Jürgen Roth

Wahrscheinlichkeiten verstehen: Kernideen und geeignete Baumdiagramme nutzen

Abstract. *Auch mit der Frage des Zugangs zur Wahrscheinlichkeit hat sich Wilfried Herget in seinem umfangreichen Werk intensiv auseinandergesetzt und wichtige stoffdidaktische Reflexionen sowie unterrichtspraktische Konzeptionen in die Diskussion eingebracht. Im Vorwort des von ihm 2008 herausgegebenen Sammelbands „Wege in die Stochastik" schreibt er dazu u.a.: „Stochastik ist angewandte Mathematik par excellence: Nicht aufs Rechnen kommt es an, sondern aufs Modellieren und Interpretieren. Stochastische Phänomene haben besondere Eigenschaften und sie erfordern eine ganz eigene Art des Denkens, um sie angemessen zu erfassen und zu beschreiben." Wenn also ein Verständnis für Wahrscheinlichkeiten entwickelt werden soll, muss die Mathematik-Brille aufgesetzt werden. Es geht um einen Zugriff auf Kernideen der Stochastik anhand von Situationen, die leicht durchschaubar sind und als Verständnisanker dienen können. Dies kann durch die Nutzung weniger, gut fassbarer Repräsentationen, wie etwa Baumdiagramme, gelingen, wenn diese flexibel als heuristische Werkzeuge genutzt und geeignet (dynamisch) dargestellt werden. Auch das muss man im Unterricht lernen. Im Beitrag wird ein Weg aufgezeigt, auf dem das Verständnis für Wahrscheinlichkeiten und Baumdiagramme gemeinsam weiterentwickelt werden, bis hin zur Nutzung von Doppelbäumen (doppelten Baumdiagrammen) im Zusammenhang mit bedingten Wahrscheinlichkeiten. Hier wird auch herausgestellt, dass die Art der Darstellung der (doppelten) Baumdiagramme eine entscheidende Rolle bei gelingender bzw. misslingender Verständnisentwicklung spielen kann.*

1. Kernideen der Stochastik

Wenn Stochastikunterricht verstehensorientiert sein soll (Bender, 1997; Herget, 1997), dann muss er sich an *Kernideen* der Inhalte und den Lernenden ausrichten (Roth, 2018). Der Fokus auf die Lernenden ist offensichtlich, da sie die Adressaten des Unterrichts sind und ihr Lernen sowie ihre kognitive Entwicklung im Mittelpunkt aller Bemühungen stehen. Kernideen umfassen nach Leuders et al. (2011) gleichzeitig kontextbezogene subjektive und fachliche Aspekte. Sie bieten einerseits *Verständnisanker* (Roth und vom Hofe, 2023) für die Vorschauperspektive der Lernenden, die das Wesen des Stoffgebiets als Ganzes schlaglichtartig erfassen und sich von da aus das Gebiet erschließen wollen. Gleichzeitig sind sie in der Rückschauperspektive der Lehrenden ein Vehikel um Konzepte der fertigen Mathematik und normative Grundvorstellungen zu Inhaltsbereichen (Siller & Roth, 2016) für lebensweltlich geprägte Vorstellungen von Lernenden fassbar zu machen. Vor diesem Hintergrund lohnt es sich, für die weitere Diskussion in diesem Beitrag wesentliche Kernideen der Stochastik zusammenzustellen.

Preprint

Roth, J. (2026). Wahrscheinlichkeiten verstehen: Kernideen und geeignete Baumdiagramme nutzen. In K. Richter & T. Fritzlar (Hrsg.), *xxx* (S. xx-xx). WTM. https://doi.org/10.xxxx/xxxx

Kernidee: Hypothetisch-prognostische Wahrscheinlichkeit

Die Kernidee der *hypothetisch-prognostischen Wahrscheinlichkeit* (vgl. Riemer, 2019, 2023; Ossadnik & Roth, 2026) umfasst zwei Perspektiven:

(1) *Prognostische Perspektive*: Wahrscheinlichkeiten sind mathematische Modelle, mit denen vorhergesagt werden kann, um welchen Wert relative Häufigkeiten bei mehrfacher Versuchswiederholung zufallsabhängig schwanken werden.

(2) *Hypothetische Perspektive*: Modelle können angezweifelt und modifiziert werden, wenn sie die Realität bei mehrfacher Versuchswiederholung nicht adäquat abbilden.

Diese zentrale Kernidee der Stochastik fokussiert insbesondere auf das Wechselspiel zwischen Realität (Empirie, relative Häufigkeit) und Modell (Theorie, Wahrscheinlichkeit). Auf der Grundlage von Zufallsversuchen bzw. aus relativen Häufigkeiten von Ereignissen in Stichproben kann man einerseits Modelle für Wahrscheinlichkeiten aufstellen, modifizieren und untersuchen. Andererseits lassen sich mit Hilfe von Wahrscheinlichkeitsmodellen in der Realität auftretende relative Häufigkeiten prognostizieren. Auch Simulationen ermöglichen auf der Basis von Wahrscheinlichkeitsmodellen virtuelle Empirie und sehr lange Versuchsreihen. Diese können bei der Reflexion und ggf. Modifizierung des Wahrscheinlichkeitsmodells unterstützen können (vgl. Abbildung 1 und Eichler, 2014).

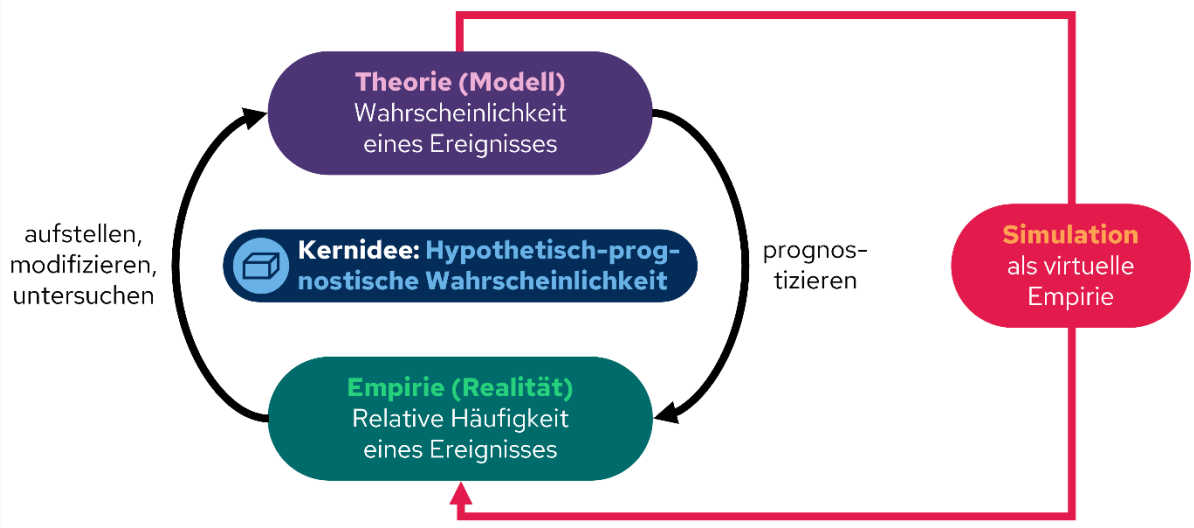

Abb. 1: Hypothetisch-prognostische Wahrscheinlichkeit als Vernetzung von Realität und Modell. (J. Roth nach Eichler, 2014)

Kernidee: Konditionale Denkfigur

Für die Stochastik ist daneben die Entwicklung der Kernidee *konditionale Denkfigur* im Sinne von „wenn …, dann …"-Strukturen wesentlich. Dies gilt etwa, wenn ich weiß, dass das Ereignis A eingetreten ist, und mich frage, wie groß dann die Wahrscheinlichkeit für das Eintreten des Ereignisses B ist. Man kann mit Wahrscheinlichkeiten auch dadurch konditional umgehen, dass man von Wahr-

scheinlichkeiten zu Anteilen (relativen Häufigkeiten) übergeht (vgl. Post und Prediger, 2020) und sich dabei jeweils fragt, was dabei das Ganze und was der Teil ist. So kann z. B. zwischen totalen Wahrscheinlichkeiten (das Ganze, die Bezugsgröße ist die Grundgesamtheit) und bedingten Wahrscheinlichkeiten (das Ganze, die Bezugsgröße ist die Teilmenge der Grundgesamtheit, die die Bedingung charakterisiert) unterschieden werden. Eine bedingte Wahrscheinlichkeit wird hier im Sinne des Anteils der Teilgruppe A an der durch B eingeschränkten Bezugsgruppe aufgefasst. Die Unterscheidung zwischen dem „Ganzen“ (der Bezugsgröße) und dem „Teil“ ist dabei von entscheidender Bedeutung.

Neben den eben genannten Kernideen gibt es mit der Kernidee *Repräsentativität und Variabilität von Stichproben-Daten* und der Kernidee *Verteilungen* noch zwei weitere für die Stochastik wichtige Kernideen, die allerdings für die Diskussionen im vorliegenden Beitrag keine entscheidende Rolle spielen. Für einen Überblick über alle vier Kernideen sei auf Ossadnik und Roth (2026) verwiesen.

2. Mehrstufige Zufallsexperimente und Baumdiagramme

Um Zugänge zur Bestimmung von Wahrscheinlichkeiten zusammen mit der Darstellungsform Baumdiagramm zu erarbeiten, bietet es sich an überschaubare Situationen zu wählen, die auch gut experimentell untersucht werden können (vgl. Herget, 1997). Im Folgenden werden drei sonderbare Würfel – ähnlich zu Efron-Würfeln (vgl. Pundsack & Schlie, 2014) – eingesetzt, die durch Bekleben von Holzwürfeln mit entsprechenden Augenzahlen leicht hergestellt und experimentell genutzt werden können (vgl. Abbildung 2). Sollte man diese Mühe scheuen, kann man notfalls auch digitale Experimente mit entsprechend gestalteten Glücksrädern (vgl. https://roth.tel/sonderbare_wuerfel) nutzen.

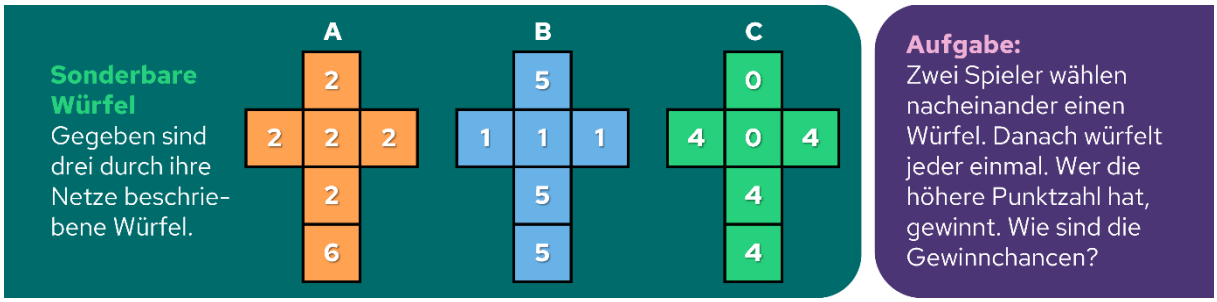

Abb. 2: Aufgabe zu sonderbaren Würfeln (J. Roth)

Zunächst sollte man das in Abbildung 2 beschriebene Spiel einige Male spielen. Die dabei gesammelten Erfahrungen legen nahe, dass es zu jedem der drei Würfel einen besseren gibt. Diese für Lernende überraschende Feststellung wird zum Anlass genommen, in Gruppen anhand von Baumdiagrammen zu bestimmen, mit welcher Wahrscheinlichkeit ein Würfel jeweils gegen einen anderen gewinnt. Die Abbildungen 3 bis 5 visualisieren die jeweiligen Ergebnisse für die Würfelpaare A-B, B-C und C-A. Dabei werden die Pfadregeln (vgl. Abbildung 6) entweder erarbeitet oder systematisch wiederholt und ihre jeweilige Herkunft festgehalten.

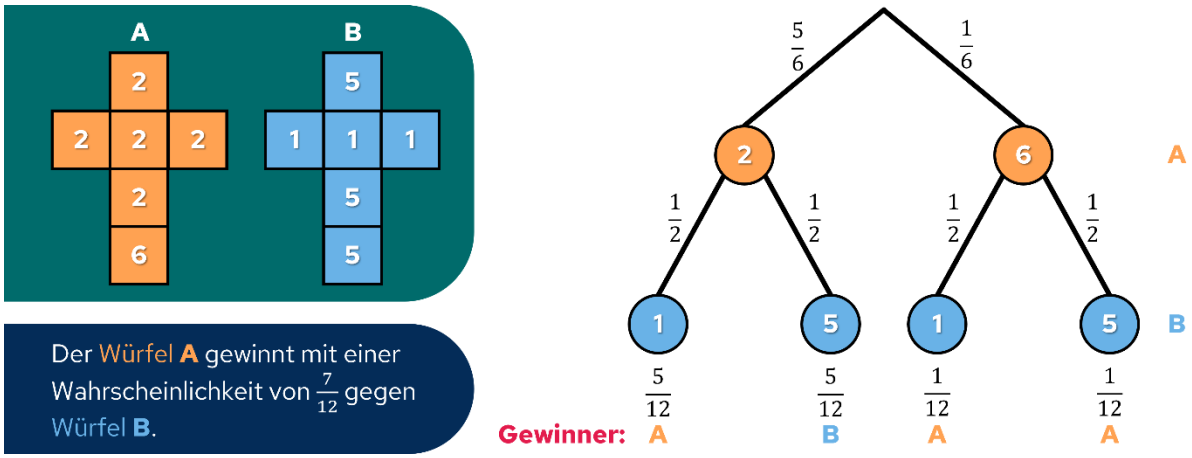

Abb. 3: Sonderbare Würfel: Baumdiagramm zu A gegen B (J. Roth)

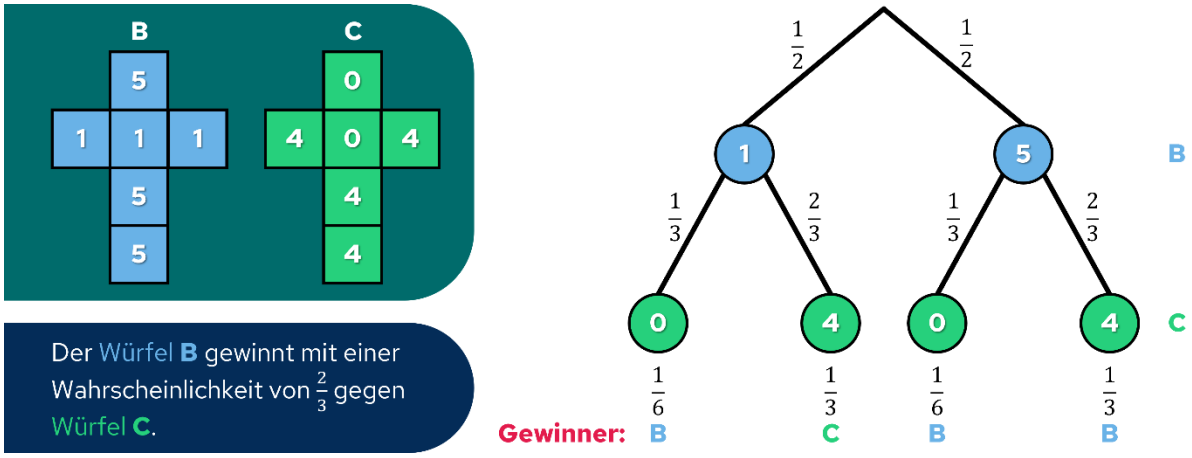

Abb. 4: Sonderbare Würfel: Baumdiagramm zu B gegen C (J. Roth)

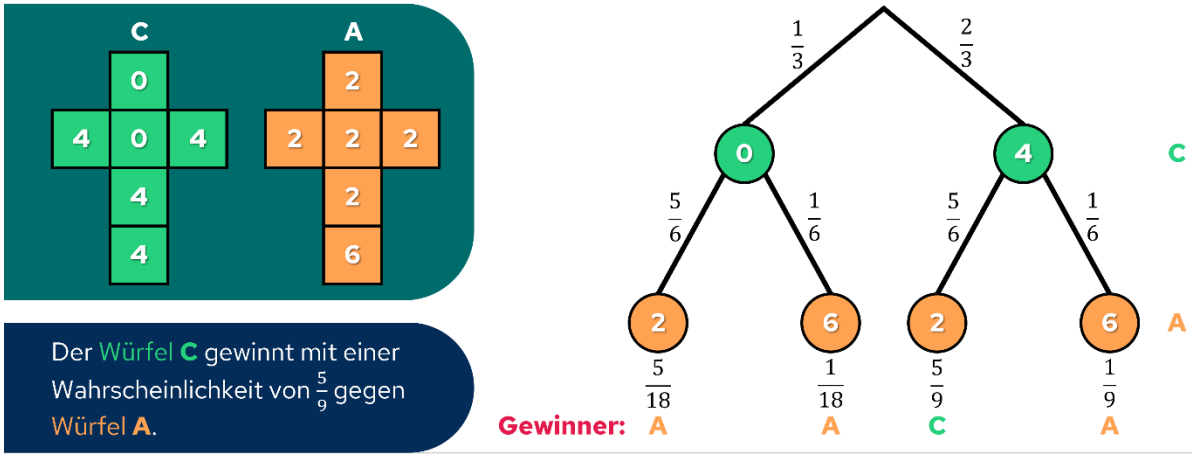

Abb. 5: Sonderbare Würfel: Baumdiagramm zu C gegen A (J. Roth)

So ergibt sich die **Produktregel** (vgl. Abbildung 6) für die Pfadwahrscheinlichkeit aus der *von-Deutung*, also der Grundvorstellung zur Multiplikation von Brüchen. So kann man sich anhand von Abbildung 4 die Pfadwahrscheinlichkeit des Pfades über 1 und 4 z. B. wie folgt klarmachen:

– Die Kernidee *hypothetisch-prognostische Wahrscheinlichkeit* hilft festzustellen, dass die Wahrscheinlichkeit mit dem **blauen Würfel** eine 1 zu würfeln $\frac{1}{2}$ ist: Da drei Seiten des Würfels mit 1 und ebenfalls drei mit 5 beschriftet

sind und wegen der Symmetrie der Spielwürfel ist das Laplace-Wahrscheinlichkeitsmodell plausibel. Analog überlegt man sich, dass die Wahrscheinlichkeit mit dem **grünen Würfel** eine **4** zu würfeln $\frac{2}{3}$ ist.

– Die Kernidee *konditionale Denkfigur* unterstützt dabei die Pfadwahrscheinlichkeit, also die Wahrscheinlichkeit $P(1 \cap 4)$, dass die Ereignisse **1** (der **blaue Würfel** zeigt die Augenzahl 1) und **4** (der **grüne Würfel** zeigt die Augenzahl 4) gemeinsam auftreten, zu identifizieren. Denken wir Wahrscheinlichkeiten im Sinne von Anteilen, dann zeigt der **blauen Würfel** in einem Anteil von $\frac{1}{2}$ aller Fälle die Augenzahl 1 und von diesem Anteil zeigt der **grüne Würfel** in einem Anteil von $\frac{2}{3}$ die Augenzahl 4. Im Sinne der von-Deutung, also der Grundvorstellung zur Multiplikation von Brüchen, ergibt sich $\frac{1}{2}$ von $1 = \frac{1}{2} \cdot 1 = \frac{1}{2}$ und $\frac{2}{3}$ von $\frac{1}{2} = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$.

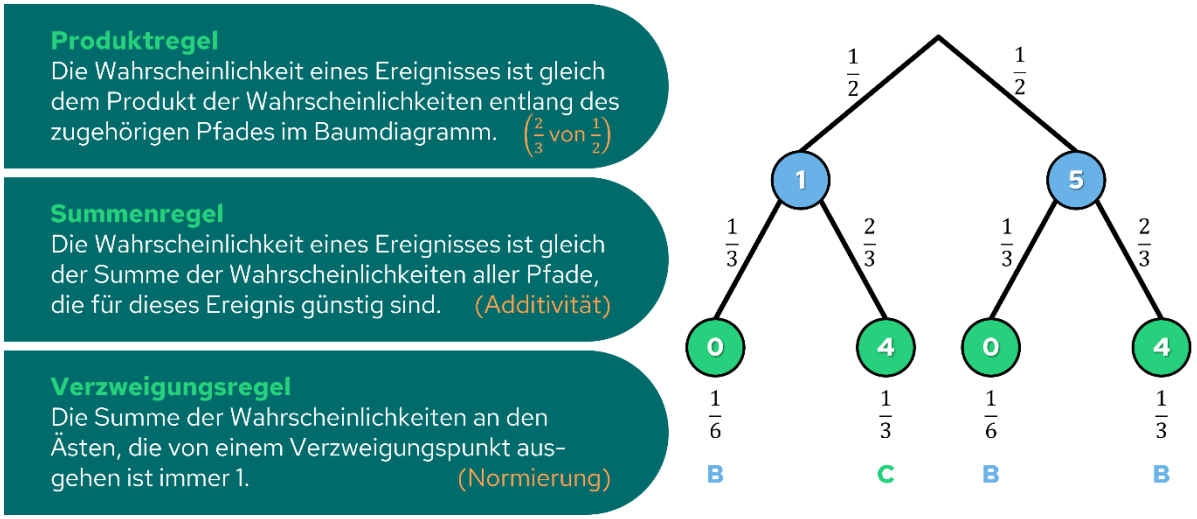

Abb. 6: Pfadregeln und ihre Herkunft (J. Roth)

Die **Summenregel** ist eine direkte Folge des Axioms der Additivität für Wahrscheinlichkeiten. Sie ist auch anschaulich plausibel, weil verschiedene Pfade, die zum selben Ereignis führen, in Abbildung 4 etwa das Ereignis, dass **Würfel B** gewinnt, sich gegenseitig nicht beeinflussen.

Die **Verzweigungsregel** (vgl. Abbildung 6) folgt direkt daraus, dass irgendetwas bei einen Zufallsversuch eintreten muss. Damit ist die Wahrscheinlichkeit, dass irgendetwas eintritt, gleich 1 (Axiom der Normierung). Die Wahrscheinlichkeiten für die einzelnen möglichen Elementarereignisse müssen sich also zu 1 summieren.

Werden die Ergebnisse der Gruppenarbeitsphase für die Gewinn-Wahrscheinlichkeiten bzgl. der Würfelpaare **A**-**B**, **B**-**C** und **C**-**A** zusammengestellt werden, zeigt sich das überraschende Gesamtergebnis von Abbildung 7: Die Gewinnwahrscheinlichkeiten sind nicht transitiv. Es gibt für jeden der drei Würfel jeweils einen besseren Würfel, dessen Gewinn-Wahrscheinlichkeit also größer ist.

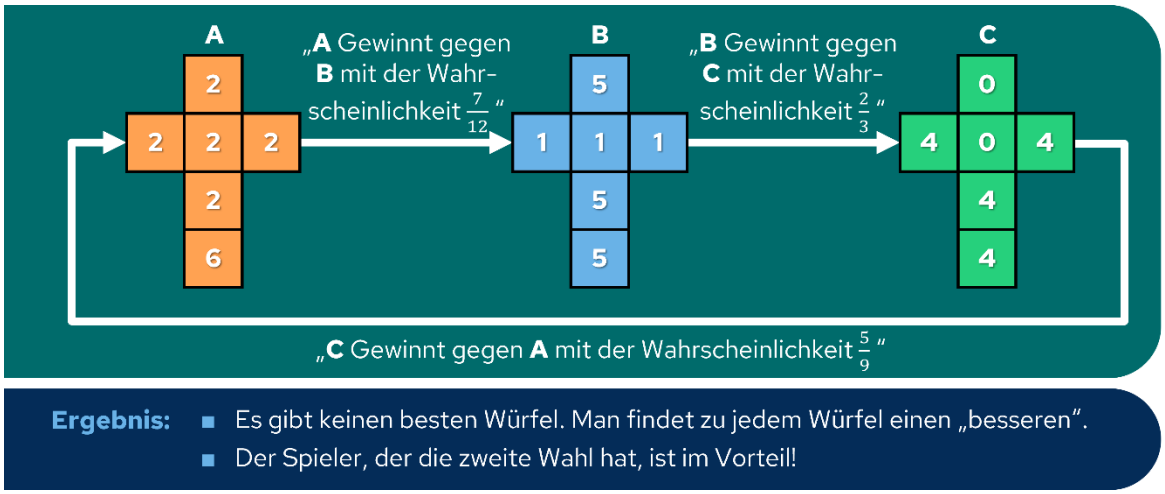

Abb. 7: Sonderbare Würfel: Wahrscheinlichkeiten sind nicht transitiv (J. Roth)

Anhand der sonderbaren Würfel lassen sich auch Problemlösestrategien für das Bestimmen von Wahrscheinlichkeiten mit Baumdiagrammen erarbeiten. Beispielsweise müsste man zur Lösung der Aufgabe aus Abbildung 8 eigentlich einen vollständiges dreistufiges Baumdiagramm zeichnen. Geht man geschickt vor und beginnt den Baum mit dem Würfel, mit der höchsten Augenzahl aller Würfel, dann kann der Teilbaum beim Blatt mit dieser Zahl (hier 6) bereits beendet werden, da die Augenzahlen aller anderen Würfel kleiner sind. Macht man anschließend mit dem Würfel weiter, der die nächsthöhere Augenzahl trägt, kann man analog vorgehen. Auf diese Weise lässt sich durch geschickte Ausführung eines unvollständigen Baumdiagramms (vgl. Abbildung 8) in manchen Fällen eine schnelle und übersichtliche Lösung für Wahrscheinlichkeitsfragen erarbeiten.

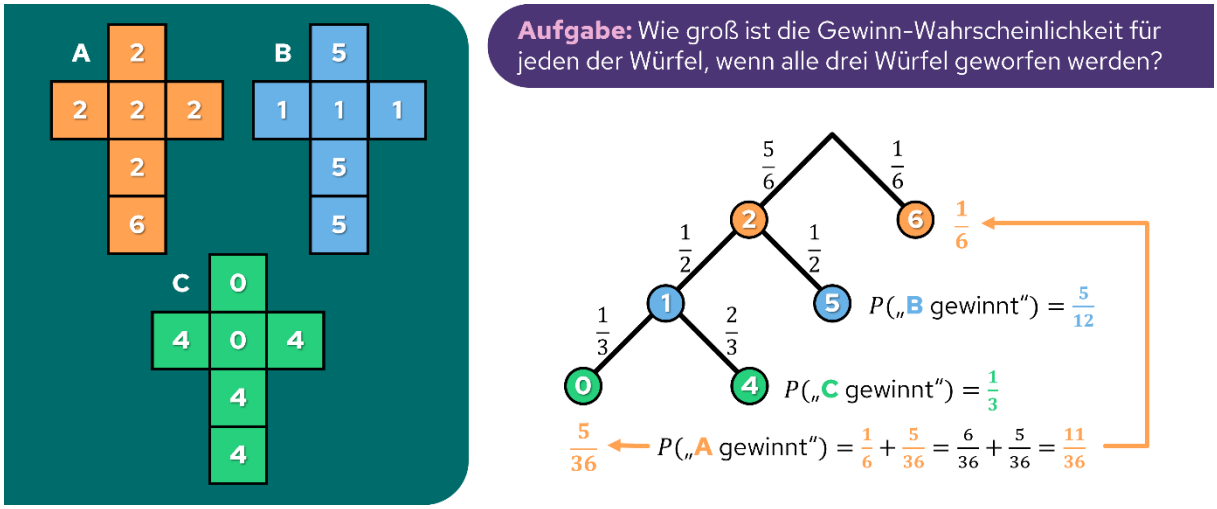

Abb. 8: Geschickt dargestelltes unvollständiges Baumdiagramm zur Problemlösung (J. Roth)

Baumdiagramme unterstützen auch bei komplexeren stochastischen Begriffen, wie stochastische (Un-)Abhängigkeit und bedingte Wahrscheinlichkeit.

3. Stochastische (Un-)Abhängigkeit und Baumdiagramme

Vielen ist intuitiv klar, dass die Wahrscheinlichkeit für Regen unabhängig von unseren Wünschen ist, aber die Zahl der Personen, die mit Regenschirm aus dem

Haus gehen, abhängig von der Wettervorhersage ist. Leider entzieht sich im Allgemeinen die Frage ob Ereignisse stochastisch abhängig oder unabhängig sind, einer intuitiven Zugriffsweise. Mit Hilfe des Urnenmodells (drei schwarze und zwei rote Kugeln) zweimaliges Ziehen mit Zurücklegen und zweimaliges Ziehen ohne Zurücklegen, bei dem der zweite Zug im ersten Fall unabhängig und im zweiten Fall abhängig vom ersten Zug ist, können Entscheidungskriterien für stochastische (Un-)Abhängigkeit identifiziert werden. Wegen des engen Zusammenhangs zur bedingten Wahrscheinlichkeit wird stochastische Unabhängigkeit zweier Ereignisse, hier S_1 „Schwarz beim ersten Zug“ und S_2 „Schwarz beim zweiten Zug“, oft ohne Nutzung der bedingten Wahrscheinlichkeit über $P(S_1 \cap S_2) = P(S_1) \cdot P(S_2)$ definiert.

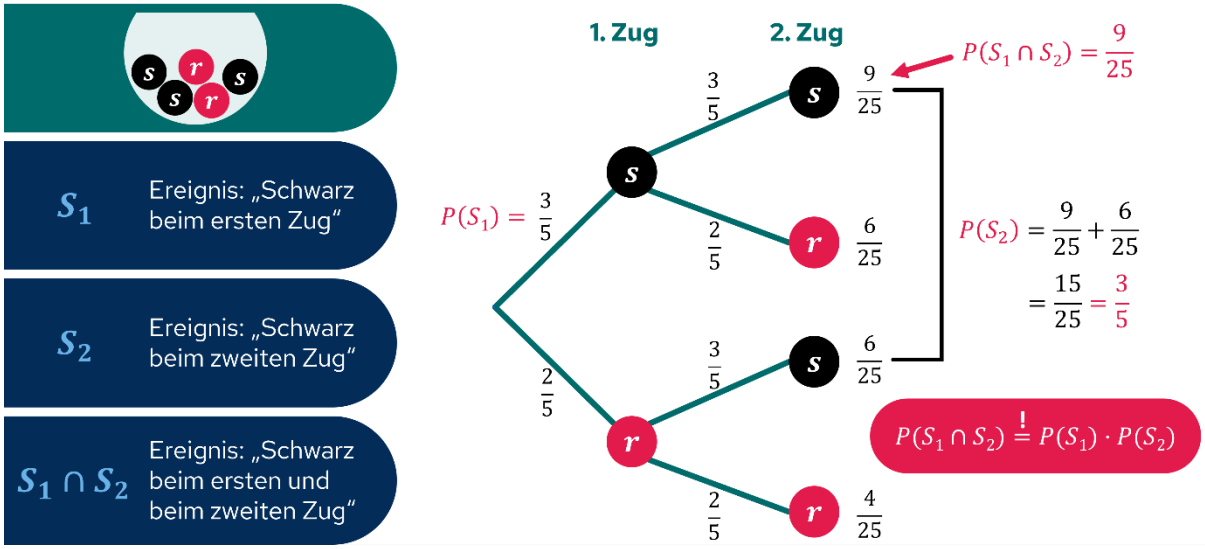

Abb. 9: *Ziehen mit Zurücklegen*: Aus einer Urne mit zwei roten und drei schwarzen Kugeln wird zweimal nacheinander eine Kugel gezogen. Vor jedem Zug wird gemischt. (Jürgen Roth)

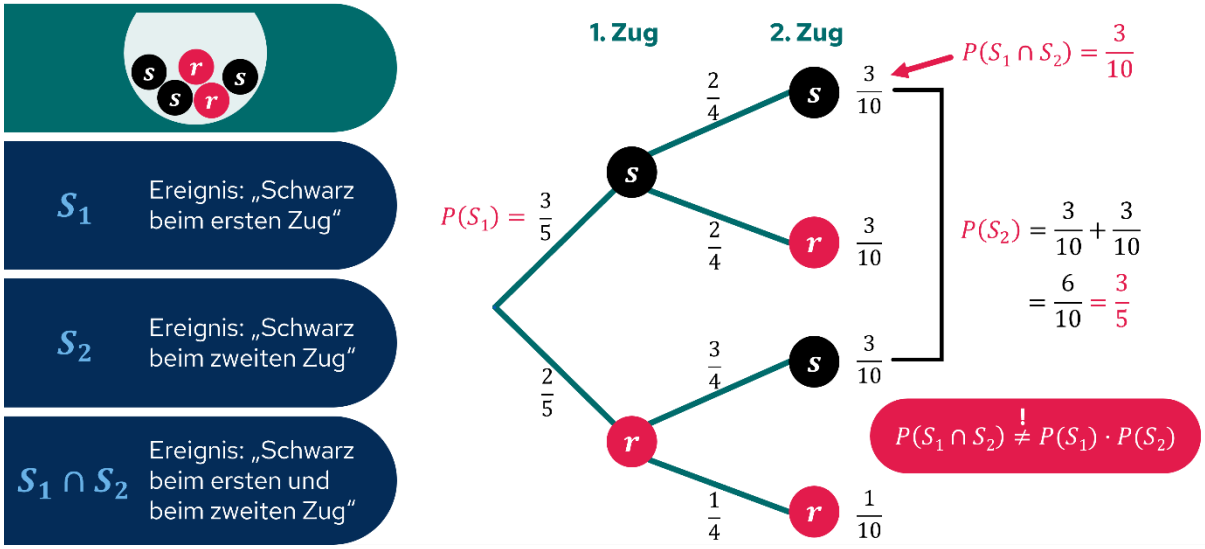

Abb. 10: *Ziehen ohne Zurücklegen*: Aus einer Urne mit zwei roten und drei schwarzen Kugeln wird zweimal nacheinander eine Kugel gezogen. Vor jedem Zug wird gemischt. (Jürgen Roth)

Anhand der beiden Baumdiagramme in den Abbildungen 9 und 10, die mit Hilfe von Laplace-Wahrscheinlichkeiten und Pfadregeln vollständig erarbeitet werden

können, und der offensichtlichen stochastischen Unabhängigkeit (Ziehen mit Zurücklegen) bzw. Abhängigkeit (Ziehen ohne Zurücklegen) für die Ereignisse S_1 und S_2 kann die Bedingung $P(S_1 \cap S_2) = P(S_1) \cdot P(S_2)$ für stochastische Unabhängigkeit plakativ herausgearbeitet und für eine Definition genutzt werden.

Die Analyse der beiden Baumdiagramme in den Abbildungen 9 und 10 zeigt noch etwas anderes: Die Wahrscheinlichkeiten, eine schwarze Kugel nach einer roten bzw. nach einer schwarzen Kugel zu ziehen, sind beim Ziehen mit Zurücklegen gleich groß und gleich der totalen Wahrscheinlichkeit für schwarz im zweiten Zug $\left(P(S_2|\bar{S}_1) = P(S_2|S_1) = P(S_2)\right)$. Beim Ziehen ohne Zurücklegen gilt das dagegen nicht. Diese Feststellung kann zum Phänomen der bedingten Wahrscheinlichkeit überleiten, die im Rückblick die Gleichung in Klammern als weitere Definition für stochastische Unabhängigkeit ermöglicht (vgl. Wiesner & Binder, 2024).

Eine interessante Zugangsweise zu stochastischer (Un-)Abhängigkeit über das Würfeln mit einem fairen Würfel haben Tomlinson und Quinn (1998) vorgeschlagen. Dabei wird einerseits mit dem leicht zugänglichen Laplace-Modell für Wahrscheinlichkeiten gearbeitet, das es ermöglicht Wahrscheinlichkeiten vergleichsweise intuitiv zu bestimmen. Andererseits werden Ereignisse ausgewählt, bei denen es intuitive nicht möglich ist, Aussagen über die stochastische Ab- bzw. Unabhängigkeit zu treffen. Es handelt sich um folgende Ereignisse bzgl. der Augenzahl: ***X***: „Gerade Zahl“, ***Y***: „Zahl größer als 2“ und ***Z***: „Primzahl“. Über die Darstellung der Seiten des Würfels in einem Rechteck können alle relevanten totalen und Schnitt-Wahrscheinlichkeiten direkt abgelesen und daraus die interessierenden stochastischen (Un-)Abhängigkeiten bestimmt werden (vgl. Abbildung 11). Wenn die Darstellungen der Würfelaugen als Venn-Diagramme interpretiert werden, kann diese Vorgehensweise grundsätzlich verallgemeinert werden.

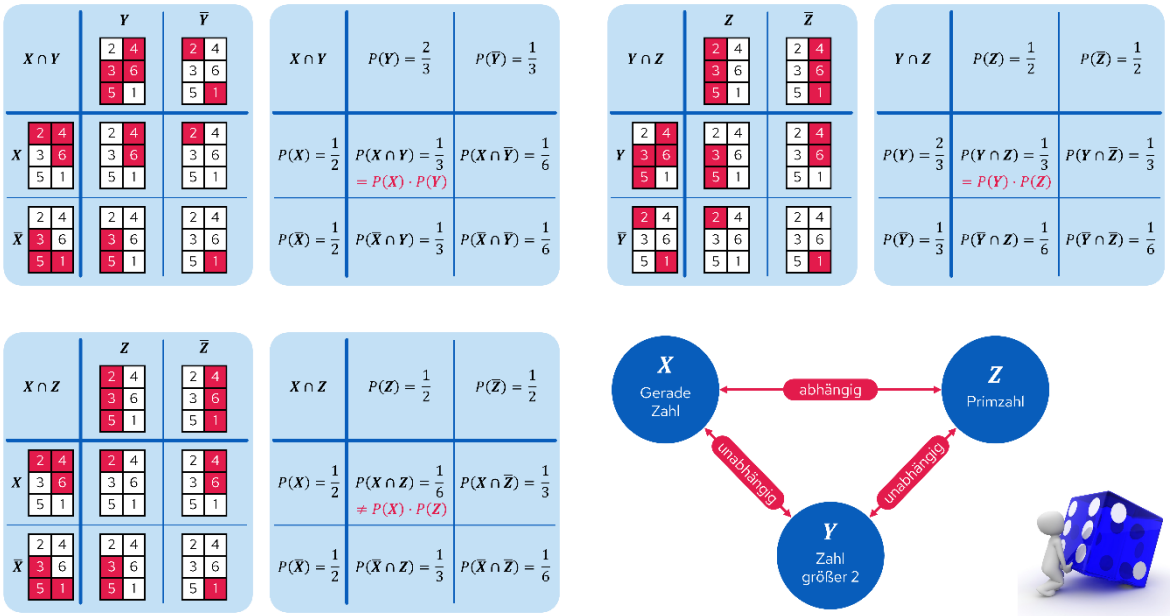

Abb. 11: *Würfeln mit einem fairen Würfel*: Stochastische Ab- bzw. Unabhängigkeit der Ereignisse bzgl. der Augenzahl: ***X***: „Gerade Zahl“, ***Y***: „Zahl größer 2“, ***Z***: „Primzahl“. (J. Roth)

4. Bedingte Wahrscheinlichkeiten in Baumdiagrammen und Doppelbäumen

Die Auseinandersetzung mit stochastischer (Un-)Abhängigkeit in Baumdiagrammen führt zur Erkenntnis, dass Wahrscheinlichkeiten an Ästen der zweiten Stufe eines Baumdiagramms vom Ereignis abhängen (können), das vorher auf der ersten Stufe eingetreten ist. Weil diese Wahrscheinlichkeiten also von Bedingungen abhängen, nennt man sie *bedingte Wahrscheinlichkeiten*. Um diesem Phänomen nachzuspüren, betrachten wir zunächst weiterhin das Beispiel des Würfelns mit einem Spielwürfel und interessieren uns für die Frage, wie groß die Wahrscheinlichkeit ist, dass die Augenzahl eine Primzahl (***Z***) ist, wenn man weiß, dass die Augenzahl eine gerade Zahl (***X***) ist.

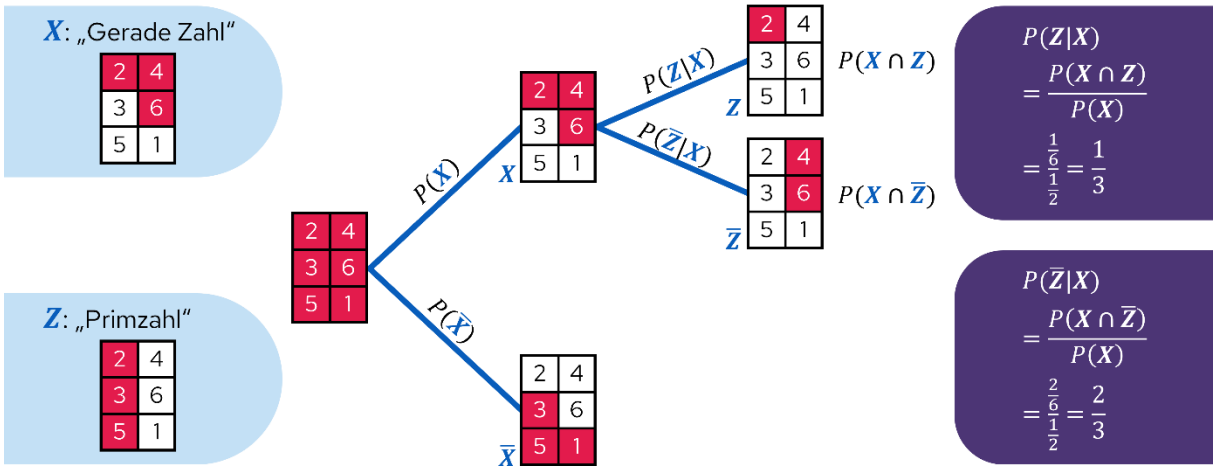

Abb. 12: *Würfeln mit einem fairen Würfel*: Bedingte Wahrscheinlichkeiten mit Baumdiagramm und Pfadregeln bestimmen. (J. Roth)

Im Baumdiagramm von Abbildung 12 kann das dabei Interessante am obersten Pfad abgelesen werden. Auf diesem Pfade finden sich drei Wahrscheinlichkeiten, die über folgende Fragen adressiert werden können (vgl. Post & Prediger, 2020):

(A) Wie groß ist die Wahrscheinlichkeit, dass beim Würfeln mit einem Spielwürfel eine geworfene Augenzahl gerade ist?

(B) Wie groß ist die Wahrscheinlichkeit, dass beim Würfeln mit einem Spielwürfel eine geworfene gerade Augenzahl eine Primzahl ist?

(C) Wie groß ist die Wahrscheinlichkeit, dass beim Würfeln mit einem Spielwürfel eine geworfene Augenzahl gerade und eine Primzahl ist?

Wie Post und Prediger (2020) betonen fällt Lernenden bereits die sprachliche Unterscheidung der drei Fragen schwer. Teilweise werden inhaltlichen Unterschiede gar nicht erkannt. Inhaltlich bezieht sich Frage **(A)** auf die *totale Wahrscheinlichkeit* $P(X)$, die am Ast der ersten Stufe des obersten Pfades steht. Frage **(B)** adressiert die *bedingte Wahrscheinlichkeit* $P(Z|X)$, also die Wahrscheinlichkeit für Z wenn man weiß, dass X eingetreten ist. Diese steht am Ast der zweiten Stufe des obersten Pfades. Frage **(C)** zielt auf die Pfadwahrscheinlichkeit, also die *Schnitt-Wahrscheinlichkeit* $P(X \cap Z)$. Das ist die Wahrscheinlichkeit, dass beide Ereignisse gleichzeitig eintreten.

In der Darstellung in Abbildung 12 können die gemeinten Wahrscheinlichkeiten im Sinne der Kernidee **Konditionale Denkfigur** durch Rückgriff auf die **rot markierten** Anteile vom jeweiligen Ganzen identifiziert werden. Für die Frage (A), bezogen auf den Ast der ersten Stufe des Pfades, ist es der Anteil von drei geraden Zahlen an allen sechs möglichen Zahlen, also $\frac{3}{6} = \frac{1}{2}$. Bei Frage (B), bezogen auf den Ast der zweiten Stufe des Pfades, geht es um den Anteil der einen Primzahl unter den drei geraden Zahlen, also $\frac{1}{3}$. Die Frage (C) adressiert die Pfadwahrscheinlichkeit, als den Anteil der einen geraden Primzahl, nämlich 2, unter allen möglichen sechs Augenzahlen des Würfels, also $\frac{1}{6}$.

Wenn nur die *bedingte Wahrscheinlichkeit* $P(\boldsymbol{Z}|\boldsymbol{X})$ aus Frage **(B)** unklar ist, kann diese aus dem Baumdiagramm über die Pfadregeln, hier die Produktregel, hergeleitet werden. Da sich die Pfadwahrscheinlichkeit $P(\boldsymbol{X} \cap \boldsymbol{Z})$ als Produkt aus den Wahrscheinlichkeiten an den Ästen des entsprechenden Pfades ergibt, gilt nach Abbildung 12 hier also konkret $P(\boldsymbol{X}) \cdot P(\boldsymbol{Z}|\boldsymbol{X}) = P(\boldsymbol{X} \cap \boldsymbol{Z})$. Durch beidseitige Division der Gleichung durch $P(\boldsymbol{X})$ folgt unmittelbar $P(\boldsymbol{Z}|\boldsymbol{X}) = \frac{P(\boldsymbol{X} \cap \boldsymbol{Z})}{P(\boldsymbol{X})} = \frac{1}{3}$.

Vervollständigt man den Baum in Abbildung 12, können an den Ästen der zweiten Stufe die vier bedingten Wahrscheinlichkeiten $P(\boldsymbol{Z}|\boldsymbol{X})$, $P(\overline{\boldsymbol{Z}}|\boldsymbol{X})$, $P(\boldsymbol{Z}|\overline{\boldsymbol{X}})$ und $P(\overline{\boldsymbol{Z}}|\overline{\boldsymbol{X}})$ abgelesen bzw. über die Produkt- oder Verzweigungsregel berechnet werden. Wie kann man aber bedingte Wahrscheinlichkeiten bestimmen, bei denen das andere Ereignis, hier $\boldsymbol{Z}$ oder dessen Gegenereignis $\overline{\boldsymbol{Z}}$, die Bedingung ist, also die Wahrscheinlichkeiten $P(\boldsymbol{X}|\boldsymbol{Z})$, $P(\overline{\boldsymbol{X}}|\boldsymbol{Z})$, $P(\boldsymbol{X}|\overline{\boldsymbol{Z}})$ und $P(\overline{\boldsymbol{X}}|\overline{\boldsymbol{Z}})$?

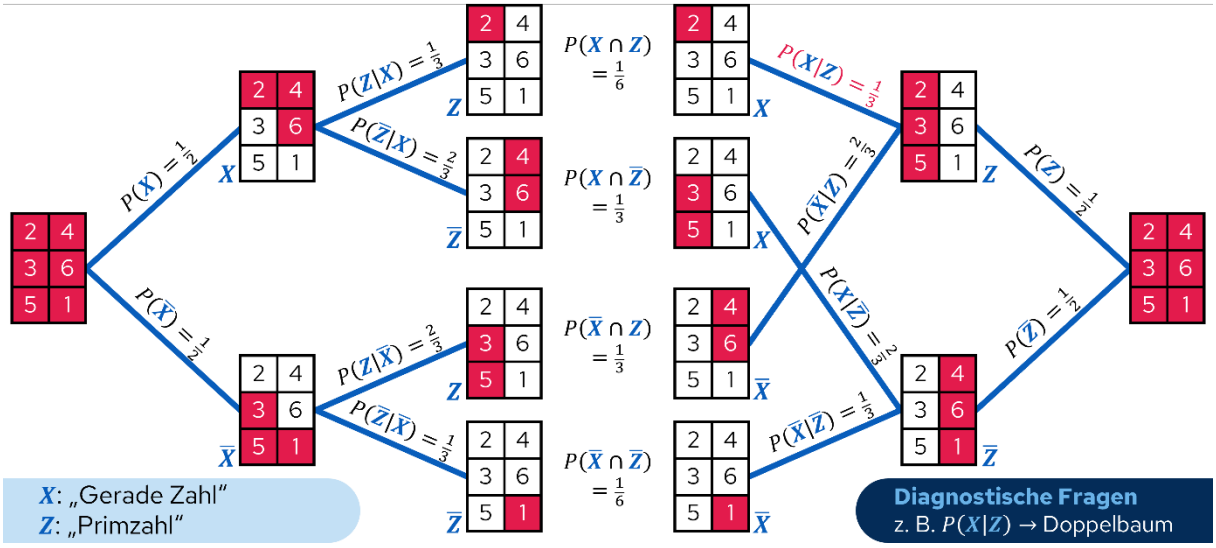

Abb. 13: *Würfeln mit einem fairen Würfel*: Doppelbaum zur Darstellung aller totalen, bedingten und Schnitt-Wahrscheinlichkeiten (J. Roth)

Offensichtlich benötigt man dafür einen anderen Baum, der in der ersten Stufe zu den Ereignissen $\boldsymbol{Z}$ bzw. $\overline{\boldsymbol{Z}}$ führt. Allerdings sind dafür oft einzelne Wahrscheinlichkeiten nicht bekannt. Wäre z. B. die Wahrscheinlichkeit $P(\boldsymbol{Z})$ nicht gegeben, so könnte sie mit der Additionsregel und den Pfadwahrscheinlichkeiten $P(\boldsymbol{X} \cap \boldsymbol{Z})$

sowie $P(\overline{\boldsymbol{X}} \cap \boldsymbol{Z})$ aus dem bisherigen Baum zu $P(\boldsymbol{Z}) = P(\boldsymbol{X} \cap \boldsymbol{Z}) + P(\overline{\boldsymbol{X}} \cap \boldsymbol{Z})$ bestimmt werden. Mit diesem Einstieg gelingt es recht gut, ein zweites Baumdiagramm zu den fehlenden bedingten Wahrscheinlichkeiten zu erstellen, wenn man den zweiten Baum entgegengesetzt zum ersten Baum zu einem doppelten Baumdiagramm, einem sogenannten Doppelbaum, aufbaut, wie in Abbildung 13. Dabei kann man sich zu nutze machen, dass die Pfadwahrscheinlichkeiten jeweils die Wahrscheinlichkeiten der Schnitte der Ereignisse auf den beiden Stufen des Baumdiagramms sind. Wenn man also einen Pfad auf der ersten Stufe beim Ereignis $\boldsymbol{Z}$ beginnt und von da aus zum Ereignis $\boldsymbol{X}$ auf der zweiten Stufe weiterläuft muss die Pfadwahrscheinlichkeit dieses Pfades die Schnitt-Wahrscheinlichkeit $P(\boldsymbol{X} \cap \boldsymbol{Z})$ sein. Diese Schnittwahrscheinlichkeit gibt es aber schon im ersten, linken Teilbaum als Pfadwahrscheinlichkeit. Der Pfad im zweiten, rechten Teilbaum sollte also bei eben dieser Schnittwahrscheinlichkeit enden (vgl. Abbildung 13). Damit hat man aber schon alles zusammen um die bedingte Wahrscheinlichkeit $P(\boldsymbol{X}|\boldsymbol{Z})$ zu bestimmen. Da sich die Pfadwahrscheinlichkeit des eben erstellten Pfades im rechten Baum nach Produktregel zu $P(\boldsymbol{Z}) \cdot P(\boldsymbol{X}|\boldsymbol{Z}) = P(\boldsymbol{X} \cap \boldsymbol{Z})$ ergibt und die Schnitt-Wahrscheinlichkeit $P(\boldsymbol{X} \cap \boldsymbol{Z})$ bereits als Pfadwahrscheinlichkeit im linken Teilbaum berechnet wurde, ergibt sich direkt $P(\boldsymbol{X}|\boldsymbol{Z}) = \frac{P(\boldsymbol{X} \cap \boldsymbol{Z})}{P(\boldsymbol{Z})}$. Alle anderen Äste und Pfade sowie die entsprechenden Wahrscheinlichkeiten des rechten Teilbaums lassen sich jeweils analog durch Ausnutzung von Pfadregeln und ggf. Rückgriff auf den linken Teilbaum bestimmen. Bei der hier genutzten Darstellung der Ereignisse als Rechteckstruktur der Augenzahlen (oder allgemeiner Venn-Diagramm) ermöglicht es die Ergebnisse auch als Anteile des jeweiligen Ganzen zu bestimmen. Dies Zugangsweise fällt einzelnen Lernenden leichter und ermöglicht eine eigenständige Überprüfung des Problemlöseprozesses mit dem Doppelbaum.

Wesentliche Überlegungen, die beim Erstellen eines Doppelbaums relevant sind und die vollständige Entwicklung eines Doppelbaums aus einem einfachen Baumdiagramm werden im Erklärvideo unter https://roth.tel/doppelbaum/video am bekannten Taxi-Problem (Tversky & Kahneman, 1975) erläutert:

> ***Taxi-Problem:*** *Ein alter Mann wird Zeuge eines Autounfalls mit Fahrerflucht und berichtet, dass das flüchtende Auto ein blaues Taxi gewesen sei (Behauptung $\boldsymbol{b}$). Es gibt zwei Taxiunternehmen in der Stadt, die 15 blaue ($\boldsymbol{B}$) bzw. 85 grüne ($\boldsymbol{G}$) Taxis haben. In der Verhandlung wird die Sehfähigkeit des Mannes geprüft und man bekommt heraus, dass er in 80% der Fälle die richtige Farbe zuordnen kann, d. h.:* $P(\boldsymbol{b}|\boldsymbol{B}) = P(\boldsymbol{g}|\boldsymbol{G}) = 0{,}8$.

Auf der Grundlage dieser Kenntnisse muss der Richter sich mit der Frage auseinandersetzen, wie groß die Wahrscheinlichkeit ist, dass wirklich ein blaues Taxi in den Unfall verwickelt war, wenn der alte Mann sagt es war ein blaues Taxi. In unserer Notation sucht der Richter also den Wert der bedingten Wahrscheinlichkeit $P(\boldsymbol{B}|\boldsymbol{b})$.

Das Ergebnis dieses Prozesses, den zugehörigen Doppelbaum, zeigt Abbildung 14. An ihm wird deutlich, dass

- an den Ästen der ersten Stufen der beiden Teilbäume *alle vier totalen Wahrscheinlichkeiten*,
- an den Ästen der zweiten Stufen der Teilbäume *alle acht bedingten Wahrscheinlichkeiten* und
- zwischen den beiden Teilbäumen als Pfadwahrscheinlichkeiten sowohl des Pfades im linken Teilbaum als auch des Pfades im rechten Teilbaum, der jeweils dorthin führt, *alle vier Schnitt-Wahrscheinlichkeiten* enthalten sind.

Im Ergebnis liefert ein so aufgebauter Doppelbaum also alle sechzehn beteiligten Wahrscheinlichkeiten. Der Doppelbaum in Abbildung 14 wurde so gestaltet, dass auch absolute Häufigkeiten vorkommen, weil eine ganze Reihe von Lernenden damit besser umgehen können als mit Wahrscheinlichkeiten. Dazu wurde eine geeignete große Zahl (hier 10.000) als Eingangsgröße, hier die Anzahl der beobachteten Taxis, gesetzt und damit weitergearbeitet. Wie diese absoluten Anzahlen beim Aufbau des Doppelbaums genutzt werden können, wird ebenfalls im Erklärvideo unter https://roth.tel/doppelbaum/video konkretisiert.

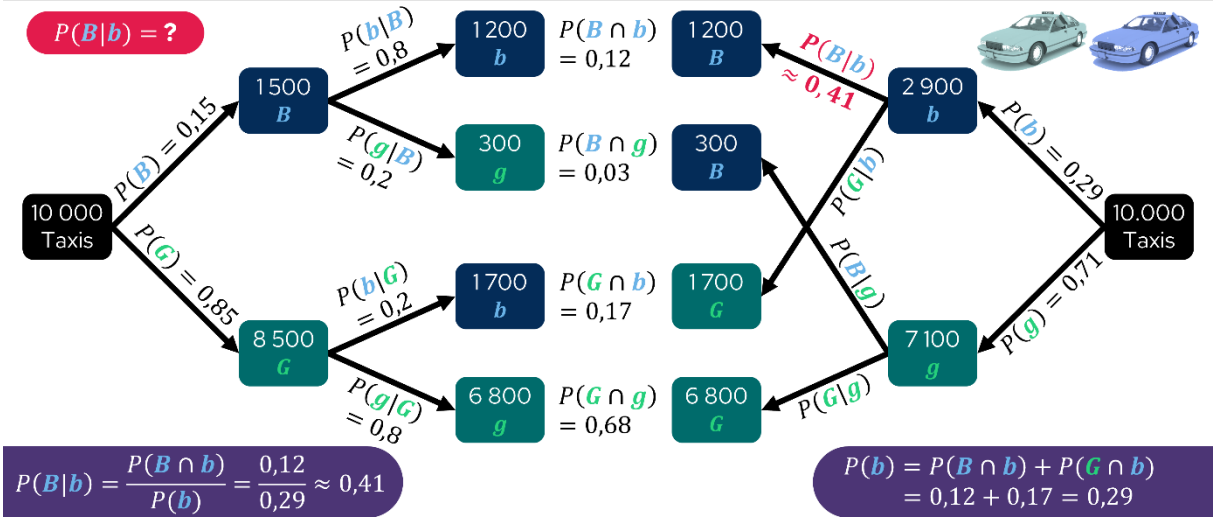

Abb. 14: Doppelbaum zum Taxi-Problem (J. Roth)

Mit Hilfe eines GeoGebra-Applets unter https://roth.tel/doppelbaum kann die Entwicklung eines Doppelbaums schrittweise visualisiert und von Lernenden als Kontrollmöglichkeit genutzt werden, nachdem sie selbstständig eigene Baumdiagramme im Rahmen des Problemlöseprozesses der Bestimmung von (bedingten) Wahrscheinlichkeiten bei zwei Ereignissen erstellt haben. Dazu können im Applet die Bezeichner für die beteiligten Ereignisse frei gewählt werden.

Sind die Ideen und Arbeitsweisen im Zusammenhang mit dem Doppelbaum verstanden, lohnt sich die Auseinandersetzung mit gesellschaftlich relevanten Fragen zu bedingten Wahrscheinlichkeiten etwa im Zusammenhang mit medizinische Tests (Pinkernell, 2006) anzugehen, die ebenfalls gut mit dem GeoGebra-Applet

zum Doppelbaum und ergänzend ggf. der dynamischen Vierfeldertafel zu medizinischen Tests (https://roth.tel/vierfeldertafel) strukturiert werden können.

5. Anmerkungen zur Darstellung von Doppelbäumen

Doppelbäume sind von Lernenden sehr gut nutzbare Diagramme, um Probleme im Zusammenhang mit bedingten Wahrscheinlichkeiten zu lösen. Empirischen Ergebnisse von Binder, Steib und Krauss (2023) zeigen, auch im Vergleich zu anderen Darstellungen, dass es sich lohnt Doppelbäume als Problemlösewerkzeuge im Unterricht zu bedingten Wahrscheinlichkeiten einzusetzen. Die positiven empirischen Ergebnisse wurden allerdings mit Doppelbäumen erzielt, die wichtige Eigenschaften der im vorliegenden Beitrag dargestellten Doppelbäume gar nicht erfüllt. Es wäre empirisch zu prüfen, ob ein Arbeiten mit der hier vorgestellten Darstellung von Doppelbäumen (vgl. die Abbildungen 13 und 14) sogar ein noch besseres Abschneiden von Lernenden bei Leistungstests ermöglichen würde als die bisher häufig in der Literatur genutzten Darstellungen von Doppelbäumen.

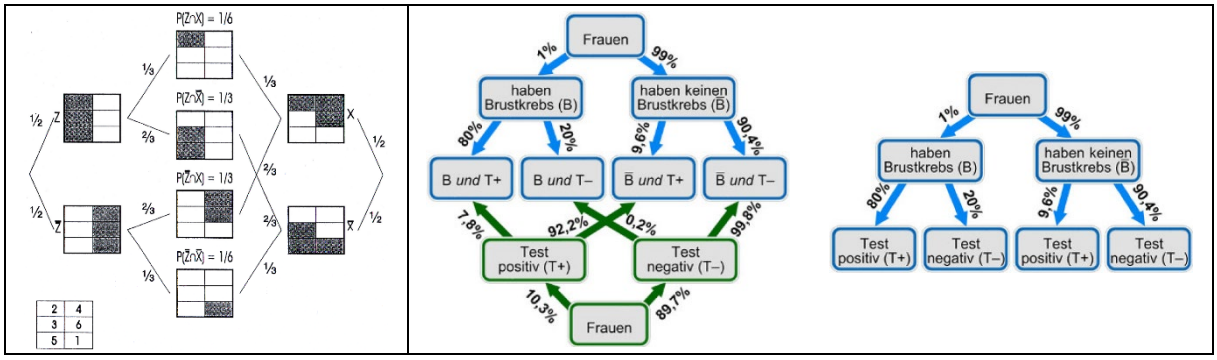

Abb. 15: Andere Doppelbäume – Links: Doppelbaum aus Tomlinson und Quinn (1998); Rechts: Doppelbaum und Baumdiagramm aus Binder et al. (2023)

Im Folgenden sollen die Unterschiede in der Darstellung von Doppelbäumen kurz erläutert werden. Abbildung 15 zeigt auf der linken Seite die Originalabbildung des Doppelbaums aus dem Artikel von Tomlinson und Quinn (1998), der inhaltlich dem Doppelbaum in Abbildung 13 des vorliegenden Beitrags entspricht. Wie man unschwer erkennt, wurden bei Tomlinson und Quinn die Blätter der zweiten Stufen der beiden Teilbäume verschmolzen. Dies hat den Nachteil, dass die Pfadwahrscheinlichkeiten, das sind die Schnittwahrscheinlichkeiten $P(\boldsymbol{X} \cap \boldsymbol{Z})$, $P(\overline{\boldsymbol{X}} \cap \boldsymbol{Z})$, $P(\boldsymbol{X} \cap \overline{\boldsymbol{Z}})$ und $P(\overline{\boldsymbol{X}} \cap \overline{\boldsymbol{Z}})$ nicht mehr als solche erkennbar sind, sondern an den Blättern „angeheftet werden“. In der Version des Doppelbaums von Binder, Steib und Krauss (2023), die ebenfalls die Blätter der zweiten Stufe verschmelzen (vgl. Abbildung 15 mittleres Diagramm), wird sogar noch weiter gegangen. Dort werden die Schnittwahrscheinlichkeiten komplett weggelassen und im Text wird sogar behaupten, dass Schnittwahrscheinlichkeiten in Doppelbäumen keinen Platz haben.

Daneben gibt es noch ein anderes Problem, das im direkten Vergleich des Baumdiagramms und des Doppelbaums aus Binder, Steib und Krauss (2023) – vgl. in

Abbildung 15 die rechten beiden Diagramme – deutlich wird: In Baumdiagrammen werden in Blättern jeweils einzelne Ereignisse bzw. deren Gegenereignisse notiert, aber keine Schnittereignisse. Warum sollte das im Doppelbaum, der sich ja aus zwei einzelnen Teilbäumen zusammensetzt, nun anders sein? Ein Überdenken der häufig gewählten Darstellung von Doppelbäumen mit Blick auf eine bessere Verständlichkeit für Lernende erscheint hier naheliegend.

Literatur

Bender, P. (1997). Grundvorstellungen und Grundverständnisse für den Stochastikunterricht. *Stochastik in der Schule, 17*(1), 8–33

Binder, K., Steib, N. & Krauss, S. (2023). Von Baumdiagrammen über Doppelbäume zu Häufigkeitsnetzen – kognitive Überlastung oder didaktische Unterstützung? *Journal für Mathematik-Didaktik, 44*(2), 471–503.

Eichler, A. (2014). Simulation als Bindeglied zwischen der empirischen Welt der Daten und der theoretischen Welt des Zufalls. In T. Wassong et al. (Hrsg.). *Mit Werkzeugen Mathematik und Stochastik lernen* (S. 251–265). Springer Spektrum.

Herget, W. (1997). Wahrscheinlich? Zufall? Wahrscheinlich Zufall .... *Mathematik lehren, 85*, 4–8.

Herget, W. (Hrsg.) (2008). *Wege in die Stochastik*. Mathematik lehren Sammelband. Friedrich.

Ossadnik, H. & Roth, J. (2026). Core Ideas for teaching hypothesis testing – Structure, concepts, and validation. In L. Birk et al. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025*, Münster. DOI: 10.52041/iase25.122

Pinkernell, G. (2006). Test positiv – Diagnose negativ. Medizinische Testergebnisse richtig interpretieren. *Mathematik lehren, 138*, 50–55

Post, M. & Prediger, S. (2020). „Nur das ‚und' ist anders, sonst ist es gleich". Mit bedingten Wahrscheinlichkeiten sprachsensibel umgehen. *Mathematik 5-10, 14*(53), 38–41.

Pundsack, F. & Schlie, S. (2014). Spielerisch zu den Pfadregeln. *Mathematik lehren, 186*, 18–22.

Riemer, W. (2019). Grundvorstellungen beurteilender Statistik: Wahrscheinlichkeit als bezweifelbares Modell. *Der Mathematikunterricht, 65*(6), 11–22.

Riemer, W. (2023). *Statistik unterrichten – Eine handlungsorientierte Didaktik der Stochastik*. Kallmeyer.

Roth, J. (2018). Wirksamer Mathematikunterricht − Ausrichtung an Kernideen der mathematischen Inhalte und den Lernenden. In M. Vogel (Hrsg.). *Wirksamer Mathematikunterricht* (S. 182–188). Schneider.

Roth, J. & vom Hofe, R. (2023). Verständnisvoll lernen – Grundvorstellungen vernetzen und Verständnisanker nutzen. *Mathematik lehren, 236*, 8–11.

Tversky, A. & Kahneman, D. (1975). Judgment under Uncertainty: Heuristics and Biases. In D. Wendt & C. Vlek (eds) *Utility, Probability, and Human Decision Making. Theory and Decision Library, vol 11* (pp. 141–162). Springer.

Tomlinson, S. & Quinn, R. (1998). Bedingte Wahrscheinlichkeit verstehen. *Stochastik in der Schule, 18*(3), 27–40.

Wiesner, P. & Binder, K. (2024). Stochastische Unabhängigkeit: Die Rolle der gewählten Definition und verschiedener Visualisierungen. *Stochastik in der Schule, 44*(3), 21–29