



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Achtung Verwechslungsgefahr: Warum man der Intuition insbesondere in der Stochastik nicht immer trauen sollte“

verfasst von / submitted by

Theresa Martha Schuh

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 190 406 353

Studieneinrichtung lt. Studienblatt/
degree programme as it appears on
the student record sheet:

Lehramtsstudium UniStG
UF Mathematik UniStG
UF Spanisch UniStG

Betreut von / Supervisor:

ao. Univ.-Prof. Mag. Dr. Peter Raith

Danksagung

In erster Linie möchte ich mich bei meinem Betreuer ao. Univ.-Prof. Mag. Dr. Peter Raith für die professionelle, nette und zuverlässige Unterstützung beim Verfassen dieser Arbeit bedanken. Da er mich schon während meines gesamten Studiums begleitete, bin ich froh, dass ich die Möglichkeit hatte auch meine Abschlussarbeit unter seiner Aufsicht zu schreiben. Mit seiner Hilfe konnte ich meinen Ideen Form geben und konnte stets auf seinen Rat zurückgreifen.

Der wohl größte Dank gebührt meiner Familie und meinen Freunden. Danke an meinen Eltern und meinen Brüdern, dass sie mich zu dem Menschen erzogen haben, der ich heute bin. Ohne sie wäre ich wohl, sowohl in finanzieller als auch in emotionaler Hinsicht, nicht so weit gekommen. Mein Freund Markus half mir in jeder Angelegenheit weiter, und es gelang ihm mich in stressigen Situationen wieder zu beruhigen - danke für alles. Gemeinsam mit meinen Freunden schafften sie es mir stets ein Lachen in das Gesicht zu zaubern und mir in schweren Zeiten den Rücken zu stärken. Außerdem bin ich dankbar für die Begegnungen bei denen aus Studienkolleginnen und Studienkollegen Freunde wurden.

Ein weiterer Dank gilt Elisabeth und Albert für ihre kompetente Hilfe und ihre Zeit, die sie mit meiner Arbeit verbracht haben.

Ich möchte auch Direktorin Mag.^a Ingrid Weltler-Müller und Direktor Mag. Robert Antoni für die Möglichkeit der Durchführung der empirischen Untersuchung an den beiden Schulen, dem BG/BRG/BORG Oberschützen und dem BORG Güssing, danken. Durch ihre Flexibilität und unkomplizierte Terminvereinbarung konnte ich interessante Ergebnisse erzielen. Ebenso gilt der Dank sowohl den Lehrerinnen und Lehrern als auch den Schülerinnen und Schülern, die dadurch betroffen waren.

Inhaltsverzeichnis

Einleitung.....	1
1. Was ist Stochastik? Eine kurze Einführung.....	3
2. Geschichtliches zum Thema Stochastik.....	7
2.1 Die ersten Probleme der Stochastik	8
2.1.1 Das Teilungsproblem und die Lösungsvorschläge von Pacioli, Pascal und Fermat.....	8
2.1.2 Das Würfelproblem und der Lösungsvorschlag von de Méré	9
2.1.3 Das St. Petersburger Paradoxon und der Lösungsvorschlag von Daniel Bernoulli.....	10
3. Ein Einblick in die beschreibende Statistik.....	12
3.1 Relative und absolute Häufigkeiten.....	12
3.2 Bernoullisches schwaches Gesetz der großen Zahlen	13
4. Der Weg zu den ersten Wahrscheinlichkeitsbegriffen.....	15
4.1 Was ist ein (Zufalls-)Experiment? Ein Einblick in die Schule.....	15
4.2 Zufallsexperimente und Mengen aus mathematischer Sicht	16
4.3 Laplace und die Wahrscheinlichkeit.....	17
4.4 Subjektive Wahrscheinlichkeit	18
4.5 σ -Algebra	20
4.6 Das Wahrscheinlichkeitsmaß.....	21
4.7 Kolomogorov und die Wahrscheinlichkeit	23
5. Was hat Zählen mit Wahrscheinlichkeiten zu tun? - Kombinatorik.....	25
5.1 Permutationen.....	27
5.2 Kombinationen.....	28
5.3 Modellvorstellung zur Kombinatorik.....	28
5.3.1 Urnen-Modell.....	29
5.3.2 Fächer-Modell.....	30
6. Erste Versuche zum Berechnen von Wahrscheinlichkeiten	32
6.1 Bedingte Wahrscheinlichkeit	32
6.2 Satz von Bayes	36
6.3 Das Ziegenproblem: das klassische Beispiel der bedingten Wahrscheinlichkeit.....	38
6.3.1 Lösung des Ziegenproblems	39
6.3.2 Interpretation des Ziegenproblems	40
6.4 Verwechslung bei der bedingten Wahrscheinlichkeit	41
6.5 Das Simpson-Paradoxon.....	45
6.6 Stochastische Unabhängigkeit	45
6.7 Baumdiagramm.....	46
6.7.1 Pfadregeln.....	48
7. Zufallsvariable.....	52

7.1 Rechenregeln von Zufallsvariablen	53
7.2 Erwartungswert, Varianz und Standardabweichung einer Zufallsvariable	54
7.3 Verteilungen	55
7.3.1 Binomialverteilung.....	55
7.3.2 Multinomialverteilung.....	57
7.3.3 Hypergeometrische Verteilung.....	58
7.3.4 Geometrische Verteilung	59
7.3.5 Poissonverteilung.....	60
7.3.6 Exponentialverteilung	63
7.3.7 Normalverteilung.....	64
8. Zufall und Simulation.....	66
9. Ein Auszug zur beurteilenden Statistik	67
9.1 Parameterschätzung	68
9.2 Hypothesentests	70
9.3 Chi-Quadrat-Test.....	73
10. Intuition und Wahrscheinlichkeit – Freund oder Feind?	75
11. Der Umgang mit den Begriffen Zufall, Statistik und Wahrscheinlichkeit in der Schule	82
11.1 Unterricht und Stochastik.....	84
11.2 Lehrplan für Mathematik in Bezug auf Stochastik.....	87
12. Erhebung der Wahrnehmung von Aufgabenstellungen der bedingten Wahrscheinlichkeit an AHS-Schülerinnen und Schülern.....	89
12.1 Ausgewählten Beispiele mit Lösungen.....	89
12.2 Analyse der Lösungen der Schülerinnen und Schüler.....	92
Literaturverzeichnis.....	100
Abbildungen.....	102
Tabellen	103
Anhang	104
Anhang 1: Elternbrief.....	104
Anhang 2: Operationalisierung der Fragestellung.....	105
Anhang 3: Vorlage des Untersuchungsmaterial.....	106
Anhang 4: Genehmigung des LSR Burgenlandes zur Durchführung der Untersuchung...	108
Abstract.....	109

Alle personenbezogenen Begriffe sind als geschlechterneutral aufzufassen. Es wird keine von der Autorin bevorzugte Form hervorgehoben. Aufgrund der Lesbarkeit wurde an manchen Stellen die grammatikalisch näherliegende Bezeichnung für den Gebrauch beider Geschlechter verwendet.

Einleitung

„*Vertrauen ist gut, Kontrolle ist besser*“ - diese bekannte Redewendung soll die Vermutung bestärken, dass man nicht alles im Leben glauben und gewisse Sachverhalte hinterfragen sollte. In diesem Sinne werden in der folgenden Arbeit genügend Argumente geliefert, warum manche Situationen nicht so sind, wie sie auf den ersten Blick scheinen. Auch der Begriff der Intuition wird in Bezug auf die Stochastik kritisch betrachtet.

Der Ausdruck Stochastik ist vielleicht nicht so bekannt wie die zu behandelnden Teilkapitel auf welchen das Thema basiert. Auf der einen Seite beruht es auf dem Unterthema der Statistik und behandelt sowohl die beschreibende, als auch die beurteilende Statistik. Und auf der anderen Seite bildet die Wahrscheinlichkeitstheorie einen wichtigen Beitrag. Das Wort Stochastik selbst lässt sich aus dem griechischen Ausdruck *stochos* ableiten und bedeutet so viel wie *die Kunst des Mutmaßens*. [19] Dem Bereich der Statistik wird hier weniger Beachtung geschenkt. Was aber nicht an ihrer Bedeutung und ihrem Nutzen zweifeln lassen soll. Vorwiegend wird auf das Thema der Wahrscheinlichkeitstheorie eingegangen.

Dennoch widmen sich die ersten Kapitel der beschreibenden Statistik und die letzten Abschnitte der beurteilenden Statistik. Es soll nämlich auch auf die wesentlichen und für die Wahrscheinlichkeitstheorie wichtigen Punkte eingegangen werden, um einen Einblick in die Materie zu bekommen. Auch ein Einblick in die Behandlung dieses Themas in der Schule wird präsentiert.

Die Wahrscheinlichkeitstheorie hat einen langen geschichtlichen Hintergrund. Vermutlich waren die Ursprünge spieltheoretische Probleme. Ausgehend von Überlegungen zum Glücksspiel haben sich dann fortlaufend Gedanken und weiterführende Theorien gebildet. Aber das wahrscheinlichkeitstheoretische Arbeiten umfasst nicht nur mathematisch korrekte Rechnungen. Im Laufe der Zeit haben sich daraus sogar Paradoxa entwickelt. Einige diesbezüglich bedeutsame Aussagen werden in den nächsten Seiten gesondert diskutiert. Es werden sowohl die Lösungen als auch die irrtümlichen Lösungsversuche erläutert.

Nicht nur Paradoxa verursachen, sondern auch die bedingte Wahrscheinlichkeit verursacht immer wieder Denkfehler und Verwirrung. Vor allem soll der Umgang mit diesem Thema in der Schule anhand der empirischen Untersuchung zum Thema „Erhebung der Wahrnehmung von Aufgabenstellungen der bedingten Wahrscheinlichkeit bei AHS-Schülerinnen und Schülern“ dargestellt werden. Es wurden gewisse Beispiele der bedingten Wahrscheinlichkeit ausgewählt für die Schülerinnen und Schüler Lösungen angeben sollten.

Ein wichtiger Faktor, der die Vorstellung der Wahrscheinlichkeit begleitet, ist die Intuition. Der Titel dieser Arbeit weist bereits auf die Relation zwischen dem Bauchgefühl und der Mathematik hin. Allerdings wird dieses Verhältnis skeptisch betrachtet, da die Intuition für Schülerinnen und Schüler durchaus behilflich bei Lernprozessen sein kann, aber auch zu Fehlvorstellungen führen kann. Im Speziellen wird der Instinkt bei der bedingten Wahrscheinlichkeit auf die Probe gestellt, da man bei der gesuchten Wahrscheinlichkeit oft

mit einem unerwarteten Endergebnis überrascht wird. Dieser Effekt tritt in einigen Beispielen auf und soll durch die empirische Untersuchung erklärt werden.

Manche Institutionen machen sich die Strategien zur Darstellung widersprüchlicher Aussagen zu Nutze. Zum Beispiel wird in den Medien, im Gerichtssaal oder sogar im Gesundheitswesen mit dem Themengebiet der Stochastik immer wieder falsch umgegangen. Oft werden nur gewisse Teile einer Studie veröffentlicht, um eine Schlagzeile zu erlangen. Auch in der Justiz und in der Medizin können die verantwortlichen Sachverständigen, Anwälte, Richter und Ärzte gelegentlich nicht richtig mit der Stochastik und vor allem nicht mit der Wahrscheinlichkeit umgehen. Daher werden in dieser Arbeit auch einige berühmte Gerichtsfälle beziehungsweise auch medizinische Fälle, die fälschlicherweise sabotiert wurden, besprochen. Auf diese Manipulationen wird später noch näher eingegangen. Es wird auch versucht die Ansicht der Fallstricke der Wahrscheinlichkeitstheorie durch eine alternative Denkweise mit Hilfe der relativen Häufigkeiten zu verbessern.

Da die Bedeutung des Themas der Stochastik in den letzten Jahren vor allem in den Schulen durch die zentrale Reifeprüfung an Bedeutung dazugewonnen hat, soll diese Arbeit den Wert der Statistik und der Wahrscheinlichkeitstheorie klarstellen. Durch die praxisorientierte Anwendung spielen persönliche Erfahrungen in diesem Bereich eine große Rolle. Genau dieses individuelle Wissen verleitet die Intuition und den Instinkt mitunter zu vorschnellen Mutmaßungen. Diese wiederum können zu falschen Annahmen führen und Verwirrungen im Hinblick auf die Berechnungen erzeugen. Doch in der Stochastik ist nicht nur die Rechnung gefragt, es wird auch der dazugehörigen Interpretationsmöglichkeit ein hoher Stellenwert zugesprochen.

Der Begriff der Intuition bietet viele Vorteile und ist ständiger Begleiter im Leben der Menschen. Daher kann man dessen Auswirkungen auf Denkweisen und Vorstellungen nur wenig beeinflussen, denn man ist durch die eigenen Vorerfahrungen in gewisser Hinsicht manipuliert. Trotzdem kann man im Speziellen in der Schule auf die Intuition der Schülerinnen und Schüler zurückgreifen, um auf deren Ansichten die Theorie aufbauen zu können.

1. Was ist Stochastik? Eine kurze Einführung

Stochastik ist ein sehr umfangreicher und dichter Begriff. Denn dieser Ausdruck beeinflusst viel mehr als nur die Mathematik. Stochastik hat viele Facetten und daher auch ein großes Aufgabengebiet. Auch wenn das Wort Stochastik im ersten Augenblick für viele fremd und kompliziert wirkt, findet man die Bereiche der Stochastik auch im alltäglichen Leben. Deshalb ist es wichtig diesen Fachausdruck zugänglicher zu machen und ihn auch in die Schule einzuführen beziehungsweise mit den Fachbereichen der Stochastik zu arbeiten.

Um Stochastik zugänglicher zu machen, ist zwischen Wahrscheinlichkeitsrechnung und Statistik zu unterscheiden. [13] Um diese beiden Untergebiete der Stochastik näher zu erläutern, wird zunächst versucht, die Begriffe zu erklären. Wobei sich diese Arbeit hauptsächlich der Wahrscheinlichkeitsrechnung widmet und die Statistik nur streift, um Grundsätzliches zu erklären.

Unter Wahrscheinlichkeitsrechnung kann man sich schon mehr vorstellen als unter dem anfänglichen Terminus Stochastik, da man diesen Ausdruck auch aus dem Alltag kennt. Aussagen wie zum Beispiel „die Regenwahrscheinlichkeit für morgen beträgt 65%“ oder „ich glaube, Emil kommt heute später nach Hause“ sind Angaben, die jede und jeder kennt. Dabei sind solche Sätze nichts anderes als bloße Vermutungen. Allerdings wird die Wahrscheinlichkeitstheorie nicht nur für einfache Aussagen benutzt, sondern hat auch einen hohen Gebrauch in anderen Bereichen, wie zum Beispiel in Gerichtsverfahren oder auch in der Spieltheorie, wobei dies zum nächsten Stichwort führt. Einfache wahrscheinlichkeitstheoretische Sätze haben einen beiläufigen Zusammenhang mit der Ungewissheit. [10, S. 133] Auch wenn der Begriff anfangs etwas abschreckend und komplex wirkt, kann man ihn langsam durch solche alltäglichen Situationen etwas greifbarer machen und vor allem in der Schule für die Schülerinnen und Schüler verständnisvoller erklären.

So gelangt man zur Folgerung, dass die Wahrscheinlichkeit etwas mit Vermutungen zu tun haben muss. Die Stochastik bietet allerdings eine Vielzahl von Facetten, die eine Varietät der Interpretationen und der Berechnungen erlaubt. In der Wahrscheinlichkeit dreht sich alles um Ereignisse und Ausgänge eines Zufallsexperimentes. Das bedeutet, man hat verschiedene Möglichkeiten mit unterschiedlichen Ereignissen und versucht mit Hilfe der Wahrscheinlichkeitsrechnung die denkbaren Chancen für ein Ereignis zu berechnen.

Das typischste Beispiel für Wahrscheinlichkeitsrechnung repräsentiert wohl die äußerst bekannte Situation des Werfens einer Münze beziehungsweise eines Würfels. Als Folgerung für einen Laplace-Würfel wird die Frage assoziiert: Wie lange muss man würfeln, um beim *Mensch ärgere dich nicht* die 6 zu bekommen? Dieses Spiel und daher auch diese Frage kennt wahrscheinlich jede und jeder. Dabei ist die Antwort sehr einfach: Die Wahrscheinlichkeit, dass man beim ersten Würfeln eine Sechs würfelt, beträgt genau $\frac{1}{6} = 0,1\bar{6}$. Wie man auf dieses Ergebnis kommt, wird später im Kapitel *Laplace und die Wahrscheinlichkeit* erklärt.

Diese Situation kann auch schrittweise anschaulich gemacht werden: Man wirft mit einen herkömmlichen Würfel. Es ist bekannt, dass ein Würfel aus sechs Seiten besteht, mit 1 bis 6 Würfelaugen. Daher sollte jeder und jedem klar sein, dass jede Zahl von 1 bis 6 mit der

gleichen Wahrscheinlichkeit $\frac{1}{6}$ möglich ist. Trotzdem hat man oft, und vor allem dann, wenn man am Verlieren ist, den Eindruck, dass die 6 nie kommt. Denn es spielt noch eine weitere Komponente eine wichtige Rolle und zwar ist das der Zufall. Denn auch der Zufall hat bestimmte Regeln, wie im Laufe dieser Arbeit noch gezeigt werden wird. [19]

Die zweite Säule der Stochastik ist, wie oben schon erwähnt worden ist, die Statistik. Auf dieses Thema wird in den folgenden Kapiteln jedoch weniger eingegangen. Nichtsdestotrotz sollen eine kurze Erklärung und einleitende Worte nicht fehlen. Denn ebenso wie der Wahrscheinlichkeitsbegriff ist auch die Statistik kein fremder Ausdruck im Leben. Mit Statistik bezeichnet man den Umgang von Daten. Dies erkennt man allein schon in einer Zeitung, wo man ständig neue Datenerhebungen finden kann. Das Thema der Statistik lässt sich wiederum zum einen in die beschreibende Statistik und zum anderen in die beurteilende Statistik teilen. [10]

Das Wort *statisticum* bedeutet den Staat betreffend, was auch eigentlich die Aufgabe dieses Fachgebietes war. Erst seit dem letzten Jahrhundert veränderte sich das Gebiet und man widmete sich mehr der Darstellung und der Bearbeitung von Daten. [19] Gerade deswegen wird der Umgang mit Statistiken immer wichtiger. Deshalb sollte auch schon in der Schule ein klarer Einstieg vollzogen werden, damit die Schülerinnen und Schüler auch im späteren Alltag und Berufsleben damit umgehen können.

Denn immer wieder findet man, vor allem in den Medien, falsch verwendete Datenerhebungen, Übertreibungen oder unkorrekte Piktogramme. Im Fernsehen und in Zeitschriften werden derart falsche Angaben aber mitunter nicht ohne Grund verwendet: Sie werden manchmal gezielt eingesetzt, um das Publikum von einer Meinung zu überzeugen, um Angst zu erzeugen oder um eine Schlagzeile zu erhalten. Solche Angaben, die aus mathematischer Sicht falsch sind, werden absichtlich erzeugt. Zum Beispiel durch das Strecken oder Kürzen der Achsen einer Jahresbilanz.

Angenommen in der folgenden Abbildung 1 wird hypothetisch die Gehaltskurve einer jungen Frau angezeigt. Die Person fängt als Praktikantin, begleitend mit der Ausbildung, in einer Firma an und verdient im ersten Jahr 900€. Aufgrund der guten Leistungen bekommt sie eine Gehaltserhöhung und verdient im zweiten Jahr ihrer Laufbahn 1100€. Im dritten Jahr wird wegen einer neuen Regierungsregelung ein Plus von 400€ von ihrem Gehalt erzielt und sie verdient nun 1500€. Nach vier Jahren als Praktikantin schließt sie ihr Studium erfolgreich ab und bekommt ein Gehalt von 2100€. Sie wird schwanger und geht im nächsten Jahr in Karenz. Während ihrer Karenzpause beträgt ihr Einkommen 600€. Nach zwei Jahren steigt sie langsam wieder in die Berufswelt ein und arbeitet Teilzeit mit einem Gehalt von 1500€. Die dazugehörige Darstellung des Einkommens während ihrer Berufskarriere würde man statistisch entsprechend der Abbildung 1 visualisieren.

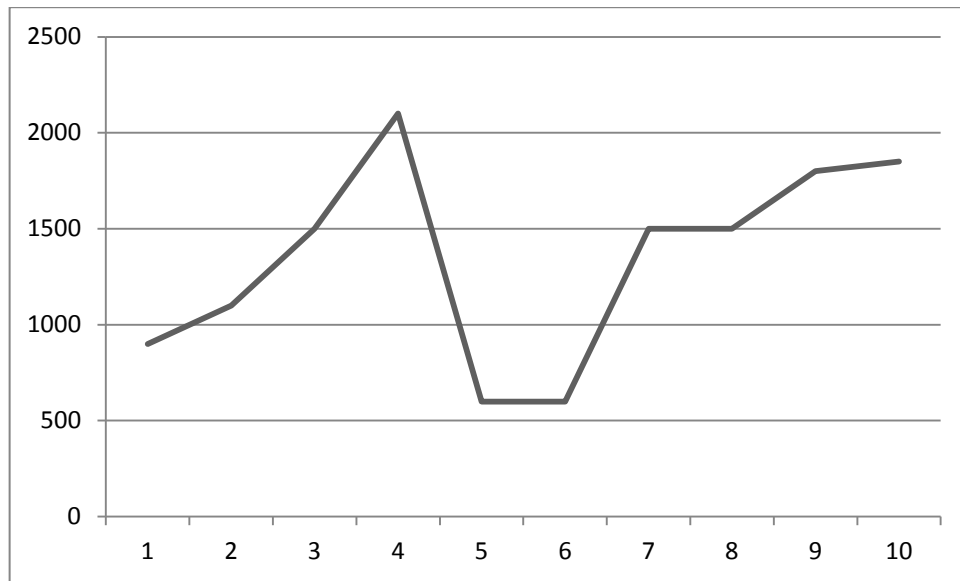


Abbildung 1: Statistik einer Gehaltskurve

Was mit dieser Grafik nun alles repräsentiert werden kann, zeigen folgende Abbildungen:

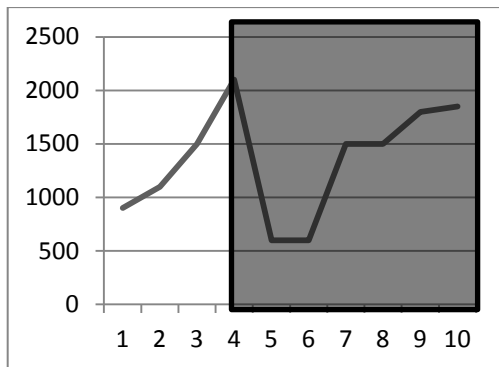


Abbildung 2: Manipulierte Statistik einer Gehaltskurve I

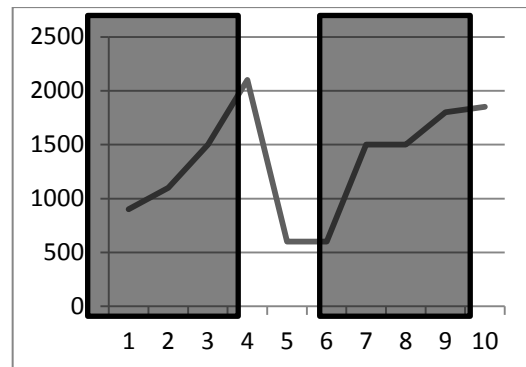


Abbildung 3: Manipulierte Statistik einer Gehaltskurve II

In der Abbildung 2 könnte man behaupten, wenn man nur den unschattierten Bereich veröffentlicht, dass das Gehalt kontinuierlich ansteigt. So hätte man einen positiven Eindruck von der gezeigten Statistik. Im Vergleich dazu ist in Abbildung 3 wieder derselbe Graph abgebildet, würde man die schattierten Bereiche wegschneiden beziehungsweise unveröffentlicht lassen, so hätte die Leserin oder der Leser das Gefühl, dass das Gehalt innerhalb kürzester Zeit sehr viel abgenommen hat.

Und genau diese Methoden, das Strecken beziehungsweise das Kürzen von Achsen wird sehr häufig von den Medien benutzt, um Schockmeldungen zu provozieren. [19] Das große Publikum glaubt oftmals diese Falschmeldungen, weil die meisten Menschen die Statistiken nicht weiter hinterfragen und den Medien durchwegs Glauben schenken. Dennoch sollte man sich immer fragen, wie viel Wahrheit in einem Graphen oder einer Statistik steckt.

Daten können in jeglicher Form erfasst und auch auf verschiedene Arten präsentiert werden. Am häufigsten werden Daten wohl in Form von Zahlen erfasst. Zum Beispiel will eine Lehrerin oder ein Lehrer die Noten einer Schularbeit erfassen und erstellt dazu eine

anschauliche Tabelle. Außerdem gibt es eine große Anzahl an Möglichkeiten, die erfassten Daten zu repräsentieren. Die wohl bekannteste Form ist das Kreis- oder Tortendiagramm. Neben dem Kreisdiagramm gibt es auch noch unter anderem das Säulen-, Balken- und Stabdiagramm. [19]

Beim Erheben von Daten zählt man gewisse Merkmale und deren Häufigkeiten. Als weitere Hilfestellung zur weiteren Verarbeitung und Deutung der erhobenen Kennzeichen werden die Merkmalsausprägungen in sogenannte Skalen übertragen. Je nach Merkmal gibt es verschiedene Zuordnungstypen. Zum einen gibt es nominal skalierte Merkmale. Ihre Charakteristik ist, dass es sich dabei nur um Namen handelt, wie zum Beispiel Berufe. Das Rechnen mit solchen Angaben ist nicht nur sinnlos, sondern auch nicht möglich. Man spricht von ordinal skalierten Merkmalen, wenn zwar eine Reihenfolge vorhanden ist, wie zum Beispiel bei Noten, aber die Abstände zwischen den Werten nicht gleichmäßig sind. Auch hier ist das Rechnen mit den Daten unlogisch. Zum anderen spricht man von metrisch skalierten Merkmalen. Bei diesem Fall der Zuordnung wird mit eindeutigen Zahlen gearbeitet. Diese Kategorie der Merkmalsausprägung teilt sich noch einerseits in die Intervallskala und andererseits die Verhältnisskala. Bei der ersteren ist es möglich, Summen beziehungsweise Differenzen zu erzeugen, und bei der zweiten Skala ist es sogar möglich, mit Quotienten zu rechnen. [19]

2. Geschichtliches zum Thema Stochastik

Der Ausgangspunkt zu stochastischen Überlegungen und Berechnungen der Wahrscheinlichkeit war wohl das Glücksspiel, welches es schon seit Menschengedenken gibt. Jedoch muss man sich bewusst machen, dass es früher in einer anderen Form abgelaufen ist als heutzutage. Denn die damaligen Mathematiker waren erst auf der Spur zu den Lösungen der Wahrscheinlichkeitsprobleme. Dazu kam noch der Faktor hinzu, dass die Wissenschaftler zu jener Zeit keine technischen Hilfsmittel zur Verfügung hatten, die bei der Entwicklung der Problematik unterstützen hätten können.

Man setzte sich schon sehr früh mit Wahrscheinlichkeitsbegriffen auseinander. Schon bei den Völkern der Inder, Babylonier und Ägypter findet man erste Gedanken und Überlegungen zu Wahrscheinlichkeiten. Um 3500 vor Christus verwendete man Astragali, welches dem heutigen Würfel vergleichbar ist. Es wurde aus Knochen der Schafe gewonnen. Vor allem die Römer kannten, ähnlich dem heutigen Würfelspiel, ein Spiel, wobei sie mit vier Astragalis, die von 1 bis 6 nummeriert waren, warfen. So entwickelte sich, ausgehend von einem Astragali der Würfel, wie man ihn heutzutage kennt. In seiner heutigen Form wurde er aber erst später entworfen. Die Erfindung des Würfels geht auf Palamedes zurück, der ihn während des trojanischen Krieges erfunden haben soll. [7, S. 56]

Im Zeitalter der Antike und der Renaissance hatte man einen anderen Begriff und eine andere Auffassung von Wahrscheinlichkeit. Zu jener Zeit war das Wort Zufall üblich. Man beschäftigte sich weniger mit den Gründen und Ergebnissen des Würfelns. Wahrscheinlich deshalb, weil zu dieser Zeit das Wort Gottes und das Schicksal als Grund für alle Ergebnisse und Ereignisse galt. Der damalige religiöse Glaube beeinflusste das Leben der Menschen sehr. So wurden Würfel verwendet, um Ereignisse vorherzusagen. [7]

Doch erst das Jahr 1654 zählt als entscheidende Wende der Wahrscheinlichkeitsrechnung. Denn zu dieser Zeit befassten sich Pierre de Fermat und Blaise Pascal in einem Briefwechsel mit dem Problem der Teilung und dem Würfelproblem von Chevalier de Méré. Seit diesem Zeitpunkt kann man von einem wissenschaftlichen Umgang mit Wahrscheinlichkeitsrechnung sprechen. [10, S. 7]

Aber nicht nur Fermat und Pascal beschäftigten sich mit den ersten Problemen der Wahrscheinlichkeitstheorie. Es zeigten auch weitere bekannte Naturwissenschaftler an diesen Themen ihr Interesse. Zum einen arbeitete Jakob Bernoulli an dem St. Petersburger Paradoxon. Zum anderen befasste sich auch der Mathematiker Simon Marquis de Laplace damit, der die Wahrscheinlichkeitstheorie mit seinem Werk *„Théorie Analytique des Probabilités“* bis heute entscheidend beeinflusste. [10, S. 7] Zu den Pionieren dieses Bereichs zählt auch der russische Mathematiker Alexander Kolmogorov, der allerdings erst im Jahr 1933 eine Definition der Wahrscheinlichkeit zu geben versuchte. [10, S. 7]

2.1 Die ersten Probleme der Stochastik

Schon früh entwickelten sich Probleme und Paradoxa zum Thema der Wahrscheinlichkeit. Viele Mathematiker beschäftigten sich mit solchen Problemstellungen und versuchten auf einen richtigen Lösungsweg zu kommen. Dieser Anreiz war ein erster Schritt zur Entwicklung einer eigenen Wahrscheinlichkeitsrechnung in der Mathematik. Und so gewann diese Theorie immer mehr an Gewicht und an Interesse. Nicht nur für Mathematikerinnen und Mathematiker selbst, sondern es entwickelte sich, ausgehend von diesen Fragestellungen, große Sympathie für den Zufall. In weiterer Folge entstand so auch ein enormer Zuwachs der Theorien für Glücksspiele und Wetten. In der obigen kurzen geschichtlichen Einführung wurden einige wenige genannt, die in dem folgenden Kapitel etwas ausführlicher beschrieben werden sollen.

2.1.1 Das Teilungsproblem und die Lösungsvorschläge von Pacioli, Pascal und Fermat

Das Teilungsproblem, mit dem sich unter anderem Fermat und Pascal beschäftigten, auch *problème des partis* genannt, ist deswegen für die Geschichte der Stochastik so wichtig, weil die Auseinandersetzung damit als der Ursprung der Wahrscheinlichkeitstheorie gilt. [24, S. 81] Behandelt wurde das Problem der Teilung bereits vor Fermat und Pascal im Jahre 1494. [9, S. 12] Jedoch wird den beiden Mathematikern die erste Lösung zugeschrieben. Ihre Überlegungen zu diesem Problem begründeten viele Theorien zur weiterführenden Entwicklung der Wahrscheinlichkeitsrechnung.

Die Aufgabenstellung lautet: Person A und Person B betreiben ein Spiel. Dabei werden sie aber durch eine höhere Macht, auch als *force majeure* bezeichnet, beim Stand $4 : 3$ unterbrochen. Die Spielregeln besagen, dass derjenige als Gewinner gilt, der zuerst fünf Spiele gewonnen hat. Nun stellt sich die Frage, wie der Spieleinsatz fair aufgeteilt werden soll. [10, S. 5]

Die ersten Lösungsvorschläge zu diesem Problem stammen von Luca Pacioli. Und zwar löste Pacioli die Frage nach dem Spieleinsatz so, dass er das Verhältnis des jeweiligen Gewinnes zuteilte. Das Verhältnis war ja $4:3$, daher solle Spieler A $\frac{4}{7}$ von seinem Einsatz zurückbekommen und, Spieler B würde demgemäß $\frac{3}{7}$ des Einsatzes erhalten. Jedoch, kritisierten viele Mathematiker zu dieser Zeit, dass Pacioli dabei nicht die fehlende letzte Runde berücksichtigte, da man den Gewinner ja erst nach fünf Spielen festlegen konnte.

Um nicht außer Acht zu lassen, warum in der Geschichte die Wahrscheinlichkeit so viel an Bedeutung dazugewonnen hat, folgt nun auch die normative klassische Lösung des Problems. Dieser Lösungsvorschlag stammt von Pascal und Fermat. Man nehme an, dass die beiden Spieler A und B gleich stark wären. Dann kann man mit Hilfe des Glücksspielmodells, laut Pascal und Fermat, annehmen, dass die Wahrscheinlichkeit von $P(A)$ für den Sieg von A $\frac{1}{2}$ beträgt. Nun, wie beim obigen Versuch von Pacioli, wird der Gewinn im Verhältnis der Gewinnmöglichkeiten berechnet.

$P(A) : P(B)$, bedeutet im Kontext des Beispiels: $P(A | 4:3) = \frac{1}{2} + (\frac{1}{2})^2 = \frac{3}{4}$. Das bedeutet Spieler A bekäme im Sinne dieses Verhältnisses $\frac{3}{4}$ des Einsatzes und Spieler B erhielte $\frac{1}{4}$ davon. [6]

2.1.2 Das Würfelproblem und der Lösungsvorschlag von de Méré

Eine andere Aufgabe, die bereits früher die Wissenschaftler beschäftigte, ist das Würfelproblem von Chevalier de Méré. Er behauptete, dass die Augensumme 11 beim Würfeln von drei Würfeln öfter vorkommt als die Augensumme 12. Und die Weiterführung des Problems: Hier überlegte de Méré, weshalb die Möglichkeiten, eine 6 bei vier Würfeln zu erhalten, wahrscheinlicher ist, als zwei Mal eine 6 bei 24 Würfeln zu erwürfeln. [10, S. 5–6]

Laut de Méré, auf den dieses Paradoxon des Würfelproblems zurückgeht, lautet die Lösung folgendermaßen:

$\Omega = \{abc \mid a, b, c \in \{1, 2, 3, \dots, 6\} \text{ und } a \leq b \leq c\}$. Nach der Zählung der Menge Ω von de Méré gibt es $\binom{6}{3} = 20$ Ergebnisse mit drei unterschiedlichen Ziffern, $2 \cdot \binom{6}{2} = 30$ für je zwei verschiedene Ziffern und sechs mögliche Ereignisse mit drei gleichen Ziffern. Daher beträgt die Kardinalität der Menge Ω , $|\Omega| = 56$. Das Ereignis der Augensumme 11 erhält man nun in folgenden sechs Fällen:

$$E_1 = \{1 + 4 + 6, 1 + 5 + 5, 2 + 3 + 6, 2 + 4 + 5, 3 + 3 + 5, 3 + 4 + 4\}.$$

Hingegen erhält man die Augensumme 12 in den Ziffernkombinationen der folgenden Ereignisse:

$$E_2 = \{1 + 5 + 6, 2 + 4 + 6, 2 + 5 + 5, 3 + 3 + 6, 3 + 4 + 5, 4 + 4 + 4\}.$$

Sowohl das Ereignis E_1 als auch E_2 hat sechs günstige Varianten. Aus dieser Überlegung heraus nimmt de Méré mit der Laplace-Wahrscheinlichkeit $\frac{|\text{günstigen}|}{|\text{möglichen}|}$ an:

$$P(E_1) = P(E_2) = \frac{6}{56} \approx 0,107.$$

Jedoch wusste der Mathematiker de Méré aus seinen eigenen Spielerfahrungen, dass das Ereignis mit der Augensumme 11 öfters fällt. Er nahm irrtümlicherweise an, dass die beiden Fälle E_1 und E_2 gleichwahrscheinlich seien, diese Überlegung folgt ja aus der Laplace-Wahrscheinlichkeit. Jedoch ist bei diesem Beispiel keine gleiche Wahrscheinlichkeit gegeben, denn es sind drei Würfel gegeben, das bedeutet, es gibt $6^3 = 216$ mögliche Ergebnisse der Würfel. Erst wenn man diesen Gedanken berücksichtigt, kann man mit der Laplace-Wahrscheinlichkeit rechnen. Dann lautet die Rechnung wie folgt:

Augensumme 11:

$$E_1 = \{1 + 4 + 6, 1 + 6 + 4, 4 + 6 + 1, 4 + 1 + 6, 6 + 1 + 4, 6 + 4 + 1, \\ 1 + 5 + 5, \dots\}$$

$\Rightarrow |E_1| = 27$, daher beträgt die Wahrscheinlichkeit $P(E_1) = \frac{27}{216} = 0,125$. [10, S. 422]

Augensumme 12:

$$E_2 = \{1 + 5 + 6, 1 + 6 + 5, 5 + 1 + 6, 5 + 6 + 1, 6 + 1 + 5, 6 + 5 + 1, \dots\}$$

$\Rightarrow |E_2| = 25$, deshalb ist hier die Wahrscheinlichkeit für $P(E_2) = \frac{25}{216} = 0,116$. [10, S. 422]

2.1.3 Das St. Petersburger Paradoxon und der Lösungsvorschlag von Daniel Bernoulli

Bei dieser Aufgabe wird eine Münze so oft geworfen bis eine Zahl erscheint. Kommt die Zahl gleich beim ersten Mal, so erhält man einen Euro. Zeigt sich die Zahl erst beim zweiten Wurf, bekommt man zwei Euro. Sollte man die Zahl erst beim dritten Wurf sehen, bekommt man vier Euro, und so weiter. Also je öfter die Münze geworfen werden muss, um die Zahl das erste Mal zu bekommen, desto höher wird der Gewinn, da er sich stetig verdoppelt. Dieses Paradoxon soll von Nicolas Bernoulli stammen. Jedoch hat sich auch Daniel Bernoulli mit diesem Problem beschäftigt. [24, S. 120]

Die Gewinnchancen dieses Spieles werden aus [5] anhand der Tabelle 1 adaptiert und sollen die Möglichkeiten eines Profites veranschaulichen.

Tabelle 1: Gewinnchancen des St. Petersburger Paradoxons [5, S. 58]

Gewinn in Euro	Wahrscheinlichkeit in Prozent
2	50
4	25
8	12,5
16	6,25
32	3,125
...	...

An diesem Paradoxon ist vor allem der Erwartungswert auf lange Sicht interessant. [5, S. 58] Wenn man das nun anhand dieses Beispiels näher betrachtet, erkennt man, dass man bei einer Gewinnchance von 1 : 2 einen Gewinn von 2€ erlangt. Dies bedeutet im Durchschnitt $2 \cdot \frac{1}{2} \text{€} = 1\text{€}$. Bei weiteren Versuchen mit 4€ beträgt die Möglichkeit 1 : 4 und man berechnet $\frac{1}{4} \cdot 4\text{€} = 1\text{€}$. Bei 8€ beschreibt man das Verhältnis 1 : 8, bei der nächsten Angabe von 16€ hat man 1 : 16 und so weiter. Jeder Euro-Wert ergibt im Durchschnitt genau 1€. Daher bekommt man, wenn man alle Werte zusammenrechnet, den Ausdruck: $1 + 1 + 1 + \dots$. Diese Summe führt zu einem unendlich großen Wert. [5, S. 61]

Nicolas Bernoulli übermittelte seine Überlegungen zum einen an den Mathematiker Gabriel Cramer und zum anderen an seinen Bruder Daniel Bernoulli, damit die Wissenschaftler ihre Gedanken zu diesem Problem austauschen können. Cramer löste die Problematik äußerst pragmatisch und meinte, dass es nur begrenzt viel Geld auf der Welt gebe. Daher hat sich das Problem des Paradoxons in seiner Sicht aufgelöst. [5, S. 63–64]

Hingegen entwickelte Daniel Bernoulli den Widerspruch weiter und gelangte zu einer anderen Lösung als sein Kollege. Und zwar führte er den unter dem heutigen Namen bekannten Grenznutzen des Geldes ein. [5, S. 64] Dieser beschreibt den Gebrauch des Geldes

je nach eigenem Besitz und persönlichem Kapital. Sprich für einen Obdachlosen symbolisiert zum Beispiel ein Defizit von 10€ etwas anderes als für einen wohlhabenden Menschen. [5, S. 64] Die Funktion, die diesen Nutzen entsprechend beschreibt, nennt man auch Logarithmus. [5, S. 64–65] Wie in Tabelle 2 erkennbar ist, erhöht sich der Nutzen nicht so rasch wie der Geldbetrag. Dies folgert im Gegensatz zur Tabelle 1 einen endlichen Erwartungswert, der ungefähr 0,60 beträgt.

Tabelle 2: Logarithmusfunktion des St. Petersburger Paradoxons [5, S. 65]

Nutzen	Wahrscheinlichkeit in Prozent
0,30103	50
0,60206	25
0,90309	12,5
1,20412	6,25
1,50515	3,125
...	...

Es ist ein relativ bedeutsamer Unterschied zwischen den beiden Lösungsversuchen von Cramer und Bernoulli zu sehen. Genau diese Varietät an Lösungsvorschlägen macht das Problem eines Paradoxons aus. Denn vor allem soll ein mathematisches Paradoxon zum Denken anregen und zu verschiedenen Überlegungen führen. Dabei sind der Fantasie und den mathematischen Strategien keine Grenzen gesetzt. Es gibt viele Mathematikerinnen und Mathematiker, die an einigen Widersprüchen der Mathematik scheitern. Dennoch repräsentieren solche komplexe Problematiken einfache Darstellungen, die jedoch durch falsches Denken zu Irreführungen führen können. Die Namensgebung dieses Paradoxons ergibt sich aus dem Umstand, dass Daniel Bernoulli seinen Lösungsversuch als einen Beitrag zu den *Kommentaren der Kaiserlichen Akademie zu St. Petersburg* publizierte. [5, S. 65]

3. Ein Einblick in die beschreibende Statistik

Auch wenn das Thema der beschreibenden Statistik nicht den Hauptteil dieser Arbeit repräsentiert, so werden dennoch fundamentale Begriffe dieses Kapitels für die weiteren Überlegungen dieser Arbeit an dieser Stelle explizit genannt. Dieser Teil soll vor allem die möglichen Hilfestellungen der Statistik für die Wahrscheinlichkeitsrechnung kurz beschreiben. Insbesondere damit gemeint sind absolute und relative Häufigkeiten. Wie der Name Häufigkeit schon verrät, hat der folgende Abschnitt mit einer Anhäufung von mehreren ausgeführten Experimenten zu tun.

Wie verhält sich der Ausgang einer Versuchsserie auf lange Sicht? Gibt es einen Versuchsausgang, der häufiger auftritt oder sind die Ereignisse ungefähr gleich wahrscheinlich? Es sind Fragen wie diese, die mit Hilfe von relativen Häufigkeiten beantwortet werden können. Auch der Zufall spielt hierfür eine wichtige Rolle. Im Gegensatz zum Alltagsgebrauch hat der Begriff des Zufalls in der Mathematik mit Gesetzmäßigkeiten zu tun. Denn, auch wenn man es anders gewohnt ist, der Zufall spielt in der Wahrscheinlichkeitstheorie nach gewissen Regeln. Wenn man einen Versuch sehr häufig unter denselben Bedingungen durchführt, erkennt man an den Ausgängen des Experimentes eine Annahme. Dadurch ist es möglich, Schlussfolgerungen aus den Erfahrungen zu ziehen.

3.1 Relative und absolute Häufigkeiten

Da sich die relativen Häufigkeiten bei der wiederholten Durchführung desselben Zufallsexperimentes stabilisieren, ist es möglich, eine Schätzung der Wahrscheinlichkeit zu machen. [7, S. 103–104] Denn ohne Zweifel besteht zwischen relativen Häufigkeiten und Wahrscheinlichkeit eine enge Beziehung. Um genauer zu sein, lassen die relativen Häufigkeiten einen Schätzwert für die Wahrscheinlichkeiten zu. Aus der Perspektive der Wahrscheinlichkeit geben die Wahrscheinlichkeitswerte Auskunft über die relativen Häufigkeiten.

Nun muss man unter den Begrifflichkeiten absolute und relative Häufigkeiten unterscheiden. Die absoluten Häufigkeiten entsprechen der möglichen Menge des auftretenden Ereignisses. Man schreibt für absolute Häufigkeiten folgende Begriffserklärung.

Definition: Für jede Ausprägung x_i wird die Anzahl der Merkmalsträger mit dieser Ausprägung mit $H_n(x_i)$ bezeichnet und absolute Häufigkeit von x_i bei n untersuchten Merkmalsträgern genannt. [10, S. 25]

Hingegen geben die relativen Häufigkeiten den Anteil an der gesamten Ergebnismenge an und man schreibt:

Definition: Der Anteil eines Merkmals wird definiert durch

$$h_n(x_i) = \frac{H_n(x_i)}{n}$$

und relative Häufigkeit von x_i bei n untersuchten Merkmalsträgern genannt. [10, S. 25]

Die weiteren Überlegungen verleiten schnell zu einem Irrtum. Denn man kann nun behaupten, dass man die Wahrscheinlichkeit einfach mit Hilfe des Grenzwertsatzes definieren kann, da bei großem n sich die Versuchsserie nach einer gewissen Zeit einpendelt und bei einem gewissen Wert ungefähr konstant bleibt. Einen ähnlichen Gedanken hatte auch Richard von Mises. Er versuchte mit einem ähnlichen Gedanken die Wahrscheinlichkeit mit Hilfe des Grenzwertes zu definieren.

Definition nach Mises: Die Wahrscheinlichkeit wird durch den Limes bezeichnet als:

$$P(X) = \lim_{n \rightarrow \infty} h_n(x). [19]$$

Zwar gilt diese Definition der Wahrscheinlichkeit als nicht zufriedenstellend, da Mises wichtige Anforderungen, die seine Definition der Wahrscheinlichkeit nicht erfüllt, außer Acht ließ. Zum einen muss ein Versuch tatsächlich unendlich oft durchführbar sein und zum anderen missachtet er dabei die Konvergenz. Denn ohne diese Anforderung und nur mit der Definition relativer Häufigkeiten wäre die Folgerung einer bedingten Wahrscheinlichkeit nur schwer möglich. [7, S. 112] Dennoch war dieser Versuch einer ersten möglichen Begriffsbestimmung äußerst inspirierend für die damalige Zeit und verursachte weitere Überlegungen zur Wahrscheinlichkeitstheorie.

3.2 Bernoullisches schwaches Gesetz der großen Zahlen

In diesem Unterkapitel soll, wie schon vorhin erwähnt, die Konvergenz der Wahrscheinlichkeitsrechnung besprochen werden. Denn mit Hilfe der relativen Häufigkeiten können Konvergenzsätze gefolgert werden. Dafür wird an dieser Stelle zuerst die Tschebyscheffsche Ungleichung formuliert.

Satz: Ist x eine eindimensionale Zufallsvariable, deren Verteilung Momente bis zur zweiten Ordnung besitzt, dann gilt für jedes a und für $b > 0$:

$$P(|x - a| \geq b) \leq \frac{E((x - a)^2)}{b^2}.$$

$$\text{Insbesondere ist für } a = E(x) = \mu: P(|x - \mu| \geq b) \leq \frac{\sigma^2}{b^2}. [17, S. 216]$$

Beweis für die stetigen Verteilungen:

$$\begin{aligned} E((x - a)^2) &= \\ &= \int_{\mathbb{R}} (x - a)^2 f(x) dx \geq \int_{x: |x-a| \geq b} (x - a)^2 f(x) dx \geq b^2 \int_{x: |x-a| \geq b} f(x) dx = b^2 P(|x - a| \geq b). \end{aligned}$$

Die Behauptung folgt durch bloße Division durch b^2 . [17, S. 216]

Wenn man nun ein Experiment E mit dem Ereignis A und der Wahrscheinlichkeit $P(A) = p$ definiert, dann zählt $z_n = H_n(A)$, wie oft das Ereignis A bei n Versuchen eintritt. Wie schon erwähnt, erhält man nun die relativen Häufigkeiten durch $h_n(A) = \frac{z_n}{n}$.

Aufgrund von $\mu_{z_n} = n \cdot p$ und $\sigma_{z_n}^2 = n \cdot p \cdot (1 - p)$ gilt $\mu_{y_n} = p$ und $\sigma_{y_n}^2 = \frac{p(1-p)}{n}$. Arbeitet man nun mit der Tschebyscheffschen Ungleichung auf $h_n(A)$, kommt man zu dem Bernoullischen Gesetz der großen Zahlen. In diesem Fall ist allerdings das Schwache Bernoullische Gesetz von Interesse.

Bernoullisches schwaches Gesetz der großen Zahlen: Für eine Zufallsvariable X sollen der Erwartungswert $E(X)$ und die Varianz $V(X) = \sigma^2 := \sigma^2(X)$ existieren. Wenn $X_1, X_2, \dots, X_n$ unabhängige Kopien von X und $t > 0$, dann gilt

$$P\left(\left\{\omega \mid \left|\frac{X_1 + X_2 + \dots + X_n}{n} - E(X)\right| \geq t\right\}\right) \leq \frac{\sigma^2}{nt^2}. [3, S. 173]$$

Wenn man dieses Gesetz nun verständlich formulieren möchte, bedeutet es konkret, dass die Wahrscheinlichkeit für den Mittelwert von $X_1, X_2, \dots, X_n$ um mindestens t vom Erwartungswert schwankt. Sie beträgt höchstens den Wert von $\frac{\sigma^2}{nt^2}$. [3, S. 173]

4. Der Weg zu den ersten Wahrscheinlichkeitsbegriffen

Die intuitive Annahme spielt in der Wahrscheinlichkeitstheorie eine große Rolle. Oft lässt man sich durch die Intuition leiten und nimmt vorschnell eine Annahme an, ohne einen angemessenen Grund dafür zu haben. Wird man vom Bauchgefühl zu sehr beeinflusst, so kann es zu Fehlschlüssen kommen und man wird mit dem Ergebnis dann oftmals überrascht. Deshalb ist es essentiell, in der Stochastik nicht immer seinem Bauchgefühl zu vertrauen, sondern deutlichen Regeln und Strukturen, die vor allem von den Theorien der Wahrscheinlichkeit angeboten werden, in Anspruch zu nehmen.

Um nun zu erklären, was man unter Wahrscheinlichkeit versteht und wie man sie berechnen kann, folgen einige wichtige Terminologien, die für das Thema der Stochastik unerlässlich sind. Da die Wahrscheinlichkeitsrechnung viel mehr ist als nur die oft damit verbundene Spieltheorie, werden im folgenden Kapitel einige Grundbegriffe erläutert und definiert. Dies erleichtert zugleich das Verständnis der folgenden Kapitel. Das wohl wichtigste Mittel zum Berechnen von Wahrscheinlichkeiten sind Experimente. Solche Versuche müssen aber nicht nur klar und strukturiert definiert werden, sondern auch auf eine angemessene Weise durchgeführt werden.

4.1 Was ist ein (Zufalls-)Experiment? Ein Einblick in die Schule

Der Begriff des Experiments begleitet eine Schülerin oder einen Schüler schon ab der frühesten Schulzeit. Immer öfter werden im Unterricht Experimente durchgeführt, vor allem in den naturwissenschaftlichen Fächern. Doch viele Jugendliche, aber auch einige Erwachsene können mit dem Wort Experiment nicht viel anfangen.

Für Experimente im Mathematikunterricht gibt es ebenso einen großen Spielraum. Jedoch wird dieser nur wenig genutzt, um den Schülerinnen und Schülern ein Problem anschaulicher zu machen und näher zu bringen. Vor allem das Kapitel der Wahrscheinlichkeitsrechnung bietet ein solches Angebot mit Experimenten in das Thema zu starten. Denn gerade in der Wahrscheinlichkeit behandelt man eben Ergebnisse und mögliche Ereignisse, die anhand von Experimenten dargestellt werden können. Jedoch wird in der Institution Schule mit dem Begriff Experiment beziehungsweise Zufallsexperiment falsch umgegangen, da der Begriff Experiment in der Schule eine etwas andere Bedeutung hat als jener eines Zufallsexperimentes in der Wahrscheinlichkeitstheorie. Oft geht man von klassischen Experimenten der Physik aus. Deren Sinn und Zweck ist jedoch ein anderer. Und zwar will man hierbei versuchen, auf Grund der Experimente, die im Normalfall auf die gleichen beziehungsweise ähnlichen Ergebnisse kommen, eine These zu bestärken und daraus eine Regel abzuleiten. Anders in der Stochastik. Im Bereich der Stochastik ist der Sinn eines Zufallsexperimentes ein anderer, wie man im folgenden Kapitel erkennen wird.

Im Hinblick auf die Schule sollte man darauf explizit eingehen und den Schülerinnen und Schülern ausdrücklich den Unterschied erklären. Man sollte den Jugendlichen die Definition eines Experimentes Schritt für Schritt näher bringen und sie selbst zur Erklärung hinführen.

Für den Anfang reichen soweit auch naive Definitionsversuche der Schülerinnen und Schüler, die man als Lehrkraft mit Hilfe der Vorschläge aus der Klasse zu einer vollständigen mathematisch korrekten Definition ergänzt.

4.2 Zufallsexperimente und Mengen aus mathematischer Sicht

Auch wenn es offensichtlich zu sein scheint, was ein Zufallsexperiment ist und welche Ansprüche dieses fordert, so ist es an dieser Stelle trotzdem wichtig, eine Definition dafür zu geben.

Definition: Ein Experiment heißt Zufallsexperiment, wenn

1. jeder Versuch unter denselben Bedingungen durchgeführt wird,
2. alle möglichen Ergebnisse bekannt sind und
3. es theoretisch möglich ist das Experiment unendlich oft unter denselben Bedingungen durchzuführen. [18, S. 3]

Daher sollte an dieser Stelle bewusst gemacht werden, dass die Stochastik eine wichtige Beziehung zur Mengenlehre hat, da die möglichen Varianten eines Ergebnisses in einer Menge, die mit Ω bezeichnet wird, angegeben werden können. Dabei gibt es verschiedene Arten einer Menge. Entweder man spricht davon, dass die Menge Ω überabzählbar ist. Oder es handelt sich dabei um eine endliche Menge. Dies ist genau dann der Fall, wenn, wie der Name schon verrät, es sich um eine Menge handelt, die bei einem bestimmten Wert endet. Typisches Beispiel dafür ist das Experiment mit einem normalen Würfel. Hier lautet die Ergebnismenge $\Omega = \{1, 2, 3, 4, 5, 6\}$. Oder es handelt sich bei der Ergebnismenge um eine abzählbar-unendliche Menge. Eine solche Menge hat den gesamten Zahlenbereich der natürlichen Zahlen als mögliche Ereignisangabe. Ein Beispiel dafür wäre das Spiel *Mensch ärgere dich nicht*. Wobei hier nicht die Würfelmöglichkeiten mit den Varianten 1, ..., 6 zählen, sondern das Zufallsexperiment, wie oft man würfeln muss, um eine 6 zum ersten Mal zu erhalten. In diesem Fall lautet $\Omega = \{1, 2, 3, \dots\}$, was dem gesamten Bereich der natürlichen Zahlen entspricht. [18, S. 3–5] Solche Zuteilungen, die einer entweder endlichen oder einer abzählbar-unendlichen Menge, werden diskrete Ergebnismengen genannt. [10, S. 136]

Ein wichtiges Teilkapitel der Stochastik ist also die Mengenlehre, weil mögliche Ereignisse eines Zufallsexperimentes mit Teilmengen angegeben werden. Man kann sagen, dass die Mengenlehre für die Stochastik eine große Hilfe ist, um den Umgang mit den Ergebnissen und die Vorstellung von den möglichen Varianten deutlicher zu machen. Es gibt dafür verschiedene Begriffe, die man in der folgenden Abbildung als diskrete Zufallsexperimente bezeichnen kann.

Tabelle 3: Ereignisse E und deren Erklärung [10, S. 137]

Begriff	Definition
Ergebnismenge	Die Menge Ω aller möglichen Ergebnisse des Zufallsexperiments (Abbildung 4)
Ereignis	Teilmenge $E \subseteq \Omega$ (Abbildung 5)
Elementarereignis	Teilmenge mit einem Element $\{\omega\} \subseteq \Omega$
Sicheres Ereignis	$E = \Omega$ (Abbildung 4)
Unmögliches Ereignis	$E = \emptyset$
Unvereinbare Ereignisse	$E_1, E_2 \subseteq \Omega$, aber $E_1 \cap E_2 = \emptyset$ (Abbildung 6)
Gegenereignis, Komplement von E	$\bar{E} = \Omega \setminus E$ (Abbildung 7)
Und-Ereignis der Ereignisse E_1 und E_2	$E_1 \cap E_2$ (Abbildung 8)
Oder-Ereignis der Ereignisse E_1 und E_2	$E_1 \cup E_2$ (Abbildung 9)

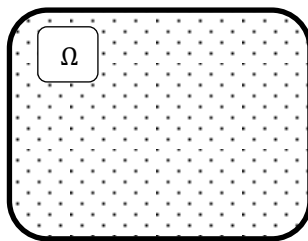
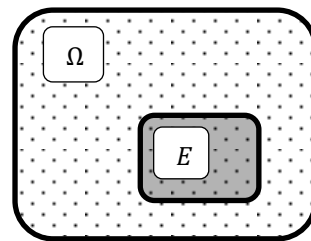
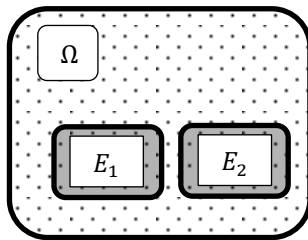

Abbildung 4: Menge Ω [18, S. 9]

Abbildung 5: Teilmenge E [18, S. 9]


Abbildung 6: Unvereinbare Ereignisse [18, S. 9]

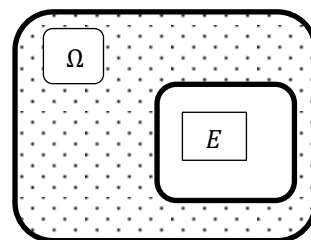


Abbildung 7: Gegenereignis [18, S. 9]

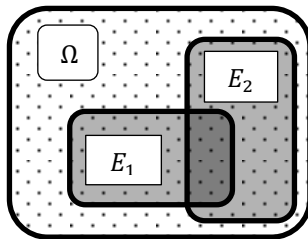


Abbildung 8: Und-Ereignis [18, S. 9]

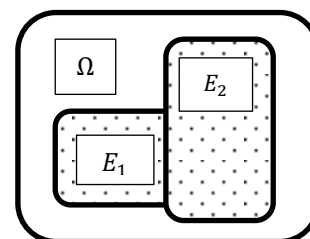


Abbildung 9: Oder-Ereignis [18, S. 9]

4.3 Laplace und die Wahrscheinlichkeit

Der Franzose Pierre Simon Laplace war einer derjenigen, durch die die Theorien der Wahrscheinlichkeit einen großen Fortschritt erzielt haben. Zwar gelang auch ihm keine Definition der Wahrscheinlichkeit, nichts desto trotz war sein 1812 veröffentlichtes Werk *Théorie Analytique des Probabilités* ein großer Gewinn für die damaligen Naturwissenschaften. [10, S. 139]

Definition: Sei Ω eine endliche Menge und alle möglichen Einzelwahrscheinlichkeiten können als gleich wahrscheinlich angesehen werden, dann kann man die Wahrscheinlichkeit berechnen durch

$$P(E) = \frac{|E|}{|\Omega|} = \frac{\text{Anzahl der Elemente von } E}{\text{Anzahl der Elemente von } \Omega} = \frac{|\text{günstige}|}{|\text{mögliche}|} \quad [10, \text{S. 139}]$$

Denn das Laplace-Modell zur Berechnung von Wahrscheinlichkeiten ist nur dann möglich, wenn unter jeder Einzelwahrscheinlichkeit Gleichwahrscheinlichkeit vorausgesetzt werden darf. Ist keine gleiche Wahrscheinlichkeit unter den einzelnen Ereignissen gegeben, so darf man auch nicht mit der Laplace-Methode rechnen. Schließlich bietet die Laplace-Wahrscheinlichkeit zwar einen praktischen Nutzen in der gesamten Wahrscheinlichkeitsrechnung. Trotzdem zählt diese Theorie von Laplace nur als ein Spezialfall in der Stochastik. Jedoch gilt die Theorie von Laplace als essentiell und enorm wichtig. Laplace hat dazu auch wichtige Eigenschaften formuliert:

Satz: Es sei ein Laplace-Experiment, welches die endliche Ergebnismenge $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ besitzt. So gilt für die Laplace-Wahrscheinlichkeiten der Ereignisse:

1. $0 \leq P(E) \leq 1 \quad \forall E \subseteq \Omega$.
(Nichtnegativität) [10, S. 140]

2. $P(\Omega) = 1; P(\emptyset) = 0$.
(Normiertheit) [10, S. 140]

3. $P(\bar{E}) = 1 - P(E) \quad \forall E \subseteq \Omega$.
[10, S. 140]

4. $P(E_1 \cup E_2) = P(E_1) + P(E_2) \quad \forall$
 $E_1, E_2 \subseteq \Omega$ mit $E_1 \cap E_2 = \emptyset$.
(Additivität) [10, S. 140]

5. $P(E_1 \cup E_2) =$
 $P(E_1) + P(E_2) - P(E_1 \cap E_2) \quad \forall$
 $E_1, E_2 \subseteq \Omega$. [10, S. 140]

6. $P(E) = \sum_{\omega \in E} P(\{\omega\})$ und
 $P(\Omega) = \sum_{\omega \in \Omega} P(\{\omega\}) = \sum_{i=1}^n P(\{\omega_i\}) = 1$.
[10, S. 140]

Ausgehend von Symmetrieeigenschaften der Zufallsexperimente erklärte Laplace seinen Ansatz zur Berechnung der Wahrscheinlichkeit. Man nehme einen herkömmlichen Würfel. Man zweifelt nicht daran, dass die Augensumme 1 gleich oft erscheint wie die Augensumme 2, da alle 6 Würfelaugen mit der gleichen Wahrscheinlichkeit erscheinen können. Diese Voraussetzung, dass alle möglichen Ausgänge eines Experimentes, gleichwahrscheinlich sein müssen, stellt gleichzeitig auch das Hindernis dieser Ansicht des Wahrscheinlichkeitsbegriffes dar. So sind Wahrscheinlichkeiten, deren Ereignisse nicht mit derselben Erwartung anzunehmen sind, aus dem Laplace-Modell ausgeschlossen.

4.4 Subjektive Wahrscheinlichkeit

Man kann sich selbst sehr schnell davon überzeugen, dass die Wahrscheinlichkeitstheorie auch immer mit der subjektiven Einstellung in Verbindung steht. Aussagen wie zum Beispiel „Morgen schneit es“ oder „Ich komme morgen um 16:00 Uhr nach Hause“, sind bloße Vermutungen über die Zukunft. Und deshalb besitzen sie auch einen gewissen Grad an Ungewissheit und Unsicherheit. Die menschliche Intuition hat jedoch nicht immer das

gewisse Gespür im Vergleich zu den mathematischen Grundgesetzen. Aus diesem Grund kommt es vor allem in der subjektiven Wahrscheinlichkeit zu Täuschungen.

Jedoch ist gerade in der Schule der intuitive Einstieg in die Wahrscheinlichkeit sehr hilfreich. Ausgehend aus einfachen und aus dem Alltag bekannten Sachverhalten der Wahrscheinlichkeit ist es durchaus erwünscht die Schülerinnen und Schüler schrittweise auf explizite Axiome und Grundgesetze hinzuweisen. Auch Fischbein ist der Meinung, dass ein Ausdruck mathematischer Theorien nur zugänglich gemacht werden kann, wenn dieser über die eigenen menschlichen Empfindungen nachvollzogen werden kann. [7, S. 7]

Man muss sich aber im Hinblick auf das Kapitel der Wahrscheinlichkeit bewusst machen, dass dieses nicht so leicht zugänglich ist wie andere Kapitel in der Mathematik. Zu komplex und ungenau scheinen die Probleme und Hypothesen der Wahrscheinlichkeit zu sein. Dieses Problem der Unsicherheit führt dazu, dass man nicht immer auf Vermutungen vertrauen kann. Anders als zum Beispiel in den Zahlenbereichen, wo man eindeutig mit den Zahlen operieren kann und so einen konkreten Umgang in der Mathematik pflegen kann, bietet die Wahrscheinlichkeitsrechnung kein solches Fundament. [7, S. 38]

Auch wenn man sich nicht immer auf die eigene Intuition verlassen sollte, so kann man trotzdem auf Vermutungen aufbauen. Laut Borovcnik unterscheidet man dazu primäre und sekundäre Intuitionen. Erstere sollen bei einem Zufallsexperiment die möglichen Ergebnisse vorhersagen. Zum Beispiel betragen bei einem Münzwurf die Wahrscheinlichkeiten $P(\text{Zahl}) = P(\text{Kopf}) = \frac{1}{2}$. [7, S. 39] Der nächste Schritt, im Bezug auf die sekundäre Intuition, liegt darin einen Zusammenhang mit dem Gesetz der großen Zahlen herzustellen. [7, S. 42] Wie bereits in einem vorigen Kapitel beschrieben wurde, arbeitet die Wahrscheinlichkeit auch mit Konvergenzsätzen. Diese beschreiben das Verhalten von wiederholten Zufallsexperimenten. Sie zeigen, dass sich die Ergebnisse der Wahrscheinlichkeit, zum Beispiel bei dem zuvor angenommenen Münzwurf, bei häufiger Ausführung desselben Versuches bei einem gewissen Wert einpendeln. In dem Fall des Münzwurfes wird der Wert bei ungefähr $\frac{1}{2}$ liegen, weil beide Ergebnisausgänge mit derselben Wahrscheinlichkeit zu erwarten sind.

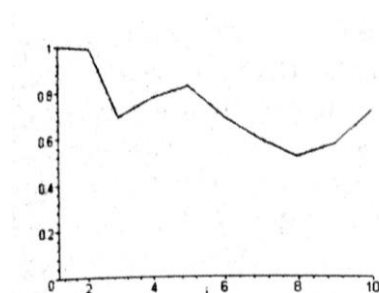


Abbildung 10: Relative Häufigkeiten eines Münzwurfs bei 10 Würfungen [10, S. 145]

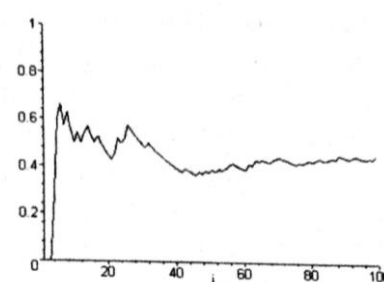


Abbildung 11: Relative Häufigkeiten eines Münzwurfs bei 100 Würfungen [10, S. 145]

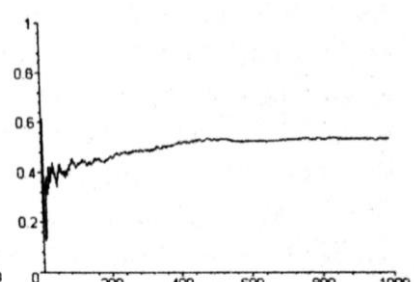


Abbildung 12: Relative Häufigkeiten eines Münzwurfs bei 1000 Würfungen [10, S. 145]

Beispielhaft dafür zeigen die drei oben gezeigten Abbildungen, wie sich die relativen Häufigkeiten und damit auch die Wahrscheinlichkeiten bei einem Münzwurf verändern. Hier wird allerdings nicht das zuvor besprochene Bernoulli'sche schwache Gesetz der großen Zahlen sondern das starke Gesetz der großen Zahlen visualisiert. Man sieht in der

Darstellung einen eindeutigen Unterschied. In Abbildung 10 erkennt man noch keine explizite Vermutung, dass die relativen Häufigkeiten bei einem gewissen Wert stabilisieren werden. Man könnte sich, im Hinblick auf zehn Wurfversuche, noch nicht festlegen, bei welchem Wert die relativen Häufigkeiten einen ungefähren Betrag der Wahrscheinlichkeit aufweisen werden. In Abbildung 11 erkennt man, dass sich ein Zufallsexperiment bei mehreren Ausführungen, sprich in diesem Zusammenhang bei 100 Würfeln, schon einem Zielwert annähern wird und man vermuten könnte, dass der Wert der relativen Häufigkeiten bei ungefähr $\frac{1}{2}$ liegen würde. Logischerweise bestätigt sich dieser Verdacht bei der Abbildung 12. Man nimmt wahr, dass sich die Vermutung vom Wert der relativen Häufigkeiten bei $\frac{1}{2}$ verifiziert hat.

Dieses Phänomen repräsentiert nichts anderes als das Gesetz der großen Zahlen, welches ja besagt, dass sich die relativen Häufigkeiten bei häufiger Durchführung des Zufallsexperimentes bei einem gewissen Ergebnis bleiben. Es ist wohl auch sehr leicht nachvollziehbar. Denn gerade bei dem Beispiel des Münzwurfes weiß man aus Erfahrung, dass sich die beiden möglichen Gegebenheiten $P(Kopf)$ beziehungsweise $P(Zahl)$ jeweils mit der gleichen Wahrscheinlichkeit ergeben können. Somit ist auch das Einpendeln der relativen Häufigkeiten eine Tatsache, die sich aus ihren Eigenschaften ergibt.

4.5 σ -Algebra

Als moderne Anschauung des Wahrscheinlichkeitsbegriffs gilt die σ -Algebra. Da es in der Wahrscheinlichkeitstheorie nicht immer möglich ist jedem Ausgang $A \in \Omega$ eines Zufallsexperimentes die entsprechende Wahrscheinlichkeit $P(A)$ zuzuverteilen, wird die Hilfe eines Systems $\mathcal{A}$ von den Ereignissen aus Ω zur Verfügung gestellt. [25, S. 3] Falls das Ereignis A geschieht, kann man somit die entsprechende Wahrscheinlichkeit formulieren.

Definition: Ein System $\mathcal{A}$ von Teilmengen aus Ω heißt σ -Algebra, falls

1. $\emptyset \in \mathcal{A}$,
2. mit $A \in \mathcal{A}$ folgt $A^c \in \mathcal{A}$ und
3. für abzählbar unendlich viele $A_1, A_2, \dots \in \mathcal{A}$ gilt $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$. [25, S. 3]

Satz: Sei $\mathcal{A}$ eine σ -Algebra von Teilmengen aus Ω . Dann gilt:

1. es gilt $\Omega \in \mathcal{A}$,
2. für endlich viele Mengen $A_1, A_2, \dots, A_n$ aus $\mathcal{A}$ folgt $\bigcup_{j=1}^n A_j \in \mathcal{A}$,
3. gehören abzählbar unendlich viele $A_1, A_2, \dots$ zu $\mathcal{A}$, so auch $\bigcap_{j=1}^{\infty} A_j$,
4. für endlich viele Mengen $A_1, A_2, \dots, A_n$ aus $\mathcal{A}$ ergibt sich $\bigcap_{j=1}^n A_j \in \mathcal{A}$. [25, S. 3]

Beweis:

1. aus dem Punkt 2 der Definition folgt $\emptyset \in \mathcal{A}$ und $\Omega = \emptyset^c$,
2. man setzt $A_{n+1} = A_{n+2} = \dots = \emptyset$. Da $j = 1, 2, \dots$ immer $A_j \in \mathcal{A}$ gilt und nach dem Punkt 3. der Definition weiß man auch $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$. Dann ist 2. bewiesen, wegen $\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^n A_j$,
3. $A_j \in \mathcal{A}$ und $A_j^c \in \mathcal{A}$, daher gilt $\bigcup_{j=1}^{\infty} A_j^c \in \mathcal{A}$. Da die σ -Algebra gegenüber Komplementbildung abgeschlossen ist, kann man $\left(\bigcup_{j=1}^{\infty} A_j^c\right)^c \in \mathcal{A}$. Und mit Hilfe der Morganschen Regel $\left(\bigcup_{j=1}^{\infty} A_j^c\right)^c = \bigcap_{j=1}^{\infty} A_j$ hat man 3. bewiesen,
4. der Beweis für 4. kann man aus 2. folgern. [25, S. 4]

Zu den typischen Exempeln, die die σ -Algebra repräsentieren, zählen $\mathcal{A} = \{\emptyset, \Omega\}$ oder $\mathcal{A} = P(\Omega)$. [25, S. 4]

Definition: Es sei $\mathcal{C} \subseteq P(\mathbb{R})$ die Gesamtheit aller abgeschlossenen endlichen Intervalle in $\mathbb{R}$, also

$$\mathcal{C} = \{[a, b]: a < b, a, b \in \mathbb{R}\}.$$

Die von $\mathcal{C}$ auf $\mathbb{R}$ erzeugte σ -Algebra $\sigma(\mathcal{C})$ heißt Borelsche σ -Algebra, kurz $\mathcal{B}(\mathbb{R})$ geschrieben. Die dazugehörigen Mengen von $\mathcal{B}(\mathbb{R})$ werden auch als Borelmengen bezeichnet. [25, S. 5]

4.6 Das Wahrscheinlichkeitsmaß

Die Wahrscheinlichkeit gibt Werte an, die die Vorstellung beim Eintreten eines Ergebnisses unterstützen sollen. Dabei werden im Normalfall Zahlen zwischen $[0, 1]$ angegeben. Je nachdem, ob das Ereignis mit ziemlicher Sicherheit oder eher unwahrscheinlicher eintritt. [25, S. 5]

Es soll ein Experiment mit n Wiederholungen durchgeführt werden, wobei jeder neue Versuch unabhängig von den anderen abläuft. Bereits aus einem vorigen Kapitel ist bekannt, dass es sich bei $H_n(x_i)$ um die absoluten Häufigkeiten handelt. Diese zählen mit, wie oft das Ereignis A eingetreten ist. Wie schon bekannt ist, lässt sich daraus der Begriff der relativen Häufigkeiten ableiten, und man schreibt dafür $h_n(x_i) = \frac{H_n(x_i)}{n}$. Bei unendlicher Wiederholung des Zufallsexperimentes, ordnet sich der Wert der relativen Häufigkeiten bei einem bestimmten Wert von $[0, 1]$ ein. Dafür muss jedoch die σ -Additivität gelten. [25, S. 6]

Definition: Die Abbildung P ist σ -additiv, wenn sich für abzählbar unendlich viele paarweise disjunkte Ereignisse $A_1, A_2, \dots$ in Ω die Wahrscheinlichkeit der Vereinigung so berechnen lässt:

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j). \quad [25, S. 6]$$

Da es bei zu großen beziehungsweise zu kleinen Grundräumen nicht immer möglich ist, die entsprechende Wahrscheinlichkeit anzugeben oder nur bedingt zu formulieren, wird eine weitere Definition benötigt.

Definition: Es sei ein beliebiger Grundraum Ω mit einer σ -Algebra ausgestattet. Die Abbildung P heißt Wahrscheinlichkeitsmaß oder Wahrscheinlichkeitsverteilung auf $(\Omega, \mathcal{A})$, wenn Folgendes gilt:

1. P bildet von $\mathcal{A}$ nach $[0, 1]$ ab mit $P(\emptyset) = 0$ und $P(\Omega) = 1$,
2. P ist σ -additiv, das bedeutet, für paarweise disjunkte Mengen $A_1, A_2, \dots$ aus $\mathcal{A}$ folgt

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j). \quad [25, S. 7]$$

Dadurch kann an dieser Stelle nun eine weitere Begriffserklärung angegeben werden.

Definition: Ein Wahrscheinlichkeitsraum ist ein Tripel $(\Omega, \mathcal{A}, P)$ mit Grundraum Ω , einer σ -Algebra $\mathcal{A}$ von Teilmengen aus Ω , der sogenannten Ereignis- σ -Algebra, und einem Wahrscheinlichkeitsmaß P definiert auf $\mathcal{A}$ und mit den Werten aus $[0, 1]$. [25, S. 8]

Um sowohl einen besseren Einblick als auch einen praktischen Umgang mit dem Begriff des Wahrscheinlichkeitsmaßes zu erlangen, werden nun noch einige wichtige Eigenschaften aus [25] angeführt, wobei für die jeweiligen Beweise auf die Literatur verwiesen wird.

Satz: Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, dann gilt:

1. die Abbildung P ist endlich-additiv,
2. für zwei Ereignisse $A, B \in \mathcal{A}$ mit $A \subseteq B$ folgt $P(B \setminus A) = P(B) - P(A)$,
3. es gilt $P(A^c) = 1 - P(A)$ für $A \in \mathcal{A}$,
4. das Wahrscheinlichkeitsmaß ist monoton, das heißt, aus der Inklusion $A \subseteq B$ für Ereignisse $A, B \in \mathcal{A}$ folgt $P(A) \leq P(B)$,
5. das Wahrscheinlichkeitsmaß ist subadditiv, das bedeutet, für beliebige $A_j \in \mathcal{A}$ gilt

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} P(A_j),$$

6. P ist stetig von unten, das heißt, es sind $A_j \in \mathcal{A}$ mit $A_1 \subseteq A_2 \subseteq \dots$, dann folgt

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{j \rightarrow \infty} P(A_j),$$

7. P ist stetig von oben in folgendem Sinn: gilt für Ereignisse $A_j \in \mathcal{A}$ die Aussage $A_1 \supseteq A_2 \supseteq \dots$, dann folgt

$$P\left(\bigcap_{j=1}^{\infty} A_j\right) = \lim_{j \rightarrow \infty} P(A_j). \quad [25, S. 9]$$

Die nachfolgende Darstellung von Kolmogorov und der Wahrscheinlichkeit scheitert für unendliche Mengen von Ω . Sie zählt nicht zu den aktuellen Betrachtungen der Wahrscheinlichkeitstheorie und gilt als Spezialfall der endlichen Mengen.

4.7 Kolmogorov und die Wahrscheinlichkeit

Aufgrund der starken Beziehung zwischen der Wahrscheinlichkeitsrechnung und der Physik, leiten sich viele Theorien und Ansätze der Wahrscheinlichkeit aus der Physik ab. Deshalb gibt es eine starke Verbindung zwischen den beiden naturwissenschaftlichen Themen. Dabei versuchte Alexander Kolmogorov zu Anfang des 20. Jahrhunderts einen axiomatischen Zugang zum Thema Wahrscheinlichkeitsrechnung. [10, S. 152] Jedoch gilt diese Auffassung der Wahrscheinlichkeitstheorie als veraltet und ist eine spezielle Ansicht für die Wahrscheinlichkeit mit endlichen Mengen.

In Kolmogorovs Theorie behandelt man eine endliche Menge $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ bestehend aus den Elementen $\omega_1, \omega_2, \dots, \omega_n$, die Ergebnisse repräsentieren. Weiters gibt es eine Potenzmenge $\mathcal{P}(\Omega)$. Nun versteht man unter der Wahrscheinlichkeitsverteilung unter der Ansicht von Kolmogorov, dass die Funktion, jeder Menge $\mathcal{P}(\Omega)$ eine reelle Zahl $P(E)$ zuweist. [10, S. 152] Dieser Wahrscheinlichkeitsverteilung unterliegen gewisse Axiome, die anweisen, wie man mit der Funktion umgehen darf.

Im Vergleich zu den Eigenschaften im Satz der Laplace-Wahrscheinlichkeiten reduziert Kolmogorov seine Axiome auf drei Aussagen, die erfüllt werden müssen.

Definition: Ein Wahrscheinlichkeitsraum (Ω, P) ist ein Paar bestehend aus einer nichtleeren endlichen Menge $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ und einer Funktion $P: \mathcal{P}(\Omega) \rightarrow \mathbb{R}$ mit den Eigenschaften

1. $P(E) \geq 0 \quad \forall$ Teilmengen E von Ω (Nichtnegativität),
2. $P(\Omega) = 1$ (Normiertheit),
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2) \quad \forall$ Teilmengen E_1, E_2 mit $E_1 \cap E_2 = \emptyset$ (Additivität).

Ω heißt Ergebnismenge, $\mathcal{P}(\Omega)$ heißt Ereignismenge, P Wahrscheinlichkeitsverteilung und $P(E)$ Wahrscheinlichkeit des Ereignisses E . [10, S. 153] Durch diese Definition von Kolmogorov können daraus weitere Eigenschaften aus der Struktur der Axiome abgeleitet werden.

Satz: Es sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum. Dann gilt:

1. $P(\emptyset) = 0$,
2. $P(E) \leq 1 \quad \forall$ Teilmenge E .
3. $P(\bar{E}) = 1 - P(E) \quad \forall$ Teilmenge E ,
4. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) \quad \forall$ Teilmengen E_1, E_2 .
5. $P(E) = \sum_{\omega \in E} P(\{\omega\})$ und $P(\Omega) = \sum_{\omega \in \Omega} P(\{\omega\}) = \sum_{i=1}^n P(\{\omega_i\}) = 1$. [10, S. 153]

Satz (Festlegung von Wahrscheinlichkeitsverteilungen): Ist $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ eine endliche Ergebnismenge, so ist eine Wahrscheinlichkeitsverteilung P durch ihre Werte für die Elementarereignisse eindeutig festgelegt. Wird hierfür jedem $\omega_i \in \Omega$ eine Zahl $P(\{\omega_i\}) \in [0; 1]$ zugeordnet, wobei $\sum_{i=1}^n P(\{\omega_i\}) = 1$ gilt, und wird festgesetzt

$P(E) = \sum_{\omega \in \Omega} P(\{\omega\})$, so ist P eine Wahrscheinlichkeitsverteilung auf $\mathcal{P}(\Omega)$. Die Umkehrung gilt genauso. [10, S. 155]

5. Was hat Zählen mit Wahrscheinlichkeiten zu tun? - Kombinatorik

Das einfache Zählen ist ein elementarer Bestandteil der Wahrscheinlichkeit. Denn schon das banale Zählen der möglichen Ereignisse oder auch das Zählen der Laplace-Wahrscheinlichkeit $\frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}}$ verlangt das grundlegende Abzählen von Möglichkeiten. Diese geschickte Zählstrategie der günstigen und der möglichen Fälle wird in der Mathematik auch mit dem Wort Kombinatorik beschrieben. Allerdings hat auch diese Technik in der Wahrscheinlichkeitsrechnung ihre Grenzen, und zwar genau dann, wenn man mit zu großen Ergebnismengen hantiert. [10, S. 190] Hier hat die Kombinatorik klar ihre Vorteile, dennoch zeigen sich auch gewisse Nachteile.

Exemplarisch dafür ist das weltweit beliebte Spiel Lotto. In Europa gibt es unterschiedliche Spielsysteme, in Österreich spielt man „6 aus 45“. [10, S. 193] Dabei muss man aus 45 anzukreuzenden Zahlen 6 kennzeichnen. [26] Die Reihenfolge ist irrelevant. Das heißt, dass es im österreichischen Lotto für die erste gezogene Kugel 45 verschiedene Möglichkeiten gibt. Für die zweite Ziehung bleiben noch 44 unterschiedliche Kugeln zur Auswahl und so weiter, bis 6 Kugeln gezogen wurden. Also gibt es insgesamt $45 \cdot 44 \cdot 43 \cdot 42 \cdot 41 \cdot 40$ Auswahlmöglichkeiten, eine Kugel zu ziehen. Und die Chance, dass eine der getippten Zahlen gezogen wird, verteilt sich auf $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ Möglichkeiten. Insgesamt ergibt dies $\frac{45 \cdot 44 \cdot 43 \cdot 42 \cdot 41 \cdot 40}{6!} = 8\,145\,060$ verschiedene Tipps. Man würde also dann bei Lotto gewinnen, wenn man 8 145 060 Tipps abgegeben hätte. [10, S. 194]

Aber auch das Toto-Spiel ist ein geeignetes Beispiel für Kombinatorik. Anders als beim Lotto gibt man bei den Sportwetten für 12 mögliche Spielausgänge jeweils X für ein Unentschieden, 1 für die Wahrscheinlichkeit eines Heimsieges, und 2 im Falle eines Auswärtssieges an. Die Spielerin oder der Spieler kreuzt dann, je nach dem wie die Chancen stehen, die zu vermutenden Spielausgänge an. In diesem Fall entsprechen die Resultate 12-Tupel, die je mit der Wahl zwischen X , 1 oder 2 gewinnen können. Daher ergibt $3^{12} = 531\,441$. [18, S. 56]

Der praktische Umgang mit dem Zählen von Zahlen hat sehr viele Vorteile. Deshalb haben sich diesen Nutzen der Kombinatorik auch die bekannten Glücksspiele zu eigen gemacht, um daraus Profit zu schlagen. Die hier genannten Modelle des Lotto oder Toto, stammen aus der Idee der elementaren Grundlagen. Dahinter versteckt sich nichts anderes als die Multiplikationsregel.

Multiplikationsregel: Es sollen k -Tupel $(a_1, a_2, \dots, a_k)$ gebildet werden, indem man die k Plätze des Tupels nacheinander von links nach rechts besetzt, die Reihenfolge wird dabei nicht beachtet. Für die Anordnung von a_1 gibt es i_1 Varianten, für die Auswahl von a_2 bleiben i_2 mögliche Stellen übrig, ...

für die Stelle vom Objekt a_{k-1} sind i_{k-1} Plätze zur Verfügung, und für die letzte a_k Stelle bleiben noch i_k Möglichkeiten. Somit ergeben sich alles in allem $i_1 \cdot i_2 \cdot \dots \cdot i_{k-1} \cdot i_k$ verschiedene *Tupel*, die man bilden kann. [18, S. 55]

Die Kombinatorik besteht zum einen aus Kombinationen und zum anderen aus Permutationen. Sie leistet in der Mathematik einen großen Beitrag im Bereich der Wahrscheinlichkeitstheorie. Und in den beiden folgenden Unterkapiteln soll der Unterschied zwischen den Sachverhalten von Kombinationen und Permutationen anschaulich anhand der verschiedenen möglichen Modelle erklärt werden. Wichtig für diese Unterscheidung zwischen Kombinationen und Permutationen ist zuerst, ob man das Experiment mit oder ohne Wiederholung durchführt. Außerdem muss man darauf achten, ob die Reihenfolge relevant oder irrelevant ist. Tupel erweisen sich in diesem Fall als ein gutes Vorstellungsmittel und repräsentieren die denkbaren Ausgänge eines Experimentes. [18]

Ein typisches Beispiel für das Zählen in der Kombinatorik ist das Geburtstagsparadoxon: Die Aufgabe lautet, wie hoch die Wahrscheinlichkeit ist, dass in einer Gruppe von 25 verschiedenen Personen, zum Beispiel einer Schulklasse, mindestens 2 davon am selben Tag Geburtstag haben, wobei das Geburtsjahr ausgeschlossen wird. Diese Chance wird intuitiv häufig als sehr gering angesehen. Aber die Lösung lautet wie folgt:

Man berechnet nach dem Laplace-Modell, denn alle 365 Tage im Jahr sind ja gleich wahrscheinlich, so gibt es 365 mögliche Fälle eines Geburtstages. Man unterstützt diesen Gedanken durch die Annahme der Gegenwahrscheinlichkeit des Ereignisses E , dass alle Geburtstage der 25 verschiedenen Personen auf einen anderen Tag fallen. Dann berechnet man $P(E) = 1 - P(\bar{E})$.

$$P(E) = 1 - \frac{365 \cdot 364 \cdot \dots \cdot 341}{365^{25}} \approx 0,568. [23, S. 381]$$

Vermutlich hätte man das Ergebnis von mehr als 50% Wahrscheinlichkeit für das Ereignis, dass 2 von 25 Personen am gleichen Tag ihren Geburtstag feiern, nicht für möglich gehalten. Daher bietet sich diese Aufgabenstellung als sehr interessant für den Schulunterricht an. Man könnte nach dem Lesen der Angabe die Schülerinnen und Schüler zuerst eine erste Angabe zu ihrer Vermutung abgeben lassen. Und danach mit ihrem Ergebnis vergleichen. Oft lernen Jugendliche durch ein solches unerwartetes Ergebnis mehr. Vor allem aber weckt man das Interesse der Studierenden und regt die Neugier auf das Thema an. Das Beispiel zählt zu den Besetzungsproblemen, welche oft in der Physik erläutert werden. [1, S. 119]

Die Verallgemeinerung für eine Gruppe von n Menschen würde

$$P(E) = 1 - \frac{365 \cdot 364 \cdot \dots \cdot (365 - n + 1)}{365^n}$$

lauten. [23, S. 381]

Ähnlich wie andere Beispiele in der Stochastik und vor allem in der Wahrscheinlichkeitstheorie zeigt das Beispiel mit dem zufälligen Geburtstagszwilling, dass die eigenen Vorstellung Grenzen besitzt und dass trotz mathematischen Wissens die eigene Intuition zu Irrtümern führen können. Durch Überprüfung und Anwendung von stochastischen Rechenmodellen gelangt man dann zu der richtigen, oft etwas überraschenden Lösungen. Dann ist die entsprechende Interpretation gefragt und es sollte versucht werden den

Ausgang des Experimentes verständlich deuten zu können. So kann man als Lehrkraft die Schülerinnen und Schüler für das Thema interessieren. [1, S. 122]

Jedoch ist es als Lehrerin beziehungsweise als Lehrer trotzdem wichtig, auf den Unterschied zwischen Permutationen und Kombinationen Rücksicht zu nehmen. Deshalb sollte man sehr wohl die Themen im Unterricht fachlich erklären und anhand von Beispielen illustrieren. Denn nur durch den richtigen Zugang und den entsprechenden Umgang mit dem Themenbereich wird das mathematisch korrekte Arbeiten und Wissen der Jugendlichen gefördert. Deshalb wird im Folgenden sowohl das Kapitel der Permutationen als auch jenes der Kombinationen fachmathematisch erläutert. Diese Erklärungen sind aber auch für den Schulgebrauch bestens geeignet.

5.1 Permutationen

M sei eine Menge mit den Einträgen von k -Tupel $(a_1, a_2, \dots, a_k)$, dann nennt man M eine k -Permutation mit Wiederholung. Mit Wiederholung bedeutet bloß, dass ein Objekt der Menge mehrmalig auftauchen kann. Diese Menge der k -Permutationen entspricht der Menge des kartesischen Produkts aus k .

$$M^k = \{(a_1, a_2, \dots, a_k) : a_j \in M \text{ für } j = 1, \dots, k\}. [18, S. 56]$$

Satz: Für die Menge $M = \{1, 2, \dots, n\}$ gibt es:

1. n^k, k -Permutationen mit Wiederholung,
2. $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)$, k -Permutationen ohne Wiederholung für $k \leq n$. [10, S. 199]

Beweis:

1. Angenommen man zieht Kugeln aus einer Menge von n Elementen aus einer Urne. Es stehen k Plätze zur Verfügung. Für jeden Platz ist es erlaubt, n unterschiedliche Kugeln zu geben, weil man ja das Experiment mit Wiederholung durchführt. Das bedeutet, bei jedem neuen Zug befinden sich gleich viele Elemente in der Urne wie beim Zug zuvor. Deshalb kommt man zurück zu der Annahme von n^k Möglichkeiten. [10, S. 199]
2. Wieder nimmt man ausgehend vom Urnen-Modell an, dass es k mögliche Plätze gibt, dieses Mal unter der Bedingung von $k \leq n$. Der erste Platz kann theoretisch mit n unterschiedlichen Kugeln besetzt werden. Die zweite Stelle kann mit allen weiteren Kugeln, bis auf die erste schon abgelegte, belegt werden, also $n-1$ Plätze. Dies führt zum Term $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1) = \frac{n!}{(n-k)!}$. [10, S. 200]

Als typisches Beispiel sollte man an dieser Stelle nochmals die oben genannten Glücksspiele erwähnen. So spiegelt das Lotto-Spiel eine Permutation ohne Wiederholung wider. In Zahlen ausgedrückt bedeutet dies aus der Menge $\{1, 2, \dots, 45\}$ wird sechs Mal, entsprechend den sechs angekreuzten Zahlen, ohne Zurücklegen gezogen. Und die Struktur einer Permutation mit Wiederholung gleicht dem Toto-Spielen. Hier wird aus der möglichen Menge $\{0, 1, 2\}$ mit den Ausgängen von 3^{12} veranschaulicht.

5.2 Kombinationen

Im Gegensatz zu den Permutationen spielt die Reihenfolge bei den Kombinationen keine entscheidende Rolle. Man sagt auch, dass jede k -Kombination eine besondere k -Permutation ist. Dies soll bedeuten, dass eine k -Permutation $(a_1, a_2, \dots, a_k)$ mit den Zahlen $1, 2, \dots, n$ und der Anordnung $a_1 \leq a_2 \leq \dots \leq a_k$ eine k -Kombination aus $\{1, 2, \dots, n\}$ mit Wiederholung ist. Wie bei den Permutationen gilt auch hier der zweite Fall für Kombinationen ohne Wiederholung. Dies ist hier genau dann der Fall, wenn $k \leq n$ und sich alle Elemente des k -Tupels unterscheiden. So bedeutet, dass $a_1 < a_2 < \dots < a_k$ ist, dann handelt es sich dabei um k -Kombinationen aus $\{1, 2, \dots, n\}$ ohne Wiederholung. [18, S. 57]

Satz: Für die Menge $M = \{1, 2, \dots, n\}$ gibt es:

1. $\binom{n}{k}$, k -Kombinationen ohne Wiederholung für $k \leq n$,
2. $\binom{n+k-1}{k}$, k -Kombinationen mit Wiederholung. [10, S. 199]

Beweis:

1. Bei der Permutation ohne Wiederholung handelt es sich um eine spezielle Anordnung der Komponenten. Denn schließlich führt uns die genannte Überlegung zum Binomialkoeffizienten. Die Kombinationen $\{a_1, a_2, \dots, a_k\}$, die ordentlich strukturiert sind, lassen sich aus den möglichen Kugeln $a_1, a_2, \dots, a_k$ erfassen. So folgert man $\binom{n}{k} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!}$. [10, S. 200]
2. Für den Beweis benötigen wir den gerade eben bewiesenen Teil von 1. Man nehme an $a_1 a_2 \dots a_k$ mit $1 \leq a_1 \leq a_2 \leq \dots \leq a_k \leq n$ entspreche einer k -Kombination mit zurücklegen. Wegen $1 \leq a_1 \leq a_2 \leq \dots \leq a_k \leq n$ führt dies weiter zu $1 \leq a_1 < a_2 + 1 < a_3 + 2 < \dots < a_k + k - 1 \leq n + k - 1$, so stellt jede k -Kombination $a_1 a_2 \dots a_k$ mit Zurücklegen einer n -elementigen Menge eindeutig eine k -Kombination $a_1(a_2 + 1)(a_3 + 2) \dots (a_k + k - 1)$ einer $(n + k - 1)$ -elementigen Menge ohne Zurücklegen dar. Die Umkehrung $1 \leq b_1 \leq b_2 - 1 \leq b_3 - 2 \leq \dots \leq b_k - k + 1 \leq n$ zu $1 \leq b_1 < b_2 < b_3 < \dots < b_k \leq n + k - 1$. [10, S. 200]

5.3 Modellvorstellung zur Kombinatorik

Um solche komplexen Zählaufgaben zu vereinfachen, gibt es besondere Methoden, damit man sich den Sachverhalt besser und leichter vorstellen kann. Es gibt einige Systeme, die den Zusammenhang des Problems zugänglicher gestalten. Die beiden wichtigsten Modelle, welche dazu zu nennen sind, sind das Urnen-Modell und das Fächer-Modell. Auf beide Vorstellungsarten wird in den nächsten Zeilen näher eingegangen.

Vor allem im Hinblick auf die Schule sind solche zugänglichen Vorstellungen von enormer Wichtigkeit. Gerade den Schülerinnen und Schülern, die besonders im Fach Mathematik mit

den Vorstellungen der komplexen Fachzusammenhänge kämpfen müssen, bieten die beiden vorhin genannten Modelle einen anschaulichen Einstieg und eine außerordentliche Stütze in der gesamten Wahrscheinlichkeitstheorie.

Besonderes Augenmerk, gerade in der Institution Schule, sollte auf genaue Formulierungen gelegt werden. Speziell sollte man die Jugendlichen auf die im vorigen Kapitel schon genannten Möglichkeiten mit beziehungsweise ohne Beachtung der Reihenfolge und mit beziehungsweise ohne Zurücklegen hinweisen. Oft liegen hier, bei den Grundlagen, schon Verwirrungen oder Fehler vor. Meistens nur deshalb, weil man in den Schulen der Grundlage des Themas zu wenig Zeit widmet und die Schülerinnen und Schüler den Unterschied zwischen den vier verschiedenen Zusammenhängen nicht selbstständig erarbeiten lässt. Infolgedessen geschehen oftmals schon schwere Fehler in den Vorstellungen der jungen Mädchen und den Burschen. An dieser Stelle sollte man als Lehrerin und als Lehrer klar differenzieren zwischen den unterschiedlichen Arten, die die Modelle bieten.

5.3.1 Urnen-Modell

Beim Urnen-Modell stellt man sich, wie der Name schon sagt, eine Urne vor, aus der man nummerierte Kugeln zieht. Dieses Verfahren kann man, wie in den obigen Zeilen bereits erwähnt wurde, mit oder ohne Zurücklegen durchführen. Im Allgemeinen bezeichnet man auch Experimente, bei denen man die Reihenfolge beachtet, als Variationen und jene, bei denen man die Reihenfolge außer Acht lässt, als Kombinationen. [19] Gerade diese Vorstellung entspricht einer guten Basis, um Wahrscheinlichkeiten richtig aufzufassen.

1. Ziehen mit Zurücklegen, Reihenfolge relevant:

Im Urnen-Modell stellt man sich nun n nummerierte Kugeln vor. Bei jeder neuen Ziehung beträgt die gesamte Summe der gegebenen Menge n , da man die gezogene Kugel ja wieder zurücklegt. So bleibt die Anzahl immer gleich. Und man gelangt zu folgender Anzahl von Auswahlmöglichkeiten

$$n^k. [19]$$

Typisches Beispiel für diese Art ist der Toto Schein. Wenn man das Spiel im Hinblick auf das Urnen-Modell interpretiert, kann man sagen, dass in der Urne die Kugeln durch X , 1 und 2 gekennzeichnet werden und man daraus 12 Mal mit zurücklegen zieht. [19]

2. Ziehen ohne Zurücklegen, Reihenfolge relevant:

Der Unterschied zum ersten Punkt ist, dass man nun ohne Zurücklegen aus einer Urne mit n nummerierten Kugeln zieht. Das bedeutet, dass man bei jedem neuen Zug um eine Kugel weniger in der Urne findet. Was uns zu dem Gesamtausdruck kommen lässt

$$n \cdot (n - 1) \cdot \dots \cdot (n - k + 1) = \frac{n!}{(n - k)!}.$$

Spricht man hat die Möglichkeit, aus einer Menge mit n Elementen k Elemente auf $\frac{n!}{(n-k)!}$ unterschiedlicher Weise auszuwählen. [19]

3. Ziehen mit Zurücklegen, Reihenfolge irrelevant:

Bei dieser Art legt man zwar im Urnen-Modell die gezogenen Kugeln nach jedem Zug wieder zurück, aber die Reihenfolge spielt hierbei keine Rolle. So liegt die Zahl der Auswahl der Chancen, die gezogen werden können bei

$$\binom{n+k-1}{k}$$

Möglichkeiten, wobei n der gesamten Anzahl der Menge entspricht und k der Stückzahl der Ziehungen. [19]

4. Ziehen ohne Zurücklegen, Reihenfolge irrelevant:

Da die Reihenfolge unwichtig ist, hat man in diesem Fall

$$k!$$

Möglichkeiten. Das bedeutet es bestehen

$$\frac{n!}{k! (n-k)!} = \binom{n}{k}$$

Möglichkeiten, aus einer Urne mit n Kugeln k auszuwählen. [19]

5.3.2 Fächer-Modell

Andere Vorstellungen braucht man beim Fächer-Modell. Hier ist der Grundgedanke kein Zug, sondern ein Belegen von Fächern mit Teilchen. Dabei gleichen die Fächer der Urne und die Teilchen ähneln den Zügen. In anderen Worten bedeutet dies, dass man k Teilchen (Züge) auf n mögliche Fächer aufteilt. Hier gibt es kein Zurücklegen, sondern hier entspricht diesem Ausdruck eine Mehrfachbelegung. Also ist es möglich, in ein Fach mehrere Teilchen zu legen. Und der nächste Unterschied bezeichnet, ob die Teilchen gleich sind oder als unterscheidbar gelten. [18, S. 65] Wie beim Urnen-Modell sollte man großes Augenmerk auf die jeweiligen Möglichkeiten und den entsprechenden Unterschied lenken.

1. Teilchen unterscheidbar, Mehrfachbelegung erlaubt:

Entsprechend dem Urnen-Modell 1 lautet hier die Anzahl der Möglichkeiten

$$n^k. [18, S. 65]$$

Jedoch interpretiert man hier in einem anderen Ausmaß. Wenn man sich das Toto-Beispiel in Erinnerung ruft, so erfasst man das Spiel folgendermaßen: Man besitzt 12 unterschiedliche Teilchen, die auf die bezeichneten Fächern X , 1, 2 aufgeteilt werden können. [19]

2. Teilchen unterscheidbar, Mehrfachbelegung verboten:

Die Anzahl der Möglichkeiten berechnet man wie beim Urnen-Modell mit dem Ausdruck

$$n \cdot (n - 1) \cdot \dots \cdot (n - k + 1) = \frac{n!}{(n - k)!} \text{ [18, S. 65]}$$

Der Unterschied in der Vorstellung ist dieser, dass man keine Fächer doppelt belegen darf. Dies führt zu der Folgerung, dass man, gleich wie beim Urnen-Modell, bei jedem neuen Teilchen ein Fach weniger zur Auswahl hat.

3. Teilchen nicht unterscheidbar, Mehrfachbelegung erlaubt:

Wie beim Urnen-Modell im Punkt 3, ist diese Art für die Anzahl der Möglichkeiten mit

$$\binom{n + k - 1}{k}$$

zu berechnen. [18, S. 65] Bei dieser Möglichkeit ist die Mehrfachbelegung der Fächer zwar zugelassen, aber die Teilchen sind nicht unterscheidbar. Dies entspricht dem Nicht-Beachten der Reihenfolge im Urnen-Modell.

4. Teilchen nicht unterscheidbar, Mehrfachbelegung verboten:

$$\frac{n!}{k!(n - k)!} = \binom{n}{k}.$$

Hier ist der gleiche Faktor wie beim Urnen-Modell beim Ziehen ohne Zurücklegen, wobei die Reihenfolge unwichtig ist. [18, S. 65]

6. Erste Versuche zum Berechnen von Wahrscheinlichkeiten

Wahrscheinlichkeiten können beziehungsweise müssen nicht nur interpretiert und vorhergesagt werden; eine weitere wichtige Aufgabe ist das Berechnen von Wahrscheinlichkeiten. Auf einige Rechenstrategien wurde in den vorigen Kapiteln dazu schon eingegangen und hingewiesen, wie zum Beispiel auf die Laplace-Wahrscheinlichkeit. Dabei zählt diese Methode zur Berechnung von Wahrscheinlichkeiten, aber mehr noch zur Entwicklung des Wahrscheinlichkeitsbegriffs. Denn erst durch die Überlegungen und Folgerungen der Wahrscheinlichkeitstheorie von Laplace gelang ein erfolgreicher Fortschritt der Wahrscheinlichkeitsrechnung.

Zudem werden die Techniken zum Berechnen von Wahrscheinlichkeiten nicht nur wie in den obigen Kapiteln genannten Glücksspielen verwendet, sondern ihr Bereich umfasst ein großes und vor allem fächerübergreifendes Gebiet. Denn die Wahrscheinlichkeitsrechnung wird in der ganzen Naturwissenschaft, zum Beispiel in der Physik, aber auch in der Wirtschaft genutzt. [11] Wie bereits bekannt ist, versucht man mit Hilfe der relativen Häufigkeiten die Wahrscheinlichkeiten immer nur zu schätzen. Deshalb behält die Wahrscheinlichkeit auch immer einen Grad der Unsicherheit, da man in der Theorie nur Vermutungen abgeben kann.

Insbesondere versucht man in der Wahrscheinlichkeitsrechnung mit bereits bekannten Wahrscheinlichkeiten neue zu berechnen. [10] Diese Methode erlaubt einen Einstieg in das Thema in der Schule, auch im Fach Mathematik. Wie man aus schon vorhandenen empirischen Werten auf eine neue unbekannte Wahrscheinlichkeit schließen kann, dieser Frage widmen sich nun die nächsten Unterkapitel. Es wird zunächst ausgehend von einem Beispiel der bedingten Wahrscheinlichkeit, erklärt wie man zur theoretischen Definition gelangt.

Dieses Kapitel soll das Fundament der späteren empirischen Untersuchung zum Thema „*Erhebung zur Wahrnehmung von Aufgabenstellungen der bedingten Wahrscheinlichkeit an AHS – Schülerinnen und Schülern*“ bilden. Dennoch wird in diesem Abschnitt nicht nur auf die bedingte Wahrscheinlichkeit beziehungsweise auf den Satz von Bayes eingegangen. Es werden auch andere nützliche und bekannte Berechnungsmöglichkeiten der Wahrscheinlichkeitslehre präsentiert und anhand von Beispielen transparent für die Leserin und den Leser dargestellt.

6.1 Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit, die man ohne weiteres als selbstverständlich verwendet, fordert nichts desto trotz ein umfangreiches Verständnis. Vor allem verlangt die bedingte Wahrscheinlichkeit eine Vorstellung dazu, wie sich Wahrscheinlichkeiten im Allgemeinen verändern können, wenn ein Sachverhalt bereits bekannt ist. Ein bereits vertrautes Ereignis, welches auf den weiteren Ausgang des Zufallsexperimentes Einfluss hat, ändert auch die Erwartungen der Beobachterin beziehungsweise des Beobachters. [3, S. 90]

Zum besseren Verständnis wird dieses Unterkapitel mit Hilfe eines Einstiegsbeispiels näher erläutert.

Beispiel: Für einen Betriebsrat wird eine Person nachgewählt. Zur Auswahl stehen dabei 5 Frauen und 8 Männer. Von den möglichen Kandidatinnen und Kandidaten sind 3 Frauen und 3 Männer Angestellte. Daraus folgt, dass die anderen 2 Frauen und die übrigen 5 Männer zu den Arbeitern gehören. Um dies übersichtlich darstellen zu können, fertigt man folgende Tabelle 4 an.

Tabelle 4: Vierteltafel [10, S. 167] adaptiert von [8, S. 29]

	Angestellte/Angestellter	Arbeiterin/Arbeiter	Summe
Weiblich	3	2	5
Männlich	3	5	8
Summe	6	7	13

Nun sollen die folgenden möglichen Ereignisse mit der Laplace – Wahrscheinlichkeit $\frac{|\text{günstige}|}{|\text{mögliche}|}$ dargestellt werden:

$$P(A) = \text{die gewählte Person ist Angestellte/Angestellter} = \frac{6}{13} \approx 0,462,$$

$$P(B) = \text{die gewählte Person ist ein Mann} = \frac{8}{13} \approx 0,615,$$

$$P(A \cap B) = \text{die gewählte Person ist ein Angestellter und ein Mann} = \frac{3}{13} \approx 0,231. [8, S. 29]$$

Es ist vor der offiziellen Bekanntgabe veröffentlicht worden, dass der Gewinner ein Mann ist. Wie hoch ist nun die Wahrscheinlichkeit dafür, dass ein Angestellter die Wahl gewonnen hat? So bleiben nur mehr 8 mögliche Gewinner übrig. Die Wahrscheinlichkeit, dass es sich dabei um einen Angestellten handelt, beträgt $\frac{3}{8}$.

Um genauer zu sein, bedeutet dieses Ergebnis exakt die Wahrscheinlichkeit, dass Ereignis B eintritt, unter der Bedingung, dass das Ereignis A eingetreten ist. Diese neue Wahrscheinlichkeit entspricht nicht dem Wert von $P(A)$. Denn nun schreibt man

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{3}{13}}{\frac{8}{13}} = \frac{3}{8} = 0,375. [8, S. 30]$$

Man sagt, die Wahrscheinlichkeit für das Ereignis A unter der Bedingung des Ereignisses B ist gleich die Wahrscheinlichkeit dafür, dass die gewählte Person ein Angestellter und Mann ist, durch die Wahrscheinlichkeit, dass es sich dabei um einen Mann handelt. Das bedeutet, dass die beiden Ereignisse voneinander abhängig sind. [10, S. 166]

Da ein Ereignis ja bereits eingetreten ist, kann man dieses mit Sicherheit weiterverwenden, um die neue unbekannte Wahrscheinlichkeit zu berechnen. Also ist es möglich, aus bekannten Wahrscheinlichkeiten empirisch ein neues, noch nicht bekanntes Ereignis zu berechnen. Es handelt sich dann um die bedingte Wahrscheinlichkeit, wenn sie den Ausgang des Ereignisses B darlegt, wenn der Vorgang A bereits geschehen ist.

Wie schon am Anfang des Beispiels erwähnt, kann man anfangs mit der Laplace-Strategie an das Problem herangehen. Dies kann man sich in der Abbildung 13 graphisch besser vorstellen.

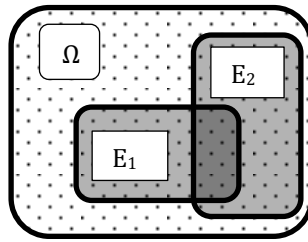


Abbildung 13: Bedingte Laplace-Wahrscheinlichkeiten [10, S. 168]

Dank der Laplace-Vorstellung kann man die bedingte Wahrscheinlichkeit definieren. Es seien $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ und $A = \{\omega_1, \omega_2, \dots, \omega_k\}$. [10, S. 168]

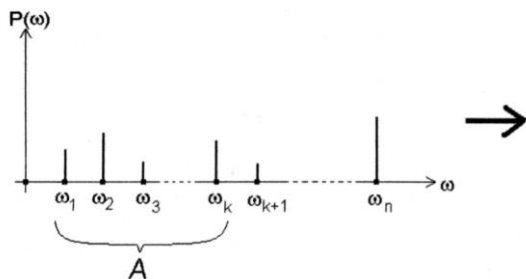


Abbildung 14: Bedingte Wahrscheinlichkeiten [10, S. 168]

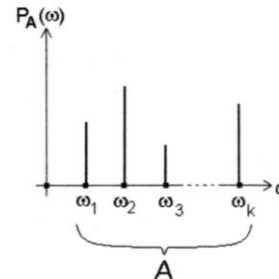


Abbildung 15: Bedingte Wahrscheinlichkeiten [10, S. 168]

In der Abbildung 14 sieht man die Wahrscheinlichkeiten der Ereignisse. Die Ergebnisse werden in einem Stabdiagramm veranschaulicht. 1 lautet die Summe, wenn man alle Stablängen addiert. Hingegen, zählt man alle Werte der zu A gehörigen Stablängen zusammen, beträgt die Summe $P(A)$. Bei der bedingten Wahrscheinlichkeit weiß man ja aus der Erfahrung, dass ein Ereignis bereits eingetreten ist. Daher kann man andere Ausgänge ausschließen. Dann bedeutet dies für die Wahrscheinlichkeiten

$P(\omega_i|A) = 0$ für $k+1, \dots, n$. In Folge dessen machen die übrigen Wahrscheinlichkeitswerte $P(\omega_i|A)$ für $i = 1, \dots, k$ alles in allem 1 aus. Wenn man alle eingetretenen Wahrscheinlichkeiten mit der Bestimmungsgröße r multipliziert, das heißt

$P(\omega_i|A) = r \cdot P(\omega_i)$, erkennt man die Abbildung 15. Da $r = \frac{1}{P(A)}$ ist, folgert man $1 = r \cdot P(\omega_1) + \dots + r \cdot P(\omega_k) = r \cdot P(A)$. [10, S. 168]

Insgesamt führen die beiden oben abgebildeten Darstellungen mit der Idee, die sie vermitteln, zuerst zur Motivation, und davon ausgehend zur Definition der bedingten Wahrscheinlichkeit.

Motivation der Definition: Man motiviert die Definition der bedingten Wahrscheinlichkeit als:

$$r_n(B|A) = \frac{\text{Anzahl der Versuche, in denen } B \text{ und } A \text{ eintreten}}{\text{Anzahl der Versuche, in denen } A \text{ eintritt}}. \quad [18, \text{S. } 100]$$

Definition: Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, für $A \in \mathcal{A}$ mit $P(A) > 0$. Dann heißt

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \text{ für } B \in \mathcal{A}$$

die bedingte Wahrscheinlichkeit für das Eintreten von B unter der Bedingung von A . [10, S. 169]

Satz: Es sei Ω eine paarweise disjunkte Vereinigung von $B_1, \dots, B_n$, das heißt $\bigcup_{j=1}^n B_j = \Omega$ und $B_i \cap B_j = \emptyset$ für $i \neq j$; wir nehmen $P(B_i) > 0$ für alle i an. Dann gilt:

$$P(A) = \sum_{i=1}^n P(A|B_i) \cdot P(B_i) \text{ für alle Ereignisse } A. \quad [3, \text{S. } 92]$$

Beweis: Die Familie $(A \cap B_i)_{i=1, 2, \dots, n}$ ist eine disjunkte Zerlegung von A . Folglich gilt

$$P(A) = \sum P(A \cap B_i) = \sum \frac{P(A \cap B_i)}{P(B_i)} P(B_i) = \sum P(A|B_i) \cdot P(B_i). \quad [3, \text{S. } 92]$$

Auch wenn die Definition der bedingten Wahrscheinlichkeit trivial wirkt, kommt es trotzdem immer wieder zu Fehlern beim Berechnen von Wahrscheinlichkeiten mit Hilfe der genannten Methode. Das größte Missverständnis, welches zu gravierenden Problemen führt, ist die Verwechslung zwischen der Bedingung und dem zu bedingenden Sachverhalt. [19]

Man führt als Grund der Fehlvorstellungen und Irrtümer, die oft auftretende Vertauschung der relevanten Faktoren $P(A|B) \neq P(B|A)$ an. Auf den ersten Blick mag die genannte Aussage trivial wirken, dennoch birgt genau diese Verwechslung der bedingten Wahrscheinlichkeit große Probleme in der Verwendung dieses Verfahrens.

Man beurteile die Aussagen „jeder Vater ist ein Mann“ und „jeder Mann ist ein Vater“. [10, S. 169] An diesem Exempel erkennt man die Verwechslung eindeutig, denn die beiden angeschriebenen Wahrscheinlichkeiten $P(\text{Mann}|\text{Vater}) \neq P(\text{Vater}|\text{Mann})$ demonstrieren einen unmissverständlichen Unterschied. [10, S. 169] Dennoch gibt es genügend andere Beispiele bei der Anwendung der bedingten Wahrscheinlichkeit, wo die Differenzierung nicht mehr transparent zu sein scheint.

Dieser typische Irrtum tritt aber, wie man vielleicht glauben mag, nicht nur in den Schulen auf, sondern auch in den Medien oder in der Kriminalstatistik. Ausgehend von den verschiedenen Institutionen, die diese falsch berechneten Wahrscheinlichkeiten dann veröffentlichen, betrifft es die ganze Menschheit. Oft werden daraus beunruhigende Schlagzeilen in der Zeitung oder im Fernsehen. Dies führt zum nächsten Kapitel, welches sich weiterführend mit dem falschen Berechnen und den Fehlinterpretationen der bedingten Wahrscheinlichkeiten beschäftigt.

6.2 Satz von Bayes

Die häufig auftretende Verwechslung zwischen den beiden Wahrscheinlichkeiten $P(A|B)$ und $P(B|A)$ geht zwar grundsätzlich von zwei verschiedenen Zusammenhängen aus, dennoch besteht auch ein informatives Verhältnis zwischen der Wahrscheinlichkeit A unter der Bedingung von B beziehungsweise der Wahrscheinlichkeit B , wenn die Bedingung A eingetreten ist. Dieser Zusammenhang wird auch der Satz von Bayes genannt, der einer Folgerung der bedingten Wahrscheinlichkeit entspricht. [3, S. 95] Benannt ist diese Regel nach dem im 18. Jahrhundert in London lebenden presbyterianischen Geistlichen Thomas Bayes. [10, S. 183] Eine, für die Regel von Bayes hilfreiche Folgerung ist folgender Satz über die totale Wahrscheinlichkeit.

Satz von Bayes: Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, wobei Ω in m paarweise disjunkte Teilmengen $A_1, A_2, \dots, A_m \in \mathcal{A}$ mit $P(A_j) > 0 \forall j$ zerlegt werden kann. Das bedeutet $\Omega = \bigcup_{i=1}^m A_i$ mit $A_i \cap A_j = \emptyset$ für $i \neq j$. Dann gilt für jede Zahl j und jedes Ereignis B mit $P(B) \neq 0$:

$$P(A_j|B) = \frac{P(A_j) \cdot P(B|A_j)}{\sum_{i=1}^m P(A_i) \cdot P(B|A_i)}. [10, S. 184]$$

Beweis: Man muss für den Beweis der Regel von Bayes einfach in die Formel der bedingten Wahrscheinlichkeit einsetzen und dann geschickt umformen.

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j) P(A_j)}{P(B)}. [27, S. 71]$$

Im Allgemeinen weist der Satz von Bayes keine komplizierten Eigenschaften auf. Denn im Prinzip überprüft man eine Auffassung ausgehend von sachlichen Daten. [31, S. 8] Obwohl diese Regel nur eine Ableitung der Formel zur bedingten Wahrscheinlichkeit darstellt, kommt es trotzdem immer wieder zu Verwirrungen bei der Bedingung. Vor allem wenn das Ereignis A_i als Ursache für die Darstellung des Ergebnisses B präsentiert wird. [18, S. 103] Man verwendet eine so genannte a priori-Wahrscheinlichkeit für A_i , die man vor dem Ablauf einer wahrscheinlichen Begebenheit den Wahrscheinlichkeiten A_i zuweist. [18, S. 103] Die Wahrscheinlichkeit für $P(B|A_i)$ stellt die bedingte Wahrscheinlichkeit dar für: Das Ereignis B tritt ein unter der Bedingung, dass das Ereignis A_i geschehen ist. [18, S. 103–104] Passiert hingegen das umgekehrte Szenario, dass das Ereignis A_i eintritt, wenn B die Bedingung darstellt, so spricht man von der a posteriori-Wahrscheinlichkeit. In diesem Fall bezeichnet man A_i als die Ursache von B . [18, S. 104]

Das heißt mit Hilfe der Regel von Bayes muss man durch eine weitere Information die Aussage über die Wahrscheinlichkeiten neu überdenken. [7, S. 177] Dabei spielt die intuitive Vorstellung der Wahrscheinlichkeit eine große Rolle. Denn die zentrale Problemstellung, die man mittels der Regel von Thomas Bayes, lösen kann, ist, wie man bei der gegebenen Wahrscheinlichkeit von $P(A|B)$ rechnerisch auf die umgekehrte unbekannte Wahrscheinlichkeit von $P(B|A)$ kommt. [27, S. 69]

Trotz dem Nutzen und der enormen Unterstützung, die die bedingte Wahrscheinlichkeit, und vor allem auch die davon abgeleiteten Regeln von Bayes, bieten, wird diese Strategie der Wahrscheinlichkeitsrechnung oftmals übergangen. Das nächste Beispiel und die zusätzlichen Fragestellungen helfen, um das Augenmerk auf die Genauigkeit und den Zweck dieser Formeln zu lenken.

Beispiel: Die Ähnlichkeit dreier Pilzsorten ist sehr gefährlich. Zum einen gibt es den äußerst giftigen grünen Knollenblätterpilz (G) und zum anderen gibt es die beiden genießbaren Sorten, den Waldchampignon (W) und das Nymphenbrot (N). Nach dem Sammeln von Pilzen werden im Durchschnitt 85% W , 10% G und 5% N gefunden. Es ist der Fall, dass sogar ein erfahrener Pilzsammler die Sorten nur mit einer Wahrscheinlichkeit von 95% richtig identifiziert. Die Sammlerin oder der Sammler verwechselt mit 4%iger Wahrscheinlichkeit einen W mit einem G und mit 2%iger Wahrscheinlichkeit ein N mit einem G . Nun ist die Überlegung, mit welcher Wahrscheinlichkeit ein geglaubter giftiger G auch tatsächlich einer ist. [27, S. 71]

Dabei ist die Menge A = alle für giftig gehaltenen Pilze G und gesucht ist die bedingte Wahrscheinlichkeit $P(G|A)$. [27, S. 72] Sprich wie hoch ist die Wahrscheinlichkeit von giftigen grünen Knollblätterpilzen zu finden unter der Bedingung, dass man glaubt, einen solchen gesammelt zu haben. Man kennt laut Angabe folgende Wahrscheinlichkeiten:

$P(W) = 0,85$; entspricht der Wahrscheinlichkeit einen Waldchampignon entdeckt zu haben,

$P(G) = 0,1$; repräsentiert die Chance einen ungenießbaren grünen Knollenblätterpilz gesammelt zu haben,

$P(N) = 0,05$; symbolisiert die Möglichkeit ein Nymphenbrot zu pflücken,

$P(A|W) = 0,04$; bedeutet die Wahrscheinlichkeit dafür, dass sich in der Menge aller für giftig gehaltenen auch Waldchampignons verstecken,

$P(A|G) = 0,95$; zeigt die Chance an, dass sich in der Menge aller für giftig gehaltenen Pilze auch ein giftiger grüner Knollenblätterpilz befindet,

$P(A|N) = 0,02$; und ebenfalls die bedingte Wahrscheinlichkeit, dass in der Menge der möglichen giftigen Pilze, ein Nymphenbrot enthalten ist.

Gesucht ist die abhängige Wahrscheinlichkeit dafür, dass ein giftiger grüner Knollenblätterpilz entdeckt wird, unter der Bedingung, dass man in der Menge aller für ungenießbarer haltenden Pilzen sucht. Die bedingte Wahrscheinlichkeit lautet nach der Formel:

$$\begin{aligned} P(G|A) &= \frac{P(G) \cdot P(A|G)}{P(G) \cdot P(A|G) + P(W) \cdot P(A|W) + P(N) \cdot P(A|N)} \\ &= \frac{0,1 \cdot 0,95}{0,1 \cdot 0,95 + 0,85 \cdot 0,04 + 0,05 \cdot 0,02} \\ &\approx 0,731. [27, S. 72] \end{aligned}$$

Wie so oft überrascht bei der bedingten Wahrscheinlichkeit das Ergebnis. Ungefähr $\frac{1}{4}$, der für giftige gehaltenen Pilze, wären essbar. Um den schmalen Grad der Genauigkeit zu erkennen, sollte man sich für die Differenzierung folgende Frage stellen: Mit welcher Wahrscheinlichkeit wurde beziehungsweise wurde der giftige Pilz G erkannt? [27, S. 72] Zwischen den beiden Verben „wird“ und „wurde“ besteht ein großer Unterschied im Hinblick auf die Wahrscheinlichkeit. Denn die Wahrscheinlichkeit, dass ein giftiger Pilz gefunden wird, ist ja bereits nach der Angabe bekannt, und zwar mit einer Wahrscheinlichkeit von 95%. [27, S. 72] Erst die Frage, mit welcher Wahrscheinlichkeit ein ungenießbarer Pilz gefunden wurde, verlangt die Regel nach Bayes.

Den Grundgedanken von der Regelung von Thomas Bayes repräsentiert seine Vorstellung eines Tisches, der quadratförmig ist. Auf diesen Tisch sollen Kugeln mit gleicher Wahrscheinlichkeit auf jedem Fleck dieses Möbelstückes landen können. Der Kern seiner Überlegung war, dass die Kugel an jedem Platz des Tisches stehen bleiben könnte. Es wurde zum Beispiel eine erste Kugel mit einer bestimmten Farbe auf den Tisch geworfen. Die Person, die das Experiment durchführt, steht allerdings mit dem Rücken zu der Oberfläche, sodass es nicht möglich ist, die Lage der liegenden Kugel zu erraten. Im nächsten Schritt wird eine weitere anders farbige Kugel auf den Tisch geworfen. Die Assistentin oder der Assistent des Versuches sieht die Standpunkte der beiden unterschiedlichen Kugeln. Der Durchführer allerdings nicht und fragt nun, in welcher Position die zwei Kugeln zueinander stehen. Informierte die Assistentin beziehungsweise der Assistent, dass die zweite geworfene Perle weiter rechts von der ersten Kugel liege, so kann man annehmen, dass sich die erste Kugel eher auf der linken Tischhälfte befinde. Umgekehrt gilt die entgegengesetzte Hypothese. Je öfter der Wurf hintereinander durchgeführt wird, desto sicherer kann man sich bezüglich des Standorts der ersten Kugel sein. Diese Spekulation ist möglich, weil durch die zusätzlichen Informationen der Sachverhalt eingeschränkt werden kann. Es war genau dieses Gedankenexperiment von Bayes, welches ihn Überlegungen im Hinblick auf das endgültige Bayes-Theorem anstellen ließ. [31, S. 7–8]

6.3 Das Ziegenproblem: das klassische Beispiel der bedingten Wahrscheinlichkeit

Das wohl bekannteste Problem, welches mit der bedingten Wahrscheinlichkeit zu lösen ist, ist das sogenannte Ziegenproblem. Berühmt geworden ist es durch die amerikanische Spielshow *Let's make a deal*. [18, S. 52] Die Moderatorin oder der Moderator bietet der Kandidatin oder dem Kandidaten eine Auswahl von drei verschlossenen Türen an. Hinter zwei dieser Türen befindet sich eine Ziege und hinter einer Tür steht ein Auto. Nun lässt der Moderator der Mitspielerin oder dem Mitspieler, die oder der sich für eine dieser drei Türen festlegt, freie Wahl. Als nächsten Schritt lässt die Moderatorin beziehungsweise der Moderator eine Tür öffnen, von der er sich sicher ist, dass dahinter eine Ziege steht. Nun bekommt die Kandidatin oder der Kandidat die Möglichkeit seine erste Wahl zu revidieren und sich noch für die andere Tür zu entscheiden. Die Frage lautet nun, sollte man bei seiner ersten Wahl bleiben oder doch die Tür wechseln, um seine Gewinnchancen zu erhöhen? [18, S. 52]

Aufgrund eines amerikanischen Magazins, in welchem eine Autorin genau dieses Problem in ihrem Beitrag behandelte, wurde weltweit über das Problem diskutiert. [18, S. 53] Denn einerseits erscheint die Wahrscheinlichkeit trivial zu wirken, aber andererseits steckt viel mehr Mathematik in diesem Beispiel als man auf den ersten Blick vermuten mag. Viele Menschen behaupten, dass die Entscheidung irrelevant sei, denn es handle es sich um eine fifty-fifty-Möglichkeit. [5, S. 67]

6.3.1 Lösung des Ziegenproblems

Um die Lösung anzugeben, versucht man zuerst einmal die Angabe etwas umzuformulieren. Der Quizmaster lässt die Kandidatin oder den Kandidaten frei aus den drei möglichen Türen wählen. Also bedeutet dies, dass die Wahrscheinlichkeit dafür, dass auf Anhieb die richtige Tür (jene mit dem Auto) gewählt wird, der folgenden Wahrscheinlichkeit entspricht,

$$P(\text{KandidatIn wählt die Tür hinter der das Auto steht beim 1. Versuch}) = \frac{1}{3}.$$

Die Chance von $\frac{1}{3}$ ist ja nach dem Laplace-Modell auch leicht nachvollziehbar. Der Anzahl der Möglichkeiten entsprechen die drei zur Verfügung stehenden Türen T_1 , T_2 und T_3 .

Unter der Annahme, dass die Kandidatin oder der Kandidat bei der ersten Möglichkeit die Tür T_1 gewählt hat, öffnet die Moderatorin beziehungsweise der Moderator eine der beiden Türen T_2 oder T_3 , mit der Gewissheit, dass sie oder er eine Türe wählt, hinter der eine Ziege wartet. Dies ist ja immer möglich. Zum einen weil der Quizmaster weiß, hinter welcher Tür der Gewinn steht, und zum anderen, um die Situation spannender zu gestalten. Angenommen er öffnet die Tür T_3 , hinter der eine Ziege steht. Nun bietet er der Mitspielerin beziehungsweise dem Mitspieler an, ihre beziehungsweise seine Wahl noch zu überdenken und ermöglicht einen Wechsel von Tür T_1 zur Tür T_2 .

Welcher nächste Schritt wäre in dieser Lage die klügste Entscheidung? Sollte man auf das erste Bauchgefühl vertrauen oder doch die zweite Chance zum Wechseln nutzen? Mit Hilfe der bedingten Wahrscheinlichkeit schaut die Lösung wie folgt aus:

$$P(\text{Auto hinter } T_2 | \text{ModertorIn öffnet } T_3) = \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 0} = \frac{2}{3}. [19]$$

Diese bedingte Wahrscheinlichkeit setzt sich zusammen aus dem Nenner, den man wie folgt interpretieren kann: $\frac{1}{3} \cdot 1$ entspricht der Wahrscheinlichkeit, dass sich hinter den beiden Türen T_1 und T_3 eine Ziege befindet und hinter der Tür T_2 das Auto; $\frac{1}{3} \cdot \frac{1}{2}$ bedeutet die Möglichkeit, dass hinter den Türen T_2 und T_3 eine Ziege wartet und der Gewinn hinter der Tür T_1 ; und $\frac{1}{3} \cdot 0$ ergibt sich aus der Wahrscheinlichkeit, dass hinter der Tür T_3 das Auto versteckt ist. Daher kann man die Lösung so interpretieren, dass es mit einer Wahrscheinlichkeit von $\frac{2}{3}$ ratsam ist, von der Tür T_1 zur Tür T_2 zu wechseln, um das Auto zu gewinnen.

6.3.2 Interpretation des Ziegenproblems

Das Ziegenproblem ist eine weltweit lange diskutierte Aufgabenstellung aus dem Fachgebiet der Stochastik. Vor allem durch das erstaunliche Ergebnis, welches man auf den ersten Blick nicht erwarten würde, gewinnt es immer wieder an Aufmerksamkeit. Auf den ersten Blick wirkt das Problem sehr trivial, aber man muss die gegebenen Bedingungen, die dieses Beispiel fordert, genau beachten. Denn sonst tappt man schnell in die Falle der berühmten Aufgabe, welches bereits in verschiedenen Kontexten zu interpretieren gibt. Außerdem wurde die Grundidee der Aufgabe bereits in verschiedene Zusammenhänge formuliert, wie man im späteren Teil dieser Arbeit noch lesen wird. Doch der Kern des Modells bleibt immer derselbe und zwar eine Auswahl aus drei Möglichkeiten und der danach angebotene Wechsel.

Der Ursprung der Frage geht zurück auf das Schubladenparadoxon, des französischen Mathematikers Joseph Bertrand aus dem Jahre 1889. [21, S. 206] Ungefähr 70 Jahre später wurde es in eine neue Geschichte verpackt und eine andere Art der Problemstellung formuliert. Die neu aufgesetzte Fragestellung wurde durch Martin Gardner als das Drei-Gefangenen-Problem bekannt. [21, S. 206]

Denn oft kommt man zum falschen Entschluss, dass die mögliche Lösung $\frac{1}{2}$ betragen muss. Dieser Gedanke wird durch das Öffnen einer Tür durch die Moderatorin oder den Moderator beeinflusst. Nun könnte die Begründung lauten, aufgrund der nun eingeschränkten Wahrscheinlichkeit von zwei möglichen Türen hat sich die Wahrscheinlichkeit auf den Gewinn von der ersten Vermutung $\frac{1}{3}$ auf $\frac{1}{2}$ erhöht. [16, S. 188] Denn es ist in diesem Beispiel sehr wichtig, die richtige Bedingung für die Wahrscheinlichkeit zu finden. Was man in dieser Fehlinterpretation des Ziegenproblems erkennt, ist wieder, dass man der bedingten Wahrscheinlichkeit keine Beachtung schenkt. Und dieses Missachten des Satzes von Bayes repräsentiert den typischen ignoranten Fehler und die damit verbundene falsche Interpretation der Aufgabenstellung. Das erklärt auch, warum dieses Problem von vielen Personen falsch gedeutet wird und warum diese auch auf eine falsche Lösung kommen.

Die Lösung kann man mit Hilfe weiteren Strategien besser nachvollziehen. Zum Beispiel kann das Beispiel auf mehr Türen erweitert werden auf mehrere Türen. Dann stehen bei der Quizshow nicht nur 3 Türen zur Verfügung sondern gleich 10. Hier wird der Unterschied aufgrund der größeren Anzahl an Türen deutlicher. Bei einer Auswahl von 10 Türen liegt die Wahrscheinlichkeit, dass man beim ersten Versuch die richtige Tür, hinter welcher der Gewinn wartet, erraten hat, bei 1 : 10. Die Gegenwahrscheinlichkeit, dass sich der Preis hinter einer der anderen Türen befindet, beträgt hingegen 9 : 10. [5, S. 68]

Ursprünglich geht das Problem auf eine andere Quizshow zurück, deren Regeln sich etwas unterschieden haben. Und zwar hieß diese Sendung *Monty Hall Show*. Im Gegensatz zum bekannten Ziegenproblem traten in dieser Version zwei Personen an. Eine Kandidatin beziehungsweise ein Kandidat suchte sich eine Tür aus, hinter der eine Ziege versteckt war. Daraufhin musste dieser Mitspieler zurücktreten. Dem anderen Kandidaten wurde der Wechsel angeboten. Auch hier beträgt die Wahrscheinlichkeit, dass sich hinter einer der drei Türen ein Wagen versteckt $\frac{1}{3}$. Man kann nun folgende Situationen unterscheiden: [5, S. 69]

Die erste Option sieht vor, dass hinter der vom Kandidaten K_1 ausgewählten Tür, zum Beispiel T_1 , der Preis versteckt ist, dann muss der andere Kandidat K_2 zurücktreten. Die nächste Eventualität könnte beschreiben, dass sich der Gewinn hinter der Tür der Wahl des Kandidaten K_2 befindet. Man nehme an dies sei die Tür T_2 . In diesem Fall verlangt der Showmaster vom Kandidaten K_1 den Spielraum zu räumen. Wenn sich der Treffer hinter der Tür T_3 befindet, so hat der Quizmaster freie Wahl der Person, die ausscheidet. [5, S. 70]

Tabelle 5: Gewinnchancen Monty Hall Paradoxon [5, S. 70]

Gewinn steht hinter	Wahrscheinlichkeit	Wahrscheinlichkeit für den Verlust	Wahrscheinlichkeit insgesamt
T_1	$\frac{1}{3}$	$K_2, 1$	$\frac{1}{3}$
T_2	$\frac{1}{3}$	$K_1, 1$	$\frac{1}{3}$
T_3	$\frac{1}{3}$	$K_2, \frac{1}{2}$	$\frac{1}{6}$
		$K_1, \frac{1}{2}$	$\frac{1}{6}$

Die Wahrscheinlichkeit, dass Kandidat K_1 ausscheidet, beträgt $\frac{1}{3}$. Zusätzlich gilt noch, dass der dritte Fall mit einer Wahrscheinlichkeit von $\frac{1}{2}$ eintreten kann. Daher $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$ und insgesamt ergibt dies $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$. Wenn der Kandidat K_1 den ersten Durchgang überstanden hat und in die nächste Partie aufsteigt, bleiben noch die erste genannte oder die letzte mögliche Variante übrig. Die erste Version mit einer Wahrscheinlichkeit von 1 : 3 und die andere mit 1 : 6 zu beachten ist. Die unterschiedlichen Möglichkeiten werden in Tabelle 5 dargestellt.

6.4 Verwechslung bei der bedingten Wahrscheinlichkeit

In den vorigen Abschnitten der Arbeit konnte man bereits ein einführendes Beispiel zur charakteristischen Verwechslung lesen. In diesem weiter ausführenden Kapitel soll diese Vertauschung der Bedingung und des zu bedingenden Sachverhaltes erläutert und anhand von bekannten Fällen aus den Medien exemplarisch belegt werden. Denn dieser Fehler begleitet sowohl die Schülerinnen und Schüler als auch später das öffentliche Leben.

Auch wenn man den Eindruck hat, dass die einfachen Wahrscheinlichkeiten ein äußerst plausibles und unkompliziertes Thema sind. Diese Vorstellung gewinnt man vor allem dann, wenn man an das Würfelbeispiel zurückdenkt. Dennoch erweist sich die Berechnung mit Hilfe der bedingten Wahrscheinlichkeit als eine der irreführendsten Ermittlungen in der Wahrscheinlichkeitstheorie. Denn man fügt noch eine weitere bekannte Information hinzu und dies führt dann oft zu einem Missverständnis des Sachverhaltes. Aus diesem Grund sammeln sich sowohl bei der Kalkulation als auch bei der Interpretation immer wieder Fehler an. [2, S. 145]

Äußerst große Aufmerksamkeit sollte man der hinzukommenden Auskunft widmen. Denn diese beeinflusst die weitere Berechnung der Wahrscheinlichkeit. Gerade dieses Feingefühl

der bedingten Wahrscheinlichkeit machen sich viele Betrüger auf Jahrmärkten mit ihren Spielen zu nutzen. [2, S. 146]

Aber es gibt auch andere Fehler, bei denen die bedingte Wahrscheinlichkeit nicht zum Vorteil genutzt wird, sondern zu Fehlentscheidungen führt, wie im Falle des amerikanischen Football-Spielers O. J. Simpson. Er wurde wegen des Verdachts, seine Frau umgebracht zu haben, angeklagt. Belastend für ihn sprach die Tatsache, dass er seine Frau geschlagen hatte. Laut der amerikanischen Kriminalstatistik weiß man, dass die Wahrscheinlichkeit von $P(\text{Mann ermordet seine Frau} | \text{er hat sie geschlagen}) = \frac{1}{2500}$ beträgt. Aufgrund dieser, zwar richtig berechneten, aber in diesem Fall falsch verwendeten bedingten Wahrscheinlichkeit wurde Simpson nicht angeklagt. Der eigentliche Sachverhalt würde an dieser Stelle lauten

$P(\text{Mann schlägt seine Frau} | \text{Mann ermordet seine Frau}) = \frac{8}{9}$, denn seine Frau ja bereits tot ist, gehört dieses Ereignis in die Bedingung. [2, S. 153–154]

Dass die bedingte Wahrscheinlichkeit nicht nur Vorteile bezwecken kann, zeigt uns das Exempel der englischen Rechtsanwältin Sally Clark. Die Frau wurde wegen Verdacht des zweifachen Kindsmordes zu zweimal lebenslänglich verurteilt. Beide Söhne starben an einen plötzlichen Kindstod. Das Gericht hatte mit der geringen Wahrscheinlichkeit $P(\text{ihre zwei Kinder sterben} | \text{Mutter ist unschuldig})$ gerechnet. Ein solcher Fall sei als äußerst unrealistisch einzuschätzen, sodass die Mutter angeklagt wurde. Jedoch wurden hier wieder Bedingung und zu bedingender Sachverhalt verwechselt. Denn die gesuchte Wahrscheinlichkeit könnte mit $P(\text{Mutter ist unschuldig} | \text{ihre zwei Kinder sterben})$ berechnet werden. Das Gericht sah den Fehler nach einer bestimmten Zeit auch ein, und ließ die zu Unrecht verurteilte Frau wieder frei. Diese konnte aber nie wieder in ihr ursprüngliches Leben zurückkehren und starb bald daraufhin. [2, S. 156]

Die Angelegenheit von Sally Clark wurde von einem Arzt geprüft, der im Fall der Anwältin, welche ein gesundes und gutes Leben führte, argumentierte, dass die Wahrscheinlichkeit des Eintretens des plötzlichen Kindstodes bei 1 : 8500 liege. Der zuständige Arzt hatte aber nicht beachtet, dass ein Gen für den plötzlichen Kindstod verantwortlich sein kann. Somit beruht der Fehler des Arztes darauf, dass er die Todesfälle der zwei Söhne nicht abhängig voneinander betrachtete. So kam er auf eine Wahrscheinlichkeit von 1 : 72 250 000, da er $8\,500 \cdot 8\,500$ rechnete. Dies ist klarerweise ein äußerst unwahrscheinlicher Vorfall und deshalb wurde die Mutter auch schuldig gesprochen. [31, S. 45]

Worauf jedoch nicht Rücksicht genommen wurde, ist der Einsatz der Bayes-Regel bei dieser Situation. Die Wahrscheinlichkeit, zu dieser Zeit an plötzlichem Kindstod zu sterben, lag laut [31] bei 1 : 1 300. Da es sich bei dieser Todesursache um einen genetischen Fehler handelte, ist das Auftreten dafür innerhalb einer Familie auch höher. Und die Wahrscheinlichkeit, dass das zweite Kind ebenfalls an dem tragischen Schicksalsschlag stirbt, beträgt ebenfalls nach [31] 1 : 100. Die Wahrscheinlichkeit $P(H)$ der Annahme, dass die Söhne der Anwältin an plötzlichem Kindstod starben, sowie die gegenteilige Lage $P(\bar{H})$ berechnet man wie folgt:

$$P(H) = \frac{1}{1\,300} \cdot \frac{1}{100} \approx 0,000008,$$

$$P(\bar{H}) = 1 - P(H) \approx 0,999992. [31, S. 46]$$

$P(D|\bar{H}) = \frac{30}{650\,000} \cdot \frac{1}{10} \approx 0,000005$, beschreibt die Wahrscheinlichkeit dafür, dass zwei Kinder überraschend sterben, allerdings aus einem anderen Grund außer plötzlichem Kindstod. [31, S. 46]

Die gesuchte bedingte Wahrscheinlichkeit für diese Situation würde bedeuten, wie hoch die Wahrscheinlichkeit sei, dass Geschwister an plötzlichem Kindstod sterben unter der Bedingung, dass der Tod überraschend eintritt. Man schreibt

$$\begin{aligned} P(H|D) &= \frac{P(H) \cdot P(D|H)}{P(H) \cdot P(D|H) + P(\bar{H}) \cdot P(D|\bar{H})} \\ &= P(H|D) = \frac{1 \cdot 0,000008}{1 \cdot 0,000008 + 0,999992 \cdot 0,000005} \approx 0,6. [31, S. 46-47] \end{aligned}$$

Da die Vorstellung zwei Kinder hintereinander durch plötzlichen Kindstod zu verlieren, so unrealistisch und unwahrscheinlich wirkt, kann man leicht verunsichert werden und der falschen Rechtssprechung zustimmen. Allerdings, nach dem Satz von Bayes erkennt man, dass die Wahrscheinlichkeit für den Fall, dass die Kinder tatsächlich an plötzlichem Kindstod starben, bei unerwarteten 60% liegt.

Um zu zeigen, dass nicht einmal ausgebildete Mediziner vor solch einer Differenzierung gefeit sind, wird in diesem Absatz ein weiteres Beispiel zur Verwirrung zwischen der Bedingung und dem bedingten Ereignis dargeboten.

Beispiel: Der Autor Gigerenzer stellte einem Arzt folgende Frage zur bedingten Wahrscheinlichkeit. Die Wahrscheinlichkeit, dass eine Frau Brustkrebs hat, beträgt 0,8%. Wenn eine Frau Brustkrebs hat, beträgt die Wahrscheinlichkeit 90%, dass ihr Mammogramm ein positives Ergebnis aufweist. Wenn eine Frau aber keinen Brustkrebs hat, gibt es dennoch eine 7%ige Wahrscheinlichkeit, dass ihr Mammogramm positiv zu sein scheint. Die Frage ist nun, wie hoch die tatsächliche Wahrscheinlichkeit ist, dass eine Frau Brustkrebs hat, wenn ihr Mammogramm positiv ausfällt? [15, S. 65]

Der befragte Arzt konnte keine richtige Lösung nennen und schätzte dann einfach, dass sich die Wahrscheinlichkeit bei 90% liegen müsse, wobei er aber immerhin den Verdacht hatte, mit dieser Antwort falsch zu liegen.

Dabei ist die Frage ganz einfach nach der Formel der bedingten Wahrscheinlichkeit beziehungsweise nach dem Satz von Bayes zu lösen. Nach der Angabe kann man sich schon ein gutes Bild der einzelnen Wahrscheinlichkeiten machen.

$P(B) = 0,008$, für die Wahrscheinlichkeit, dass man Brustkrebs hat;

$P(\bar{B}) = 0,992$, entspricht der Gegenwahrscheinlichkeit, dass man nicht erkrankt ist;

$P(\text{Mammogramm positiv}|B) = 0,9$, repräsentiert die Chance, dass man Brustkrebs bei der Untersuchung erkennt;

$P(\text{Mammogramm positiv}|\bar{B}) = 0,07$, ist die Chance dafür, dass man ein positives Ergebnis bekommt, obwohl man keinen Brustkrebs hat. [15, S. 71]

Mit diesen Daten kann man sich die fragende Wahrscheinlichkeit nach der Formel berechnen.

$$\begin{aligned}
 &P(B|\text{Mammogramm positiv}) \\
 &= \frac{P(B) \cdot P(\text{Mammogramm positiv}|B)}{P(B) \cdot P(\text{Mammogramm positiv}|B) + P(\bar{B}) \cdot P(\text{Mammogramm positiv}|\bar{B})} \\
 &= \frac{0,008 \cdot 0,9}{0,008 \cdot 0,9 + 0,992 \cdot 0,07} \approx 0,09. [15, S. 72]
 \end{aligned}$$

Das heißt, die Wahrscheinlichkeit, dass eine Frau wirklich an Brustkrebs erkrankt ist, wenn sie ein positives Ergebnis bekommt, beträgt rund 9%. Ein anschaulicher Lösungsversuch wird ebenfalls im späteren Kapitel Baumdiagramm repräsentiert. Dort wird auch auf die Hilfe von relativen Häufigkeiten bei der Vorstellung näher eingegangen.

Ein Grund für diese Fehlinterpretationen und falschen Rechnungen soll laut Kahneman die so genannte Repräsentationsheuristik sein. [2, S. 156] Dieser Terminus beschreibt die Klassifizierung von Stereotypen, die jeder Mensch im Unterbewusstsein hat und die, ohne großes Überlegen, im Leben den Alltag beeinflussen. Eine klischeehafte Meinung kann zwar das tägliche Tun und Handeln erleichtern, aber sie bringt nicht nur Vorteile mit sich, wie man in den obigen Beispielen erkennen konnte. In Schubladen zu denken beeinflusst den menschlichen Verstand in jeglicher Hinsicht und führt so zu Irrtümern und Fehlvorstellungen.

Es gibt noch unendlich viele solcher Fehlberechnungen, wie zum Beispiel, dass Fußball der gefährlichste Sport sei, weil es hier die meisten Verletzungen gäbe. Oder, dass Schäferhunde die aggressivste Hunderasse seien, weil die meisten Bissvorfälle auf sie zurückzuführen seien. [20, S. 139–140] Genau solche Aussagen findet man regelmäßig in den Medien. Man sollte solchen Schlagzeilen im Allgemeinen aber nicht zu viel Glauben schenken, denn wenn man diese Angaben akzeptiert, kommt es genau zu den typischen Verwechslungen. Sonst käme man noch zu dem Entschluss, dass man beim Autofahren den Gurt lieber weglasse, um für seine Sicherheit zu garantieren. Wie es in einem Bericht der ADAC-Motorwelt zu lesen war: Vier von zehn tödlich verunglückten Autofahrern trugen keinen Sicherheitsgurt. [20, S. 137] Bei genauerem Betrachten dieses Titels bedeutet die Aussage, dass sechs Personen, trotz Anschnallen des Sicherheitsgurtes, tödlich umkamen. [20, S. 137] Aber auch in diesem genannten Beispiel liegt der Fehler wieder in der Verwechslung zwischen Bedingung und des dadurch bedingten Ereignisses. [20, S. 138] Die hier gefragte Wahrscheinlichkeit heißt nicht $P(\text{fahren ohne Gurt}|\text{tödlicher Unfall})$, sondern eigentlich $P(\text{tödlicher Unfall}|\text{fahren mit Gurt})$.

Man kann nur nochmals daran appellieren, sich zuerst klarzumachen, bei welchem Ereignis es sich um die Bedingung handelt, sprich welcher Ausgang bereits bekannt beziehungsweise gegeben ist. Und welches die, auf die beschränkte Wahrscheinlichkeitsmenge, gesuchte Begebenheit ist.

6.5 Das Simpson–Paradoxon

Auslösend für das so genannte Simpson–Paradoxon war die Universität in Berkeley, Kalifornien. Denn die Universität stand unter Verdacht, Frauen zu diskriminieren, weil sie 44% der Männer, aber nur 35% der Frauen aufnahm. [20, S. 146] Dies wurde aber zu Unrecht vermutet, denn insgesamt nahm die Universität mehr Frauen auf als Männer. Die Täuschung lag darin, dass die Vereinigung von Teilmengen einen anderen Eindruck macht als die Gesamtmenge.

Man kann durchaus noch weitere Beispiele nennen, die das Simpson-Paradoxon verdeutlichen, da es auch in vielen Alltagssituationen vorkommt. Bei einem Elfmeterschuss muss sich der Trainer zum Beispiel entscheiden, welcher Spieler die höhere Trefferquote hat. Spieler 1 hat zwar eine bessere Bilanz beim Treffen, dennoch spielt Spieler 2 einfach besser Fußball. [20, S. 145]

Oft kommt es auch zu Verwirrungen, da sich der menschliche Verstand hinter dem Simpson–Paradoxon Ursache und Wirkung wünscht. [20, S. 147] So wie zum Beispiel: Aus der Bedingung B folgt das Ereignis A , was ein natürlicher Gedankengang leichter nachvollziehen kann. Der entscheidende Faktor, der an dieser Stelle verlangt wird, lautet Kausalität. Der kausale Zusammenhang, der zwischen der Bedingung und dem bedingten Ereignis zu bestehen scheint, ist der Anlass der zu der oft erlittenen Verwirrung führt.

6.6 Stochastische Unabhängigkeit

Wie der Name des Kapitels schon verrät, wird sich dieser Abschnitt der Arbeit mit der Unabhängigkeit von Ereignissen befassen. Im vorigen Teil der Arbeit, der über das Thema der bedingten Wahrscheinlichkeit handelte, definierte man

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

für das Eintreten des Experimentausgangs von B unter der Bedingung, dass das Ergebnis A bereits geschehen sei. Daraus schließt man, dass die Wahrscheinlichkeit des Durchschnitts der beiden Ereignismengen $A \cap B$ nicht das gleiche Resultat liefert wie die Wahrscheinlichkeit von $P(B)$. Das Gleiche kann man über den umgekehrten Sachverhalt behaupten, wenn die zu berechnenden Wahrscheinlichkeiten $P(A|B)$ und $P(A)$ gegeben sind. Wenn jedoch die Gleichungen $P(B|A) = P(B)$ beziehungsweise $P(A|B) = P(A)$ gültig wären, dann würde das Ereignis A das Ereignis B oder umgekehrt auf keine Weise beeinflussen. [18, S. 115] Wenn man statt $P(B|A) = P(B)$ und $P(A|B) = P(A)$ die Definitionen der bedingten Wahrscheinlichkeiten $\frac{P(B \cap A)}{P(A)}$ und $\frac{P(A \cap B)}{P(B)}$ umformt, so ist zu erkennen, dass die beiden Terme zu folgendem Ausdruck führen:

$$P(A \cap B) = P(A) \cdot P(B).$$

Dann nennt man diese Formel der zwei Ereignisse A und B stochastisch unabhängig. [18, S. 116]

Definition: Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Dann heißen die beiden Ereignisse A und B stochastisch unabhängig, wenn $P(A) \cdot P(B) = P(A \cap B)$ gilt. [10, S. 170] Vor allem ist die Definition auch für die beiden Fälle $P(A) = 0$ beziehungsweise $P(B) = 0$ gültig.

Interpretiert man diese Definition nun in einem verständnisvolleren Kontext, so beschreibt die stochastische Unabhängigkeit, wenn jeweils endlich viele dieser Ereignisse unabhängig untereinander sind. Dies zeichnet eine sehr nützliche Eigenschaft in der Stochastik aus. Denn auch wenn man aus der eigenen Erfahrung erkennt, dass Experimente unabhängig voneinander ablaufen, lässt man sich dennoch leicht durch bereits bekannte Ergebnisse beeinflussen.

Angenommen ein Zufallsexperiment bestünde aus dem Werfen von zehn gleichen Münzen. Neun Münzen werden von einer Freundin oder einem Freund geworfen und nur das Ergebnis der zehnten Münze sieht man selbst. Nun behauptet das Gegenüber, dass sie beziehungsweise er neun Mal Zahl hatte. Intuitiv weiß man, dass die Wahrscheinlichkeit für ein weiteres Mal Zahl gleich $\frac{1}{2}$ beträgt, da jeder Wurf unabhängig von dem Wurf davor ist. Trotzdem wird man durch die Erfahrung verunsichert und bewertet die Experimente abhängig voneinander. [27, S. 40–41]

Definition: Es seien $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A_1, A_2, \dots, A_n \in \mathcal{A}$ für ein $n \in \mathbb{N}$. Die Ereignisse heißen unabhängig, falls für jedes $1 \leq m \leq n$ und jede Teilmenge $\{i_1, i_2, \dots, i_m\} \subseteq \{1, 2, \dots, n\}$ gilt $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_m})$. [27, S. 42] Bei unendlich vielen Ereignissen nennt man diese unabhängig, wenn jeweils endlich viele dieser Ereignisse unabhängig sind. [3, S. 103]

6.7 Baumdiagramm

Als Baumdiagramm wird eine Darstellungsform in der Stochastik bezeichnet, die mit Ästen die einzelnen Elementarereignisse anschaulich wiedergibt. Es erlaubt so einen leichten und transparenten Ansatz für das Verständnis des Berechnens von Wahrscheinlichkeiten. Vor allem die Beschreibung von mehreren ausgeführten Experimenten durch Baumdiagramme bietet viele Vorteile, wobei man dann von einem mehrstufigem Zufallsexperiment spricht. Diese Idee einer Skizzierung von Wahrscheinlichkeiten stammt von Christiaan Huygens. [10, S. 173]

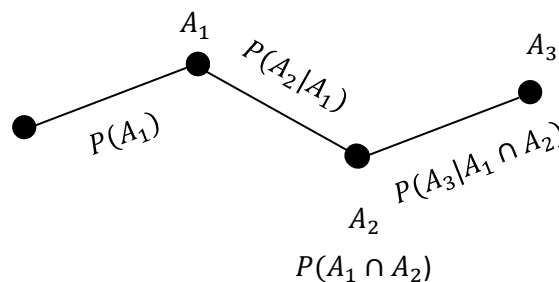


Abbildung 16: Allgemeines Baumdiagramm [18, S. 175]

In der oben stehenden Abbildung 16 wird ein allgemeines Baumdiagramm mit den einzelnen Wahrscheinlichkeiten präsentiert. Dabei erklären die gezeichneten Pfade jeweils die eingetretenen Ereignisse. Bis zum ersten Punkt, welcher das Eintreten des Ereignisses A_1 darstellt, wird die Wahrscheinlichkeit $P(A_1)$ angezeigt. Der zweite Knoten repräsentiert den möglichen Ausgang A_2 unter der Bedingung, dass A_1 bereits eingetreten ist. Und das letzte Ereignis A_3 zeigt die abhängige Wahrscheinlichkeit dafür, dass A_3 genau dann eintritt, wenn die beiden vorherigen Ereignisse A_1 und A_2 passiert sind. Insgesamt wird durch dieses Baumdiagramm die Wahrscheinlichkeit des Ereignisses $A_1 \cap A_2 \cap A_3$ illustriert. [18, S. 175]

Um ein Baumdiagramm anhand eines schon bekannten Beispiels anschaulich zu erklären, wird die Lösung des Ziegenproblems an dieser Stelle mit Hilfe von Pfaden visualisiert. Das nachstehende Baumdiagramm in der Abbildung 17 erklärt einen weiteren Lösungsversuch.

Der erste Ast symbolisiert die erste Entscheidung der Kandidatin beziehungsweise des Kandidaten, die mit einer Wahrscheinlichkeit von $\frac{1}{3}$ gefällt wird. Denn hinter den möglichen drei Türen besteht nach dem Laplace-Modell, $\frac{|\text{günstige}|}{|\text{mögliche}|}$, eine günstige Wahl, hinter der das Auto wartet. Angenommen die Moderatorin oder der Moderator öffnet dann eine Tür, hinter der nicht der Gewinn, sondern eine Ziege steht. Nun gibt der Quizmaster die Möglichkeit die Entscheidung zu überdenken und bietet einen Wechsel an. Da eine Tür bereits geöffnet wurde, gibt es nur mehr zwei mögliche Fälle, den Gewinn zu erraten. Daher ergibt sich die Wahrscheinlichkeit $\frac{1}{2}$.

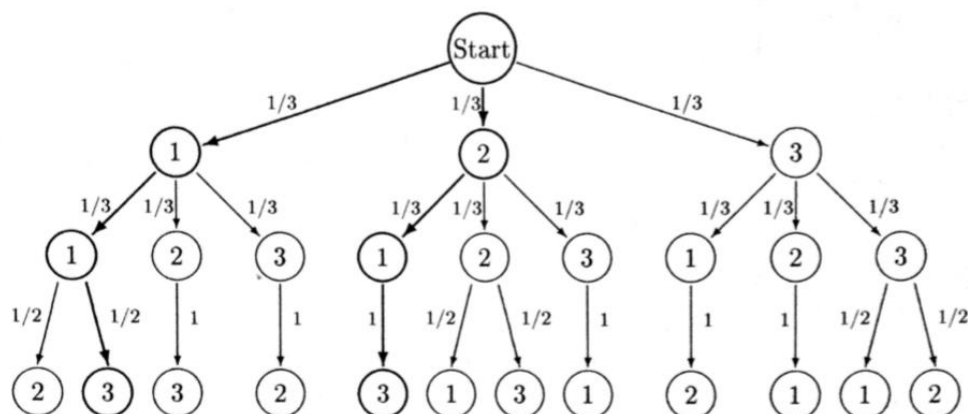


Abbildung 17: Ziegenproblem anhand eines Baumdiagramms [18, S. 105]

Ein etwas anderes Modell eines Baumdiagramms liefert die Abbildung 18. Der Unterschied zum vorherigen Baumdiagramm sind die beschrifteten Daten, die diese Grafik kennzeichnen. Handelte es sich beim Ziegenbeispiel um Wahrscheinlichkeiten, geben die Äste an diesem Exempel die absoluten Häufigkeiten an. Das Beispiel, welches im vorigen Kapitel zur Berechnung bedingte Wahrscheinlichkeiten mit dem Risiko an Brustkrebs zu erkranken, besprochen wurde, wird dieses Mal mit Hilfe absoluter Häufigkeiten und anhand eines Baumdiagramms demonstriert. Eine solche Vorstellungsart ist besonders für praktische Berufe, wie Ärzte oder Juristen ratsam, da man durch die absolute Häufigkeit einen verständlicheren Eindruck erlangt als mit Prozentangaben beziehungsweise Brüchen.

Zur Abbildung 18 sei gesagt, dass die notwendigen Zahlen jeweils vom linken Ast repräsentiert werden. Für diese Fragestellung beachtet man die fett dargestellten Äste. Bei 8 erkrankten Frauen identifiziert der Test 7 mit einem positiven Ergebnis. Ebenso bekommen bei 992 nicht erkrankten Frauen rund 70 Frauen ein positives Ergebnis, obwohl sie gesund sind. Die Problemstellung lautete wie hoch die Wahrscheinlichkeit sei, dass man krank ist unter der Bedingung, dass man ein positives Ergebnis erhalten hat. Das heißt 77 Frauen haben ein positives Ergebnis bekommen, und unter ihnen gibt es 7, die tatsächlich krank sind. Also

$$P(\text{Brustkrebs}|\text{Mammogramm positiv}) = \frac{7}{7 + 70} \text{ [15, S. 72]}$$

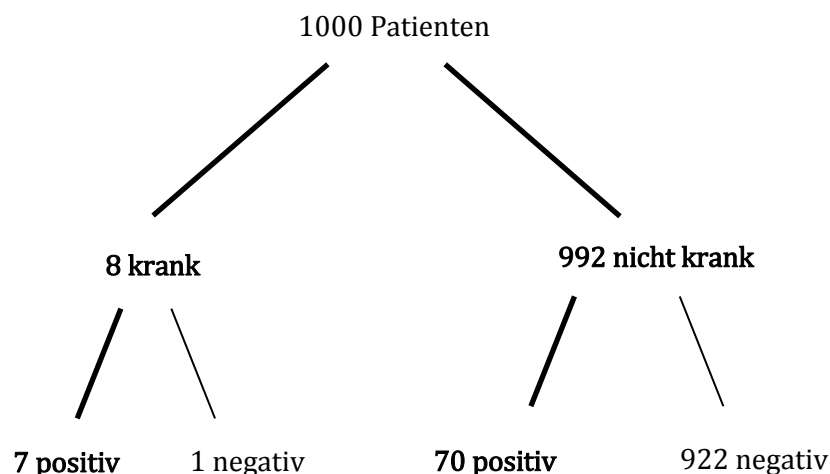


Abbildung 18: Baumdiagramm absolute Häufigkeiten [15, S. 72]

6.7.1 Pfadregeln

Die Pfadregel stellt eine äußerst einfache, aber gleichzeitig auch hilfreiche Rechenmethode von Wahrscheinlichkeiten dar. Vermutlich, weil sie eine Kombination aus der Formel der bedingten Wahrscheinlichkeit und dem dritten Axiom von Kolmogorov darstellt. [10, S. 179] Und einen weiteren Vorteil stellt die erfassbare Übersicht von mehrstufigen Experimenten dar. Dadurch, dass sie so verständlich und anschaulich den Sachverhalt eines stochastischen Zusammenhangs repräsentieren, können ebenso die Schülerinnen und Schüler ebenso problemlos die Strategie der Pfadregeln anwenden. Nichts desto trotz muss man diese auch klar formulieren.

Pfadregel: In einem mehrstufigen (mit mehreren Zufallsversuchen) Zufallsexperiment ist die Wahrscheinlichkeit eines ω das Produkt der Wahrscheinlichkeiten entlang seines Pfades im zugehörigen Baumdiagramm. [27, S. 47]

Beweis: Der Beweis folgt aus der bedingten Wahrscheinlichkeit. Man multipliziere die Wahrscheinlichkeiten entlang eines Pfades $\omega_{i_1}^{(1)}, \omega_{i_2}^{(2)}, \dots, \omega_{i_n}^{(n)}$:

$$[\omega_{i_1}^{(1)}] [\omega_{i_2}^{(2)}] [\omega_{i_1}^{(1)}] [\omega_{i_3}^{(3)} | \omega_{i_1}^{(1)}, \omega_{i_2}^{(2)}] \cdot \dots \cdot [\omega_{i_k}^{(k)} | \omega_{i_1}^{(1)}, \dots, \omega_{i_{k-1}}^{(k-1)}] \cdot \dots \cdot [\omega_{i_n}^{(n)} | \omega_{i_1}^{(1)}, \dots, \omega_{i_{n-1}}^{(n-1)}]$$

Da das Produkt des ersten Faktors dem Produkt des zweiten Faktors entspricht

$$P(\{\omega \in \Omega | \omega_j = \omega_{i_j}^{(j)} \text{ für } 1 \leq j \leq 2\}),$$

daraus ergibt sich auch, dass das Produkt des dritten Faktors gleich ist

$$P(\{\omega \in \Omega | \omega_j = \omega_{i_j}^{(j)} \text{ für } 1 \leq j \leq 3\})$$

und so weiter. Bis man erkennt, dass jeder Faktor dem anderen entspricht

$$P(\{\omega \in \Omega | \omega_j = \omega_{i_j}^{(j)} \text{ für } 1 \leq j \leq n\}) = P(\{\omega_{i_1}^{(1)}, \dots, \omega_{i_n}^{(n)}\}). [27, S. 47]$$

Die folgende Abbildung 19 zeigt ein übliches Baumdiagramm mit den Wahrscheinlichkeiten $\omega_1, \omega_2, \dots, \omega_k$, wenn man nun die gegebenen Wahrscheinlichkeiten berechnen will, dann muss man die Werte entlang eines Pfades beachten. Dies führt uns dann sowohl zur ersten als auch zur zweiten Pfadregel.

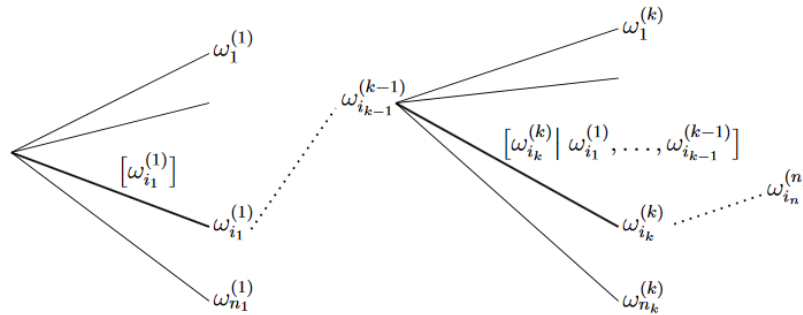


Abbildung 19: Baumdiagramm Pfadregel [27, S. 46]

1. Pfadregel (Multiplikationsregel): Die Wahrscheinlichkeit eines zusammengesetzten Ergebnisses ist gleich dem Produkt der Wahrscheinlichkeiten entlang des Pfades, der zu dem Ergebnis führt. [21, S. 135]

Beispiel: Bei einem Spiel darf sich eine Frau beziehungsweise ein Mann zwischen zwei Toren, Tor 1 und Tor 2, entscheiden. Hinter diesen Pforten haben sich ein roter, ein gelber und ein blauer Ring versteckt. Die beteiligte Person zieht dabei keine der beiden Türen vor. Mit welcher Wahrscheinlichkeit wählt sie den gelben Ring hinter dem Tor 2? [21, S. 134]

In der Abbildung 20 werden die folgenden möglichen Ereignisse anschaulich abgebildet. Für die Lösung des Beispiels folgt man den dicken Pfaden. Daher ist das Ergebnis nach der ersten Pfadregel:

$$P(2. \text{ Tor und gelber Ring}) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}. [21, \text{ S. 136}]$$

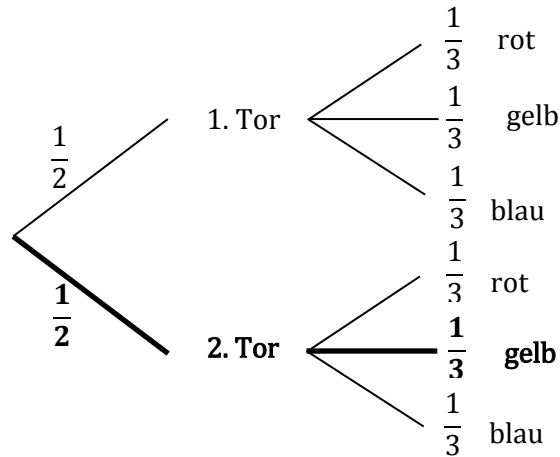


Abbildung 20: 1. Pfadregel anhand eines Beispiels [21, S. 135]

2. Pfadregel (Additionsregel): Wenn ein Ereignis aus mehreren zusammengesetzten Ergebnissen besteht, dann ergibt sich die dazugehörige Wahrscheinlichkeit aus der Summe der Wahrscheinlichkeiten dieser Ergebnisse. [21, S. 138]

Beispiel: Ein Möbelhaus verkauft einen Kasten, der eingepackt von den Kundinnen und Kunden selbst genommen wird. Dieser Vorgang entspricht einem Zufallsexperiment mit Ziehen ohne Zurücklegen. Es ist bekannt, dass rund 1% der Einzelteile des Schrankes ein fehlerhaft gebohrtes Loch haben, ungefähr 2% der Stücke einen Mangel am Lack aufweisen und circa 4% der nicht maßgetreu sind. Nun stellt sich die Frage, wie hoch die Wahrscheinlichkeit ist, dass die gewählte Ware genau einen Fehler aufweist? [21, S. 137]

Die Abbildung 21 erklärt das Beispiel anhand der zweiten Pfadregel. Die gesuchte Wahrscheinlichkeit, dass man bei dem gekauften Produkt einen Fehler entdeckt, wird berechnet, indem man die Einzelwahrscheinlichkeiten jedes Astes, der dick markiert wurde, zusammenzählt. Und nach der folgenden Rechnung wird mit einer Wahrscheinlichkeit von 7% ein Kasten mit mindestens einem Fehler gekauft.

$$P(\text{gekauft Produkt hat mindestens einen Fehler}) = 0,01 + 0,02 + 0,04 = 0,07$$

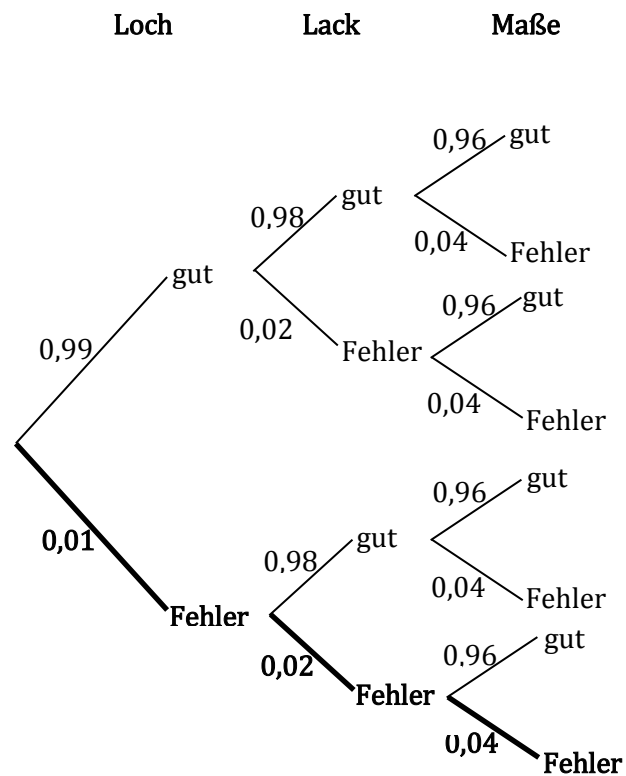


Abbildung 21: 2. Pfadregel anhand eines Beispiels [21, S. 139]

7. Zufallsvariable

Beim Namen der Zufallsvariable lässt man sich gerne von der Intuition leiten und denkt an Zufälle. Allerdings hat die Zufallsvariable aus mathematischer Sicht nur wenig Zusammenhang mit dem Zufall. Wenn man von Zufallsvariablen spricht, so assoziiert man damit Funktionen, die Elementen aus der Ergebnismenge eine reelle Zahl zuordnen. Zufallsvariablen bieten somit eine Hilfestellung zum Rechnen von Wahrscheinlichkeiten. [10, S. 229]

Definition: Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, dann nennt man eine Abbildung die $X: \Omega \rightarrow \mathbb{R}^n \forall \alpha \in \mathbb{R}: \{\omega \in \Omega, X(\omega) \leq \alpha\} \in \mathcal{A}$ erfüllt, eine Zufallsvariable. [11, S. 188]

Dabei wird bei Zufallsvariablen generell noch zwischen diskreten und stetigen Zufallsvariablen differenziert. Es handelt sich um eine diskrete Zufallsvariable, wenn es sich um eine diskrete Menge handelt, sprich wenn X nur endlich oder abzählbar unendlich viele Werte annehmen kann. [35, S. 113] Man spricht von einer stetigen Zufallsvariable, falls es eine Funktion f , sodass $P(X \leq \alpha) = \int_{-\infty}^{\alpha} f(t) dt \forall \alpha \in \mathbb{R}$ gilt. [35, S. 113]

Angenommen wir haben zwei herkömmliche Würfel mit den jeweiligen Augensummen 1 bis 6. Würfel A habe die Ergebnisse a und für Würfel B bezeichnen wir die möglichen Ausgangsfälle als b . Die Ergebnismenge laute dann

$$\Omega = \{(a, b): a, b \in \{1, 2, 3, 4, 5, 6\}\}.$$

Man kann dann die möglichen Augensummen der beiden Würfelwurfergebnisse schreiben als

$$X(\omega) = a + b, \omega = (a, b).$$

So versteht man in diesem Einstiegsbeispiel die Zufallsvariable X als Augensumme der zwei Würfel. Die möglichen Varianten für die Zufallsvariable können $\{2, 3, \dots, 11, 12\}$ sein, wobei nicht jede Augensumme mit der gleichen Wahrscheinlichkeit auftritt. Zum Beispiel kommt die Augensumme 2 mit den Varianten $\{1, 1\}$, während die Augensumme 5 durch vier denkbare Ereignisse möglich ist, $\{2, 3\}, \{3, 2\}, \{1, 4\}, \{4, 1\}$. Das führt zur Folgerung, dass X nichts über das Ergebnis des Zufallsexperimentes aussagen kann, weil die Zufallsvariable X keinen merklichen Unterschied zwischen den Ausgängen ω mit gleicher Augensumme $X(\omega)$ macht. Das heißt $X(\omega)$ beschreibt die möglichen Ergebnisse des Würfelwurfes. Diese werden folgendermaßen formuliert

$$\{X = k\} = \{\omega \in \Omega: X(\omega) = k\}.$$

Werden der Zufallsvariable X endlich unterschiedliche Werte $X(\omega)$ zugeordnet, so spricht man von einer endlichen Menge. [18, S. 12–13] Und man nennt diese Zufallsvariable $X(\omega)$, die nur endlich viele beziehungsweise abzählbar-unendlich viele Werte annehmen kann, eine diskrete Zufallsvariable. Eine weitere wichtige Funktion, die bereits von der Verteilungsfunktion für Datenreihen bekannt ist, lässt sich leicht in den Bereich der Zufallsvariablen übernehmen. [10, S. 232]

Definition: Die Funktion

$$f_x(x) = \begin{cases} P(X = x) & \text{für } x = x_i, i = 1, 2, \dots \\ 0 & \text{sonst.} \end{cases}$$

heißt Wahrscheinlichkeitsfunktion der diskreten Zufallsvariablen X . Diese teilt jeder Wahrscheinlichkeit einen möglichen Wert der Zufallsvariable zu. [35, S. 114]

Eigenschaften:

$$1. 0 \leq f_x(x) \leq 1 \text{ (Wahrscheinlichkeit),} \quad 2. \sum_{i=1}^{\infty} f_x(x_i) = 1 \text{ (normiert). [35, S. 114]}$$

Definition: Es sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum und Z eine Zufallsvariable auf Ω . Dann heißt die durch

$$F: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto F(x) := P(Z \leq x) = \sum_{\substack{a \in Z(\Omega) \\ a \leq x}} P(Z = a)$$

definierte Funktion F die Verteilungsfunktion von Z . Es ist bekannt, dass es sich bei einer Verteilungsfunktion um eine Treppenfunktion handelt. [10, S. 232]

Eigenschaften:

$$1. 0 \leq F_Z(x) \leq 1,$$

$$2. \lim_{x \rightarrow \infty} F_Z(x) = 1,$$

$$3. \lim_{x \rightarrow -\infty} F_Z(x) = 0,$$

$$4. F_Z(x_1) \leq F_Z(x_2), \text{ falls } x_1 < x_2, \text{ (monoton nicht abnehmend),}$$

$$5. \lim_{\varepsilon \rightarrow 0} F_Z(x + \varepsilon) = F_Z(x), \varepsilon > 0, \text{ (rechtsseitig stetig). [35, S. 114–115]}$$

7.1 Rechenregeln von Zufallsvariablen

Einfach ausgedrückt kann man mit Zufallsvariablen arbeiten wie mit gewöhnlichen Zahlen. Man kann die Summe, die Differenz und das Produkt von zwei Zufallsvariablen bilden. Darüber hinaus kann man auch ein Vielfaches von Zufallsvariablen bilden.

Definition (Summe): $(X + Y)(\omega) = X(\omega) + Y(\omega), \omega \in \Omega$. [18, S. 14]

Definition (Differenz): $(X - Y)(\omega) = X(\omega) - Y(\omega), \omega \in \Omega$. [18, S. 14]

Definition (Produkt): $(X \cdot Y)(\omega) = X(\omega) \cdot Y(\omega), \omega \in \Omega$. [18, S. 14]

Definition (Vielfaches): $(a \cdot X)(\omega) = a \cdot X(\omega), \omega \in \Omega$. [18, S. 14]

7.2 Erwartungswert, Varianz und Standardabweichung einer Zufallsvariable

Um den Erwartungswert $E(Z)$ einer Zufallsvariable zu erhalten, muss man die relativen Häufigkeiten für die Berechnung des arithmetischen Mittels einfach durch die Wahrscheinlichkeiten der Zufallsvariable austauschen. [10, S. 235] Man kann sagen, dass es sich beim Erwartungswert um das Analogon zum arithmetischen Mittel handelt. [10] Ebenso werden die Auffassungen zu der Varianz und der Standardabweichung übernommen. [10]

Definition: Es sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum und Z eine Zufallsvariable auf Ω . Dann heißt die Zahl

$$E(Z) := \sum_{x \in Z(\Omega)} P(Z = x) \cdot x$$

der Erwartungswert von Z . Und die definierten Zahlen

$$E(Z) := \sum_{x \in Z(\Omega)} P(Z = x) \cdot (x - E(Z))^2 \text{ und } \sigma_Z := \sqrt{V(Z)}$$

heißen Varianz und Standardabweichung von Z . [10, S. 236]

Für eine bessere Vorstellung werden dafür anhand eines demonstrativen Beispiels die nun eingeführten Begriffe Erwartungswert, Varianz und der Standardabweichung erläutert.

Ein typisches Roulettefeld besteht aus 37 Feldern, von denen 18 schwarz gefärbt sind und 18 die Farbe Rot haben. Zusätzlich ist ein grünes Feld vorhanden, welches der 0 entspricht. Die übrigen Zahlen sind von 1 bis 36 auf dem Spielfeld abgebildet. Spieler A setzt 10€ auf eine konkrete Zahl, zum Beispiel die 17. Spieler B hingegen legt 10€ Einsatz auf alle roten Felder. Die entsprechenden Wahrscheinlichkeiten für Spieler A und Spieler B sind, nach dem Laplace-Modell, welches sich für dieses Glücksspiel sehr eignet, $\frac{1}{37}$ beziehungsweise $\frac{18}{37}$. Die Investition von Spieler A von 10€ würde sich im Falle eines Treffers verfünfunddreißigfachen, die des Spielers B würde sich nur verdoppeln. Man berechne nun den Gewinn, der sich durch die Zufallsvariablen A und B beschreiben lässt. [10, S. 237]

$$P(A = 350) = \frac{1}{37}, P(A = -10) = \frac{36}{37}: E(A) = -10 \cdot \frac{36}{37} + 350 \cdot \frac{1}{37} = -\frac{10}{37}. [10, S. 237]$$

$$P(B = 10) = \frac{18}{37}, P(B = -10) = \frac{19}{37}: E(B) = -10 \cdot \frac{19}{37} + 10 \cdot \frac{18}{37} = -\frac{10}{37}. [10, S. 237]$$

Man sieht an den beiden Lösungen für Spieler A und Spieler B , dass sie dieselbe Gewinnerwartung von $-\frac{10}{37}$ haben. Das bedeutet, dass beide auf einen längeren Zeitraum einen Verlust erleiden würden. [10, S. 238] Als Nächstes wird die Standardabweichung berechnet:

$$\sigma(A) = \sqrt{\left(-10 + \frac{10}{37}\right)^2 \cdot \frac{36}{37} + \left(350 + \frac{10}{37}\right)^2 \cdot \frac{1}{37}} = \frac{2160}{37} \approx 58,4. [10, S. 238]$$

$$\sigma(B) = \sqrt{\left(-10 + \frac{10}{37}\right)^2 \cdot \frac{19}{37} + \left(10 + \frac{10}{37}\right)^2 \cdot \frac{18}{37}} \approx 10,0. [10, S. 238]$$

Wie man am Ergebnis sieht, liefert die Standardabweichung einen unterschiedlichen Wert. Dieser Unterschied deutet darauf hin, dass sowohl das Verlustrisiko als auch die Gewinnchancen von Spieler A höher einzuschätzen sind. [10, S. 238]

7.3 Verteilungen

In der Wahrscheinlichkeitsrechnung ist der Begriff der Verteilung im Hinblick auf die Zufallsvariablen unerlässlich. Jedoch gibt es nicht die einzig richtige Verteilung, sondern seit jeher wurden unterschiedliche Verteilungen je nach Frage und zugehörigem Experiment verwendet. In den folgenden Teilkapiteln werden einige ausgewählte Verteilungen und deren Anwendungen, die als wichtig gesehen werden, präsentiert und beschrieben. Ein wichtiges Entscheidungskriterium für die Wahl einer bestimmten Verteilung ist das zu der Wahrscheinlichkeit gehörige, entsprechend gewählte Zufallsexperiment. [10]

7.3.1 Binomialverteilung

Die Binomialverteilung kann man anhand des Experiments leicht erklären und somit auch relativ einfach erkennen. Für die Binomialverteilung benötigt man ein Zufallsexperiment mit zwei möglichen Ergebnissen, welches unabhängig voneinander unter denselben Umständen n -Mal auftreten. [10, S. 253] Da nur zwei Ausgangsmöglichkeiten des Zufallsexperiments erfüllbar sind, kann man die Menge Ω der Ergebnisse mit $\Omega = \{0, 1\}$ beschreiben, wobei 0 für das Nichteintreten und 1 für das Eintreten des Ergebnisses gedeutet werden kann. Das heißt, man schreibt die Wahrscheinlichkeiten entsprechend $P(1) = p$ und $P(0) = q = 1 - p$. [10, S. 253] Man spricht demnach von einem Bernoulli-Experiment. [10, S. 253] Wie schon erwähnt, wenn das Bernoulli-Experiment n -Mal ausgeübt wird und die einzelnen Experimente jeweils unabhängig voneinander durchgeführt werden, dann wird von einer Bernoulli-Kette der Länge n mit der Ergebnismenge $\Omega_n = \{0, 1\}^n$ gesprochen. [10, S. 253]

Die entsprechende Verteilung der Wahrscheinlichkeiten dieser eben genannten Bernoulli-Kette mit dem Ausmaß n nennt man Binomialverteilung. [10, S. 253] Als Begründer der Binomialverteilung und damit ausschlaggebend für die damit beginnende neuzeitige Stochastik gilt Jacob Bernoulli. Das Wort binomial lässt sich auf den lateinischen Ausdruck *binominis*, was übersetzt zweinamig bedeutet, zurückführen. Damit werden die zwei möglichen Ergebnisausgänge des Zufallsexperiments assoziiert. [10, S. 253]

Definition: Sei (Ω, P) ein Wahrscheinlichkeitsraum und Z eine Zufallsvariable auf Ω , die die Werte von $0, 1, 2, \dots, n$ annehmen kann. Dann heißt Z binomialverteilt mit den Parametern n und p , kurz auch $B(n, p)$, wenn es ein $p \in [0, 1]$ gibt mit

$$P_Z(k) = P(Z = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \text{ für } k = 0, 1, 2, \dots, n.$$

Die Verteilung P_Z heißt Binomialverteilung. [10, S. 254]

Kennzeichnend für die Binomialverteilung ist ihre Darstellung. Denn sie hat eine typische glockenkurvige Form, wie man in der nachstehenden Abbildung 22 erkennen kann. Die Maße von $B(n, p, k)$ wachsen monoton von $k = 0$ bis zu ihrem Maximum, welches sie bei ungefähr $k \approx n \cdot p$ besitzen, an. [10, S. 254]

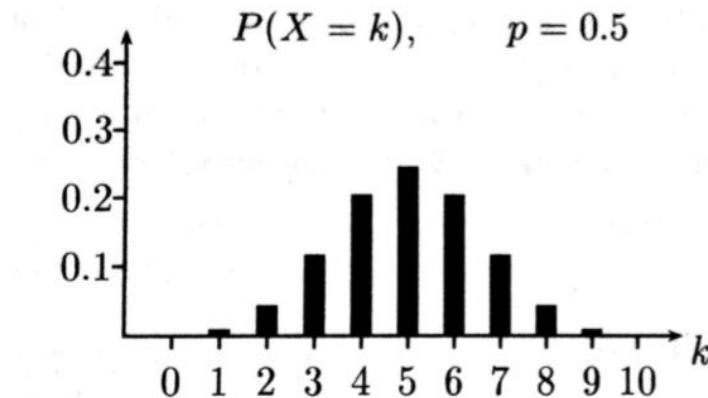


Abbildung 22: Glockenkurvige Darstellung einer Binomialverteilung mit unterschiedlichen n [18, S. 140]

Auch für die Binomialverteilung sind Erwartungswert und Varianz wichtige zu berechnende Kennzahlen.

Satz: Für eine $B(n, p)$ - verteilte Zufallsvariable Z gilt

$$E(Z) = n \cdot p$$

und

$$V(Z) = n \cdot p (1 - p). [10, S. 255]$$

Um einen verständlichen Einstieg in dieser Verteilung zu gewähren, werden die Beschreibung und die Anwendung binomialverteilter Zufallsvariablen anhand von Beispielen verdeutlicht.

Als erstes Beispiel gilt der klassische Wurf einer Münze und die dabei gesuchte Wahrscheinlichkeit vom Ereignis A , nämlich dass die Münze Zahl zeigt. Die dazugehörigen Wahrscheinlichkeiten schreibt man als $P(A) = p$ und $P(\bar{A}) = 1 - p$. Die dazugehörige Zufallsvariable zählt dabei die Anzahl des Eintretens des Ergebnisses A bei insgesamt n Wiederholungen. [10, S. 251] Nach der Pfadregel lautet die entsprechende Berechnung der Wahrscheinlichkeit für k -Mal Zahl dann

$$\left(\frac{1}{2}\right)^k \cdot \left(\frac{1}{2}\right)^{n-k} = \left(\frac{1}{2}\right)^n [10, S. 251]$$

Denn sowohl die Wahrscheinlichkeit, dass die Münze Zahl zeigt als auch die Gegenwahrscheinlichkeit, dass sie auf die andere Seite fällt, beträgt logischerweise $\frac{1}{2}$. Es ist möglich, dass die Münze k -Mal Zahl aufweist und die Ereignisse dabei an n verschiedenen

Positionen der Pfade liegen. Diese Überlegung führt zurück zu Kombinationen ohne Wiederholung. [10, S. 251] Und zwar handelt es sich um k -Kombinationen, die Zahl zeigen können. Dementsprechend liegen $\binom{n}{k}$ derartiger Wege vor. [10, S. 251] Und somit beträgt die gefragte Wahrscheinlichkeit mit Hilfe der zweiten Pfadregel

$$P(X = k) = \binom{n}{k} \cdot \left(\frac{1}{2}\right)^n \cdot [10, S. 251]$$

Ein weiterer charakteristischer Vertreter für eine Binomialverteilung ist das Experiment von Francis Galton aus England. Sein Interesse galt den Erbanlagen eines Menschen. Für seine Untersuchung, die von seinem Vater angeregt und beeinflusst wurde, analysierte er, wie sich die Fingerabdrücke von Menschen im Laufe des Lebens verändern. [10, S. 252]

Für seine Forschung entwickelte er das Galtonbrett. Man erkennt in Abbildung 23, dass in das Brett von oben Kugeln gegeben werden, welche durch die an dem Brett befestigten Nägel abgelenkt werden. Zufällig fallen die Bälle weiter nach rechts beziehungsweise mehr nach links in das unten abgeteilte Gefäß. Die Fallrichtungen der Kugeln können als mögliche Ergebnisse angesehen werden und als gleichwahrscheinlich betrachtet werden. Zusätzlich muss sich die Kugel n -Mal zwischen links oder rechts ausweichen entscheiden. Dies führt wieder zur geforderten Annahme, dass das Experiment binomialverteilt ist. [10]

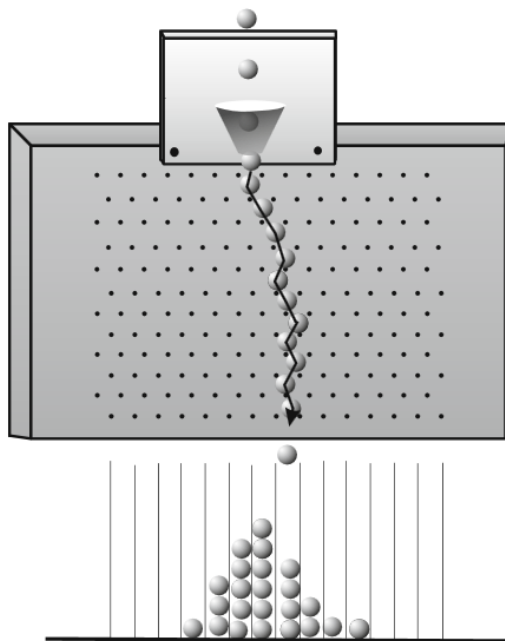


Abbildung 23: Galtonbrett [12, S. 22]

7.3.2 Multinomialverteilung

Von der Binomialverteilung ausgehend kann man die Multinomialverteilung deduzieren. Waren bei Binomialverteilungen nur zwei Ereignisse für den Ausgang eines Zufallsexperimentes denkbar, so sind bei der Multinomialverteilung, wie das Lexem multi schon verrät, viele Ereignisse möglich. Folglich besitzt die Ereignismenge $\Omega = \{1, 2, \dots, z\}$ viele Ergebnisse mit der dazugehörigen Wahrscheinlichkeit von $P_0(i) = p_i$ für

$i = 1, 2, \dots, z$. [10, S. 257] Wieder wird das Zufallsexperiment von n -Mal realisiert. Die passende Zufallsvariable der Verteilung zählt wieder, wie oft das Ereignis k bei n maliger Wiederholung verwirklicht wird. [10, S. 257]

Für die Definition lässt sich die Motivation, dass für den ersten Platz von k_1 noch alle n Möglichkeiten zur Verfügung stehen, erklären. Für die zweite Situation gibt es eine Gelegenheit weniger zur Auswahl. Sprich für die zweite Auswahl hat man nur mehr die Chance aus $n - k_1$ auszuwählen. Diese Struktur geht so lange weiter, bis man bei der letzten Stelle z angekommen ist, für die nur mehr ein Platz frei ist. In der Form von Binomialkoeffizienten lässt sich das wie folgt anschreiben

$$\binom{n}{k_1, k_2, \dots, k_z} := \binom{n}{k_1} \cdot \binom{n-k_1}{k_2} \cdot \dots \cdot \binom{n-k_1-k_2-\dots-k_{z-1}}{k_z} = \frac{n!}{k_1! \cdot k_2! \cdot \dots \cdot k_z!}. \quad [10, S. 257]$$

Definition: Wenn für alle z nicht negativen ganzen Zahlen $k_i, i = 1, 2, \dots, z$ mit $k_1 + k_2 + \dots + k_z = n$ gilt

$$P(X_1 = k_1 \wedge X_2 = k_2 \wedge \dots \wedge X_z = k_z) = \frac{n!}{k_1! \cdot k_2! \cdot \dots \cdot k_z!} \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot \dots \cdot p_z^{k_z},$$

dann heißt der Vektor $(X_1, X_2, \dots, X_z)$ von Zufallsvariablen multinomialverteilt mit den Parametern $n \in \mathbb{N}$, $p_1, p_2, \dots, p_z \in [0, 1]$ und $p_1 + p_2 + \dots + p_z = 1$. [10, S. 257]

Das Beispiel, welches die Multinomialverteilung am besten beschreiben kann, ist das Urnenbeispiel, wobei die Urne unterschiedliche Kugeln beinhaltet und aus ihr n -Mal mit Zurücklegen gezogen wird. [10, S. 257] Außerdem wird die Multinomialverteilung später auch noch beim Chi-Quadrat-Test eine wichtige Rolle spielen. [10, S. 258]

7.3.3 Hypergeometrische Verteilung

Die zwei bereits eingeführten Verteilungen sind Modelle für Zufallsexperimente mit Wiederholungen, die unabhängig voneinander vollzogen werden. Was tut man nun, wenn ein Experiment keine Wiederholungen erlaubt? Wie zum Beispiel beim Lotto-Spiel. Hier werden nacheinander aus einer Gesamtmenge nummerierte Kugeln gezogen, die nicht mehr zurückgelegt werden. Hierfür benutzt man dann die hypergeometrische Verteilung.

Definition: Eine Zufallsvariable Z in einem Wahrscheinlichkeitsraum (Ω, P) heißt hypergeometrisch verteilt mit den Parametern n, w und s , wobei $n, w, s \in \mathbb{N}$, wenn die Zufallsvariable Z nur Werte in $\{0, 1, 2, \dots, n\}$ annimmt und für $k \in \{0, 1, 2, \dots, n\}$ gilt

$$P(Z = k) = H(n, w, s; k) = \frac{\binom{w}{k} \cdot \binom{s}{n-k}}{\binom{w+s}{n}}. \quad [10, S. 259]$$

Man schreibt kurz $H(n, w, s)$ -verteilt.

Satz: Für eine $H(n, w, s)$ -verteilte Zufallsvariable Z gilt mit $p = \frac{w}{w+s}$

$$E(Z) = n \cdot p,$$

der Erwartungswert einer hypergeometrisch verteilten Zufallsvariable,

$$V(Z) = n \cdot p \cdot (1 - p) \cdot \frac{w + s - n}{w + s - 1},$$

die Varianz einer hypergeometrisch verteilten Zufallsvariable. [10, S. 259]

Beweis für den Satz vom Erwartungswert: Der Beweis ist sehr anschaulich anhand des altbekannten Urnen-Modell. Die Urne enthält sowohl weiße Kugeln, die mit $1, 2, \dots, w$ nummeriert sind, als auch schwarze Kugeln in der Größenanordnung von s . Im Falle, dass es sich bei der i -ten Kugel um eine weiße handelt, besitze die Zufallsvariable Z_i den Wert 1 für $i = 1, 2, \dots, n$, sonst den Wert 0. Da es sich um die gleiche Wahrscheinlichkeit handelt, eine weiße Kugel zu ziehen, beträgt die zu berechnende Wahrscheinlichkeit

$$P(Z_i = 1) = \frac{w}{w + s}.$$

Weiters gilt wegen $Z = Z_1 + Z_2 + \dots + Z_n$, $E(Z) = 1 \cdot \frac{w}{w + s} + 0 \cdot \left(1 - \frac{w}{w + s}\right) = p$ und $E(X + Y) = E(X) + E(Y)$, dass der Erwartungswert durch $E(Z) = n \cdot p$ geschrieben werden kann. [10, S. 259]

7.3.4 Geometrische Verteilung

Definition: Es sei X eine diskrete Zufallsvariable auf einem Wahrscheinlichkeitsraum (Ω, P) mit $X(\Omega) = \mathbb{N}$ heißt dann geometrisch verteilt mit dem Parameter $p \in (0, 1)$, wenn für die Wahrscheinlichkeitsverteilung

$$P(X = k) = (1 - p)^{k-1} \cdot p$$

gilt. [10, S. 260]

Satz: Der Erwartungswert für eine geometrisch verteilte Zufallsvariable X und den Parameter p wird durch

$$E(X) = \frac{1}{p},$$

und die dazugehörige Varianz durch

$$V(X) = \frac{1 - p}{p^2}$$

berechnet. [10, S. 261]

Beweis für den Satz vom Erwartungswert: Um einen Beweis führen zu können, benötigt man die Formel

$$\sum_{k=1}^{\infty} k \cdot a = \frac{a}{(1-a)^2},$$

wobei $|a| < 1$ gilt. Daher folgt dies den gesuchten Erwartungswert

$$E(X) = \sum_{k=1}^{\infty} k (1-p)^{k-1} \cdot p = \frac{p}{(1-p)} \cdot \sum_{k=1}^{\infty} k \cdot (1-p)^k = \frac{p}{1-p} \cdot \frac{1-p}{p^2} = \frac{1}{p}.$$

[10, S. 261]

Als Abschluss für die geometrische Verteilung fehlt noch ein klassisches Beispiel. Anhand dieses Beispiels kann man die Verteilung gut charakterisieren. Hierfür bietet sich das Würfelexperiment wieder an. Dabei stellt sich die Frage, wie lange man würfeln muss, bis alle 6 möglichen Seiten eines Würfels gefallen sind und sich jede Augensumme gezeigt hat. Als erstes legt man die Zufallsvariablen fest. X_1 sei die Zufallsvariable für die erste gefallene Augensumme z_1 . Dieser erste Zufallsvariable, die den Erwartungswert $E(X_1) = 1$ besitzt, nimmt dabei nur 1 an. Weiters gibt die Zufallsvariable X_2 eine Information über jeden zusätzlichen Wurf an, der nicht gleich dem Wurf z_1 ist. Daher beschreibt X_2 ein geometrisch verteiltes Ereignis mit der Wahrscheinlichkeit von $p_2 = \frac{5}{6}$ mit dem Erwartungswert von $E(X) = \frac{6}{5}$. Die nächste Zufallsvariable X_3 gibt Auskunft über jeden weiteren Wurf, der ungleich den anderen beiden erschienen Zahlen z_1 und z_2 ist. Bei dieser Zufallsvariable X_3 liegt erneut eine geometrische Verteilung vor, die die Wahrscheinlichkeit von $p_3 = \frac{4}{6}$ und den Erwartungswert $E(X_3) = \frac{6}{4}$ hat. Mit dieser Struktur folgt das System bis zur letzten Zufallsvariablen X_6 . Nun lautet die Formel für alle möglichen Augensummen

$$E(X) = E(X_1) + E(X_2) + \dots + E(X_6) = 1 + \frac{6}{5} + \frac{6}{4} + \dots + \frac{6}{1} = 14,7. [10, S. 261]$$

7.3.5 Poissonverteilung

Ein wichtiger Zusammenhang zur Binomialverteilung stellt die Poissonverteilung dar. Denn es handelt sich bei der erwähnten Verteilung um eine Approximation der Binomialverteilung. Speziell lässt sich diese Relation am gleichen Erwartungswert erkennen. [10, S. 261]

Ausschlaggebend für die Herleitung der Poissonverteilung war der Franzose Siméon Denis Poisson. Seine Entdeckung geht aber noch viel weiter zurück in die Geschichte, weil sein Modell bereits von Abraham de Moivre im Anfang des 18. Jahrhunderts benutzt wurde. Für Poisson galt das Problem der schlechten Approximation der Binomialverteilung an die Normalverteilung als der entscheidende Anlass für seine Arbeit im Werk *Recherches sur la probabilité des jugements en matière criminelle et in matière civile*. [10, S. 264]

Satz der Poisson-Näherung: Für $\mu \in \mathbb{R}$, $n \in \mathbb{N}$ und $p = \frac{\mu}{n}$ gilt

$$\lim_{n \rightarrow \infty} B(n, p, k) = \frac{\mu^k}{k!} e^{-\mu}. [10, S. 262]$$

Beweis: Wegen $E(X) = \mu = n \cdot p$, was $p = \frac{\mu}{n}$ bei der Binomialverteilung entspricht, gilt

$$\begin{aligned}
 B(n, p, k) &= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!} \cdot \frac{\mu^k}{n^k} \cdot \frac{\left(1 - \frac{\mu}{n}\right)^n}{\left(1 - \frac{\mu}{n}\right)^k} \\
 &= \frac{\frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \dots \cdot \frac{n-k+1}{n}}{\left(1 - \frac{\mu}{n}\right)^k} \cdot \frac{\mu^k}{k!} \cdot \left(1 - \frac{\mu}{n}\right)^n \\
 &= \frac{1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \dots \cdot \left(1 - \frac{k-1}{n}\right)}{\left(1 - \frac{\mu}{n}\right)^k} \cdot \frac{\mu^k}{k!} \cdot \left(1 - \frac{\mu}{n}\right)^n \\
 &= \underbrace{\frac{1}{1 - \frac{\mu}{n}} \cdot \frac{1 - \frac{1}{n}}{1 - \frac{\mu}{n}} \cdot \dots \cdot \frac{1 - \frac{k-1}{n}}{1 - \frac{\mu}{n}}}_{\substack{\rightarrow 1 \cdot 1 \cdot \dots \cdot 1 \\ n \rightarrow \infty}} \cdot \frac{\mu^k}{k!} \cdot \underbrace{\left(1 - \frac{\mu}{n}\right)^n}_{\substack{\rightarrow e^{-\mu} \\ n \rightarrow \infty}} \xrightarrow{n \rightarrow \infty} \frac{\mu^k}{k!} \cdot e^{-\mu}. [10, S. 263]
 \end{aligned}$$

Durch diese Annäherung kann man für große n und für kleine p die Binomialverteilung durch den Exponentialausdruck austauschen. Das bedeutet

$$B(n, p, k) \approx \frac{\mu^k}{k!} \cdot e^{-n \cdot p}. [10, S. 263]$$

Definition: Man nennt dann eine Zufallsvariable X mit den Werten auf $\mathbb{N}_0$ poissonverteilt mit dem Parameter $\mu > 0$, wenn

$$P(X = k) = \frac{\mu^k}{k!} \cdot e^{-\mu}$$

gilt. [10, S. 264]

Satz: Wenn die Zufallsvariable X poissonverteilt mit dem Parameter μ ist, dann gilt

$E(X) = \mu$ ist der dazugehörige Erwartungswert und

$V(X) = \mu$ entspricht der Varianz. [10, S. 264]

Beweis für den Satz vom Erwartungswert und der Varianz: Laut Definition ist

$$\begin{aligned}
 E(X) &= \sum_{k=0}^{\infty} k \cdot \frac{\mu^k}{k!} \cdot e^{-\mu} = e^{-\mu} \cdot \sum_{k=0}^{\infty} k \cdot \frac{\mu^k}{k!} \\
 &= e^{-\mu} \cdot \mu \cdot \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!} = e^{-\mu} \cdot \mu \cdot \sum_{k=0}^{\infty} \frac{\mu^k}{k!} = e^{-\mu} \cdot \mu \cdot e^{\mu} = \mu. [10, S. 265]
 \end{aligned}$$

Mit Hilfe des Verschiebungssatzes kann man dann auch die Varianz beweisen.

$$V(X) = E(X^2) - E(X)^2$$

$$E(X^2) = \sum_{k=0}^{\infty} k^2 \cdot \frac{\mu^k}{k!} \cdot e^{-\mu} = \sum_{k=1}^{\infty} k \cdot \frac{\mu^k}{(k-1)!} \cdot e^{-\mu}$$

$$= \mu \cdot \left[\underbrace{\sum_{k=1}^{\infty} (k-1) \cdot \frac{\mu^{k-1}}{(k-1)!} \cdot e^{-\mu}}_{E(X)=\mu} + \underbrace{\sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!} \cdot e^{-\mu}}_{e^{\mu}} \right]$$

$$\mu [\mu + e^{-\mu} \cdot e^{\mu}] = \mu^2 + \mu.$$

Daraus folgt aus $V(X) = \mu^2 + \mu - \mu^2 = \mu$. [10, S. 265]

Ein sehr anschauliches Beispiel, mit welchem die Poissonverteilung nahezu identisch beschreibbar ist, ist der Versuch von Ernest Rutherford und Hans Wilhelm Geiger im Jahre 1910. [18, S. 190] Dabei gingen die beiden Wissenschaftler einerseits von den α -Strahlen, die die radioaktive Strahlung eines Polynoms widerspiegeln, aus. Und andererseits differenzierten sie dazu die β -Strahlen. Bei dem radioaktiven Element, welches Rutherford und Geiger bearbeiteten, handelte es sich um einen α -Strahler. Die dabei abgegebenen Strahlen, die durch Lichtblitze verwirklicht wurden, wurden dann mit Hilfe eines Zinksulfidschirms erfasst. Für ihren Versuch konzentrierten sie sich, auf die in einem gewissen Zeitintervall produzierten Zerfälle. So entstanden folgende Erkenntnisse, wobei k der Betrag der entsprechenden Lichtblitze in einem Zeitintervall zu je 7,5 Sekunden ist und a_k der Menge der Intervalle mit exakt k Lichtblitzen entspricht. [10, S. 265]

Tabelle 6: Ergebniswerte zum Rutherford–Geiger Experiment [10, S. 265]

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	≥ 15
a_k	57	203	383	525	532	408	273	139	45	27	10	4	0	1	1	0

Um eine bessere Vorstellung von diesen Werten zu erlangen, wird die Situation im Folgenden in einem Diagramm in der Abbildung 24 präsentiert. Hier wird durch die empirischen relativen Häufigkeiten und das Gleichsetzen des arithmetischen Mittelwertes mit dem Erwartungswert die Poissonverteilung angenähert. [18, S. 191] Um sich langsam an das Poissonmodell zu nähern, wird zuerst die Unabhängigkeit zwischen dem Nichtzerfall und dem Zerfall in Betracht gezogen. Daraus folgert man die hypergeometrische Verteilung. Aufgrund der indirekten Proportionalität, die zwischen der Menge der Atome und der möglichen Zerfälle besteht, lässt sich die Binomialverteilung assoziieren. Denn es besteht eine große Zahl an Atomen, aber eine geringe Zerfallsanzahl, was durch die Binomialverteilung zur Poissonverteilung gelangen lässt.

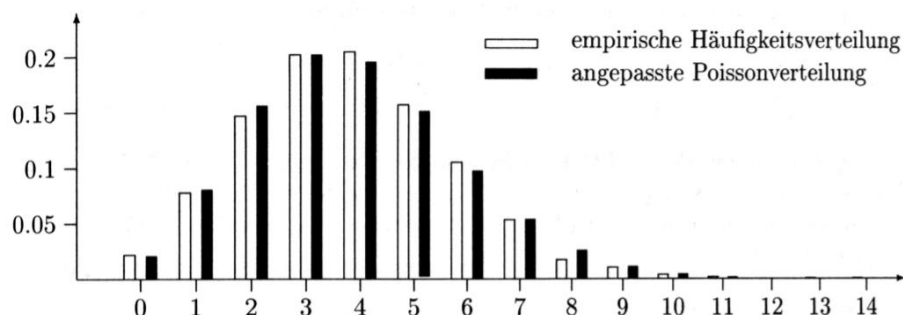


Bild 24.2 Zerfallshäufigkeiten beim Rutherford–Geiger–Versuch mit angepasster Poisson–Verteilung

Abbildung 24: Rutherford – Geiger Experiment mit adaptierter Poissonverteilung [18, S. 191]

7.3.6 Exponentialverteilung

Bei dieser Art von Verteilung spielt, wie im Leben so oft, das Warten eine große Rolle. Warten prägt das Leben eines jeden Einzelnen und ist ständiger Begleiter im Laufe der Zeit. Dass das Warten auch einen mathematischen Vorführeffekt hat, wissen wohl nur die Wenigsten. Auch die Unterschiede beim Warten, werden in diesem Kapitel verdeutlicht und anhand von Beispielen erläutert. Was in diesem Abschnitt besonders hervorgehoben wird, ist das Warten ohne Gedächtnis. [4, S. 186] Denn in vielen Situationen kann man die zu wartende Zeit nicht einschätzen. Deshalb kann man die Wartezeit wohl gut mit dem Intervall $[0, +\infty[$ beschreiben. [4, S. 188]

Um sich ein genaueres Bild der gedächtnislosen Wartezeit bilden zu können, werden nun deutlich unterschiedliche Situationen des Wartens einander gegenübergestellt. Dabei hat die Erwartungshaltung, die eine Person den Umständen entgegenbringt, einen hohen Stellenwert und verdeutlicht den Unterschied.

Man nehme an, nach einer Bestellung bekommt man vom Lieferservice eine Nachricht, dass das Essen in der nächsten Stunde geliefert wird. Als Kundin beziehungsweise als Kunde erwartet man nun die Lieferung innerhalb der 60 Minuten. Man rechnet mit der Ankunft der Bestellung gleichmäßig verteilt binnen einer Stunde. Sind bereits 35 Minuten dieser Stunde vergangen, so ändert sich auch die Haltung der oder des Wartenden gegenüber der Anforderung. So kommt man zu dem Punkt, dass sich die Erwartung im Hinblick auf die schon gewartete Zeit auf die noch zu wartende Zeit ändert. [4, S. 187–188]

Im Gegensatz dazu ein anderes Beispiel. Man ruft bei einem Kundenservice an und landet in der Warteschleife. Im Durchschnitt gelangt man nach drei Minuten in eine offene Leitung. Jedoch kann man dies im Allgemeinen nicht bestätigen. Immerhin bekommt man keine konkrete Auskunft über die Wartezeit. In diesem Beispiel liegt der Unterschied im Vergleich zur ersten Annahme darin, dass sich die Erwartungshaltung, trotz der schon gewarteten Zeit, nicht ändert. [4, S. 188]

Definition: Sei T die Wartezeit, die im Intervall $T: \Omega \rightarrow [0, +\infty[$ zufällige Zufallsvariable ist. Dann bedeutet T ist gedächtnislos, wenn

$$P(\{T \geq t\}) = P(\{T \geq s + t\} | \{T \geq s\})$$

für beliebige $s, t \geq 0$ ist. [4, S. 188]

Außerdem sollte man die Faustregel bei der Exponentialverteilung immer im Gedächtnis haben. Diese sagt aus, dass in Fällen, die eine gedächtnislose Wartezeit annehmen lassen, wie zum Beispiel beim Arzt zu warten oder im Stau zu stehen, die entsprechende Wartezeit im Mittel μ ergibt. Und genau, wenn solche Umstände gegeben sind, kann man die Exponentialverteilung mit dem Parameter $\lambda := \frac{1}{\mu}$ anwenden. [3, S. 149]

Beispiel: Man will die beste Freundin anrufen, aber wie oft ist es wieder besetzt. Aus Erfahrung weiß man, dass es 20 Minuten dauern kann, bis man sie wieder erreichen kann. Daraus kann man aber nicht im Einzelfall darauf schließen, dass es in diesem Fall auch eine gute Viertelstunde dauern könnte. Daher kann man durchaus auf das Modell der

gedächtnislosen Wartezeit zurückgreifen. Infolgedessen kann man die eben genannte Faustregel anwenden. Man möchte in dieser Situation fragen, wie hoch die Wahrscheinlichkeit ist, dass man sie schon nach zehn Minuten erreichen kann. [4, S. 191]

$$P(\{T \leq 10\}) = \frac{1}{20} \int_0^{10} e^{\frac{-t}{20}} dt = -e^{\frac{-t}{20}} \Big|_0^{20} = 1 - e^{-0,5} \approx 0,39. [4, S. 191]$$

Also könnte man mit ungefähr 40% Wahrscheinlichkeit rechnen, dass man bereits nach zehn Minuten beim Anruf auf ein Freizeichen stößt.

7.3.7 Normalverteilung

Als Letztes wird die Normalverteilung behandelt. Diese Verteilung wird in vielen Fachbüchern oft als eigenes großes Kapitel deklariert, weil es sich dabei um eine sehr umfangreiche Verteilung handelt. Um es genauer auszudrücken, gibt es nicht die eine, sondern viele Normalverteilungen. Doch die Unterscheidung zwischen den Normalverteilungen ist oft nur der verwendete Maßstab. [3, S. 161] In diesem Teilkapitel wird mit der „klassischen“ Normalverteilung, mit der es möglich ist, Wahrscheinlichkeiten mit Hilfe der Binomialverteilung zu approximieren, gearbeitet. [3, S. 161] Oft wird dieser Typ von Verteilung auch einfach Gauß'sche Glockenkurve genannt, und beschreibt konkret diskrete Zufallsvariablen. [10, S. 271] Obwohl de Moivre und auch Laplace die Normalverteilung entwickelten und ihnen viel Respekt gebührt, wurde sie dennoch nach dem Mathematiker Carl Friedrich Gauß benannt. Denn er war derjenige, der die in der Natur auffällig ähnlichen Prozessen zur Gauß'schen Glockenkurve entdeckte. [10, S. 272]

Für die Normalverteilung wichtig sind zwei Parameter. Zum einen der Erwartungswert μ und zum anderen die Standardabweichung σ beziehungsweise die Varianz σ^2 . [33, S. 112] Als Definition der Dichtefunktion der Normalverteilung gilt

$$f_{Normal; \mu; \sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. [33, S. 113]$$

Wie am Anfang des Kapitels schon erwähnt, gibt es viele so genannte Normalverteilungen. In diesem Kapitel wird vor allem die Standardnormalverteilung mit den Parametern Mittelwert $\mu = 0$ und der Standardabweichung $\sigma = 1$ behandelt. Man kann diese Art der Normalverteilung kurz als $N(0, 1)$ mit der zugehörigen Dichtefunktion

$$f_{Normal; 0; 1}(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

schreiben. [33, S. 113]

Die Abbildung 25 zeigt eine charakteristische Wahrscheinlichkeitsdichtefunktion der Normalverteilung. Wie man in der unteren Darstellung erkennen kann, handelt es sich bei der Glockenkurve um eine symmetrische Anordnung der Daten. Das bedeutet ebenfalls, dass die Hälfte der Information reicht, um eine Standardnormalverteilung zu konstruieren. [33, S. 114]

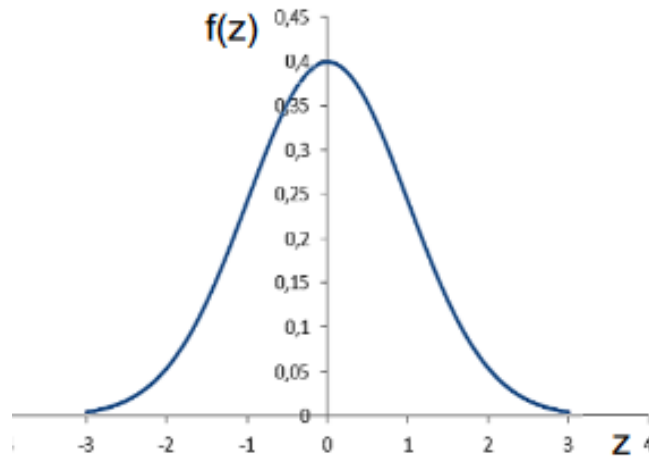


Abbildung 25: Gauß'sche Glockenkurve [33, S. 112]

Für die Normalverteilung haben Grenzen einen großen Stellenwert. Diese werden mit Hilfe der Formel

$$x = \mu + \sigma \cdot z$$

und der Intervalle angegeben. Und danach wird eine Tabelle mit den zugehörigen z -Werten zu Hand genommen, um die Wahrscheinlichkeiten zu erfahren. [33, S. 114–116]

Der wohl größte Vorteil der Normalverteilung ist die mögliche Annäherung durch die Binomialverteilung. Oft wird gesagt, dass für solche Approximationen

$$n \cdot p \geq 5, n \cdot (1 - p) \geq 5 \text{ und } n > 20$$

gelten müsse. Allerdings hat dies keine mathematische Grundlage. In dieser Situation fällt aber auf, dass es sich hier einerseits um eine diskrete und andererseits um eine stetige Verteilung handelt. Daher ist es von Vorteil, wenn man die eine Wahrscheinlichkeit von der Binomialverteilung $P_{binom}(X = k)$ durch die andere von der Normalverteilung $P_{normal}(k - \frac{1}{2} \leq X \leq k + \frac{1}{2})$ annähert. Also man passt $P_{binom}(X \leq k)$ mit $P_{normal}(X \leq k + \frac{1}{2})$ an. [33, S. 118]

Als nennenswerte Wissenschaftler an dieser Stelle gelte der Franzose Abraham de Moivre und Pierre Simon Laplace. De Moivre erweiterte die Annäherung der Binomialverteilung durch die Normalverteilung für $p = \frac{1}{2}$. Und später entwickelte Laplace diesen Gedanken weiter für ein $p \in [0; 1]$. [10, S. 272]

Zentraler Grenzwertsatz: Es sei die Zufallsvariable $Z = Z_1 + Z_2 + \dots + Z_n$ die Summe von n unabhängigen und identisch verteilten Zufallsvariablen Z_i mit positiver Varianz. Ebenso sei $T = \frac{Z - \mu_Z}{\sigma_Z}$ die standardisierte Zufallsvariable zu Z mit $\mu_T = 0$, $\sigma_T = 1$. Dann gilt

$$\lim_{n \rightarrow \infty} P(T \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz. [10, S. 288]$$

8. Zufall und Simulation

In der Wahrscheinlichkeitstheorie wird man ständig vom Wort Zufall begleitet. Man weiß, dass es im Vergleich zum Alltag eine Differenzierung zu diesem Ausdruck gibt. Im mathematischen Sinn besteht ein Zusammenhang mit der Simulation. Solche Simulationen dienen in der Stochastik dazu, mit Hilfe technischen Einsatzes Zufallsereignisse zu ergründen. Durch Computer und andere technische Mittel ist es heutzutage möglich, tatsächliche Bedingungen nachzuahmen und ein mögliches Ergebnismodell zu konstruieren. Das Anwendungsgebiet ist äußerst bedeutend für die technische Entwicklung. [18, S. 148]

Trotz der ausgedehnten Verwendung des Wortes Zufalls bleibt eine genauere Definition für den Begriff aus. Nicht deshalb, weil eine Begriffserklärung unversucht worden lassen wäre, aber auf lange Sicht hatte der Definitionsversuch aufgrund der unklaren Bedingungen keinen Erfolg. So versuchten zum Beispiel Andrei Nikolajewitsch Kolmogorov und Gregory Chaitin unabhängig voneinander eine Folge von Zahlen mit der Eigenschaft zufällig zu definieren. Und zwar meinten die Wissenschaftler, dass es sich genau dann um eine zufällige Zahlenfolge handelt, wenn es nicht möglich sei, diese durch eine kürzere Version der Zahlenserie zu ersetzen. [10, S. 293]

Da es unmöglich ist, ein Zufallsexperiment in der Realität wirklich unendlich oft zu wiederholen, um zu einer Annahme zu gelangen, greift man zur Simulation anhand von Zufallszahlen. Man führt ein Experiment beliebig oft mit dem Computer durch und beobachtet die erzeugten Ereignisse, um auf ein Ergebnis zu kommen. Dieses Modell, welches durch stochastische Hilfestellungen Zufallszahlen produziert, wird auch als die Monte-Carlo-Methode bezeichnet. Der Name lässt sich vom Casino in Monte-Carlo ableiten. [23, S. 216] Ein berühmtes, unzählige Zufallszahlen beinhaltendes Werk stammt von der *Rand Corporation* und trägt den Titel *A Million Random Digits with 100 000 Normal Deviates*. [23, S. 214]

Um in der Wahrscheinlichkeit simulieren zu können, braucht man erst einmal eine Problemstellung, die man mit Hilfe ein stochastisches Musters gliedert. Mit diesem nachgebildeten Modell werden Zufallsexperimente wiederholt ausgeübt und die Ereignisse beschrieben. So gelangt man dann zu einer Behauptung, die man aus den resultierenden Ergebnissen weiter interpretieren kann. [23, S. 214]

Um Zufallszahlen erzeugen zu können, verwendete man früher wie heute spezielle Algorithmen. Früher wurden Rechenschemen verwendet, wie zum Beispiel die *Middle-Square-Methode* oder der *Fibonacci-Algorithmus*. Jedoch mit der sich entwickelnden Technik, änderte sich auch die Form der Erzeugung von Zufallsziffern. Das Rechenverfahren, welches aktuell weit verbreitet verwendet wird, ist wohl die lineare Kongruenzmethode. [23, S. 216]

Satz linearer Kongruenzgeneratoren: Ein linearer Kongruenzgenerator erzeugt genau dann eine Zahlenfolge der maximalen Periode m , wenn folgende Bedingungen gelten: jeder Primteiler p von m teilt $a - 1$. Falls 4 ein Teiler von m ist, so auch von $a - 1$. Die Zahlen b und m sind teilerfremd. [10, S. 296]

9. Ein Auszug zur beurteilenden Statistik

Das Themengebiet der Stochastik bietet viele Facetten. Irrtümlicherweise wird der Begriff manchmal synonym zu Wahrscheinlichkeit verwendet, jedoch bietet das Fachgebiet der Stochastik mehr Bereiche an. Wie aus den anderen Kapiteln bereits erkenntlich wurde, bildet die Wahrscheinlichkeitstheorie bloß einen Bruchteil der gesamten Stochastik. Denn die drei Unterkapitel, die zusammen das Thema der Stochastik ergeben, sind die beschreibende Statistik, die Wahrscheinlichkeitstheorie und die nun zu behandelnde beurteilende oder auch schließende Statistik. [19]

Der praktische Hintergrund der beurteilenden Statistik ist der, dass man aus einzelnen Daten, die man zum Beispiel mit den Grundlagen der beschreibenden Statistik erhebt, auf eine größere Gesamtheit schließen kann. Die zu erhebenden Informationen werden dann als nützlich angesehen, wenn die befragte Gruppe ein angemessenes Bild der Gesamtheit bildet. Dabei ist das längerfristige Ziel einer solchen Schätzung das Vorhersagen eines bestimmten aufweisenden Charakteristikums oder eines in der Zukunft liegenden Ergebnisses. [11, S. 235–236] Diese Kombination, bestehend aus der beschreibenden Statistik und der Wahrscheinlichkeitsrechnung, ist das typische Merkmal der beurteilenden Statistik. [10, S. 310]

Eine relevante Rolle bei der Erhebung von Stichproben und deren Intervallen spielen die unterschiedlichen Verfahrensmöglichkeiten, eine Annahme aufzufassen. Dabei gibt es zum einen die Punkteinschätzung, die eine konkrete Information auffasst. Und zum anderen die Intervallschätzung. Wie der Name schon verrät, gibt man bei dieser Darstellungsform ganze Intervalle von Wahrscheinlichkeiten an. Diese Art bietet den Vorteil eines Spielraumes an. Und zuletzt trennt man von den bereits genannten Gebieten die des Hypothesentests. Hierfür werden Behauptungen aufgestellt und die richtige Hypothese wird mit einer geringen Chance abgelehnt. [11, S. 239]

Anhand eines Beispiels sollten die Eigenschaften der diversen Verfahrensarten klar gestellt werden. Angenommen man untersucht ein spezielles Merkmal an verschiedenen Personen. Dabei beschreibt die Wahrscheinlichkeit $P \in [0, 1]$, ob die oder der Untersuchte dieses Charakteristikum besitzt oder nicht. [11, S. 239] Die Stichprobe umfasst die Menge von $X_1, X_2, \dots, X_{10}$, daher $n = 10$. Die Untersuchung ist binomialverteilt mit $\text{bin}(1, p)$.

Interpretiert man dieses Beispiel jetzt entsprechend der Punktschätzung, dann würde nach dieser Art die Lösung

$$p = \frac{1}{10} \sum_{i=1}^{10} x_i$$

lauten. [11, S. 239]

Gemäß der Intervalleinschätzung würde man die Angabe

$$P(P \in [u, o]) \geq 0,95$$

machen. [11, S. 239]

Und im Hinblick auf den Hypothesentest würde man zum Beispiel die Annahme meinen, dass die Wahrscheinlichkeit P , dass die Person das spezielle Kennzeichen aufweist, weniger als 5% beträgt. [11, S. 239]

Als ein solches Beispiel, welches für die Anwendung der beurteilenden Statistik charakteristisch ist, kann man folgendes verwenden: In der Pharmabranche wird ein neuer Impfstoff gegen eine gewisse Krankheit getestet. Dabei stellt man sich die Frage, ob die neu entwickelte Substanz hilfreich gegen die Symptome ist und die Zahl an Erkrankten verringern kann. Es sind für den modernen Test keine neuen Nebenwirkungen bekannt und er soll deshalb an Kindern getestet werden. Nun lauten die zu behandelnden Fragen, zum einen, wie man eine Untersuchung durchführen soll, und zum anderen, ab welchen Grenzen der neue Impfstoff einen Erfolg liefert. [10, S. 309]

9.1 Parameterschätzung

Im einführenden Teil dieses Kapitels wurden bereits erste Erklärungen zur Punktschätzung und Intervallschätzung gegeben. An dieser Stelle werden die beiden Themenpunkte noch weiter ausgeführt und konkretisiert. Dennoch muss man zuerst den Begriff der Parameterschätzung etwas genauer erklären. Bei dieser Methode ist es das Ziel, dass man aus einer Gesamtmenge von Parametern die Informationen schließt, die man aus der Stichprobe erlangt. [10, S. 311]

Um eine Punktschätzung durchzuführen, wird mit der *Maximum-Likelihood-Methode* gearbeitet. Dieses Verfahren geht auf Sir Ronald Aylmer Fisher zurück, dem ein großer Teil der Entwicklung der Mathematik im vorigen Jahrhundert zugeschrieben wird. Zwar war die Anwendung schon anderen, wie Gauß bekannt, aber Fisher analysierte das Vorgehen auch. [10, S. 321]

Angenommen man untersucht 100 Glühbirnen, die man zufällig auswählt, aus einer 15 000 Stück Produktion. Es handelt sich dabei um ein binomialverteiltes Modell mit den Parametern $n = 100$ und dem noch unbekannten Wert α , der angibt, ob es sich um eine funktionstüchtige Glühbirne handelt oder nicht. Unter diesen 100 zufällig ausgewählten Glühbirnen verstecken sich 3, die kaputt sind. [10, S. 319] Man beschreibt das binomialverteilte Ereignis mit der Wahrscheinlichkeit

$$P(Z = 3) = \binom{100}{3} \cdot \alpha^3 \cdot (1 - \alpha)^{100-3}, [10, S. 319]$$

Z repräsentiert die Menge der nicht funktionierenden Glühbirnen. Man erkennt, dass diese Unbekannte α eine entscheidende Rolle für die Berechnung der Wahrscheinlichkeit spielt. Da die gefragte Wahrscheinlichkeit nach $Z = 3$ fragt, wäre es plausibel, für α einen Messwert anzugeben, der diese gesuchte Wahrscheinlichkeit am ehesten beschreiben könnte. Dadurch entsteht eine Extremwertaufgabe, die abhängig von der Unbekannten α ist und man gelangt zur folgenden Likelihood-Funktion

$$L_3: (0; 1) \rightarrow (0; 1), \alpha \mapsto L_3(\alpha) := \binom{100}{3} \cdot \alpha^3 \cdot (1 - \alpha)^{100-3}. [10, S. 320]$$

Wie die Funktion L_3 aussieht, erkennt man in der Abbildung 26. Die Werte für α werden im Intervall von $[0; 0,15]$ visualisiert, weil die Wahrscheinlichkeit für ein weiteres Intervall gegen 0 konvergieren würde. Die Likelihood-Funktion abgeleitet ergibt

$$\begin{aligned} L_3(\alpha) &= \binom{100}{3} \cdot 3 \cdot \alpha^2 \cdot (1 - \alpha)^{97} + \binom{100}{3} \cdot \alpha^3 \cdot 97 \cdot (1 - \alpha)^{96} \cdot (-1) \\ &= \binom{100}{3} \cdot \alpha^2 \cdot (1 - \alpha)^{96} \cdot (3 \cdot (1 - \alpha) - \alpha \cdot 97) \\ &= \binom{100}{3} \cdot \alpha^2 \cdot (1 - \alpha)^{96} \cdot (3 - 100 \cdot \alpha). \quad [10, S. 320] \end{aligned}$$

Und so kommt man zur Lösung, dass die erste Ableitung nur dann 0 entspricht, wenn $\alpha = \frac{3}{100}$ annimmt. [10, S. 320]

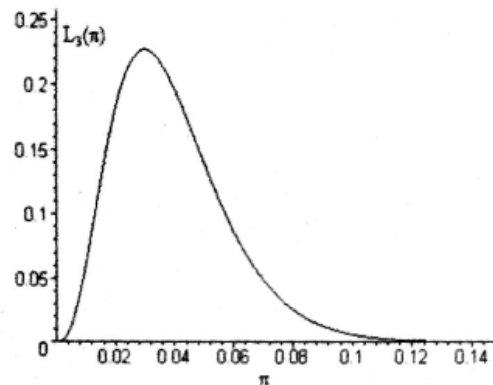


Abbildung 26: Maximum-Likelihood-Methode [10, S. 320]

Es gibt aber noch zusätzliche Alternativen zu der bisherigen Methode, die man beim Suchen der Wahrscheinlichkeit mit unbekannten Parametern benutzen kann. Dabei gibt es aber bestimmte Kriterien, die zu erfüllen sind. Man sollte für eine ideale Schätzung den unbekannten Parameter so weit wie möglich treffen können (Erwartungstreue). Bei zunehmender Größe der Stichprobe sollte das Schätzen immer pointierter werden (Konsistenz) und man muss erfolgreicher beim Schätzen sein als bei alternativen Annahmen (Effizienz). [10, S. 324]

Definition Erwartungstreue und Konsistenz einer Schätzfunktion: T_n sei eine Schätzfunktion bei Stichprobenumfang n für den Parameter η .

1. T_n heißt erwartungstreu oder unverzerrt, wenn $E(T_n) = \eta$ gilt,
2. T_n heißt konsistent, wenn für jedes $\varepsilon > 0$ gilt $\lim_{n \rightarrow \infty} P(|T_n - \eta| \leq \varepsilon) = 1$. [10, S. 324]

Auch wenn die gerade beschriebene Funktion der Punktschätzung in gewisser Hinsicht Vorteile aufweist, ist die Leistung nicht ganz zufriedenstellend. Denn in der Praxis erhält man nur in einzelnen Fällen die wirklich genaue Lösung für den gesuchten Parameter. Um diesen Nachteil zu umgehen, wird die Intervallschätzung verwendet. [10, S. 325]

Wenn man das Beispiel mit den 15 000 zu untersuchenden Glühbirnen nochmals betrachtet: Die Lösung war $\alpha = \frac{3}{100}$ für den unbekannten Parameter. Aber durch dieses Ergebnis kann man nur außer Acht lassen, dass unter dieser Gesamtanzahl von Glühbirnen weniger als eine Menge von 3 und mehr als 14 903 unbrauchbar sind. Daher kann man das Intervall

$$I_U := \frac{3}{15\,000} \leq \alpha \leq \frac{14\,903}{15\,000} =: I_O$$

angeben und die entsprechende Wahrscheinlichkeit schreibt man als

$$P(I_U \leq \alpha \leq I_O) = 1$$

an. [10, S. 325] Der Sinn der Intervallschätzung ist es, dass man ein Intervall angibt, in welches der gesuchte Parameter mit hoher Wahrscheinlichkeit hineinfällt. [10, S. 326]

Man kann innerhalb der Intervallschätzungen zwischen ganzseitigen, linksseitigen und rechtsseitigen Grenzen sprechen, je nachdem wie die Fragestellung lautet. Bei der Frage, wie hoch die Wahrscheinlichkeit der kaputten Glühbirnen mindestens sei, kann auf eine rechtsseitige Intervallschätzung gedeutet werden. Oft weisen zweiseitige Grenzen eines Intervalls auf eine symmetrische Angabe hin.

Ebenso wichtig bei der Intervallschätzung ist die Wahrscheinlichkeit δ , die angibt, wie hoch die Wahrscheinlichkeit ist, dass das Intervall den unbekannten Parameter umfasst. Dementsprechend kann man die Wahrscheinlichkeiten

$$\begin{aligned} P(I_U \leq \eta \leq I_O) &\geq \delta, \\ P(-\infty < \eta \leq I_O) &> \delta \text{ und} \\ P(I_U \leq \eta < \infty) &\geq \delta \end{aligned}$$

annehmen. Die Wahrscheinlichkeit δ bezeichnet man auch als das Vertrauensniveau beziehungsweise Konfidenzniveau. Demgemäß hat das Intervall mit der Wahrscheinlichkeit δ den Namen Vertrauensniveau und auch Konfidenzintervall. [10, S. 326]

9.2 Hypothesentests

Beispiel: *Tea tasting lady*. Eine Dame aus England behauptete, dass sie bei ihrem Tee genau erkennt, ob zuerst die Milch oder der Tee eingegossen wurde. Da sie diesen Unterschied der Mischung am Geschmack merken würde. Um ihre Behauptung zu überprüfen, wird ihr n Mal jeweils eine Tasse mit der Mischung zuerst Milch dann Tee und zuerst Tee dann Milch präsentiert.

Diese Durchführung des Experiments kann man mit dem eines Münzwurfs gleichsetzen und man kann die Wahrscheinlichkeit P , dass sie die Teemischung erkennt, als $H_0: P = \frac{1}{2}$ oder $H_1: P > \frac{1}{2}$ schreiben. Für das Experiment werden der Dame $n = 20$ Paare Tassen gezeigt und man glaubt ihrer Behauptung genau dann, wenn sie mindestens $k = 14$ Mal die richtige Anordnung durchschaut. Errät sie nur weniger als 14 Kombinationen, so zweifelt man an ihrer Glaubwürdigkeit. Daher kann man die Hypothese H_1 mit der Aussage

$S_{20} \geq 14$ und andernfalls die Nullhypothese H_0 mit $S_{20} \leq 13$ interpretieren. [18, S. 233–234] Die dazugehörige Wahrscheinlichkeit sei gegeben durch

$$g_{n,k}(p) := P_p(S_n \geq k) = \sum_{j=k}^n \binom{n}{j} \cdot p^j \cdot (1-p)^{n-j}. \quad [18, S. 234]$$

Der Bereich der Hypothesentests war vor allem für die Geschichte der Statistik in der Mathematik sehr bedeutsam. Eine wesentliche Rolle spielt das Testen von Hypothesen insgeheim in wissenschaftlichen Gebieten. Denn man versucht durch eine Hypothese eine Annahme durch Erfahrungen und Versuchsausgänge zu verallgemeinern und zu bestätigen. Dabei ist es naheliegend, dass man eine Hypothese verwirft beziehungsweise beibehält, wenn sie den Informationen aus den Stichproben entweder nicht entspricht oder eben schon entspricht. [10]

Grundsätzlich formuliert man zunächst einmal für eine solche Anwendung eine Nullhypothese. Diese Aussage gibt das Gegenteil der Hypothese an und man gibt ihr den Namen H_0 . Jedoch ist es in manchen Fällen gar nicht so leicht, zwischen Nullhypothese und alternativer Hypothese zu unterscheiden. Das Problem wird auch als statistisches Testproblem bezeichnet. [11, S. 267–268] Ein Beispiel wäre für die Nullhypothese die Aussage, dass die durchschnittliche Füllmenge kleiner oder gleich μ_0 ist. Dementsprechend würde die Hypothese, die das Gegenteil davon aussagt, lauten, dass die durchschnittliche Füllmenge größer als μ_0 ist. [11, S. 268]

Im Allgemeinen laufen Hypothesentests immer nach demselben Schema ab. Es werden Nullhypothese und alternative Hypothese formuliert. Ausgehend von einer Stichprobe werden die Daten T berechnet. Zusätzlich kennt man eine Grenze c , die entscheidend für das Verwerfen oder Beibehalten einer Hypothese sind. Dann kann man gemäß der Annahme verwerfe H_0 , falls $T > c$ entscheiden. [11, S. 270]

Dennoch bleibt immer ein gewisser Grad an Ungenauigkeit bei solchen Tests haften und so kann man sich nie wirklich ganz sicher sein, da diese Aussagen eine gewisse Anfälligkeit für Fehler mit sich bringen. Man differenziert sie allerdings untereinander, wie man in der Tabelle 7 erkennen kann.

Tabelle 7: Fehlerarten der Hypothesentests [11, S. 270]

	Nullhypothese H_0	Alternative H_1
H_0	Richtig	Fehler 2. Art beziehungsweise β -Fehler
H_1	Fehler 1. Art beziehungsweise α -Fehler	Richtig

Man spricht also von einem Fehler der 1. Art, wenn man fälschlicherweise die Hypothese H_1 als richtig annimmt, obwohl die Nullhypothese H_0 die korrekte Entscheidung wäre. Umgekehrt, tritt der Fall des Fehlers 2. Art ein, wenn man statt der Hypothese H_1 die falsche Nullhypothese H_0 verwendet. [11, S. 270] Man sollte an dieser Stelle die Aufmerksamkeit auf die verwendete Wahrscheinlichkeit lenken. Immerhin handelt es sich bei der Deutung der Fehler um das Berechnen von bedingten Wahrscheinlichkeiten. [10, S. 341]

Um einen beschränkten Bereich zu erhalten, ermittelt man eine obere Schranke $\alpha \in (0, 1)$. Diese Schranke wird im Falle für die Wahrscheinlichkeit eines Fehlers 1. Art angewandt und muss

$$g(v) \leq \alpha$$

erfüllen. Diese Vorschrift wird Signifikanztest des Signifikanzniveaus α genannt. Normalerweise sind die Angaben von 0,01 bis 0,1 hier geläufig. [18, S. 238–239]

Außerdem muss man im Hinblick auf Hypothesentests die Gütefunktion g formulieren. Man definiert sie als $g(x) := P(T \in V \mid p = x)$. Die Gütefunktion ist zuständig für die Fehler 1. Arts. [10, S. 347]

Wenn man anhand des vorigen Beispiels, welches mit 15 000 Glühbirnen angegeben war, nun nochmals auf das Testen eingeht, so würde zum Beispiel die Nullhypothese $H_0: P \leq 0,01$ lauten. Denn so kann der Verkäufer behaupten, dass höchstens 1% seiner Bestellung nicht funktionstüchtig ist. Demnach würde man den Alternativtest $H_1: P > 0,01$ beschreiben. Zusätzlich benötigt man ein Signifikanzniveau α . Angenommen in diesem Beispiel beträgt $\alpha = 0,04$. Dabei spricht man in diesem Fall von einem rechtsseitigen Test, weil man die Hypothese mit „höchstens ein Prozent sind kaputt“ formuliert hat. Angenommen man untersuche von der Lieferung 100 Stück der Glühbirnen. [10, S. 348]

$$P(T \geq k \mid p = 0,01) = \sum_{i=k}^{100} \binom{100}{i} \cdot 0,01^i \cdot 0,99^{100-i} \leq 0,04. [10, S. 349]$$

Mit Hilfe der entsprechenden Tabelle erhält man den korrekten Wert für $k = 4$

$$P(T \geq 4 \mid p = 0,01) \approx 0,0184. [10, S. 349]$$

Daher kennt man jetzt auch den Verwerfungsbereich $V := \{4; \dots; 100\}$. Die Wahrscheinlichkeit, dass sich unter den 100 getesteten Glühbirnen mindestens drei defekte befinden, liegt bei ungefähr 7,9%. [10, S. 349]

Wichtig zu beachten beim Testen von Hypothesen sind die Fehler, die auftreten können. Vor allem beim Fehler 1. Art muss ein besonderes Augenmerk auf das Signifikanzniveau gelegt werden, da das Intervall dadurch eine obere Grenze erhält. [10, S. 353]

In den letzten Entwicklungen hat der Einfluss der Alternativtests oder auch Signifikanztests genannt, sehr an Interesse gewonnen, sodass andere Bereiche der beurteilenden Statistik, wie zum Beispiel die Parameterschätzung, an Bedeutung verloren haben. [10, S. 357] Oft wird in den Wissenschaften argumentiert, dass das Arbeiten mit Hypothesen konzeptlos wirkt, da man von der eigenen Intuition ausgehend eine Behauptung aufstellt. In der Geschichte hat sich bis heute die Methode der Falsifizierbarkeit bewährt. Denn diese Art des Denkens fordert, dass man bei jeder Situation auch einen Rückschlag erleiden könnte. Eine solche Betrachtung des Zusammenhangs geht auf Sir Karl Popper und den kritischen Rationalismus zurück. [10, S. 358]

9.3 Chi-Quadrat-Test

Der Chi-Quadrat-Test, kurz auch χ^2 -Test geschrieben, wurde von Karl Pearson hervorgebracht. Dieser Test ist einer der am längsten verwendeten Tests in der Statistik, da er bereits Ende des 19. Jahrhunderts entwickelt wurde. Für das Testverfahren wird geprüft, wie die relativen Häufigkeiten mit hypothetischen Wahrscheinlichkeiten multinomial miteinander harmonisieren. [18, S. 247] Der χ^2 -Test basiert auf einer eigenen Verteilung und zwar der so genannten χ^2 -Verteilung. [10, S. 360]

Wenn man sich nochmal den Versuch einer Münzwurfsimulation vor Augen ruft, dann kann man die beiden Wahrscheinlichkeiten für $P(\text{Kopf}) = P(\text{Zahl}) = \frac{1}{2}$ festlegen. In diesem Fall beobachtet man ein Experiment nun mit s denkbaren Versuchsausgängen und stellt dies mit der Menge $\Omega = \{1, 2, \dots, s\}$ dar. Es wäre möglich, dass es sich dabei um das Beispiel mit einem Würfel und der Ereignismenge $\Omega = \{1, 2, \dots, 6\}$ handelt, wobei die jeweilige Wahrscheinlichkeit für die möglichen Augensummen

$P(1) = P(2) = \dots = P(6) = \frac{1}{6}$ beträgt. Dabei benutzt man in vielen Fällen zur Bearbeitung solcher Angaben den χ^2 -Test. [10, S. 360]

Als Erklärung soll ein Beispiel mit Würfeln behandelt werden. Es gibt die möglichen Wahrscheinlichkeiten $P_1, P_2, \dots, P_6$ für die möglichen Augensummen 1, 2, ..., 6. Die Nullhypothese H_0 beschreibt die Situation, dass es sich dabei um einen Laplace-Würfel handelt und es wird entsprechend geschrieben

$$H_0: "P_i := P(Z = i) = \frac{1}{6} \text{ für } i = 1, 2, \dots, 6". [10, S. 361]$$

Der Würfel wird $n = 600$ -mal geworfen und man erhält folgende Tabelle 8:

Tabelle 8: Ausgänge des Werfens mit einem Würfel [10, S. 361]

i	1	2	3	4	5	6
Anzahl	121	92	76	109	95	107

Dabei wird die multinomiale Wahrscheinlichkeit durch

$$P(X_1 = k_1 \wedge X_2 = k_2 \wedge \dots \wedge X_6 = k_6 \mid p_1 = p_2 = \dots = p_6 = \frac{1}{6}) \text{ verkörpert. [10, S. 361]}$$

Wenn die Nullhypothese H_0 als richtig angesehen wird, dann schreibt man

$$\begin{aligned} P(X_1 = k_1 \wedge X_2 = k_2 \wedge \dots \wedge X_6 = k_6 \mid p_1 = p_2 = \dots = p_6) &= \frac{1}{6} \\ &= \binom{n}{k_1, k_2, \dots, k_6} \cdot \left(\frac{1}{6}\right)^{k_1} \cdot \left(\frac{1}{6}\right)^{k_2} \cdot \dots \cdot \left(\frac{1}{6}\right)^{k_6} \\ &= \binom{n}{k_1, k_2, \dots, k_6} \cdot \left(\frac{1}{6}\right)^n = \binom{600}{k_1, k_2, \dots, k_6} \cdot \left(\frac{1}{6}\right)^{600}. [10, S. 361] \end{aligned}$$

Eine praktische Methode zur Berechnung ist die Annahme, dass die möglichen Ausgänge $1, 2, \dots, s$ mit der Wahrscheinlichkeit $p_1, p_2, \dots, p_s$ berechnet werden können und dabei die Zahlen für $i = 1, 2, \dots, s$ exakt a_i oft geschrieben werden können. [10, S. 362] Im Allgemeinen kann behauptet werden, dass aufgrund des Erwartungswertes und dessen Abhängigkeit von der Anzahl s der Ereignisse, der Ausdruck $s - 1$ die Menge der Freiheitsgrade angibt. [10, S. 364]

Vor allem der Annäherung der Wahrscheinlichkeit $P(T_n \geq c)$ mit Hilfe des Integrals

$$\lim_{n \rightarrow \infty} P(T_n \geq c) = \int_c^\infty f_{s-1}(t) dt \text{ mit } f_{s-1}(t) := K \cdot e^{-\frac{t}{2}} \cdot t^{\frac{s-3}{2}}$$

wird eine besondere Bedeutung durch Pearson zugeschrieben. Und mit der reellen Konstante, die durch

$$\int_0^\infty f_{s-1}(t) dt = 1$$

bestimmt ist, wird die Wahrscheinlichkeitsverteilung χ^2 -Verteilung mit $s - 1$ Freiheitsgraden ermittelt. Voraussetzung für eine solche Anwendung ist ein genügend großes n . [10, S. 364]

10. Intuition und Wahrscheinlichkeit – Freund oder Feind?

Oft ist die menschliche Intuition eine große Hilfe. In der Mathematik ist dieses Phänomen des Instinktes keinesfalls ausgeschlossen. Es ist in bestimmten Fragestellungen durchaus eine Unterstützung und führt zu einem richtigen Ansatz das Problem zu lösen. Mit solchen Vorkenntnissen wird jede und jeder quasi im täglichen Leben konfrontiert. Sie sind die Ursache dafür, dass die Intuition einen wichtigen Beitrag zum Verständnis und der Handhabung von Problemen leistet. Denn es ist bekannt, dass Sachverhalte und mathematische Zusammenhänge leichter anhand des eigenen Arbeitens und mit Hilfe des Verständnisses und der Logik erklärt werden können.

Einerseits wirken die Begriffe der Intuition und der Stochastik beziehungsweise allgemein ausgedrückt der Mathematik nicht kompatibel. Andererseits stehen die beiden Ausdrücke in Verbindung miteinander. Dabei begegnen sich einander Naturwissenschaft und menschliche Einschätzung auf Augenhöhe. Auch wenn der erste Eindruck täuscht, so können die Mathematik und die Intuition durchaus miteinander in Verbindung gebracht werden. Dabei sollen die korrekten mathematischen Definitionen und Beweise nicht zu Verwirrungen führen, sondern, im Gegenteil sollen sie helfen die Intuition und die naiven Vorstellungen der Schülerinnen und Schüler zu bestärken. [7, S. V] Insbesondere als Lehrperson trägt man für diesen Prozess eine große Verantwortung. Denn man darf in dieser Phase des Lernprozesses die Auffassung der Jugendlichen nicht beschränken oder gar unterdrücken.

Die allgemein üblichen Begriffserklärungen sind zu akzeptieren und dahingehend formuliert worden, weil das Bauchgefühl doch nicht wissenschaftlich bewiesen ist und zu Fehlern führen kann. Solche Irrtümer, die man durch ein falsches Gefühl, verursacht, sind vor allem in der Stochastik weit verbreitet. Die Fachgebiete der Statistik und der Wahrscheinlichkeitsrechnung zählen zu den unterschätzten Themen der Mathematik. Deshalb wird oft die Meinung vertreten, dass man durchaus ausschließlich mit Hilfe der Intuition zu einer richtigen Deutung der Problemstellung gelangt.

Dessen ungeachtet, benötigt man in der Stochastik beide Komponenten, um auf eine Antwort zu kommen. Sowohl die statistischen und wahrscheinlichkeitstheoretischen Elemente als auch den Instinkt. Nur durch ein Wechselspiel beider Gegenstände kann man mit dem Thema richtig umgehen. Denn nicht nur die Mathematik, sondern auch die Intuition und somit die Philosophie verbindet die Existenz der Dinge. [7, S. 19] Zum einen spiegelt diese Aussage im Hinblick auf die Mathematik die Realität und das Modellbilden wider. Und zum anderen repräsentiert der Instinkt die damit verbundenen Vorstellungen, die man mit dem Fachgebiet intuitiv erstmals in Verbindung bringt. [7, S. 19–20]

Besonders in der Statistik und der Wahrscheinlichkeitsrechnung wird man mit dem Ausdruck der Intuition mehr als in anderen Fachgebieten in der Mathematik konfrontiert. Dies wird durch die Differenzierung innerhalb der Wahrscheinlichkeitstheorie begründet. Auf der einen Seite gibt es Unterstützer der subjektivistischen Ansicht, auch Bayes-Ansatz genannt, und zum anderen wird die Auffassung der Objektivisten, die auch als klassischer Ansatz bekannt ist, akzeptiert. [7, S. 21]

Die beiden verschiedenen Meinungen unterscheiden sich in ihrer Interpretation. Die Antwort auf eine Problemstellung sollte zum gleichen Ergebnis führen. Nur die jeweilige Deutung des Ergebnisses variiert je nach Anschauung. Bei der Auffassung der subjektivistischen Art basiert der Ansatz auf Informationen, die vor der Verteilung beziehungsweise dem Experiment angenommen werden. Hingegen gilt im klassischen Ansatz nur das Erfassen von objektiven Fakten als gewinnbringend. Dadurch wird dem Arbeiten mit relativen Häufigkeiten ein hoher Stellenwert zugeschrieben. [7, S. 357–358]

Trotzdem führen die intuitiven Vorstellungen oft zu Irrtümern und Fehlvorstellungen. Der Grund für solche Irreführungen wird durch die besonders in der Wahrscheinlichkeitstheorie auftretenden Paradoxa, intensiviert. [7, S. 30] Auch wenn widersprüchliche Aufgaben in den meisten Fällen für den Betrachter komplex und unsinnig erscheinen mögen, so haben sie einen tiefen Hintergrund. Auf Grund der vernunftwidrigen Aufgabenstellungen wird aber gleichzeitig der Denkvorgang gefordert und gefördert. Gerade in dieser Phase des Lernprozesses spielt die Intuition eine große Rolle für das Verständnis. Denn solche Paradoxa besitzen eine starke Verbindung mit dem Instinkt. Jedoch sollte man dahingehend immer beachten, dass das Bauchgefühl in vielen Fällen zwar hilft, aber gerade in Angelegenheiten, die die Mathematik betreffen, man der Intuition allein nicht trauen sollte. Werden jedoch die beiden Elemente, zum einen der Instinkt und zum anderen das mathematische Modell, welches hinter dem Problem steckt, kombiniert, so wird es keine Probleme geben, eine ungewöhnliche Fragestellung lösen zu können.

Allerdings ist zu beachten, dass der Einfluss der Intuition im Leben und vor allem in der Mathematik eine bedeutende Rolle spielt. Es ist nicht generalisierbar, dass die Vorstellung eines jeden Menschen die gleichen Gedanken mit sich bringt. Die Macht der Intuition ist vielfältig und deshalb bietet sie auch einen großen Anwendungsbereich an. Dennoch sind intuitive Ideen leitend für die Mathematisierung des stochastischen Denkens.

Was die Stochastik von anderen Fachgebieten der Mathematik wie der Algebra oder der Analysis unterscheidet, ist das Fehlen der Operationalisierbarkeit. [7, S. 37] Das Arbeiten mit natürlichen Zahlen, zum Beispiel, präsentiert ein anschauliches Handhaben mathematischer Objekte. [7, S. 37] Das Addieren oder Subtrahieren mit ihnen erscheint logisch nachvollziehbar. Anders verhält es sich in der Stochastik. Vor allem im Bereich der Wahrscheinlichkeitstheorie erweisen sich die Aussagen über Sachverhalte als unsicher und konturlos. Die Probabilität bei einem Gewinnspiel zu siegen, besitzt immer den Beigeschmack einer gewissen Unsicherheit und Vagheit. Aus diesem Grund erlaubt es der Bereich der Wahrscheinlichkeit nicht mit Ansätzen der Vorstellung eine Situation dementsprechend zu erklären. Es fehlt die Sicherheit und die Gewissheit dafür, dass man mit stochastischen Modellen hantieren kann. [7, S. 38]

Einen wichtigen Beitrag für das Vertrauen in die zweifelhaften Annahmen, die man für die Wahrscheinlichkeit gebrauchen kann, steuert die Erfahrung der bereits bekannten Ergebnisse von Experimenten bei. Zwar bleibt trotzdem ein Grad an Ungewissheit an der Theorie haften, aber nichts desto trotz kann man sich an ein sich wiederholendes Muster klammern. Daher ist es möglich die Heuristik ausgehend von bestimmten harmonischen Rhythmen zu bestärken. [7, S. 41]

Dabei kann man klare Strukturen in der Wahrscheinlichkeitstheorie definieren, die bei der Berechnung und der Interpretation von stochastischen Problemen einen wichtigen Beitrag leisten. Ausgehend von symmetrischen Erscheinungen kann man von einer Behauptung auf die Gesamtheit erschließen. Diese wiederum ermöglicht eine Schlussfolgerung, bevor ein Ergebnis bekannt wird. Es handelt sich dabei um eine *a priori* Einschätzung der Wahrscheinlichkeit. Andererseits kann man mit Hilfe von Häufigkeiten die Chance, dass ein bestimmtes Ereignis eintritt, vermuten. Weiters kann man mit Hilfe des individuellen Glaubens eine Möglichkeit ausdrücken. Dieses Konzept ist stark von subjektiven Erfahrungen geprägt und beansprucht das persönliche Denkvermögen. Eine weitere Methode Wahrscheinlichkeiten zu interpretieren ist die durch Axiome definierte Betrachtung des Kapitels. Den Aufbau der formalen Anschauung kann man einerseits als sehr strukturiert und organisiert beschreiben, aber andererseits vernachlässigt diese Art die Deutung der Chancen. [7, S. 87–91]

Vor allem der Bayes-Ansatz wird von der Intuition bedeutsam geprägt und beeinflusst. Insbesondere ist es möglich das Phänomen der bedingten Wahrscheinlichkeit anhand des Chancenverhältnisses sinnvoll zu erläutern. Durch die zusätzliche Information, die bei bedingten Wahrscheinlichkeiten als Hilfestellung bekannt ist, wird der Zusammenhang neu organisiert und dies ermöglicht eine einschränkende Darstellung des Problems. [7, S. 177] Die Formel der bedingten Wahrscheinlichkeit, aber auch der weiterführende Term des Bayes-Theorems, unterstützt durch die intuitive Vorstellung das Verstehen und Interpretieren eines auf den ersten Blick komplex wirkenden Zusammenhang. [7, S. 178]

In diesem Kontext wird die bedingte Wahrscheinlichkeit auch dazu eingesetzt, die Begriffe der stochastischen Unabhängigkeit und der Multiplikationsregel zu illustrieren und zu simplifizieren. Als eigenständige Begriffserklärungen wirken die beiden Begriffe für sich äußerst komplex und unbrauchbar. Jedoch betrachtet mit der Definition der bedingten Wahrscheinlichkeit wird die Folgerung daraus erkenntlich. [7, S. 179]

Unumgänglich für die bedingte Wahrscheinlichkeit, aber auch allgemein für die gesamte Wahrscheinlichkeitstheorie ist der Begriff Unabhängigkeit. Zum einen verlangt man in der Wahrscheinlichkeit Unabhängigkeit für das Ausführen von Zufallsexperimenten. Und zum anderen wird bei der bedingten Wahrscheinlichkeit die Abhängigkeit zweier Ereignisse gefordert. [7, S. 188]

Eine Folge die, insbesondere die bedingte Wahrscheinlichkeit mit sich bringt, ist die Spekulation über mehrere Ergebnisse eines Zufallsexperimentes. Zum Beispiel erschien beim Roulette-Tisch bereits mehrere Male hinter einander die Farbe Rot. Als Spielerin beziehungsweise als Spieler wird man vermuten, dass nach so vielen roten Runden nun bald die Farbe Schwarz eintreten wird. [7, S. 191] Dieser Gedankengang präsentiert das typische Denken, welches durch den Einfluss der vorherigen Ergebnisse beeinträchtigt wurde. Denn als Besucherin oder Besucher eines Casinos weiß man im Allgemeinen, dass die Wahrscheinlichkeit, dass die Kugel auf ein schwarzes oder auf ein rotes Feld fällt, unabhängig von den Ausgängen davor ist. Die dazugehörige Wahrscheinlichkeit ist ja durch $P(\text{rotes Feld}) = \frac{1}{2} = P(\text{schwarzes Feld})$ gegeben. In vielen Fällen wird sich in der Situation eine Casinogeherin oder ein Casinogeher nach einer gewissen Zeit gefragt haben, wie hoch die Wahrscheinlichkeit für Schwarz ist, wenn in den letzten Runden Rot erschien.

Diese Frage repräsentiert allerdings nichts anderes als die bedingte Wahrscheinlichkeit, für die man schreibt

$P(\text{Kugel fällt auf Schwarz} | \text{Kugel ist auf Rot gefallen})$. Und diese Wahrscheinlichkeit beschreibt eine ganz neue Situation, die von einem bereits eingetretenen Ereignis beeinflusst wird. Gerade diese Eigenschaft macht den Vorteil der bedingten Wahrscheinlichkeit aus. Gleichzeitig wird dieser Charakter jedoch von einigen in der Praxis nicht erkannt. Oder anders gesagt, empfinden viele Menschen diesen Nutzen nicht als Unterstützung, sondern oft sorgt er für Verwirrungen und Missverständnisse.

Die Bayes-Regel ist vor allem für den subjektivistischen Ansatz der Stochastik ein wichtiger Bestandteil. Jedoch wird der Satz von Bayes auch in der klassischen Auffassung des Themas verwendet. [7, S. 226] Dadurch, dass die bedingte Wahrscheinlichkeit sowohl für die individuelle Intuition als auch für die formale Operationalisierung einen essentiellen Beitrag leistet, bringt diese Methode der Berechnung von gewissen Chancen einen großen. Darüber hinaus liefert der Satz von Bayes noch weitere Ergänzungen im Hinblick auf die beurteilende Statistik. Denn es werden Annahmen beziehungsweise Hypothesen aufgestellt, die mit Hilfe von bereits eingetretenen Ereignissen in ihrer Wahrscheinlichkeit beschränkt werden. [7, S. 226]

Trotzdem bleiben die Fehlvorstellungen in der Stochastik, und im Speziellen die durch die bedingte Wahrscheinlichkeit verursachten in der Praxis nicht aus. Dies ist wohl darauf zurückzuführen, dass die mit der Mathematik in Verbindung gebrachten Intuitionen oft nicht viel über den Sachverhalt aussagen. Für diesen Nachteil des Begriffs der Wahrscheinlichkeit sind die grundlegenden Axiome unterstützend. [7, S. 260] Allerdings weist dieser Zugang wiederum einen Widerspruch in sich auf. Auf der einen Seite wird durch fundamentale Begriffserklärungen ein Einstieg in das Thema erleichtert. Auf der anderen Seite wird der Ansatz, besonders in Bezug auf die Wahrscheinlichkeitstheorie, intuitiv erläutert. [7, S. 261] Denn durch das unterbewusste Verständnis und die Vorstellung wird der Zugang zu einem neu gelernten Thema leichter geformt.

Beispiel: Beim Ort eines Verbrechens werden auf einer Leiche Blutspuren entdeckt. Das Blut kann mit einer Wahrscheinlichkeit von 100% durch das DNA-Programm der gesuchten Person richtig zugeordnet werden. Die Chance, dass bei dieser Analyse eine Unschuldige oder ein Unschuldiger zur Rechenschaft gezogen wird, beträgt 0,0001%. In dieser Situation passt die Probe des Angeklagten mit der gefundenen Blutspur zusammen. Nun stellt sich die Frage, mit welcher Wahrscheinlichkeit hat man tatsächlich den Schuldigen gefunden? [34, S. 123–124] Es wird mit Hilfe der Formel von Bayes die Annahme überprüft. Es sind folgende Aussagen zur Berechnung bekannt:

$P(V) = 0,0001\%$, beschreibt die Wahrscheinlichkeit, dass eine zufällige Person unter Verdacht steht;

$P(\bar{V}) = 99.9999\%$, beträgt die Gegenwahrscheinlichkeit, dass die verdächtige Person nicht der Täter war;

$P(B|V) = 100\%$, gibt die Wahrscheinlichkeit dafür an, dass unter der Bedingung, dass es sich um die verdächtige Person handelt, auch die Blutprobe übereinstimmt;

$P(B|\bar{V}) = 0,0001\%$, entspricht der Wahrscheinlichkeit, dass die Blutspur eines Unschuldigen mit der Probe am Tatort übereinstimmt. [34, S. 124]

Die gesuchte Wahrscheinlichkeit lautet: Wie hoch ist die Chance, dass die Verdächtige beziehungsweise der Verdächtige am Tatort war und die Blutspuren vor Ort mit ihrer oder seiner Blutgruppe übereinstimmen? [34, S. 124] Kurz:

$$P(V|B) = \frac{P(B|V) \cdot P(V)}{P(B|V) \cdot P(V) + P(B|\bar{V}) \cdot P(\bar{V})}$$

$$= \frac{0,1 \cdot 0,000001}{0,1 \cdot 0,000001 + 0,000001 \cdot 0,999999} = 0,5. [34, S. 124]$$

Überraschenderweise beträgt die Wahrscheinlichkeit, dass die Blutspur mit der DNA-Probe des Verdächtigen übereinstimmt lediglich 50%. [34, S. 124] Dieses Beispiel visualisiert die oft vorkommende unerwartete Wendung. Rein intuitiv hätte man vermutlich wegen der geringen Wahrscheinlichkeit von 0,0001% eines Unschuldigen das Ergebnis für viel sicherer eingeschätzt.

Doch hat diese Methode, einen stochastischen Sachverhalt zu berechnen, nichts mit Vertrauen zu tun. Es ist die bloße Anwendung der Formel der bedingten Wahrscheinlichkeit oder besser gesagt der Einsatz des Bayes-Theorems. Durch diese Regel wird das Ergebnis eben von bereits bekannten Informationen beeinflusst. Das bedeutet aber noch lange nicht, dass man den intuitiven Einblick und Vorstellungen gegenüber der Wahrscheinlichkeit unberücksichtigt lassen darf.

Besonders vernachlässigt wird der Zugang zur Stochastik über die eigene Intuition in der Schule. Doch gerade für Schülerinnen und Schüler zeigt sich ein intuitiver Einstieg in das Thema als eingängig. Es sollte auf die Vorstellungen der Jugendlichen explizit eingegangen werden und von diesen Anhaltspunkten ausgehend wäre der Aufbau der formalen Theorie und Definitionen optimal. [7, S. 262] Die Lehrperson muss an dieser Stelle sensibel und behutsam mit den anfänglichen Ideen der Lernenden umgehen, damit die Entwicklung zu einer brauchbaren Anwendung gelingt.

In vielen Situationen fällt es schwer, zwischen einer Modellebene und der Realitätsebene zu differenzieren. [34, S. 15] Daher wäre das Modellbilden eine weitere Unterstützung, die es ermöglichen könnte die Intuition zu bestärken. [34, S. 13] Modellbilden kann helfen die Vorstellungen zu organisieren. Dadurch wird ermöglicht, dass die Gedanken geordnet werden und eine Relation zu den Anwendungsbereichen erststellt wird. [34, S. 13] Denn es ist typisch für die stochastischen Probleme, dass die Theorie von der Anwendung in sachbezogenen Situationen isoliert betrachtet werden muss. [34, S. 14]

Modellbilden wird durch die Umgangssprache gefördert. Mit Hilfe von bekannten Ausdrücken und vertrauten Situationen ist es möglich, einen mathematischen Sachverhalt in einfachen Worten widerzugeben. Schon im Alltag wird man von Wahrscheinlichkeiten begleitet. Oft sind es aber unschwellige Aussagen, die man nicht auf den ersten Blick mit der Wahrscheinlichkeitstheorie verbindet. [34, S. 165] Solche Ausdrücke, die eine Verbindung zur Stochastik bilden sind zum Beispiel: Die Möglichkeit oder die Chance zu haben, das war Zufall oder aus Überzeugung. [34, S. 166] Diese Begriffe repräsentieren allesamt dem Zusammenhang zwischen Alltag und Wahrscheinlichkeit, der aber in den meisten Fällen des Gebrauchs unbewusst bleibt.

Außerdem steht nicht nur die Alltagssprache in Verbindung zu wahrscheinlichkeitstheoretischem Denken, sondern auch die Logik. [34, S. 225] Logik und Intuition gehen ebenfalls eine transparente Verbindung miteinander ein, die es ermöglichen kann, aus instinktiven Annahmen auf konsequente Folgerungen zu schließen. Die beiden Faktoren spielen somit eine wichtige Rolle in der Statistik und der Wahrscheinlichkeitstheorie, weil sie miteinander kohärent sind.

Die Erklärung durch die unterschiedliche Wahrnehmung der Umwelt kann auf bestimmte Darstellungen zurückgeführt werden. Zum einen wird behauptet, dass die menschliche Kognition fähig ist, sich der Umgebung und der damit voranschreitenden Evolution anzupassen. Hinter dieser Theorie steckt das Fachgebiet der Entwicklungspsychologie. Und zum anderen wird die Auffassung von Vorstellungen mit dem Prinzip von Lernen durch Erfahrungen und expliziten Lernprozessen begründet. Als anschauliches Beispiel kann man hierfür den Satz von Bayes nennen. Die Bayes-Regel veranschaulicht deutlich, dass die anfänglichen Intuitionen falsifizierbar sein können und man erst durch die Berechnung des Bayes-Theorems auf ein unvermutetes Ergebnis gelangt. [34, S. 227]

Nicht nur der Satz von Bayes führt zu Verwirrungen und falschen Vorstellungen, auch der Zusammenhang zwischen Korrelation und Kausalität wird in vielen Bereichen falsch interpretiert. Wenn ein Zusammenhang bestehend aus zwei Kennzeichen verglichen wird, kommt es oftmals, vor allem in den Medien, zu falschen Deutungen und führt dann zu Falschmeldungen. In den meisten Fällen verursacht das Ursache-Wirkung-Verhältnis Probleme. [34, S. 262] Zum Beispiel wird häufig in Zeitschriften oder im Fernsehen veröffentlicht, dass sich eine lange Studienzeit rentiert, da man auf dem Arbeitsmarkt mehr Geld verdiene. [19] Natürlich macht sich eine hohe Anzahl der Semester nicht bezahlt. Was in solchen Beiträgen nicht berücksichtigt wird, ist das Verhältnis der beiden Komponenten. In diesem Fall wird nicht beachtet, dass es sich bei längeren Studien oftmals um kompliziertere Fachgebiete aus der Technik handelt. Das Publikum verbindet allerdings mit dieser Schlagzeile die Vorstellung, dass sich langzeitstudieren auszahlt.

Ein weiteres Beispiel für Fehlinterpretationen im Hinblick auf Korrelation und Kausalität ist die falsche Annahme, dass eine Glatze das Einkommen erhöhe. [34, S. 263] Bei diesem Fall wird ignoriert, dass nicht die Haare das Gehalt steigen lassen, sondern das Alter. Denn im Verhältnis verlieren ältere Männer vergleichsmäßig mehr Haare als jüngere, verdienen aber dementsprechend mehr aufgrund ihrer Erfahrung. Eine solche Fehlinterpretation wird auch als Moderatoreffekt oder Drittvariable bezeichnet. [34, S. 263]

Die Argumente, welche versuchen, die Verwirrungen und Komplexität des stochastischen Denkens zu ergründen, weisen alle daraufhin, dass das stochastische Denken ein umfangreiches Spektrum anbietet. Die Aufgabe, die die Stochastik verlangt, ist jene, dass man diese Diversität systematisiert, damit es möglich ist, statistische und wahrscheinlichkeitstheoretische Methoden darauf anzuwenden. [32, S. 45] Eine wichtige Rolle dafür bilden die jeweiligen Vorstellungen, die jedes Individuum mit sich bringt und die den Verstand intuitiv zu weiteren Folgerungen führen. Die entsprechenden Theorien, die man in den Situationen benützen muss, können durchaus noch durch weitere Einwirkungen geformt werden. Man soll die Methoden und Anschauungsmöglichkeiten des stochastischen Denkens nicht als Nachteil ansehen, was allerdings leider oft der Fall ist. Sondern die Konzepte, welche hinter den formalen Berechnungen stehen, sollen durch die Intuition unterstützt werden.

11. Der Umgang mit den Begriffen Zufall, Statistik und Wahrscheinlichkeit in der Schule

Auf das umfangreiche, sehr mathematiklastige Fachwissen im ersten Teil der Arbeit, folgt nun ein vielleicht doch sehr konträrer Bereich der Wahrscheinlichkeitstheorie und dem Bereich der Statistik. Denn man muss durchaus zugeben, dass, obwohl es sich um dasselbe Thema handelt, die jeweiligen Disziplinen sich sehr unterschiedlich mit dem Gebiet der Stochastik befassen.

Einerseits gibt es in der Fachmathematik eine sehr detaillierte und klare Handhabung. Und andererseits wird das Thema der Stochastik in der Schule sehr breit gefächert und wirkt für die Schülerinnen und Schüler oftmals sehr komplex. Aber warum gibt es eine solche Differenzierung zwischen der Fachmathematik und der Schulmathematik im Hinblick auf die Wahrscheinlichkeitstheorie?

Wie im ersten Teil erkenntlich ist, zeigt die Stochastik im Fachbereich Mathematik viele verschiedene Zugänge zum Wahrscheinlichkeitsbegriff. Diese Varietät, die den Einstieg zum Verständnis erschwert, ist wohl ein Grund dafür, dass in der Schule mit diesem doch komplexen Thema anders umgegangen wird. Ebenso wird die Berechnung in den beiden Fachrichtungen anders behandelt. Besser gesagt nicht nur das Kalkül unterscheidet sich in den Gebieten, sondern viel mehr die Interpretation. Denn gerade das Thema der Stochastik lässt, zum einen in der Fachmathematik, aber zum anderen ebenso in der Schulmathematik, ein großes Spektrum an Interpretationsmöglichkeiten zu.

Insbesondere der Begriff des Zufalls weist eine solche Eigenschaft der Unterscheidung auf. Es scheint, dass gerade der Zufall einen leichten Einstieg in das Thema und für die Weiterführung der Wahrscheinlichkeitsrechnung anbietet. Jedoch trügt der Schein dabei. Besitzt der Ausdruck des Zufalls im Alltag einen sehr gängigen Gebrauch, beschreibt er im Fachwissen der Mathematik eine eher spezielle Situation. Wie bereits im einführenden Kapitel zu lesen war, gilt der Zufall im Normalfall als etwas Ungewöhnliches im Leben, ohne jeglichen Rhythmus. Hingegen stellt der Begriff in der mathematischen Perspektive einen Zustand dar, der über deutliche Eigenschaften verfügt. So hat der Zufall in diesem Sinne jene Charakteristik, die ihm erlaubt, mit einer gewissen Gesetzmäßigkeit zu spielen. Im Gegensatz dazu beschreibt eine Zufälligkeit im täglichen Leben etwas Unregelmäßiges ohne eine bestimmte Ordnung.

Das Beispiel des Zufalls vermittelt wohl den verständlichsten und naheliegendsten Widerspruch. Doch es gibt noch einige andere gegensätzliche Begrifflichkeiten zwischen der pädagogischen und der naturwissenschaftlichen Sicht im Hinblick auf die Theorie der Wahrscheinlichkeit.

Wenn man sich nochmals die Gegenüberstellung der Berechnungen beziehungsweise die Interpretationsmöglichkeiten der Wahrscheinlichkeitsbegriffe ins Gedächtnis ruft, so zeigt dieser Vergleich ein weiteres gutes Modell für die Unterscheidung zwischen Schule und Universität. Die Stochastik bietet einige sehr unterschiedliche Zugänge zur Kalkulation oder Deutung eines Beispiels. So wird in der Mathematik nicht wie in der Schule nach Berechnungsart differenziert, sondern nach der Interpretation. [19]

Schließlich bevorzugen Schülerinnen und Schüler Rechnungen in der Mathematik, die stets die gleiche Struktur und dasselbe Prinzip aufweisen, um einen systematischen Vorgang einzuüben. Sie wollen keine Möglichkeiten haben, die ihnen Platz für Interpretationen lassen. Solche Freiräume im Schulfach der Mathematik verwirren die Jugendlichen oft nur. Dennoch ist dies der falsche Zugang, den komplexen Begriff der Stochastik den Schülerinnen und Schülern beizubringen.

Allerdings hat man sich in vielen Schulbüchern dazu entschieden, den Kindern das Thema so näher zu bringen. Bereits beim Begriff Stochastik beziehungsweise Wahrscheinlichkeit scheitert der Versuch, eine klare Begriffserklärung zu geben. Zu umfangreich und vielschichtig sind die Teilgebiete der Wahrscheinlichkeitsrechnung, um eine einzige richtige Definition für den Begriff Wahrscheinlichkeit formulieren zu können.

Diese fehlende Präzisierung des doch sehr umfassenden Themas der Stochastik führt die Schülerinnen und Schülern zu einer schwammigen Vorstellung gegenüber dem Gegenstand. Wie kann man den Lernenden nun als Lehrkraft Hilfestellung dazu leisten, damit sie mit dem Begriff der Stochastik mehr anfangen können als bloßes einstudiertes Rechnen? Auch wenn man glaubt, dass die Wahrscheinlichkeit ein Fachgebiet der Mathematik ist, welches für die Schülerinnen und Schüler naheliegend beziehungsweise anwendungsorientiert ist, gibt es trotzdem immer wieder Fehlvorstellungen und Missverständnisse im Gedankengang der Jugendlichen.

Allerdings darf man nicht außer Acht lassen, dass die Stochastik mit ihren Teilgebieten stetig an Bedeutung und Interesse zunimmt. Die steigende Aufmerksamkeit für das Fachgebiet kann man wohl durchaus auch auf die Entwicklung der Gesellschaft zurückführen. Denn man wird ständig mit Aussagen der Wahrscheinlichkeit oder mit Situationen, in welchen stochastisches mathematisches Wissen gefragt ist, konfrontiert. [21, S. 1]

Aber nicht nur im Alltag hat sich der Umgang mit dem Thema der Stochastik geändert, sondern auch im Schulleben wurde vor allem in 1980er Jahren das Gebiet der Stochastik erweitert. Galt die Vorliebe anfangs noch eingeschränkt dem Unterkapitel der beschreibenden Statistik, änderte sich beziehungsweise vergrößerte sich die Aufmerksamkeit zu einem breiteren Wissen. [21, S. 2]

Das weitere Kapitel über die Wahrscheinlichkeitstheorie entwickelte sich erst im 17. Jahrhundert. Zu dieser Zeit besaß die Disziplin jedoch noch kein eigenes Standbein. Erst in den 1930er Jahren etablierte sich das Fachgebiet. Ähnlich hat die schließende Statistik erst ab dem 19. Jahrhundert durch ihre reiche Anwendung an Bedeutung gewonnen. [21, S. 3]

Jedoch lässt sich über die Begrenzung der Unterkapitel streiten. Da unterschiedliche Autorinnen und Autoren den Begriff entweder kompakter oder doch weiter auffassen, gilt das Einschränken der Themen als kontrovers. So sehen manche Mathematikerinnen und Mathematiker die beschreibende beziehungsweise die beurteilende Statistik als einen Teil der gesamten Stochastik an und manche nicht. Ebenso wird mit der Einordnung des Themas der Kombinatorik unterschiedlich umgegangen. Daher sucht man dieses Kapitel vergebens in den Bildungsstandards unter dem Bereich Daten und Zufall. Allerdings findet man den Gegenstand der Kombinatorik unter dieser Überschrift. [21, S. 3]

Dennoch wird immer wieder ein Zusammenhang zwischen den drei Teilkapitel, der beschreibenden Statistik, der Wahrscheinlichkeitstheorie und der beurteilenden Statistik bemerkt. Daher wird vermutet, dass diese Komponenten doch in einer gewissen Art in einer Beziehung zueinander stehen. [21, S. 4] Auf diese Verbindung wird in den Schulen relativ wenig eingegangen. Wohl auch deshalb stellt die Vernachlässigung des Verhältnisses zwischen den einzelnen Teilkapiteln ein Problem dar und führt zu Missverständnissen. [21, S. 5]

11.1 Unterricht und Stochastik

Wie schon erwähnt, handelt es sich bei der Stochastik um ein breitgefächertes und äußerst umfangreiches Themengebiet. Aus diesem Grund betrifft das Arbeiten mit stochastischen Methoden nicht nur den Mathematikunterricht, sondern wird im besten Fall fächerübergreifend unterrichtet. [21, S. 6] Insbesondere das Themengebiet der beschreibenden Statistik, welches oft sehr trivial wirkt, ist in vielen anderen Fächern einsetzbar. [21, S. 7] Oft wird ein Kapitel mit einem anschaulichen Beispiel eingeführt. Gerade der Bereich der Stochastik bietet die Option, anhand von alltäglichen und gebräuchlichen Versuchen einen Sachverhalt zu erklären und zu demonstrieren. [21, S. 7–8]

Aber es gibt durchaus noch mehr, was die Stochastik in der Unterrichtsplanung ermöglichen kann. Ausgehend von bekannten Beispielen, für welche die Schülerinnen und Schüler bereits eigene Erfahrungen mitbringen, zeigen sich viele Vorteile für den Einstieg in das Thema. Dennoch können auch fiktive Probleme, die trotzdem einen sachlichen Kern besitzen, zweckvoll eingesetzt werden. So werden vielfältige Zugänge zu den statistischen und wahrscheinlichkeitstheoretischen Aufgabenstellungen geboten. Deshalb gibt es Möglichkeiten, durch unterschiedliche Experimente und Nachahmungen den Jugendlichen verschiedenste Ansichten und Berechnungsmöglichkeiten von Aufgabenstellungen darbieten zu können. [21, S. 8]

Besonders wird die Stochastik mit Interpretation und Deutung in Verbindung gebracht. Zwar gehören das Berechnen der Wahrscheinlichkeit, das Zuordnen von Daten und das Testen von Vermutungen zur Aufgabe der Stochastik, aber trotzdem gilt durch die Auffassung des Problems und vor allem die Lösung eine wichtige Rolle. Denn eben die Interpretation des Zusammenhangs erweist sich oft als Herausforderung. Insgeheim zählt diese Schwierigkeit zu den Tücken im Schulalltag. Denn durch das Deuten der Fragestellung und der dazugehörigen Frage werden größtenteils die Kompetenzen des Argumentierens und Begründens verlangt. [21, S. 8–9]

Vor allem aufgrund der zentralbasierten Reifeprüfung hat das Themengebiet der Stochastik einen Aufschwung erhalten. Denn früher wurden die stochastischen Kapitel in den Schulen von den Lehrerinnen und Lehrern teilweise nur überflogen. Doch schon seit längerer Zeit findet man Aufgaben und die dazugehörige Theorie sowohl in den Lehrplänen als auch in den Schulbüchern. [1, S. 115]

Dennoch gilt die Stochastik für viele Jugendliche nicht als ein einfaches Fachgebiet der Mathematik, da die Sachverhalte oft zu komplex und unvorstellbar auf die Lernende wirken.

Deshalb führen die Vorstellungen der Schülerinnen und Schüler aus gutem Grund zu gewissen Fehlvorstellungen. [1, S. 122]

Zum einen gibt es die Behauptung von Jugendlichen, dass beim mehrmaligen Ergebnis einer 6 beim Werfen eines Würfels die Wahrscheinlichkeit, beim nächsten Wurf erneut eine 6 zu erhalten, geringer ist. Es wird nicht beachtet, dass der Würfel kein Gedächtnis besitzt und jeder neuer Versuch wieder dieselbe Wahrscheinlichkeit von $\frac{1}{6}$ besitzt, da jeder Zufallsversuch unabhängig von dem anderen betrachtet werden kann. [1, S. 122]

Zum anderen gibt es die Fehlvorstellung, dass sich bei mehrmaligem Wurf einer Münze die Ergebnisse der beiden möglichen Varianten Kopf oder Zahl auspendeln. Doch in Wirklichkeit ist das Gegenteil der Fall, nämlich, dass sich die Ungleichheit der zwei Ergebnisse bei wiederholtem Versuch erhöht. [1, S. 126] Der menschliche Verstand versucht Begründungen und Erklärungen für mathematische Kontexte zu ergründen. Dabei sind oft, vor allem wegen der Unkenntnis von bestimmten Grundlagen, wie der beschreibenden Statistik, Fehlvorstellungen und Irrtümer miteinbezogen. [1, S. 129]

Aus diesem Grund wird oft das Stichwort des Modellbildens mit Stochastik in Verbindung gebracht. Schülerinnen und Schüler müssen gewohnt sein, aus einer Angabe ein entsprechendes, zunächst einmal theoretisches vereinfachtes Modell zu bilden. Und von einem solchem Konstrukt ausgehend soll den Jugendlichen ermöglicht werden, die Berechnung zu illustrieren und eine angemessene Interpretation darzubieten. Bei dem ersten Versuch ein Modell zu konstruieren; soll dabei die fachliche Mathematik vollkommen ausgeschlossen werden und es sollen nur mit trivialen Gegenständen die wichtigsten Komponenten aus dem Kontext repräsentiert und in einem adäquaten Zusammenhang wiedergegeben werden. [21, S. 12] Genau so wird das Verständnis und die oft noch reduzierte Vorstellungskraft der Lernenden gefördert und die selbstständige Arbeit gefordert. Das Umformulieren von Aufgabenstellungen für Schülerinnen und Schüler bekannte Sachverhalte erleichtert das Lernen. Wenn die Lernenden das Problem eigenständig erklären können und den Kern der Aufgabe erkannt haben, ist bereits ein wichtiger Teil des Lösungsweges geschafft. Auf das triviale Modell, welches die eigentliche Fragestellung verkörpert, wird dann das mathematische stochastische Wissen angewandt. [21, S. 12] Das eben formulierte Konzept wird durch die Abbildung 27 visualisiert.

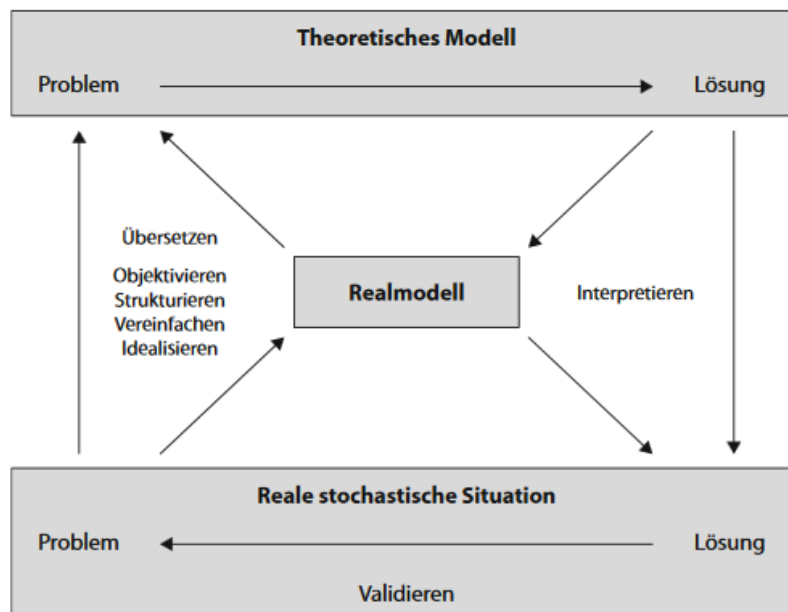


Abbildung 27: Modellbilden stochastischer Probleme

Das Problem der Stochastik und des Denkens in wahrscheinlichkeitstheoretischen Kontexten liegt wohl darin, dass man aus eigener Erfahrung viele Vorurteile gegenüber der Wahrscheinlichkeit mitbringt. Auch wenn das Lösen von Problemen aus der Wahrscheinlichkeitstheorie gekannt wird, wird das Berechnen immer durch einen Hauch von Eigenerfahrungen beeinflusst. Eine solche Voreingenommenheit präsentiert das typische Warten auf die 6 beim beliebten Spiel *Mensch ärgere dich nicht*. Man hat den Eindruck, dass ewig auf die Würfelseite gewartet werden muss, um überhaupt seinen ersten Zug starten zu können. [21, S. 14]

Trotz der Problematik, die das stochastische Denken so anspruchsvoll werden lässt, gibt es genügend plausible Gründe, die die Bedeutung des Themas, vor allem in der Schule, darstellen. Wie den Vorteil der praxisnahen Anwendung und die eigene Erfahrung, die unterstützend für das Verständnis wirkt. [22, S. 21] Ein weiterer Aspekt, der die Vielfalt der Stochastik repräsentiert, ist die Erweiterung, die das Thema besitzt. In der Stochastik sind viele Gebiete der Mathematik von Nutzen. Zum Beispiel wird die Kompetenz zum Prozentrechnen und auch zum Bruchrechnen gefordert. [22, S. 25]

Aufgrund der vielen Bereiche, die das Kapitel der Statistik und der Wahrscheinlichkeitstheorie umfasst, wird der Stochastik in der Schule ein immer bedeutsamerer Stellenwert zugeschrieben. Insbesondere durch die zentrale Reifeprüfung hat die Stochastik durch ihre breite Anwendungsmöglichkeit an Bedeutung dazugewonnen. Heutzutage ist der alleinige Wert der Wahrscheinlichkeit eine 6 zu würfeln nicht mehr gefragt. Die Strukturen und die Anwendungsbereiche haben sich modifiziert und neuen Gegebenheiten angepasst. Daher müssen sich auch die Schülerinnen und Schüler, aber noch mehr die Lehrerinnen und Lehrer entsprechend anpassen.

11.2 Lehrplan für Mathematik in Bezug auf Stochastik

Um einen Einblick in den Lehrstoff der Schule zu gewährleisten, werden im folgenden Abschnitt einige für das Thema Statistik und Wahrscheinlichkeitstheorie bedeutsame, Ausschnitte aus dem Lehrplan des Bundesministeriums für Bildung, Wissenschaft und Forschung aus [30] präsentiert. Dabei wird vor allem auf den Inhalt der Oberstufe für allgemeinbildende höhere Schulen eingegangen, da vorwiegend auch dieses Fachgebiet in der Diplomarbeit behandelt wird.

Der Lehrstoff weist für den Lehrplan der Oberstufe für AHS-Schulen andere Kennzeichen und Merkmale des Unterrichts auf als die Unterstufe. Es fällt auf, dass die einzelnen Themengebiete der jeweiligen Schulstufen ziemlich umfangreich sind und komplex wirken. Außerdem wird in der fünften Klasse kein Themenpool mit Stochastik angeboten. Dieses Kapitel fängt erst ab der sechsten Klasse an. Das bedeutet aber nicht, dass man die in der Unterstufe gelernten statistischen Anwendungen nicht für andere Kapitel benötigt.

Da in der fünften Klasse kein stochastisches Wissen im Lehrplan explizit verlangt wird, wird in der sechsten Klasse unter der Unterüberschrift Stochastik folgender Stoff gefordert:

- *Arbeiten mit Darstellungsformen und Kennzahlen der beschreibenden Statistik*
- *Kennen des Begriffes Zufallsversuch, Beschreiben von Ereignissen durch Mengen*
- *Kennen der Problematik des Wahrscheinlichkeitsbegriffs; Auffassen von Wahrscheinlichkeiten als relative Anteile, als relative Häufigkeiten und als subjektives Vertrauen*
- *Berechnen von Wahrscheinlichkeiten aus gegebenen Wahrscheinlichkeiten; Arbeiten mit der Multiplikations- und der Additionsregel; Kennen des Begriffs der bedingten Wahrscheinlichkeit*
- *Arbeiten mit dem Satz von Bayes [30, S. 5]*

In der siebten Klasse werden die in der sechsten Klasse, erworbenen Kenntnisse erweitert zu diesen Anforderungen:

- *Kennen der Begriffe diskrete Zufallsvariable und diskrete Verteilung*
- *Kennen der Zusammenhänge von relativen Häufigkeitsverteilungen und Wahrscheinlichkeitsverteilungen; von Mittelwert und Erwartungswert sowie von empirischer Varianz und Varianz*
- *Arbeiten mit diskreten Verteilungen (Binomialverteilung) in anwendungsorientierten Bereichen [30, S. 5]*

Und zuletzt folgt der Lehrstoff für eine abschließende achte Klasse einer allgemeinbildenden höheren Schule:

- *Kennen der Begriffe stetige Zufallsvariable und stetige Verteilungen*
- *Arbeiten mit der Normalverteilung in anwendungsorientierten Bereichen*
- *Kennen und Interpretieren von statistischen Hypothesentests und von Konfidenzintervallen [30, S. 6]*

Der Lehrplan wird ständig erweitert und angepasst. Denn früher waren darin weniger Kapitel aus dem Fachgebiet der Stochastik zu finden. Wenn die einzelnen Aufgaben jeder Schulstufe betrachtet werden, ist zu erkennen, dass es einen großen Sprung zwischen der Unterstufe und der Oberstufe für allgemeinbildende höhere Schulen gibt. Der Schwerpunkt wird in der Unterstufe auf die Beschreibung und Darstellung von Daten gelegt. Hingegen konzentriert man sich in der Oberstufe umfassender auf die verschiedenen Begriffe der Wahrscheinlichkeitsrechnung.

An dieser Stelle soll kurz zusammenfassend auf den Inhalt des Lehrplanes in der Unterstufe eingegangen werden. In der ersten Klasse wird der Fokus im entsprechenden Kapitel zwar auch auf Proportionen gelegt, aber im Hinblick auf die weiterführende Stochastik, sind vor allem die Unterpunkte des Modells und das Herauslesen aus graphischer Veranschaulichung von Informationen wichtig. Auch in der zweiten Klasse gibt es einen Bereich, der sich mit Proportionen auseinandersetzt. Allerdings, für die Stochastik relevant sind die, mit einem eigenen Absatz gekennzeichneten Anforderungen der relativen Häufigkeiten und das Interpretieren von Tabellen und Illustrationen. In der dritten Klasse werden in Bezug auf die Statistik und die Wahrscheinlichkeitstheorie nur das Analysieren und das Abbilden von Datenmengen gefordert. Als Abschluss der Unterstufe werden erstmals wichtige Merkmale der Statistik behandelt, die die Schülerinnen und Schüler auch später in der Schullaufbahn benötigen. [29, S. 7]

Nach der doch umfangreichen einführenden Behandlung von Anfängen der Stochastik werden in der fünften Klasse keine Kapitel zum Themengebiet angeboten. Das Thema wird erst wieder in der sechsten Klasse aufgegriffen. Dort werden wichtige grundlegende Begriffe eingeführt, wie der des Zufallsversuches und der Wahrscheinlichkeit. Zusätzlich werden auch erste einfache Berechnungen gefordert. In der nächsten Schulstufe werden die Themen der Stochastik bereits etwas komplexer und die Schülerinnen und Schüler müssen Verteilungsarten kennen. Zuletzt wird in der achten Klasse noch auf das Deuten von Hypothesentests, was vielen Jugendlichen schwer fällt, und auf den Begriff der Stetigkeit im Zusammenhang mit der Stochastik eingegangen. [30, S. 6]

12. Erhebung der Wahrnehmung von Aufgabenstellungen der bedingten Wahrscheinlichkeit an AHS-Schülerinnen und Schülern

Nun soll den Missverständnissen und den Verwechslungen vor allem in Bezug auf die bedingte Wahrscheinlichkeit auf den Grund gegangen werden. Dazu wurden ausgewählte Beispiele der bedingten Wahrscheinlichkeit und der Anwendung des Satzes von Bayes ausgesucht, die von 120 Schülerinnen und Schülern aus AHS Oberstufenklassen gelöst wurden. Bei den getesteten Schulklassen handelte es sich um Oberstufengymnasien im Südburgenland. Zum einen wurde die Untersuchung am BG/BRG/BORG Oberschützen und zum anderen am BORG Güssing durchgeführt. Im folgenden Abschnitt werden die Analyse der gelösten Aufgaben und die Lösungsansätze präsentiert. Ebenso werden die mathematisch korrekten Lösungen der gestellten Aufgaben dargelegt. Weiters wird versucht anhand des richtigen Lösungsweges die Lösungsversuche der Schülerinnen und Schüler zu verstehen und etwaige Gedankenfehler der Jugendlichen zu korrigieren.

12.1 Ausgewählten Beispiele mit Lösungen

In diesem Unterkapitel werden die Aufgaben, welche die Schülerinnen und Schüler präsentiert wurden, wiedergegeben. Nach der Angabe des Problems wird die korrekte Lösung des Beispiels folgen.

1. Eine Urne enthält 100 Kugeln. 70 Kugeln bestehen aus dem Material Holz und 30 Kugeln sind aus Kunststoff. 25 der Holzkugeln sind mit der Farbe Rot gestrichen und 45 sind grün. 10 der Kunststoffkugeln sind rot und 20 sind grün. Folgende Ereignisse werden definiert:

A: Holzkugel	$\bar{A}$: Kunststoff-Kugel
B: Rote Kugel	$\bar{B}$: Grüne Kugel

Jemand zieht eine Kugel und spürt mit der Hand, dass es sich um eine Kunststoffkugel handelt. Wie groß ist die Wahrscheinlichkeit dafür, dass die Kugel in der Hand grün ist? [28]

Wie im Anhang 3 zu erkennen ist, hatten die Schülerinnen und Schüler eine zusätzliche Tabelle als Hilfestellung bei der Angabe dabei. Um auf die korrekte Wahrscheinlichkeit zu kommen, muss man sich stückweise die Wahrscheinlichkeiten ausrechnen, dazu wird eine Tabelle bereitgestellt.

Man berechnet zunächst einmal die Wahrscheinlichkeiten für die etwaigen Möglichkeiten, dass es sich um eine Holzkugel, eine Kunststoffkugel, eine rote oder eine grüne Kugel handelt. Danach kann man aus der Tabelle 9 die entsprechenden Wahrscheinlichkeiten herauslesen.

$$\begin{array}{ll} P(A) = 0,7 & P(\bar{A}) = 0,3 \\ P(B) = 0,35 & P(\bar{B}) = 0,65 \end{array}$$

12. Erhebung der Wahrnehmung von Aufgabenstellungen der bedingten Wahrscheinlichkeit an AHS-Schülerinnen und Schülern

Tabelle 9: Wahrscheinlichkeiten Urnenbeispiel [28]

	B: rot	$\bar{B}$: grün	Σ
A: Holz	$P(A \cap B) = 0,25$	$P(A \cap \bar{B}) = 0,45$	$P(A) = 0,7$
$\bar{A}$: Kunststoff	$P(\bar{A} \cap B) = 0,1$	$P(\bar{A} \cap \bar{B}) = 0,2$	$P(\bar{A}) = 0,3$
Σ	$P(B) = 0,35$	$P(\bar{B}) = 0,65$	1

Es sind durchaus unterschiedliche Lösungsmethoden vorstellbar. Zum Beispiel kann man mit Hilfe eines Baumdiagramms auf die korrekte Lösung gelangen. Es ist aber auch legitim, mit der Formel von Bayes zu arbeiten. Gefragt ist die Wahrscheinlichkeit dafür, dass man eine grüne Kugel zieht, unter der Bedingung, dass man beim Ziehen der Kugel das Material Kunststoff wahrnimmt. Kurz wird geschrieben

$$P(\bar{B}|\bar{A}) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{A})} = \frac{0,2}{0,3} = \frac{2}{3} \approx 0,6.$$

Die Wahrscheinlichkeit eine grüne Kugel zu ziehen, nachdem man schon weiß, dass es sich um eine Kunststoffkugel handelt, beträgt $\frac{2}{3}$.

2. Viele Internetnutzer klagen über Spam-Mails. Nehmen wir an, in 1% der guten und 40% der Spam-Mails komme das Wort "Viagra" vor. Außerdem seien 10% der Mails gut und 90% Spam. Wie groß ist die Wahrscheinlichkeit, dass eine Mail, von der man weiß, dass in ihr das Wort "Viagra" vorkommt, eine Spam-Mail ist? [28]

$P(G) = 0,1$; bedeutet die Wahrscheinlichkeit für eine gute Mail,

$P(\bar{G}) = 0,9$; ist die Wahrscheinlichkeit für eine Spam-Mail

Gesucht ist die Wahrscheinlichkeit, die angibt, ob es sich um eine Spam-Mail handelt, wenn man bereits weiß, dass in der Nachricht das Wort Viagra vorkommt. Also kann man formulieren:

$$\begin{aligned}
 & P(\bar{G}|\text{Wort Viagra}) \\
 &= \frac{P(\bar{G}) \cdot P(\text{man findet das Wort Viagra}|\bar{G})}{P(\bar{G}) \cdot P(\text{Wort Viagra}|\text{Spam-Mail}) + P(G) \cdot P(\text{Wort Viagra}|G)} \\
 &= \frac{0,9 \cdot 0,4}{0,9 \cdot 0,4 + 0,1 \cdot 0,01} \approx 0,997.
 \end{aligned}$$

Mit rund 99,7% Wahrscheinlichkeit handelt es sich um eine Spam-Mail, wenn das Wort Viagra in der Nachricht vorkommt.

3. Drei einer ernsten Krankheit verdächtigen Personen A, B und C wurde eine Blutprobe entnommen. Das Untersuchungsergebnis sollte vorläufig nicht bekannt gegeben werden. A erfuhr, dass sich nur bei einer Person der Verdacht bestätigte, und bat den Arzt, ihm im Vertrauen den Namen einer der Personen B oder C zu nennen, die gesund ist. Der Arzt lehnt die Auskunft mit der Begründung ab, dass damit die Wahrscheinlichkeit dafür, dass A erkrankt ist, von $\frac{1}{3}$ auf $\frac{1}{2}$ ansteigen würde. A bestreitet dies. Schlichten Sie den Streit unter

der Annahme, dass der Arzt, wenn A an der ansteckenden Krankheit leidet, mit gleicher Wahrscheinlichkeit (nämlich $\frac{1}{2}$) B oder C nennen würde! [19]

Das Problem stellt einen Bezug zum schon behandelten Ziegenproblem dar. Allerdings wurde das Beispiel hier in einen anderen Kontext gekleidet. Anfangs beträgt die Wahrscheinlichkeit für jeden der Patienten $\frac{1}{3}$, dass der Bluttest positiv ausfällt. Würde der Arzt zum Beispiel Person B nennen, würde die dazugehörige Wahrscheinlichkeit wie folgt lauten:

$$P(\text{Person } A \text{ ist gesund} | \text{Arzt sagt } B \text{ gesund}) = \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0} = \frac{2}{3}.$$

Also lässt sich daraus schließen, dass die Wahrscheinlichkeit für Person A einen negativen Befund zu erhalten gleich $\frac{2}{3}$ und nicht $\frac{1}{2}$ ist.

4. Sie wissen, dass eine Familie zwei Kinder hat, und Sie wissen zusätzlich, dass eines davon ein Mädchen ist. Wie groß ist die Wahrscheinlichkeit, dass beide Kinder Mädchen sind? [14, S. 22]

Dieses Beispiel kann je nach Interpretation anders verstanden werden. Die Lösungen differenzieren sich nach der entsprechenden Auffassung der Problemstellung. Um die Wahrscheinlichkeit zu berechnen, dass die Familie zwei Töchter hat, kann man von dem Fall ausgehen, dass die Menge $\{MM, MJ, JM, JJ\}$ die Basis der möglichen Fälle bildet, wobei die Konstellation $\{JJ\}$ ausgeschlossen werden kann, da man ja schon weiß, dass es sicher ein Mädchen in der Familie gibt. Daraus kann man annehmen, dass die Wahrscheinlichkeit für zwei Mädchen $\frac{1}{3}$ beträgt, da es dafür drei mögliche Konstellationen gibt. [18, S. 112]

Eine andere Lösung liefert die Version mit dem Satz von Bayes. Da man bereits weiß, dass die Familie ein Mädchen hat, kann man aus dieser Information daraus auf die bedingte Wahrscheinlichkeit schließen, dass die Familie zwei Töchter hat. Man fragt dann nach der Wahrscheinlichkeit für zwei Mädchen, unter der Bedingung, dass man weiß, dass in der Familie ein Mädchen ist. Kurz:

$$P(MM|M) = \frac{P(MM) \cdot P(M|MM)}{P(M)} = \frac{1 \cdot \frac{1}{4}}{1 \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{4}} = \frac{1}{2}.$$

Mit Hilfe des Bayes-Theorems beträgt die Wahrscheinlichkeit $\frac{1}{2}$. Daran ist zu erkennen, dass auch die jeweilige Sichtweise eine wichtige Rolle spielt. Und man im Allgemeinen auch keine der beiden Lösungsansätze als falsch beurteilen kann. Die zwei unterschiedlichen Lösungen sind beide mathematisch korrekt. Demgemäß kommt es auf die Denkweise an.

12.2 Analyse der Lösungen der Schülerinnen und Schüler

Die Untersuchung wurde innerhalb einer Woche an beiden Schulen unabhängig voneinander durchgeführt. Dabei waren jeweils die sechsten und siebten Klassen von der Untersuchung betroffen. Das Verhältnis der Geschlechter verteilte sich genau 1 : 1. In der Abbildung 28 erkennt man die dazugehörige Verteilung differenziert nach Alter und Geschlecht. Es ist deutlich zu erkennen, dass die meisten Probandinnen und Probanden im Alter von 16 und 17 Jahren waren. Einige wenige Studierende waren etwas älter beziehungsweise jünger. Dennoch ergab das Geschlechterverhältnis ein sehr ausgewogenes. Es nahmen 60 Jungen und 60 Mädchen an der Umfrage teil.

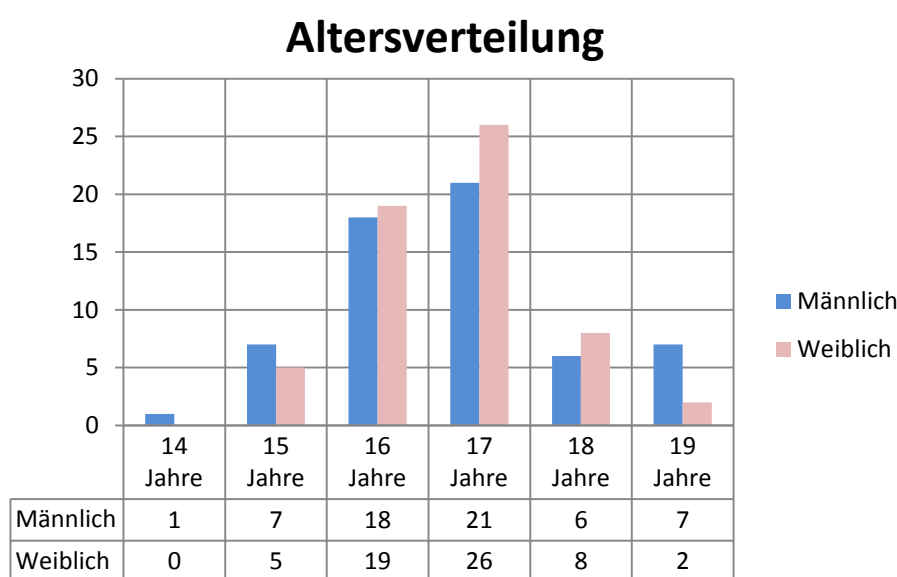


Abbildung 28: Altersverteilung

Die nächsten Abbildungen und Diagramme sind etwas spezifischer dargestellt. Damit diese nachvollziehbar sein können, folgt zunächst einmal eine kurze Erläuterung zur Klassifizierung der behandelnden Beispiele. Die Lösungen der Schülerinnen und Schüler wurden unterteilt nach „richtig“, „falsch“ und „intuitive“. Als „richtig“ gelten alle Antworten, die mit einem Lösungsweg zu einem richtigen Ergebnis geführt haben. „Falsch“ wurden solche Angaben gewertet, die keinen richtigen Wert anführten oder das Beispiel gar nicht behandelten. Und als „intuitive“ Antworten gelten jene, die einen richtigen Schätzwert in Form eines Bruches oder einer Prozentangabe ohne Rechenweg als Lösung anschraben.

Um einen ersten Eindruck von der Gesamtleistung zu bekommen, werden in der folgenden Kreisdiagrammen die Verteilungen der Antworten der Schülerinnen und Schüler dargestellt. Abbildung 29 zeigt die Aufgabe 1, welches von 66,67% aller Schülerinnen und Schüler richtig gelöst wurde. 32,5% haben das Beispiel falsch gelöst und 0,83% der Schülerinnen und Schüler haben es intuitiv gelöst.

Zu den Lösungsstrategien lässt sich erwähnen, dass die meisten Jugendlichen den Wert direkt aus der zur Verfügung gestellten Tabelle ablesen konnten und somit auf die richtige Lösung gelangten. Für einige wiederum stellte die Tafel ein Hindernis dar, da viele aus einer falschen Zeile beziehungsweise aus einer falschen Spalte die gegebenen Werte nahmen. Die

12. Erhebung der Wahrnehmung von Aufgabenstellungen der bedingten Wahrscheinlichkeit an AHS-Schülerinnen und Schülern

intuitive Antwort kommt daher zustande, dass ein Schüler oder eine Schülerin sich wohl bei der Lösung nicht sicher war und ein $\sim$ vor dem Ergebnis schrieb. Eine Schülerin oder ein Schüler gab auch die Angabe zu diesem Beispiel, dass genau so eine Aufgabe zur Schularbeit kam. Allerdings hatte er beziehungsweise sie die Lösungsmethode schon wieder vergessen.

Man kann diese Angaben noch genauer betrachten und weiter in Details aufspalten. Dann kann man aus der Tabelle 10 alle Prozentangaben nach Alter und Geschlecht erkennen. In einer Altersgruppe wurde das Beispiel richtig gelöst. Bei den 15- und 19-jährigen Burschen haben es ungefähr 71% richtig gelöst. Von den 16- und 17-jährigen Jungen wurde es nur von zweidrittel korrekt gelöst. Überraschend ist, dass zweidrittel der 18-jährigen Schüler nicht auf die korrekte Lösung kamen. Bei den Teilnehmerinnen zeigt sich, dass von den 15-jährigen Mädchen 80% die erste Aufgabe lösen konnten. Hingegen wurde es nur von rund 60% der 16-jährigen Schülerinnen korrekt gelöst. Bei den 17- und 18-jährigen Frauen liegt der Anteil der richtigen Lösungen zwischen 73% und 75%. Und nur 50% der 19-jährigen wurde des Problem entsprechend gelöst.

Tabelle 10: Prozentangaben von Beispiel 1

Geschlecht	Alter in Jahre	Richtig gelöst in Prozent	Falsch gelöst in Prozent	Intuitiv gelöst in Prozent
Männlich				
	14	100	0	0
	15	71,4	28,6	0
	16	66,7	33,3	0
	17	66,7	28,6	4,8
	18	33,3	66,7	0
	19	71,4	28,6	0
Weiblich				
	15	80	20	0
	16	57,9	42,1	0
	17	73,1	26,9	0
	18	75	25	0
	19	50	50	0

Beispiel 1

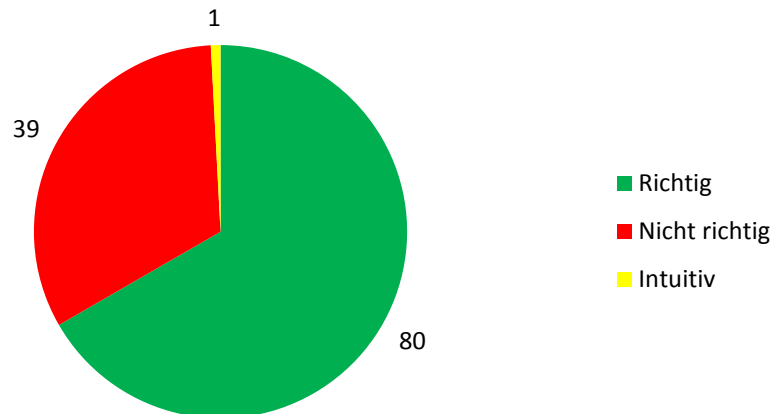


Abbildung 29: Verteilung von Beispiel 1

Das zweite Beispiel war wohl am schwierigsten für die Jungen und Mädchen. Denn es wurde von 99,17% falsch gelöst. Nur ein Schüler schaffte dieses Beispiel zu lösen. Er löste es mit Hilfe von dem Satz von Bayes und kam dementsprechend auf die richtige Lösung. Die meisten Jugendlichen probierten bei dieser Aufgabe die Wahrscheinlichkeiten 40% und 90% einfach zu multiplizieren. Die am häufigsten genannte falsche Antwort war daher, dass die Wahrscheinlichkeit eine Spam-Mail zu erhalten, wenn das Wort Viagra darin enthalte ist, bei 36% lag.

Beispiel 2



Abbildung 30: Verteilung von Beispiel 2

Ein ähnlicher Trend erwies sich beim Beispiel 3. Auch hier ist der Anteil an falschen beziehungsweise nicht berechneten Angaben sehr hoch. Man muss dazu sagen, dass viele Schülerinnen und Schüler bei diesem Beispiel nicht einmal versucht haben einen Lösungsweg zu finden und das Beispiel ausließen. Dabei handelt es sich bei dieser Angabe um das bekannte Ziegenproblem, welches allerdings unter einem anderen Kontext gegeben war.

Viele Schülerinnen und Schüler haben dieses Beispiel wohl aufgrund des längeren Textes nicht gelöst. Oft wirken lange Zeilen für Lernende abschreckend und deshalb geben Studierende bereits beim Durchlesen auf, da sie nicht gewohnt sind entsprechend umfangreiche Texte konzentriert durchzulesen.

Allerdings haben sowohl 1,7% der Mädchen als auch 1,7% der Jungen dieses Problem erkannt und sind mit dem Problem richtig umgegangen. Sie haben sogar vermerkt, dass es sich dabei um das Ziegenproblem beziehungsweise um das Monty-Hall-Problem handelt. Insgesamt haben 3,3% zumindest erkannt, um welches berühmte Problem es sich handelt, aber konnten nicht mehr mit der Aufgabe anfangen.



Abbildung 31: Verteilung von Beispiel 3

Bei Beispiel 3 lässt sich bei den Angaben der Schülerinnen und Schüler ein Muster erkennen. Denn viele haben zwar das Problem nicht erkannt, aber gaben trotzdem einen intuitiven Schätzwert zur Lösung an. Unter den als falsch beurteilten Angaben der Jugendlichen lässt sich der Trend zu den Antwortmöglichkeiten von $\frac{1}{2}$ und $\frac{1}{3}$ erkennen, wie in Abbildung 32 erkennbar ist. So meinten 30% der Schülerinnen und Schüler, dass der Arzt Recht habe und die Wahrscheinlichkeit auf $\frac{1}{2}$ steigen würde, wenn er einen der Namen *B* oder *C* nennen würde. Und 7% der Befragten meinten, dass die Wahrscheinlichkeit von $\frac{1}{3}$ gleich bliebe.

Beispiel 3 Antwortmöglichkeiten

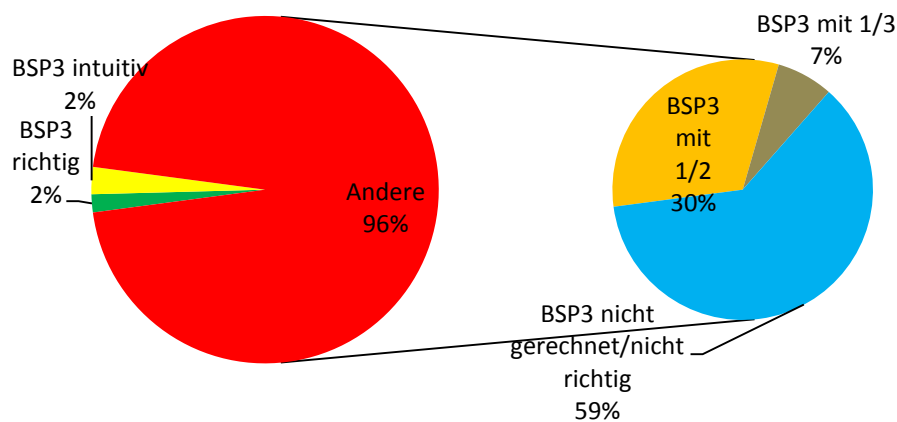


Abbildung 32: Antwortmöglichkeiten von Beispiel 3

Und zuletzt präsentiert Abbildung 33 die Verteilung der Antwort von Frage 4. Dieses Problem konnten insgesamt 11,7% der Jungen und 33,3% der Mädchen richtig lösen. Allerdings wurde dieses Beispiels von niemand mit Hilfe der Bayes Regel gelöst, sondern alle arbeiteten mit einem Baumdiagramm. Daher ist auch keiner der Studierenden auf die Lösung von $\frac{1}{3}$ gekommen, sondern die richtig gewerteten Lösungen waren mit dem Wert von $\frac{1}{2}$. Es waren auch intuitiv Angaben von $\frac{1}{2}$, die von den Probandinnen und Probanden nur geschätzt wurden. Hier beträgt der Anteil der Burschen bei 55% und bei den Mädchen bei 36,7%.

Beispiel 4

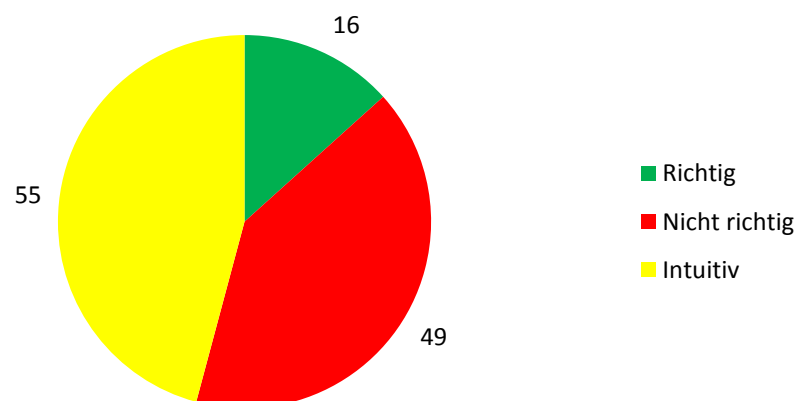


Abbildung 33: Verteilung von Beispiel 4

Auch bei Beispiel 4 lässt sich ähnlich wie beim Beispiel 3 ein Muster zu anderen häufigen Lösungen, die aber falsch sind, erkennen. Abbildung 34 zeigt den Anteil von 34% der als falsch gelösten Angaben, die die Wahrscheinlichkeit von $\frac{1}{4}$ einschätzten. Was diese

Schülerinnen und Schüler jedoch nicht beachteten, war, dass die Konstellation von $\{JJ\}$ wegfällt, da man ja bereits weiß, dass ein Mädchen in der Familie ist.

Beispiel 4 Antwortmöglichkeiten

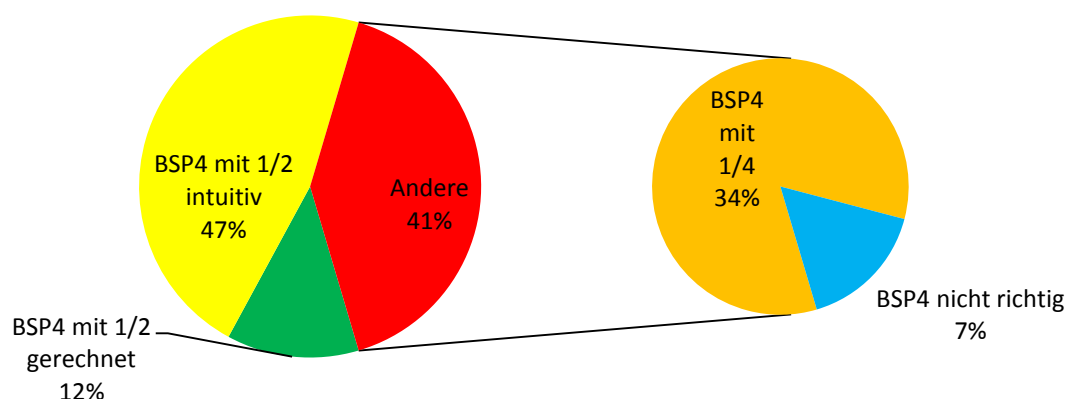


Abbildung 34: Antwortmöglichkeiten von Beispiel 4

Sehr auffällig ist, dass einige Schülerinnen und Schüler ihre Lösungen nicht hinterfragt haben. Denn das Beispiel 2 lässt mit Überlegung einen höheren Schätzwert als 36% zu. Es sollte relativ bewusst sein, dass in einer normalen E-Mail nicht das Wort Viagra vorkommt. Auch beim ersten Beispiel haben einige Jugendliche das Ergebnis nicht hinterfragt. So sind sie zu einem ziemlich unwahrscheinlichen Resultat von $\frac{1}{3}$ oder ähnliches gekommen. Beispiel 3 ist ein ziemlich komplexer Sachverhalt, welches auch Fachmathematiker manchmal in Unsicherheit bringen lässt. Dennoch war es enttäuschend, dass einige Schülerinnen und Schüler nicht einmal eine intuitive Antwort angaben, allerdings ist dies wohl auf die längere Angabe zurückzuführen. Bei Beispiel 4 war die Verteilung im Vergleich zu den anderen Aufgaben etwas ausgeglichener. Trotzdem erkennt man im Allgemeinen, dass der Satz von Bayes bloß von einzelnen Fällen richtig verwendet wurde. Die Mehrheit arbeitete mit Baumdiagrammen oder schätzte die Wahrscheinlichkeit nur ungefähr ab.

Zusätzlich zu den Rechenaufgaben wurden am Ende des Fragebogens folgende offene Fragen gestellt. An dieser Stelle werden weitere Diagramme für die Beantwortung dieser Fragen gezeigt.

Kommen dir diese Beispiele aus dem Unterricht bekannt vor? Wenn ja: um welches Thema handelt es sich in der Mathematik?

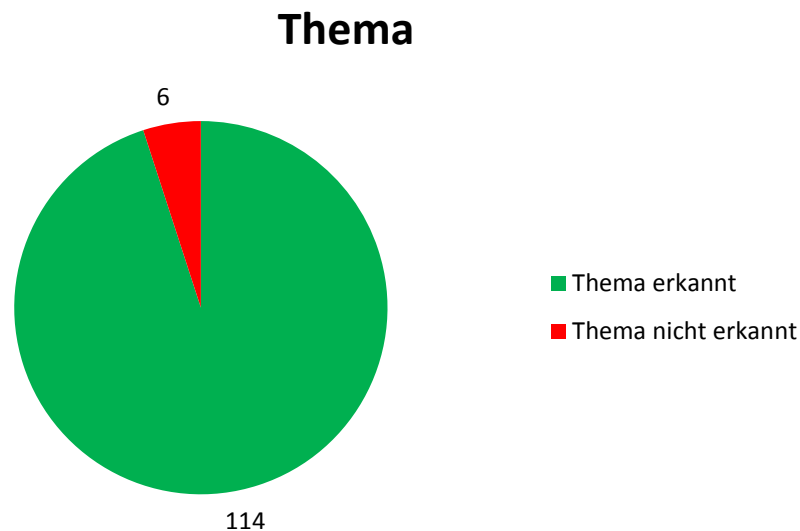


Abbildung 35: Verteilung Thema

Anhand der Abbildung 35 ist deutlich erkennbar, dass die meisten Jugendlichen, das in den Fragebögen behandelnde Thema, erkannt haben. Zwei der Teilnehmer erkannten sogar, dass es sich dabei explizit um die bedingte Wahrscheinlichkeit handelte. Sechs der Untersuchten war das Thema unbekannt.

Habt ihr bereits ähnliche Beispiele im Unterricht gerechnet?

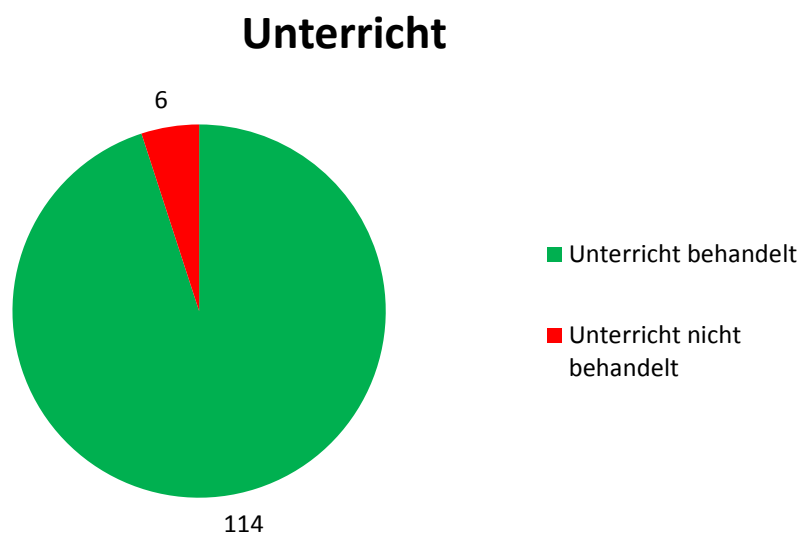


Abbildung 36: Verteilung Unterricht

Gleich wie das vorherige Diagramm fällt die Abbildung 36 aus. Mehrheitlich wurde das Thema im Unterricht schon behandelt. Viele Lernende kommentierten, dass es schon länger her ist und sie deshalb nicht mehr wüssten wie es ginge. Oder es wurde auch behauptet, dass in der Schule leichtere Beispiele geübt wurden.

Findest du diese Beispiele nützlich? Ja/nein, warum/warum nicht?

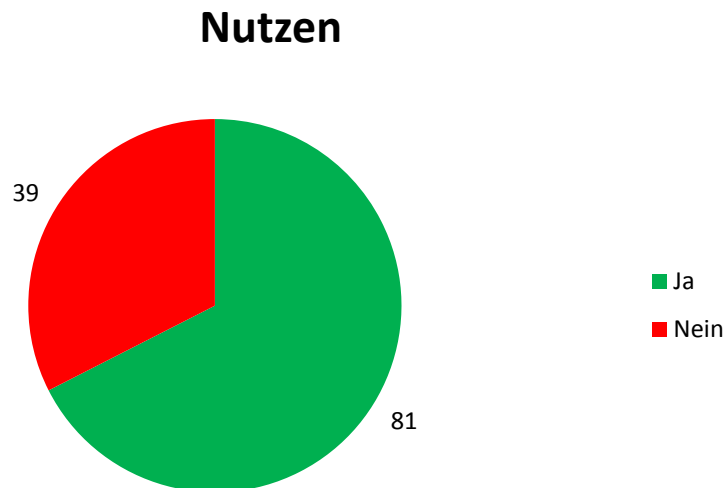


Abbildung 37: Verteilung Nutzen

Auf die Frage, ob die Schülerinnen und Schüler einen Nutzen in dieser Thematik der Mathematik erkennen würden, wurden folgende Daten erhoben. Dabei lässt sich in der Abbildung 37 erkennen, dass fast dreiviertel der Untersuchten Aufgaben der Wahrscheinlichkeit als nützlich empfinden.

Wo wird diese Art der Wahrscheinlichkeitsrechnung deiner Meinung nach im Alltag verwendet/gebraucht?

Auf die weiterführende Frage von Frage 3 wurden einige sehr umfangreiche Einsatzgebiete der Wahrscheinlichkeitstheorie der Schülerinnen und Schüler genannt. Die am häufigsten genannte Antwort war der Gebrauch bei Gewinnspielen, dicht gefolgt von Alltagsnutzen wie in Abbildung 38 zu erkennen ist.

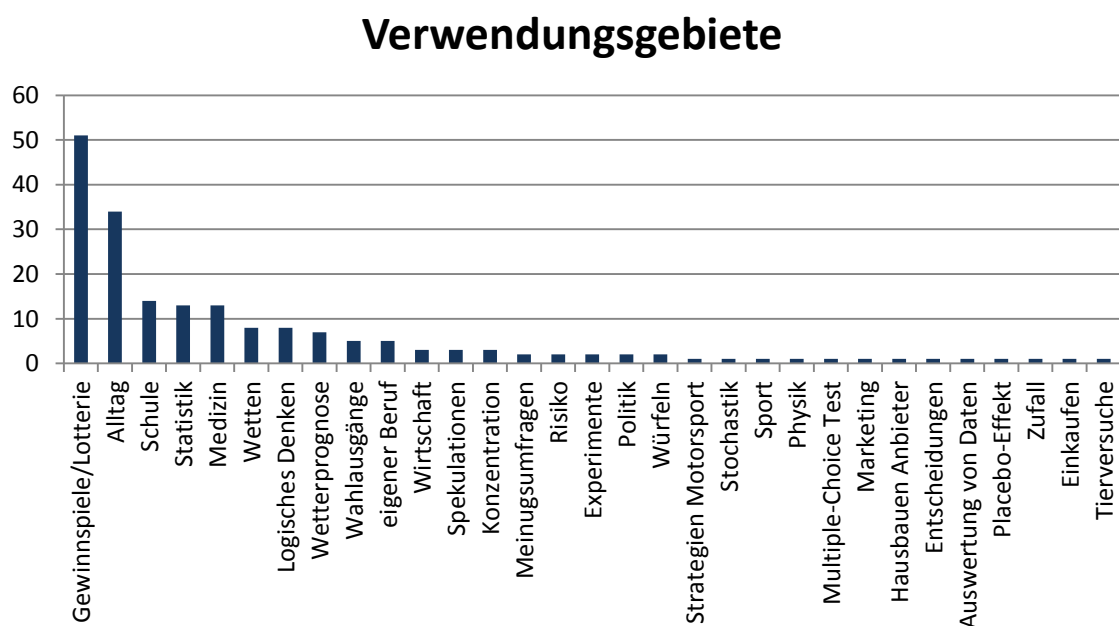


Abbildung 38: Verwendungsgebiete

Literaturverzeichnis

- [1] Allmendinger, H, Lengnink, K, Vohns, A, Wickel, G (2013): Mathematik verständlich unterrichten. Perspektiven für Unterricht und Lehrerbildung. Springer Fachmedien Wiesbaden, Wiesbaden, s.l.
- [2] Bauer, TK, Gigerenzer, G, Krämer, W (2014): Warum dick nicht doof macht und Genmais nicht tötet. Über Risiken und Nebenwirkungen der Unstatistik. Campus, Frankfurt am Main.
- [3] Behrends, E (1994): Überall Zufall. Eine Einführung in die Wahrscheinlichkeitsrechnung. BI-Wiss.-Verl., Mannheim.
- [4] Behrends, E (2013): Elementare Stochastik. Ein Lernbuch - von Studierenden mitentwickelt. Vieweg+Teubner Verlag, Wiesbaden.
- [5] Berger, J (2010): Paradoxien aus Naturwissenschaft, Geschichte und Philosophie. Anaconda, Köln.
- [6] Borovcnik, M: Zum Teilungsproblem.
https://www.researchgate.net/profile/Manfred_Borovcnik/publication/305073570_Zum_Teilungsproblem_-_On_the_Division_of_Stakes/links/587510c508aebf17d3b3f29f/Zum-Teilungsproblem-On-the-Division-of-Stakes.pdf.
- [7] Borovcnik, M (1992): Stochastik im Wechselspiel von Intuitionen und Mathematik. BI-Wiss.-Verl., Mannheim.
- [8] Bosch, K (2011): Elementare Einführung in die Wahrscheinlichkeitsrechnung. Mit 82 Beispielen und 73 Übungsaufgaben mit vollständigem Lösungsweg. 11. Auflage. Vieweg+Teubner Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden, Wiesbaden.
- [9] Bronder, T (2016): Spiel, Zufall und Kommerz. Theorie und Praxis des Spiels um Geld zwischen Mathematik, Recht und Realität. 1. Auflage. Springer, Berlin, Heidelberg.
- [10] Büchter, A, Henn, H-W (2005): Elementare Stochastik. Eine Einführung in die Mathematik der Daten und des Zufalls. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg.
- [11] Cramer, E, Kamps, U (2014): Grundlagen der Wahrscheinlichkeitsrechnung und Statistik. Ein Skript für Studierende der Informatik, der Ingenieur- und Wirtschaftswissenschaften. 3. Auflage. Springer Spektrum, Berlin.
- [12] Dürr, D, Froemel, A, Kolb, M (2017): Einführung in die Wahrscheinlichkeitstheorie als Theorie der Typizität. Mit einer Analyse des Zufalls in Thermodynamik und Quantenmechanik. Springer Spektrum, Berlin, Heidelberg.
- [13] Eckhardt, K (2013): Stochastik. Statistik und Wahrscheinlichkeitsrechnung in der Landwirtschaft ; 44 Tabellen. Ulmer, Stuttgart.
- [14] Eckhardt, U (19.01.2010): Mathematik - Das Werkzeug der Erkenntnis 2. Wahrscheinlichkeit & Statistik, Hamburg.
- [15] Gigerenzer, G, Zillgitt, M: Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken. 3. Auflage.
- [16] Grams, T (2016): Klüger irren - Denkfallen vermeiden mit System. Springer, Berlin, Heidelberg.

- [17] Hafner, R (1989): Wahrscheinlichkeitsrechnung und Statistik. Springer, Wien u.a.
- [18] Henze, N (2004): Stochastik für Einsteiger. Eine Einführung in die faszinierende Welt des Zufalls. 5. Auflage. Vieweg+Teubner Verlag, Wiesbaden, s.l.
- [19] Humenberger, Johann Univ.-Prof. Mag. Dr. (Sommersemester 2017): 250043 VO Schulmathematik Stochastik (2017S).
- [20] Krämer, W: Denkste! Trugschlüsse aus der Welt des Zufalls und der Zahlen. 2. Auflage.
- [21] Krüger, K, Sill, H-D, Sikora, C (2015): Didaktik der Stochastik in der Sekundarstufe I. Springer Spektrum, Berlin.
- [22] Kütting, H (1994): Didaktik der Stochastik. BI-Wiss.-Verl., Mannheim.
- [23] Kütting, H, Sauer, MJ (2014): Elementare Stochastik. Mathematische Grundlagen und didaktische Konzepte. 3. Auflage. Springer Spektrum, Berlin.
- [24] Kwech, A (1992): Darstellung der Geschichte der Wahrscheinlichkeitstheorie anhand ausgewählter Kapitel, Institut für Mathematik, Universität Wien.
- [25] Linde, W (2014): Stochastik für das Lehramt. De Gruyter Oldenbourg, Berlin.
- [26] Lotto 6 aus 45: Regeln 6 aus 45. <https://www.lotto6aus45.com/regeln>. Abgerufen am 14.03.2018.
- [27] Löwe, M, Knöpfel, H (2011): Stochastik - Struktur im Zufall. 2. Auflage. Oldenbourg, München.
- [28] Mathe-Brinkmann: Bedingte Wahrscheinlichkeit. https://www.brinkmann-du.de/mathe/gost/stoch_01_08.htm. Abgerufen am 12.04.18.
- [29] Mathematik. Bildungs- und Lehraufgabe (Unterstufe). https://bildung.bmbwf.gv.at/schulen/unterricht/lp/ahs14_789.pdf?61ebzm. Abgerufen am 04.04.18.
- [30] Mathematik. Bildungs- und Lehraufgabe (Oberstufe). https://bildung.bmbwf.gv.at/schulen/unterricht/lp/lp_neu_ahs_07_11859.pdf. Abgerufen am 04.04.18.
- [31] McGrayne, SB (2014): Die Theorie, die nicht sterben wollte. Wie der englische Pastor Thomas Bayes eine Regel entdeckte, die nach 150 Jahren voller Kontroversen heute aus Wissenschaft, Technik und Gesellschaft nicht mehr wegzudenken ist. Springer Spektrum, Berlin.
- [32] Schnell, S (2014): Muster und Variabilität erkunden. Konstruktionsprozesse kontextspezifischer Vorstellungen zum Phänomen Zufall. Springer Spektrum, Wiesbaden.
- [33] Schuster, T, Liesen, A (2017): Statistik für Wirtschaftswissenschaftler. Ein Lehr- und Übungsbuch für das Bachelor-Studium. 2. Auflage. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [34] Sproesser, U, Wessolowski, S, Wörn, C (2014): Daten, Zufall und der Rest der Welt. Didaktische Perspektiven zur anwendungsbezogenen Mathematik. Springer Spektrum, Wiesbaden.
- [35] Vogel, F (2005): Beschreibende und schließende Statistik. Formeln, Definitionen, Erläuterungen, Stichwörter und Tabellen. 13. Auflage. De Gruyter Oldenbourg, München.

Abbildungen

Abbildung 1: Statistik einer Gehaltskurve	5
Abbildung 2: Manipulierte Statistik einer Gehaltskurve I	5
Abbildung 3: Manipulierte Statistik einer Gehaltskurve II.....	5
Abbildung 4: Menge Ω [18, S. 9].....	17
Abbildung 5: Teilmenge E [18, S. 9].....	17
Abbildung 6: Unvereinbare Ereignisse [18, S. 9].....	17
Abbildung 7: Gegenereignis [18, S. 9]	17
Abbildung 8: Und-Ereignis [18, S. 9]	17
Abbildung 9: Oder-Ereignis [18, S. 9]	17
Abbildung 10: Relative Häufigkeiten eines Münzwurfs bei 10 Würfeln [10, S. 145]	19
Abbildung 11: Relative Häufigkeiten eines Münzwurfs bei 100 Würfeln [10, S. 145].....	19
Abbildung 12: Relative Häufigkeiten eines Münzwurfs bei 1000 Würfeln [10, S. 145]	19
Abbildung 13: Bedingte Laplace-Wahrscheinlichkeiten [10, S. 168].....	34
Abbildung 14: Bedingte Wahrscheinlichkeiten [10, S. 168]	34
Abbildung 15: Bedingte Wahrscheinlichkeiten [10, S. 168]	34
Abbildung 16: Allgemeines Baumdiagramm [18, S. 175]	47
Abbildung 17: Ziegenproblem anhand eines Baumdiagramms [18, S. 105].....	47
Abbildung 18: Baumdiagramm absolute Häufigkeiten [15, S. 72]	48
Abbildung 19: Baumdiagramm Pfadregel [27, S. 46]	49
Abbildung 20: 1. Pfadregel anhand eines Beispiels [21, S. 135]	50
Abbildung 21: 2. Pfadregel anhand eines Beispiels [21, S. 139]	51
Abbildung 22: Glockenkurvige Darstellung einer Binomialverteilung mit unterschiedlichen n [18, S. 140].....	56
Abbildung 23: Galtonbrett [12, S. 22]	57
Abbildung 24: Rutherford – Geiger Experiment mit adaptierter Poissonverteilung [18, S. 191]	62
Abbildung 25: Gauß'sche Glockenkurve [33, S. 112].....	65
Abbildung 26: Maximum-Likelihood-Methode [10, S. 320].....	69
Abbildung 27: Modellbilden stochastischer Probleme	86
Abbildung 28: Altersverteilung	92
Abbildung 29: Verteilung von Beispiel 1	94
Abbildung 30: Verteilung von Beispiel 2	94
Abbildung 31: Verteilung von Beispiel 3	95
Abbildung 32: Antwortmöglichkeiten von Beispiel 3	96
Abbildung 33: Verteilung von Beispiel 4	96
Abbildung 34: Antwortmöglichkeiten von Beispiel 4	97
Abbildung 35: Verteilung Thema.....	98
Abbildung 36: Verteilung Unterricht.....	98
Abbildung 37: Verteilung Nutzen	99
Abbildung 38: Verwendungsgebiete	99

Tabellen

Tabelle 1: Gewinnchancen des St. Petersburger Paradoxons [5, S. 58]	10
Tabelle 2: Logarithmusfunktion des St. Petersburger Paradoxons [5, S. 65]	11
Tabelle 3: Ereignisse E und deren Erklärung [10, S. 137]	17
Tabelle 4: Vierteltafel [10, S. 167] adaptiert von [8, S. 29].....	33
Tabelle 5: Gewinnchancen Monty Hall Paradoxon [5, S. 70]	41
Tabelle 6: Ergebniswerte zum Rutherford–Geiger Experiment [10, S. 265]	62
Tabelle 7: Fehlerarten der Hypothesentests [11, S. 270].....	71
Tabelle 8: Ausgänge des Werfens mit einem Würfel [10, S. 361].....	73
Tabelle 9: Wahrscheinlichkeiten Urnenbeispiel [28]	90
Tabelle 10: Prozentangaben von Beispiel 1.....	93

Anhang

Anhang 1: Elternbrief

Liebe Eltern,

ich bin Studierende der Universität Wien für das Lehramtsstudium Mathematik und Spanisch. Im Zuge meiner Diplomarbeit führe ich eine Untersuchung an Schulen durch. Für den experimentellen Teil meiner Arbeit habe ich ausgewählte Beispiele aus der Wahrscheinlichkeitsrechnung ausgesucht, die Ihr Kind in den nächsten Tagen lösen soll. Dabei werden die von den Schülerinnen und Schülern erbrachten Lösungen selbstverständlich anonym behandelt und die erhobenen Daten nur für die Analyse meiner Diplomarbeit verwendet. Ebenso wenig werden die Lösungen Ihres Kindes an die Lehrerin oder an den Lehrer weitergegeben und haben somit auch keine Auswirkung auf die Note Ihres Kindes.

Ich bitte Sie höflichst den nächsten Abschnitt auszufüllen und der zuständigen Lehrerin oder dem zuständigen Lehrer zukommen zu lassen.

Vielen Dank für Ihre Mitarbeit!

Theresa Schuh

Hiermit bestätige ich, dass mein Kind _____ an der empirischen Untersuchung teilnehmen darf.

Datum/Unterschrift des/der Erziehungsberechtigten:

Anhang 2: Operationalisierung der Fragestellung

- Die Schülerinnen und Schüler sollen die Aufgaben selbstständig und alleine bearbeiten!
- Zum Berechnen der Aufgaben ist genügend Platz auf den Angaben vorhanden. Falls es nötig ist, werden zusätzliche Blätter zur Verfügung gestellt!
- Zum Berechnen sind Hilfsmittel, wie Taschenrechner, Computer (falls vorhanden) und Lineal (falls benötigt) erlaubt!
- Falls die Schülerinnen und Schüler nicht wissen, wie die Aufgaben zu lösen sind, wird um eine intuitive Antwort gebeten!
- Wichtig: Es geht bei diesen Aufgabenstellungen nicht um das schulische Können der Schülerinnen und Schüler! Gefragt ist nur der Umgang mit diesen Aufgaben aus der Wahrscheinlichkeit in der Schule!
- Die Leistung der Schülerinnen und Schüler bei diesen Aufgaben beeinflusst keineswegs ihren Erfolg in der Schule!
- Darum wird gebeten, dass die Schülerinnen und Schüler die Aufgaben so gut es möglich ist berechnen und eine Lösung angeben.

Vielen Dank für die Mitarbeit!

Anhang 3: Vorlage des Untersuchungsmaterial

Löse die Aufgaben, so gut du kannst! Versuche mit einem Lösungsweg bis zur Lösung zu kommen. Falls du nicht weißt, wie eine Aufgabe zu lösen ist, versuche eine intuitive Lösung anzugeben.

☐ Männlich	☐ Weiblich
Alter:	

- 1.) Eine Urne enthält 100 Kugeln. 70 Kugeln bestehen aus dem Material Holz und 30 Kugeln sind aus Kunststoff. 25 der Holzkugeln sind mit der Farbe Rot gestrichen und 45 sind grün. 10 der Kunststoffkugeln sind rot und 20 sind grün. Folgende Ereignisse werden definiert:

A: Holzkugel $\bar{A}$: Kunststoff Kugel
 B: Rote Kugel $\bar{B}$: Grüne Kugel

	B: rot	$\bar{B}$: grün	Σ
A: Holz	25	45	70
$\bar{A}$: Kunststoff	10	20	30
Σ	35	65	100

Jemand zieht eine Kugel und spürt mit der Hand, dass es sich um eine Kunststoffkugel handelt. Wie groß ist die Wahrscheinlichkeit dafür, dass die Kugel in der Hand grün ist? [28]

- 2.) Viele Internetnutzer klagen über Spam-Mails. Nehmen wir an, in 1% der guten und 40% der Spam-Mails komme das Wort "Viagra" vor. Außerdem seien 10% der Mails gut und 90% Spam. Wie groß ist die Wahrscheinlichkeit, dass eine Mail, von der man weiß, das in ihr das Wort "Viagra" vorkommt, eine Spam- Mail ist? [28]
- 3.) Drei einer ernsten Krankheit verdächtigen Personen A , B und C wurde eine Blutprobe entnommen. Das Untersuchungsergebnis sollte vorläufig nicht bekannt gegeben werden. A erfuhr, dass sich nur bei einer Person der Verdacht bestätigte, und bat den Arzt, ihm im Vertrauen den Namen einer der Personen B oder C zu nennen, die gesund ist. Der Arzt lehnt die Auskunft mit der Begründung ab, dass damit die Wahrscheinlichkeit dafür, dass A erkrankt ist, von $\frac{1}{3}$ auf $\frac{1}{2}$ ansteigen würde. A bestreitet dies. Schlichten Sie den Streit unter der Annahme, dass der Arzt, wenn A an der ansteckenden Krankheit leidet, mit gleicher Wahrscheinlichkeit (nämlich $\frac{1}{2}$) B oder C nennen würde! [19]
- 4.) Sie wissen, dass eine Familie zwei Kinder hat, und Sie wissen zusätzlich, dass eines davon ein Mädchen ist. Wie groß ist die Wahrscheinlichkeit, dass beide Kinder Mädchen sind? [14, S. 22]

- Kommen dir diese Beispiele aus dem Unterricht bekannt vor? Wenn ja: um welches Thema handelt es sich in der Mathematik?
- Habt ihr bereits ähnliche Beispiele im Unterricht gerechnet?
- Findest du diese Beispiele nützlich? Ja/nein, warum/warum nicht?
- Wo wird diese Art der Wahrscheinlichkeitsrechnung deiner Meinung nach im Alltag verwendet/gebraucht?

Danke für die Mitarbeit!

Anhang 4: Genehmigung des LSR Burgenlandes zur Durchführung der Untersuchung



Zahl: **LSR/2-24/15-2018**
Bei Antwortschreiben Bezugszahl anführen!

Sachbearb.: LSI Mag. Jürgen Neuwirth
Telefon/DW: (02682) 710-1223
E-Mail: juergen.neuwirth@lsr-bgld.gv.at

Theresa Martha Schuh
E-Mail: theresa.schuh@gmx.at

Eisenstadt, 03. April 2018

Genehmigung zur Durchführung wissenschaftlicher Untersuchungen

Der Landesschulrat für Burgenland genehmigt die Durchführung der empirischen Untersuchung zum Thema „Erhebung der Wahrnehmung von Aufgabenstellungen der bedingten Wahrscheinlichkeit an AHS Schülerinnen und Schülern“ am BG/BRG/BORG Oberschützen, Hauptplatz 7, 7432 Oberschützen und am BORG Güssing, Schulstraße 17, 7540 Güssing, mit der Berücksichtigung unten angeführten Kriterien.

Über die Teilnahme an dieser Fragebogenaktion kann die Direktion selbst entscheiden.

Die Vorgabe dieses Fragebogens darf nur außerhalb des Unterrichts erfolgen.

- Es muss sichergestellt werden, dass den angesprochenen Schülerinnen und Schülern die Teilnahme an der Untersuchung vollkommen freigestellt wird.
- Verbindliche Übersendung einer Kurzfassung der Evaluierung.
- Die erhobenen Daten dürfen ausschließlich für die von Ihnen angeführte Untersuchung verwendet werden.
- Die Anonymität der Beteiligten ist zu gewährleisten.
- Auf die Einhaltung der einschlägigen Bestimmungen des Schulrechtes wird verwiesen.

Die Direktionen der oben angeführten Schulen wurden von diesem Schreiben in Kenntnis gesetzt.

Für den Bildungsdirektor:
Mag.^a Steiner

Elektronisch gefertigt

7000 Eisenstadt, Kernaustieg 3 • Telefon: (02682) 710 • Telefax: (02682) 710-1009 • office@lsr-bgld.gv.at
• www.lsr-bgld.gv.at •

Abstract

Abstract in Deutsch:

Der Umgang mit stochastischen Problemen wird immer wieder unterschätzt. Vor allem wird in der Schule das Thema Stochastik oft nur unzureichend besprochen. Jedoch erweist sich dieses Teilkapitel der Mathematik als äußerst hilfreich und praxisorientiert. Es gibt eine unendliche Anzahl an Beispielen aus dem Alltag, die eine Verbindung zur Statistik und zur Wahrscheinlichkeitstheorie knüpfen. Diese Relation zwischen der formalen Mathematik und der bekannten Vertrautheit aus dem täglichen Leben, die das Thema mit sich bringt, bedingt gleichzeitig auch die Irrtümer und Fehlschlüsse der Stochastik, auf die in dieser Arbeit besonders Wert gelegt wird. Es werden aber nicht nur die bekannten Denkfehler erläutert, sondern es wird vor allem auch auf den richtigen Umgang mit den Theorien und Interpretationen der Stochastik hingewiesen.

Die Arbeit beginnt unter anderem mit einführenden und geschichtlichen Tatsachen der stochastischen Entwicklung. Im nächsten Kapitel werden für die Arbeit wichtige Zugänge der beschreibenden Statistik kurz behandelt. Ein wichtiger Teil der Ausarbeitung widmet sich den unterschiedlichen Berechnungsarten und den diversen Interpretationsmöglichkeiten der Wahrscheinlichkeitstheorie.

Da sich viele Missverständnisse und Paradoxa in der bedingten Wahrscheinlichkeit verstecken, wird darauf ebenfalls Augenmerk gelegt. Außerdem wird auch die beurteilende Statistik besprochen. In einem weiteren Kapitel werden die Unterschiede in der Handhabung der Stochastik zwischen der Pädagogik und der Fachmathematik verdeutlicht. Abschließend wird eine empirische Untersuchung an AHS-Schulen präsentiert, die sich mit Lösungen von Schülerinnen und Schülern von ausgewählten Aufgaben der bedingten Wahrscheinlichkeit auseinandersetzt.

Abstract in English

Stochastic problems tend to be underrated. Especially in schools stochastics is often not discussed sufficiently. Yet this part of mathematics has proven to be quite helpful and reality-tested. Countless examples in everyday life prove its connection to statistics and to probability theory. This relation between formal mathematics and the familiarity with everyday life inherent to stochastics also causes stochastic errors and mistakes, which are specially covered in this thesis. But not only the most common fallacies are presented, it is also explained how to properly deal with the theories and interpretations of stochastics.

The thesis starts with, amongst other things, introductory and historical facts concerning the development of stochastics. The next chapter shortly covers approaches of descriptive stochastics which are important for this thesis. Another important part is dedicated to various ways of calculating and to the many possible ways of interpreting the probability theory.

As many a misunderstanding and paradox is hidden in conditional probability, this aspect is covered too – as well as assessing stochastics. Another chapter explains the differences in handling stochastics inside and outside teaching. Finally, an empirical study of high schools is presented. It deals with students' solutions of selected conditional probability tasks.

