Modellieren im Stochastikunterricht

Annahmen hinterfragen, Ergebnisse validieren

Das Nutzen von Wahrscheinlichkeiten eröffnet die Chance, im Stochastikunterricht gemäß der ersten Winter'schen Grunderfahrung authentische Fragestellungen mit mathematischen Mitteln zu untersuchen. Zentral dafür sind aber Aufgaben, die mehr sind als zweckmäßige, realitätsfremde Einkleidungen.

ROLF BIEHLER, BIRGIT GRIESE

Aufgaben spielen eine Schlüsselrolle beim Lernen – auch im Stochastikunterricht. Doch viele Aufgaben in Schulbüchern und Prüfungen konterkarieren das Erreichen des Kompetenzziels „Mathematisch Modellieren", weil die Sachsituation nicht wirklich ernst genommen wird und die angestrebte schulische Bearbeitung für die Situation keine befriedigende Lösung darstellt. Die Aufgaben folgen anderen Lernzielen, etwa dem Üben von Verfahren – und ergeben sich eben nicht in sinnvoller Weise aus dem Sachkontext. Realitätsprüfungen lassen dann implizit getroffene Modellannahmen oft obsolet werden.

Wir möchten für solche problematischen Aufgaben sensibilisieren und Alternativen für den Unterricht aufzeigen. Dabei beschränken wir uns auf unterrichtliche Aktivitäten, in denen Wahrscheinlichkeitsmodelle genutzt werden; Aufgaben aus der Beschreibenden Statistik per se bleiben unberücksichtigt. Reale Daten werden im Kontext der Modellierungen einbezogen.

Was hinter Stochastikaufgaben steht – aber niemand sagt

Wahrscheinlichkeitsberechnungen basieren immer auf Modellannahmen, die im Unterricht aber oft implizit bleiben. So kann deren Angemessenheit gar nicht überprüft werden. Das betrifft Annahmen der stochastischen Unabhängigkeit, der Konstanz von Wahrscheinlichkeiten, der zufälligen Auswahl sowie der Annahme einer Binomial- oder Normalverteilung. In der außerschulischen Praxis wird der Vergleich von Modell und Daten mit statistischen Tests gemacht, die nicht zum Unterrichtsstoff gehören. Didaktische Elementarisierungen für den Modell-Daten-Abgleich sind also gefragt.

Stochastisch unabhängig?

Wenn mehrere Würfel oder Münzen geworfen werden oder wenn aus einer Urne mit Zurücklegen gezogen wird, werden in der Schule in der Regel Berechnungen durchgeführt, die implizit stochastische Unabhängigkeit voraussetzen. Das mag beim Umgang mit idealen Wurfgeräten und Urnen noch didaktisch vertretbar sein, aber wenn diese Annahme unreflektiert auf andere Situationen übertragen wird, wird es problematisch. Ein typisches Beispiel für eine solche (durchaus schulübliche) Aufgabe lautet: Eine Maschine wiederholt einen Arbeitsgang oder stellt Bauteile her – und ermittelt werden soll die Fehlerwahrscheinlichkeit. Aber hier ist a priori offen, inwiefern dieser Prozess stochastisch unabhängig passiert: Es wird stochastische Unabhängigkeit vorausgesetzt, aber vielleicht entstehen Fehler bei den Bauteilen immer dann, wenn Dreckpartikel zufällig in bestimmte Teile der Maschine gelangen, und enden erst, wenn sich die Dreckpartikel zufällig gelöst haben.

Ein möglicher Weg, dies im Unterricht zu thematisieren, ist, ausgehend von einer Aufgabenstellung, welche stochastische Unabhängigkeit implizit voraussetzt, diese implizit gemachten Modellannahmen herauszuarbeiten und zu diskutieren, ob die Annahmen im vorgestellten Kontext vertretbar sind (vgl. Biehler/Eichler 2015, S. 77, Parkplatzaufgabe). Im Beitrag **Grüne Welle – was geht?** wird das Szenario zweier Ampeln mit vorgegebenen Grün- und Rotphasen systematisch daraufhin untersucht, ob die Annahme der stochastischen Unabhängigkeit, die für diesen Kontext in manchen Schulbüchern implizit unterstellt wird, angemessen ist bzw. prinzipiell angemessen sein kann.

Konstante Wahrscheinlichkeit?

In der Schule wird sehr häufig die *Binomialverteilung* als Standardmodell für wiederholte Zufallsexperimente mit zwei Ausgängen genutzt. Das Galton-Brett wird oft als materielle Verkörperung dieses Modells gesehen. Die Frage „Wie verteilen sich die Kugeln in

Steinbring 1985) wird oft nicht gestellt. Die Annahme einer konstanten Ablenkungswahrscheinlichkeit mit $p = 0{,}5$ gilt als gesetzt. Hier wird selten angezweifelt, ob die Modellannahme eines unveränderten Wertes von p in allen Stufen des Experiments mit der Realität im Einklang steht. Auch die stochastische Unabhängigkeit wird nicht infrage gestellt. Der Beitrag **Echt schräg – Galton revisited** arbeitet dies auf, indem reale Daten aus Experimenten mit Galton-Brettern herangezogen werden, um die Modellierung als Binomialverteilung zu validieren und auch für schräg gestellte Galton-Bretter zu untersuchen. Dabei werden vorhergesagte Häufigkeiten und beobachtete Häufigkeiten durch Vergleiche grafischer Darstellungen beurteilt, um die Angemessenheit des Modells einzuschätzen.

Einsatz von Simulationen

Realitätsnähere Modelle, die besser als die Binomialverteilung passen, sind meist deutlich komplexer, können jedoch oft nur durch Simulationen untersucht werden, so etwa bei Flugüberbuchungen. Diese werden in der Schule typischerweise mit Binomialverteilungen modelliert, wobei eine konstante Stornierungswahrscheinlichkeit und die stochastische Unabhängigkeit der Stornierungen angenommen werden. Mit Simulationen kann beispielsweise berücksichtigt werden, dass Familien gemeinsam reisen und stornieren und dass Geschäftsreisende eher einmal eine Reise nicht antreten (vgl. Biehler/Eichler/Löding/Stender 2015 und den Beitrag **„Die Differenz 5 kommt ja nur einmal vor!“**).

Alles normalverteilt?

Sobald eine Messgröße nicht mehr diskret ist, wird gerne die Normalverteilung herangezogen. Schülerinnen und Schülern wird vermittelt, dass die Normalverteilung zur Beschreibung von Größen wie dem Kopfumfang von Neugeborenen, der Befüllung von Milchtüten oder dem Gewicht von Äpfeln einer Sorte universell taugt. Es wird aber fast nie anhand von vorliegenden Daten geprüft, ob das auch wirklich der Fall ist.

Es ist deshalb interessant und lehrreich, an einem authentischen Datensatz zu untersuchen, ob wirklich eine Normalverteilung (annähernd) vorliegt. Im Beitrag **Alles normal?!** wird an echten Daten aus gesammelten Geburtsanzeigen untersucht, inwiefern das Gewicht und die Körpergröße von Neugeborenen als Normalverteilung modelliert werden kann. Mit einbezogen werden auch Überlegungen zu möglichen Ursachen von Abweichungen. In der Wissenschaft erfolgt der Vergleich von Modell und Daten oft mit statistischen Tests wie dem χ^2- oder dem Kolmogorov-Smirnov-Test. Doch man kann auch mit dem Vergleich geeigneter Diagramme und mit schulüblichen Mitteln wie 95 %-Prognoseintervallen zu relevanten Beurteilungen kommen.

Gleichverteilt?

Wählt man zum Beispiel eine Laplace-Verteilung, um Würfel- oder Münzwurf-Experimente zu beschreiben, so werden der im Experiment konkret benutzte Gegenstand und der Wurfvorgang in der Regel idealisiert: Die Prägung und Dicke der Münze, die Vertiefungen der Würfelpunkte, Materialvariationen, Abnutzungseffekte und verschiedene Wurftechniken oder -höhen werden nicht beachtet. Auch in der Sekundarstufe I können mathematisch reichhaltige Überlegungen dazu angestellt werden, ob ein Würfel gezinkt sein könnte, wie der Beitrag **Fairer oder unfairer Würfel?** zeigt. Die Schülerinnen und Schüler lernen über eigene Würfelexperimente und deren digitale Entsprechungen ein anschaulich nachvollziehbares Verfahren kennen, das ein Kriterium liefert, ob ein Würfel (in Bezug auf die „Sechs“) gezinkt sein könnte.

Wettervorhersagen beruhen auf stochastischer Modellierung unter Einbeziehung umfangreicher Daten und sind heute deutlich verlässlicher als früher

IgorZh/stock.adobe.com

KARL CHARON

Wenn der Würfel fehlt

Auf der Suche nach einem Ersatzwürfel

LERNGRUPPE:	6. Schuljahr
IDEE:	Zufallsgeneratoren untersuchen
ARBEITSBLATT:	Gleichverteilung prüfen
WEITERES MATERIAL:	Korkuntersetzer, Büroklammern, Pinnadeln, Rommékarten
ZEITBEDARF:	1 Doppelstunde

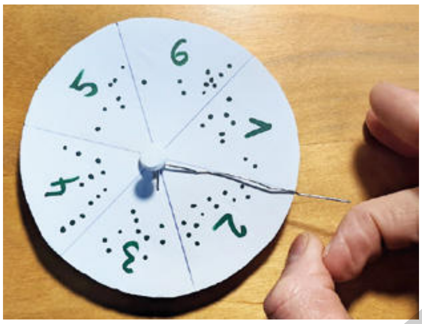


Abb. 1: Glücksrad aus Korkuntersetzer (d = 10 cm) mit Markierungen

Stellen Sie sich vor, Sie bauen das Spiel „Mensch ärgere dich nicht" auf, haben alle Figuren verteilt, sind auf eine spannende Partie eingestellt und dann: Der Würfel fehlt! Für meine Schülerinnen und Schüler war diese Geschichte ein guter Aufhänger einer Unterrichtstunde, in der es darum ging, den Spielwürfel durch äquivalente stochastische Generatoren zu ersetzen – und erste Einblicke in deren Tauglichkeit zu gewinnen.

Den sechsseitigen Würfel haben wir schon ausführlich behandelt und dabei eine mit der Anzahl der Versuche zunehmende Gleichverteilung als Kriterium für die Güte eines Laplace-Zufallsexperimentes kennengelernt – und auch die Abweichungen von den erwarteten Anzahlen einzelner Ausfälle genau betrachtet.

Aufbau der Doppelstunde

Ich starte mit der Idee des verlorenen Würfels. Die Schülerinnen und Schüler teile ich in Dreiergruppen ein und weise diesen je einen konkreten Würfelersatz zu (selbst gebasteltes Glücksrad, Spielkarten oder – ganz ohne Material – Würfeln mit Fingern, vgl. **Arbeitsblatt**):

→ *Beschreibt euer Zufallsexperiment. Worauf müsst ihr bei der Durchführung achten? Ist es ein guter Würfelersatz für „Mensch ärgere dich nicht"?*

Zunächst sammeln die Lernenden Argumente, weshalb ihr Experiment Ergebnisse mit zufälligem Ausgang liefert. Dazu formulieren sie auch Regeln für den Umgang mit dem Experiment, die einen zufälligen Ausgang sichern sollen. Dann folgt die Datenerhebung (wenn auch aus Praktikabilitätsgründen mit kleiner Stichprobe). Der Abgleich zwischen den vorab formulierten Argumenten und Durchführungsregeln und dem realen Verhalten des Zufallsgerätes mündet in einer Neubewertung des Zufallsgerätes, die in einer Pro-und-Kontra-Liste erfasst wird.

Erfahrungen beim Glücksrad

Die Glücksrad-Korkscheiben sind mit Papier beklebt, dies verringert die Reibung des Zeigers. Die Gruppen einigen sich auf „Würfel"-Regeln: Der Zeiger wird so fest angeschnipst, dass er mindestens zwei Umdrehungen vollzieht. Intuitiv sorgen sie damit zumindest für eine Basis stochastischer Unabhängigkeit – ein Umstand, der hier nicht diskutiert wird, aber später aufgegriffen werden könnte. Eine der Gruppen bittet mich zu sich. „Schauen Sie mal, der Zeiger landet meistens links." Ich schlage vor, die Lagen des Zeigers mit Punkten auf der Scheibe zu markieren.

Die Verteilung der 60 Punkte könnte zufällig sein (**Abb. 1**). „Da sind aber viel mehr Punkte bei 2 als bei 5", meint Alex. Ich erinnere an unsere Datenerhebung zum Würfel. Auch dort beobachteten wir solche Abweichungen von der erwarteten Anzahl. Dass dies noch nicht genügt, um verlässliche Aussagen zu treffen, thematisiere ich nicht (dazu s. Nieszporek/Griese 2022). Es bleibt Skepsis: „Wenn da immer gleich stark geschnipst wird, muss das doch öfter auf derselben Stelle landen – und dann hängt es auch davon ab, wo gestartet wird." „Da müssten wir noch viel mehr testen." Als Vorteil jedenfalls zeigt sich das schnelle „Würfeln" mit dem Glücksrad.

Die Spielkarten

Bei Rommékarten stehen (wird Ass = eins gesetzt) die Würfelwerte direkt auf den Karten. Man einigt sich, wie gezogen und vor allem wie „richtig" gemischt werden soll. Denn Sarah kennt einen Trick, wie die oberste Karte nach dem Mischen wieder oben liegt. Diese Anfälligkeit des Experimentes wird bei den Pros und Kontras notiert. Nun beendet Person 1 das Mischen, wenn Person 2 „Stop!" ruft. Dann zieht Person 3 eine Karte.

Die Auswertung der 60 Ziehungen nach ordnungsgemäßem Mischen verläuft ähnlich wie bei den Glücksrädern. Die Schülerinnen und Schüler beobachten bei den Differenzen zu den erwarteten Anzahlen keine auffällig größeren Abweichungen als die ihnen vom fairen Würfeln bekannten. Fazit der Karten-Gruppen: Es ergeben sich keine Anhaltspunkte, an einer Gleich-

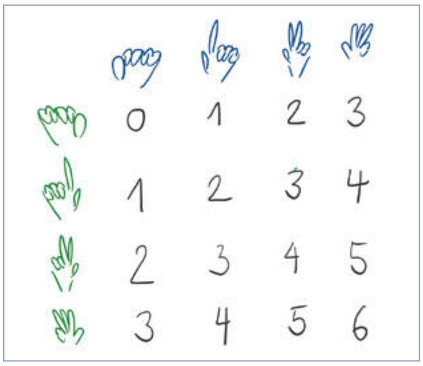


Abb. 2: Finger-Würfeln – kein guter Würfelersatz

Finger-Würfel

Zwei Spieler zeigen auf ein Signal hin 0, 1, 2 oder 3 Finger. Dann wird addiert. „Da gibt es ja eine Null“ bemerkt Tim. Und das wird in der Arbeitsphase nicht der einzige Haken bei diesem Experiment bleiben. Bei Null, so die Entscheidung, soll wiederholt werden. Auf den ersten naiven Blick wirkt diese Anordnung (**Abb. 2**) wie ein guter Würfelersatz. Wie sieht es mit der Zufälligkeit aus? Zwar nimmt jeder Einfluss (Wahl der gezeigten Finger), doch die Ergänzung durch den zweiten Spieler liefert ein unvorhersehbares Element. Wobei: „Wenn ich nicht will, dass eine Sechs rauskommt, zeige ich einfach nur 2 Finger“, meint Esra.

Wie die Datenerhebung zeigt, kann von einer Gleichverteilung keine Rede sein. Die willentliche Beeinflussung der Gleichverteilung ist schnell erkannt. Eine ganz anders gelagerte Problematik taucht auf, als eine der Gruppen auf meinen Impuls hin alle möglichen Fingerkombinationen in einer Tabelle ermittelt. Nun wird auch sichtbar, weshalb eine Gleichverteilung der „Augenzahlen“ selbst bei zufälliger Auswahl der gezeigten Finger nicht erfolgt.

Austausch im Plenum

Die Gruppen berichten von ihren Ideen und Erfahrungen. Sarah stellt ihren Kartentrick vor. Der Bericht einer Finger-Würfel-Gruppe zeigt: Durch die beobachtete ungleiche Verteilung der „Fingersummen“ entstehen neue Spielideen („Blitz-Mensch-ärgere-dich-nicht“ mit einem Glücksrad, dessen halbe Fläche nur für die Sechs vorgesehen ist. Was das wohl wird? …)

Die Lernenden haben immer wieder erfahren, dass sie ihre eigene Intuition prüfen sollten – und dies auch können. Im Moment noch mit einem eher groben Kriterium: Ist die Verteilung auffällig (im Sinne einer größeren Abweichung von erwarteten Werten als beim fairen Würfeln)? Die Annahmen zu den Modellen wurden immer wieder infrage gestellt, die Modellierungen sind an Bedingungen geknüpft: Bedingungen, die sie zum Teil kontrollieren können, und solche, deren Einfluss manchmal ausgeblendet werden muss und kann.

Literatur

Nieszporek, R./Griese, B.: Fairer oder unfairer Würfel. Elementare Modellbewertung – In: mathematik lehren 232, S. 13 – 18.

Grafik: K. Charon

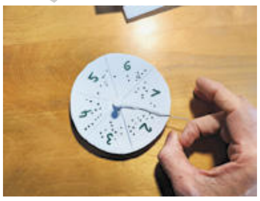
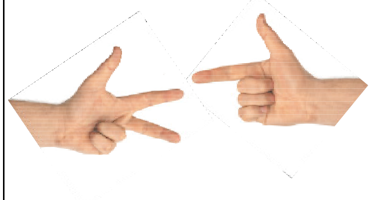
1

Arbeitsblatt

Name: ____________________ Datum: ____________________

Würfel weg – was nun?

In dieser Aufgabe geht es darum, einen sechsseitigen Spielwürfel zu ersetzen.
Arbeitet in einer Dreiergruppe mit einem Würfel-Ersatz aus der folgenden Liste.

A Glücksrad	**B** Spielkarten	**C** Finger
		
Die Büroklammer wird mit einer Pinnadel drehbar als Zeiger auf der Korkscheibe fixiert. Mit einem „Fingerschnippen“ wird gedreht.	Sechs Karten werden aus dem Stapel genommen. Das Ass zählt als eins. Eine Karte wird verdeckt gezogen	Zwei Personen zeigen mit einer Hand eine Anzahl an Fingern zwischen Null und Drei. Die Summe ist der Würfelwert.

a) Gebt Argumente an, weshalb euer Würfel-Ersatz zufällige Zahlen liefert. Entscheidet gemeinsam über Regeln, wie er verwendet werden soll.
b) „Würfelt“ mit eurem Zufallsgerät 60-mal. Führt dazu eine Strichliste.
c) Gebt eine Einschätzung ab: Ist euer Würfel-Ersatz eine Alternative zum sechsseitigen Würfel?
d) Vergleicht eure Strichliste mit denen anderer Grupp
Führt die Listen zusammen.
e) Formuliert Pro- und Contra-Argumente für den Ein

KARIN BINDER, WERNER BLUM, STEFAN KRAUSS

Gesichtserkennung

Wie verlässlich sind die Ergebnisse?

LERNGRUPPE: 9.-10. Schuljahr

IDEE: Modellieren und Validieren bei realitätsbezogenen Aufgaben zu bedingten Wahrscheinlichkeiten

MATERIAL 1: Sachinformationen (Online: Auswertungsvarianten)

ZEITBEDARF: 2 Unterrichtsstunden

Bedingte Wahrscheinlichkeiten werden im Unterricht oft im Kontext medizinischer Diagnoseverfahren (z. B. Corona-Tests) thematisiert. Das Beispiel der technischen *Gesichtserkennung* (Gigerenzer 2018) bietet dazu eine gesellschaftspolitisch relevante Alternative.

In diesem Unterrichtsvorschlag durchlaufen wir einen vollständigen Modellierungskreislauf und zeigen, dass hier die Modellannahmen besonders diskussionswürdig sind. Die Unterrichtsgespräche dazu können beitragen, das Validierungsbewusstsein auszuschärfen. Im Fokus stehen dabei die Leitfragen: *Warum überrascht das Ergebnis so sehr? Welchen Einfluss haben dabei die getroffenen Modellannahmen?*

Der Kontext: Projekt zur Gesichtserkennung erfolgreich

Im Oktober 2018 vermeldete das Bundesinnenministerium, der Test der technischen Gesichtserkennungssysteme am Bahnhof Berlin-Südkreuz sei erfolgreich verlaufen. So äußerte sich Bundesinnenminister Seehofer mit den Worten „Die Systeme haben sich in beeindruckender Weise bewährt, sodass eine breite Einführung möglich ist“ (BMI 2018). Ziel der Gesichtserkennung ist die Erfassung potenzieller Gefährder (die Amokläufe planen) an Bahnhöfen durch Abgleich der Kamerabilder mit eingespeisten Fahndungsfotos. Die Schülerinnen und Schüler erkunden dazu folgende Fragen:

→ *Chancen und Probleme der Einführung einer flächendeckenden Gesichtserkennung: Wie verlässlich sind die Ergebnisse?*

→ *Mit welcher Wahrscheinlichkeit ist eine Person, die Alarm ausgelöst hat, tatsächlich ein Gefährder?*

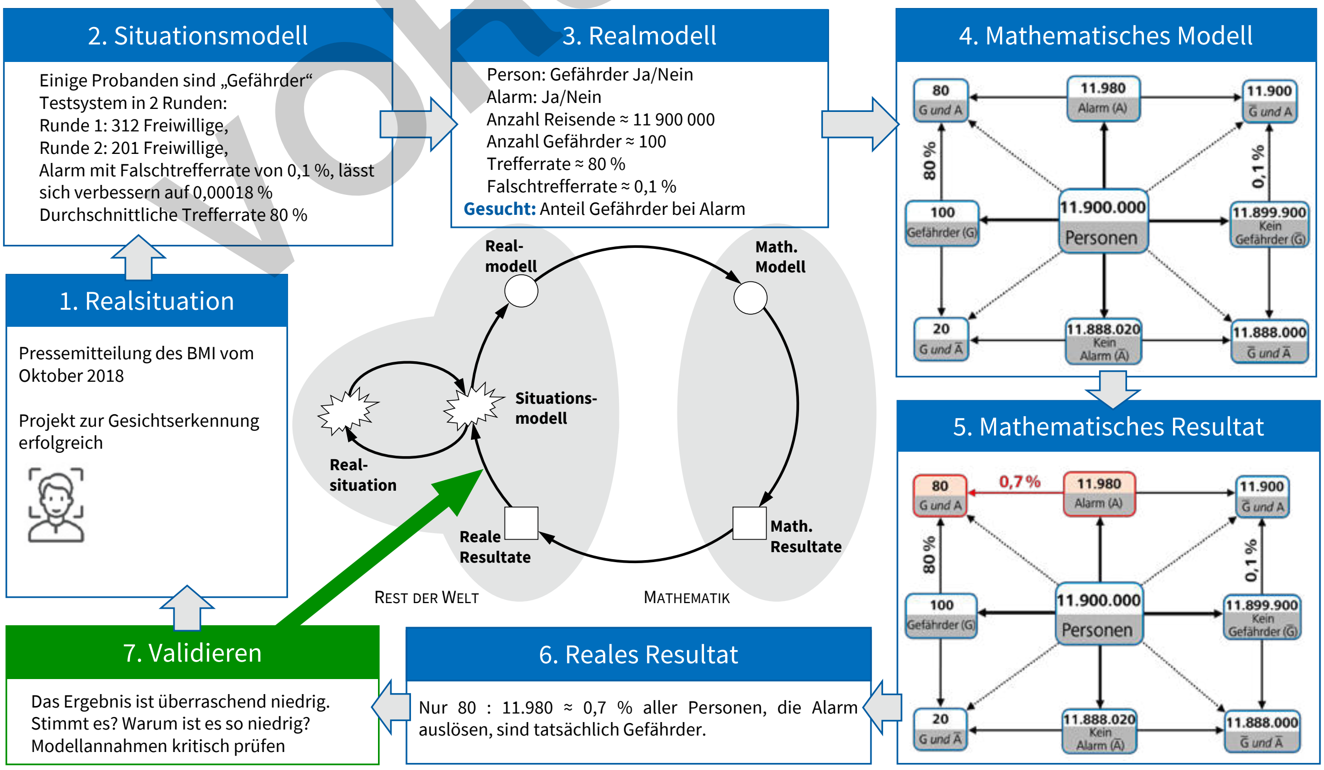


Abb. 1: Wie verlässlich sind die Ergebnisse des Gesichtserkennungs-Projekts in Berlin 2018? Ein exemplarischer Modellierungskreislauf (nach Blum/Leiß 2005, s. Blum 2006) für den Unterricht.

Dahinter steht die Suche nach einer „bedingten Wahrscheinlichkeit“, was aber unterrichtlich nicht expliziert werden muss.

Modellierungskreislauf

Ausgehend von der obigen Meldung oder dem zugehörigen Abschlussbericht der Bundespolizei (s. **Material 1**) kann ein vollständiger Modellierungsprozess (Abb. 1) durchlaufen werden, in dem verschiedene Modellparameter kritisch hinterfragt werden.

Bilden eines Realmodells

Der erste Schritt besteht im Verstehen des Arbeitsauftrags mit dem Ziel, ein Situationsmodell zu bilden. In der Pressemitteilung (BMI 2018) und dem Abschlussbericht (Material 1) finden sich einerseits statistische Angaben wie die durchschnittliche Trefferrate von 80 % und die Falschtrefferrate von 0,1 %, jedoch auch zahlreiche Informationen, die für eine spätere Mathematisierung unbeachtet bleiben können. Die Trefferrate gibt an, wie groß unter den Gefährdern der Anteil der Personen ist, bei denen die Gesichtserkennung korrekterweise Alarm schlägt (richtig-positives Testergebnis). Die Falschtrefferrate gibt hingegen unter den Nicht-Gefährdern den Anteil der Personen an, bei denen die Gesichtserkennung fälschlich alarmiert (falsch-positives Testergebnis). Darüber hinaus werden im Abschlussbericht auch Angaben gemacht, die erst bei einem erneuten Durchlaufen des Modellierungskreislaufs Berücksichtigung finden (etwa eine verbesserte Falschtrefferrate).

An dieser Stelle bietet sich eine Schätzung an:

→ *Stellt euch vor, heute würden am Frankfurter Bahnhof Kontrollen durchgeführt und bei 150 Personen hätte die Gesichtserkennung Alarm ausgelöst.*
Wie groß schätzt ihr den Anteil der tatsächlichen Gefährder unter diesen 150 Personen?
Wie groß müsste dieser Anteil sein, damit ihr die Gesichtserkennung als erfolgreich anseht?

Im zweiten Schritt kann auch noch eine Schätzung auf Basis der Zahlenangaben für Trefferrate usw. erfolgen. Die Schätzergebnisse werden schriftlich festgehalten und später mit den Berechnungen verglichen.

Über das mathematische Modell zur Wirklichkeit

Wie die Lernenden verstehen, können die Bahnreisenden einerseits aufgeteilt werden in Gefährder und Nicht-Gefährder und andererseits in Personen, die Alarm auslösen, und solche, die keinen Alarm auslösen. Überdies sind hierzu verschiedene Anteilswerte bekannt. Eine derartige Situation kann mithilfe einer Vierfeldertafel, eines (Doppel-)Baumes, eines Einheitsquadrates oder eines Häufigkeitsnetzes (unsere Präferenz) modelliert werden. **Abb. 1** zeigt die Modellierung mithilfe eines Häufigkeitsnetzes (zu dessen Vorteilen s. Binder u. a. 2021).

In einem Gedankenexperiment kann man nun überlegen, was es bedeuten würde, wenn morgen an allen deutschen Bahnhöfen die Gesichtserkennung eingeführt würde. Da sich jeden Tag etwa 11 900 000 Personen an deutschen Bahnhöfen aufhalten, von denen etwa 100 Gefährder sind, werden aus dieser Information die