Denkfallen vermeiden – Am Beispiel des Geschwisterproblems

STEFAN BARTZ, MECKEL

***Zusammenfassung**: Anhand von konkreten Aufgaben werden typische Denkfallen der Stochastik vorgestellt und Hilfen angeboten, wie sich diese vermeiden lassen. Im vorliegenden Artikel geht es um typische Fehlerquellen bei abhängigen Ereignissen.*

Das Geschwisterproblem

> Sie wissen, dass Ihr Kollege 2 Kinder hat. Zufällig treffen Sie ihn in der Stadt und er hat eine Tochter dabei. Wie wahrscheinlich ist es, dass das andere Kind auch eine Tochter ist? (Motzer 2008)

Wir betrachten im Folgenden also nur Väter, die in Begleitung *eines* Kindes sind und genau 2 Kinder haben. Uns interessiert die bedingte Wahrscheinlichkeit, dass das andere Kind eine Tochter ist, falls der Vater eine Tochter dabei hat, kurz: $P_{T\text{-}dabei}\,(aKiT)$.

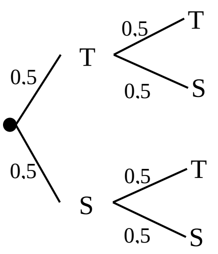
Argumentation 1 $P_{T\text{-}dabei}\,(aKiT) = 1/2$

Da das Geschlecht der gesehenen Tochter keinen Einfluss auf das Geschlecht des anderen Kindes hat, sind beide Ereignisse (*T-dabei* und *aKiT*) voneinander unabhängig. Und da Jungen- und Mädchengeburten ungefähr gleichwahrscheinlich sind, gilt:

$$P_{T\text{-}dabei}\,(aKiT) = P(aKiT) = 0{,}5.$$

Argumentation 2 $P_{T\text{-}dabei}\,(aKiT) = 1/3$

Nach der „gesehenen" Tochter gibt es nur noch 3 mögliche Fälle: (T-T), (T-S), (S-T); sie sind mit 25% alle gleichwahrscheinlich. Uns interessiert von diesen 3 möglichen Fällen nur der (T-T)-Fall. Dessen Wahrscheinlichkeit muss somit $^1/_3$ betragen:



Fragen

- Beide Argumentationen sind falsch. Wo genau liegen die Denkfehler?
- Wie kann man Schülern helfen, solche Aufgaben so zu lösen, dass sie nicht in derartige Denkfallen tappen?

Denkfehler

Leichtfertig unterstellte Unabhängigkeit

Vertreter von Argumentation 1 kennen vermutlich Aufgaben vom Typ „Wie groß ist die Wahrscheinlichkeit, dass das 6. Kind einer Familie mit 5 Töchtern wieder eine Tochter wird?" und übertragen diese Lösung fälschlicherweise auf das vorliegende Problem. Dort beeinflussen die bereits vorhandenen 5 Töchter nicht die Wahrscheinlichkeit für die 6. Tochter – die Ereignisse sind voneinander unabhängig[1]. Hier ist es jedoch anders. Dadurch, dass der Vater ein Kind nach gewissen Kriterien zum Spaziergang auswählt, stehen die beiden Ereignissen *T-dabei* und *aKiT* in einer Beziehung zueinander, über die sie abhängig sein *könnten*. Bei solchen Aufgabenkonstellationen darf man Unabhängigkeit nur dann behaupten, wenn man sie auch korrekt nachweist. Das geschieht jedoch nicht und somit ist es reine Glückssache, dass hier tatsächlich Unabhängigkeit vorliegt und die Gleichung aus Argumentation 1 angewendet werden darf. Und es verwundert nicht, dass Argumentation 1 bei den Problemvarianten "Treffpunkt Modenshow", "Ohne Auswahlvorgang" und "Kästchenproblem" (s.u.) sofort zu falschen Resultaten führt.

Leichtfertiges Drauflloszählen

Schüler kennen am besten die Laplace-Wahrscheinlichkeit (Anzahl der interessierenden durch Anzahl aller möglichen Ausgänge). Folglich versuchen sie – manchmal krampfhaft – interessierende und mögliche Fälle zu identifizieren und abzuzählen. Dass diese "Fälle" gleichwahrscheinlich sein müssen und alle bedingenden Vorereignisse mit einbezogen werden müssen, wird bei komplexen Situationen leicht übersehen. So sind die Fälle (T-T), (T-S), (S-T) zwar für sich genommen mit je 25% gleichwahrscheinlich. Nicht jedoch, wenn man das *dazu* abhängige Vorereignis *T-dabei* mit einbezieht.
Der gleiche Fehler wird beim Ziegenproblem gemacht: Natürlich gibt es nach der Hilfe des Moderators nur noch 2 Tore, von denen eines das interessierende Auto enthält. Trotzdem ist die Wahrscheinlichkeit, das Auto zu treffen, nicht ½. Dies würde nur gelten, wenn *keine* zusätzlichen Informationen vorhanden wären (Indifferenzprinzip). Der Kandidat hat jedoch zusätzliche Informationen erhalten.
Das (leichtfertige) Abzählen von Ausgängen ist auch deshalb so beliebt, weil in der Regel keine andere Lösungsstrategie bekannt ist.

[1] Untersuchungen zeigen, dass selbst dabei keine vollkommene stochastische Unabhängigkeit vorliegt.

Interessierendes Ereignis verfehlt

In Argumentation 2 wird das interessierende Ereignis nicht exakt getroffen. Dort hat man sich auf alle spazierengehenden 2-Kind-Väter bezogen, die mindestens eine Tochter *haben*, und nicht auf all diejenigen, die eine Tochter *dabei haben*. 25% aller spazierengehenden, von *einem* Kind begleiteten, 2-Kind-Väter sind *2T*-Väter und 50% sind *1T*-Väter. Jedoch haben von diesen 50% der *1T*-Väter nur die Hälfte ihre Tochter dabei, die anderen gehen mit ihrem Sohn spazieren. Bei den *2T*-Vätern haben *alle* ihre Tochter dabei. Insgesamt trifft man also mit gleicher Wahrscheinlichkeit auf tochterbegleitete *2T*-Väter und tochterbegleitete *1T*-Väter. Die gesuchte Wahrscheinlichkeit $P_{\text{T-dabei}}(aKiT)$ muss folglich $^1/_2$ betragen. In Argumentation 2 wurde fälschlicherweise $P_{\text{meKiT}}(aKiT)$ statt $P_{\text{T-dabei}}(aKiT)$ bestimmt (s.u. "Ohne Auswahlvorgang").

Empfohlener Lösungsweg

Die oben besprochenen, typischen Denkfallen lassen sich durch folgendes Vorgehen vermeiden.

1) Formuliere das interessierende Ereignis exakt

Zunächst sollte immer das interessierende Ereignis, möglichst exakt und entsprechend des Aufgabentextes formuliert werden:

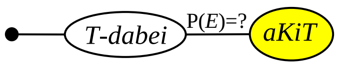
E: anderes Kind ist Tochter, falls ein von einem Kind begleiteter 2-Kind-Vater eine Tochter dabei hat.

Erst danach sollte man knappere Formulierungen bzw. Abkürzungen einführen:

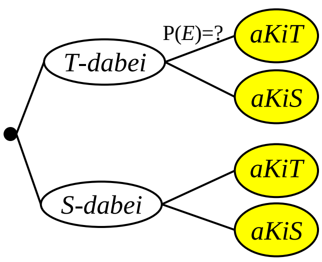
E: aKiT, falls T-dabei

2) Lasse das interessierende Ereignis anhand eines Baumpfades schrittweise "passieren"

Hier ist es wichtig, dass genau das oben formulierte Ereignis anhand eines Baumpfades entsteht:

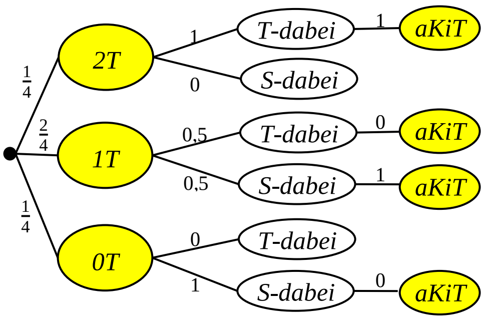


3) Vervollständige den Baum



4) Ermittle die Astwahrscheinlichkeiten

Da man keine stochastische Unabhängigkeit voraussetzen darf, können zunächst keine Astwahrscheinlichkeiten gefunden werden. Eine Umkehrung oder eine Erweiterung (Fallunterscheidung) des Baums hilft dann in der Regel weiter:



Jetzt sind alle Astwahrscheinlichkeiten bekannt, wenn man davon ausgeht, dass die *1T*-Väter, wenn sie ein Kind in die Stadt mitnehmen, gleichhäufig Sohn und Tochter wählen. (Die Anwendung des Indifferenzprinzips ist hier erlaubt, da keine weiteren Informationen diesbezüglich vorliegen.)

5) Berechne die gesuchte Wahrscheinlichkeit

Die gesuchte Wahrscheinlichkeit lässt sich nun mit Hilfe des Satzes von Bayes bestimmen:

$$P(E) = P_{T\text{-}dabei}(aKiT) = \frac{P(T\text{-}dabei \cap aKiT)}{P(T\text{-}dabei)}$$

$$= \frac{\frac{1}{4}\cdot 1\cdot 1 + \frac{2}{4}\cdot 0{,}5\cdot 0}{\frac{1}{4}\cdot 1 + \frac{2}{4}\cdot 0{,}5 + \frac{1}{4}\cdot 0} = \frac{1}{2}$$

Hält man sich an diesen Lösungsweg, können kaum Unsicherheiten bzw. Fehler entstehen. Die gesuchte Wahrscheinlichkeit, dass das andere Kind ebenfalls eine Tochter ist, beträgt 50%.[1]

Problemvarianten zum Üben

Treffpunkt Modenshow

Wie ist die Aufgabe zu lösen, wenn Sie den Kollegen mit der Tochter nicht in der Stadt sondern bei einer Modenshow treffen?

Da man davon ausgehen kann, dass wesentlich mehr Mädchen als Jungs an Modenshows interessiert sind, darf beim obigen Baum in der mittleren Zeile nicht mehr das Indifferenzprinzip angewendet werden, es sind ja nun zusätzliche Informationen vorhanden. Die Astwahrscheinlichkeit für ***T-dabei*** muss von 0,5 auf nahezu 1 erhöht werden. Die Wahrscheinlichkeit für eine weitere Tochter sinkt dann auf fast $^1/_3$ ab.[2]

Ohne Auswahlvorgang

Es ist überraschend, dass die Wahrscheinlichkeit für die zweite Tochter davon abhängt, *wo* man den Kollegen antrifft, ob in der Stadt oder bei einer Modenshow. Die 2. Stufe des Baumdiagramms, der Auswahlvorgang, ist dafür verantwortlich. In der Stadt muss man davon ausgehen, dass der Vater die begleitende Toch-

[1] Anhand des Baumdiagramms lässt sich auch die in Argumentation 1 leichtfertig behauptete Unabhängigkeit korrekt nachweisen:

$0{,}25 = P(T\text{-}dabei \cap aKiT) = P(T\text{-}dabei)*P(aKiT) = 0{,}5\cdot 0{,}5$

Dagegen sind die Ereignisse *T-dabei* und *T-Zahl* voneinander abhängig.

[2] Dies gilt allerdings nur, wenn 2T-Väter und 1T-Väter ungefähr gleich häufig Modenshows in Begleitung *eines* Kindes besuchen.

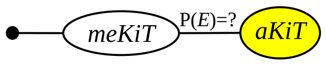
ter zufällig ausgewählt (50%). Bei Modenshows kann man dagegen annehmen, dass eher Töchter als Söhne zur Begleitung „ausgewählt“ werden.

Das Geschwisterproblem lässt sich nun so abwandeln, dass dieser Auswahlvorgang komplett entfällt. Manchmal ist nur (z.B. durch Hörensagen) bekannt, dass mindestens eines der beiden Kinder eine Tochter ist. Dadurch findet in der 2. Stufe des Baumdiagramms keine Auswahl mehr statt:

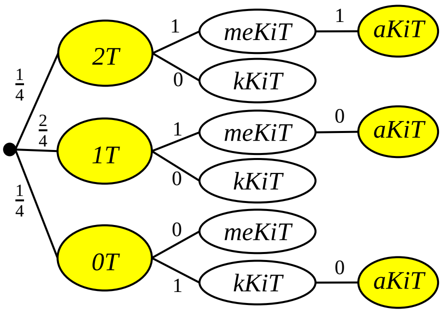
> Sie wissen, dass Ihr Kollege 2 Kinder hat und dass eines davon[1] eine Tochter ist. Wie wahrscheinlich ist es, dass das andere Kind auch eine Tochter ist?

E: das andere Kind ist auch eine Tochter, falls mindestens ein Kind eine Tochter ist

E: aKiT, falls meKiT



Erweiterter Baum:



$$P(E)^1 = P_{meKiT}(aKiT) = \frac{\frac{1}{4}\cdot 1\cdot 1 + \frac{2}{4}\cdot 1\cdot 0}{\frac{1}{4}\cdot 1 + \frac{2}{4}\cdot 1 + \frac{1}{4}\cdot 0} = \frac{1}{3}.$$

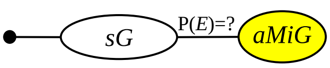
Kästchenproblem

> Ein Kästchen hat 3 Schubfächer, die je 2 Gold-, 2 Silber- bzw. 1 Gold- und 1 Silbermünze enthalten. Jedes Schubfach lässt sich nur soweit aufschieben, dass lediglich die vorderste der beiden enthaltenen Münzen erkennbar ist. Ich öffne zufällig ein Schubfach und sehe eine Goldmünze. Wie groß ist die Wahrscheinlichkeit, dass die andere Münze auch eine Goldmünze ist?

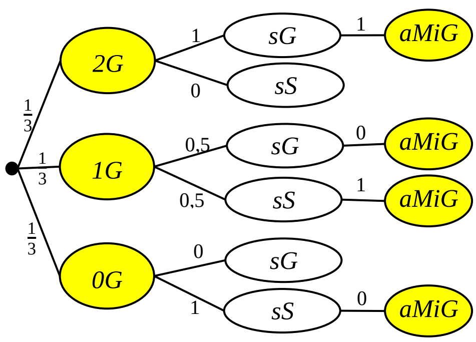
Dieses (Bertrand'sche) Kästchenproblem stimmt mit dem Geschwisterproblem überein, was am nachfolgenden Baudiagramm direkt ersichtlich ist:

E: andere Münze ist auch Goldmünze (aMiG), falls ich eine Goldmünze sehe (sG)

E: aMiG, falls sG



Erweiterter Baum:



$$P(E) = P_{sG}(aMiG) = \frac{\frac{1}{3}\cdot 1\cdot 1 + \frac{1}{3}\cdot 0{,}5\cdot 0}{\frac{1}{3}\cdot 1 + \frac{1}{3}\cdot 0{,}5 + \frac{1}{3}\cdot 0} = \frac{2}{3}$$

Fazit

Beim Bestimmen von Wahrscheinlichkeiten sollten die beiden Grundfehler, (1) leichtfertig stochastische Unabhängigkeit annehmen, (2) leichtfertig irgendwelche Fälle abzählen, unbedingt vermieden werden. Wesentlich besser ist es, wenn bei *allen* Wahrscheinlichkeitsproblemen zuerst einmal das interessierende Ereignis exakt formuliert und dieses dann anhand eines Baumpfades in Teilschritte zerlegt wird. Dadurch ergeben sich viele Vorteile:

- Das interessierende Ereignis kann so kaum aus den Augen verloren und verfehlt werden.
- Es wird sichergestellt, dass alle Vor-Informationen mit einbezogen werden.
- Man sieht beim Eintragen der Astwahrscheinlichkeiten genau, auf welche Vorereignisse sich diese beziehen, und kann Abhängigkeiten besser erkennen und nachweisen.
- Man erhält durch das Baumdiagramm einen klaren Überblick über das gesamte Problem. Komplexe Zusammenhänge können besser erfasst und einzelne Ausgänge weniger schnell übersehen werden.
- Man hat weniger Schwierigkeiten festzustellen, ob Reihenfolgeaspekte beachtet werden müssen.
- Man erkennt am Diagramm direkt, ob die Binomial- oder die hypergeometrische Verteilung zu Hilfe gezogen werden kann (Bartz 2008).
- Auch Probleme, bei denen kombinatorische Hilfsmittel oder Markov-Ketten benötigt werden, lassen sich über diesen Ansatz gut identifizieren und meistern (Bartz 2009).

Wahrscheinlichkeitsprobleme anhand von Baumdiagrammen zu lösen, ist nicht immer der eleganteste, meist jedoch der sicherste Weg.

[1] Ist dagegen das Geschlecht eines *konkreten* Kindes – z.B. des jüngsten – bekannt, beträgt die Astwahrscheinlichkeit im mittlere Pfad wieder ½, was wiederum zu $P_{jKiT}(aKiT)$ = ½ führt (bitte selbst am entsprechenden Baumdiagramm überprüfen).

Literatur

Bartz S. (2008): Baumdiagramme als roter Faden der Schulstochastik. In: *Stochastik in der Schule* 28(1). www.stefanbartz.de/materialien.htm

Bartz S. (2009): Was tun bei Mammutbäumen? In: *Stochastik in der Schule* 29(1).

Motzer, R. (2008): Zum Paradoxon der beiden Kinder. In: *Stochastik in der Schule* 28(1)

Riehl G. (2009): Ergänzungen zum Paradoxon der beiden Kinder. In: *Stochastik in der Schule* 29(3).

Wikipedia (2010): de.wikipedia.org/wiki/Ziegenproblem