Abwesenheit von Evidenz ist nicht Evidenz für Abwesenheit.

Martin Rees, Astronomer Royal

Liebe Leserin, lieber Leser,

unter der Rubrik „Mathematische Unterhaltungen“ war in einer bekannten populärwissenschaftlichen Zeitschrift Familie Schmidt abgebildet. Neben Mama und Papa ist deren Tochter zu sehen. Ein zweites Kind spielt mit Schmidts Hund. Ob Mädchen oder Junge ist nicht auszumachen. Trotzdem ließ es sich Ian Stewart, der Autor, nicht nehmen, nach der Wahrscheinlichkeit zu fragen, dass es sich ebenfalls um eine Tochter handelt. Nach Auflisten aller Möglichkeiten kommt man auf ein Drittel. Doch ich hatte die Wichtigkeit des Hundes verkannt. Lautet die Botschaft nämlich, das Kind, das nicht mit dem Hund spielt, ist ein Mädchen, erhöht sich die Wahrscheinlichkeit auf ein Halb. (Der Hinweis, die Tochter sei das *ältere* Kind, hätte das Gleiche bewirkt.) Der Anlass, diese Denksportaufgabe zu erwähnen, ist ein denkwürdiges Ereignis der Wissenschaftsgeschichte. Vor 250 Jahren, das Jahr 1763 neigte sich dem Ende zu, stellte auf einer Sitzung der Royal Society in London ein Prediger, Versicherungsmathematiker und Mitglied dieser honorigen Gesellschaft die Schrift eines kürzlich verstorbenen Presbyterianischen Geistlichen und Akademiekollegen vor. Dieser, Thomas Bayes (1701?–1761), dilettierte in Wahrscheinlichkeitstheorie und war auf einen Satz über bedingte Wahrscheinlichkeiten gestoßen. Der Bedeutung war sich damals niemand bewusst, auch der gelehrte Nachlassverwalter nicht. Und dabei ist für die Wahrscheinlichkeitsrechnung der Bayes’sche Satz, folgt man dem Mathematiker, Geophysiker und Astronomen Sir Harold Jeffreys (1891–1989), so etwas wie der Pythagoras in der Geometrie. Bekannt gemacht hat den Bayes’schen Satz kein geringerer als der französische Mathematiker und Himmelsmechaniker Pierre-Simon Laplace (1749–1827). Das war um 1812. Doch das Blatt wendete sich. Der Bayes’sche Satz sei zwar richtig, seine Anwendung auf Aussagen aber verwerflich[1]. Damit kratze man am Objektivitätsanspruch der

[1] Wer den Wahrscheinlichkeitsbegriff ohne Not auf Häufigkeit beschränkt, auf die Häufigkeit, mit der beispielsweise eine Sechs gewürfelt wird, wird sich weigern, einer wis-

Wissenschaft! Hinzu kam, dass sich der Bayes'sche Zugang zur Statistik als rechenaufwändig erwies. Heute ist das Adjektiv „Bayesianisch“ aus wissenschaftlichen Schriften nicht mehr wegzudenken, meist taucht es bereits in der Überschrift auf, weil sich der Autor ein Mehr an Aufmerksamkeit davon verspricht. Ursache dieser Renaissance ist der Computer! Er erlaubt das Spielen mit den Möglichkeiten im großen Stil.

Doch auch außerhalb der Wissenschaft ist Bayesianisches Schlussfolgern nützlich: bei der Wahrheitsfindung (Gericht!) und als Entscheidungshilfe (Klimapolitik!). Wir Menschen sind nämlich keine geborenen Bayesianer – vermutlich ist richtiges Denken biologisch zu teuer – und neigen zu Fehlurteilen[2].

Die langen Abende der Advents- und Weihnachtszeit laden dazu ein, sich auf die Spielwiese der „Bayesianer“ zu begeben und über Wege zur Erkenntniss zu spekulieren. Hals- und Beinbruch wünscht

Hans-Erich Fröhlich

Der Himmel im Dezember

Obwohl sich die Venus von uns aus gesehen der Sonne am Himmel wieder nähert, verbessert sich in der ersten Monatshälfte noch ihre Sichtbarkeit am Abendhimmel. Sie verschwindet gegen 19 Uhr am Westhorizont. Im hellsten Glanz steht sie am 10. Dezember. Sie erreicht dann -4,7te Größe.

Jupiter ist bereits fast die ganze Nacht über sichtbar – was kein Wunder ist, am 5. Januar steht er in Opposition zur Sonne –, Mars und Saturn gehören in die zweite Nachthälfte.

Am 15. Dezember durchquert der Mond die Hyaden. So etwas verspricht immer einen reizvoller Anblick.

Astronomisch gesehen wird es bei uns Winter, erreicht die Sonne mittags ihren niedrigsten Stand. Jahrestiefpunkt ist der 21. Dezember 18 Uhr 11.

Was macht eigentlich ISON, der „Weihnachtskomet“? Anscheinend hat er den Sonnenvorbeisaus schlecht vertragen. Falls er doch noch reanimiert werden

---

senschaftlichen Aussage eine Wahrscheinlichkeit zuzuordnen, wahr zu sein. Der hat aber auch Probleme mit der Wettervorhersage. Was mag eine Regenwahrscheinlichkeit von 30 % bedeuten?

[2]Unsere Schwäche, *intuitiv* richtig zu denken, tritt beim sog. „Ziegenproblem“ zutage. Es geht um die Frage: Lohnt es sich bei einem Ratespiel, eine bereits gefällte Entscheidung nach Erhalt eines Tipps zu revidieren? Bayes' Theorem gibt darauf die Antwort.

sollte, wird er zunächst in den Morgenstunden zu finden sein, nach Weihnachten dann die ganze Nacht über. Der Komet steigt steil auf gen Norden, (westlich) vorbei am Kugelsternhaufen M 13 im Herkules (22./23.) auf den Polarstern zu, an dem er im neuen Jahr, Anfang Januar, vorbeiziehen wird.

Hochwürden mathematisiert den gesunden Menschenverstand

Wie gelangt man zu einer Einschätzung der Glaubwürdigkeit wissenschaftlicher Behauptungen? Auch wenn es nicht jeder so sieht, der Weg zum begründeten Urteil führt über Bayes' Theorem.

Wurde etwas beobachtet oder gemessen, so kann man sich unschwer ausrechnen, wie wahrscheinlich in Anbetracht von Messfehlern die Daten (D) sind, gesetzt man hat eine Vorstellung, wie sie zustande gekommen sein könnten, sprich ein physikalisches Modell. Statt Modell kann man auch von einer Arbeitshypothese (H) sprechen. Wir kürzen diese *bedingte* Datenwahrscheinlichkeit mit $P(\mathrm{D}|\mathrm{H})$ ab. Bedingt ist sie, weil die Hypothese H dabei als wahr vorausgesetzt ist. (Hinterm Strich steht die Voraussetzung.) Die Crux: Es gibt immer Alternativhypothesen, die die Daten ebenfalls beschreiben können. Um nun von $P(\mathrm{D}|\mathrm{H})$ auf die Wahrscheinlichkeit $P(\mathrm{H}|\mathrm{D})$ unserer Lieblingshypothese schließen zu können, bedarf es der Anwendung des Bayes'schen Theorems, was rechnerisch aufwändig ist. Weshalb wohl betreibt die NSA riesige Rechenzentren? Um Telefongespräche bloß aufzuzeichnen braucht man die nicht.

Nicht nur Wissenschaft und Geheimdienst bedürfen des Hypothesenrankings. Da das Bayes'sche Theorem den „gesunden Menschenverstand" in die Sprache der Mathematik übersetzt, sollte man ihm öfters begegnen.

Richter sind sich des Unterschieds zwischen $P(\mathrm{D}|\mathrm{H})$ und $P(\mathrm{H}|\mathrm{D})$ bewusst. Die Indizienwahrscheinlichkeit $P(\mathrm{I}|\mathrm{S})$ ist nicht die Schuldwahrscheinlichkeit $P(\mathrm{S}|\mathrm{I})$! Wir haben hier Hypothese durch Schuld (S) und Daten durch Indizien (I) ersetzt. Ein Beispiel verdeutliche das:

Am Tatort inmitten der Großstadt werden DNS-Spuren gefunden. Einen Verdächtigen, auf den das DNS-Profil passt, gibt es auch. Dennoch ist zu bedenken, dass in einer Millionenstadt einige Bürger mit dem gleichen DNS-Profil herumlaufen. Hat der Tatverdächtige biologische Verwandte dort, sind es sogar etliche.

Entwarnung ist insbesondere dann angesagt, ist die sog. Prior- oder Basiswahrscheinlichkeit eines unerwünschen Ereignisses gering. Wird bei jemandem eine *seltene* Krankheit diagnostiziert, so heißt das zunächst nur, dass die Wahrscheinlichkeit, erkrankt zu sein, deutlich höher ist als im Bevölkerungsdurchschnitt, sagen wir verzehnfacht. Aber das Zehnfache einer kleinen Zahl ist immer noch eine kleine Zahl! Man sollte sich in so einem Fall nach der Häufigkeit von Fehldiagnosen erkundigen. (Ein Beduine in der Sahara wird auch nicht gleich nach dem Regenschirm greifen, bloß weil ihm seine Wetterstation wieder einmal Regen prophezeit!)

In allen solchen Fällen von Umkehrschlüssen, $P(\mathrm{D}|\mathrm{H}) \Rightarrow P(\mathrm{H}|\mathrm{D})$, ist der Satz von Bayes das einzige Hilfsmittel, das wir haben.

Top-down **oder** ***Bottom-up*****?**

Aus noch so vielen Fakten kann nie aufs allgemeine Gesetz geschlossen werden. Der Schluss vom Speziellen aufs Allgemeine, die *Induktion*, ist logisch nicht vertretbar[3]. Zwingend ist nur die *Deduktion*. Die Bibel hat, zumindest was die Erkenntnislogik anbelangt, recht: Am Anfang ist die Idee, die Intuition. Wissenschaft beginnt im Kopf mit der makellosen Vorstellung, die wir von den Dingen haben, dem *Vor*-Urteil.

Der Pfad zur Erkenntnis ist keine Einbahnstraße: Auf dem fragwürdigen Weg der Induktion (*bottom-up*) wird die Hypothese erraten und auf dem Wege der Deduktion (*top-down*) begutachtet. Man darf sich natürlich nichts vormachen und die Hypothese ausgerechnet an dem Datenmaterial testen wollen, das einem zur Hypothese verholfen hat, wie es Pseudowissenschaften tun.

Sind wir offen, wird die Innenschau belästigt: Sinnes- oder Messdaten aus der realen Welt strömen auf uns ein. Wir filtern und interpretieren sie im Rahmen von bewährten Vorstellungen. Vertragen sich die Eindrücke mit den mentalen Erwartungen, bestärkt dies das Vorurteil. Man fühlt sich wohl – und hat nicht an Wissen zugenommen. Nur wenn die Daten sich quer stellen und es geistiger Verrenkung bedarf, sie in Einklang zu bringen mit dem, was wir zu wissen meinen, lehrt uns das etwas: Wir lagen falsch! Es ist dann an der Zeit, sich von diesem Vorurteil zu verabschieden und ein neues auszuprobieren.

[3]Sie lesen die Zahlenfolge 1, 2, 4, 8, 16 und sollen die folgende Zahl erraten. Vorsicht! Ein denkbar einfaches geometrisches Modell generiert die Folge 1, 2, 4, 8, 16, 31, . . .

Mit zunehmender Messgenauigkeit wurde die Ptolemäische Weltsicht unglaubwürdiger. Immer mehr Epizykel waren vonnöten, den Abstand zwischen Vorstellung und Wirklichkeit zu minimieren. Erst Kepler hatte nach langem Herumprobieren *den* Einfall: *eine* Ellipse statt vieler verschachtelter Kreise.

Es ist wie bei der Evolution. Auch die Gene machen über viele Generationen ihre „Erfahrungen" mit der Außenwelt des Organismus allein durch Versuch und Irrtum[4]. Die Außenwelt kann die Erbinformation nicht im Lamarckschen Sinne zielgerichtet umschreiben. (Sie hat höchstens Einfluss darauf, welche Informationen von der DNS transkribiert werden.)
Darwin (1809–1882) beantwortet die Kantsche Frage nach dem *a priori*: Dass der Raum dreidimensional ist, diese Vor-Erfahrung ist seit Dutzenden von Jahrmillionen in den Lebewesen fest verdratet. Spätestens der Urvogel hatte sie. Einen Irrtum bei der Dimensionalität des Raumes hätte er nicht überlebt! Der Dichtermediziner Gottfried Benn hat das sehr schön in Worte gefasst: „In unserem Gehirn liegt die Vorwelt gesammelt [...]".

Man könnte sagen, die Wissenschaft erforscht nicht das „Ding an sich". Sie spielt „nur" gedanklich (heutzutage mit Computer) mit dem „Ding an sich" und bewertet, *à la* Bayes, anhand der Daten die Glaubwürdigkeit von Modellen, sprich Hypothesen.

Dies geschieht in Schritten. Zunächst wird für jede Einzelhypothese aus einem Satz von einander ausschließenden Hypothesen $\{\mathrm{H}_i\}$ die jeweilige Wahrscheinlichkeit der Daten $P(\mathrm{D}|\mathrm{H}_i)$ berechnet, was ein Leichtes ist. Diese wird mit der Basiswahrscheinlichkeit $P(\mathrm{H}_i)$ der Hypothese gewichtet. Um die Hypothesenwahrscheinlichkeit $P(\mathrm{H}_i|\mathrm{D})$ zu erhalten ist noch durch einen Normierungsfaktor zu dividieren, der dafür sorgt, dass die Summe aller Hypothesenwahrscheinlichkeiten am Ende Eins ergibt. Die Wahrheit muss sich ja irgendwie auf die Hypothesen verteilen, falls man alle überhaupt denkbaren aufgelistet hat. Dieser letzte Schritt ist selbst für Hochleistungscomputer eine Zumutung[5].

Die paar Zahlen! Kann das so schwer sein?

Es geht eben nicht nur um ein paar Zahlen. Hypothesen enthalten fast immer irgendwelche freien Parameter, die man ebenfalls aus den Daten bestimmen möchte. So eine Hypothese splittet sofort in viele Unterhypothesen auf.

[4]Der Irrtum kann tödlich sein, er wird mit dem Aussterben bestraft. Der Angepasste ist der Glückliche, der bei diesem Spiel bisher überlebte.

[5]Geht es preiswerter? Ja, für „arme" Wissenschaftler gibt's sog. Bayes'sche Informationskriterien, die als kostengünstiger Ersatz herhalten können. Befriedigend ist das nicht.

Bilden die Parameterwerte ein Kontinuum, d. h. liegen sie dicht an dicht, hat man es gleich mit unendlich vielen Unterhypothesen zu tun. Das läuft mathematisch auf eine Integration hinaus, was in hochdimensionalen Parameterräumen rechenintensiv ist. (Der Kosmo-Bote hätte keine Probleme damit, die NSA-Computer mit dem Testen von Hypothesen auszulasten! Der Rechenpowerhunger der Wissenschaft ist gewaltig, sofern sie sich nicht mit „Indizienbeweisen" begnügt.)

Sie haben die irregulär erscheinende Lichtschwankung[6] eines Sterns verfolgt. Der Stern pulsiert, und Sie vermuten die Überlagerung mehrerer Sinusschwingungen. Also gibt's mehrere Hypothesen zu vergleichen: keine Schwingung (H_0), eine Schwingung (H_1), zwei Schwingungen (H_2) usf. Bei der Hypothese H_0 wird angenommen, dass der Stern konstante Helligkeit hat und die Schwankungen Messfehler sind. H_1 wird durch vier freie Parameter spezifiziert, H_2 bereits durch sieben.

Hypothesen mit vielen freien Parametern haben bei Bayesianern schlechte Karten, es sei den die Daten sind wirklich exzellent. Man wird für jeden freien Parameter bestraft, der nicht wesentlich ist! Die Bayes'sche Methode befolgt das Sparsamkeitsgebot des mittelalterlichen Gelehrten William von Ockham (um 1288–1347): Bei mehreren möglichen Erklärungen entscheide man sich für diejenige, welche mit den wenigsten Annahmen auskommt!

Die sparsamste Formulierung von „Ockhams Rasiermesser" gab Albert Einstein: „Man muss die Dinge so einfach wie möglich machen. Aber nicht einfacher."

Erst dann, nach umfänglichen Recherchen, lässt man aufgrund der Evidenzen $P(\mathrm{H}_i|\mathrm{D})$ die erfolglosen Hypothesen sterben.

Da Wissenschaftler Menschen sind und – bezüglich Vorurteil – kaum über ihren Schatten springen können, ist der Ideenwettbewerb zwischen den Forscherhirnen wichtig, die Kontroverse. Allein sachliche Kritik garantiert, dass am Ende die beste Hypothese überlebt, also die – nach allem was man so weiß

[6] Bei der Periodensuche zeigt sich die Leistungskraft Bayes'scher Verfahren. Eine kurze, stark verrauschte Zeitreihe enthalte ein Signal, bestehend aus *zwei* Frequenzen, die dicht beieinander liegen und einander überlagern. Ein Auswerteprogramm „von der Stange" verkündet nach nur einer Sekunde Rechenzeit seine Lösung: *eine* Schwingung schlecht definierter Frequenz. Die Bayes'sche Periodensuche findet nach Stunden *beide* Schwingungen! Mit Beispielen zu astronomischen Anwendungen (und Fehlanwendungen) ließen sich zig Newsletter füllen. Astronomen hatten schon immer ein Faible für Bayes. Sie waren die ersten. Gute Daten zu beschaffen ist teuer. Man sollte also bei der Interpretation nicht knausrig sein.

– wahrscheinlichste. Durch Eliminierung des Falschen nähert man sich der Wahrheit. Erkenntnis ist trotz angebrachter Skepsis machbar, zumindest in den Naturwissenschaften! Sir Karl Popper (1902–1994) hat recht: Das Falsche kann man als falsch erkennen. Das genügt! Wahrheitsbeweise gibt es nicht.

Stein des Anstoßes: das Vor-Urteil

Bayes' Formel beschreibt, wie sich ein Anfangsverdacht durch Messdaten zu Gewissheit verdichtet oder in Rauch auflöst. Das Problem: Am Anfang ist der Glaube, der Verdacht. Wer sich dem hehren Objektivitätsideal der Wissenschaft verpflichtet fühlt, mag das anstößig finden. Auch hier bietet Popper die Lösung. Die Objektivität liege in der kritischen Methode. Was der einzelne Wissenschaftler so glaubt, ist am Ende belanglos. Das „Menschlich-Allzumenschliche" mittelt sich heraus.

Sind die Daten aussagekräftig, ist die Subjektivität zu Anfang von minderer Bedeutung. Die Wahrheit setzt sich durch, sofern die wahre Hypothese im Satz der betrachteten Hypothesen enthalten ist. Die „vergessene" Hypothese könnte die wahre gewesen sein! Der Physiker Richard Feynman (1918–1988) geht in seiner Generosität so weit, dass er selbst Pseudowissenschaften wie der Astrologie eine Basischance von 1 : 1 Million zubilligt. Mehr würde das Selbstbewusstsein des Physikers untergraben. Ist irgendetwas dran an der Pseudowissenschaft, wird sie sich trotz der schlechten Ausgangsposition in der Welt der Wissenschaft durchsetzen.

Beunruhigender ist ein anderer Gedanke: Wir könnten phantasiebegrenzt sein und die wahre Hypothese nie erraten.

Die Formel

Hier ist sie, die Formel, die aus der Daten- bzw. Indizien-Wahrscheinlichkeit Hypothesen- bzw. Schuldwahrscheinlichkeiten zu berechnen gestattet!

$$P(\mathrm{H}_i|\mathrm{D}) = P(\mathrm{D}|\mathrm{H}_i) \cdot P(\mathrm{H}_i)/P(\mathrm{D})$$

$P(\mathrm{H}_i)$ ist die Prior-Wahrscheinlichkeit der Hypothese H_i, also *bevor* die Daten eintreffen, $P(\mathrm{D})$ die über alle Hypothesen gemittelte Wahrscheinlichkeit der Daten. Die Formel beschreibt, wie durch Daten aus der Dingwelt die Einschätzung von Hypothesen in der Geistwelt aktualisiert wird: Der „geupdatete“ Kenntnisstand ist die Posterior-Wahrscheinlichkeit $P(\mathrm{H}_i|\mathrm{D})$. Es gibt Gewinner und Verlierer. Die eine oder andere Hypothese wird an Glaubwürdigkeit zunehmen, andere werden von den Daten abgestraft. Treffen neue Daten ein, wird der Posterior von neulich zum Prior von heute und unter erneuter Anwendung der Formel ein nochmals aktualisierter Posterior berechnet. So einfach ist das.

Der geniale Thomas Bayes dürfte sich über seinen triumphalen Empfang im Himmel gewundert haben. Es war nicht seine Profession, wofür ihm die Engel huldigten, es war seine Passion! Auf Erden ist sein Name kaum bekannt.

Um auf Familie Schmidt zurückzukommen, die Hypothese H_1 lautet *zwei Töchter*. Der der Schmidtschen Familienverhältnisse Unkundige weiß von *zumindest einer Tochter*. Es gibt vier gleichwahrscheinliche Fälle (M = Mädchen, B = Bube): MM, MB, BM, BB.
Die MM-Wahrscheinlichkeit[7] ergibt sich zu $(1 \cdot 1/4)/(3/4) = 1/3$.
Die Alternativhypothese H_2, *das andere Kind ist ein Bube*, ist doppelt so wahrscheinlich: $(1 \cdot 1/2)/(3/4) = 2/3$.

Dem Kosmos-Boten ist bewusst, dass das Mädchen-Junge-Paradoxon jede Menge Zündstoff birgt. Es ist, wie er zu spät bemerkt hat, vertrackter als das Ziegenproblem! Wenn Sie also – auch ohne den Hund als Spielkameraden – gute Gründe[8] für Wahrscheinlichkeit 1/2 finden, ist das in Ordnung! Nur bitte deshalb keine Zuschriften oder gar Beschimpfungen! (Die soll es gegeben haben, nachdem der Mathematiker und Kolumnist Martin Gardner (1914–2010) 1959 mit dem „Mädchen oder Junge?“ eine Zuschriftenlawine im *Scientific American* losgetreten hatte.)

Vor Beschimpfungen per Email schützen Spam-Filter. Die basieren auf einem Satz aus der Wahrscheinlichkeitstheorie, verlesen 1763 auf einer Sitzung der Royal Society in London.

[7] In einem Viertel aller Fälle hat eine Zwei-Kind-Familie zwei Töchter, in drei Viertel aller Fälle wenigstens eine Tochter. Wir sehen davon ab, dass Knabengeburten etwas häufiger sind als Mädchengeburten, was übrigens Pierre-Simon Laplace bereits wusste.

[8] Hinweis: Ist eines der Kinder durch ein Merkmal, Alter, Hund . . . , irgendwie ausgezeichnet, gibt's acht gleichwahrscheinliche Paarungen!