Natürliche Häufigkeiten zur intuitiven Einführung der bedingten Wahrscheinlichkeiten

Eine Idee für den Mathematikunterricht der gymnasialen Oberstufe

Axel Müller

7. Oktober 2017

Der Begriff der bedingten Wahrscheinlichkeit, wie er meist im Mathematik-Unterricht der gymnasialen Oberstufe eingeführt wird, bietet viel Anlass für Missverständnisse. Eine alternative Einführung ist durch die von Gigerenzer bevorzugte Darstellung mit natürlichen Häufigkeiten möglich. Ich zeige, wie die bedingte Wahrscheinlichkeit hier intuitiv als Betrachtung natürlicher Häufigkeiten auftaucht.

1 Bedingte Wahrscheinlichkeit - bekannte Schwierigkeiten

Im Stochastik-Unterricht der gymnasialen Oberstufe nimmt der Begriff der bedingten Wahrscheinlichkeit eine zentrale Stelle ein. Zur Berechnung bedingter Wahrscheinlichkeiten werden den Schülerinnen und Schülern Baumdiagramme und Vierfeldertafeln zur Verfügung gestellt. Nach einiger Übung gelingt es vielen Schülern bedingte Wahrscheinlichkeiten zu berechnen. Vorausgesetzt, sie erkennen, dass nach einer solchen gefragt ist! Das Erkennen, wann eine bedingte Wahrscheinlichkeit gesucht ist und wann nicht, bleibt jedoch schwierig. Auch die Interpretation des Begriffs in Sachsituationen ist schwierig. Entscheidendes bleibt unverstanden.
Nach meiner Ansicht liegt dies darin begründet, dass oft unklar bleibt, was die bedingten Wahrscheinlichkeit beschreibt. Der Begriff ist wenig intuitiv.
Vor einigen Jahren hörte ich einen Vortrag von Prof. Gerhard Gigerenzer zu diesem Thema. Seine Beispiele belegen, wie wenig intuitiv die bedingte Wahrscheinlichkeit ist und wie viele Fehleinschätzungen auf Grund eines falschen Verständnisses geschehen. Die gesellschaftliche Bedeutung ist frappierend (vgl. [Gigerenzer2002]). Gigerenzer macht sich dafür stark , entsprechende Situationen durch natürliche Häufigkeiten statt durch bedingte Wahrscheinlichkeiten zu beschreiben.
Ich habe den Versuch gestartet, diese Idee im Mathematikunterricht der gymnasialen Oberstufe zu nutzen und die bedingten Wahrscheinlichkeit mit Hilfe der natürlichen Häufigkeiten einzuführen. Erst nachrangig wird die Definition und die Berechnung mit Baumdiagrammen nachgeschoben. In diesem Setting tritt die bedingte Wahrscheinlichkeit völlig intuitiv auf.

Ich beschreibe zunächst die Spielsituation „Gesucht ist Person ’X’“ und erläutere danach die Umsetzung im Unterricht.

2 Gesucht ist Person ’X’!

Wir spielen ein Spiel, ein Detektivspiel. Gesucht ist Person ’X’!
Verdeckt auf einem Zettel hat der Spielleiter Geschlecht und Schuhgröße der gesuchten Person notiert. Wir sollen schätzen, ob die Person Mann oder Frau ist und welche Schuhgröße die Person hat. Dazu erhalten wir nach und nach Informationen über die Person und passen unsere Prognosen an.
Kandidaten für Person ’X’ sind 200 Personen, genau 100 Männer und 100 Frauen. Von

diesen Personen kennen wir noch die Verteilung der Schuhgrößen. Sie sind in den Tabellen 1 und 2 gezeigt.

Schuhgröße	Anzahl Männer
38	10
39	15
40	25
41	25
42	15
43	10

Tabelle 1: Schuhgrößen der Männer

Schuhgröße	Anzahl Frauen
36	10
37	20
38	25
39	25
40	15
41	5

Tabelle 2: Schuhgrößen der Frauen

Schuhgröße	Anzahl Personen
36	10
37	20
38	35
39	40
40	40
41	30
42	15
43	10

Tabelle 3: Schuhgrößen aller Personen

Wir können die Schuhgrößen aller 200 Personen auch in der gemeinsamen Tabelle 3 sammeln.

Wir legen noch drei Merkmale festlegen, die Person X haben kann:

$$A = \text{„X ist ein Mann“} \tag{1}$$
$$B = \text{„X ist eine Frau“} \tag{2}$$
$$C = \text{„X hat Schuhgröße 38“} \tag{3}$$

Aus den oben dargestellten Daten der 200 Personen können wir nun leicht die Wahrscheinlichkeiten dieser Merkmale bestimmen:

$$P(A) = \frac{100}{100+100} = \frac{100}{200} = 50\% \tag{4}$$

$$P(B) = \frac{100}{100+100} = \frac{100}{200} = 50\% \tag{5}$$

$$P(C) = \frac{35}{200} = 17{,}5\% \tag{6}$$

Bei diesen Wahrscheinlichkeiten haben wir im Zähler des Bruches die Anzahl der Personen geschrieben, die durch das Merkmal festgelegt sind und im Nenner des Bruches steht die Zahl der Personen, die insgesamt in Betracht kommen. Es sind Laplace-Wahrscheinlichkeiten.

Wenn wir nichts weiteres über Person X wissen, ist die Wahrscheinlichkeit für einen Mann ebenso hoch wie für eine Frau, jeweils 50% . Ändert sich unsere Einschätzung, wenn wir nun mehr über die Person erfahren?

Wenn ich zum Beispiel weiß, dass Person X ganz sicher die Schuhgröße 38 hat, glaube ich dann immer noch daran, dass sie genau so gut ein Mann wie eine Frau sein könnte? Nein! Jede vierte Frau hat 38, aber nur jeder zehnte Mann. Durch die zusätzliche Information ändert sich meine Einschätzung. Aus $P(A)$ wird $P_C(A)$. Aus der einfachen Wahrscheinlichkeit für A wird die **bedingte** Wahrscheinlichkeit für A, **wenn wir schon wissen, dass Merkmal C zutrifft**.

$P_C(A)$ zu berechnen ist sehr einfach. Wir müssen nur zählen, wie viele Personen durch die Merkmale C **und** A festgelegt sind und die mit allen vergleichen, die Merkmal C haben. Wir betrachten also nur noch die Personen mit Schuhgröße 38, die haben ja Merkmal C. 10 Männer haben diese Größe und 25 Frauen, also 35 Personen insgesamt. Somit gilt

$$P_C(A) = \frac{10}{10+25} = \frac{10}{35} \approx 29\% \tag{7}$$

Wieder vergleichen wir hier im Zähler die Anzahl der Personen die die Merkmale A und C haben mit allen Personen die Merkmal C haben.

Die Wahrscheinlichkeit, dass X ein Mann ist verringert sich also deutlich, wenn wir wissen, dass X Schuhgröße 38 hat. Die Wahrscheinlichkeit, dass X eine Frau ist, sollte sich demnach erhöhen, wenn wir wissen: „X hat Schuhgröße 38". So ist es auch. Hier müssen wir die 25 Frauen mit Schuhgröße 38 mit allen 35 Personen vergleichen, die diese Größe haben:

$$P_C(B) = \frac{25}{10+25} = \frac{25}{35} \approx 71\% \tag{8}$$

Bei diesen bedingten Wahrscheinlichkeit schauen wir also nicht mehr auf alle 200 Personen, sondern nur noch auf die, die Schuhgröße 38 haben. Die Bedingung verändert die relevante Grundmenge.

Dies wird noch deutlicher, wenn wir $P(C)$ mit $P_A(C)$ vergleichen. Ein Blick in Tabelle

3 liefert sofort das Ergebnis: Von 200 Personen haben 35 diese Größe, die Wahrscheinlichkeit ist also $P(A) = 17,5\%$. $P_A(C)$ ist die Wahrscheinlichkeit für „X hat Schuhgröße 38“, **wenn wir schon wissen, dass X ein Mann ist**. Um diese Wahrscheinlichkeit zu bestimmen, schauen wir nicht in Tabelle 3 nach, sondern natürlich in der Tabelle für Männer, Tabelle 1. Hier sehen wir sofort, dass sind 10 %, also $P_A(C) = 10\%$

Diese Beschreibung der bedingten Wahrscheinlichkeiten über die natürlichen Häufigkeiten macht sehr deutlich, dass das Ereignis auf die wir bedingen, uns zusätzliche Information gibt. Nicht mehr die Grundgesamtheit ist relevant sondern nur noch die Personen, die das Merkmal auf das ich bedinge erfüllen. Die Zahlen die hier auftreten können wir sofort intuitiv verstehen, es handelt sich um natürliche Häufigkeiten, wie Gerhard Gigerenzer ausführt.[1]

Auch die bekannte Formel aus der Definition der bedingte Wahrscheinlichkeit können wir hier wiederentdecken, wenn wir wollen. Wenn wir in Zeile 7 Zähler und Nenner mit 200 erweitern, so entsteht

$$P_C(A) = \frac{10}{10+25} = \frac{\frac{10}{200}}{\frac{10+25}{200}} = \frac{P(A \cap C)}{P(C)} \tag{9}$$

Das intuitive Verständnis ist hier allerdings eingeschränkt, denn wir sind von den natürlichen Häufigkeiten wieder zu den unübersichtlicheren Wahrscheinlichkeiten übergegangen.

3 Eine mögliche Umsetzung im Unterricht

Ich schreibe die Daten über die fiktive Person auf die Rückseite der umgeklappten Tafel, meist noch ergänzt durch einen Namen, Alter und einen Beruf, also zum Beispiel „Lea Braun, 32 Jahre, Kaufmännische Angestellte, Schuhgröße 38“. Danach stelle ich die Frage: „Wer ist Person X hinter der Tafel?“ und präsentiere ich die Zahlen der in Frage kommenden Personen: 100 Frauen, 100 Männer und die Tabellen über die Schuhgrößen. Anschließend notiere ich an der Tafel drei Merkmale:

$$\begin{aligned} A &= \text{„X ist ein Mann“} \\ B &= \text{„X ist eine Frau“} \\ C &= \text{„X hat Schuhgröße 38“} \end{aligned}$$

Die Schüler können die Wahrscheinlichkeiten sofort angeben. Ich halte sie an der Tafel fest und dabei auch die Berechnung als Bruch, wie oben dargestellt.
Nun frage ich, ob sich die Einschätzung für die Schuhgröße 38 verändern würde, wenn ich das Geschlecht von Person X mitteilen würde. Die Antwort ist intuitiv klar. Die entsprechende Wahrscheinlichkeit für dieses Ereignis mit Zusatzinformation schreibe ich als $P_A(C)$ bzw. $P_B(C)$ an die Tafel. Eine Berechnung ist zunächst gar nicht nötig, da

[1] vgl. Gigerenzer2002

wir die Wahrscheinlichkeiten direkt aus den Tabellen der Schuhgrößen ablesen können. Ich notiere dennoch ergänzend an der Tafel:

$$P_A(C) \quad = \quad \frac{10}{100} = 10\% \tag{10}$$

und erläutere, dass dies gelesen werden kann als „10 Männer haben 38“ und „100 Männer kommen in Frage“. Für $P_B(C)$ gehe ich genau so vor.

Nun schließt sich die Frage an, ob sich die Einschätzung des Geschlechtes ändern würde, wenn ich die Schuhgröße von Person X bekannt gäbe. Auch hier gibt es eine klare Aussage. Ja, natürlich! Bei Größe „38“ glauben wir eher an eine Frau, bei 41 eher einen Mann. Da bleiben wir nicht bei fivty-fivty.
Die Berechnung der Wahrscheinlichkeit für „X ist Mann“, **wenn ich schon weiß, „X hat Schuhgröße 38“** ist leicht. Die Schreibweise $P_C(A)$ habe ich ja schon vorher eingeführt:

$$P_C(A) \quad = \quad \frac{10}{10+25} = \frac{10}{35} \approx 28,6\%\% \tag{11}$$

Im Zähler stehen die Personen, die Mann sind und Größe 38 haben, im Zähler stehen alle Personen die 38 haben. Die Berechnung und Erläuterung von $P_C(B)$ übernehmen die Schüler, die Situation ist transparent genug.
Anschließend lasse ich von den Schülern in einer Murmelphase besprechen, worin in dieser Sachsituation die Unterschiede von $P(A)$ und $P_C(A)$ bzw. $P(C)$ und $P_A(C)$ bestehen. Vieles ist hier intuitiv klar und wir sammeln die Erkenntnisse nochmals kurz im Plenum.

Sehr übersichtlich wird die Situation auch im Baumdiagramm (siehe Abbildung 1) mit natürlichen Häufigkeiten. Dies zeichne ich an die Tafel. Das Aufdecken der umgeklappten Tafel folgt, ist aber gar nicht mehr entscheidend, das Spiel diente nur der Vermittlung des mathematischen Inhaltes.

Die Einführung des Terms $P_A(C)$ und der Unterscheidung von $P(A)$ geschieht hier auf so naheliegende Weise, das kaum Verwirrung entstehen kann. Natürliche Häufigkeiten sind intuitiv.
Nun spanne ich den Bogen weiter und erläutere, wie die bedingten Wahrscheinlichkeiten verstanden und berechnet werden können, wenn mir die natürlichen Häufigkeiten nicht direkt zugänglich sind.[2]. Dazu erkläre ich zunächst wie die bedingte Wahrscheinlichkeit allgemein definiert werden kann:

$$P_C(A) = \frac{P(A \cap C}{P(C)} \tag{12}$$

Hier wird deutlich, dass bei der bedingten Wahrscheinlichkeit die Wahrscheinlichkeiten der Ereignisse, die A **und** C erfüllen mit der Wahrscheinlichkeit **aller** Ereignisse,

[2]Gerhard Gigerenzer macht sich stark dafür, dass wir stets versuchen sollten, die bedingten Wahrscheinlichkeiten in natürliche Häufigkeiten zu übersetzen, um so die Situation viel besser verstehen zu können. In der Schule bin ich aber in der Zwangslage, meinen Schülern auch die Berechnung ohne natürliche Häufigkeiten vermitteln zu müssen, da dies eine zentrale Anforderung in der Abiturprüfung darstellt

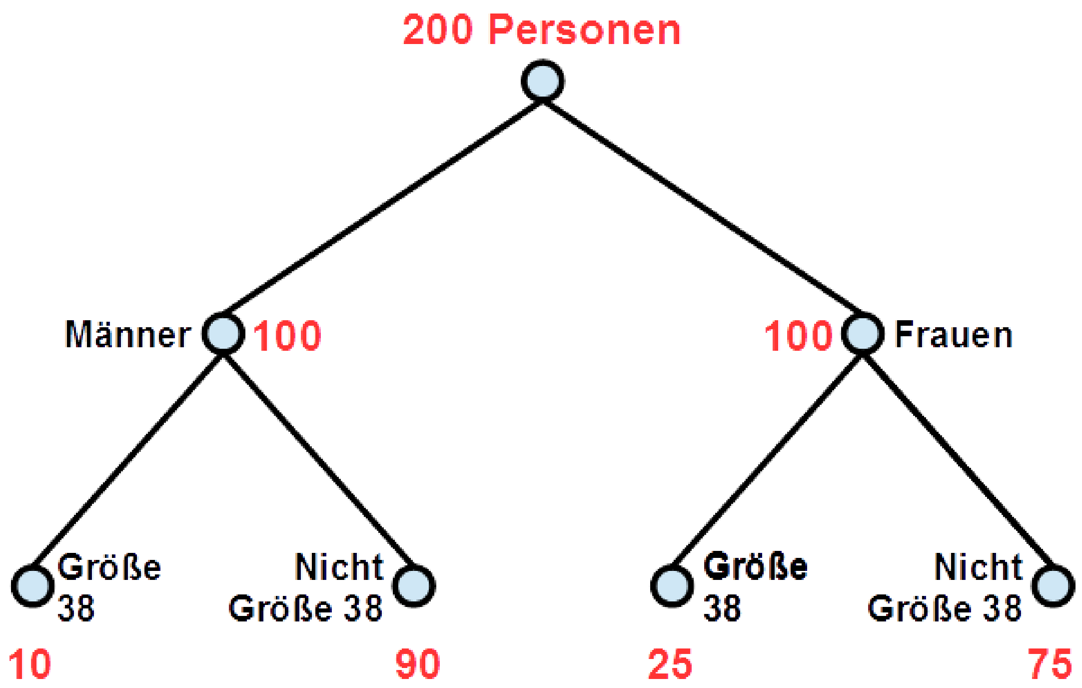

Abbildung 1: Baumdiagramm der natürlichen Häufigkeiten

die zu C gehören verglichen werden. Ganz analog zu der obigen Betrachtung mit den natürlichen Häufigkeiten.
Schließlich zeige ich noch, wie die beiden Wahrscheinlichkeiten $P(A \cap C)$ und $P(C)$ aus einem entsprechenden Baumdiagramm (siehe Abbildung 2)gewonnen werden können. Dieses Baumdiagramm steht neben dem Baumdiagramm der natürlichen Häufigkeiten. Hier offenbart der Vergleich der beiden Baumdiagramme, welche Bedeutung die einzelnen Wahrscheinlichkeiten haben und wo die bedingten Wahrscheinlichkeiten auftauchen.

4 Fazit

Die Einführung der bedingten Wahrscheinlichkeit über die natürliche Häufigkeit ist intuitiv und macht den Unterschied zwischen bedingter und einfacher Wahrscheinlichkeit offensichtlich. Da die Prüfungsaufgaben weiterhin die Berechnung von bedingten Wahrscheinlichkeiten über Formel oder Baumdiagramme fordert, muss in Bezug auf die Berechnung hier auch weiterhin ein Schwerpunkt des Unterrichts sein.
Für die Begriffsbildung sind die natürlichen Häufigkeiten aber eine ungemeine Bereicherung. Wir brauchen mehr Beispiele dieser Art im Unterricht und in den Schulbüchern.

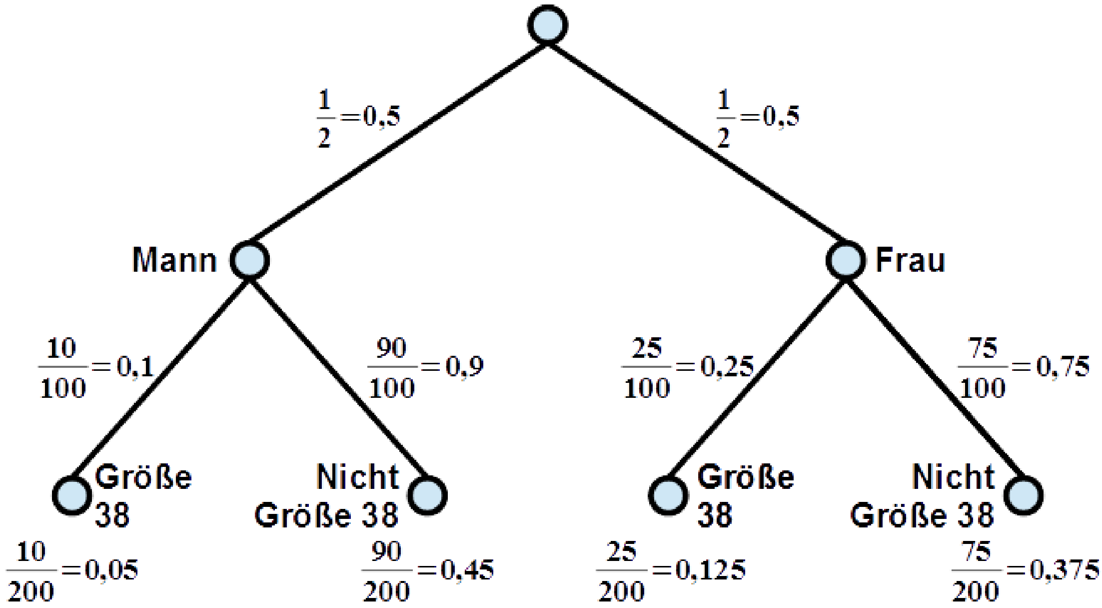

Abbildung 2: Baumdiagramm der Wahrscheinlichkeiten

Literatur

[Gigerenzer2002] “Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken“; Gigerenzer, Gerd; 2002; Berlin Verlag; Berlin

[freqformhyp] “Frequency format hypothesis“; wikipedia.org; `https://en.wikipedia.org/wiki/Frequency_format_hypothesis`; abgerufen am 6. Oktober 2017