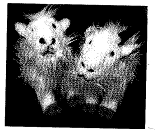
14. Wechseln oder nicht wechseln?

Das Ziegenproblem

Die Wahrscheinlichkeitsrechnung ist reich an Paradoxien, es gibt viele Aussagen, die dem „gesunden Menschenverstand“ widersprechen. So wurde vor einigen Jahren einer breiteren Öffentlichkeit das so genannte Ziegenproblem bekannt.

Zur Erinnerung: Der Quizmaster lässt den in die Endrunde gekommenen Kandidaten eine von drei Türen wählen: Hinter einer ist der Hauptgewinn versteckt, hinter den anderen zwei Türen sind Nieten in Form von Ziegen. Der Kandidat wählt Tür 1, der Quizmaster öffnet Tür 3, hinter der eine Ziege zu sehen ist. Nun kommt die Pointe: Der Kandidat wird gefragt, ob er noch einmal wechseln möchte, d.h. statt Tür 1 nun Tür 2 favorisieren will. Für „Nein“ spricht, dass sich ja die Position des Hauptgewinns durch die bisherigen Aktionen nicht geändert hat, für „Ja“, dass durch das Öffnen von Tür 3 eine neue Situation entstanden ist. Die Frage nach der richtigen Antwort spaltete ganze mathematische Fachbereiche, das Problem fand seinen Weg in die großen Zeitschriften und wurde auch unter Nichtmathematikern intensiv diskutiert. Die Ja-Partei fand die Meinung der Nein-Partei naiv/lächerlich/vorwissenschaftlich, und umgekehrt galt das auch. Die Angelegenheit hatte auch noch einen geschlechterspezifischen Aspekt.



Eine der ersten Verfechter des „Ja“ war nämlich die amerikanische Journalistin Marilyn vos Savant, eine Frau, die durch einen besonders hohen IQ berühmt geworden war. Es gab nicht wenige Stimmen aus dem Mathematikerlager, die ihr rieten, sich nicht in Sachen einzumischen, bei denen sie als Frau sowieso keinen Durchblick haben könnte.

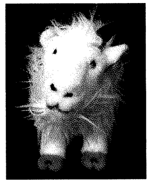
Wer hatte nun Recht? Marylin lag richtig, der Kandidat sollte wechseln, denn seine Gewinnchancen erhöhen sich durch das Wechseln von $1/3$ auf $2/3$. (Die Begründung findet man nachstehend.)

Die Analyse: Warum ist es günstiger zu wechseln?

Die Aussage „Beim Ziegenproblem ist das Wechseln zur anderen Tür günstiger“ ist nur die erste Annäherung an die Wahrheit. Es folgt eine detaillierte Analyse des Problems, um diese Aussage zu verstehen und am Ende die ganze Wahrheit kennen zu lernen. Der Weg zu diesem Ziel ist nicht ganz einfach, denn es geht um ein ziemlich komplexes Phänomen.

Wahrscheinlichkeiten

Zunächst müssen wir uns ein bisschen um wahrscheinlichkeitstheoretische Grundbegriffe kümmern. Glücklicherweise ist eine philosophische Diskussion der Frage „Was ist Wahrscheinlichkeit?“ hier nicht erforderlich.



Wir stellen uns ein Verfahren vor, das Zufallsergebnisse produziert. Man könnte zum Beispiel einen Würfel werfen oder eine Karte aus einem gut gemischten Skatspiel ziehen. Wenn man das sehr oft wiederholt, lassen sich gewisse „Tendenzen“ beobachten: In etwa einem Sechstel der Würfelwürfe wird eine „Vier“ geworfen, in etwa einem Viertel der Fälle wird „Herz“ aus dem Kartenstapel gezogen usw. Man drückt das so aus, dass man sagt, dass beim Würfel die Wahrscheinlichkeit für das Werfen der Vier gleich ein Sechstel ist, und für das Ziehen einer Herz-Karte ist die Wahrscheinlichkeit ein Viertel anzusetzen. Allgemein:

> Die Wahrscheinlichkeit für ein bei einer Zufallswahl mögliches Ergebnis E ist die Zahl p mit der folgenden Eigenschaft:
>
> Wiederholt man die Zufallsauswahl sehr oft, so wird in einem Anteil p der Durchgänge das Ergebnis E erzielt werden. Wenigstens stimmt das ungefähr, und umso besser, je mehr Wiederholungen durchgeführt werden. Man schreibt dann auch $P(E) = p$, gesprochen wird das „P von E gleich p“. Das große P soll dabei an „Wahrscheinlichkeit“ erinnern[14].

In den Beispielen war $P(\text{Vier})=1/6$ und $P(\text{Herz})=1/4$.

Da Anteile immer zwischen 0 und 1 liegen, gilt das auch für Wahrscheinlichkeiten. Auch sind einige einfache Eigenschaften aufgrund des Ansatzes klar. Wenn zum Beispiel ein Ergebnis E stets insbesondere die Bedingung erfüllt, die man an ein Ergebnis F stellt, so sollte die Wahrscheinlichkeit für F eher größer als die von E sein, auf keinen Fall aber kleiner. Zum Beispiel ist die Vier eine gerade Zahl. Deswegen ist es wenig überraschend, dass die Wahrscheinlichkeit, eine gerade Zahl zu würfeln (sie ist gleich 0.5), größer ist als die, eine Vier zu erhalten.

Beim Ziegenproblem spielen viele Wahrscheinlichkeiten eine Rolle. Zum Beispiel wäre es interessant zu wissen, mit welchen Wahrscheinlichkeiten das Auto hinter den verschiedenen Türen platziert wird. Kann man für jede der drei Türen die gleiche Wahrscheinlichkeit (also $1/3$) ansetzen? Oder hat die Tür, die näher am Bühneneingang ist, bessere Chancen, gewählt zu werden? (Immerhin ist so ein Auto gar nicht so leicht zu schieben.)

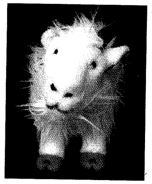
Bedingte Wahrscheinlichkeiten

Als Nächstes geht es um das wichtige Prinzip „Informationen verändern Wahrscheinlichkeiten“. Ein Beispiel: Die Wahrscheinlichkeit, eine Vier zu würfeln, ist ein Sechstel. Wenn man aber nach dem Würfeln (und vor dem Bekanntgeben des Ergebnisses) gesagt bekommt, dass eine gerade Zahl gewürfelt wurde, sieht das anders aus. Nun ist nämlich klar, dass es eine der Zahlen Zwei, Vier oder Sechs sein wird, und damit steigt die Wahrscheinlichkeit der Vier auf ein Drittel. Es hätte auch

[14] Wozu man allerdings die englische oder französische Bezeichnung für Wahrscheinlichkeit kennen muss: *probability* bzw. *probabilité*.

sein können, dass die Information lautet: „Es ist eine ungerade Zahl!". Dann ist es bestimmt keine Vier, die Wahrscheinlichkeit ist also auf Null gesunken.

Kurzum: Es kann alles Mögliche mit Wahrscheinlichkeiten passieren, wenn es Zusatzinformationen gibt. Sie können gleichbleiben, steigen oder fallen.



Das gleiche Phänomen kennen wir aus dem wirklichen Leben[15]. Angenommen etwa, Sie fahren an jedem Arbeitstag die gleiche Strecke. Auf der linken Spur fließt der Verkehr etwas weniger zäh, deswegen würden Sie gern dort fahren. Interessant wäre dann zu wissen, ob der Wagen vor Ihnen an der nächsten Kreuzung links abbiegen wird. (Falls ja, sollten Sie besser die Spur wechseln, denn Sie wollen geradeaus weiter fahren.) Das passiert selten, nur etwa jeder zwanzigste Fahrer will nach links, und deswegen kann man die Wahrscheinlichkeit für „Abbiegen" mit $1/20$ ansetzen. Nun kann es aber doch sein, dass das Nummernschild des Vordermanns (der Vorderfrau) zu einer Stadt gehört, zu der es an genau dieser Kreuzung nach links abgeht. Und deswegen wird die Wahrscheinlichkeit für „Abbiegen" in diesem Fall sicher höher sein.

Es wird sich als günstig erweisen, das Ganze etwas formaler aufzuschreiben. Wenn E ein Ereignis ist, so bezeichnet doch $P(E)$ die Wahrscheinlichkeit für das Eintreten von E. Und wenn F irgendeine Zusatzinformation ist, so schreibt man $P(E \mid F)$ für die neue Wahrscheinlichkeit (also unter Berücksichtigung von F) für E. Man spricht das als „P von E unter F" aus, die Zahl $P(E \mid F)$ wird die *bedingte Wahrscheinlichkeit von E unter F* genannt.

Im einleitenden Beispiel war E das Ergebnis, eine Vier zu Würfeln und F die Information, dass die gewürfelte Zahl gerade ist. Wir haben uns überzeugt, dass in diesem Beispiel $P(E \mid F) = 1/3$ gilt.

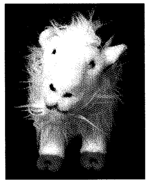
Allgemein geht man so vor. Man bestimmt $P(F)$ (die Wahrscheinlichkeit für F) und $P(E \text{ und } F)$ (die Wahrscheinlichkeit, dass E und F gleichzeitig auftreten). Anschließend wird $P(E \mid F)$ wie folgt definiert:

$$P(E \mid F) := \frac{P(E \text{ und } F)}{P(F)}.$$

Testen wir das an unserem ersten Beispiel. $P(F) = 1/2$, denn die Hälfte der möglichen Ergebnisse beim Würfelwurf besteht aus geraden Zahlen. „E und F" entspricht dem Ergebnis, dass eine Zahl gewürfelt wurde, die gleichzeitig mit der Vier übereinstimmt und gerade ist. Das leistet als Einzige die Zahl Vier, und deswegen ist $P(E \text{ und } F) = 1/6$. Wir erhalten so wirklich

$$P(E \mid F) = \frac{P(E \text{ und } F)}{P(F)} = \frac{1/6}{1/2} = \frac{1}{3}.$$

[15] Ich wage sogar die Behauptung, dass das richtige Anpassen von Wahrscheinlichkeiten aufgrund neuer Informationen in der stammesgeschichtlichen Entwicklung des Menschen eine ganz wichtige Rolle spielte und inzwischen im Gehirn fest „verdrahtet" ist.



Als weiteres Beispiel betrachten wir ein gewöhnliches Skatspiel. Die Wahrscheinlichkeit für E =„Pik As" ist gleich $1/32$, denn diese Karte gibt es nur einmal. Verrät man uns aber, dass eine schwarze Karte gezogen wurde, so steigt die Wahrscheinlichkeit für „Pik As" auf $1/16$: Mit F =„schwarze Karte" ist doch $P(F) = 1/2$ (die Hälfte der Karten ist schwarz) und $P(E \text{ und } F) = 1/32$ (Pik As ist die einzige Karte, für die „E und F" gilt); so folgt

$$P(E \mid F) = \frac{P(E \text{ und } F)}{P(F)} = \frac{1/32}{1/2} = \frac{1}{16}.$$

Die Bayes-Formel

Bemerkenswerterweise lassen sich bedingte Wahrscheinlichkeiten quasi umkehren, dafür verwendet man die Bayes-Formel. Die Fragestellung ist aus dem Leben bekannt:

> Angenommen, Sie stellen nach dem Besuch Ihrer Freunde fest, dass Ihre Lieblings-DVD verschwunden ist. Sie wissen, dass Ihre Freunde eine unterschiedliche Tendenz haben, Sachen ohne zu fragen einfach „auszuborgen". Wen sollten Sie nun verdächtigen?

Für unsere Zwecke ist es am günstigsten, die Bayes-Formel wie folgt zu formulieren. Wir betrachten ein Zufallsexperiment, bei dem sich jedes mögliche Ergebnis in eine von drei vorher festgelegten Klassen einordnen lässt. Wir nennen sie B_1, B_2, B_3. Wichtig ist, dass sich die Klassen nicht überlappen.

> Ein Beispiel für den Würfelwurf: Man könnte
>
> B_1 : „Es wurde eine 1 oder eine 2 gewürfelt."
>
> B_2 : „Es wurde eine 3 oder eine 4 gewürfelt."
>
> B_3 : „Es wurde eine 5 oder eine 6 gewürfelt."
>
> definieren.

Nun geht es um irgendein bei diesem Experiment mögliches Ergebnis, wir nennen es A. (Etwa: „Es wurde eine Primzahl gewürfelt.") Sind dann die bedingten Wahrscheinlichkeiten $P(A \mid B_1), P(A \mid B_2), P(A \mid B_3)$ und die Wahrscheinlichkeiten $P(B_1), P(B_2), P(B_3)$ bekannt, so kann man die „umgekehrten" bedingten Wahrscheinlichkeiten, also die $P(B_1 \mid A), P(B_2 \mid A), P(B_3 \mid A)$ berechnen:

$$P(B_1 \mid A) = \frac{P(A \mid B_1)P(B_1)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3)},$$

$$P(B_2 \mid A) = \frac{P(A \mid B_2)P(B_2)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3)},$$

$$P(B_3 \mid A) = \frac{P(A \mid B_3)P(B_3)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3)}.$$

Diese Formeln heißen die *Bayes-Formeln* [16].

Beweisen wollen wir die Formeln hier nicht, zur Illustration gibt es aber ein Beispiel: Die B_1, B_2, B_3 sollen wie vorstehend sein (B_1=„1 oder 2 gewürfelt", usw.). Als A betrachten wir „Die gewürfelte Zahl ist größer als 3". Mit den im vorigen Abschnitt beschriebenen Rechnungen erhält man dann

$$\begin{aligned} P(A \mid B_1) &= 0, \\ P(A \mid B_2) &= \frac{1}{2}, \\ P(A \mid B_3) &= 1, \end{aligned}$$

auch gilt offensichtlich $P(B_1) = P(B_2) = P(B_3) = 1/3$.

Nun wird gewürfelt, und das Ergebnis liegt in A (die gewürfelte Zahl ist also größer als 3). Mit welcher Wahrscheinlichkeit liegt es – zum Beispiel – in B_2? Dazu verwenden wir die Bayes-Formel:

$$\begin{aligned} P(B_2 \mid A) &= \frac{P(A \mid B_2)P(B_2)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3)} \\ &= \frac{(1/2) \cdot (1/3)}{0 \cdot (1/3) + (1/2) \cdot (1/3) + 1 \cdot (1/3)} \\ &= \frac{1}{3}. \end{aligned}$$

Ganz analog hätte man mit der Bayes-Formel ausrechnen können [17], dass $P(B_1 \mid A) = 0$ und $P(B_3 \mid A) = 2/3$ gilt.

Die beste Strategie beim Ziegenproblem: Standardvariante

Nach diesen Vorbereitungen können wir entscheiden, ob Wechseln günstiger ist. Wir diskutieren zunächst die Wahrscheinlichkeiten, die mit dem Verstecken des Autos zusammenhängen. Die folgenden Abkürzungen werden verwendet:

B_1 : Das Auto wird hinter Tür 1 versteckt.
B_2 : Das Auto wird hinter Tür 2 versteckt.
B_3 : Das Auto wird hinter Tür 3 versteckt.

[16] Hat man nicht in drei Klassen B_1, B_2, B_3, sondern allgemeiner in n Klassen $B_1, B_2, \ldots, B_n$ aufgeteilt, so lauten die Bayes-Formeln

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A \mid B_1)P(B_1) + \cdots + P(A \mid B_n)P(B_n)},$$

wobei man für die Zahl i die Zahlen $i = 1, 2, \ldots, n$ einsetzen kann.

[17] In diesem Fall ist auch eine direkte Rechnung leicht möglich; sie führt natürlich zum gleichen Ergebnis.

Wir wollen uns auf den Standpunkt stellen, dass diese drei Möglichkeiten die gleiche Wahrscheinlichkeit haben, dass also

$$P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$$

gilt. Das ist vielleicht etwas naiv, aber solange nichts Gegenteiliges bekannt ist, kann man ja einmal von dieser Annahme ausgehen.

Nun kommt der spannende Moment, wo eine Entscheidung zu treffen ist: Der Spieler hat Tür 1 gewählt, der Quizmaster zeigt die Ziege hinter Tür 3, und es ist nicht klar, ob man von Tür 1 zu Tür 2 wechseln sollte. Hier die Analyse.

Das Ereignis „Der Quizmaster zeigt die Ziege hinter Tür 3" soll mit A abgekürzt werden. Unter Verwendung dieser Information interessiert uns dann die Antwort auf die Frage: Ist das Auto eher hinter Tür 1 oder hinter Tür 2? Verwenden wir die vorstehend eingeführten Begriffe, so wollen wir wissen, wie sich die Zahlen $P(B_1 \mid A)$ und $P(B_2 \mid A)$ zueinander verhalten: Sind sie beide gleich (dann lohnt Wechseln nicht), oder ist die zweite die größere (dann ist Wechseln sinnvoll)?

Das ist ein typischer Fall für die Anwendung der Bayes-Formel. Um sie anwenden zu können, benötigen wir noch die Zahlen $P(A \mid B_1), P(A \mid B_2), P(A \mid B_3)$.

Wie groß ist $P(A \mid B_1)$? In Worten: Wenn das Auto hinter Tür 1 steht, wie groß ist dann die Wahrscheinlichkeit, dass der Quizmaster Tür 3 öffnet? Er könnte natürlich auch Tür 2 öffnen (oder gleich die Lösung verraten). Wir wollen hier annehmen, dass er sich mit gleicher Wahrscheinlichkeit für das Öffnen von Tür 2 oder Tür 3 entscheidet, und deswegen setzen wir $P(A \mid B_1) = 1/2$.

$P(A \mid B_2)$ ist leichter zu ermitteln: Das Auto steht hinter Tür 2, wie groß ist dann die Wahrscheinlichkeit, dass Tür 3 geöffnet wird? Sie ist sicher gleich Eins, denn der Quizmaster kann ja nicht Tür 2 öffnen (da steht das Auto), und Tür 1 ist auch tabu, denn die wurde vom Spieler gewählt.

Ähnlich leicht ist $P(A \mid B_3)$ zu behandeln: Da kommt natürlich Null heraus, denn es wird sicher nicht die Tür mit dem Auto geöffnet. Wir fassen zusammen:

$$P(A \mid B_1) = \frac{1}{2}, \; P(A \mid B_2) = 1, \; P(A \mid B_3) = 0.$$

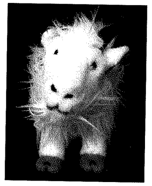
Und nun kann die Bayes-Formel angewendet werden:

$$\begin{aligned} P(B_1 \mid A) &= \frac{P(A \mid B_1)P(B_1)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3)} \\ &= \frac{(1/2) \cdot (1/3)}{(1/2) \cdot (1/3) + 1 \cdot (1/3) + 0 \cdot (1/3)} \\ &= \frac{1}{3}, \end{aligned}$$

$$
\begin{aligned}
P(B_2 \mid A) &= \frac{P(A \mid B_2)P(B_2)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3)} \\
&= \frac{1 \cdot (1/3)}{(1/2)\cdot(1/3) + 1\cdot(1/3) + 0\cdot(1/3)} \\
&= \frac{2}{3}.
\end{aligned}
$$

Fazit: Da $P(B_1 \mid A)$ (bzw. $P(B_2 \mid A)$) die Wahrscheinlichkeit ist, mit der Strategie „Nicht wechseln!“ (bzw. „Wechseln!“) das Auto zu gewinnen, so heißt das, dass Wechseln die Chancen verdoppelt.

Das Ziegenproblem: Die ganze Wahrheit



Wer die vorstehende Analyse aufmerksam verfolgt hat, wird bemerkt haben, dass für der Aussage „Wechseln ist besser, die Gewinnwahrscheinlichkeit wird verdoppelt!“ einige Annahmen erforderlich waren. Wir haben zum Beispiel $P(A \mid B_1) = 1/2$ gesetzt. Das ist eigentlich nicht zwingend: Vielleicht öffnet der Quizmaster immer Tür 3, wenn es möglich ist (wenn also nicht gerade das Auto dahinter steht). Um das allgemein zu untersuchen, setzen wir $P(A \mid B_1) = p$, wobei p eine Zahl zwischen 0 und 1 ist. Dann ergibt unsere Analyse:

$$P(B_1 \mid A) = \frac{p}{1+p}, \; P(B_2 \mid A) = \frac{1}{1+p}.$$

Es ist dann zwar immer noch so, dass Wechseln günstiger ist (denn die erste der beiden Zahlen ist nie die größere), der Unterschied kann aber wesentlich unspektakulärer sein.

Es ist auch eine *andere Herangehensweise* möglich[18]. Dabei verfolgt der Spieler eine andere Strategie, er kümmert sich nämlich überhaupt nicht um die Vorlieben des Quizmasters und *wechselt auf jeden Fall.* Man kann dann so argumentieren:

- Bei der ursprünglichen Wahl (und damit auch dann, wenn man bei der Entscheidung bleibt) wird das Auto mit $1/3$ Wahrscheinlichkeit gewonnen, denn beim Verstecken hatten alle Türen die gleiche Wahrscheinlichkeit.
- Beim Wechseln wird genau dann gewonnen, wenn die erste Wahl falsch war, also mit Wahrscheinlichkeit $2/3$.

Das kann man noch ein bisschen verfeinern. Bezeichnet man nämlich mit p_1 bzw. p_2 bzw. p_3 die Wahrscheinlichkeit, dass das Auto hinter Tür 1 bzw. Tür 2 bzw. Tür 3 versteckt wurde, so ist – bei Wahl von Tür 1 – die Wahrscheinlichkeit für „Auto gewinnen, vorher nicht wechseln“ gleich p_1 und die Wahrscheinlichkeit für „Auto gewinnen, vorher wechseln“ gleich $p_2 + p_3$.

[18] Ich danke Herrn Professor Dieter Puppe aus Heidelberg für diesen Hinweis.

Es ist nicht auszuschließen, dass manche Leser nun verwirrt sind, denn in der zweiten Analyse spielt das Verhalten des Quizmasters scheinbar keine Rolle. Man muss schon genau hinsehen, um einzusehen, dass beide Vorgehensweisen korrekt waren.

In der *ersten Analyse* war die Ausgangssituation so gegeben: Tür 1 gewählt, Tür 3 (mit Ziege) geöffnet. Und aufgrund dieser Situation sollte man die Wahrscheinlichkeiten ausrechnen.

Bei der *zweiten Analyse* war die Situation anders: Das Verhalten des Quizmasters war völlig unerheblich, gewechselt wird in jedem Fall. Trotzdem ist nur schwer intuitiv einzusehen, dass diese unterschiedliche Information für unterschiedliche Wahrscheinlichkeiten verantwortlich ist.

P.S.: Alle, die es noch genauer wissen wollen, sollten sich das sehr lesenswerte Buch „Das Ziegenproblem“ von Gero von Randow besorgen (rororo Science, 7.50 Euro).