Aus einer Vorlesung von Henn, Uni Dortmund:

Die vier Wahrscheinlichkeitsbegriffe und die Riemerquader

Die vier folgenden Auffassungen von Wahrscheinlichkeit könnten an den Anfang gesetzt werden, zur Entwicklung tragfähiger Begriffe ist der letzte Ansatz, der prognostische, am tragfähigsten.

a) Laplacescher Wahrscheinlichkeitsbegriff

In der Schulpraxis dominiert ab der 5. Klasse traditionsgemäß nahezu ausschließlich der Laplacesche Wahrscheinlichkeitsbegriff, bei dem sich die Wahrscheinlichkeiten aufgrund von Symmetrien als Quotienten „günstiger“ und „möglicher“ Fälle berechnen lassen. Hier gibt es zu einer richtigen Hypothese keine sinnvollen Alternativen – man denke etwa an einen Würfel. Die meisten Zufallsexperimente lassen sich durch geeignete Wahl von Elementarereignissen als Laplace-Experimente beschreiben, das Abzählen der für ein interessierendes Ereignis günstigen Fälle erfordert nur mehr oder weniger geschickte kombinatorische Überlegungen. Damit verschwindet der hypothetische Charakter von Wahrscheinlichkeiten, nämlich das Erzeugen von Erwartungshaltungen und Aufstellen von Hypothesen vor dem Experiment und der Bewertung und ggf. Revision der Prognosen durch Gegenüberstellung von Voraussagen und Ergebnissen erst im nachhinein. Dies verklammert erst statistische und wahrscheinlichkeitstheoretische Aspekte.

Zwei weitere Gesichtspunkte sprechen dagegen, diesen Wahrscheinlichkeitsbegriff an den Anfang zu stellen:

Diese „Definition“ erfüllt nicht die Ansprüche an Exaktheit und Wohldefiniertheit von heute. Im Dunkeln bleibt, was „gleichwahrscheinlich“ heißt. Dies hat die Mathematik zu Laplaces Zeit nicht gestört. Heute wäre aber dieser Begriff als erste Definition eine Verfälschung im Sinne des Spiralprinzips.

Diese Definition unterstützt auch die problematische Denkweise vieler Menschen, nur in absoluten, nicht in relativen Dimensionen zu denken (d. h. nur den Zähler zu sehen).

Beispiele:

- Bei Hausgeburten hat man weniger Komplikationen als in der Klinik. Es sei M die Anzahl der Problemgeburten, N die Anzahl aller Geburten,

Klar: M(Haus) < M(Klinik), aber N(H) << N(Kl.).

Außerdem gehen i. A. Risikopatientinnen von vorne herein in die Klinik.

- „Ich stehe stets in der längeren Schlange“.

Klar: Da in der längeren Schlange mehr Personen stehen, haben auch mehr Personen dieses Gefühl.

- 33.000 Kinder verunglücken in den USA jährlich beim Rodeln, also sollte eine so gefährliche Sportart verboten werden.

Klar: Ca. 33 Millionen Kinder sind im Rodelalter 4 – 14 Jahre. Wenn z. B. 50 % rodeln, so verunglückt nur jedes 500ste Kind.

b) Frequentistischer Wahrscheinlichkeitsbegriff

Ebenfalls nur bedingt geeignet zur Entwicklung eines adäquaten Wahrscheinlichkeitsbegriffs sind Experimente zum <u>frequentistischen Wahrscheinlichkeitsbegriff</u>, der Wahrscheinlichkeiten als Grenzwerte relativer Häufigkeiten in unendlich langen Versuchsserien beschreibt; Standardbeispiel hierfür ist das Werfen eines Reißnagels. Wegen der Endlichkeit aller Versuchsserien und des Wunsches, diesen in der Laplaceschen Welt eindeutig bestimmten Grenzwert möglichst genau zu ermitteln, sind Zufallsschwankungen, so paradox dies klingt, vom Lehrer emotional unerwünscht. Dabei wäre es aus statistischer Sicht viel sinnvoller, ein Gefühl für die Schwankungen der relativen Häufigkeiten zu entwickeln (die zum „schwachen Gesetz der großen Zahlen“ führen) als den Blick auf die viel komplexere Konvergenz in „unendlichen“ Versuchsserien zu fixieren (also auf das „starke Gesetz der großen Zahlen“).

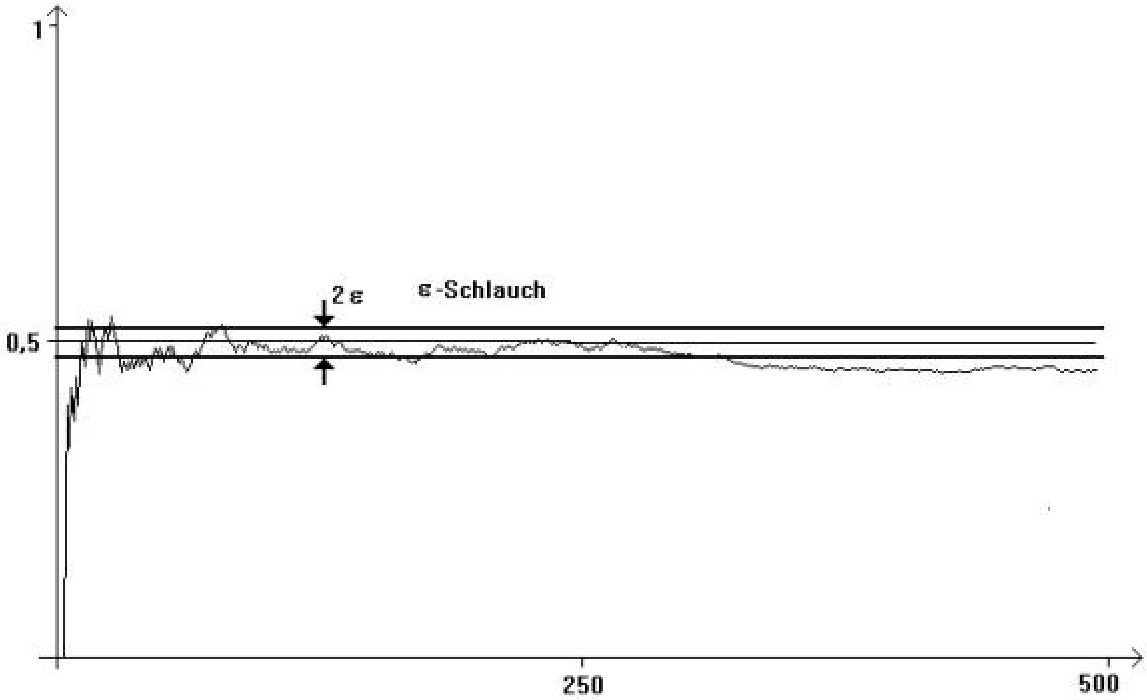
Ausgehend vom frequentistischen Wahrscheinlichkeitsbegriff versuchte Richard v. Mises eine <u>Definition</u> des Wahrscheinlichkeitsbegriffs, die auch Nicht-Laplace-Experimente umfassen sollte. Er definierte die Wahrscheinlichkeit eines Ereignisses E als Limes der relativen Häufigkeiten:

$$p(E) = \lim_{n \to \infty} h_n(E)$$

Leider ist diese Definition mit dem üblichen analytischen Grenzwertbegriff nicht verträglich. Dann müsste nämlich gelten:

$$p(E) = \lim_{n \to \infty} h_n(E) \Leftrightarrow \underset{\varepsilon > 0}{\forall} \; \underset{n_0}{\exists} \; \underset{n \geq n_0}{\forall} \; |p(E) - h_n(E)| < \varepsilon$$

Gerade dies kann aber <u>nicht</u> gewährleistet werden: nach Wahl eines e > 0 kann sehr wohl die relative Häufigkeit wieder aus dem -Schlauch p(E) ± e herauslaufen. In der folgenden Abbildung ist die Computersimulation eines Münzenwurfs abgebildet, bei dem natürlich p(Kopf) = $\frac{1}{2}$ anzusetzen ist. Zunächst wandert der Graph der relativen Häufigkeiten in den gewählten -Schlauch, wandert dann aber nach etwa 300 Würfen wieder hinaus.



Zwar zeigt das Bernoullische Gesetz der großen Zahl

$$p(|h_n(E) - p(E)| > \varepsilon) \leq \frac{1}{4n\varepsilon^2},$$

dass Abweichungen um mehr als e immer unwahrscheinlicher werden. Es handelt sich aber deutlich um einen stochastischen, keinen zur Definition geeigneten analytischen Grenzwert.

c. Axiomatischer Wahrscheinlichkeitsbegriff

Erst Kolmogoroff gab die inhaltliche Deutung des Begriffs „Wahrscheinlichkeit" auf. Der Begriff wurde von ihm uminterpretiert zu dem axiomatischen Ansatz einer „Wahrscheinlichkeitsfunktion" $P: \mathcal{P}(\Omega) \to \mathbb{R}$, deren geforderte Eigenschaften sich an den deduzierten Eigenschaften der relativen Häufigkeiten orientierte:

$P: \mathcal{P}(\Omega) \to \mathbb{R}$ heißt Wahrscheinlichkeitsverteilung auf dem Ergebnisraum Ω, wenn gilt:

(p1) $\forall E \in \mathcal{P}(\Omega)$ gilt $P(E) \geq 0$,

(p2) $P(\Omega) = 1$,

(p3) $\forall E, F \in \mathcal{P}(\Omega)$ gilt: $E \cap F = \emptyset \Rightarrow P(E \cup F) = P(E) + P(F)$.

Diese Definition entspricht für endliche Mengen W der Kolmogoroff-Definition. Alle „üblichen" Eigenschaften lassen sich deduzieren, es fehlt aber jegliche inhaltliche Interpretation! Zum Aufbau stochastischer Grundvorstellungen ist der axiomatische Ansatz nicht geeignet.

d. Prognostischer Wahrscheinlichkeitsbegriff

Dieser Begriff setzt bei den Erfahrungen aus dem „täglichen Leben" an. Bewertungen mit „mehr" oder „weniger wahrscheinlich" sind allenfalls qualitativ und stets subjektiv. Bei diesen Erfahrungen holen wir die Kinder ab. Ein tragfähiger Wahrscheinlichkeitsbegriff muss

- prognostischen Charakter haben, d. h. die Wahrscheinlichkeit stellt eine Prognose dar, um die die relativen Häufigkeiten schwanken werden.

- Er muss hypothetischen Charakter haben, d. h. Wahrscheinlichkeitsprognosen, die sich in Versuchen nicht bewähren, müssen verworfen und durch bessere ersetzt werden.

- Insbesondere kann es durchaus verschiedene, manchmal gleichberechtigte, manchmal verschieden glaubwürdige Wahrscheinlichkeiten geben. Aussagen wie „nach Lage der Versuchsergebnisse halte ich Hypothese A für glaubwürdiger als die Hypothese B" müssen erlaubt sein.

Statt zu Beginn des Unterrichts formal zu erklären, was Wahrscheinlichkeiten sind, sollten die Begriffe mit konkretem Material festgemacht und ihre Bedeutung mit den Kindern ausgehandelt werden. Begriffe sollten nicht als Selbstzweck oder aus formallogischen Gründen, sondern zur Strukturierung dessen, was Schüler tun, eingeführt werden. Als Material haben sich Quader und andere Objekte mit teilweisen Symmetrien bewährt. Einerseits lassen sie im Gegensatz zu Würfeln und Glücksrädern in gewissen Grenzen mehrere sinnvolle Wahrscheinlichkeitsannahmen zu, so dass der hypothetische Charakter von Wahrscheinlichkeiten sinnfällig wird. Andererseits wird – im Gegensatz etwa zum Reißnagel – der Unterschied zwischen Wahrscheinlichkeit und relativer Häufigkeit greifbar: alle akzeptablen Wahrscheinlichkeitshypothesen sind partiell symmetrisch, die relativen Häufigkeiten nur in Ausnahmefällen. Der prognostische Aspekt, der das Verhalten auf lange Sicht beschreibt, prägt sich ein. Dass die von Schülern festgelegten vermuteten Wahrscheinlichkeiten den Kolmogoroff-Axiomen genügen, also implizit eine Wahrscheinlichkeitsverteilung darstellen, ist nahezu selbstverständlich. Ganz von allein erfahren Kinder, dass die Festlegung einer Wahrscheinlichkeitsfunktion kein Problem der Wahrscheinlichkeitstheorie, sondern das im allgemeinen viel kompliziertere Problem, die Realität in ein mathematisches Modell abzubilden, darstellt. Die Wahrscheinlichkeitstheorie lehrt dann, aus den vorhandenen Wahrscheinlichkeiten neue zu berechnen.

Riemer-Quader als Zugang zum Wahrscheinlichkeitsbegriff

Schon bei Kindern im Primarstufenalter lassen sich adäquate Vorstellungen zum prognostischen Wahrscheinlichkeitsbegriff vorbereiten. Bei den üblichen Würfelspielen werden anstatt normaler Laplace-Würfel drei verschiedene Holzquader verwendet (vgl. S. 17, 2.2.2). Vor jedem Wurf darf der Spieler denjenigen „Würfel" wählen, der in der momentanen Spielsituation den größten Erfolg verspricht. Kinder erkennen schnell, dass sich einerseits bei keinem der Quader ein Ergebnis vorhersagen lässt, dass es aber sinnvoll ist, erwünschte Zahlen auf möglichst großen Flächen stehen zu haben.

Die Präzisierung dieser Vorstellungen zur Entwicklung des Wahrscheinlichkeitsbegriffs im Sinne des Spiralprinzips erfordert höchstens Prozentrechnung, hat also schon früh einen Platz in der Sekundarstufe I. Der Lehrer teilt an seine Schüler Quader mit gleicher Nummerierung aus, und lässt zuerst schätzen, in wie viel Prozent aller Fälle auf lange Sicht die einzelnen Augenzahlen zu erwarten sind. Die Ergebnisse werden an der Tafel oder auf einer Folie gesammelt und mit einer subjektiven Wertung der „Glaubwürdigkeit" der jeweiligen Hypothese versehen (machen Sie auch eine eigene Hypothese!).

Name	1	2	3	4	5	6	Glaubwürdigkeit
Stefan	10%	5%	35%	35%	5%	10%	32%
Klara	15%	10%	25%	25%	10%	15%	18%
Sarah	10%	12%	55%	20%	15%	8%	0%

Danach würfelt (oder besser "quadert") jeder 100 mal, um die Hypothesen zu überprüfen:

Name	1	2	3	4	5	6
Stefan	13	10	32	27	5	13
Klara	13	4	34	36	8	7
Sarah	11	4	41	35	5	4

Jetzt kann man diskutieren, welche der Hypothesen am besten von den Ergebnissen unterstützt wird.

Die Werte geben ein Gefühl für die Größe der Schwankungen bei je 100 Versuchen. Einige der ersten Hypothesen werden sich verwerfen lassen, andere wird man revidieren. Das Bild zeigt die 1. Revision der Ausgangshypothesen

Name	1	2	3	4	5	6
Stefan	13%	7%	30%	30%	7%	13%
Klara	10%	6%	34%	34%	6%	10%
Sarah	8%	4%	38%	38%	4%	8%

Die Zusammenfassung der Werte in Gruppen zu 500 oder das Gesamtergebnis der Klasse führen dann zu Häufigkeitsverteilungen, die zu besseren Wahrscheinlichkeitshypothesen führen. Allerdings wird es durchaus verschiedene, teilsymmetrische Hypothesen geben, die brauchbar sind. Auch noch so viele Versuche können daran nichts ändern.

Es ist für Kinder eine wichtige Erkenntnis, dass man Wahrscheinlichkeiten <u>nicht berechnen</u> kann. Man sollte allerdings versuchen, bei den Riemer-Quadern die Flächen der Quader in Beziehung zu den Wahrscheinlichkeiten für die Flächen durchaus nachgehen. Eine oft genannte Hypothese ist es die Wahrscheinlichkeiten proportional zu den Flächeninhalten anzusetzen.

Messen der Würfelkanten zu 2,3 cm, 2,0 cm und 1,3 cm ergibt:

Augenzahl 1, 6, Flächeninhalt 1,3 × 2,3 cm² = 2,99 cm²

Augenzahl 2, 5, Flächeninhalt 2,0 × 1,3 cm² = 2,60 cm²

Augenzahl 3, 4, Flächeninhalt 2,0 × 2,3 cm² = 4,60 cm²

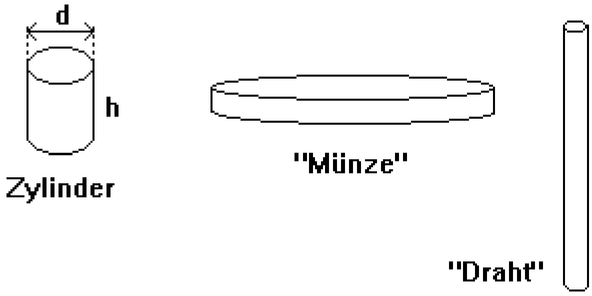
Setzt man diese Werte in Relation zur Gesamtfläche von 20,38 cm², so erhält man die Wahrscheinlichkeitshypothese

1	2	3	4	5	6
14,7%	12,8%	22,5%	22,5%	12,8%	14,7%,

die sich in keiner Weise mit den experimentellen Ergebnissen verträgt. Trotz der schönen Rechnung ist diese Hypothese zu verwerfen.

Weitere analoge Experimente und Vorschläge

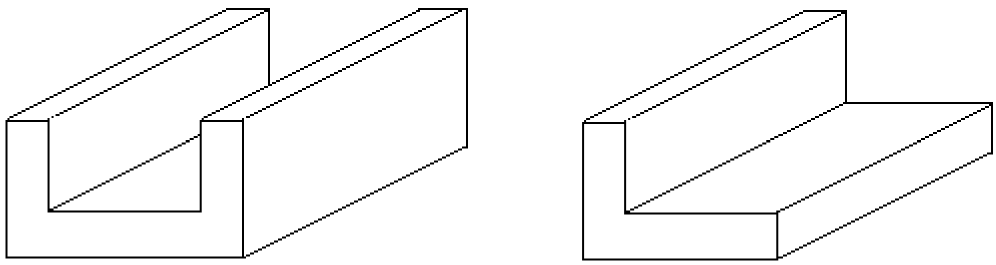
- Zylinder („Stangen“ und „Münzen“)



Man kann jeweils drei Ausfälle ansetzen, die Stange liegt auf der Seite (L) oder auf einer der kreisförmigen Flächen (K, Z), also W = {K, Z, L}. Klar ist, dass die Wahrscheinlichkeiten von dem Verhältnis $\frac{d}{h}$ abhängen. Für d >> h ist p(L) = 0, für d << h ist p(K) = p(Z) = 0. Und sonst?

- Verwendung von Lego-Steinen (vgl. z. B. Lambacher-Schweizer, Bd. 8, Klett 1996, S. 156)

- „Riemer-„Würfel“ in L- und U-Form.



- Martin Bruns, Eine Lehrsequenz zum Stochastikunterricht in der SI. Ziel: den subjektiven Wahrscheinlichkeitsbegriff herausbilden. - In: Beiträge zum Lehren und Lernen von Mathematik (Glatfeld Festschrift, Hrsg. Padberg), S. 27 – 42

Für die Zuordnung von Schätzwerten (subjektiven Wahrscheinlichkeiten) zu Ereignissen durch die Schüler könnte nach einer Idee von Prof. Kroll eine „subjektive Wahrscheinlichkeitsskala“ hilfreich sein. Dies ist eine Strecke, auf der Ereignisse entsprechend ihrer Wahrscheinlichkeit angeordnet werden können. Der linke Streckenendpunkt symbolisiert ein unmögliches Ereignis, der rechte ein sicheres Ereignis. Man kann dann jedem Ereignis intuitiv einen Punkt auf der Skala zuordnen, und zwar entsprechend des subjektiven Zutrauens in sein Eintreffen. Von links nach rechts steigt die Wahrscheinlichkeit an. Die Skala wird anschließend normiert, das heißt, der linke Streckenendpunkt erhält die Wertzuweisung 0, der rechte die Wertzuweisung 1. Durch einen Messprozess kann dann jedem

eingetragenen Ereignis eine Zahl zwischen 0 und 1 – seine Wahrscheinlichkeit – zugeordnet werden.

Aufgabenbeispiele:

Zeichne eine Wahrscheinlichkeitsskala (Länge 6 cm) und trage deine Schätzwerte zur Wahrscheinlichkeit der folgenden Ergebnisse ein! Formuliere weitere Ergebnisse!

a) Mit 20 bist du größer als 1,70 m.

b) Beim Würfeln fällt eine 6.

c) Du bekommst noch Geschwister.

d) Morgen scheint die Sonne.

e) Im nächsten Sommer gehst du baden.

f) Beim Hochsprung schaffst du 1m im 1. Versuch.

g) Eine Frau bekommt ein Kind und es wird ein Mädchen.

h) Bei eurer nächsten Klassensprecherwahl wird ein Mädchen gewählt.