam sind ?“

Von den 1000 eingegangenen e-mails sind 99 im Ordner „Posteingang“ und 45 im Ordner „Spamverdacht“ keine Spam. Die gesuchte Wahrscheinlichkeit beträgt: $\frac{99}{99+45} = \frac{99}{144} \approx 0{,}6875$. Hier erweist sich der Filter als nicht besonders wirksam.



Allgemein erhält man für diesen Fall bei einem Anteil p von Spam:

$$\frac{0{,}99\cdot(1-p)}{0{,}99\cdot(1-p)+0{,}05\cdot p} = \frac{99-99\cdot p}{99-94\cdot p}.$$

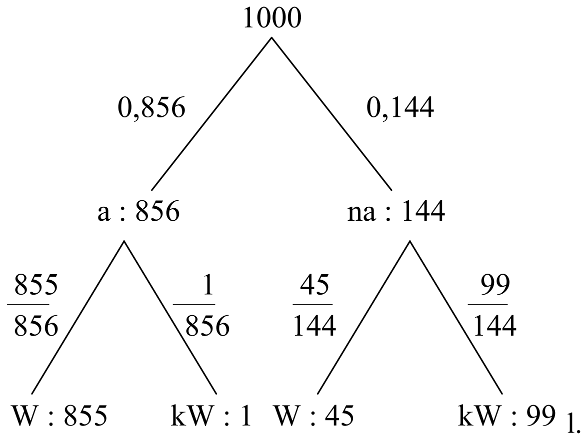
Von einem Anteil p = 0,688 ab sinkt der Anteil der erwünschten e-mails (kein Spam) auf unter 90%. Auch hier lässt sich elementar begründen, dass der Graph streng monoton fällt.

Interessant und besonders motivierend für den Unterricht in der letzten Lerngruppe war die Tatsache, dass die Zeitschrift ComputerBild in ihrer Ausgabe 22/2004 einen Spamfilter als Testsieger herausgestellt hat, der die Eigenschaften des hier betrachteten Filters hatte. Das verlieh dem Unterricht große Aktualität. Das hat die Computerzeitschrift nicht beschrieben: Es gibt zwei Gesichtspunkte, unter denen der Filter betrachtet werden kann. Die Lernenden sind stolz, dies entdeckt zu haben. Wenn mich der Anteil Spam an den als Spam deklarierten interessiert, dann arbeitet der Filter bereits bei geringem Spamanteil von 10% gut. Interessiert mich aber der Anteil der erwünschten e-mails (kein Spam) an den als Nicht-Spam aussortierten, dann wirkt er nur bis zu einem Anteil Spam von ca. 68% gut. Bei einem größerem Anteil an Spam wird die interessierende Wirkung immer geringer, der Graph fällt rapide ab. So muss also jeder Nutzer selbst sein Hauptinteresse definieren und danach entscheiden, ob ein Spamfilter bei ihm Hilfe verspricht und eingesetzt werden soll.

„Setzen wir voraus, dass wir nie einen Filter bekommen, der 100% exakt trennt.“ Unter dieser Leitlinie wurde in den Lerngruppen diskutiert, was eine Verbesserung von 95% auf zum Beispiel 98% bei sonst gleichen Bedingungen bringt. Die Enttäuschung, dass die Verbesserung minimal ist, möge jeder Leser doch bitte selbst einmal durch eigenes Tun erleben.

2. Arbeit mit der Vierfeldertafel

Aus den beiden unterschiedlichen Gesichtspunkten sollen nun zwei völlig unterschiedliche Beschreibungen entwickelt werden, bei denen man auf den ersten Blick nicht erkennen kann, dass sie sich auf den gleichen Spamfilter beziehen. Hier ein Vorschlag aus einer Lerngruppe: „Merle arbeitet in einem Betrieb, der täglich im Mittel 1000 e-mails erhält. Sie ärgert sich über den neuinstallierten Spamfilter; denn jetzt muss sie mühsam alle für die Firma wichtigen e-mails aus den beiden Ordnern „Eingang“ und „Spamverdacht“ herausfinden und in den Ordner „Zu bearbeiten“ verschieben. Merle hat die mails vom Wochenende gezählt. Im Ordner „Eingang“ hat sie die nicht als Spam aussortierten 45 Werbe-mails von 99 für die Firma wichtigen mails trennen müssen. Im Ordner „Spamverdacht“ hat sie eine einzige für die Firma wichtige mail aus den insgesamt 856 dort befindlichen herausfinden müssen. Merle meint, es sei für sie einfacher, die eingehenden Mails der Reihe nach ohne Einsatz des Filters auf Spam oder Nicht-Spam zu untersuchen.“ Zu Merles Darstellung könnte das folgende Baumdiagramm gehören, das rechts oben abgedruckt ist. Gegenüber dem Baumdiagramm aus Abschnitt 1 ist die Reihenfolge der Stufen vertauscht. Aber ansonsten erhalten wir aus diesem Baumdiagramm, das wir als zum ersten „dual“ betrachten, die gleiche Vierfeldertafel.



Die andere Notiz wird von der Lerngruppe so formuliert: „Selbst bei einem so hohen Anteil von 90% Spam an den eingehenden e-mails schafft es der Spamfilter, dass nur 1% der erwünschten e-mails nicht im richtigen Ordner landen und per Hand verschoben werden müssen. Von 900 Werbemails werden nur 5% fehlsortiert.“ Zu diesem Text passt das Baumdiagramm aus Abschnitt 1 besser. Man muss schon beide Baumdiagramme erstellen, um zu erkennen, dass beide Texte Aussagen über den gleichen Spamfilter machen.

Bis hierher wurde aus einem gegebenen Kontext die Vierfeldertafel entwickelt. Nun wird der Spieß umgedreht. Es wird eine Vierfeldertafel vorgegeben, ohne dass den Zahlen irgendeine Bedeutung unterlegt wird und eine **Aufgabe** gestellt:

„Gegeben sei folgende Vierfeldertafel:

95	5
18	882

Schreibe zwei kurze Zeitungsnotizen, die beide auf dieser Vierfeldertafel beruhen, die aber mit so unterschiedlichen Zahlen argumentieren, dass man erst nach einer mathematischen Untersuchung die gemeinsame Datenquelle entdeckt.“

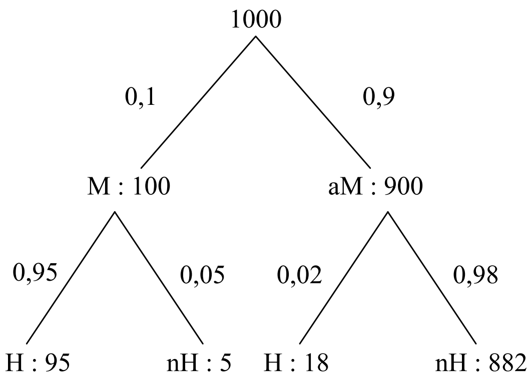
Annike schreibt zum Stichwort „Zwei-Klassen-Gesellschaft in der Medizin“: „Es wurde ein neues Verfahren zur Behandlung einer Krankheit entwickelt. Die Krankenkassen lehnen eine Übernahme der hohen Kosten ab, so dass bisher nur 10% der Erkrankten dieses neue Verfahren in Anspruch nehmen konnten. Davon wurden 95% geheilt. Wer mit herkömmlichen Methoden behandelt wird, hat dagegen nur eine 2%-Chance auf Heilung.“

Zum Stichwort „Zu wenig Erfolge bei der Krankheitsbehandlung“ schreibt sie: „Es ist alarmierend, dass nur 11,3% der Patienten, die ein bestimmtes Leiden haben, bisher geheilt werden. Dabei gibt es eine neue Behandlungsmethode, die bisher nur in 5 von 887 statistisch erfassten Fällen zu keinem Erfolg führte. Mit herkömmlichen Methoden konnten dagegen nach dieser Statistik nur 144 von 904 Erkrankten geheilt werden.“

Wir erweitern die Vierfeldertafel. Die Bezeichnungen orientieren sich an der Lösung von Annike:

	geheilt	nicht geheilt	Summe
neue Methode	95	5	100
alte Methode	18	882	900
Summe	113	887	1000

Zum ersten Artikel von Annike entwickeln wir folgendes Baumdiagramm:

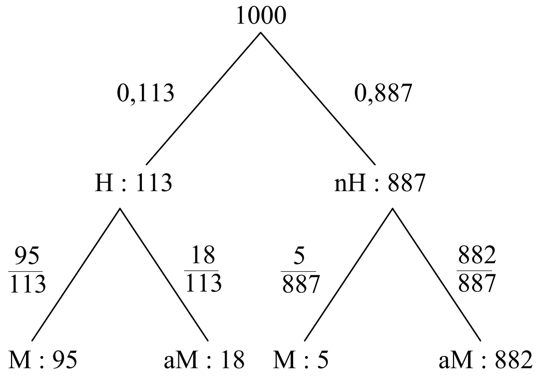


M: neue Methode
aM: alte Methode

H: geheilt
nH: nicht geheilt

An die Spitze des Baumdiagramms schreiben wir die Gesamtsumme der Vierfeldertafel 1000 und an die Enden der Pfade der ersten Stufe die beiden Zeilensummen 100 bzw. 900. Daraus berechnen wir die beiden Pfadwahrscheinlichkeiten 0,1 und 0,9. An die Enden der Pfade der zweiten Stufe schreiben wir die Elemente der ersten Zeile (95 und 5) sowie der zweiten Zeile (18 und 882). Daraus lassen sich wieder die Pfadwahrscheinlichkeiten (0,95; 0,05; 0,02; 0,98) als Anteile 95 von 100, 5 von 100, 18 von 900 bzw. 882 von 900 berechnen.

Zum zweiten Artikel von Annike entwickeln wir das folgende Baumdiagramm:



An die Spitze des Baumdiagramms schreiben wir wieder 1000 und an die Enden der Pfade der ersten Stufe die beiden Spaltensummen 113 bzw. 887. Daraus berechnen wir die beiden Pfadwahrscheinlichkeiten 0,113 und 0,887. An die Enden der Pfade der zweiten Stufe schreiben wir die Elemente der ersten Spalte (95 und 18) sowie der zweiten Spalte (5 und 882). Daraus lassen sich wieder die vier Pfadwahrscheinlichkeiten als Anteile berechnen.

Damit ist auch das Verfahren beschrieben, wie Lernende die beiden Sichtweisen aus einer Vierfeldertafel entwickeln können: Einmal zeilen-, das andere Mal spaltenorientiert.

Es können leicht weitere Aufgaben aus dem Umfeld des Satzes von Bayes zum Interpretieren oder Erstellen von Vierfeldertafeln oder zum Rückwärtserschließen einer Wahrscheinlichkeit in Baumdiagrammen nach diesen Vorbildern erstellt werden.

3. Zwischenbilanz

In den beiden vorhergehenden Abschnitten wurde dargestellt, dass und wie mit dem Baumdiagramm und den beiden Pfadregeln als alleinigen Hilfsmitteln die Inhalte des Bausteins „Rückwärtsschließen im Baumdiagramm“ in Klasse 9 erarbeitet werden können. Bedingte Wahrscheinlichkeit gilt als schwieriger Begriff, der leicht Missverständnissen ausgesetzt ist. Aber in vielen interessanten Problemstellungen werden bedingte Wahrscheinlichkeiten angesprochen, das macht die Integration in den Unterricht so reizvoll. Lernende haben jedoch häufig vor allem bei offen formulierten Aufgabenstellungen Schwierigkeiten, bedingte Wahrscheinlichkeiten zu erkennen. Beim hier dargestellten Vorgehen haben Lernende an Hand der Baumdiagramme mit natürlichen Anzahlen eine gute Möglichkeit, das Vorgehen bei Problemstellungen im Umfeld des Satzes von Bayes selbst zu erkennen. Insbesondere ist es dann nicht erforderlich, eine Definition „bedingte Wahrscheinlichkeit“ zu erarbeiten, den Satz von Bayes und den von der totalen Wahrscheinlichkeit zu formalisieren. Beim anderen Typ Baumdiagramm, bei dem nur mit

Wahrscheinlichkeiten gearbeitet wird, stellt sich das Erkennen eines Lösungswegs und das Entwickeln einer Lösung viel schwieriger dar.

4. Zum Alternativtest

Ein erster Übergang zum Teilbaustein „Alternativtests“ soll mit folgender **Aufgabe** erfolgen:

„Martin wählt aus den beiden Zufallsgeräten Laplace-Würfel und langer U-Würfel eins aus. Er würfelt die Ergebnisse 5, 4, 3, 4, 4 in genau dieser Reihenfolge.

a. Welchen Würfel hat er gewählt? Schreibe zuerst eine spontane Vermutung auf, ohne lange nachzudenken.

b. Untersuche, welchen Würfel er gewählt hat. Für den langen U-Würfel wird folgende aus einer langen Versuchsreihe entwickelte Wahrscheinlichkeitsverteilung angenommen:

Ergebnis	1	2	3	4	5	6
Wahrscheinlichkeit	0,12	0,06	0,22	0,42	0,06	0,12

c. Untersuche, was sich an den Ergebnissen aus Aufgabe b ändert, wenn die Reihenfolge der Wurfergebnisse nicht bekannt ist, es also lediglich heißt: Es wurde einmal die „3“, dreimal die 4“ und einmal die „5“ gewürfelt.“

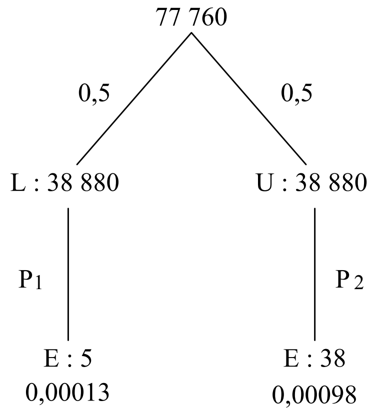
Hinweis: Der lange „U-Würfel“ gehört zu den sogenannten Riemer-Würfeln und ist in Riemer (1985) und Riemer (1988) beschrieben. Aus einer im Querschnitt U-förmigen Profilleiste (das „U" ist 2 cm breit und 1,2 cm hoch) werden 2,5 cm lange Stücke ausgesägt. Die 6 Wurflagen werden wie beim normalen Würfel mit 1 bis 6 beschriftet.

Lösungsskizzen zu a:

Anke schreibt: „Ganz spontan entscheide ich mich für den langen U-Würfel, weil für ihn die Ergebnisse „3“ und „4“ häufiger als beim L-Würfel auftreten.“

Lösungsskizzen zu b:

Anke schreibt: „E bedeutet das Ereignis „Es wurde zuerst eine „5“, dann eine „4“, danach „3“, dann „4“ und zuletzt „4“ gewürfelt. Ich stelle mir 77 760 solcher Versuche vor. Als neutraler Beobachter erwarte ich, dass Martin 38 880 Mal den L-Würfel (L) und 38 880 Mal den langen U-Würfel (U) benutzt. Das habe ich in der ersten Stufe des Baumdiagramms ausgedrückt.



Die Wahrscheinlichkeit für die beobachtete Ergebnisfolge beträgt beim L-Würfel

$$P_1 = \left(\frac{1}{6}\right)^5 \approx 0{,}00013\,.$$

Bei 77 760 Versuchen ist das Ereignis 5 Mal zu erwarten. Beim langen U-Würfel beträgt die Wahrscheinlichkeit für die beobachtete Ergebnisfolge $P_2 = 0{,}06 \cdot 0{,}22 \cdot 0{,}42^3 \approx 0{,}00098$. Das ist bei 77 760 Versuchen rund 38 Mal zu erwarten. Das habe ich in der zweiten Stufe des Baumdiagramms dargestellt. Eingetreten ist das Ereignis E mit

$P(E) = 0{,}5 \cdot P_1 + 0{,}5 \cdot P_2 \approx 0{,}00055$. Der Anteil der Wahrscheinlichkeit, dass der L-Würfel benutzt wurde, an P(E) ist $\frac{0{,}5 \cdot P_1}{P(E)}$ und beträgt $\frac{5}{43} \approx 0{,}12$, der Anteil der Wahrscheinlichkeit, dass der lange U- Würfel benutzt wurde, an P(E) ist $\frac{0{,}5 \cdot P_2}{P(E)}$ und beträgt $\frac{38}{43} \approx 0{,}88$. Ich behalte meine spontane Vermutung bei, dass der lange U-Würfel benutzt wurde. Dafür spricht eine Wahrscheinlichkeit von 88%.“

Lösungsskizzen zu c:

Es gibt 5 Möglichkeiten, wann das Ergebnis „3“ gewürfelt werden kann. Danach gibt es nur noch 4 Möglichkeiten, „5“ zu platzieren. Die Ergebnisse „4“ werden an die restlichen Stellen gebracht. Also gibt es $5 \cdot 4 = 20$ Möglichkeiten für die von Martin angegebene Folge der Wurfergebnisse. Wir müssen daher P_1, P_2 und P(E) mit 20 multiplizieren. Bei den beiden Quotienten $\frac{0{,}5 \cdot P_1}{P(E)}$ und $\frac{0{,}5 \cdot P_2}{P(E)}$ kürzen sich der Faktor 20 und die Vorbewertung 0,5 heraus, so dass sich an den Ergebnissen aus Aufgabe b nichts ändert.

Kommentar:

Diese Aufgabe bildet bei mir die Brücke zwischen dem Baustein „Rückwärtsschließen im Baumdiagramm“ und dem Teilbaustein „Alternativtests“. Ich kann mit ihr den Unterricht zum ersten Baustein abschließen oder den Unterricht zum zweiten Baustein beginnen. Die Aufgabe wird bewusst so gestellt, dass zunächst die Reihenfolge der Wurfergebnisse vorgegeben wird. So sind zunächst keine kombinatorischen Überlegungen erforderlich. Erst in Teil c erfolgt der Übergang. Aber selbst hier ist es nicht erforderlich, die Binomialverteilung vorher zu erarbeiten; denn diese elementaren Überlegungen lassen sich auch ohne Kenntnis des Binomialkoeffizienten oder Behandlung der kombinatorischen Grundaufgaben anstellen. Wir müssen noch nicht einmal wissen, wie viele Wege im Baumdiagramm genau zu diesem Ereignis (einmal die „3“, dreimal die „4“ und einmal die „5“) gehören. Es reicht völlig aus, wenn wir z schreiben, wobei die Variable z für die Anzahl der Wege im Baumdiagramm, die zu diesem Ereignis gehören, steht. Da sich diese Variable bei der Berechnung von $P(H_1)$ und $P(H_2)$ wegkürzt, ist bei diesem „Testen nach Bayes“ genannten Verfahren eine vorhergehende Behandlung der Binomialverteilung, wie sie in den Rahmenrichtlinien (2003) vorgesehen ist, nicht zwingend erforderlich. Dass das Wegkürzen des Binomialkoeffizienten aber auch zu Problemen mit dem CA-Taschenrechner führen kann, und wie man bei diesen Problemen vorgeht, wird in Wirths (2005) beschrieben. Auch wenn ich mit dieser Aufgabe allein bereits den Teilbaustein „Alternativtests“ voll erfüllen kann, soll noch eine weitere **Aufgabe** betrachtet werden:

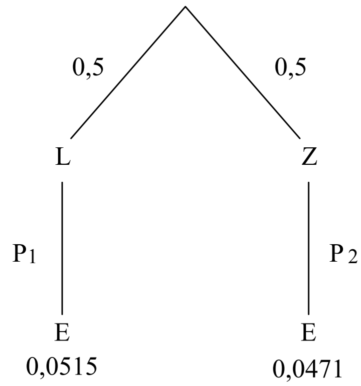
„Kai hat von seinem Onkel einen Würfel geerbt. Er weiß, dass alle Würfel seines Onkels bis auf eine Ausnahme L-Würfel waren. Nur ein Würfel ist gezinkt. Bei ihm kommt die „6“ mit einer Wahrscheinlichkeit von 42%. Leider sind die anderen Würfel verloren gegangen. Nun möchte Kai wissen, ob dies der gezinkte Würfel ist oder nicht.

a. Kai hat 18 Versuche gemacht und dabei 5 Mal die „6“ gewürfelt. Untersuche, wie sich nach diesem Versuch die Situation darstellt.

b. Kai hat bei 180 Versuchen 50 Mal „6“ gewürfelt. Arbeitsauftrag wie in Aufgabe a.“

Lösungsskizzen zu a:

Hier soll nicht mehr das Baumdiagramm mit den natürlichen Anzahlen, sondern das sonst übliche betrachtet werden:



Als neutraler Beobachter erwarte ich, dass in der Hälfte der Fälle ein L-Würfel (L) und in der Hälfte der Fälle ein gezinkter Würfel (Z) vorliegt. Das wird in der ersten Stufe des Baumdiagramms ausgedrückt. Es gilt: $P_1 = \binom{18}{5} \cdot (\frac{1}{6})^5 \cdot (\frac{5}{6})^{13} \approx$ 0,103,

$$P_2 = \binom{18}{5} \cdot 0{,}42^5 \cdot 0{,}58^{13} \approx 0{,}094.$$

Die Pfadwahrscheinlichkeiten $0{,}5 \cdot P_1$ und $0{,}5 \cdot P_2$ wurden am Ende der beiden Pfade angegeben. Es ist ein Ereignis E (bei 18 Versuchen 5 Mal „6“) eingetreten, für das gilt: $P(E) = 0{,}5 \cdot P_1 + 0{,}5 \cdot P_2 \approx 0{,}0985$. Der Anteil der Wahrscheinlichkeit, dass der L-Würfel vorliegt, an P(E) beträgt ca. 52%, der Anteil der Wahrscheinlichkeit dafür, dass der Würfel gezinkt ist, an P(E) ist rund 48%. Kai kann also genau so gut eine Münze in einem leicht unsymmetrischen Design entscheiden lassen, welcher Hypothese er Vertrauen schenken soll. Eine klare Entscheidung ist also hier nicht möglich.

Lösungsskizzen zu b:

Es gilt: $P_1 = \binom{180}{50} \cdot (\frac{1}{6})^{50} \cdot (\frac{5}{6})^{130} \approx 0{,}000064$

und $P_2 = \binom{180}{50} \cdot 0{,}42^{50} \cdot 0{,}58^{130} \approx 0{,}000026.$

Nun beträgt der Anteil der Wahrscheinlichkeit, dass der L-Würfel vorliegt, an P(E) etwa 71%, der Anteil der Wahrscheinlichkeit, dass der Würfel gezinkt ist, an P(E) ist rund 29%. Kai kann sich jetzt dafür entscheiden, dass der Würfel nicht gezinkt ist. Die Unsicherheit ist mit rund 29% jedoch noch recht hoch.

Es wird mit beiden Aufgabenteilen eine Tendenz deutlich, dass eine proportionale Vergrößerung der Anzahl an Versuchen und der Würfe mit Ergebnis „6“ zu einer immer stärkeren Hinwendung zur Hypothese, dass der Würfel nicht gezinkt ist, führt, was intuitiv auch von vornherein klar ist.

5. Abschluss

Wie mit den für viele Mathematik-Lehrerinnen und -Lehrer neuen Bausteinen „Rückwärtsschließen im Baumdiagramm“ und „Alternativtests“ gearbeitet werden kann, und welche stochastischen Voraussetzungen dazu benötigt werden, war Anliegen dieses Beitrags. Zu beiden Bausteinen benötigt man nur Vorkenntnisse im Umgang mit Baumdiagrammen und den Pfadregeln. Sie sind daher gut geeignet, bereits in der Mittelstufe des Gymnasiums behandelt zu werden. Stochastisches Denken kann dort mit interessanten und motivationsstarken Beispielen entwickelt werden.

Leider muss hier auch über Probleme im Umgang mit den neuen niedersächsischen Richtlinien (2003) berichtet werden. In den Richtlinien (2003) findet man beim Teilbaustein „Alternativtests“ in der Rubrik „Inhalte und Verfahren“ als verbindlich „Hypothesen und ihre Wahrscheinlichkeiten“, „Fehler und ihre Wahrscheinlichkeiten“ und „Interpretationen von Fehlern“. Man findet noch den Hinweis: „Die Durchführung von Alternativtests basiert auf der Bernoulliformel.“ Dass diese Einschränkung unnötig ist, wurde bereits oben gezeigt. Welche Alternativtests behandelt werden sollen, darüber sagen die Richtlinien nichts aus, der zugehörige Materialband (2003) stellt nur das klassische Verfahren vor.

„Hypothesen und ihre Wahrscheinlichkeiten“ machen nur beim Alternativtest nach Bayes Sinn, nicht aber beim klassischen Alternativtest. Damit fordern die Richtlinien etwas, was Lehrende nicht leisten können; denn für beide Konzeptionen steht in der Jahrgangsstufe 10 in Niedersachsen nicht genug Zeit zur Verfügung, erst recht nicht, wenn wir in Zukunft auf dem Gymnasium nur noch 8 Jahrgänge unterrichten. Wir müssen uns daher für einen Weg entscheiden.

Der Alternativtest nach Bayes liefert auf vernünftig gestellte Fragen Antworten. Demgegenüber ist der klassischen Alternativtest erfahrungsgemäß sehr zeitaufwändig und kann wegen seiner komplizierten und missverständlichen Ergebnisinterpretation dazu von Lernenden auch noch abgelehnt werden, was sie dann häufig sehr drastisch ausdrücken..

Wäre es nicht eine Diskussion in dieser Zeitschrift wert zu folgenden Fragen ?

Wieviel Testen braucht der Mensch zur Allgemeinbildung?

Wann soll welcher Inhalt unterrichtet werden?

Meine Meinung habe ich hier und in anderen Beiträgen vorgebracht. Ich würde mich über eine rege Diskussion in dieser Zeitschrift sehr freuen.

Literatur

Rahmenrichtlinien (2003): Rahmenrichtlinien für das Fach Mathematik: Schuljahrgänge 7 - 10. Gymnasium. Hannover: Schroedel 2003

Materialband (2003): nli-Bericht 69: Unterrichtsbeispiele zur Stochastik in den Schuljahrgängen 7 - 10 des Gymnasiums. NLI: Hildesheim 2003

Riemer (1985): Neue Ideen zur Stochastik. Mannheim: BI 1985

Riemer (1988): Riemer-Würfel. Stuttgart: Klett 1988

Wirths (2005): Hat Gregor Mendel seine Daten „frisiert“? erscheint in Stochastik in der Schule

Anschrift des Verfassers

Helmut Wirths
Cäcilienschule Oldenburg, Haarenufer 11,
26122 Oldenburg;
CvO Universität Oldenburg, Institut für Mathematik,
26015 Oldenburg
e-mail: helmut.wirths@uni-oldenburg.de