Jörg Meyer

Zweistufige Zufallsexperimente – spannender, als man vermutet[1]

Die Bildungsstandards sehen vor, dass die Schülerinnen und Schüler Daten interpretieren sollen sowie auf einer Datenanalyse basierende Argumente reflektieren und bewerten sollen. Dies wird etwa in Niedersachsen dahingehend interpretiert, dass die Lernenden Kenntnisse über zweistufige Zufallsexperimente nutzen sollen, um statistische Aussagen mit Hilfe von Baumdiagramm oder Vierfeldertafel zu interpretieren. Daher soll es hier um folgende Fragen gehen: Was kann bei Situationen, die mit zweistufigen Zufallsexperimenten zu beschreiben sind, alles vorkommen? Wie kann man quantitative Schlüsse möglichst einsichtig machen? Was kann man alles falsch machen? Wie lässt sich das Lernen aus Erfahrung in diesen Zusammenhang einordnen?

Phänomene bei bivariaten Situationen

Zweistufige Zufallsexperimente treten in zwei Varianten auf: Entweder wird zweimal nacheinander dasselbe Merkmal ermittelt (Beispiel: Zweifaches Würfeln, das Merkmal ist die jeweils gewürfelte Augenzahl) oder es werden zwei unterschiedliche Merkmale ermittelt (Beispiel: Gleichzeitiges Würfeln mit einem roten fairen Würfel und einem grünen gezinkten Würfel). Hier soll es um den Fall gehen, dass zwei unterschiedliche Merkmale vorhanden sind. Vor einer quantitativen Behandlung lasse ich die Schülerinnen und Schüler sich erst einmal qualitativ mit den Phänomenen vertraut machen, die bei zwei relevanten Merkmalen auftreten können. Dabei bleibe ich vorerst im Bereich der beschreibenden Statistik. Wir werden sehen, dass es bei zwei unterschiedlichen Merkmalen Phänomene gibt, die durchaus paradox anmuten können.

Qualitativer Einstieg

<u>Beispiel 1</u>: Zwei Lehrer unterhalten sich über eine Klasse, die sie gerade neu übernommen haben. Ein Lehrer sagt: „In dieser Klasse tragen fast alle Jungen eine Brille“. Darauf der andere: „Mir ist aufgefallen, dass fast alle Brillenträger Mädchen sind.“
Wer hat Recht? Vielleicht alle beide? Sind dann fast alle Jungen in Wirklichkeit Mädchen?

Diese Fragestellung (oder ähnliche) haben sich für den Einstieg in die Thematik als geeignet erwiesen; die Aussagen der beiden Lehrer werden als offensichtlich widersprüchlich erkannt. Der weiterführende Impuls „Vielleicht haben beide Lehrer Recht“ führt zu dem Versuch, eine Klassenzusammensetzung zu finden, die den beiden Lehrer-Aussagen entspricht. Erfahrungsgemäß problematisieren die Schülerinnen und Schüler erst einmal die Bedeutung von „fast alle“. Hier kann die Lehrperson durchaus sagen, dass mit „fast alle“ immer „mindestens 80 %“ gemeint sein sollen (die 80 % sind nicht wesentlich; sie lassen sich auch durch 55 % oder durch 90 % ersetzen). Nachdem die Schülerinnen und

---

[1] Eine leicht veränderte Version erschien in Praxis der Mathematik **39** (Juni 2011 / 53. Jahrgang); S. 19 - 24.

Schüler etwas gepuzzelt haben, entsteht nach einiger Zeit etwa eine Lösung wie die in Abbildung 1 gezeigte.

	Mädchen	Jungen
Brille	16	4
keine Br.		1

Abb. 1: Eine mögliche Lösung zum Beispiel 1 in Form einer Vierfelder-Tafel

Es ist dabei völlig egal, wie viele Mädchen ohne Brille es gibt. Die Lernenden erkennen, dass der erste Lehrer eine Aussage über die zweite Spalte (der Jungen) gemacht hat und der zweite Lehrer über die erste Zeile (der Brillenträger) geredet hat. Beide Aussagen sind durchaus miteinander verträglich; beide Aussagen sind sogar voneinander unabhängig. Das erkennen die Schülerinnen und Schüler schon dadurch, wenn sie erst die Jungenspalte und dann die Mädchenzeile ausfüllen.

Das Beispiel 1 lässt sich vielfältig variieren (viele sind Beispiele in Cukrowicz et al., 2008 und Meyer, 2008 enthalten). Dabei ist es auch sinnvoll, die Lernenden aufzufordern, sich ähnliche Gegebenheiten wie in Beispiel 1 auszudenken; erst dadurch wird die Besonderheit des Beispiels deutlich. Genau hierin liegt nämlich die Schwierigkeit der Schülerinnen und Schüler: Ein einziges Beispiel lässt noch nicht die allgemeine Struktur erkennen. Aus Schüler-Beispielen wie „Fast alle Jungen tragen eine Brille; fast alle Mädchen auch“ lassen sich zwar passende Vierfelder-Tafeln konstruieren, die prinzipiell mögliche doppelte Sichtweise kommt hier aber noch nicht zum Tragen.

Erster Ausbau: unzulässige transitive Schlussfolgerungen

Aus den Angaben in Beispiel 1 folgt natürlich nicht, dass fast alle Jungen in Wirklichkeit Mädchen sind. Daher ist es sinnvoll, etwa in der Folgestunde den Lernenden folgenden „Schluss“ zu präsentieren:

Beispiel 2: Die meisten Kölner sind Deutsche. Die meisten Deutschen waren noch nie in Köln. Also „folgt“: Die meisten Kölner waren noch nie in Köln.

Dies Beispiel ist nur dann instruktiv, wenn die Schülerinnen und Schüler die Gemeinsamkeiten zu Beispiel 1 herausarbeiten; in beiden Fällen hat man Aussagen der Art

Fast alle A sind B.

Fast alle B sind nicht A.

Schön ist es, wenn die Schülerinnen und Schüler möglichst weitere Beispiele solch unzulässiger transitiver Schlüsse finden.

Zweiter Ausbau: verschiedene Darstellungsformate

Andere Variationen zu Beispiel 1 sollen helfen, vorschnelle Schlüsse in Zukunft zu vermeiden:

<u>Beispiel 3</u>: In der 8b sind 20 Jungen und 4 Mädchen. 75 % der Mädchen und 25 % der Jungen lernen auf einer Musikschule ein Instrument. Alice aus der 8a hört, wie die Klassenlehrerin der 8b auf dem Schulflur einem Mann hinterher eilt und ihm zuruft: „Ihre Frau hat eben angerufen und lässt ausrichten, dass Sie nachher zur Musikschule sollen, um Ihr Kind abzuholen.“. „Naja“, denkt Alice, „das Kind wird wohl ein Mädchen sein, denn die meisten Mädchen gehen auf die Musikschule“. Ob sie mit der Vermutung wohl Recht hat?

Hier bieten sich zur Darstellung der in der Aufgabe enthaltenen Information mehrere Formate an; mit der *Vierfelder-Tafel* (die durch die Randsummen ergänzt wird) geht es so:

	Mädchen	Jungen	
Instrument	3	5	8
kein Instr.	1	15	16
	4	20	24

Abb. 2: Lösung in Form einer Vierfelder-Tafel

Nun ist die Übersicht leicht: 8 Kinder spielen ein Instrument, und davon sind mehr als die Hälfte männlich. Alice hat also falsch gelegen mit ihrer Vermutung. Sie hat die beiden Aussagen

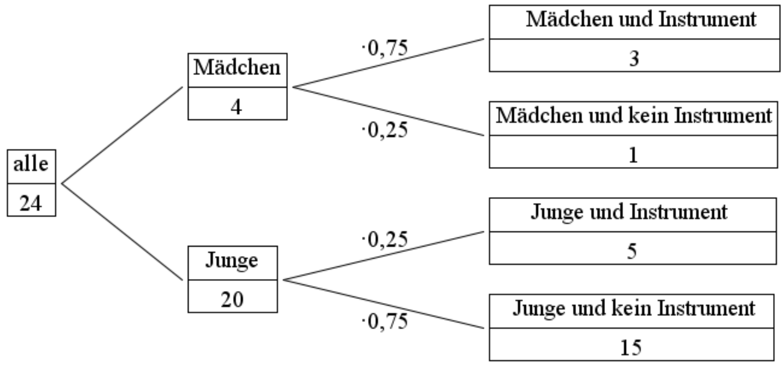
Die meisten Mädchen spielen ein Instrument

(dies bezieht sich auf die Mädchenspalte) und

Die meisten Instrumentspieler sind Mädchen

(dies bezieht sich auf die Instrumentzeile) miteinander verwechselt. Die erste Aussage ist richtig, die zweite falsch. Erfahrungsgemäß haben viele Schülerinnen und Schüler zunächst Schwierigkeiten, den entscheidenden Unterschied zwischen den beiden umrahmten Aussagen auf den Punkt zu bringen; diese Schwierigkeiten werden nicht geringer oder gar behoben, wenn die Lehrperson vorschnell zu quantitativen Berechnungen übergeht.

Statt mit Vierfelder-Tafeln lassen sich die in Beispiel 3 enthaltenen Informationen auch mit einem *Baumdiagramm* oder einem *Doppelbaum* darstellen:



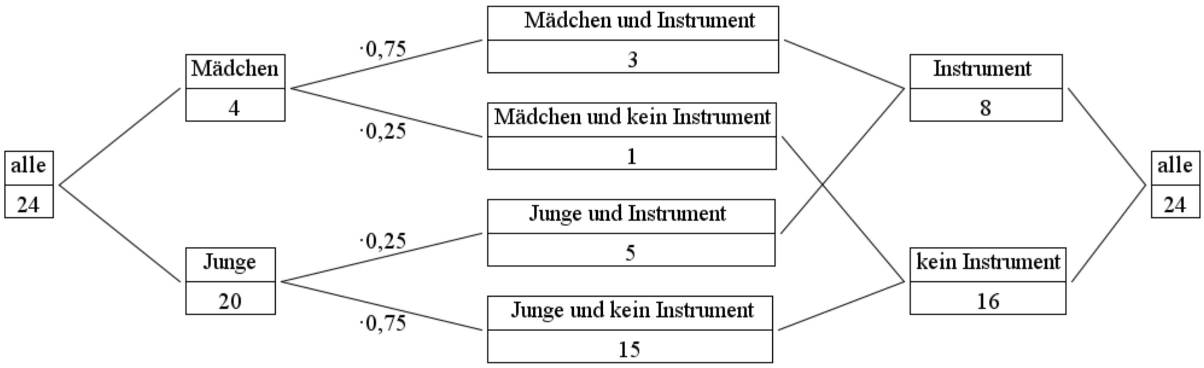


Abb. 3: Lösung in Form eines Baumdiagramms und alternativ in einem Doppelbaum

Bei der Vierfelder-Tafel war eine Sichtweise spaltenorientiert und die andere zeilenorientiert. Beim Doppelbaum kann man von links argumentieren (bei der Vierfelder-Tafel würde man sich erst einmal um die Summenzeile unten kümmern) oder auch von rechts (das entspricht der Summenspalte rechts).

Im *Kreisdiagramm* treten insbesondere die jeweils betrachteten Anteile graphisch hervor:

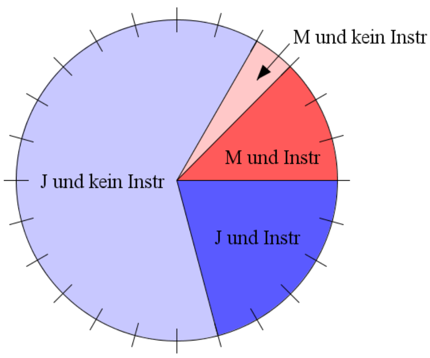


Abb. 4: Lösung in Form eines Kreisdiagramms

Beim Kreisdiagramm in Abbildung 4 sind die Instrumentenspieler dunkelfarbig und die anderen hellfarbig. Dass die meisten Mädchen ein Instrument spielen, ergibt sich aus dem Verhältnis der unterschiedlich rot gefärbten Flächen. Dass die meisten Instrumentenspieler Jungen sind, können die Schülerinnen und Schüler aus dem Verhältnis der dunkelfarbigen Kreisteile herauslesen.

Das *Einheitsquadrat* (Abb. 5; vgl. Eichler & Vogel, 2009) liefert ebenfalls eine visuelle Umsetzung der Information (sogar in zweierlei Gestalt):

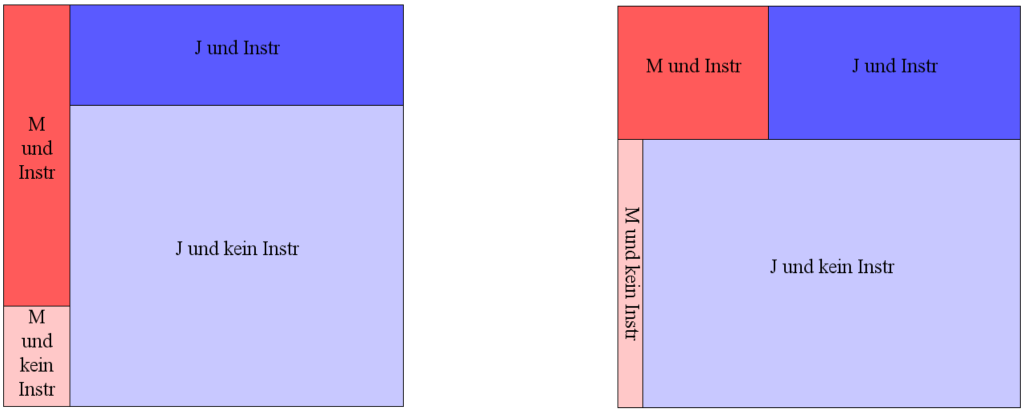


Abb. 5: Betrachtung der Zahlenverhältnisse im Einheitsquadrat

Das Kreisdiagramm hat wie das Einheitsquadrat den Nachteil, wenig variabel zu sein: Immer, wenn die Eingangsdaten geändert werden, ist ein neues Diagramm zu erstellen. Gleichwohl: Ein fertiges Diagramm mag einigen Schülerinnen und Schülern das Verständnis erleichtern. Dies gilt auch für das folgende *Piktogramm*:

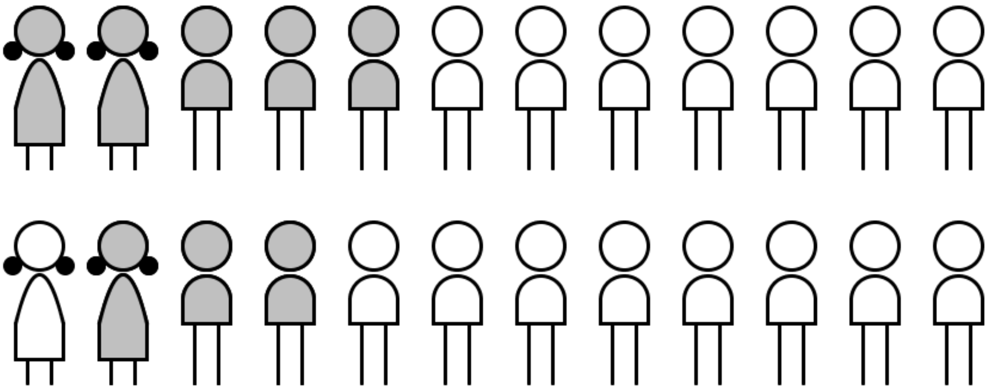


Abb. 6: Betrachtung der Zahlenverhältnisse in einer Piktogramm-Darstellung

Der Unterricht sollte auf jeden Fall Vierfelder-Tafel und Baumdiagramm pflegen, um die Schülerinnen und Schüler nicht auf eine einzige Lösungsmöglichkeit zu „trimmen“. Im Unterricht zeigte sich nämlich immer wieder, dass manche Lernenden lieber mit dem einen Format arbeiten und andere mit dem anderen Format.

Wie beim Eingangsbeispiel bietet es sich auch bei Beispiel 3 an, die Schülerinnen und Schüler selber Aufgaben entwickeln zu lassen, die eine unvermutete Lösung haben. Hier ist darüber zu reden, woran die Unvermutetheit bei Beispiel 3 liegt: 75 % der Mädchen und nur 25 % der Jungen spielen ein Instrument. Sofort wird man daraus schließen, dass es dann wohl viel mehr Spielerinnen eines Instruments gibt als Instrumentalisten. Das aber ist dann nicht aufrecht zu erhalten, wenn die Anzahl der Mädchen deutlich geringer ist als die Gesamtanzahl der Jungen. Die relativen Angaben führen in die Irre; sie müssen durch die Angabe der absoluten Grundgesamtheiten ergänzt werden:

75 % der Mädchen sind weniger Personen als 25 % der Jungen.

Das ist auch ein entscheidender Punkt in dem folgenden Kapitel.

Quantitative Schlüsse

Erst nach der qualitativen Anbahnung behandle ich Aufgaben, in denen quantitative Schlüsse zu ziehen sind. Eine typische Aufgabe (mit zunächst ganz überraschendem Ergebnis) lautet wie folgt:

<u>Beispiel 4</u>: Früher mussten sich alle Lehrerinnen und Lehrer jedes Jahr röntgen lassen, um festzustellen, ob sie Tuberculose (Tbc) hatten. Manche Leute haben dann später vom Gesundheitsamt einen Brief bekommen: „Sie haben vermutlich Tbc. Kommen Sie sofort vorbei".

Die Röntgentests waren nur zu 90 % zuverlässig: Bei 90 % der Leute, die in Wirklichkeit gesund waren, sagte, der Test, dass sie „gesund" seien; 10 % der in Wahrheit Gesunden hielt der Test für „krank". Bei 90 % der Leute, die in Wirklichkeit krank waren, sagte, der Test, dass sie „krank" seien; 10 % der in Wahrheit Kranken hielt der Test für „gesund". In der gesamten Bevölkerung war Tbc nur zu 1 % verbreitet.

Mit welcher Wahrscheinlichkeit ist jemand, bei dem der Test „krank" sagt, tatsächlich krank?

Diese vermutlich bekannte „Aids-Aufgabe" habe ich mit Absicht nicht mit Aids formuliert (man braucht einen Massentest, bei dem die Anzahl der in Wahrheit Kranken recht klein ist). Ebenfalls mit Absicht verwende ich im Unterricht auch keine Begriffe wie „Sensitivität" und „Spezifizität", um den Blick der Schülerinnen und Schüler auf das Wesentliche zu lenken. Mit der Verwendung der Anführungszeichen bei der Unterscheidung von „krank" / krank habe ich hinsichtlich eines intuitiven Verstehens bessere Erfahrungen im Unterricht gemacht als mit der Gegenüberstellung von Test-positiv / Test-negativ (die einen kontraintuitiven Charakter hat: Es ist positiv, wenn jemand Test-negativ ist). Um die Lernenden dazu zu bringen, die Problemhaftigkeit gewisser Situationen besser zu erkennen, wird man im Unterricht die abschließende Fragestellung in der Aufgabe weglassen.

Entsprechend wie schon im vorausgehenden Beispiel liegt hier ist die Verwechslung der beiden Aussagen

Wenn wirklich krank, dann mit 90 %-iger Wahrscheinlichkeit Test-„krank"

und

Wenn Test-„krank", dann mit 90 %-iger Wahrscheinlichkeit wirklich krank

nahe. Für den Unterricht stellt sich die Frage: Wie lassen sich die im Text gegebenen Informationen übersichtlich darstellen? Hier handelt es sich nicht mehr (wie noch in Abschnitt 1) um ein Problem der beschreibenden Statistik, was es meiner Erfahrung nach noch schwerer macht.

Gute Erfahrungen habe ich im Unterricht damit gemacht, zunächst so zu tun, als handele es sich bei Beispiel 4 um beschreibende Statistik. Dazu geht man zum Beispiel von einer Bevölkerung mit 1000 Leuten aus. Wenn davon 1 % krank sind, dann haben erwartungsgemäß etwa 10 Leute Tbc. Das werden nicht immer genau 10 Leute sein, aber etwa 10 Leute. Bei den 990 gesunden Leuten sagt der zu 90 % zuverlässige Test, dass erwartungsgemäß etwa 99 „krank“ sind. Auch hier gilt: Das werden nicht genau 99 sein, aber sicherlich auch nicht 50 oder 200. (Die relativen Abweichungen von den erwarteten Werten werden kleiner, wenn man die Gesamtpopulation vergrößert.) Mit diesen Überlegungen lassen sich die in Beispiel 4 gegebenen Informationen in einer Vierfelder-Tafel bzw. einem Doppelbaum übersichtlich darstellen (die Benutzung von Kreisdiagramm oder Einheitsquadrat führt hier zu unübersichtlichen Resultaten, da sich die Flächenanteile 891 und 1 so krass unterscheiden.):

	krank	gesund	
„krank“	≈ 9	≈ 99	≈ 108
„gesund“	≈ 1	≈ 891	≈ 892
	≈ 10	≈ 990	1000

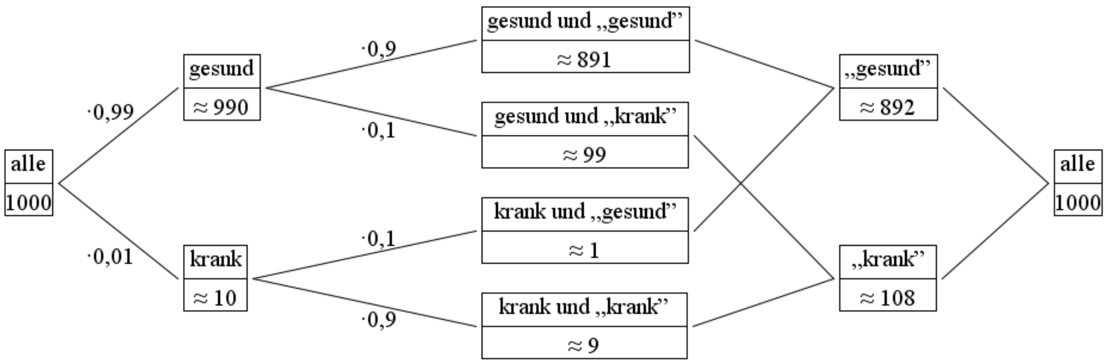


Abb. 7: Vierfelder-Tafel und Doppelbaum zur „Tbc“-Aufgabe

Es wird deutlich: Von den etwa 100 Leuten, die laut Test „krank“ sind, sind nur etwa 10 Personen wirklich krank. (Dass das wiederum etwa 10 % sind, ist reiner Zufall.) Trotzdem sind solche Tests sinnvoll: Sie sind billig, und die relativ wenigen „Kranken“ können sich anschließend noch einmal sorgfältiger (und damit aufwändiger) testen lassen. Auch dieses Beispiel lässt sich vielfältig variieren (wie schon oben sei auf viele weitere Beispiele in Cukrowicz et al., 2008 und Meyer, 2008 verwiesen). Die Erfahrung, dass das Arbeiten mit fiktiven ganzen Zahlen den Schülerinnen und Schülern leichter fällt als das Rechnen mit Wahrscheinlichkeiten, habe ich im Unterricht mittlerweile immer wieder gemacht. Rechnet man konsequent mit Wahrscheinlichkeiten, so kommt man natürlich zum gleichen Ergebnis:

$$\mathrm{prob}(k\,|\,"k") = \frac{\mathrm{prob}(k)\cdot \mathrm{prob}("k"\,|\,k)}{\mathrm{prob}(g)\cdot \mathrm{prob}("k"\,|\,g) + \mathrm{prob}(k)\cdot \mathrm{prob}("k"\,|\,k)} = \frac{0{,}01\cdot 0{,}9}{0{,}99\cdot 0{,}1 + 0{,}01\cdot 0{,}9} \approx 8{,}3\%$$

(Statt „prob“ kann man natürlich auch „p“ oder „P“ schreiben; ich habe mich für „prob“ entschieden, da „p“ in dieser Art von Aufgaben häufig ein Parameter ist. Ich habe nicht „P“ genommen, weil es schreibtechnisch sehr ähnlich zu „p“ ist.)

Absolute Häufigkeiten helfen schon bei den Pfadregeln

Das Rechnen mit fiktiven absoluten Häufigkeiten ist schon bei der Einführung der Pfadregeln in früheren Klassenstufen sinnvoll:

Beispiel 5: Jens hat eine Urne mit 3 roten und 5 blauen Kugeln und zieht zweimal ohne Zurücklegen. Mit welcher Wahrscheinlichkeit bekommt Jens zwei rote Kugeln?

Ich gebe anfangs den Schülerinnen und Schülern zunächst vor (das sollte später variiert werden), dass sie sich vorstellen sollen, dass Jens dieses Urnenexperiment 560-mal durchführt. Dann wird Jens bei der ersten Ziehung erwartungsgemäß etwa 210-mal eine rote Kugel ziehen. In diesen etwa 210 Fällen gilt: Es sind noch 2 rote und 5 blaue Kugeln in der Urne. Dann wird Jens bei der zweiten Ziehung erwartungsgemäß etwa 60-mal wieder eine rote Kugel ziehen.

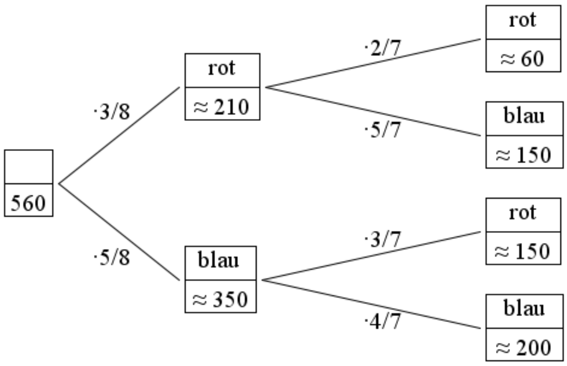


Abb. 8: Absolute Häufigkeiten bereits bei Einführung der Pfadregeln

Über diese Betrachtung von absoluten Häufigkeiten wird die Wahrscheinlichkeit für zwei rote Kugeln mit $\frac{60}{560} = \frac{6}{56}\left(= \frac{3}{8} \cdot \frac{2}{7}\right)$ leichter berechnet.

Variation der Parameter in Beispiel 4

Zurück zur Tbc-Aufgabe! Hier bieten sich Parameter-Variationen an, um zu untersuchen, welche der Variablen eine Rolle spielt. Belässt man die Zuverlässigkeit der Tests bei 10 % und ändert den Anteil der in Wahrheit Kranken von 1 % auf x %, so ergibt sich: $\text{prob}(k \mid \text{"k"}) = \frac{9 \cdot x}{8 \cdot x + 100}$. Hierbei ist $x = \text{prob}(k)$.

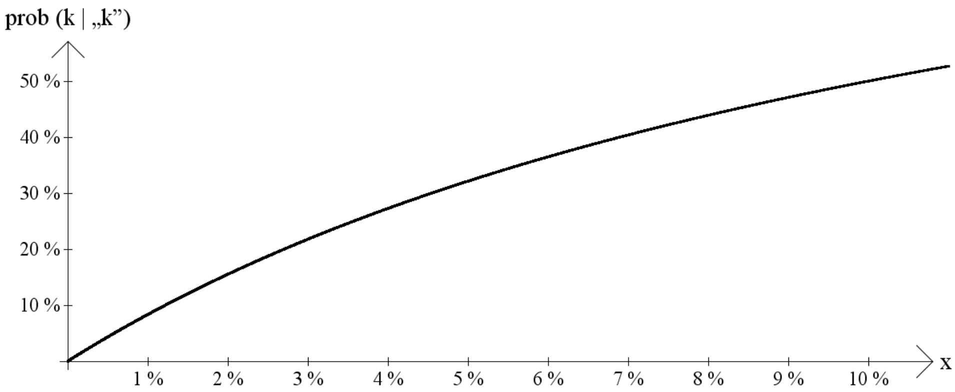


Abb. 9: Wie hängt die Lösung der Tbc-Aufgabe vom „wahren“ Krankenstand ab?

Ein Diagramm wie in Abbildung 9 ist im Unterricht für die Schülerinnen und Schüler eine wichtige Ergänzung zur bloß formelhaften Berechnung von Einzelwerten. Sie können leichter Aussagen überblicken, die hinter der Formel stehen, wie z.B.: Erst wenn $\text{prob}(\text{k}) > 10\%$ ist, ist die Wahrscheinlichkeit, wirklich krank zu sein(, wenn der Test „krank“ sagt), auf knapp 50% gestiegen. Belässt man dagegen $\text{prob}(\text{k})$ bei 1 % und ändert die Zuverlässigkeit des Tests von 90 % ab auf y %, so ergibt sich

$$\text{prob}(\text{k} \mid \text{"k"}) = \frac{y}{9900 - 98 \cdot y}.$$

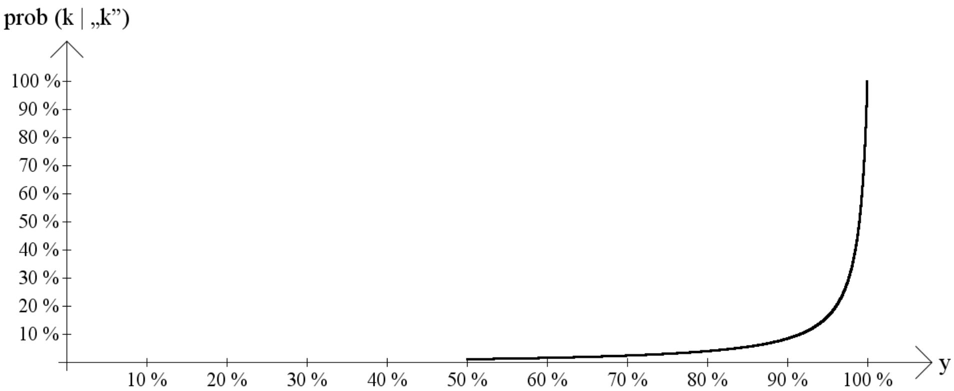


Abb. 10: Wie hängt die Lösung der Tbc-Aufgabe von der Zuverlässigkeit des Tests ab?

Das Ergebnis ist auch intuitiv klar, und das sollte auch abseits der Berechnungen und Diagrammbetrachtungen besprochen werden: Je schlechter die Tests sind, umso kleiner ist die Wahrscheinlichkeit, wirklich krank zu sein, wenn der Test „krank“ sagt. (Biologen oder Mediziner werden hier einwenden, dass die Zuverlässigkeit eines Tests aus den beiden Bestandteilen $\text{prob}(\text{"k"} \mid \text{k})$ und $\text{prob}(\text{"g"} \mid \text{g})$ bestehe (g steht hier für gesund), die man nicht unabhängig voneinander variieren könne. Dieser Einwand besteht zu Recht. Im Sinne einer einfachen Modellierung kann er erst einmal unberücksichtigt bleiben. Wenn Lehrperson und Lernende wollen, können sie die Modellierung verfeinern; das Verständnis der funktionalen Zusammenhänge wird dadurch m.E. eher erschwert.)

Ein Missverständnis

Vor einiger Zeit saß ich hinten im Unterricht, während ein Referendar vorne folgendes Beispiel behandelte:

Beispiel 6: Man hat zwei Urnen. Urne 1 enthält 1 rote und 3 weiße Kugeln. Urne 2 enthält 3 rote und 1 weiße Kugel. Meike wählt eine Urne verdeckt aus und zieht aus ihr eine Kugel; diese ist weiß. Mit welcher Wahrscheinlichkeit hatte Meike Urne 1 gewählt?

Er rechnete gleich mit Wahrscheinlichkeiten und argumentierte so: Man hat insgesamt 8 Kugeln. Mit Wahrscheinlichkeit $\frac{4}{8}$ hat Meike eine weiße Kugel gezogen. Es gibt 4 weiße Kugeln. 3 davon sind aus Urne 1. Also ist $\frac{3}{4}$ die gesuchte Wahrscheinlichkeit. Der Vollständigkeit halber skizzierte er auch das Baumdiagramm (Abb. 11).

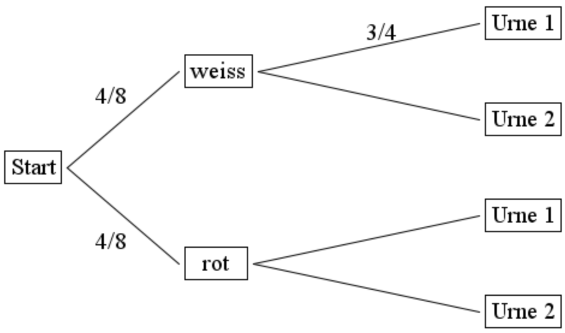


Abb. 11: Das Baumdiagramm zum falschen Lösungsweg

Warum ist dieser Lösungsweg falsch? Das Ergebnis ist doch richtig! Das Baumdiagramm in Abb. 12 liefert unter Verwendung von $\text{prob}\left(\text{Urne }1\,|\,\text{weiss}\right)=\dfrac{\frac{1}{2}\cdot\frac{3}{4}}{\frac{1}{2}\cdot\frac{3}{4}+\frac{1}{2}\cdot\frac{1}{4}}=\dfrac{3}{4}$ das gleiche Resultat.

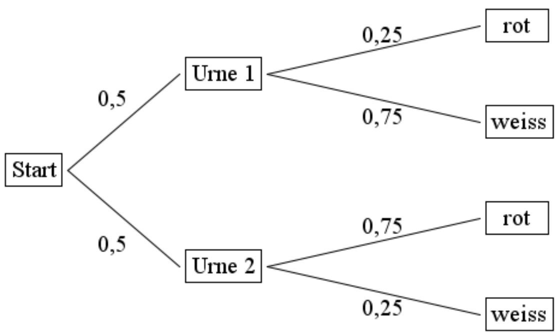


Abb. 12: Das Baumdiagramm zum richtigen Lösungsweg

Auch in der Stochastik können falsche Wege zum (zufällig) richtigen Resultat führen! Um deutlich zu machen, dass der Weg in Abb. 11 tatsächlich falsch sein muss, kann man daher nicht mit einem abweichenden Resultat argumentieren. Stattdessen schafft eine Variation der Aufgabenstellung Klarheit:

<u>Variation zu Beispiel 6</u>: Man hat zwei Urnen. Urne 1 enthält 1 rote und 1 weiße Kugeln. Urne 2 enthält 996 rote und 1 weiße Kugel. Meike wählt eine Urne verdeckt aus und zieht aus ihr eine Kugel; diese ist weiß. Mit welcher Wahrscheinlichkeit hatte Meike Urne 1 gewählt?

Der an Abb. 11 orientierte falsche Lösungsweg geht so: Man hat insgesamt 1000 Kugeln. Mit Wahrscheinlichkeit $\frac{2}{1000}$ hat Meike eine weiße Kugel gezogen. Es gibt 2 weiße Kugeln. Eine davon ist aus Urne 1. Also ist $\frac{1}{2}$ die gesuchte Wahrscheinlichkeit. Das Ergebnis ist offensichtlich falsch: Da Urne 2 fast nur rote Kugeln enthält, ist es sehr unwahrscheinlich, aus ihr eine weiße Kugel zu ziehen. Selbst wenn die Schülerinnen und Schüler (aus Zufall) den falschen Weg von Abb. 11 nicht beschreiten, würde ich ihn als Lehrer thematisieren.

Dahinter steckt das Problem: Womit fängt man bei dem Baumdiagramm an, mit Urne 1/Urne 2 oder mit der gezogenen Farbe? Es ist immer eine gute Idee, chronologisch vorzugehen: Erst wird die Urne ausgewählt, damit fängt man also an. Beispiel 6 lässt sich ausbauen und verschafft einen ersten Einblick in das Lernen aus Erfahrung.

Lernen aus Erfahrung

Die Frage, welche Urne in Beispiel 6 ursprünglich ausgewählt worden war, lässt sich zuverlässiger beantworten, wenn viele Ziehungen vorgenommen werden. Um die Sache nicht zu einfach werden zu lassen, sollen die Einzelziehungen jeweils mit Zurücklegen erfolgen; sonst könnte man ja die beiden Urnen auch einfach auskippen. Sehen wir uns ein Szenario an:

<u>Zweite Variation zu Beispiel 6</u>: Meike wählt insgeheim eine verdeckte Urne aus (und bleibt die nächsten Ziehungen immer bei ihrer gewählten Urne). Meike zieht eine weiße Kugel und legt sie wieder zurück. Mit Wahrscheinlichkeit 0,75 stammt sie aus Urne 1 und mit der Gegenwahrscheinlichkeit 0,25 aus Urne 2.

Diese Information können die Schülerinnen und Schüler verarbeiten, bevor ein zweites Mal gezogen wird:

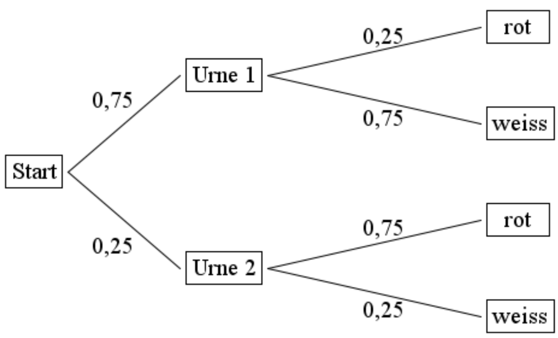


Abb. 13: Baumdiagramm zur ersten Ziehung in Beispiel 6

Nun zieht Meike eine rote Kugel und legt sie wieder zurück. Jetzt weiß jeder: Zu Beginn hatte Meike mit Wahrscheinlichkeit $\frac{0,75 \cdot 0,25}{0,75 \cdot 0,25 + 0,25 \cdot 0,75} = \frac{1}{2}$ die Urne 1 gewählt. In diesem Fall weiß man nach der zweiten Ziehung genauso viel wie vor der ersten Ziehung. Wenn dagegen Meike als zweite Kugel eine weiße zieht, so weiß jeder, dass sie zu Beginn mit Wahrscheinlichkeit $\frac{0,75 \cdot 0,75}{0,75 \cdot 0,75 + 0,25 \cdot 0,25} = \frac{9}{10}$ die Urne 1 gewählt hatte. Auch diese Information wird gleich verarbeitet:

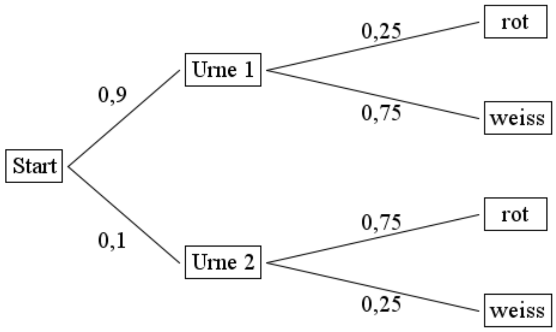


Abb. 14: Baumdiagramm zur zweiten Ziehung in Beispiel 6

Die Fortsetzung dieses Verfahrens wird nach meiner Erfahrung den Lernenden schnell offensichtlich. Wenn Meike zu Beginn Urne 1 ausgewählt hatte, so sind im Laufe der Zeit 3-mal so viel weiße wie rote Kugeln zu erwarten. Eine Simulation zeigt, dass man sich dann aufgrund der jeweils gezogenen Kugeln auch immer sicherer sein kann, dass tatsächlich Urne 1 gewählt wurde; allerdings kann es vorkommen und die Schülerinnen und Schüler sollten diese Beobachtung unbedingt machen, dass man nach den ersten Ziehungen vorübergehend der Urne 2 zuneigt. Im Diagramm ist nach rechts die Anzahl der Ziehungen und nach oben die Wahrscheinlichkeit, dass anfänglich Urne 1 gewählt worden war, aufgetragen.

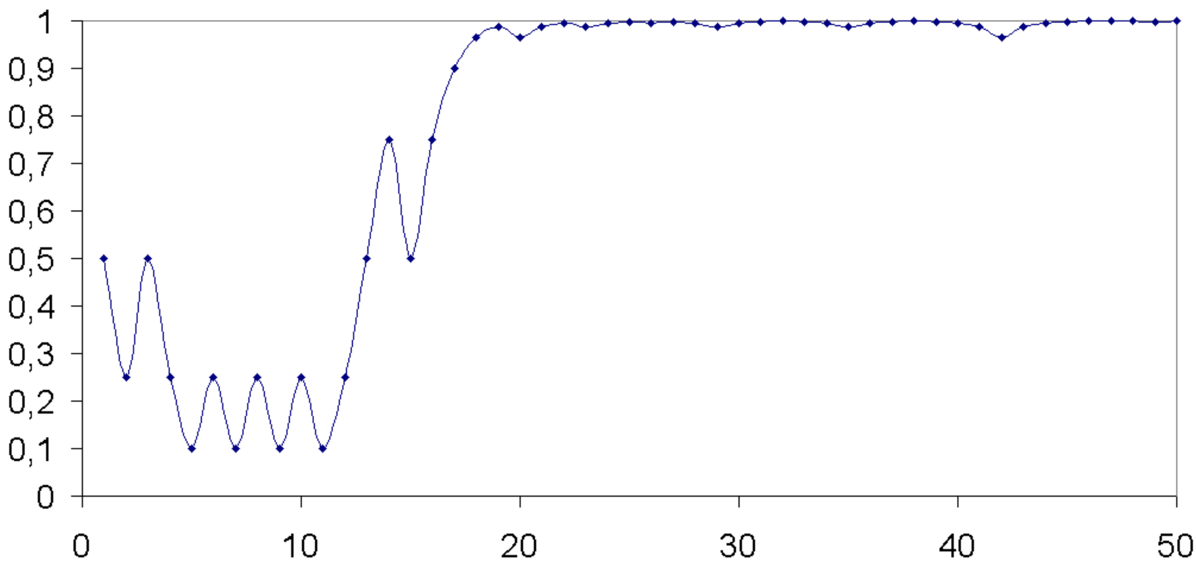


Abb. 15: Nach vielen Ziehungen hat man ein klareres Bild

Hier könnte man gleich „frequentistisch“ über die Häufigkeiten des Auftretens argumentieren: Wenn man nach sehr vielen Ziehungen etwa 3-mal so viele weiße wie rote Kugeln hat, so spricht das für Urne 1; wenn man 3-mal so viele rote wie weiße Kugeln hat, für Urne 2. Was macht man aber, wenn nach 50 Ziehungen kein klares Ergebnis vorliegt? Eine Möglichkeit besteht darin, weitere Ziehungen zu beobachten.

Das hier beschriebene Verfahren hat den didaktischen Vorteil, dass die Schülerinnen und Schüler schon bei wenigen Ziehungen wenigstens die Wahrscheinlichkeit dafür angeben können, dass Urne 1 gewählt wurde. Dabei ist diese Wahrscheinlichkeit als „Grad der inneren Überzeugtheit“ aufzufassen. Das muss auch unbedingt besprochen werden, ebenso dass auf diese Weise auch Lernen aus Erfahrung funktioniert: Man kann (aus technischen Gründen) nicht zu jedem strittigen Problem sehr viele Experimente veranstalten, sondern muss sich mit dem vorliegenden begrenzten Erfahrungsschatz begnügen. Daraus sind dann mehr oder weniger überzeugende Ansichten zu entwickeln, um „welche Urne“ es sich im übertragenen Sinn wohl handeln könne. Beispiele dafür sind den Schülerinnen und Schülern leicht nahe zu bringen: Dass die Wahrscheinlichkeit für höhere Temperaturen im Sommer groß ist, haben sie aus Erfahrung gelernt, aber auch den Charakter eines Menschen lernt man nur durch Erfahrung kennen. Ob die so entwickelte Theorie der „Wahrheit“ entspricht, wird man nie beweisen können, aber andere Methoden stehen nicht mehr zur Verfügung, seit Eva den vermaledeiten Apfel gegessen hat.

Rückblick

Wir haben einen langen und hoffentlich abwechslungsreichen Weg zurückgelegt: Ausgehend von paradox anmutenden bivariaten Situationen über das Rückwärtsschließen in Baumdiagrammen oder Vierfelder-Tafeln hin zum Lernen aus Erfahrung. Es ist vielleicht sichtbar geworden, dass der Titel dieses Aufsatzes nicht zu viel versprochen hat, dass nämlich zweistufige Zufallsexperimente durchaus interessante Eigenschaften haben können.

Literatur

Cukrowicz, Jutta/ Theilenberg, Joachim / Zimmermann, Bernd (Hrsg.) (2008): MatheNetz Klasse 9. Ausgabe N. Westermann Schulbuchverlag,

Eichler, Andreas/ Vogel, Markus (2009): Leitidee Daten und Zufall. Vieweg+Teubner, Wiesbaden.

Meyer, Jörg (2008): Bayes in Klasse 9. In: Eichler, Andreas / Meyer, Jörg (Hrsg.): Anregungen zum Stochastikunterricht, Band 4. Franzbecker, Hildesheim, S. 123 - 135.