Karin BINDER, Regensburg & Stefan KRAUSS, Regensburg

Das Häufigkeitsnetz – Wahrscheinlichkeiten einfach und verständlich kommunizieren

Einleitung

Täglich werden wir mit einer Vielzahl statistischer Informationen, Daten, Diagrammen, Tabellen und Fakten konfrontiert. Werden statistische Informationen falsch verstanden, kann dies in manchen Fällen folgenschwer sein. Im Bereich der Medizin gibt es beispielsweise regelmäßig Fälle von falschen Diagnosen und unnötigen Operationen (Wegwarth & Gigerenzer, 2013) und auch in der Rechtsprechung treten immer wieder Fehlurteile auf, die zur Verurteilung Unschuldiger oder dem Freispruch schuldiger Personen führen können (Schneps & Colmez, 2013). Diese Fehlinterpretationen hängen dabei in vielen Fällen mit einem falschen Verständnis *bedingter Wahrscheinlichkeiten* oder *Schnittwahrscheinlichkeiten* zusammen, wenn Situationen mit zwei Merkmalen mit dichotomen Ausprägungen betrachtet werden.

Baumdiagramme und Vierfeldertafeln

Im Stochastikunterricht werden Situationen mit zwei dichotomen Merkmalen in der Regel mithilfe von Vierfeldertafeln oder Baumdiagrammen visualisiert. Abbildung 1 zeigt, dass diese beiden Visualisierungen mit absoluten Häufigkeiten oder Wahrscheinlichkeiten ausgefüllt werden können.

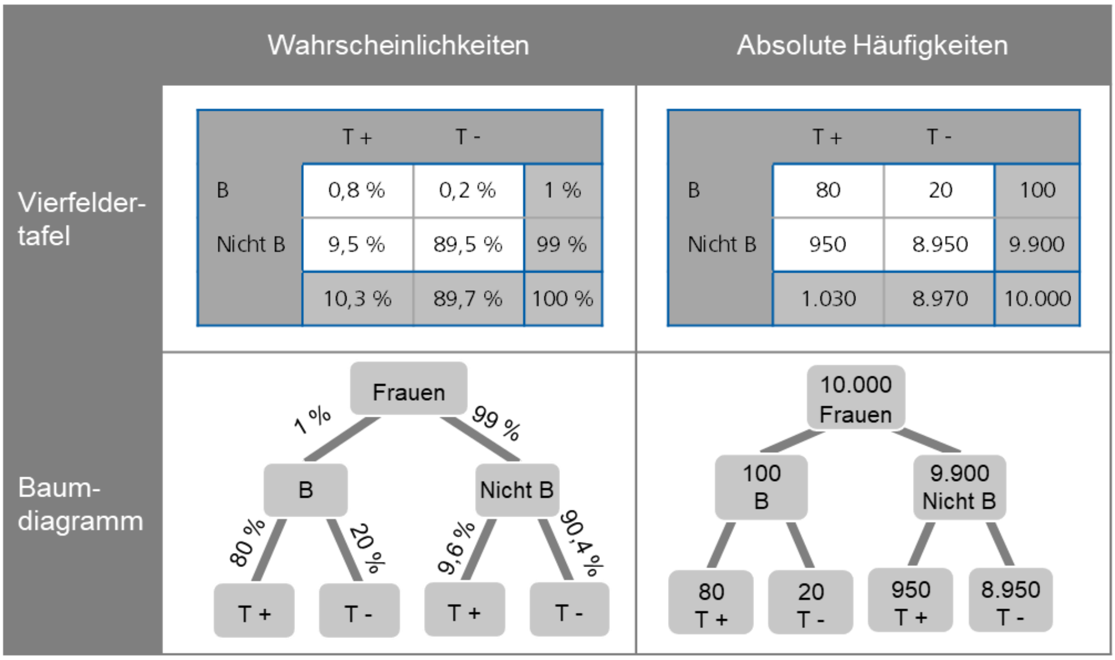

Abb. 1: Vierfeldertafeln und Baumdiagramme mit Wahrscheinlichkeiten oder absoluten Häufigkeiten für das Mammographie-Problem (B: Brustkrebs, T+: Test positiv)

Siller, H.-S., Weigel, W. & Wörler, J. F. (Hrsg.). *Beiträge zum Mathematikunterricht 2020.* Münster: WTM-Verlag, 2020.
doi: 10.37626/GA9783959871402.0

Frühere Studien haben gezeigt, dass bei Bayesianischen Aufgaben, in denen invertierte bedingte Wahrscheinlichkeiten berechnet werden müssen, Baumdiagramme und Vierfeldertafeln mit absoluten Häufigkeiten besonders hilfreich für Schülerinnen und Schüler sind, während Baumdiagramme und Vierfeldertafeln mit Wahrscheinlichkeiten zu erstaunlich niedrigen Performanzen führen (Binder, Krauss & Bruckmaier, 2015). Insgesamt zeigt sich (auch bei Aufgaben ohne Visualisierungen) ein positiver Effekt sogenannter „natürlicher Häufigkeiten" im Vergleich zu Wahrscheinlichkeiten, wenn Bayesianische Aufgaben gestellt werden (Gigerenzer & Hoffrage, 1995; McDowell & Jacobs, 2017).

Didaktisch ist es wünschenswert, dass in einer Visualisierung Wahrscheinlichkeiten und absolute Häufigkeiten gleichzeitig dargestellt werden, da so das Häufigkeitskonzept genutzt werden kann, um Wahrscheinlichkeiten besser zu verstehen. In Vierfeldertafeln ließe sich dies zwar realisieren, indem man in jede der Zellen beispielsweise oben die absoluten Häufigkeiten einträgt und darunter die zugehörigen Wahrscheinlichkeiten; in Baumdiagrammen haben absolute Häufigkeiten und Wahrscheinlichkeiten aber aufgrund ihrer *Knoten-Ast-Struktur* jeweils einen festen eigenen Platz, was deutlich übersichtlicher ist: Absolute Häufigkeiten können in die Knoten eingetragen werden und die Wahrscheinlichkeiten finden an den Ästen Platz.

Die Grenzen von Baumdiagrammen und Vierfeldertafeln

Konzentriert man sich nicht nur auf Bayesianische Aufgaben, können in Situationen mit zwei Ereignissen A und B (und deren Gegenereignissen $\bar{A}$ und $\bar{B}$) grundsätzlich 16 elementare Wahrscheinlichkeiten betrachtet werden: Die vier Randwahrscheinlichkeiten $P(A)$, $P(B)$, $P(\bar{A})$ und $P(\bar{B})$, die vier Schnittwahrscheinlichkeiten $P(A \cap B)$, $P(\bar{A} \cap B)$, $P(A \cap \bar{B})$ und $P(\bar{A} \cap \bar{B})$, und acht bedingte Wahrscheinlichkeiten $P(A|B)$, $P(\bar{A}|B)$, $P(A|\bar{B})$, $P(\bar{A}|\bar{B})$, $P(B|A)$, $P(\bar{B}|A)$, $P(B|\bar{A})$, $P(\bar{B}|\bar{A})$.

Vierfeldertafeln enthalten *Schnittwahrscheinlichkeiten*, allerdings gibt es keinen festen Platz, um bedingte Wahrscheinlichkeiten einzutragen. Abbildung 2a zeigt, dass Vierfeldertafeln unübersichtlich werden, wenn man versucht, die noch fehlenden acht bedingten Wahrscheinlichkeiten dort einzuzeichnen.

Baumdiagramme dagegen enthalten *bedingte Wahrscheinlichkeiten*. Während in „einfachen" Baumdiagrammen nur vier bedingte Wahrscheinlichkeiten abgebildet werden, finden in „Doppelbäumen" alle acht bedingten Wahrscheinlichkeiten Platz. Allerdings werden in (Doppel)bäumen üblicherweise keine Schnittwahrscheinlichkeiten gezeigt. Zeichnet man hierzu noch entsprechende Äste ein, wird auch diese Visualisierung schnell unübersichtlich (siehe z.B. Abbildung 2b).

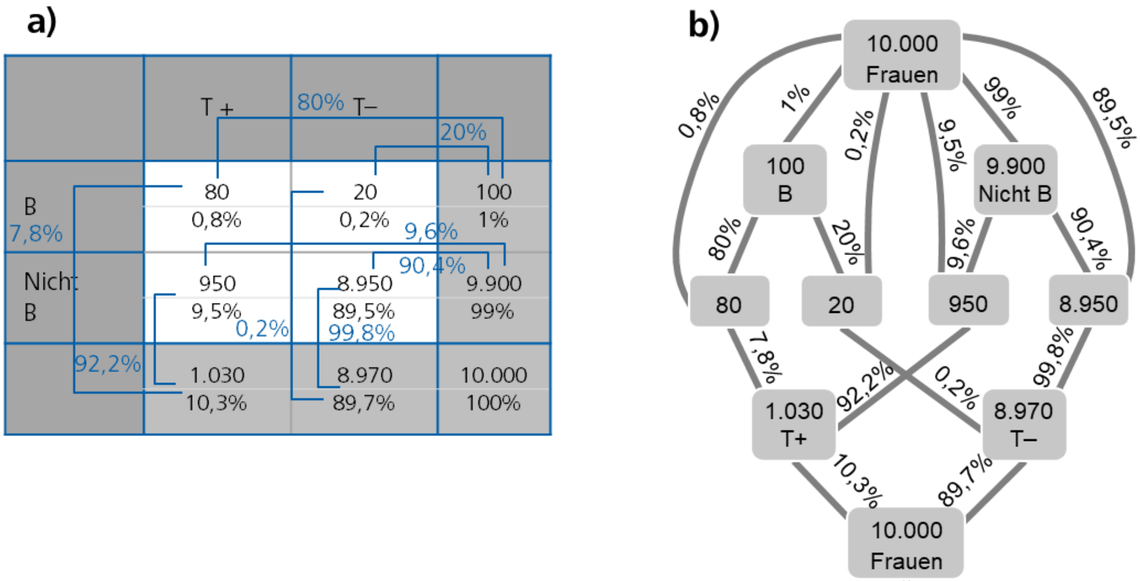

Abb. 2: a: Vierfeldertafel, in der zusätzlich bedingte Wahrscheinlichkeiten eingetragen sind; b: Doppelbaum mit zusätzlichen Ästen für die Schnittwahrscheinlichkeiten

Für den Unterricht bedeutet dies konkret, dass nur manche Aufgaben mit Vierfeldertafeln und (Doppel-)Bäumen direkt mit der Visualisierung gelöst werden können. Bei entsprechend anderen Aufgabentypen muss die Arbeit mit der Abbildung unterbrochen werden, um mithilfe einer Nebenrechnung zunächst fehlende Zahlenwerte zu ermitteln. Erst dann kann die entsprechende Visualisierung vollständig ausgefüllt werden und das Ergebnis gegebenenfalls abgelesen werden.

Wünschenswert wäre daher eine übersichtliche *Knoten-Ast-Struktur*, in der 1) im Vergleich zum Doppelbaum noch zusätzliche Äste für die vier Schnittwahrscheinlichkeiten vorgesehen sind, 2) sich keine Äste wie im Doppelbaum überschneiden und 3) keiner der Knoten doppelt vorkommt (vgl. oberster und unterster Knoten im Doppelbaum, Abbildung 2b).

Vom Doppelbaum zum Häufigkeitsnetz

Um diese drei Eigenschaften zu erfüllen, wurde als neue Visualisierung das *Häufigkeitsnetz* entwickelt (siehe Abbildung 3). Es stellt eine Knoten-Ast-Struktur dar und kann somit gleichermaßen absolute Häufigkeiten und Wahrscheinlichkeiten repräsentieren. Außerdem werden alle vier Randwahrscheinlichkeiten, vier Schnittwahrscheinlichkeiten und acht bedingten Wahrscheinlichkeiten abgebildet. Im Netz gibt es weiterhin keinerlei Überschneidungen oder doppelte Knoten mehr.

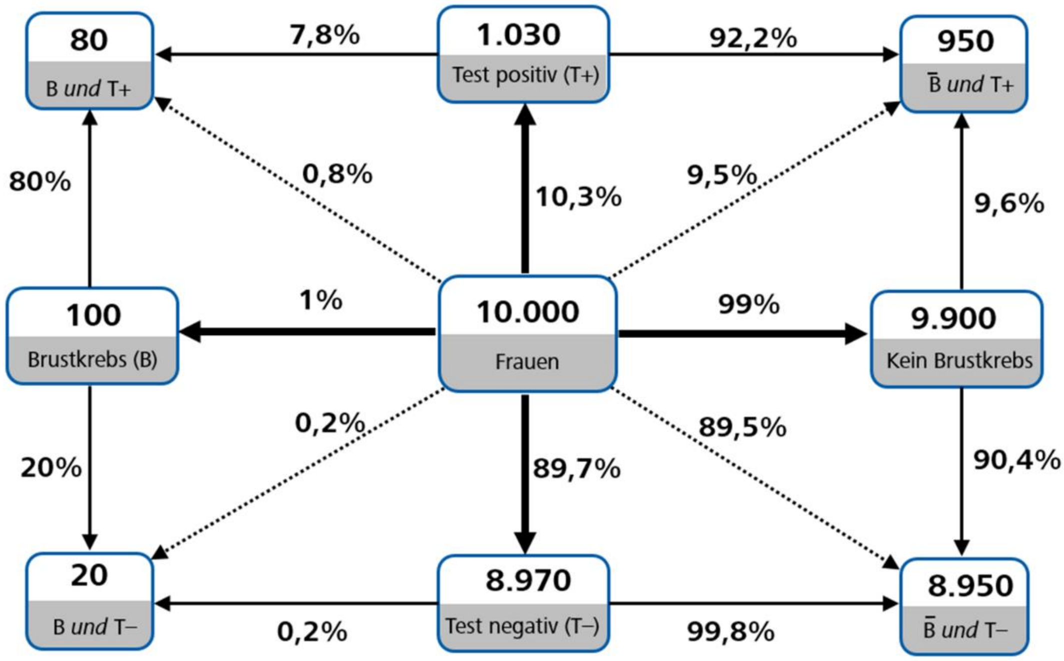

Abb. 3: Häufigkeitsnetz zur Darstellung aller vier Randwahrscheinlichkeiten, vier Schnittwahrscheinlichkeiten und acht bedingten Wahrscheinlichkeiten

Das Häufigkeitsnetz wurde bereits empirisch untersucht (Binder, Krauss & Wiesner, in review) und ein konkreter Vorschlag unterbreitet, wie die neue Visualisierung unterrichtlich genutzt werden kann (Binder, Krauss & Steib, in review). Die Ergebnisse zweier empirischer Untersuchungen zum Häufigkeitsnetz wurden auf der GDM als Poster präsentiert (siehe Steib, Binder & Krauss und Wiesner, Binder & Krauss in diesem Tagungsband).

Literatur

Binder, K., Krauss, S. & Bruckmaier, G. (2015). Effects of visualizing statistical information – an empirical study on tree diagrams and 2×2 tables. *Frontiers in psychology*, 6(1186).

Binder, K., Krauss, S. & Wiesner, P. (in review). A new visualization for probabilistic situations containing two binary events – the frequency net.

Binder, K., Krauss, S. & Steib, N. (in review). Bedingte Wahrscheinlichkeiten und Schnittwahrscheinlichkeiten GLEICHZEITIG visualisieren: Das Häufigkeitsnetz

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102(4), 684–704.

McDowell, M. & Jacobs, P. (2017). Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychological bulletin*, 143(12), 1273–1312.

Schneps, L. & Colmez, C. (2013). Math on trial: How Numbers Get Used and Abused in the Courtroom. New York, NY: Basic Books.

Wegwarth, O. & Gigerenzer, G. (2013). Overdiagnosis and overtreatment: evaluation of what physicians tell their patients about screening harms. *JAMA internal medicine*, 173(22), 2086–2088.